Exploring the energy landscape of GFP by single-molecule mechanical experiments

Hendrik Dietz and Matthias Rief*

Physik Department E22, Technische Universität München, James-Franck-Strasse, D-85748 München, Germany

Edited by George H. Lorimer, University of Maryland, College Park, MD, and approved October 4, 2004 (received for review June 24, 2004)

We use single-molecule force spectroscopy to drive single GFP molecules from the native state through their complex energy landscape into the completely unfolded state. Unlike many smaller proteins, mechanical GFP unfolding proceeds by means of two subsequent intermediate states. The transition from the native state to the first intermediate state occurs near thermal equilibrium at ≈35 pN and is characterized by detachment of a seven-residue N-terminal α-helix from the beta barrel. We measure the equilibrium free energy cost associated with this transition as 22 k_BT. Detachment of this small α-helix completely destabilizes GFP thermodynamically even though the β-barrel is still intact and can bear load. Mechanical stability of the protein on the millisecond timescale, however, is determined by the activation barrier of unfolding the β-barrel out of this thermodynamically unstable intermediate state. High bandwidth, time-resolved measurements of the cantilever relaxation phase upon unfolding of the β-barrel revealed a second metastable mechanical intermediate with one complete β-strand detached from the barrel. Quantitative analysis of force distributions and lifetimes lead to a detailed picture of the complex mechanical unfolding pathway through a rough energy landscape.

Mechanical stability is a property of protein structure that has recently become accessible through single-molecule experiments (1–3). The importance of protein stability under a mechanical force becomes immediately obvious for proteins bearing mechanical function *in vivo*, like in muscle (4) and cytoskeletal proteins (5). Because mechanical single-molecule experiments exhibit single amino acid-length resolution and can thus also provide important structural information (5–7), they have become a valuable tool to explore the energy landscape of proteins (8). Most proteins investigated in folding studies to date are small, containing <200 amino acid residues. Folding studies of larger proteins in classical bulk experiments are generally difficult because their folding kinetics are slow and many large proteins tend to aggregate in the unfolded state. Moreover, growing structural complexity causes more intermediate states to become populated; therefore, analysis and interpretation of multiexponential kinetics becomes difficult. For single-molecule mechanical experiments, however, large protein sizes can be handled easily (9).

Here, we use GFP to investigate the energy landscape of a complex protein in single-molecule force experiments. GFP is among the most important proteins used in biotechnology (10). Numerous applications ranging from localization studies of proteins in living cells to fluorescent pH or Ca sensors have been developed (10). Nevertheless, folding studies of this protein are rare (10–12) for the above reasons.

A detailed knowledge of the mechanical stability of GFP would open the door to an application of GFP as a molecular force sensor. Because fluorescence is closely coupled to an intact GFP structure (13), understanding the determinants of mechanical GFP stability will allow us to develop mutants with known unfolding forces that can report by means of their fluorescence signal the forces acting on the single-molecule level inside a living cell. The many known fluorescent GFP analogs, like red fluorescent protein and yellow fluorescent protein (10) may already exhibit different mechanical stability.

Materials and Methods

Cloning and Protein Expression. We investigated Cycle3-GFP (14) with an additional point mutation, S2G. The Ig8 titin construct, containing the domains I27–I34 from human cardiac titin was as described in ref. 9. The shortened actin crosslinker dictyostelium filamin (DdFLN1-5) construct was as described in ref. 5. To insert a single Cycle3-GFP domain into the Ig8 construct, two restriction sites (*Nco*I and *Bst*EII) were introduced into the Ig8 expression vector between the Ig30 and Ig31 domain sequences by using the QuikChange mutagenesis kit (Stratagene). The same restriction sites were introduced into the DdFLN1-5 expression vector between domains DdFLN3 and DdFLN4 by using the QuikChange Multi Site-Directed Mutagenesis kit (Stratagene). The vectors were digested, and the GFP sequence was introduced and relegated. Protein purification was done by using an N-terminal His-6-Tag. The presence of the inserted GFP domain in the two constructs was verified by DNA sequencing and protein gel electrophoresis. The sample solutions showed the typical GFP fluorescence.

Force Spectroscopy of Single Proteins. All single-molecule force measurements were performed on a custom-built atomic force microscope. Calibration of cantilevers was done in solution by using the equipartition theorem (15, 16). This method provides a resolution in force of ≈10%. Two types of gold-coated cantilevers (Bio-Levers, Olympus, Tokyo) with spring constants and resonance frequencies of either 30 pN/nm and 8.5 kHz (type A), respectively, or 6 pN/nm and 1.5 kHz (type B) were used. For the measurements, ≈20 μl of protein solution (PBS buffer, pH 7.4; protein concentration, ≈0.5 g/liter) were applied on a freshly evaporated gold surface and incubated for 5 min at room temperature. The force curves on the Ig8–GFP construct were collected at pulling speeds ranging from 100 to 1,000 nm/s. The force curves on the DdFLN1-5–GFP construct used in the analysis of unfolding forces and contour lengths as well as in the lifetime analyses were collected at a pulling speed of 300 nm/s with Olympus type A Bio-Levers. All experiments were conducted at room temperature.

Measurement of Contour Lengths and Unfolding Forces. For quantitative analysis of contour lengths and unfolding forces the force-extension traces were fit to an interpolation formula of the worm-like chain model (WLC), $F(x) = (k_BT/p)[0.25(1 - x/L)^{-2} - 0.25 + x/L]$ as introduced by Bustamante *et al.* (17). L denotes the contour length of the stretched protein, p is the persistence length, k_B is Boltzmann's constant, T is the temperature in Kelvin, and x is the distance between attachment points of the protein (extension or end-to-end distance). A value of 0.5 nm for p provided the best fit for the investigated force regime (20–150 pN) and was held fixed at this value for all fits. Unfolding forces were determined from the points where the WLC fits intersect the cantilever relaxation phase after unfolding of a protein domain. All fits and calculations were performed with IGOR PRO 4.01 (WaveMetrics, Lake Oswego, OR).

