FEBS Letters 476 (2000) 124–128 FEBS 23827

States and transitions during forced unfolding of a single spectrin repeat

P.-F. Lenne*, A.J. Raae, S.M. Altmann, M. Saraste, J.K.H. Hörber

European Molecular Biology, Cell Biology and Biophysics Programme, P.O. Box 10.2209, D-69012 Heidelberg, Germany

Received 10 May 2000; received in revised form 29 May 2000

Edited by Giorgio Semenza

Abstract Spectrin is a vital and abundant protein of the cytoskeleton. It has an elongated structure that is made by a chain of so-called spectrin repeats. Each repeat contains three antiparallel α-helices that form a coiled-coil structure. Spectrin forms an oligomeric structure that is able to cross-link actin filaments. In red cells, the spectrin/actin meshwork underlying cell membrane is thought to be responsible for special elastic properties of the cell. In order to determine mechanical unfolding properties of the spectrin repeat, we have used single molecule force spectroscopy to study the states of unfolding of an engineered polymeric protein consisting of identical spectrin domains. We demonstrate that the unfolding of spectrin domains can occur in a stepwise fashion during stretching. The force–extension patterns exhibit features that are compatible with the existence of at least one intermediate between the folded and the completely unfolded conformation. Only those polypeptides that still contain multiple intact repeats display intermediates, indicating a stabilisation effect. Precise force spectroscopy measurements on single molecules using engineered protein constructs reveal states and transitions during the mechanical unfolding of spectrin. Single molecule force spectroscopy appears to open a new window for the analysis of transition probabilities between different conformational states.

Key words: Force spectroscopy; Single molecule; Spectrin repeat; Unfolding intermediate state

1. Introduction

The modular structure of proteins in natural fibres and the cytoskeleton seems to be a general strategy for resistance against mechanical stress [1]. One of the most abundant modular proteins in the cytoskeleton is spectrin. In erythrocytes, spectrin molecules are part of a two-dimensional network that is assumed to provide the red blood cells with special elastic features [2]. The basic constituent of spectrin subunits is the repeat which has about 106 amino acids and is made of three antiparallel α-helices, folded into a left-handed coiled-coil [3]. The repeats are connected by a helical linker [4,5].

The mechanical properties of several modular proteins have already been investigated by atomic force microscopy (AFM) [6], especially titin [7]. Such experiments have demonstrated that the elongation events observed during stretching of single proteins may be attributed to the unfolding of individual domains; experiments with optical tweezers have corroborated these results [8,9]. These studies suggest that single domains unfold one at a time in an all-or-none fashion when subjected to directional mechanical stress. More recently, Oesterhelt et al. [10] have shown that individual bacteriorhodopsins can unfold in a stepwise manner.

In a typical experiment, single proteins are attached to properly prepared surfaces and subsequently captured on AFM tips. They can be stretched to more than 10 times their folded lengths, reaching almost the total contour length (Fig. 1). The force–extension curves show a characteristic sawtooth-like pattern. The reaction coordinate of unfolding is imposed by the direction of pulling, and unfolding events occurring in a single protein can be studied in this way. Each peak is attributed to the breakage of a folded structure.

Rief et al. [11] have shown that under mechanical stress the repeats of the native spectrin break as well defined units. As these repeats are homologous but not identical, we have engineered a polymeric molecule consisting of four identical repeats referred to as $(R16)_4$ in the following, in order to get a better understanding of the mechanical properties of the spectrin repeat. The unfolding features in the motif should thus be multiplied and could be compared within the frame of one repeat [12,13]. The polymeric protein was constructed from the repeat R16 of the α-spectrin chain since the structure had already been determined by nuclear magnetic resonance spectroscopy [3]. The essential features of its structure are confirmed by the recent crystallographic study [5].

We used AFM to unfold the polymeric protein consisting of identical domains. We shall demonstrate that the unfolding of a single spectrin repeat can occur in a stepwise fashion during stretching.

