JMB

Mechanical Unfolding Intermediates Observed by Single-molecule Force Spectroscopy in a Fibronectin Type III Module

Lewyn Li, Hector Han-Li Huang, Carmen L. Badilla and Julio M. Fernandez*

Department of Biological Sciences, Columbia University New York, NY 10027, USA

Domain 10 of type III fibronectin (^{10}FNIII) is known to play a pivotal role in the mechanical interactions between cell surface integrins and the extracellular matrix. Recent molecular dynamics simulations have predicted that ^{10}FNIII, when exposed to a stretching force, unfolds along two pathways, each with a distinct, mechanically stable intermediate. Here, we use single-molecule force spectroscopy combined with protein engineering to test these predictions by probing the mechanical unfolding pathway of ^{10}FNIII. Stretching single polyproteins containing the ^{10}FNIII module resulted in sawtooth patterns where ^{10}FNIII was seen unfolding in two consecutive steps. The native state unfolded at $100(\pm20)$ pN, elongating ^{10}FNIII by $12(\pm2)$ nm and reaching a clearly marked intermediate that unfolded at $50(\pm20)$ pN. Unfolding of the intermediate completed the elongation of the molecule by extending another $19(\pm2)$ nm. Site-directed mutageneses of residues in the A and B β-strands (E9P and L19P) resulted in sawtooth patterns with all-or-none unfolding events that elongated the molecule by $19(\pm2)$ nm. In contrast, mutating residues in the G β-strand gave results that were dependent on amino acid position. The mutation I88P in the middle of the G β-strand resulted in native like unfolding sawtooth patterns showing an intact intermediate state. The mutation Y92P, which is near the end of G β-strand, produced sawtooth patterns with all-or-none unfolding events that lengthened the molecule by $17(\pm2)$ nm. These results are consistent with the view that ^{10}FNIII can unfold in two different ways. Along one pathway, the detachment of the A and B β-strands from the body of the folded module constitute the first unfolding event, followed by the unfolding of the remaining β-sandwich structure. Along the second pathway, the detachment of the G β-strands is involved in the first unfolding event. These results are in excellent agreement with the sequence of events predicted by molecular dynamics simulations of the ^{10}FNIII module.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: fibronectin; correlation effects; AFM; mechanical unfolding; single-molecule

*Corresponding author

Introduction

Fibronectin (FN) is a giant dimeric protein of more than 40 modules. FN modules belong to three

distinct structural types, and type III FN modules (FNIII) are immunoglobulin-fold β-sandwiches that contain binding sites for cell surface receptors. As an important component of the extracellular matrix (ECM), FN is involved in tissue elasticity, cell adhesion and cell migration. Since the ECM is under mechanical stress *in vivo*, it is natural to ask if the force-induced deformation of FN is important to the ECM. Three independent lines of research support this view. First, fluorescence studies show that cells extend FN fibrils up to four times their

equilibrium length.[1] Recent experiments suggest that at least one module in the FN fibril unfolds when cells are stretched.[2] Second, mechanical tension is an important physiological trigger for cells to build the ECM.[3–7] Various studies support the existence of "cryptic binding sites" in FN.[7–10] Elongation of FN is believed to expose the cryptic binding sites, which are necessary for FN polymerization and fibril formation. Finally, cell adhesion requires the binding of the RGD-loop on domain 10 of FNIII ([10]FNIII) to transmembrane integrin receptors.[11] Integrin-[10]FNIII binding creates a network that mechanically connects the cytoskeleton to the ECM. Interestingly, the binding force between [10]FNIII and integrin has been reported to be ≈ 60 pN,[12] which is similar to the unfolding force of [10]FNIII (≈ 75 pN).[13] So, [10]FNIII unfolding may be coupled to integrin unbinding, if the direction of the applied force is similar in unfolding and binding.[14,15]

The [10]FNIII module of fibronectin (Figure 1(a)) has a characteristic seven-stranded β-sandwich tertiary structure.[16] The hairpin loop between the F and G β-strands includes the RGD sequence that plays a key role in integrin binding. The effect of a mechanical force on the structure of the [10]FNIII module has been modeled extensively. Molecular dynamics simulations reported by Paci & Karplus[17] and Gao *et al*.[18] have predicted the existence of an unfolding intermediate, which contains extended A and B-strands and an intact RGD-loop. Paci & Karplus observed another unfolding intermediate in [10]FNIII, where the A and G-strands have detached from the remainder of the protein.[17] In order to test these predictions and to gain further insights into [10]FNIII unfolding, we have engineered polyprotein chimeras containing native and mutant forms of the [10]FNIII module. We have examined the mechanical properties of these polyprotein chimeras using single-molecule force spectroscopy techniques.[19–23] Our results are consistent with the view that, under an applied mechanical force, [10]FNIII unravels in two different ways. Along the first pathway, there is an unfolding intermediate in [10]FNIII that results from the detachment of the A and B β-strands, in good agreement with computer simulations. Along the second pathway, the G β-strand dissociates from the folded module in the first unfolding step, again in agreement with computer simulations.

Results

Mechanical unfolding of polyprotein chimeras containing the [10]FNIII module

Atomic-force microscopy (AFM) is a useful method to investigate the mechanical unfolding of a protein at the single-molecule level,[19–27] and has stimulated many studies with analytical theory[28–30] and computer simulations.[17,18,31] We use protein engineering to construct polyproteins made of
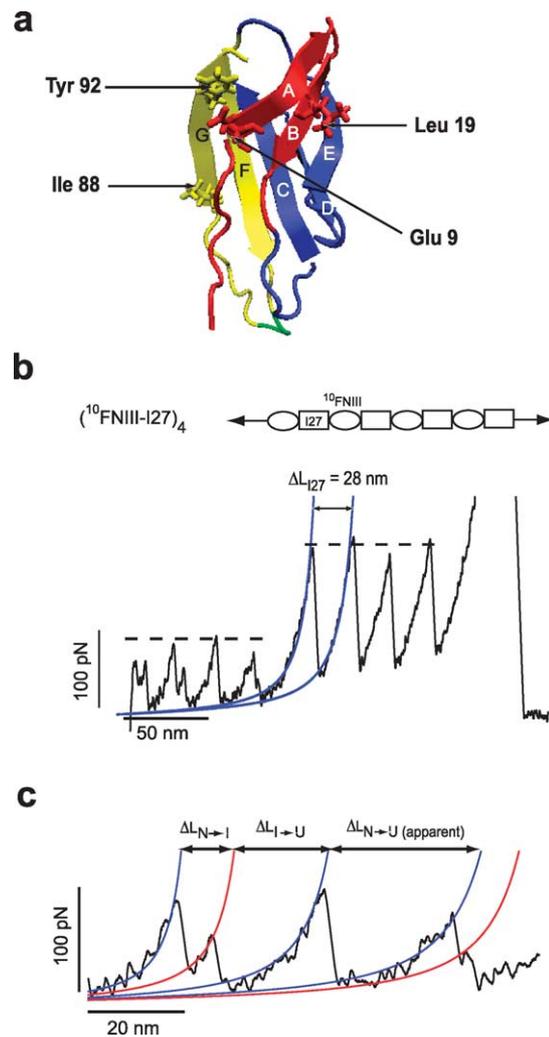
Figure 1. Domain 10 of type III fibronectin ([10]FNIII) and the chimera approach allows the unambiguous identification and characterization of [10]FNIII unfolding. (a) The three-dimensional structure of [10]FNIII (from Main *et al*.[16]) and the amino acids mutated in this study. The A and B-strands are in red. The C, D, and E-strands are in blue, and the F and G-strands are in yellow. The RGD-loop (Arg78, Gly79 and Asp80) is shown in green. (b) The wild-type ([10]FNIII-I27)$_4$ chimera gave a regularly spaced sawtooth pattern. The last four peaks are the I27 fingerprints, with a difference in contour length (ΔL_c) ≈ 28 nm and an unfolding force ≈ 200 pN. So, the first four peaks, with an unfolding force ≈ 90 pN, represent the unraveling of the [10]FNIII modules. The blue lines are worm-like chain (WLC) fits to the I27 modules. (c) Force-extension curve for three-state and apparently two-state unfolding of wild-type [10]FNIII. Blue and red lines are WLC fits to the native state and the unfolding intermediate, respectively.

