

Single protein misfolding events captured by atomic force microscopy

Andres F. Oberhauser, Piotr E. Marszalek, Mariano Carrion-Vazquez and Julio M. Fernandez

Department of Physiology and Biophysics, Mayo Foundation, Rochester, Minnesota 55905, USA.

Using single protein atomic force microscopy (AFM) techniques we demonstrate that after repeated mechanical extension/relaxation cycles, tandem modular proteins can misfold into a structure formed by two neighboring modules. The misfolding is fully reversible and alters the mechanical topology of the modules while it is about as stable as the original fold. Our results show that modular proteins can assume a novel misfolded state and demonstrate that AFM is able to capture, in real time, rare misfolding events at the level of a single protein.

Individually folded protein modules, such as the immunoglobulin and fibronectin type III (FN III) modules, are placed in tandem in a wide variety of proteins^{1,2}. These tandem modular proteins are thought to undergo repeated cycles of mechanical unfolding and refolding events during physiological activity, while preserving their tertiary structure^{3–8}. However, given the tendency of polymers to spontaneously form a random coil, it is unlikely that after unfolding all its modules, a mechanically extended protein can always refold correctly.

Recent experiments using single molecule atomic force microscopy (AFM) techniques have demonstrated that tandem modular proteins can be forced to undergo many

unfolding/refolding cycles^{3,4,9}. Here we examine the fidelity of mechanical refolding by repeatedly unfolding and refolding a tandem repeat protein constructed of eight repeats of an immunoglobulin domain. We have used recombinant DNA techniques to construct eight direct tandem repeats of the I27 domain of human cardiac titin (I27^{RS}, Fig. 1a). The three-dimensional structure of the I27 module from the cardiac I-band titin has been determined by NMR¹⁰ and shown to consist of 89 amino acids that fold into a characteristic β -sandwich with seven strands². Construction of a polyprotein with tandem repeats of the I27 module is advantageous for single molecule AFM because stretching the polyprotein generates a periodic sawtooth pattern that amplifies the features of the individual modules and allows for measurements of contour length with a high signal:noise ratio⁹. Stretching the I27^{RS} polyprotein resulted in force/extension curves with peaks that varied randomly in amplitude with an average value of ~ 200 pN (Fig. 1b).

Mechanically unfolded I27 polyproteins readily refold when relaxed to their resting length. In a typical experiment (Fig. 1b), a single polyprotein remained attached to an AFM tip allowing for repeated extension and relaxation cycles following a double-pulse protocol⁹ (Fig. 1b; inset). Three consecutive traces are shown. A first extension of the protein revealed six regularly spaced force peaks corresponding to the unfolding of six identical I27 immunoglobulin modules (Fig. 1b, trace i). Then, by moving the tip of the AFM cantilever back, the protein was relaxed to its original length (Fig. 1b, trace ii). Upon relaxation, the domains refolded and a subsequent extension showed that the refolded modules unfolded again at a similar force (Fig. 1b, trace iii). This extension/relaxation cycle can be repeated hundreds of times with the same protein.

During repeated extension/relaxation cycles, the sawtooth pattern of some of the force/extension curves showed missing force peaks (Fig. 1c,d). In one example (Fig. 1c), a first extension