2. Materials and methods

2.1. Construction of polymeric R16

A modified R16 repeat was produced by PCR using the pET R16 plasmid [14] as template. The 5′ primer sequence, CCGGATCC-TCCATGGAAATGAGCGCTGCCCGCGCGAAGCTCAACGAG-TCG, contains a *Bam*HI and a *Bsp*HI site, the initiation codon as well as an N-terminal five-residue extension of the original R16 sequence. The 3′ primer, GGGCGGCCGCTTATTAGCAACAGGTCATGA-TGTACTCCAGGGACTCCTCCAGACGTTGCCCCCGAGC, contains *Nco*I and *Not*I sites as well as two stop codons. The 3′ primer adds an eight-residue extension to the R16 repeat as well as two cysteine residues to the end of the expressed protein products. The gene sequence is extended at both ends so that the polymeric protein product contains a 13-residue linker between the consecutive R16 units, and has two cysteine residues at the C-terminus. The PCR product was cut with *Bam*HI and *Not*I, ligated into a pBluescript SK plasmid and amplified in DH-5α cells. The modified R16 sequence was cut from the vector with *Nco*I and *Not*I, and ligated in frame into the *Nco*I restriction site of a domestic pET24d derivative that confers a six-residue histidine tag and a TEV protease site to the N-terminus of the protein. This vector was cut with two different combinations of restriction enzymes, *Bsp*HI/*Xba*I and *Nco*I/*Not*I, respectively. After purification, the fragments containing the R16 sequences were joined by ligation at the *Bsp*HI and *Nco*I sites, and simultaneously ligated into a pET24d vector which was opened with *Xba*I and *Not*I, and

*Corresponding author. Fax: (49)-6221-387 306.
E-mail: lenne@embl-heidelberg.de

PII: S0014-5793(00)01704-X

dephosphorylated with shrimp alkaline phosphatase. After transformation, DH-5α cells containing the two-repeat construct were identified by PCR analysis of the bacterial colonies. The four-repeat construct $(R16)_4$ was subsequently made by iterating the same procedure.

2.2. Protein expression and purification

For expression, 0.1 μg of plasmid DNA was used to transform competent BL-21 cells. After growth at 37°C to $A_{600} = 0.7$, the cells were equilibrated at 30°C before isopropyl-1-thio-β-D-thiogalactopyranoside was added to a final concentration of 0.4 mM. After induction for 8 h, the cells were harvested by centrifugation. The cells were suspended in lysis buffer containing 50 mM Tris–HCl, pH 8.0, 30 mM imidazole, 200 mM NaCl, 0.2% NP-40, 0.15 mM phenylmethylsulfonyl fluoride and 0.1 mM 4-(2-aminoethyl)-benzolsulfonyl fluoride, and lysed by sonication. The lysate, 40 ml from a 500 ml cell culture, was mixed with 1.5 ml Ni-NTA resin and kept 10 min on ice. The Ni-NTA/lysate suspension was filtered through a 10 ml Bio-Rad polyprep column. The filtrate was reloaded on the column, and the resin was first washed with the lysis buffer, then with the lysis buffer without detergent containing 1 M NaCl, and with 30 mM imidazole. The bound protein was eluted in 1 ml 200 mM imidazole. The four-repeat construct was almost quantitatively recovered in a soluble state after expression at 30°C.

2.3. AFM

The experiments were performed with an AFM developed at EMBL working in the so-called force spectroscopy mode [15]. The cantilevers (model MLCT-AUHW, Park Scientific, Sunnyvale, CA, USA) were calibrated by using the equirepartition theorem as described by e.g. Florin et al. [16]. The proteins were kept in buffer 50 mM Tris–HCl pH 8.0, 50 mM NaCl. A drop of 40 μl of protein solution (100 μg/ml) was deposited on freshly sputtered gold surfaces as proteins have cysteine residues at the C-terminus to bind covalently to gold. The proteins were allowed to adsorb for 10 min before the sample was rinsed with buffer and brought into the fluid experimental chamber of the AFM. The single proteins were picked up at random by approaching the sample towards the tip while monitoring the deflection amplitude. After contact with the surface was established, the maximum normal forces during the spectroscopy cycles were kept as low as possible. Approach/retraction cycles of the sample with respect to the AFM tip were used to probe the force–extension curves of single molecules of spectrin. We used pulling speeds between 0.3 and 15 nm/ms. The AFM controls distance with 0.1 nm accuracy and the applied force with pN precision.

3. Results and discussion

3.1. Single molecule force spectroscopy on $(R16)_4$

We used recombinant DNA techniques to construct tandem repeats from a single spectrin domain (see Section 2). The method extends the monomer (R16) at both ends so that the polymeric protein product contains a 13-residue linker between the consecutive R16 units and has two cysteine residues at the C-terminal end.