chimeras of [10]FNIII and the I27 immunoglobulin module of titin. The polyprotein chimeras allow for the unambiguous identification of the engineered polyprotein from a background of spurious interactions. When a polyprotein chimera is picked up and pulled with an AFM, the resulting sawtooth

pattern contains the unique fingerprint of I27 unfolding: i.e. unfolding peaks of 200 pN with a contour length increase of 28 nm after every peak.[13, 20] In addition to the I27 fingerprint, the sawtooth pattern contains the mechanical unfolding events of the protein module under study (e.g. ^{10}FNIII), arranged hierarchically. Furthermore, the presence of two or more I27 unfolding events guarantees that the unfolding of at least one ^{10}FNIII protein module is included in the observed sawtooth pattern.[20] These techniques allow for a positive identification of the unfolding and extension of the ^{10}FNIII, even in such cases when mutagenesis or other manipulations render ^{10}FNIII mechanically unstable. Figure 1(b) shows the design of the polyprotein chimeras used throughout this study. A typical force-extension relationship that results from mechanically stretching a polyprotein chimera (^{10}FNIII-I27)$_4$ is shown in the trace below. In this example, the four clearly distinguished I27 unfolding peaks are preceded by four equally spaced unfolding events that undoubtedly correspond to the mechanically weaker ^{10}FNIII.

Wild-type ^{10}FNIII unfolds with a mechanically stable intermediate

Figure 1(b) and (c) show a typical force-extension curve of the wild-type (^{10}FNIII-I27)$_4$ chimera. In 138 out of 216 unfolding events, ^{10}FNIII unravels in two distinct steps: native-to-intermediate (N→I) and intermediate-to-unfolded (I→U), producing a doublet of peaks for each ^{10}FNIII unfolding (Figure 1(c)). We also observed 26 apparently isolated I→U transitions, which suggests that parts of ^{10}FNIII may already be unfolded at zero force.[13] The first peak in Figure 1(c) represent transitions from the native to the intermediate state, with average $\Delta L_{N \to I} = \langle \Delta L_{N \to I} \rangle = 12(\pm 2)$ nm (Figure 2(a)). The second peak in Figure 1(c) correspond to transitions from the intermediate to the unfolded state, with $\langle \Delta L_{I \to U} \rangle = 19(\pm 2)$ nm (Figure 2(a)). Importantly, the total contour length ($= \Delta L_{N \to I} + \Delta L_{I \to U}$) increases by 31($\pm 4$) nm, in agreement with the value of 32.5 nm expected for a ^{10}FNIII module (94 residues × 0.38 nm/residue minus the folded end-to-end distance of 3.2 nm). The agreement proves that the doublet of peaks corresponds to a single ^{10}FNIII molecule.

In 78 of 216 unfolding events for ^{10}FNIII, we cannot detect a distinct intermediate-to-unfolded peak (third peak in Figure 1(c)). We designate these events as apparently two-state and label them as N→U (apparent) in Figure 2 and Table 1. The lack of a clear intermediate-to-unfolded peak does not necessarily mean that ^{10}FNIII has a choice between two unfolding pathways (two-state *versus* three-state). It is possible that ^{10}FNIII has only a three-state pathway and that, due to the stochastic nature of mechanical unfolding, the intermediate occasionally unravels at too low a force to be seen. Further work is required to determine if ^{10}FNIII has a choice between two-state and three-state pathways. Reassuringly, $\Delta L_{N \to U \ (apparent)} = \Delta L_{N \to I} + \Delta L_{I \to U}$ (Figure 2(a)), proving that the apparent two-state events still represent the unfolding of ^{10}FNIII modules. The average unfolding forces are (Figure 2(b)):

$$\langle F_{N \to I} \rangle = 90(\pm 20) \text{ pN}$$

$$\langle F_{I \to U} \rangle = 50(\pm 20) \text{ pN}$$

$$\langle F_{N \to U \ (apparent)} \rangle = 100(\pm 20) \text{ pN}$$

The unfolding forces for the native state of ^{10}FNIII in (^{10}FNIII-I27)$_4$, reported here, are slightly higher than the value of 75 pN in (^{10}FNIII)$_8$.[13] So, the I27 module may have stabilized ^{10}FNIII against mechanical stresses. Similar stabilization was found when the I28 module was stretched in the (I28)$_8$ and (I27–I28)$_4$ chimera.[20]

a **b**
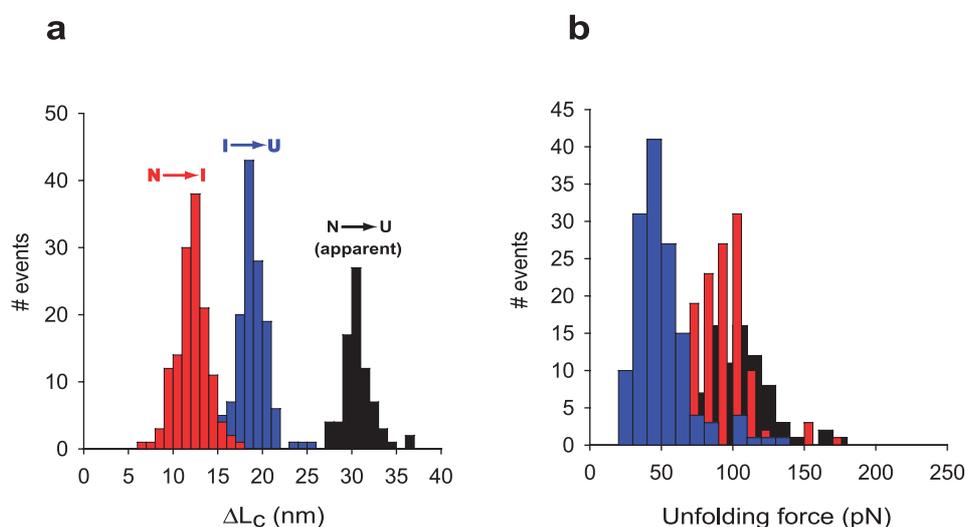
Figure 2. Histograms of the contour length increment (ΔL_c) and unfolding forces of wild-type ^{10}FNIII. (a) Histogram for ΔL_c. (b) Histogram for the unfolding forces, with the same color scheme as (a).