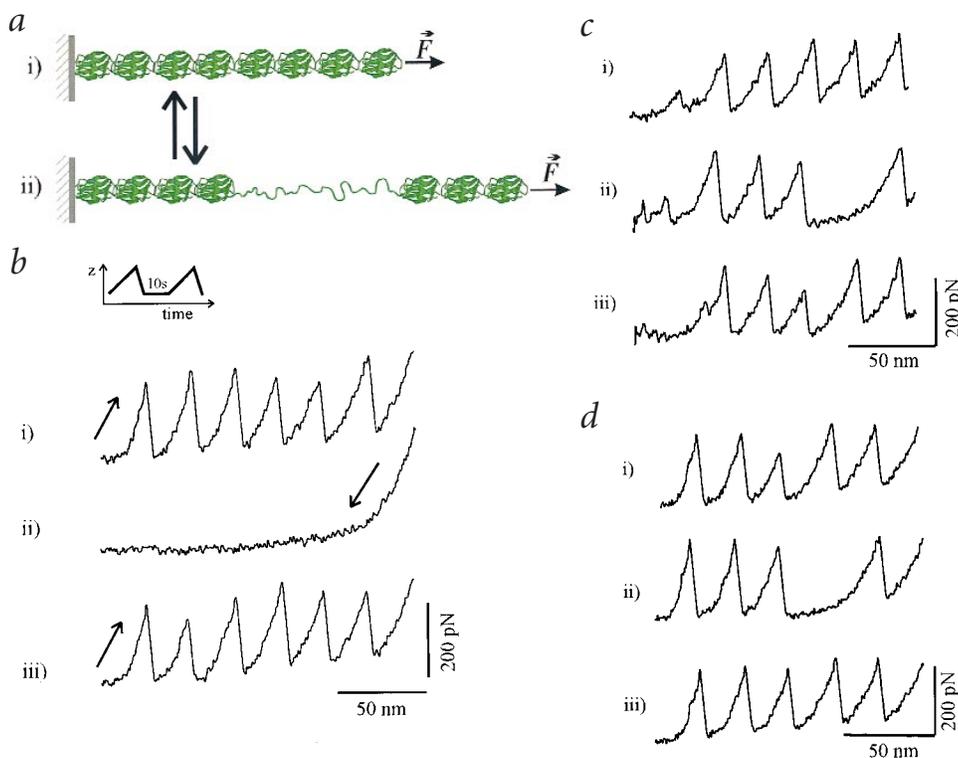


Fig. 1 Reversible misfolding events captured by single protein AFM recordings of a polyprotein composed of tandem repeats of an immunoglobulin domain. **a**, Schematic diagram of a I27^{RS} polyprotein. (i) When a force is applied to the folded polyprotein (ii) the domains unfold. Upon relaxation, the protein rapidly regains its folded conformation. **b**, Refolding kinetics probed with a double-pulse protocol (inset). The traces show force/extension curves obtained by consecutively extending (sawtooth pattern; trace i), relaxing (trace ii) and after 10 s in the relaxed state, re-extending (sawtooth pattern; trace iii) an I27^{RS} polyprotein. The traces show that all of the modules that unfolded in the first pull, spontaneously refolded during the relaxed period. **c**, **d**, During these repeated unfolding/refolding cycles, the sawtooth pattern of the force/extension curves occasionally show 'skips,' for example, missing force peak in trace ii in (c) and (d). The missing unfolding events observed during the refolding cycles occur with a frequency of about 2% (13 skips/608 refolding cycles; 46 separate experiments).