In Fig. 2, an SDS–PAGE analysis of the mono-, di- and tetrameric R16 repeats is shown together with the repeat 13–18 section from the chicken brain α-spectrin chain. The apparent molecular weight of the four-repeat construct $(R16)_4$ is 64 kDa. Due to the extensions in the connecting regions and at the N- and C-terminal ends, the total molecular size is in the same range as that of the native hexameric spectrin repeat 13–18 that was studied by Rief et al. [11].

The force–extension relationships for $(R16)_4$ were measured using an AFM developed at EMBL (Section 2). The proteins formed covalent bonds with the gold surfaces through cysteine residues which were introduced at the C-terminal end. The protein-coated surfaces were then carefully brought into contact with the tip and single proteins were picked up by unspecific binding. By retracting the sample holder away from the AFM tip, single proteins could be completely extended.

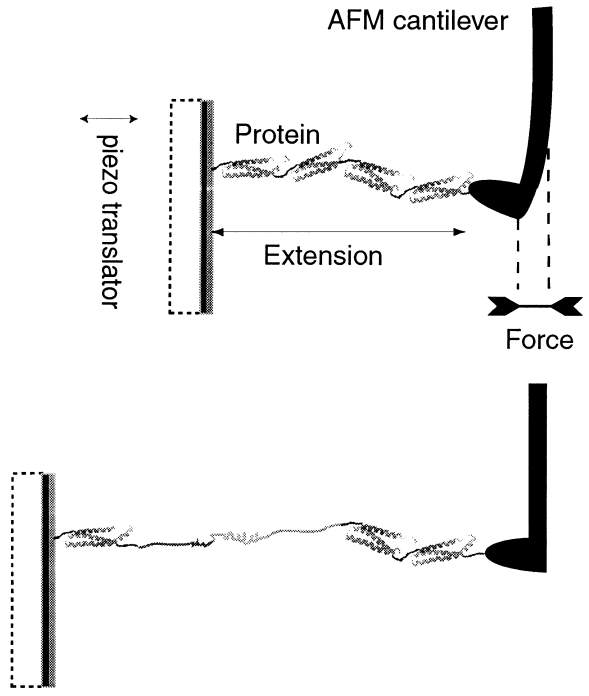

Fig. 1. AFM force measurements. Experimental scheme for protein 'fishing'.

Force–extension curves present the sawtooth-like pattern typical for modular proteins (Fig. 3). Each peak is interpreted as an unfolding event. The loss of each folded structure in the spectrin chain allows a gain of length for the unfolded part of the chain. After each peak, the force drops as a new length of the polypeptide chain is gained but has not yet been extended. The protein is then stretched further until the force is sufficient for another part of the folded structure to unfold. The maximum extension is variable as the polymer is picked up at random positions and not necessarily at its end.

The force–extension curves can be well described by a worm-like chain model (WLC) [17]. It predicts the isothermal restoring force of a flexible polymer, acting as an entropic spring during extension. Fits of the force–extension curves to the WLC model [18] resulted in a persistence length of

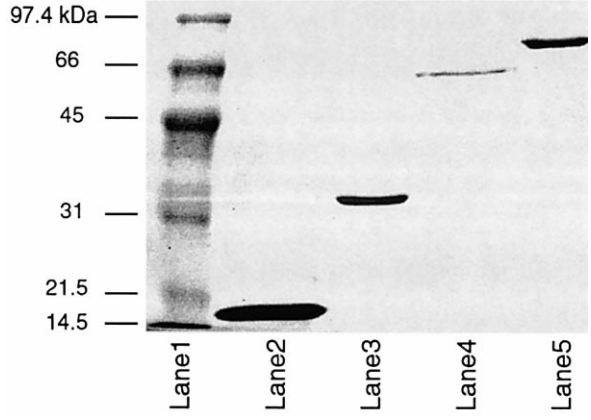

Fig. 2. An SDS–PAGE analysis of the spectrin R16 polymeric products. Lane 1: Bio-Rad LMW markers. Lane 2: monomeric R16 spectrin repeat. Lane 3: dimeric R16 spectrin. Lane 4: tetrameric R16 spectrin. Lane 5: the repeats 13–18 from the α-spectrin chain.