Table 1. The average change in contour length ($\langle\Delta L_c\rangle$), unfolding force ($\langle F\rangle$) and kinetic parameters for wild-type ^{10}FNIII and proline mutants

Protein	Transition	N	$\langle\Delta L_c\rangle$ (nm)[a]	$\langle F\rangle$ (pN)[a]	k_0 (s^{-1})[b]	Δx (nm)[c]
Wild-type	N→U (apparent)	78	31 ± 2	100 ± 20	–	–
	N→I	138	12 ± 2	90 ± 20	0.02[d]	0.38[d]
	I→U	138	19 ± 2	50 ± 20	0.5	0.45
E9P mutant	I→U	103	19 ± 2	70 ± 40	–	–
I88P mutant	N→U (apparent)	94	30 ± 2	100 ± 20	–	–
	N→I	132	12 ± 2	90 ± 20	–	–
	I→U	132	19 ± 2	50 ± 20	–	–
Y92P mutant	I→U	12	17 ± 2	60 ± 30	–	–

[a] Average $\pm$ standard deviation.
[b] Zero-force unfolding rate coefficient using the model described by Rief *et al.*[46]
[c] Distance to the transition state using the model described by Rief *et al.*[46]
[d] Adopted from Oberhauser *et al.*[13]

Structural mapping of the unfolding intermediate and pathway

Molecular dynamics simulations have identified two unfolding intermediates in ^{10}FNIII. Paci & Karplus[17] and Gao *et al.*[18] have predicted the existence of an unfolding intermediate, which contains extended A and B-strands, and an intact RGD-loop. Paci & Karplus observed another unfolding intermediate in ^{10}FNIII, where the A and G-strands have detached from the remainder of the protein.[17] In order to test these predictions, we selectively disrupted the A, B and G β-strands with proline mutations. Proline mutations delete backbone hydrogen bonds and restrict backbone rotation about the N–C$^\alpha$ bond.[21,32] E9P removes the backbone H-bond between Glu9 (A-strand) and Ser21 (B-strand); L19P eliminates the H-bond between Leu19 (B-strand) and Val11 (A-strand). So, E9P and L19P are expected to disrupt the H-bond network between the A and B-strands. I88P deletes the backbone H-bond between Ile88 (G-strand) and Val72 (F-strand); Y92P removes the backbone H-bond between Tyr92 (G-strand) and Tyr68 (F-strand). So, I88P and Y92P should disturb the H-bond network between the F and G-strands (Figure 1(a)). The mutation Y92P should also test the importance of the C terminus in fibronectin unfolding: Hamill *et al.* have demonstrated that extending the C terminus of ^{3}FNIII from human tenascin by only two residues reduced the unfolding rate 40-fold.[33]

Disrupting A and B-strands brought ^{10}FNIII to a mechanically stable intermediate state

The E9P mutant unfolds at an average force of $70(\pm40)$ pN (Figure 3(a) and (d)). The unfolding forces of the L19P mutant are similar (Figure 3(b)). Strikingly, E9P and L19P reduce ΔL_c from $31(\pm2)$ nm (=fully folded wild-type ^{10}FNIII) to $19(\pm2)$ nm (Figure 3(c)). So, E9P and L19P appear to have unfolded and extended the protein by $12(\pm3)$ nm at zero force (the "zero-force extension"). Considering the positions of Glu9 and Leu19

in ^{10}FNIII (Figure 1(a)), the simplest explanation for the zero-force extension is that E9P or L19P has unraveled the A and B-strands in the absence of force, leaving only strands C to G intact. In support of this explanation, there are 27 residues from the N terminus to the end of the B-strand, which yields $\Delta L_c = 10.3$ nm, consistent with the zero-force extension.

Importantly, the wild-type unfolding intermediate has an extension of $\langle\Delta L_{N\to I}\rangle = 12(\pm2)$ nm (Figure 2(a)), which agrees with the zero-force extension of the E9P mutant ($=12(\pm3)$ nm). This result is consistent with the prediction of an A-B-stretched intermediate from molecular dynamics simulations.[17,18]

Disrupting the middle of the G-strand has little mechanical effect

The mutation I88P in the middle of the G-strand has no significant effect on ^{10}FNIII unfolding (Figure 4(a) and Table 1). We found 16 apparently isolated I→U transitions. $\Delta L_{N\to U}$ (apparent), $\Delta L_{N\to I}$ and $\Delta L_{I\to U}$ of the I88P mutant are identical with the wild-type (Figure 5(c)), and the mutation has a negligible effect on the unfolding forces (Figure 5(d)).

Disrupting the C-terminal end of the G-strand resulted in another mechanical stable intermediate