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: DdFLN, dictyostelium filamin; WLC, worm-like chain model.

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

Interpretation of Unfolding Forces. To interpret the distributions of observed unfolding forces, Monte Carlo simulations were performed as described in ref. 18. Experimental conditions could be simulated by including the measured distribution of contour lengths at which GFP unfolding was observed. The simulated unfolding force distributions were then fitted to those obtained from experiment (see Fig. 1*e*) to yield the unloaded lifetime $\tau_0(F = 0)$ under zero force and Δx, which denotes the distance from the folded state to the transition state, along the reaction coordinate defined by the end-to-end distance of the protein and by the direction of the applied force (see supporting information, which is published on the PNAS web site).

Elastically Coupled Two-State Model. To analyze the observed force plateaus in the rising slopes before the unfolding of GFP$\Delta\alpha$, we fitted force-extension traces calculated by an equilibrium elastically coupled two-state model (18, 19) to our data (see Fig. 3). From the fits we obtain the free energy cost $\Delta G = 22 \pm 4\ k_BT$ for the transition and the contour length increase $\Delta L = 3.2 \pm 1$ nm.

Calibration and Interpretation of Contour Lengths. The contour lengths L, measured by means of the WLC fits to the data, had to be calibrated to obtain the number of amino acids involved in an unfolding event. For calibration, we used Ig domains of DdFLN, which contain exactly 100 amino acids per domain (20). The average contour length increase after the unfolding of a single DdFLN domain is $\Delta L = 32.5 \pm 0.3$ nm. Taking into account an N-terminal-to-C-terminal distance, $d_{N;C,}$ of 4 ± 0.1 nm of the folded domain from the NMR structure (21) we arrive at a total unfolded contour length of $L = \Delta L + d_{N;C} = 36.5$ nm. From this we derive a contour length, L_{aa}, of 0.365 nm per unfolded amino acid residue. We used this calibration factor for all contour length analyses. A value of 0.365 nm is a lower limit for L_{aa}, because in our estimate we assume all 100 amino acids in the DdFLN domains are tightly folded.

The distributions of contour length increases, ΔL, we show in this paper are mainly even distributions around a well defined average value, which indicates that the structural transitions are well defined. The width of these contour length increase distributions is due to a measurement error of 3 nm for a single trace. The error of the mean for a sample of 87 measurements is therefore 0.33 nm.

In this study we use the full structural information available from the Cycle3–GFP crystal structure (Protein Data Bank ID code 1B9C) (22) to map measured contour length increases to changes in the protein structure. The contour length of a protein is determined by two contributions: the contour length of the unfolded amino acids and the distance between the first structured residue at the N terminus and the last structured residue at the C terminus (see supporting information). The first part is given by $m{\cdot}L_{aa}$, assuming m unfolded amino acids. The second part, $d_{N;C,}$ can be extracted from the coordinates in the crystal structure. If a protein undergoes an unfolding transition from a state I to a state II, during which m amino acid residues unfold, the measured change in contour length will be given by $\Delta L = m{\cdot}L_{aa} + d_{N;C}(II) - d_{N;C}(I)$. Both distances, $d_{N;C}(I)$ and $d_{N;C}(II)$, can be calculated from the coordinates of the backbone C^{α} atoms of the first and last amino acid of the folded structure provided in x-ray diffraction or NMR data. As an example, if in a first transition seven N-terminal amino acids (the N-terminal α-helix) of GFP would detach from the folded portion (residues 4–230), we would expect a length increase of $\Delta L = 7{\cdot}L_{aa} + d_{11;230}(II) - d_{4;230}(I) = 2.9 \pm 0.2$ nm. The indices denote the respective folded residues. In Fig. 2*a*, we identify structural candidates for GFP$\Delta\alpha$. For unfolding events leading to complete unfolding of GFP $d_{N;C}(II) = 0$. We calculated the expected contour length increase for GFP structures lacking an increasing number of amino acids from either the N or C terminus, respectively. This contour length increase is given then by $\Delta L = (227 - i){\cdot}L_{aa} - d_{4+i;230}(I)$ if i amino acid residues detach from the N terminus or by $\Delta L = (227 - i){\cdot}L_{aa} - d_{4;230-i}(I)$ if detachment occurs from the C terminus. The result is plotted in Fig. 2*a* as a function of index i.

Lifetime Measurements. In this study, we observed a stable GFP intermediate (GFP$\Delta\alpha\Delta\beta$) with a lifetime under load in the microsecond to millisecond range. To investigate the lifetimes we performed measurements with a fast atomic force microscope cantilever (Bio-Lever type A, Olympus) by using a sampling rate of 20 kHz. We analyzed the cantilever relaxation phase after the unfolding of GFP$\Delta\alpha$ for deviations from the temporal relaxation of an overdamped cantilever. Under our experimental conditions, this relaxation phase is completed within three data points. Relaxation phases of three data points were therefore counted as zero lifetime events, whereas longer relaxation phases with intermediate levels were counted with a precision of $\pm 50\ \mu$s (compare with Fig. 4*b*) to yield the lifetime histogram of GFP$\Delta\alpha\Delta\beta$ (see Fig. 5).