letters

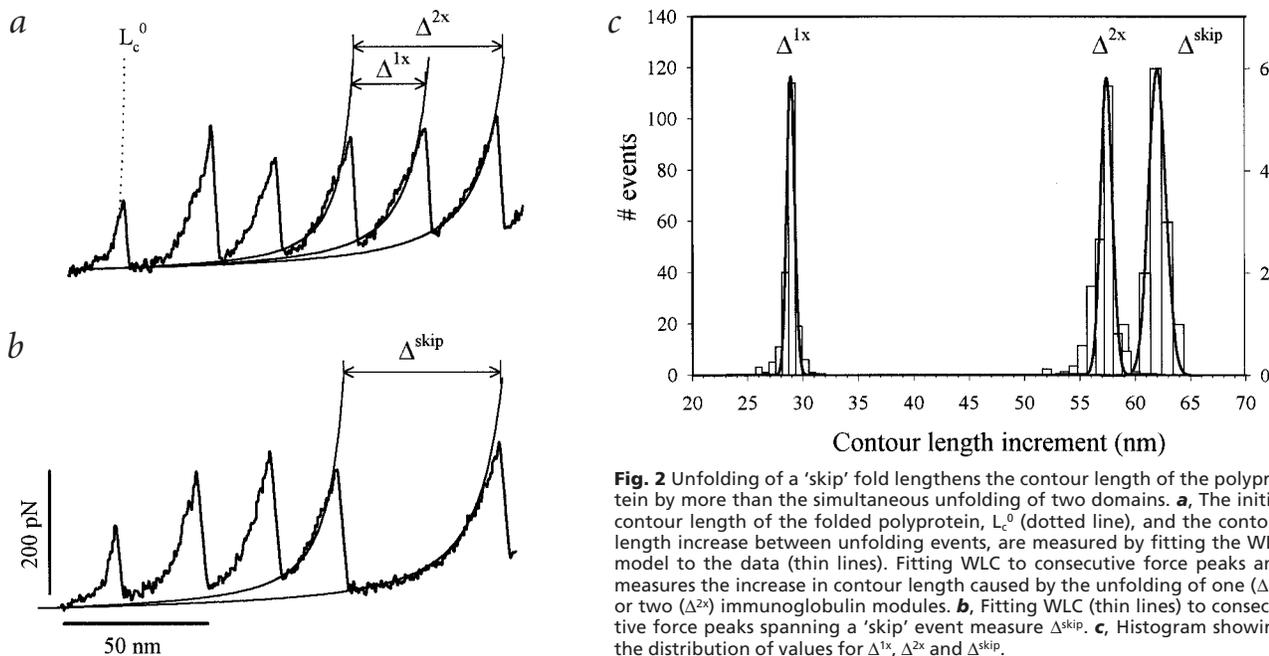


Fig. 2 Unfolding of a 'skip' fold lengthens the contour length of the polyprotein by more than the simultaneous unfolding of two domains. **a**, The initial contour length of the folded polyprotein, L_c^0 (dotted line), and the contour length increase between unfolding events, are measured by fitting the WLC model to the data (thin lines). Fitting WLC to consecutive force peaks and measures the increase in contour length caused by the unfolding of one (Δ^{1x}) or two (Δ^{2x}) immunoglobulin modules. **b**, Fitting WLC (thin lines) to consecutive force peaks spanning a 'skip' event measure Δ^{skip} . **c**, Histogram showing the distribution of values for Δ^{1x} , Δ^{2x} and Δ^{skip} .

(Fig. 1c, trace i) shows five peaks. A second pull, after the protein was relaxed to its resting length and allowed to refold, shows only four peaks (Fig. 1c, trace ii; the 'skip' occurred after the third peak). Again the protein was allowed to refold to the relaxed state. A third pull (Fig. 1c, trace iii) again reveals five peaks, showing that the protein recovered its original number of folds. A similar example is shown in Fig. 1d. Hence, the 'skip' event corresponds to a fold that contributes to a much larger increase in contour length than that corresponding to the unfolding of a single I27 module.

In order to understand the origin of the 'skip' events, we must examine the mechanical topology of the I27 module, which can be obtained from the force/extension relationships of an I27 polyprotein. The mechanical topology of the I27 module can be described by two sets of amino acids: (i) hidden amino acids that do not experience the applied force because they are mechanically isolated by the fold, and (ii) force-bearing amino acids that support the applied force and are excluded from the body of the fold. The force-bearing amino acids form the contour length of the resting, folded protein.

The force/extension curves for I27^{RS}₈ are well described by the wormlike chain model (WLC) of elasticity (thin lines, Fig. 2a,b), which predicts the entropic restoring force generated upon the extension of a polymer^{3,4,9,11}. Fitting WLC to the force/extension curve that precedes the first force peak gave a measure of the contour length of the protein when all its modules are still folded, L_c^0 (dotted line, Fig. 2a). The size of an individual folded module can be calculated as $L_c^0/m = 4.6 \pm 1.7$ nm ($n = 21$), where m corresponds to the number of folded modules in the protein segment that was picked up by the AFM tip ($3 < m < 8$). The size of the folded module measured by AFM is similar to the folded length of the I27 module determined by NMR (4.4 nm/module; ref. 12). By contrast, the hidden amino acids become exposed only after an unfolding event takes place and they determine the contour length increment observed upon unfolding. Fitting WLC to the force/extension curve between consecutive peaks measures the contour length increment of the protein when a