The mutation Y92P near the C terminus of ^{10}FNIII (Figure 1(a)) resulted in regular sawtooth patterns with $\Delta L_c = 17(\pm2)$ nm and an average unfolding of $60(\pm30)$ pN (Figure 4(b) and Table 1). These results are consistent with the idea that Y92P has converted ^{10}FNIII to a mechanically stable intermediate state. Our data on Y92P indicate that ^{10}FNIII could unfold *via* an intermediate containing a detached G-strand, in agreement with the computer simulations reported by Paci & Karplus.[17] However, from our data, we cannot conclude if the Y92P consists of detached A and G-strands, as predicted by Paci & Karplus.[17]. Stretching A and G-strands would give
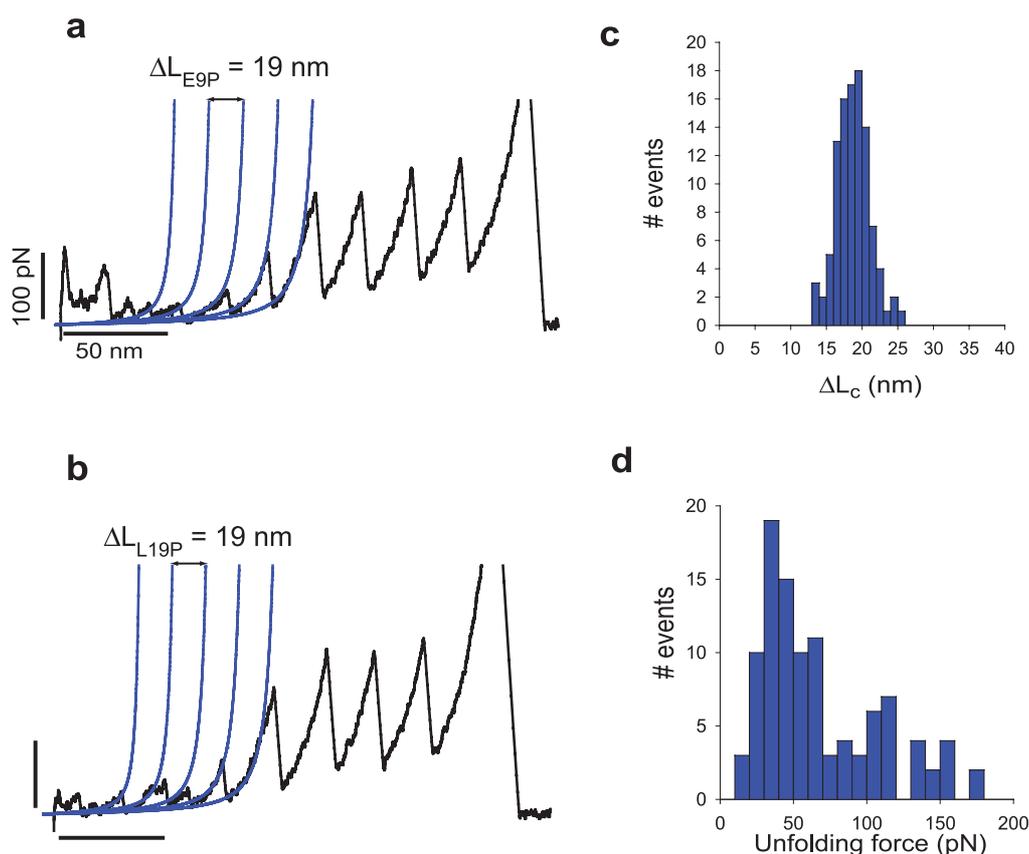
Figure 3. The unfolding of ^{10}FNIII$_{E9P}$ and ^{10}FNIII$_{L19P}$. (a) Force-extension curve of (^{10}FNIII$_{E9P}$-I27)$_4$. The blue lines are worm-like chain (WLC) fits to ^{10}FNIII$_{E9P}$ (b) Force-extension curve of (^{10}FNIII$_{L19P}$-I27)$_4$, with spacing identical with that for ^{10}FNIII$_{E9P}$ (c) Histogram for ΔL_c of ^{10}FNIII$_{E9P}$ (d) Unfolding force histogram for ^{10}FNIII$_{E9P}$.

a similar extension of ≈ 12 nm as stretching F and G-strands. Therefore, either an A-G-detached intermediate or an F-G-detached intermediate is consistent with the observed extension of $17(\pm 2)$ nm.

Instead of proline mutation, the insertion of polyglycine into various loops of a protein could also map the structure of an unfolding intermediate.[23,27] An intermediate similar to that in ^{10}FNIII has been reported in domain 4 of the actin-crosslinking protein filamin, using polyglycine insertion and single-molecule AFM.[27] A loop extension of five glycine residues would increase the contour length extension of the intermediate by only 1.9 nm ($= 5 \times 3.8$ Å) per ^{10}FNIII molecule. On the other hand, the E9P, L19P and Y92P mutations reduced the contour length increment of each ^{10}FNIII module by > 10 nm, which was detected much more readily. So, in ^{10}FNIII at least, proline mutations yielded conclusive results that are identified easily in AFM.

Detailed picture of ^{10}FNIII unfolding

Our data support the idea that there are two ways ^{10}FNIII could unfold. In the first way, the protein unfolds initially by unraveling A and B-strands, bringing the system to a mechanically stable intermediate. The remaining C to G-strands then unfold in the subsequent step. This sequence is in broad agreement with the predictions made by Paci & Karplus[17] and by Gao *et al.*[18] In the second way, ^{10}FNIII unfolding begins with stretching the G-strand, lengthening the protein by ≈ 12 nm and converting the protein to an intermediate state. The existence of a G-stretched intermediate is in accordance with the predictions made by Paci & Karplus.[17]

Correlation between the native state and the intermediate rupture forces

Whenever there are two distinct unfolding force peaks for ^{10}FNIII, we observed a correlation between $F_{N \to I}$ and $F_{I \to U}$, with a correlation coefficient of $R = 0.6$ (Figure 5(d)). To further explore this issue, we performed Monte Carlo (MC) simulations assuming that $N \to I$ and $I \to U$ are sequential and independent Markov processes (see Methods for details). We found that $F_{N \to I}$ and $F_{I \to U}$ in the MC simulations displayed a weak correlation of $R = 0.2$ (Figure 5(b)). The reason why the correlation coefficient is non-zero in the MC simulations is explained in Discussion. The unraveling of the native state and the intermediate in ^{10}FNIII are separated by ≈ 8 nm ($=$ the x-axis separation between the first and second peaks in Figure 5(c)), which translates to a time delay of $\tau \approx 20$ ms at the experimental pulling rate of 400 nm s^{-1}.
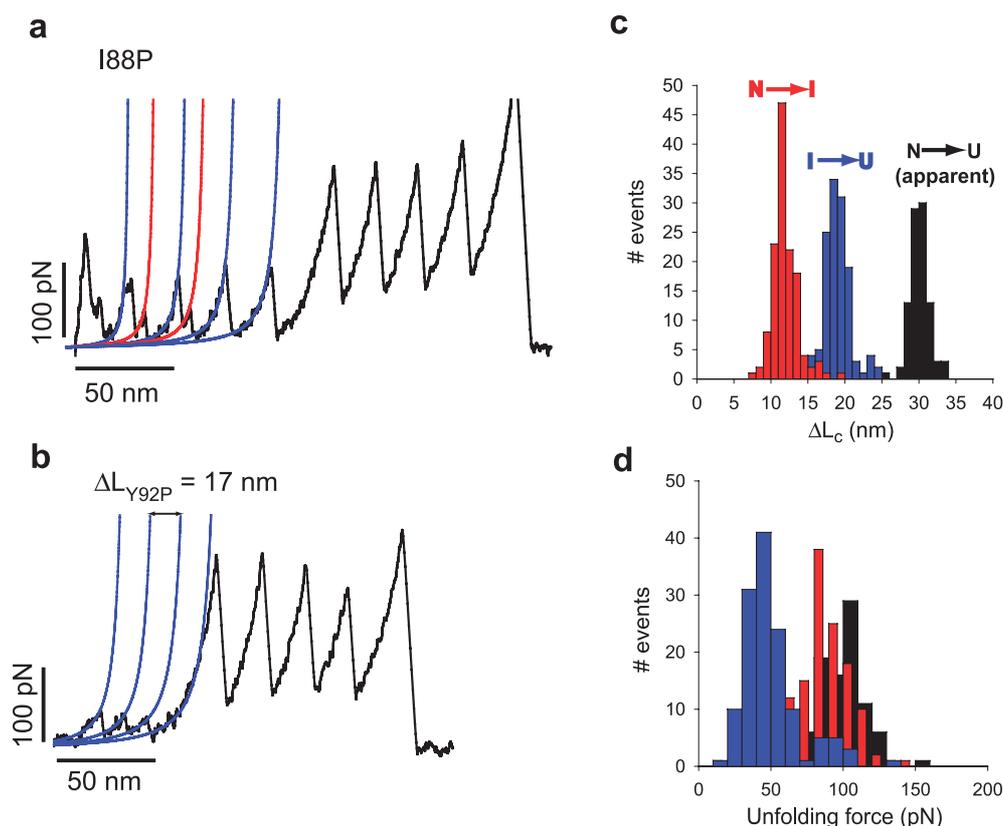
Figure 4. The unfolding of ^{10}FNIII$_{I88P}$ and ^{10}FNIII$_{Y92P}$. (a) Force-extension curve of (^{10}FNIII$_{I88P}$-I27)$_4$, with the same color scheme as Figure 1(c). (b) Force-extension curve of (^{10}FNIII$_{Y92P}$-I27)$_4$. (c) Histogram of ΔL_c for ^{10}FNIII$_{I88P}$. (d) Unfolding force histogram of ^{10}FNIII$_{I88P}$, with the same color scheme as (c).