Interpretation of Lifetimes. The average measured lifetime of GFP$\Delta\alpha\Delta\beta$ is $<t> = 1.3$ ms. The increase in force during this period at a pulling velocity of 300 nm/s is only 2 pN, which is in the range of 1% of the absolute forces acting on the intermediate. Thus, the forces acting on the GFP intermediate can be readily assumed constant. To interpret the distribution of observed lifetimes, the force dependence of lifetimes had to be included into the analysis. Poisson statistics predicts the probability to measure a certain lifetime t of the intermediate to

$$P(t, F) = \frac{1}{\tau(F)} \exp\left(-\frac{t}{\tau(F)}\right).$$

$P(t, F)$ is a function of the applied force because the average lifetime, $\tau(F)$, is a function of force. This average lifetime is predicted by Bell's formula, $\tau(F) = \tau_0 \exp(-F\Delta x/k_BT)$, where τ_0 denotes the unloaded lifetime and Δx denotes the potential width as described above (23). This formula has recently been shown to describe protein unfolding kinetics very well (24). In our experiments, the forces acting on GFP$\Delta\alpha\Delta\beta$ obey a broad distribution, $g(F)$, determined by the unfolding forces of the preceding peak. We derived this force distribution, $g(F)$, from the force distribution of the unfolding of GFP$\Delta\alpha$ shown in Fig. 1*e*. We corrected the forces of this distribution by the calculated force drop in a WLC polymer when its contour length is increased by 6.8 nm. By including the force distribution $g(F)$ of acting forces (see Fig. 5 *Inset*), we arrive at a weighted probability to measure a certain lifetime of the structure:

$$P(t) = \frac{1}{C} \int_0^{\infty} \frac{g(F)}{\tau(F)} \exp\left(-\frac{t}{\tau(F)}\right) dF.$$

The probabilities $P(t)$ to measure a certain lifetime were calculated and fitted to the measured distribution of lifetimes (see Fig. 5) to obtain the unloaded lifetime τ_0 and the potential width Δx of the intermediate state GFP$\Delta\alpha\Delta\beta$.

Estimating Transition Barrier Heights. The barrier height ΔG for a transition state from a conformational state can be estimated by means of the Arrhenius equation $\Delta G = -k_BT{\cdot}\ln(\tau_A/\tau_0)$, where τ_0 denotes the unloaded lifetime of the state and $1/\tau_A$ is the Arrhenius frequency factor. For protein dynamics, τ_A has a value of 10^{-9} s (25). We chose this value for all barrier height estimations.

Results and Discussion

Single Domain Force Spectroscopy. In mechanical, single-molecule unfolding experiments, it is crucial to distinguish unfolding events from multiple molecule interactions and nonspecific interactions between tip and substrate. The design of modular proteins with

multiple repeating subunits leading to a characteristic sawtooth-shaped mechanical unfolding pattern has proven a robust strategy to detect and measure unfolding forces of the individual subunits (9, 26–28). In the present experiment, we have combined this strategy with the possibility to investigate single protein domains (5). Therefore, we sandwiched GFP between four N-terminal and four C-terminal Ig domains of human cardiac titin to create an Ig8–GFP chimera protein (Fig. 1*a*). Force extension curves of this construct should exhibit a clear sawtooth pattern because of the unfolding of the Ig domains (9) and an additional unfolding event reflecting GFP unfolding. Force extension curves from this construct are shown in Fig. 1*b*. In all traces at extensions >80 nm, the well studied sawtooth pattern of titin domain unfolding is visible (colored in blue). In contrast to curves obtained from an Ig8 construct lacking the GFP insert (9), at shorter extensions and lower forces, an additional unfolding event appears (colored in green). The length increase associated with this unfolding event is 2.5 times larger than that of an individual Ig domain. Comparing the number of amino acids constituting GFP (238 amino acids) with that folded in an Ig domain (89 amino acids), we arrive at a length for the unfolded polypeptide chain of GFP longer by a factor of 2.7 as compared with Ig domains, which suggests that the observed unfolding event reflects GFP unfolding. Titin Ig domains exhibit a much higher unfolding force than GFP; thus, GFP unfolding events always occur at small extensions close to the surface where nonspecific interactions often mask true unfolding events. To rule out any nonspecific contributions, we designed a construct comprising five Ig domains from the actin crosslinker DdFLN with GFP sandwiched between domain 3 and 4 (DdFLN1-5–GFP) (Fig. 1*c*). The Ig domains of DdFLN exhibit unfolding forces in a range between 40 and 100 pN (5), and GFP unfolding in this chimera construct should occur at higher extensions after the weakest Ig domains have already unfolded. In force extension traces obtained with DdFLN1-5–GFP (Fig. 1*d*), the GFP unfolding events are indeed shifted toward higher extensions. Nonetheless, both the length gain and the unfolding force are identical to Ig8–GFP. Results for the two constructs are therefore in agreement, and we conclude that at pulling velocities of 300 nm/s GFP on average can bear mechanical loads of 104 ± 40 pN (Fig. 1*e*) before it unfolds.

Unfolding of the β-Barrel Occurs from an Intermediate. Evans and Ritchie (29) have shown that the distribution of unfolding forces (Fig. 1*e*) contains information about the underlying energy profile, that is, both the lifetime of the domain in the absence of force τ_0 and the N-terminal -to-C-terminal distance from the folded structure to the transition state Δx. We have used a Monte Carlo simulation to extract τ_0 and Δx from the data (open bars in Fig. 1*e*). We get a good fit to our histogram by using a value of 0.28 ± 0.03 nm for Δx and 14 ± 6 s for τ_0. A value of 0.3 nm for Δx is typical for β-sheet domains (9, 30, 31). However, a lifetime of 14 s stands in stark contrast to values of 10^{10} s obtained from solution experiments (12). Moreover, comparing the folding time of GFP of 600 s (10, 11) with our measured lifetime of 14 s suggests the state before mechanical unfolding has a positive free energy of at least 3.7 k_BT in contrast to $\Delta G = -16\ k_BT$, which is determined from bulk measurements (12). This seeming discrepancy can be resolved if we consider that the GFP structure unfolding at 104 pN may not be the native state but an intermediate state, during which part of the structure has already unfolded at lower forces. Clear evidence for the existence of such an intermediate comes from measuring the contour length increase ΔL upon GFP unfolding (see *Materials and Methods*). Force spectroscopy has an excellent length resolution down to the single amino acid (5, 6). The average length increase, ΔL, we measure is 76.6 ± 0.3 nm, $n = 87$ (Fig. 1*f*). If all of the 227 amino acids resolved in the x-ray structure of Cycle3–GFP (amino acids 4–230) by Battistutta *et al.* (22) contributed to the measured length increase, the expected value for ΔL would be 79.4 nm. This difference is well resolvable with our instrument, and it shows that the force-bearing structure unfolding at 104 pN is smaller than the native structure. A more detailed analysis, for which we plotted the calculated ΔL for structures that lack a certain amount of amino acids from either the N terminus (red) or the C terminus (blue) of the crystal structure of GFP is shown in Fig. 2*a* (see *Materials and Methods* for details). This analysis suggests two possible structures for the shorter intermediate state. The intersection points of the horizontal line indicating our measured value of $\Delta L = 76.6 \pm 0.3$ nm with the calculated curve in Fig. 2*a* lead to either seven amino acids removed from the N terminus or eight to nine amino acids removed from the C terminus. Intriguingly, the seven N-terminal amino acids constitute an α-helix that ends at the first β-sheet of the