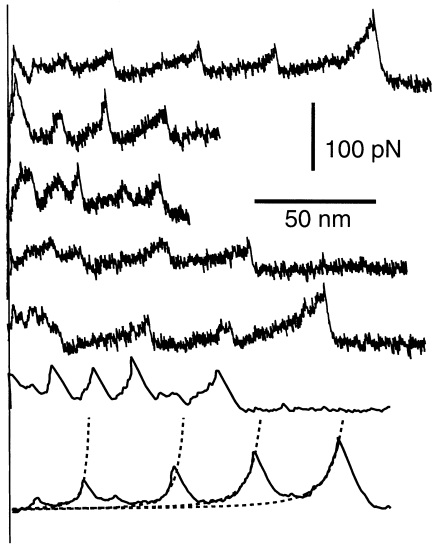

Fig. 3. Unfolding traces of a spectrin construct made of four identical tandem domains $(R16)_4$. The five upper curves were recorded at a pulling speed of 0.3 nm/ms and the two last at 3 nm/ms. The WLC model fits to the left-hand slopes of the unfolding patterns with a persistence length of $p = 0.59 \pm 0.22$ nm (see last curve). As the proteins are picked up at random, the maximum extension varies from one curve to the other but never exceeds the total contour length of the protein.

$p = 0.59 \pm 0.22$ nm ($n = 20$). Two consecutive force peaks are spaced by a distance that reflects the gain of length produced by a single unfolding event. These distances were measured for 250 unfolding events.

A histogram of distances (Fig. 4a) reveals a length distribution with two major peaks, at 15.5 nm (± 3.5 nm) and 31 nm (± 7 nm) when fitted with a mixture of two Gaussians. The force distributions associated with the short and long elongation events are also clearly distinct (Fig. 4b,c). The mean force for the short and long elongation events are 60 and 80 pN, respectively, at a pulling speed of 3 nm/ms.

General features of the unfolding events are provided by a detailed statistical analysis of the whole set of measurements:

1. The short unfolding events are most likely for a short extension, i.e. at the first stage of pulling (80% below 35 nm, $n = 87$) of the protein whereas the long unfolding events are observed at longer extension (50% above 35 nm, $n = 94$).
2. When two short unfolding events occur, they are consecutive in 90% of the cases.

We measured the same unfolding patterns for different speeds (0.3, 0.75 and 1.5 nm/ms). The mean unfolding force was decreased by approximately 20% when the pulling speed was reduced by a factor of 10 (see insets in Fig. 4b,c). The strength of the bonds that stabilise protein structures can depend on the speed of retraction as it has been shown for the biotin/streptavidin system [19].

3.2. Reversibility of unfolding

After reaching the extended state and before rupturing the bond between the protein and the tip, the protein was relaxed to its initial length (Fig. 5). After waiting for 5 s, the protein was stretched again and four force peaks were observed. This suggests that the four domains spontaneously refold during

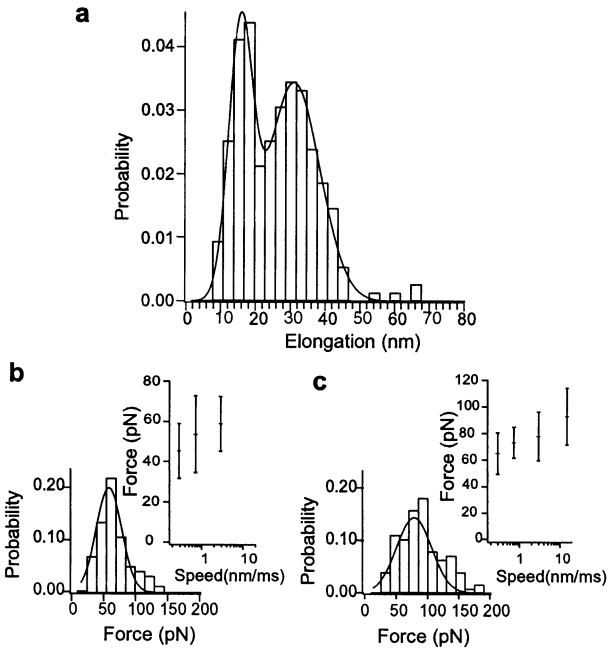

Fig. 4. Elongation events and unfolding forces of the spectrin construct $(R16)_4$. (a) Histogram probability of elongation after an unfolding event ($n = 250$). The line corresponds to a fit by a normalised function, the sum of two Gauss functions. The Gauss functions are centred on 15.5 nm and 31 nm. Unfolding force probability histogram and force–speed dependence for short elongation events (b) and long elongation events (c).