module unfolds, $\Delta^{1x} = 28.5 \pm 0.6$ nm ($n = 200$; Fig. 2a). Thus, if we consider that each amino acid contributes 0.38 nm to the contour length of a protein, then we calculate that the folded protein has 12 ± 4 force-bearing amino acids. Unfolding of a module exposes 75 ± 2 hidden amino acids that become force bearing. This simplified mechanical topology can be used to examine the 'skip' events. Fitting WLC to the consecutive force peaks of a 'skip' event shows a contour length increment of $\Delta^{skip} = 61.4 \pm 1.3$ nm ($n = 13$; Fig. 2b). This contour length increase is ~ 4.4 nm larger than the increase expected for the simultaneous unfolding of two modules ($2 \times \Delta^{1x} = 57$ nm; Fig. 2a,c). Hence, a 'skip' event reveals a fold that contains twice the number of hidden amino acids and contains an extra length of ~ 4.4 nm. This added length compares well with the contour length of a folded domain (~ 4.6 nm). These results indicate that a 'skip' fold corresponds to the fold of two neighboring I27 modules, where the normally force bearing amino acids linking these modules are hidden by becoming part of the 'skip' fold. These observations also predict that a folded protein that harbors a 'skip' fold should have a resting length that is ~ 4.4 nm shorter than that of a normally folded protein. Indeed, this can be easily observed by comparing the sawtooth patterns obtained, during and after a 'skip' event. A superposition of the sawtooth patterns obtained before and during a 'skip' event (Fig. 3) reveals that the force peaks that occur before the 'skip' are shifted to the left by $\delta = 6.2 \pm 1$ nm ($n = 6$). By contrast, the unfolding events that occur after the 'skip' event superimpose on the unfolding events observed in the control (Fig. 3). Hence, a folded protein that harbors a misfolded domain has a resting contour length that is shorter by ~ 6 nm, until the 'skip' fold unravels. The observed shift of ~ 6 nm is larger than the expected ~ 4.4 nm. This discrepancy may result from twisting of the protein or other structural rearrangements caused by the misfold. In the example shown in Fig. 3c the traces before (in black) and after (in blue) superimpose perfectly on the trace containing a 'skip' event (in green). This indicates that the 'skip' fold is a reversible structure that readily converts into the normal fold.

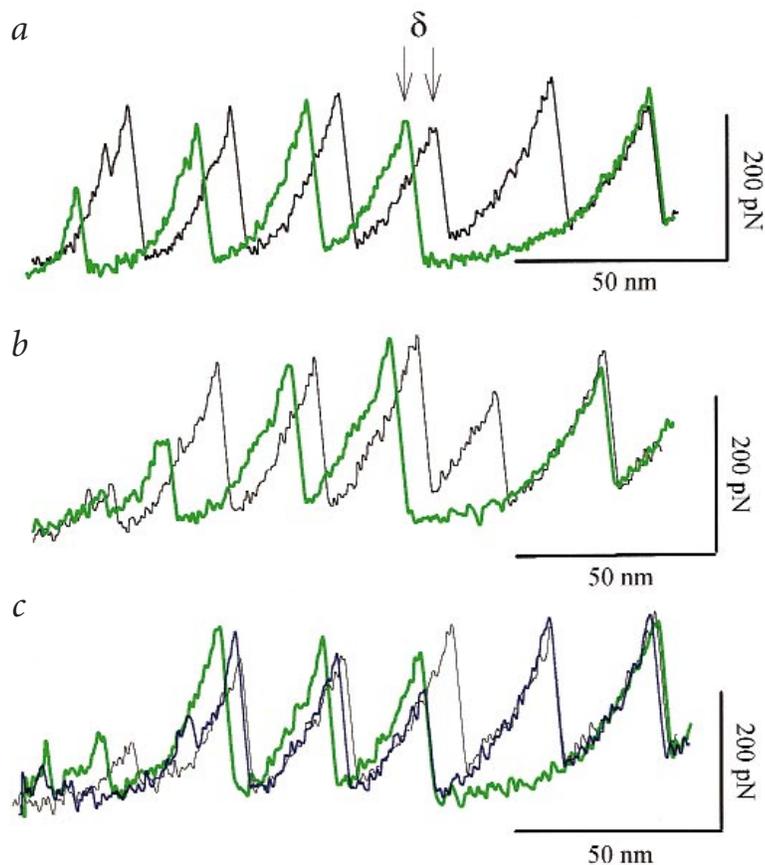


Fig. 3 Folded polyproteins that harbor a 'skip' fold are shorter than correctly folded ones. **a–c**, Superimposition of force/extension curves of control traces (in black) and traces containing missing force peaks (in green) during unfolding/refolding cycles. The force peaks that precede a 'skip' event are shifted to the left by -6 nm with respect to the control trace, as indicated by the arrows in (a). Subsequent force peaks coincide exactly with the control trace. In (c) three recordings are superimposed: the control traces just before (in black) and after (in blue) the trace containing the missing force peak (in green).