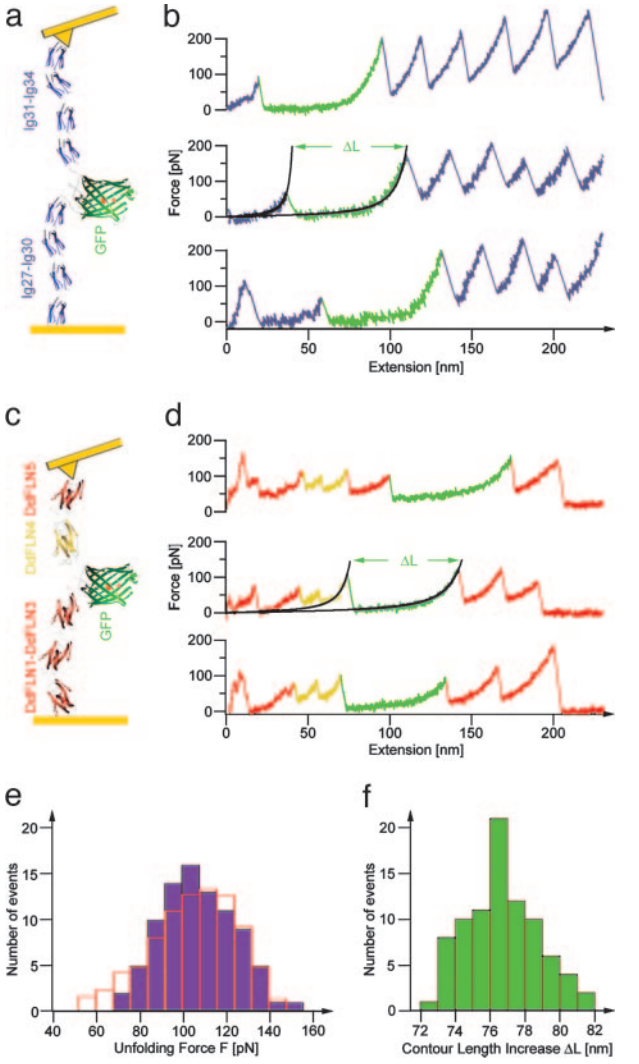

Fig. 1. Single-domain force spectroscopy of GFP. (*a*) Scheme of Ig8–GFP chimera protein, stretched between a gold surface and a gold-coated cantilever tip. (*b*) Three typical force-extension traces measured with Ig8–GFP. GFP unfolding and the subsequent stretching phase of the unfolded polypeptide now lengthened by ΔL is marked in green. Black lines show WLC fits by using a persistence length, *p*, of 0.5 nm and contour lengths *L* and *L* + Δ*L*. (*c*) Scheme of DdFLN1-5–GFP chimera protein. (*d*) Three typical force-extension traces measured with DdFLN1-5–GFP. The GFP unfolding pattern is marked in green, and the double peak indicating DdFLN domain 4 unfolding (5) is marked in yellow. Black lines show WLC fits as in *b*. (*e*) Distribution of GFPΔα unfolding forces *F* (see *Materials and Methods*) obtained from measurements with DdFLN1-5–GFP. Open red bars show results from a Monte Carlo unfolding kinetics simulation of GFPΔα with $\tau_0 =$ 14 s and $\Delta x = 0.28$ nm. (*f*) Distribution of the contour length increase, Δ*L*, because of GFPΔα unfolding as measured by WLC fits to DdFLN1-5–GFP traces (compare to *d*).

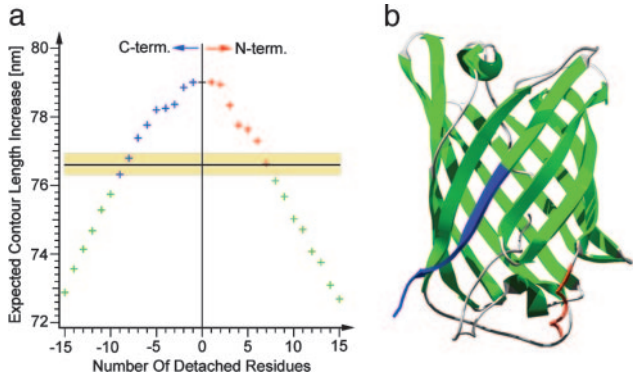

Fig. 2. Contour length mapping. (*a*) Expected contour length increase because of unfolding of GFP as a function of the number of residues detached from either the N or C terminus from the native Cycle3–GFP structure (Protein Data Bank ID code 1B9C) (22). The black line with the shaded yellow region marks the measured contour length increase of 76.6 ± 0.3 nm. (*b*) Cycle3–GFP crystal structure (22). The regions marked in blue and red correspond to the C-terminal and N-terminal solution of the calculation in *a*.

GFP barrel. The other possibility of eight to nine C-terminal amino acids detaching would require removal of half a β-strand (Fig. 2*b*). This possibility can be ruled out given the cooperative nature of β-strand breakage so far observed in mechanical experiments (5, 6). Here, we refer to the unfolding intermediate with the N-terminal α-helix detached as GFP$\Delta\alpha$. It is important to note that single-molecule force spectroscopy allows us to measure precisely the number of amino acid residues constituting the folded part of an unfolding intermediate. However, we have no indication whether the structures of these intermediates will be native-like, as we assume in our study.