We superimposed three separate experimental force-extension curves of three-state unfolding events of ^{10}FNIII in Figure 5(c). The three curves correspond to the unfolding of the first ^{10}FNIII in three (^{10}FNIII-I27)$_4$ chimeras. The force decreases to near-zero after the first peak, because the protein chain is still short after the first ^{10}FNIII module unravels. All three experimental curves (red, black and blue) obey the same worm-like chain (WLC) fit. Mechanical protein unfolding is a stochastic process and occurs within a range of forces with significant probability.[19] Because the mechanical force grows as the protein extends, a native state that breaks at a higher force (red * in Figure 5(c)) is necessarily more extended than a native state that breaks at a lower force (black * in Figure 5(c)). The same reasoning applies to the unfolding intermediate (black and red # in Figure 5(c)). So, the unfolding peaks cannot all occur at the same protein extension, and therefore do not align vertically in Figure 5(c). Reassuringly, the same WLC fits all three curves, proving that the force-extension behavior of the protein can be described by a single model, regardless of the unfolding force. More importantly, an extrapolation of WLC to infinite force gives the increase in contour length between the native state and the unfolding intermediate ($\Delta L_{N \rightarrow I}$). We found that $\langle \Delta L_{N \rightarrow I} \rangle = 12(\pm 2)$ nm (Figure 2(a)), in agreement with molecular

dynamics simulations[17,18] and our own data on the proline mutants.

Discussion

Physiological importance of the unfolding intermediates in ^{10}FNIII

There are two ways in which the A-B-detached unfolding intermediate could be physiologically relevant. First, the intermediate state could protect the ^{10}FNIII-integrin interaction against mechanical shock, by unfolding the A and B-strands without sacrificing the RGD-loop. Second, the intermediate state could expose cryptic binding sites that are necessary for FN polymerization. This possibility is supported by two recent studies. Hocking *et al.* reported the presence of a conformation-dependent binding site for ^{1}FNIII in the N-terminal region of ^{10}FNIII.[8] In another study, anastellin, which consists of ^{1}FNIII minus the A and B-strands, induces fibril formation *in vitro*.[34] The N-terminal A and B-strands are exposed in the A-B-detached intermediate and may contain the cryptic binding site. It is possible that the site is located in the hydrophobic core of ^{10}FNIII, which is partially exposed in the intermediate.

A specific experiment will distinguish between
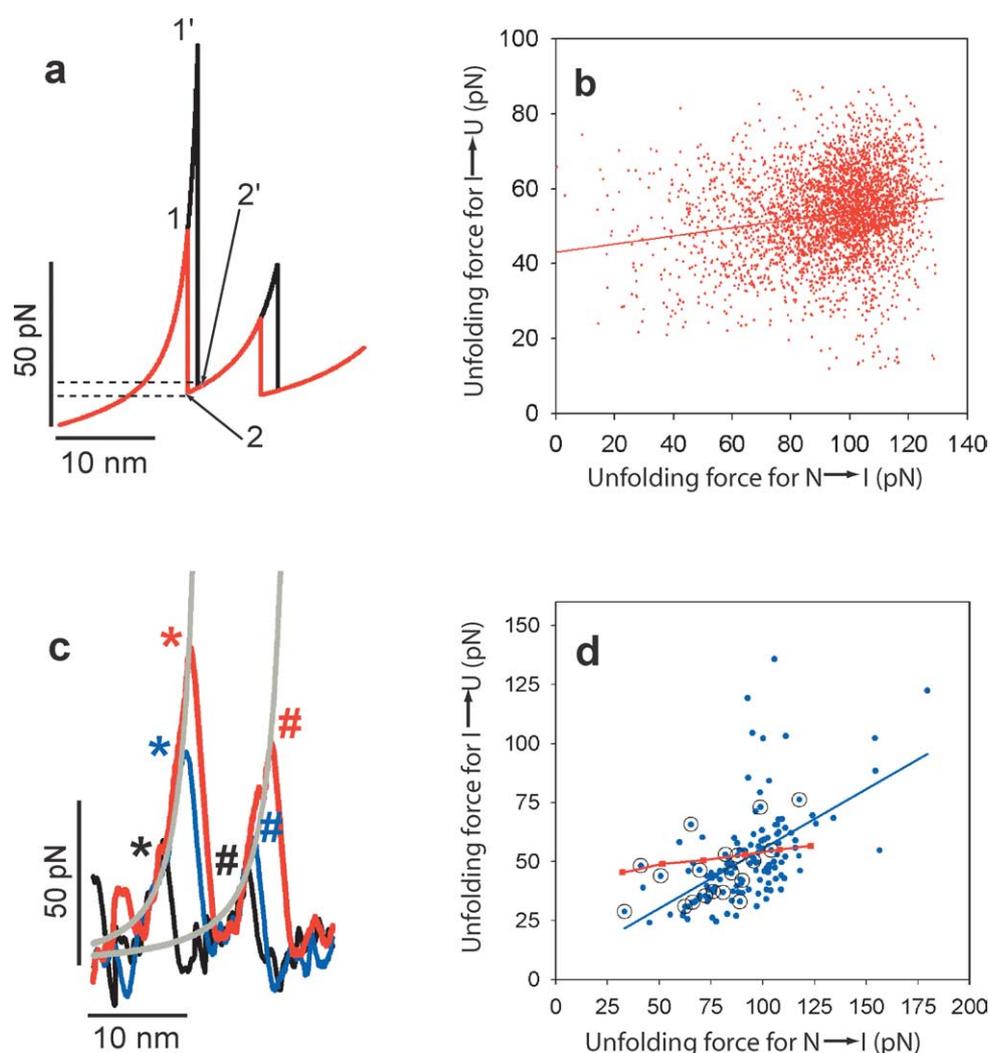
Figure 5. Correlation between the rupture forces of the native and the intermediate state in wild-type [10]FNIII. (a) Superimposition of two separate three-state unfolding events in Monte Carlo (MC) simulations. The residual strain is larger when the first unfolding peak is higher, as shown by the fact that point 2′ lies above point 2. This is probably the reason why there is a non-zero correlation between the N→I and I→U forces in MC simulations. See Discussion for details. (b) MC simulations of [10]FNIII unfolding with two-state Markovian kinetics. The abscissa and coordinate represent the N→I and I→U unfolding forces, respectively, from MC simulations. The linear regression correlation coefficient is $R=0.2$ ($N=3659$). See Methods for details. (c) Superimposition of three separate three-state unfolding events in actual AFM unfolding of [10]FNIII. The three traces correspond to the unfolding of the first [10]FNIII in three ([10]FNIII-I27)$_4$ chimeras. The force decreases to near-zero after the first peak because the protein chain is still short after the first [10]FNIII module unravels. A higher N→I force (*) is correlated with a higher I→U force (#). The grey lines are worm-like chain (WLC) fits. (d) Correlation between the N→I and I→U unfolding forces. The blue circles are the observed unfolding forces ($N=138$), and the first [10]FNIII unfolding in a chimera is additionally circled in black. The straight line is the result of linear regression of all 138 experimental unfolding forces ($R=0.6$). The red squares are the averaged unfolding forces from MC simulations using two-state Markovian kinetics. The discrepancy between the blue and red lines shows that simple two-state Markovian kinetics is inconsistent with the observed correlation, and that there may be additional sources of correlation in [10]FNIII unfolding. Each square is generated by averaging the data in bins of 20 pN for $F_{N→I}$.