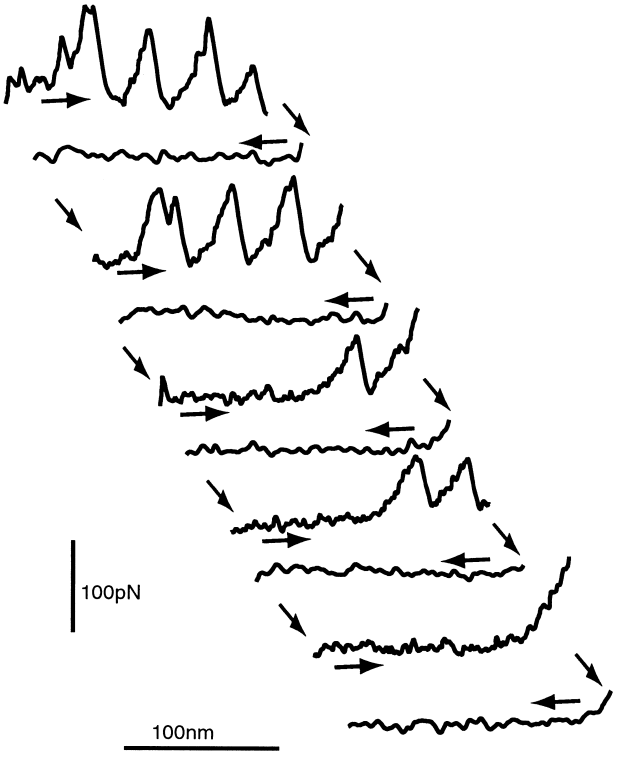

Fig. 5. Cycles of unfolding–refolding of a single molecule of spectrin (see text).

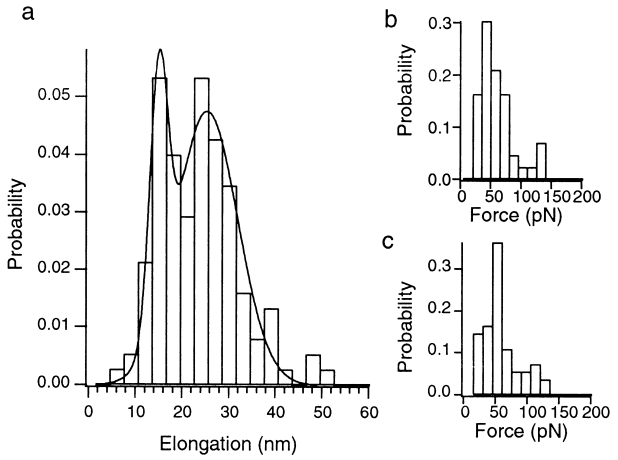

Fig. 6. Elongation events and unfolding forces of the natural hexameric 13–18 spectrin repeat (from chicken brain α-spectrin). (a) Histogram probability of elongation after an unfolding event ($n = 125$). The line corresponds to a fit by a normalised function, the sum of two Gauss functions. The Gauss functions are centred on 15 nm and 26 nm. Unfolding force probability histogram and for short elongation events (b) and long elongation events (c) at a pulling speed of 0.3 nm/ms.

relaxation. If the relaxation time is decreased to 1 s, the number of refolded domains is decreased in each cycle. This demonstrates that refolding occurs on a time scale of seconds and is therefore significantly slower than the unfolding rate.

3.3. Comparison with the native spectrin repeats

We have carried out equivalent measurements with another construct made of six homologous but non-identical repeats (the repeats 13–18 from chicken brain α-spectrin, see also [11]). Our measurements result in a similar distribution of elongation events; the most probable events occur at 15 and 26 nm (Fig. 6). The mean unfolding force is 55 pN for both the short and long elongation events with a pulling speed of 0.3 nm/ms. The long elongation event is in agreement with the results reported by Rief et al. [11]. However, because we were able to collect much more data than Rief et al., we could clearly observe short elongation events in our measurements both on $(R16)_4$ as well as R13–R18 when plotting their elongation length distributions.

3.4. Intermediates

The origin of the two distinct elongation events observed can have several causes. One can argue that short elongations are due to multiple-molecule pick-up. We showed that short elongations correspond to low forces and long elongations to large forces. Multiple pick-ups could indeed lead to shorter elongations but this would automatically have to go along with higher forces than for the long elongations which fit the unfolding length of a single repeat very well. We carefully excluded the possibility of multiple pick-ups by discarding any force curves displaying an off-set in the force indicating multiple parallel springs, i.e. multiple-molecule pick-up, from our analysis.