Can we observe detachment of the N-terminal α-helix directly in our data? To address this question, we filtered our data in the rising slope of the GFP unfolding peak to increase the force resolution. Both in the data of Ig8–GFP and DdFLN1-5–GFP we now observe consistently a hump-like transition at forces between 30 and 40 pN (Fig. 3). This transition is absent in curves obtained with DdFLN1-5 and Ig8 constructs lacking the inserted GFP domain (5, 9). In force curves with DdFLN1-5–GFP, on average, two to three weaker Ig domains unfold before the GFP domain unfolds (Figs. 1*d*, 3*b*, and 4*a*). These unfolding events before GFP unfolding exhibit the same

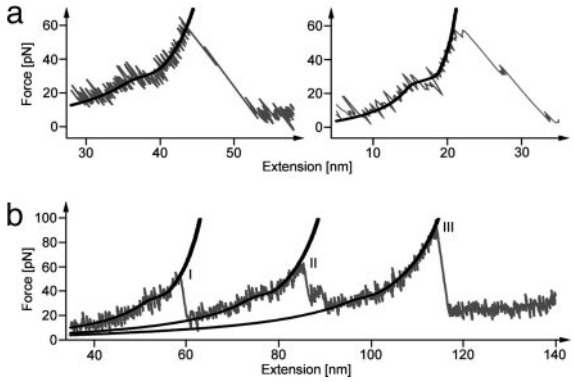

Fig. 3. Detachment of the N-terminal α-helix. (*a*) Gray lines show detail of the region before GFP$\Delta\alpha$ unfolding of two Ig8–GFP force-extension traces. Black lines shows a fit with an elastically coupled two-state model (for fit parameters see text) (*b*) Gray line shows detail of a DdFLN1-5–GFP force-extension trace. Peak I corresponds to the unfolding of a DdFLN domain. Peak II corresponds to the unfolding of DdFLN domain 4. Peak III corresponds to the unfolding of GFP$\Delta\alpha$. Black lines show a fit with an elastically coupled two-state model.

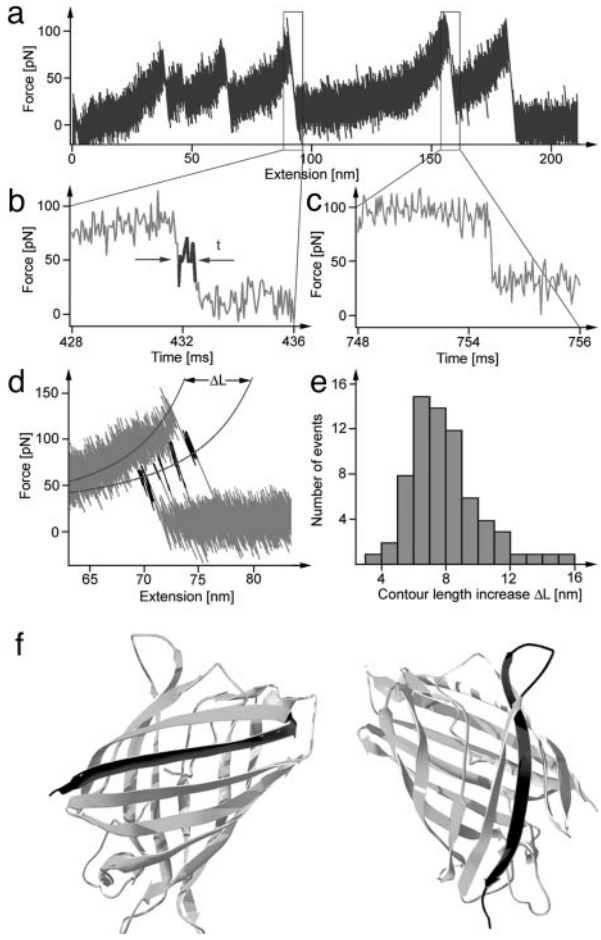

Fig. 4. Another short-lived intermediate. (*a*) DdFLN1-5–GFP force-extension trace measured at 10 kHz bandwidth with an Olympus type A Bio-Lever (resonance in water at 8.5 kHz). (*b*) Detail of the time course of the cantilever relaxation phase after unfolding of GFP$\Delta\alpha$. The region marked in black indicates the presence of a short-lived intermediate state. (*c*) Detail of the time course of the cantilever relaxation phase after unfolding of a DdFLN domain. No deviations from normal damped oscillator relaxation were detected. (*d*) Six force-extension traces measured with DdFLN1-5–GFP and superimposed in the region of GFP$\Delta\alpha$ unfolding. The regions marked in black indicate the short-lived intermediate. All six intermediate levels fall on a WLC curve with a contour length increased by 6.8 nm. (*e*) Distribution of the contour length increase, ΔL, determined by WLC fits as shown in *d*. (*f*) Two possible structures of GFP$\Delta\alpha\Delta\beta$ (light gray). Marked in dark gray are residues that unfold from the GFP barrel in the transition of GFP$\Delta\alpha$ to GFP$\Delta\alpha\Delta\beta$. (*Left*) A full β-strand detaches from the N terminus, and the folded GFP core comprises residues 25–230. (*Right*) A full β-strand detaches from the C terminus, and the folded GFP core comprises residues 11–209.

hump-like transition (Fig. 3*b*), indicating that the structure that gives during the transition recovers quickly as soon as the force drops below 20 pN (within <1 ms). The hump vanishes after GFP has unfolded. Near-equilibrium transitions are often characterized by hump-shaped force plateaus in force-extension traces (6, 19, 32), similar to the one we observe here. We used an equilibrium two-state model (19) to fit the hump-like transition (black curves in Fig. 3 *a* and *b*). The fit yields a contour length increase ΔL of 3.2 ± 1 nm ($n = 44$) and an equilibrium free energy for this transition of $\Delta G = 22 \pm 4\ k_BT$ ($n = 44$). A value of 3.2 ± 1 nm for the length increase during this transition is very close to the expected length gain upon unfolding of the N-terminal α-helix (2.9 nm) that is calculated from the GFP structure. This finding strongly supports our picture that the N-terminal α-helix detaches before the GFP barrel breaks. Remarkably, the energy required for the detachment