the above two hypotheses: mutation of Glu9 to proline in the natural FN dimer. If the unfolding intermediate is a "shock absorber" that protects the RGD-loop, then E9P would reduce cell adhesion, because the mutation abolishes the native-to-intermediate transition (Figure 3(a)). On the other hand, if the intermediate reveals the cryptic binding site in [10]FNIII, then E9P is expected to enhance cell adhesion as well as [10]FNIII–[1]FNIII binding, because

the mutation converts [10]FNIII to a state similar to the unfolding intermediate (Figure 3(a)).

The reaction coordinate in [10]FNIII unfolding

In [10]FNIII unfolding, the stronger link ($F_{N→I}=90(\pm20)$ pN) breaks before the weaker link ($F_{I→U}=50(\pm20)$ pN). This "reverse hierarchy" means that the unfolding of the native state and the

intermediate might have different reaction coordinates. If a single reaction coordinate (i.e. a single set of collective interactions such as hydrogen bonds and hydrophobic interactions among the protein atoms) governs the unfolding of the native state and the intermediate, we would expect the weaker link to break first. The possibility of two distinct reaction coordinates for ^{10}FNIII unfolding is perhaps unsurprising: the proline mutant data have shown that the unfolding of the native state involves the A, B and G-strands, whereas the unraveling of the intermediate involves β-strands not yet stretched after the first unfolding barrier has been overcome.

According to the conventional picture of a multistate reaction, the applied force distorts the free energy landscape linearly by $-F \cdot x$, where F is the applied force and x is the reaction coordinate.[26,35] This picture assumes that the unraveling of the native state and the intermediate share the same reaction coordinate. In light of our results on ^{10}FNIII, the assumption of a single reaction coordinate may not be as applicable as generally assumed in mechanical unfolding of proteins. For example, stretching a polyubiquitin chain along different force vectors caused the protein to explore distinct unfolding pathways and transition states, resulting in different mechanical stabilities.[14]

Possible causes for the correlation between the native state and the intermediate rupture forces

We observed a weak correlation of $R = 0.2$ between $F_{N \to I}$ and $F_{I \to U}$ in the MC simulations (Figure 5(b)). If the sequential unfolding of the native state and intermediate of ^{10}FNIII are independent events in the MC simulations, why is the correlation non-zero? The reason may lie in the residual strain after the native state has unfolded. If the native state unravels at a low force (=point 1 in Figure 5(a)), the protein relaxes to the WLC of the intermediate and a low residual strain (=point 2 in Figure 5(a)). An intermediate unfolding at any force higher than point 2 will produce a noticeable force peak, but an intermediate unfolding at a force lower than the residual strain will not give a force peak and therefore cannot be detected. If the native state unravels at a higher force (=point 1′ in Figure 5(a)), the residual strain is higher (=point 2′ in Figure 5(a)) and the intermediate can be detected only at forces exceeding this higher residual strain value. This could explain why there is a weak positive correlation between $F_{N \to I}$ and $F_{I \to U}$ in the MC simulations. Part of the correlation observed in experiments with ^{10}FNIII probably comes from the effect described above.

However, the correlation between the native state and the intermediate rupture forces from actual experiments ($R = 0.6$ in Figure 5(d)) is three times as strong as the correlation from MC simulations. So, there appear to be additional factors responsible for the correlation in actual ^{10}FNIII unfolding. One possibility is multiple non-interconverting

unfolding pathways, where some pathways have higher barriers for the unfolding of the native state and the intermediate, and the other pathways have lower barriers. It is noteworthy that detecting multiple unfolding pathways has been a longstanding challenge in ensemble protein folding,[36] and single-molecule experiments such as those described here have the potential to address this challenge in a novel and unambiguous way.

Another possible explanation for Figure 5(d) is that there is only a single unfolding pathway but, due to large thermal fluctuations, different ^{10}FNIII molecules are found in states with different mechanical stabilities, and that the unfolding intermediates from different ^{10}FNIII molecule do not interconvert within the time-scale of the experiment. We shall call this scenario "mechanical memory". The mechanically strong ^{10}FNIII molecule would then have high values for $F_{N \to I}$ and $F_{I \to U}$, and the weak ^{10}FNIII would unravel easily in both $N \to I$ and $I \to U$. In the course of unfolding, if the intermediate has enough time to relax to steady-state, then the correlation between $F_{N \to I}$ and $F_{I \to U}$ should decrease to $R = 0.2$, the value observed in the MC simulations. This is because relaxation to steady-state would destroy any memory between the native state and the intermediate, and decouple $N \to I$ and $I \to U$ into two independent reactions, as they are in the MC simulations. Therefore, Figure 5(d) is consistent with the notion that, during unfolding, the intermediate state is out of equilibrium.