Our results demonstrate that after mechanical denaturation, neighboring I27 modules can coalesce into a single fold with an altered mechanical topology. The force peaks corresponding to the unfolding of a 'skip' fold have an average force, $F^{\text{skip}} = 194 \pm 57$ pN ($n = 13$), that is similar to the unfolding force measured from native I27^{RS} polyproteins. Hence, the 'skip' fold has a mechanical stability that is similar to that of a correctly folded I27 module.

It is likely that the misfolding events reported here occur in a wide variety of tandem modular proteins that are typically placed under stress during their biological functions. For example, tenascin is an extracellular matrix protein that is composed of tandem epidermal growth factor and FN III domains¹³ and mediates mechanical interactions during cell rolling¹⁴. It has been proposed that the mechanical forces resulting from these interactions trigger module unfolding in the FN III region of tenascin⁴. We have found that, during repeated unfolding/refolding cycles of the FN III region of tenascin¹³ (TNfnALL), the saw-tooth pattern of the force/extension curves also show 'skip' misfolding events with a frequency of about 4% (Fig. 4a–d). The tenascin misfolding events showed an average increase in contour length of $\Delta^{\text{skip}} = 65.2 \pm 5.5$ nm ($n = 22$), which is ~ 9 nm longer than the contour length increment expected from the simultaneous unfolding of two FN III modules. Likewise, the

unfolding events preceding the 'skip' event are shifted to the left by $\delta = 7.6 \pm 2$ nm ($n = 6$; Fig. 4a–c), consistent with a shorter folded protein. We have also observed three cases where the tenascin 'skip' events showed two missing peaks (Fig. 4d). The presence of mechanical misfolds in native tenascin suggests that these misfolds may be a common feature of tandem modular proteins, regardless of their construction. Intracellular chaperones are known to prevent the misfolding of consecutive modules, for example during translation¹⁵. However, it is not clear how the misfoldings that we report here would be prevented in proteins of the extracellular matrix, which are exposed to repeated mechanical unfolding cycles. Nonetheless, misfolding events may occur whenever two consecutive modules unfold regardless of the unfolding mechanism.

It has been shown that the immunoglobulin domains of neighboring CD2 cell adhesion proteins can interact and form a stable structure that encompasses two intertwined immunoglobulin domains^{16–18}. X-ray crystallography and NMR spectroscopy were used to demonstrate that recombinant immunoglobulin domains from the CD2 protein either folded as individual domains (85%) or combined to fold into a stable single module with twice the length of a single immunoglobulin domain¹⁶ ($\sim 15\%$). Furthermore, it was demonstrated that the probability of forming these dimeric folds was dependent on some highly conserved amino acids, which are exposed to the hydrophilic interface of the immunoglobulin modules of CD2 (ref. 17). These metastable immunoglobulin dimers involve the immunoglobulin domains of two neighboring proteins. By contrast, the 'skip' misfolds include neighboring immunoglobulin domains in a single tandem modular protein. However, it is possible

that there is some structural resemblance between the two.

Our work demonstrates, for the first time, the ability of single protein AFM recording to capture protein misfolding events in real time. A 'skip' misfold could significantly affect the function of an extracellular matrix protein. A protein harboring a misfold will be shorter, by several nanometers, as compared to the correctly folded protein (Fig. 4d), affecting the binding of ligands that cannot extend to bind a shorter receptor. Furthermore, a shorter protein will be less elastic and will decrease the pulling force to a greater extent upon unfolding of the 'skip' fold. This altered elasticity is likely to change the lifetime and affinity of the receptor–ligand bond¹⁹.