BIOPHYSICS

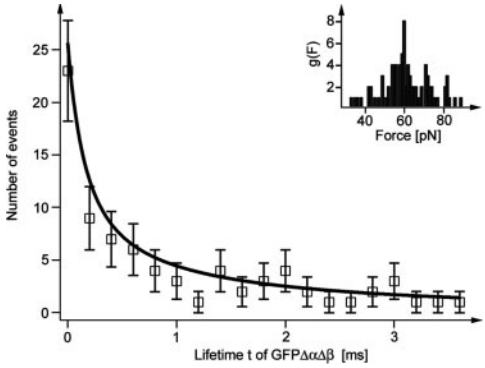

Fig. 5. Lifetime distribution. Open squares indicate the distribution of measured lifetimes of GFPΔαΔβ. The black solid line shows a fit with a two-state model with $\Delta x = 0.55$ nm and $\tau_0 = 10$ s (see *Materials and Methods*). (*Inset*) Distribution of forces acting on GFPΔαΔβ.

of the N-terminal α-helix is even slightly higher than the total free energy (16 k_BT) of the GFP fold (12). The GFP structure lacking the α-helix (GFPΔα) is therefore thermodynamically unstable. This finding is in good agreement with its short unloaded lifetime of 14 ± 6 s determined above from the distribution of unfolding forces. Moreover, deletion studies by Li *et al.* (33) of GFP have shown that removal of the N-terminal α-helix leads to nonfluorescent protein. Our results now provide an explanation for this result. Because GFPΔα has a positive folding free energy of at least 3.7 k_BT above the unfolded state, a construct lacking the α-helix will not fold into a barrel-like structure. Thermodynamic stability of GFP is dominated by the presence of the N-terminal α-helix. Mechanical stability on the subsecond timescale yet depends on the integrity of the β-barrel. For Ig domain 27 from titin, Marszalek *et al.* (6) have reported an unfolding intermediate during which a short β-strand detaches before the major mechanical unfolding event. In that study, however, the intermediate state governing mechanical stability was still thermodynamically stable (34). In contrast, our data demonstrate that mechanical experiments can be used to drive a protein into metastable states that are not populated thermally.

β-Strand Breakage Leads to Another Short-Lived Intermediate. In the data, when displayed with the full measurement bandwidth (10 kHz) of our cantilevers (Fig. 4*a*), we consistently observe deviations from the expected overdamped time course in the cantilever relaxation phase after the unfolding of GFPΔα (Fig. 4*b*). These deviations occur both in Ig8–GFP (see traces in Fig. 3*a*) and in DdFLN1-5–GFP (Fig. 4 *b* and *d*). We did not observe such deviations in the relaxation phase of two-state unfolders, like the Ig domains from titin or DdFLN (Fig. 4*c*). These deviations therefore indicate the existence of yet another short-lived intermediate state during GFP unfolding, which is stable only on the submillisecond timescale at the forces applied. In Fig. 4*d* we superimposed six traces exhibiting this intermediate level (marked in black) in their relaxation phase. A similar superposition strategy led to the detection of an unfolding intermediate in bacteriorhodopsin (35). Because the unfolding of GFPΔα has a broad force distribution (Fig. 1*e*), the intermediate level must also occur at different forces and extensions. Yet all six intermediate levels fall on a force extension curve predicted by the WLC of polymer elasticity, indicating that the structured portion of this additional intermediate also contains a well defined number of residues. A histogram of contour length increases during the transition from GFPΔα to the new short-lived intermediate shows an average lengthening of 6.8 ± 0.6 nm (Fig. 4*e*). Mapping this length increase to the structure of GFP (analogous to the method shown in Fig. 2*a*), we arrive at two possible structures for the intermediate state. Again, in our analysis we only consider possible structures in which a complete secondary structural element detaches from the barrel, and we thus disregard structures with partially frayed β-sheets at both ends. The two structures are shown in Fig. 4*f* and correspond to the GFP barrel lacking a full β-strand from either the N or the C terminus. Choosing between those two structures requires further experiments with GFP mutants. We will call the folded portion of this intermediate GFPΔαΔβ.

The observation of the short-lived intermediate GFPΔαΔβ in the relaxation phase of the cantilever sheds new light on the interpretation of unfolding forces commonly used in force spectroscopy experiments. First, it becomes obvious that the relaxation phase of the cantilever spring does not impose a "blind window" on our experiments as assumed in earlier studies (9). In contrast, every intermediate state exhibiting structural stability will interfere with the free relaxation of the cantilever and leave a signature in the force curve. Lifetimes of these intermediates can be extremely short, because they are subject to forces determined by the force of the previous unfolding event, the length increase, polymer elasticity, and the cantilever spring constant. We anticipate that increasing the measurement bandwidth into the microsecond timescale (36) or below will lead to the discovery of intermediate states, even in domains hitherto believed to be two-state unfolders.