Using single-molecule techniques, more and more proteins have been discovered to retain memory of their structures:[37–42] these proteins seem to be trapped transiently in distinct conformational substates that interconvert within a correlation time-scale of milliseconds to seconds. Because the function of a protein depends critically on its structure, there are growing interests and efforts to decipher the origin and biological importance of protein memory. For example, millisecond-to-second memory has been reported in single-enzyme kinetics of staphylococcal nuclease,[37] cholesterol oxidase,[38] flavin reductase[39] and horseradish peroxidase.[40] Different metastable conformational substates of the enzyme have been proposed to possess different catalytic activities.[37–40] Moreover, memory is not restricted to enzyme kinetics, and has been reported in the chemical (un)folding of individual two-stranded coiled coil GCN4-Pf[41] and adenylate kinase,[42] as well as the docking and undocking of ribozyme.[43] Many studies reveal a wide range of correlation times from 1 ms to >1000 ms.[38–42] Our results here on ^{10}FNIII raise the tantalizing prospect that there may be memory effects in the mechanical unfolding of proteins as well. So, memory effects could be a general mechanism for Nature to fine-tune the structure, and thereby the function, of important biological molecules.

Methods

Construction of the wild-type and mutant ^{10}FNIII-I27 chimera

We build the $(^{10}\text{FNIII-I27})_4$ chimeras with published methods.[19] Site-directed mutations are introduced by standard PCR techniques. Proteins are expressed in the recombination-defective BLR (DE3) strain (Novagene), purified by Ni^{2+} or Co^{2+} affinity chromatography and stored at 4 °C in phosphate buffered saline (PBS) with 5 mM DTT and 0.2 mM EDTA.

Atomic-force microscopy

Details of the custom-made AFM apparatus have been described.[19–21] Calibration of the cantilevers (Si_3N_4 from Digital Instruments, Inc.) is done in solution using the equipartition theorem.[44] The typical string constant is ~ 50 mN m^{-1}, and the typical pulling rate is 0.4 µm s^{-1}.

Data analysis

The force-extension curves are fitted to the WLC model of polymer elasticity.[45] We use a constant persistence length of $p = 0.4$ nm. The increase in contour length (ΔL_c) between two WLC fits is caused by partial or complete unfolding of a protein module.

Monte Carlo simulations

We model the unfolding of wild-type ^{10}FNIII as two independent and sequential Markov processes:[46]

$$N \xrightarrow{k_{N \to I}} I \xrightarrow{k_{N \to I}} U$$
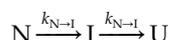
where:

$$k_{N \to I} = k_{0(N \to I)} \exp(F\Delta x_{N \to I}/kT)$$

$$k_{I \to U} = k_{0(I \to U)} \exp(F\Delta x_{I \to U}/kT)$$

k_0 is the zero-force rate coefficient, F is the applied force determined by the WLC model with $p = 0.4$ nm, and $\Delta x_{A \to B}$ is the distance between state A and state B in the dimension of the reaction coordinate. L_c is calculated from the number of folded ($L_c = 34$ nm), intermediate ($L_c = 16$ nm) and unfolded ($L_c = 4$ nm) modules. At each time-step, the probabilities of the N→I and I→U transitions are $P_{N \to I} = N_f k_{N \to I} \Delta t$ and $P_{I \to U} = N_I k_{I \to U} \Delta t$, respectively. N_f is the number of folded modules, N_I is the number of intermediates and $\Delta t = 10^{-6}$ seconds. We neglect U→I and I→N, because protein refolding is suppressed completely in our experiments.[46] $k_{0(N \to I)}$ and $\Delta x_{N \to I}$ in Table 1 are from Oberhauser et al.[13] We fit $k_{0(I \to U)}$ and $\Delta x_{I \to U}$ to the average value and standard deviation of $F_{I \to U}$. The k_0 and Δx values in Table 1 are only approximate, because the estimation of k_0 and Δx by MC simulations has a low level of sensitivity.[19]

Acknowledgements

References

1. Ohashi, T., Kiehart, D. P. & Erickson, H. P. (1999). Dynamics and elasticity of the fibronectin matrix in living cell culture visualized by fibronectin-green fluorescent protein. *Proc. Natl Acad. Sci. USA*, **96**, 2153–2158.
2. Baneyx, G., Baugh, L. & Vogel, V. (2002). Fibronectin extension and unfolding within cell matrix fibrils controlled cytoskeletal tension. *Proc. Natl Acad. Sci. USA*, **99**, 5139–5143.
3. Halliday, N. L. & Tomasek, J. J. (1995). Mechanical properties of the extracellular matrix influence fibronectin assembly *in vitro*. *Expt. Cell Res.* **217**, 109–117.
4. Shaub, A. (1999). Unraveling the extracellular matrix. *Nature Cell Biol.* **1**, E173–E175.
5. Schwarzbauer, J. E. & Sechler, J. L. (1999). Fibronectin fibrillogenesis: a paradigm for the extracellular matrix assembly. *Curr. Opin. Cell Biol.* **11**, 622–627.
6. Hynes, R. O. (1999). The dynamic dialogue between cells and matrices: implications of fibronectin's elasticity. *Proc. Natl Acad. Sci. USA*, **96**, 2588–2590.
7. Zhong, C., Chrzanowska-Wodnicka, M., Brown, J., Shaub, A., Belkin, A. M. & Burridge, K. (1998). Rho-mediated contractility exposes a cryptic site in fibronectin and induces fibronectin matrix assembly. *J. Cell Biol.* **141**, 539–551.
8. Hocking, D. C., Smith, R. K. & McKeown-Longo, P. J. (1996). A novel role for the integrin binding III-10 module in fribronectin matrix assembly. *J. Cell Biol.* **133**, 431–444.
9. Sechler, J. L., Rao, H., Cumiskey, A. M., Vega-Colon, I., Smith, M. S., Murata, T. & Schwarzbauer, J. E. (2001). A novel fibronectin binding site required for fibronectin fibril growth during matrix assembly. *J. Cell Biol.* **154**, 1081–1088.
10. Ingham, K. C., Brew, S. A., Huff, S. & Litvinovitch, S. V. (1997). Cryptic self association sites in type III modules of fibronectin. *J. Biol. Chem.* **272**, 1718–1724.
11. Ruoslahti, E. & Pierschbacher, M. D. (1987). New perspectives in cell adhesion: RGD and integrins. *Science*, **238**, 491–497.
12. Li, F., Redick, S. D., Erickson, H. P. & Moy, V. T. (2003). Force measurements of the α5β1 integrin–fibronectin interaction. *Biophys. J.* **84**, 1252–1262.
13. Oberhauser, A. F., Badilla-Fernandez, C., Carrion-Vazquez, M. & Fernandez, J. M. (2002). The mechanical hierarchies of fibronectin observed with single-molecule AFM. *J. Mol. Biol.* **319**, 433–447.
14. Carrion-Vazquez, M., Li, H., Lu, H., Marszalek, P. E., Oberhauser, A. F. & Fernandez, J. M. (2003). The mechanical stability of ubiquitin is linkage dependent. *Nature Struct. Biol.* **10**, 738–743.
15. Brockwell, D. J., Paci, E., Zinober, R. C., Beddard, G. S., Olmsted, P. D., Smith, D. A. *et al*. (2003). Pulling geometry defines the mechanical resistance of a β-sheet protein. *Nature Struct. Biol.* **10**, 731–737.
16. Main, A. L., Harvey, T. S., Baron, M., Boyd, J. & Campbell, I. D. (1992). The three dimensional structure of the tenth type III module of fibronectin: an insight into RGD-mediated interactions. *Cell*, **71**, 671–678.
17. Paci, E. & Karplus, M. (1999). Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations. *J. Mol. Biol.* **288**, 441–459.
18. Gao, M., Craig, D., Vogel, V. & Schulten, K. (2002). Identifying unfolding intermediates of FN-III$_{10}$ by steered molecular dynamics. *J. Mol. Biol.* **323**, 939–950.