Methods

Protein engineering. Recombinant DNA techniques were used to synthesize and express direct tandem repeats of the I27 monomer as described⁹. Directional DNA concatemerization was done using a multistep cloning technique that makes use of four restriction sequences (*Bam*HI, *Bgl*II, *Sma*I and *Kpn*I) to build eight tandem repeats of I27. The synthetic I27^{RS} was cloned in an *Escherichia coli* recombination-defective strain, SURE-2 (Stratagene), expressed in the M15 strain, and purified by Ni²⁺-affinity chromatography under non-denaturing conditions. The recombinant human tenascin protein comprising the 15 FNIII domains (TNfnALL)¹³ was a generous gift from H.P. Erickson (Duke University).

letters

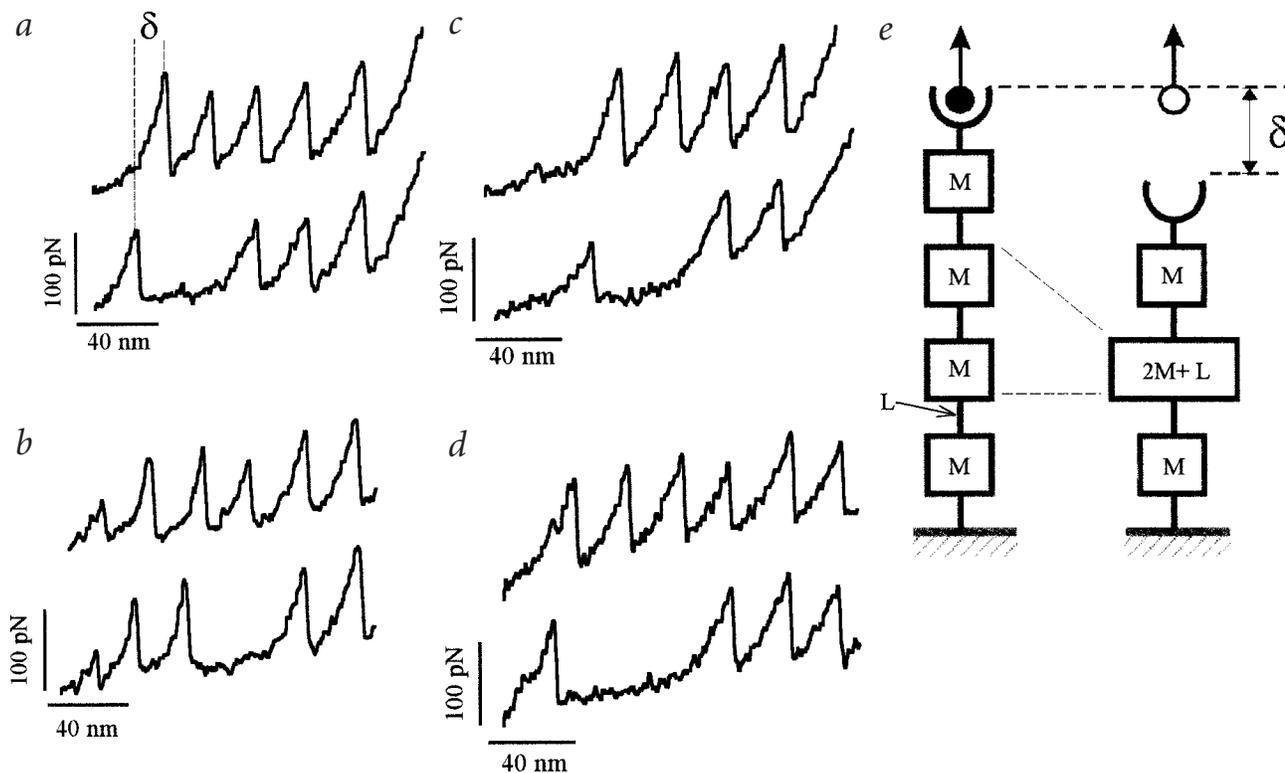


Fig. 4 Misfolding events observed during mechanical unfolding/refolding cycles of the extracellular matrix protein tenascin. Similarly to the 'skip' events observed in the $(I27)_8$ polyprotein, repeated unfolding/refolding cycles of the tandem FN III region of tenascin (TNfnALL)¹³ show misfolds with one (a–c) and two skip events d. In all these cases the unfolding peaks preceding the 'skip' event were shifted to the left, indicating a shortened folded protein (δ in a). e, Schematic diagram illustrating a correctly folded tandem modular protein with a receptor at its end being pulled by a ligand (arrow). The protein modules (M) are spaced by a linker region (L). A misfold comprising two neighboring modules plus the linker shortens the resting length of the folded protein by $\delta = L$, which results in failure of the interaction between the receptor and its ligand.