Another aspect of these measurements concerns the analysis of unfolding forces for such short-lived intermediates. During a lifetime in the submillisecond range of the intermediate, the pulling experiment can be considered quasistatic. Therefore, the force will be approximately constant during this time. Consequently, unfolding of a short-lived intermediate must be treated as a constant force

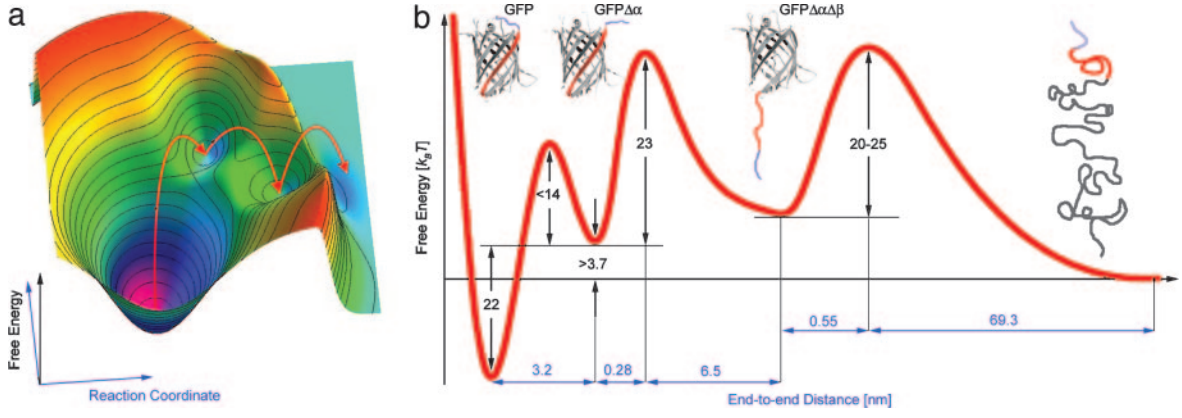

Fig. 6. Free energy surface. (*a*) Cartoon of the multidimensional energy landscape of GFP. The red arrows indicate the course of the mechanical unfolding pathway. (*b*) Projection of the energy landscape along the unfolding pathway onto one reaction coordinate.

experiment. Unfolding force histograms in such a case will lead to trivial force distributions because the force remains constant during the short lifetime of the intermediate. All of the important information now resides in the lifetime distribution of the intermediate state (37).

The lifetime distribution of GFP$\Delta\alpha\Delta\beta$ is given in Fig. 5. It appears strongly multiexponential. This distribution, however, does not reflect nonexponential protein unfolding kinetics but can be entirely explained by a single Poisson process. The multiexponential behavior is a consequence of the broad distribution of forces by which the structure is loaded until it breaks. This distribution (Fig. 5 *Inset*) is entirely governed by the unfolding force distribution of GFP$\Delta\alpha$ (Figs. 1*e* and 4*d*). The solid line in Fig. 5 is a superposition of lifetimes calculated for the measured force distribution. Our analysis yields values of 0.55 $\pm$ 0.1 nm for the position of the transition state Δx and a range of 1–100 s for the unloaded lifetime at zero force τ_0.

GFP Unfolding Is Governed by a Rough Energy Landscape. By combining our results, we are now able to reconstruct detailed features of a section through the energy landscape of GFP along a path determined by the application of mechanical force. A cartoon of this path is shown in Fig. 6*a*. Starting from the completely folded structure in the global minimum, the pathway is biased along the N-terminal-to-C-terminal direction. Upon unfolding of the N-terminal α-helix, we arrive at a local minimum (GFP$\Delta\alpha$). It is important to note that now the direction of the pathway changes, because the protein will respond to the detachment of the α-helix by turning and reorienting along the new N-terminal-to-C-terminal direction determined by the connecting line between the first structured residue at the N terminus and the last structured residue at the C terminus of the folded portion of the intermediate. The new bias induces detachment of a β-strand leading to the second local minimum (GFP$\Delta\alpha\Delta\beta$). Again, the protein will reorient and the direction of the path changes. From here, the protein will transition directly into the completely unfolded state, at least at the current temporal resolution of our experiment. In Fig. 6*b* we show a projection of the zig-zag-shaped unfolding pathway onto one coordinate. This projection summarizes our current picture of the barrier heights and positions along the mechanical GFP unfolding pathway. Yet, it is important to bear in mind that the reaction coordinate changes from minimum to minimum as explained above. For simplicity, the distances in this diagram contain our directly measured contour length increases upon detachment of amino acids. Of course, these numbers will be smaller at small stretching forces, at which the polymer will not be completely stretched.

The first transition from GFP to GFP$\Delta\alpha$ occurs close to equilibrium. From our data, we can extract a difference of $\Delta G = 22 \pm 4\,k_BT$ between GFP and GFP$\Delta\alpha$. The contour length increase upon detachment of seven amino acids is $\Delta L = 3.2 \pm 1$ nm. From our equilibrium data, we cannot conclude directly the position or height of the transition barrier. However, from our estimate for the refolding time of the α-helix of <1 ms we can estimate an upper limit for this barrier height of 14 k_BT.

The position of the barrier for the transition from GFP$\Delta\alpha$ to GFP$\Delta\alpha\Delta\beta$ is 0.28 $\pm$ 0.03 nm from the minimum, with an activation barrier of 23 k_BT. Detachment of the β-strand from the barrel leads to an additional contour length increase of 6.8 $\pm$ 0.6 nm. From the GFP$\Delta\alpha\Delta\beta$ state, GFP unfolds into the completely denatured state leading to a contour length increase of 69.8 nm. The associated barrier height is 20–25 k_BT at a position of 0.55 nm.

Conclusion

Our experimental results show that the mechanical unfolding of GFP proceeds by means of at least two metastable intermediate states, with unloaded lifetimes on the second timescale. The rough energy landscape we reconstructed from our data shows deep local minima reflecting largely structured, however metastable, intermediates with a positive free energy compared with the unfolded state. It appears that such intermediates are much more likely to occur in topologically complex proteins, like GFP, as compared with simpler, fast-folding two-state proteins. We propose that force spectroscopy is an ideal tool to investigate energy landscapes of complex proteins not amenable to bulk studies because of their slow kinetics or tendency of aggregation. The detailed picture we have obtained for the mechanical unfolding pathway will allow the tailoring of GFP mutants with altered stability that are useful for application as molecular force sensors. Such mechanical stability mutations have already been successfully realized with other proteins (31, 34, 38). A precise correlation of the breakdown of fluorescence with the breakdown of structure is an important task for the future.

We thank Ingo Schwaiger and Angelika Kardinal for help with instrumentation and cloning. This project was funded by Deutsche Forschungsgemeinschaft Grant SFB 413.