The second possible cause of a 'premature' unfolding of the R16 motif could be due to destabilising long range surface forces. These would have to reach further than 10 nm from the surface, as we observed partial unfolding at a distance out to more than 50 nm.

This is not the case. We recorded force spectra of levers in water and phosphate-buffered saline at distances between 5 and 30 nm from the surface clearly showing that the electrostatic forces can be completely ignored at a few nanometers from the surface [20].

A third possibility could be that the proteins denature at least partially on the surface. This would, in addition to the cysteine–gold bond, increase the binding forces. It would also preclude the observation of a total of more than three repeats unfolding partially or fully in one trace. In several measurements where partial unfolding occurred, the unfolding lengths sum to more than these three repeats. If the protein was only partially denatured, it seems very unlikely that reversibility, as we observed it, would be possible. Refolding of the rest of such a partially denatured domain seems very unlikely. Also the rest of the domain can be expected to be distorted greatly and increasingly by the recurrence of the surface forces influence which led to the partial denaturing.

These reasons make multiple pick-ups and at least direct surface effects quite unlikely, and we believe that the existence of the different elongation events is due to at least one stable intermediate between the folded state and completely unfolded state during forced unfolding.

Thermodynamic studies have not yet documented such an intermediate: fluorescence spectroscopy and circular dichroism measurements have shown that single spectrin repeats undergo a reversible two-state transition [21,22]. Statistical analysis of the force curves shows that well defined short elongation events of 15.5 nm take place. The 15.5 nm elongation is half of that produced by the total unfolding of one domain. It may thus correspond either to the gain of length between a folded and unfolded domain or to the gain of length between a partially and completely unfolded domain. The variation in our length distributions is large but comparable with data of Yang et al. [23]. This broadness is partly attributed to the fact that we chose to measure the distance between peaks rather than the gain of contour lengths between consecutive peaks which are obtained by multiple-parameter fits [23]. Moreover, spectrin unfolds at very low forces, and therefore unavoidably any distribution is broadened in comparison to the unfolding of titin for instance. It could also be a consequence that several possible intermediates contribute to the distributions.

We assume that, as two short elongation events are consecutive in most cases, the unfolding of one domain results from at least two transitions, each producing on average a similar elongation. An intriguing observation is then that the rest of the partially unfolded domain is stable on a time scale of 1/10 of a second as the pulling speed is in the nm/ms range. This is an unexpected result and needs more in-depth study.

A molecular dynamics simulation [24] has recently suggested that protein domains such as the fibronectin-type 3 module can unfold in a stepwise fashion. In this case, unfolding follows a pathway in which a 'β-sandwich' first separates from the fold, and the individual β-sheets unfold secondly. The molecular pathway of unfolding is unknown for the spectrin domain, but it may well involve a partial opening of the bundle and/or the loss of the secondary structure at the ends of the helices which may be elongated before the helical bundle finally opens.

The different unfolding events are not only distinct in elon-

gation lengths but also in the unfolding forces. The mean unfolding force for short elongation events is 60 pN whereas the force associated to long elongation events is 80 pN (at a pulling speed of 3 nm/ms). This shows that the transition from a native folded conformation to an intermediate is mechanically different from the transition from a totally folded conformation to a completely unfolded conformation.

The feature that the unfolding of the domains depends on the extension of the chain is also interesting by itself. At short extensions of the polypeptide chain, short unfolding events are more likely than long ones. This behaviour suggests that the selection of the unfolding pathway may be influenced by changes in the mechanical properties during the extension of polypeptide chain. Relaxation processes following the breakage of folded structures possibly govern the access to intermediates. This may indicate as well that the unfolding of tandem repeats in spectrin chains is not completely independent [25]. The tandem spectrin repeats are not classical independent structural domains, and have stabilising interactions. A folded repeat may have a different stability depending on whether its neighbouring repeats are folded or unfolded. As suggested by our measurements, a folded repeat unfolds more likely by steps when its neighbouring repeats are folded (at small extension of the chain) than when they are unfolded (at large extension). The observed intermediates are conformational states which are stable from milliseconds to seconds, i.e. on time scales accessible in AFM measurements.