Single protein atomic force microscopy. Our custom made single molecule AFM apparatus as well as its mode of operation are identical to those previously described^{4,9}. The spring constant of each individual AFM cantilever (Si_3N_4 tip NPS, Digital Instruments) was calibrated using the equipartition theorem²⁰ and varied between 70 and 120 pN nm⁻¹. The proteins were allowed to adsorb onto freshly evaporated gold coverslips for about 10 min and then rinsed with PBS. We could stretch and relax a single protein repeatedly by limiting the extension so that the molecule did not detach from the AFM cantilever or the substrate. The time interval between consecutive unfolding events was between 10 and 30 s, to allow the full refolding of the unfolded domains^{4,9}.

Contour length measurements. The initial contour length of the folded protein (L_c^0) and the contour length increments (ΔL_c) caused by domain unfolding were measured using the wormlike chain (WLC) equation¹¹. The adjustable parameters of the WLC model are the persistence length, p and the contour length of the polymer, L_c . We measured L_c^0 by fitting the first force peak of the sawtooth pattern to the WLC equation (dotted line in Fig. 2a); the zero length point was defined as the point where the AFM cantilever tip contacts the coverslip.

Acknowledgments

We would like to thank H.P. Erickson (Duke University) for providing the human tenascin protein. This work was supported by NIH grants to J.M.F., P.E.M. and A.F.O.

Correspondence should be addressed to J.M.F. email: fernandez.julio@mayo.edu

Received 26, May 1999; accepted 22 July, 1999.

- Bork, P., Downing, A.K., Kieffer, B. & Campbell, I.D. *Q. Rev. Biophys.* **29**, 119–167 (1996).
- Chothia, C. & Jones, E.Y. *Annu. Rev. Biochem.* **66**, 823–862 (1997).
- Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J.M. & Gaub, H.E. *Science* **276**, 1109–1112 (1997).
- Oberhauser, A.F., Marszalek, P.E., Erickson, H.P. & Fernandez, J.M. *Nature* **393**, 181–185 (1998).
- Ohashi, T., Kiehart, D.P. & Erickson, H.P. *Proc. Natl. Acad. Sci. USA* **96**, 2153–2158 (1999).
- Erickson, H.P. *Science* **276**, 1090–1092 (1997).
- Keller, T.C. *Nature* **387**, 233–235 (1997).
- Hynes, R.O. *Proc. Natl. Acad. Sci. USA* **96**, 2588–2590 (1999).
- Carrion-Vazquez, M. et al. *Proc. Natl. Acad. Sci. USA* **96**, 3694–3699 (1999).
- Improta, S., Politou, A.S. & Pastore, A. *Structure* **4**, 323–337 (1996).
- Marko, J.F. & Siggia, E.D. *Macromolecules* **28**, 8759–8770 (1995).
- Improta, S. et al. *J. Mol. Biol.* **284**, 761–777 (1998).
- Aukhil, I., Joshi, P., Yan, Y. & Erickson, H.P. *J. Biol. Chem.* **268**, 2542–2552 (1993).
- Clark, R.A., Erickson, H.P. & Springer, T.A. *J. Cell. Biol.* **137**, 755–765 (1997).
- Netzer, W. & Hartl, U.F. *Nature* **388**, 343–349 (1997).
- Murray, A.J., Lewis, S.J., Barclay, A.N. & Brady, R.L. *Proc. Natl. Acad. Sci. USA* **92**, 7337–7741 (1995).
- Murray, A.J., Head, J.G., Barker, J.J. & Brady, R.L. *Nature Struct. Biol.* **5**, 778–782 (1998).
- Hayes, M.V., Sessions, R.B., Brady, R.L. & Clarke, A.R. *J. Mol. Biol.* **285**, 1857–1867 (1999).
- Evans, E. & Ritchie, K. *Biophys. J.* **76**, 2439–2447 (1999).
- Florin, E.L. et al. *Biosensors & Bioelectronics* **10**, 895–901 (1995).