1. Clausen-Schaumann, H., Seitz, M., Krautbauer, R. & Gaub, H. E. (2000) *Curr. Opin. Chem. Biol.* **4,** 524–530.
2. Best, R. B. & Clarke, J. (2002) *Chem. Commun.* 183–192.
3. Fisher, T. E., Marszalek, P. E. & Fernandez, J. M. (2000) *Nat. Struct. Biol.* **7,** 719–724.
4. Li, H., Linke, W. A., Oberhauser, A. F., Carrion-Vazquez, M., Kerkvliet, J. G., Lu, H., Marszalek, P. E. & Fernandez, J. M. (2002) *Nature* **418,** 998–1002.
5. Schwaiger, I., Kardinal, A., Schleicher, M., Noegel, A. A. & Rief, M. (2004) *Nat. Struct. Mol. Biol.* **11,** 81–85.
6. Marszalek, P. E., Lu, H., Li, H., Carrion-Vazquez, M., Oberhauser, A. F., Schulten, K. & Fernandez, J. M. (1999) *Nature* **402,** 100–103.
7. Carrion-Vazquez, M., Marszalek, P. E., Oberhauser, A. F. & Fernandez, J. M. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11288–11292.
8. Williams, P. M., Fowler, S. B., Best, R. B., Toca-Herrera, J. L., Scott, K. A., Steward, A. & Clarke, J. (2003) *Nature* **422,** 446–449.
9. Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M. & Gaub, H. E. (1997) *Science* **276,** 1109–1112.
10. Zimmer, M. (2002) *Chem. Rev.* **102,** 759–781.
11. Tsien, R. (1998) *Annu. Rev. Biochem.* **67,** 509–544.
12. Fukuda, H., Arai, M. & Kuwajima, K. (2000) *Biochemistry* **39,** 12019–12024.
13. Bokman, S. H. & Ward, W. W. (1981) *Biochem. Biophys. Res. Commun.* **101,** 1372–1380.
14. Crameri, A., Whitehorn, E. A., Tate, E. & Stemmer, W. P. (1996) *Nat. Biotechnol.* **14,** 315–319.
15. Butt, H. J. & Jaschke, M. (1995) *Nanotechnology* **6,** 1–7.
16. Florin, E. L., Rief, M., Lehmann, H., Ludwig, M., Dornmair, C., Moy, V. T. & Gaub, H. E. (1995) *Biosens. Bioelectron.* **10,** 895–901.
17. Bustamante, C., Marko, J. F., Siggia, E. D. & Smith, S. (1994) *Science* **265,** 1599–1600.
18. Rief, M., Fernandez, J. M. & Gaub, H. E. (1998) *Phys. Rev. Lett.* **81,** 4764–4767.
19. Schwaiger, I., Sattler, C., Hostetter, D. R. & Rief, M. (2002) *Nat. Mater.* **1,** 232–235.
20. Fucini, P., Koppel, B., Schleicher, M., Lustig, A., Holak, T. A., Muller, R., Stewart, M. & Noegel, A. A. (1999) *J. Mol. Biol.* **291,** 1017–1023.
21. Fucini, P., Renner, C., Herberhold, C., Noegel, A. A. & Holak, T. A. (1997) *Nat. Struct. Biol.* **4,** 223–230.
22. Battistutta, R., Negro, A. & Zanotti, G. (2000) *Proteins* **41,** 429–437.
23. Bell, G. I. (1978) *Science* **200,** 618–627.
24. Schlierf, M., Li, H. & Fernandez, J. M. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 7299–7304.
25. Bieri, O., Wirz, J., Hellrung, B., Schutkowski, M., Drewello, M. & Kiefhaber, T. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 9597–9601.
26. Carrion-Vazquez, M., Oberhauser, A. F., Fisher, T. E., Marszalek, P. E., Li, H. & Fernandez, J. M. (2000) *Prog. Biophys. Mol. Biol.* **74,** 63–91.
27. Li, H., Oberhauser, A. F., Fowler, S. B., Clarke, J. & Fernandez, J. M. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 6527–6531.
28. Fisher, T. E., Carrion-Vazquez, M., Oberhauser, A. F., Li, H., Marszalek, P. E. & Fernandez, J. M. (2000) *Neuron* **27,** 435–446.
29. Evans, E. & Ritchie, K. (1997) *Biophys. J.* **72,** 1541–1555.
30. Oberhauser, A. F., Marszalek, P. E., Erickson, H. P. & Fernandez, J. M. (1998) *Nature* **393,** 181–185.
31. Carrion-Vazquez, M., Oberhauser, A. F., Fowler, S. B., Marszalek, P. E., Broedel, S. E., Clarke, J. & Fernandez, J. M. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 3694–3699.
32. Smith, S. B., Cui, Y. & Bustamante, C. (1996) *Science* **271,** 795–798.
33. Li, X., Zhang, G., Ngo, N., Zhao, X., Kain, S. R. & Huang, C. C. (1997) *J. Biol. Chem.* **272,** 28545–28549.
34. Fowler, S. B., Best, R. B., Toca Herrera, J. L., Rutherford, T. J., Steward, A., Paci, E., Karplus, M. & Clarke, J. (2002) *J. Mol. Biol.* **322,** 841–849.
35. Muller, D. J., Kessler, M., Oesterhelt, F., Moller, C., Oesterhelt, D. & Gaub, H. (2002) *Biophys. J.* **83,** 3578–3588.
36. Viani, M. B., Schäffer, T. E., Chand, A., Rief, M., Gaub, H. E. & Hansma, P. K. (1999) *J. Appl. Phys.* **86,** 2258–2262.
37. Oberhauser, A. F., Hansma, P. K., Carrion-Vazquez, M. & Fernandez, J. M. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 468–472.
38. Brockwell, D. J., Beddard, G. S., Clarkson, J., Zinober, R. C., Blake, A. W., Trinick, J., Olmsted, P. D., Smith, D. A. & Radford, S. E. (2002) *Biophys. J.* **83,** 458–472.