Although the molecular complexity of unfolding pathways can be very high, our AFM measurements demonstrate that force spectroscopy of single properly engineered proteins can give important clues on thermally averaged energy landscapes relevant on laboratory time scales. In the future, we expect an increasing contribution by force spectroscopy measurements to the understanding of protein folding as proteins will be engineered in order to systematically perturb the unfolding pathways imposed by the real space directionality in such measurements.

Acknowledgements: We thank M. Nilges and J. Linge for discussions, L. Serrano and T. Gibson for reading of this manuscript, and G. Stier for his technical advice. Support was provided by a grant to P.-F.L. from the European Commission (TMR) and a grant to A.J.R. from the Norwegian Research Council.

References

[1] Smith, B.L. et al. (1999) Nature 399, 761–763.
[2] Elgsaeter, A., Stokke, B.T., Mikkelsen, A. and Branton, D. (1986) Science 234, 1217–1223.
[3] Pascual, J., Pfuhl, M., Walther, D., Saraste, M. and Nilges, M. (1997) J. Mol. Biol. 273, 740–751.
[4] Djinovic-Carugo, K., Young, P., Gautel, M. and Saraste, M. (1999) Cell 98, 537–546.
[5] Grum, V.L., Li, D., MacDonald, R.I. and Mondragon, A. (1999) Cell 98, 523–535.
[6] Fisher, T.E., Oberhauser, A.F., Carrion-Vazquez, M., Marszalek, P.E. and Fernandez, J.M. (1999) Trends Biochem. Sci. 24, 379–384.
[7] Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J.M. and Gaub, H. (1997) Science 276, 1109–1112.
[8] Kellermayer, M.S., Smith, S.B., Granzier, H.L. and Bustamante, C. (1997) Science 276, 1112–1116.
[9] Tskhovrebova, L., Trinick, J., Sleep, J.A. and Simmons, R.M. (1997) Nature 387, 308–312.
[10] Oesterhelt, F., Oesterhelt, D., Pfeiffer, M., Engel, A., Gaub, H.E. and Muller, D.J. (2000) Science 288, 143–146.
[11] Rief, M., Pascual, J., Saraste, M. and Gaub, H. (1999) J. Mol. Biol. 286, 553–561.
[12] Carrion-Vazquez, M., Oberhauser, A.F., Fowler, S.B., Marszalek, P.E., Broedel, S.E., Clarke, J. and Fernandez, J.M. (1999) Proc. Natl. Acad. Sci. USA 96, 3694–3699.
[13] Yang, G. et al. (2000) Proc. Natl. Acad. Sci. USA 97, 139–144.
[14] Pascual, J., Pfuhl, M., Rivas, G., Pastore, A. and Saraste, M. (1996) FEBS Lett. 383, 201–207.
[15] Rief, M., Oesterhelt, F., Heymann, B. and Gaub, H.E. (1997) Science 275, 1295–1297.
[16] Florin, E.-L., Rief, M., Lehmann, H., Ludwig, M., Dornmair, C., Moy, V.T. and Gaub, H.E. (1995) Biosens. Bioelectron. 10, 895–901.
[17] Fixman, M. and Kovac, J. (1973) J. Chem. Phys. 56, 1564–1568.
[18] Bouchiat, S.M., Wang, M.D., Allemand, J.-F., Strick, T., Block, S.M. and Croquette, V. (1999) Biophys. J. 76, 409–413.
[19] Merkel, R., Nassoy, P., Leung, A., Ritchie, K. and Evans, E. (1999) Nature 397, 50–53.
[20] Butt, A.-J. (1991) Biophys. J. 60, 1438–1444.
[21] Lusitani, D., Menhart, N., Keiderling, T.A. and Fung, L.W.-M. (1998) Biochemistry 37, 16546–16554.
[22] DeSilva, T.M., Harper, S.L., Kotula, L., Hensley, P., Curtis, P.J., Otvos Jr., L. and Speicher, D.W. (1997) Biochemistry 36, 3991–3997.
[23] Carrion-Vazquez, M., Marszalek, P.E., Oberhauser, A.F. and Fernandez, J.M. (1999) Proc. Natl. Acad. Sci. USA 96, 11288–11292.
[24] Paci, E. and Karplus, M. (1999) J. Mol. Biol. 288, 441–459.
[25] Menhart, N., Mitchell, T., Lusitani, D., Topouzian, N. and Fung, L.W.-M. (1996) J. Biol. Chem. 271, 30410–30416.