19. Carrion-Vazquez, M., Oberhauser, A. F., Fowler, S. B., Marszalek, P. E., Broedel, S. E., Clarke, J. & Fernandez, J. M. (1999). Mechanical and chemical unfolding of a single protein: a comparison. *Proc. Natl Acad. Sci. USA*, **96**, 3694–3699.

20. Li, H., Oberhauser, A. F., Fowler, S. B., Clarke, J. & Fernandez, J. M. (2000). Atomic force microscopy reveals the mechanical design of a modular protein. *Proc. Natl Acad. Sci. USA*, **97**, 6527–6531.

21. Marszalek, P. E., Lu, H., Li, H., Carrion-Vazquez, M., Oberhauser, A. F., Schulten, K. J. & Fernandez, J. M. (1999). Mechanical unfolding intermediates in titin modules. *Nature*, **402**, 100–103.

22. Rief, M., Gautel, M., Oesterholt, F., Fernandez, J. M. & Gaub, H. E. (1997). Reversible unfolding of individual titin immunoglobulin modules by AFM. *Science*, **276**, 1109–1112.

23. Carrion-Vazquez, M., Marszalek, P. E., Oberhauser, A. F. & Fernandez, J. M. (1999). Atomic force microscopy captures length phenotypes in single proteins. *Proc. Natl Acad. Sci. USA*, **96**, 11288–11292.

24. Furuike, S., Ito, T. & Yamazaki, M. (2001). Mechanical unfolding of single filamin A (ABP-280) molecules detected by atomic force microscopy. *FEBS Letters*, **498**, 72–75.

25. Law, R., Carl, P., Harper, S., Dalhaimer, P., Speicher, D. W. & Discher, D. E. (2003). Cooperativity in forced unfolding of tandem spectrin repeats. *Biophys. J.* **84**, 533–544.

26. Williams, P. M., Fowler, S. B., Best, R. B., Toca-Herrera, J. L., Scott, K. A., Steward, A. & Clarke, J. (2003). Hidden complexity in the mechanical properties of titin. *Nature (London)*, **422**, 446–449.

27. Schwaiger, I., Kardinal, A., Schleicher, M., Noegel, A. A. & Rief, M. (2004). A mechanical unfolding intermediate in an actin-crosslinking protein. *Nature Struct. Mol. Biol.* **11**, 81–85.

28. Geissler, P. L. & Shakhnovich, E. I. (2002). Mechanical response of random heteropolymers. *Macromolecules*, **35**, 4429–4436.

29. Hummer, G. & Szabo, A. (2001). Free energy reconstruction from nonequilibrium single molecule pulling experiments. *Proc. Natl Acad. Sci. USA*, **98**, 3658–3661.

30. Hyeon, C. & Thirumalai, D. (2003). Can energy landscape roughness of protein and RNA be measured by using mechanical unfolding experiments? *Proc. Natl Acad. Sci. USA*, **100**, 10249–10253.

31. Klimov, D. K. & Thirumalai, D. (2000). Native topology determines force-induced unfolding pathways in globular proteins. *Proc. Natl Acad. Sci. USA*, **97**, 7254–7259.

32. Creighton, T. E. (1993). *Proteins: Structures and Molecular Properties* (2nd edit.), W.H. Freeman and Company, New York.

33. Hamill, S. J., Meekhof, A. E. & Clarke, J. (1998). The effect of boundary selection on the stability and folding of the third fibronectin type III domain from human tenascin. *Biochemistry*, **37**, 8071–8079.

34. Briknarová, K., Åkerman, M. E., Hoyt, D. W. & Ruoslahti, E. (2003). Anastellin, an FN3 fragment with fibronectin polymerization activity, resembles amyloid precursors. *J. Mol. Biol.* **332**, 205–215.

35. Evans, E. (1998). Energy landscapes of biomolecular adhesion and receptor anchoring at interfaces explored with dynamic force spectroscopy. *Faraday Discuss.* **111**, 1–16.

36. Wright, C. F., Lindorff-Larsen, K., Randles, L. G. & Clarke, J. (2003). Parallel protein unfolding pathways revealed and mapped. *Nature Struct. Biol.* **10**, 658–662.

37. Ha, T., Ting, A. Y., Liang, J., Caldwell, W. B., Deniz, A. A., Chemla, D. S. *et al*. (1999). Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proc. Natl Acad. Sci. USA*, **96**, 893–898.

38. Lu, H. P., Xun, L. & Xie, X. S. (1998). Single-molecule enzymatic dynamics. *Science*, **282**, 1877–1882.

39. Yang, H., Luo, G., Karnchanaphanurach, P., Louie, T.-M., Rech, I., Cova, S. *et al*. (2003). Protein conformational dynamics probed by single-molecule electron transfer. *Science*, **302**, 262–266.

40. Edman, L. & Rigler, R. (2000). Memory landscape of single-enzyme molecules. *Proc. Natl Acad. Sci. USA*, **97**, 8266–8271.

41. Talaga, D. S., Lau, W. L., Roder, H., Tang, J., Jia, Y., DeGrado, W. F. & Hochstrasser, R. M. (2000). Dynamics and folding of single two-stranded coiled-coil peptides studies by fluorescent energy transfer confocal microscopy. *Proc. Natl Acad. Sci. USA*, **97**, 13021–13026.

42. Rhoades, E., Gussakovsky, E. & Haran, G. (2003). Watching proteins fold one molecule at a time. *Proc. Natl Acad. Sci. USA*, **100**, 3197–3202.

43. Zhuang, X., Kim, H., Pereira, M. J., Babcock, H. P., Walter, N. G. & Chu, S. (2002). Correlating structural dynamics and function in single ribozyme molecules. *Science*, **296**, 1473–1476.

44. Florin, E. L., Rief, M., Lehmann, H., Ludwig, M., Dornmair, C., Moy, V. T. & Gaub, H. E. (1995). Sensing specific molecular interactions with the atomic force microscope. *Biosens. Bioelectr.* **10**, 895–901.

45. Marko, J. F., Siggia, E. D. & Stretching, D. N. A. (1995). *Macromolecules*, **28**, 8759–8770.

46. Rief, M., Fernandez, J. M. & Gaub, H. E. (1998). Elastically coupled two-level systems as a model for biopolymer extensibility. *Phys. Rev. Letters*, **81**, 4764–4767.

Edited by W. Baumeister