

Unfolding a Linker between Helical Repeats

Vanessa Ortiz^{1,2,3}, Steven O. Nielsen^{1,3}, Michael L. Klein^{1,3} and Dennis E. Discher^{2,3*}

¹Center for Molecular Modeling
Department of Chemistry
University of Pennsylvania
Philadelphia, PA 19104, USA

²Biophysical Engineering
Laboratory, Department of
Chemical and Biomolecular
Engineering, University of
Pennsylvania, Philadelphia
PA 19104, USA

³Laboratory for Research on the
Structure of Matter, University
of Pennsylvania, Philadelphia
PA 19104, USA

In many multi-repeat proteins, linkers between repeats have little secondary structure and place few constraints on folding or unfolding. However, the large family of spectrin-like proteins, including α -actinin, spectrin, and dystrophin, share three-helix bundle, spectrin repeats that appear in crystal structures to be linked by long helices. All of these proteins are regularly subjected to mechanical stress. Recent single molecule atomic force microscopy (AFM) experiments demonstrate not only forced unfolding but also simultaneous unfolding of tandem repeats at finite frequency, which suggests that the contiguous helix between spectrin repeats can propagate a cooperative helix-to-coil transition. Here, we address what happens atomistically to the linker under stress by steered molecular dynamics simulations of tandem spectrin repeats in explicit water. The results for α -actinin repeats reveal rate-dependent pathways, with one pathway showing that the linker between repeats unfolds, which may explain the single-repeat unfolding pathway observed in AFM experiments. A second pathway preserves the structural integrity of the linker, which explains the tandem-repeat unfolding event. Unfolding of the linker begins with a splay distortion of proximal loops away from hydrophobic contacts with the linker. This is followed by linker destabilization and unwinding with increased hydration of the backbone. The end result is an unfolded helix that mechanically decouples tandem repeats. Molecularly detailed insights obtained here aid in understanding the mechanical coupling of domain stability in spectrin family proteins.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: spectrin repeats; steered molecular dynamics; mechanical unfolding; alpha-helical linkers; cooperative domain unfolding

*Corresponding author

Introduction

Multi-domain proteins prove to be a dominant fraction of the proteome,¹ but the physical roles of linkers between domains as domain separators and/or connectors are not always clear. Multi-repeat proteins have now proven accessible to elegant single molecule experiments with atomic force microscopy (AFM)² and optical trapping,³ which allow the study of domain unfolding and folding⁴ while lending insight into linker stability. The various immunoglobulin domains of the cytoskeletal protein titin^{2,3} as well as related fibronectin-III domains^{5,6} unfold under force one domain at a time. The unfolding force clearly depends on the extension or loading rate,⁷ but

there is no significant dependence on the linkers between domains.⁸ In contrast, a much more cooperative coupling of domain unfolding under force is exhibited by spectrin family proteins.⁹

Spectrin, dystrophin, α -actinin, and related proteins share a serial repeat structure (called a spectrin repeat or domain), bind actin filaments, and clearly contribute resilience to the cell and its membrane.^{10–13} Evidence for dynamic dissociation of spectrins under cell stress is just emerging,¹⁴ as is the possibility of spectrin unfolding *in situ*.¹⁵ The characteristic spectrin repeats fold into a bundle of three antiparallel α -helices,^{12,16–20} with any two adjacent repeats linked by an \sim 8–9 nm helix that extends from one domain to the next.^{12,18–21} Even where non-helical linkers have been predicted by various secondary structure algorithms, the connecting linker is helical in crystal structures of tandem repeats.^{21,22} Forced unfolding of spectrin repeats occurs at five to ten times smaller forces^{9,23,24} than the forces measured for unfolding

Abbreviations used: AFM, atomic force microscopy; SMD, steered molecular dynamics.

E-mail address of the corresponding author:
discher@seas.upenn.edu

immunoglobulin and related β -sheet domains, with the difference in force explained by the weak, hydrophobic interactions that bundle spectrin repeats. A further difference is that adjacent repeats in spectrin family proteins stabilize each other²⁵ and frequently unfold together in single molecule extension⁹ as well as in solution denaturation.²² While this collective unfolding has been attributed to a cooperative helix-to-coil transition that propagates through the connecting helix,⁹ structural insights are needed into the mechanical coupling or decoupling of repeats.

AFM extension of a tandem repeat spectrin construct shows two unfolding pathways (Figure 1(a)). Sequential unfolding of tandem repeats is evident in a force-extension sawtooth pattern with equally spaced peaks where the period approximates the contour length of a single repeat; tandem unfolding of repeats involves simultaneous unfolding of two repeats (or more), yielding a sawtooth pattern where some peaks are spaced by about twice the contour length of a single repeat. No additional force is required for the cooperative unfolding of tandem repeats.⁹ However, the relative frequency of tandem repeat unfolding decreases with increasing temperature.²⁶ This has suggested that helical linkers propagate unfolding but also soften thermally before the repeats melt.

We focus here on forced unfolding pathways of linkers between tandem spectrin repeats. Illustration of extensive tandem repeat unfolding in AFM experiments is first provided for spectrin constructs that include dystrophin which, when missing from muscle, causes muscular dystrophy. Steered molecular dynamics (SMD) trajectories with explicit water²⁷ are then obtained for two different tandem spectrin repeats from α -actinin. As

with all other tandem spectrin structures, the linkers are initially helical which, upon closer inspection, appear shielded from hydration by intra-repeat loops. Past computational studies of spectrins have been limited to single spectrin repeats in implicit solvent.^{24,28} While the past studies have suggested unfolding intermediates within one repeat, the simulations in water here reveal pathway variations in unfolding the linker and are clearly based on hydration. The role of water-peptide interactions as well as interactions of the linkers with proximal loops in spectrin repeats are highlighted in Discussion. Overall, the results provide new insight into the unique cooperative unfolding of spectrin family proteins.

Results

Multi-repeat unfolding in AFM

Simultaneous unfolding of tandem repeats is not limited to just a pair of repeats as per Figure 1(a). For a four repeat β -spectrin construct as well as an eight repeat dystrophin construct (Figure 1(b)), histograms of the peak-to-peak unfolding length obtained from AFM experiments show an increasing number of three and even four repeat unfolding events. It is clear, from Figure 1, that the relative frequency of tandem unfolding events increases with the number of repeats, and that this increase is non-linear. The possible roles of the linkers in either propagating tandem repeat unfolding or else limiting it are elaborated below with SMD simulations on repeats 1–2 and repeats 2–3 of α -actinin; these repeats prove highly homologous to the first four repeats of β -spectrin presented here (see Methods). Implications on dystrophin are discussed later in order to gain a better understanding of these novel unfolding results.

Unfolding tandem repeats in two ways

In the simulations we focus only on the initial extension of tandem repeats. At least two pathways are evident, depending on both sequence and rate of extension. The distinction is whether the helical linker is disrupted or remains intact, as per Figure 2. While the three helices in each repeat tend to stay intact under these stresses, the linker proves susceptible to disruption in a large fraction of the trajectories. When the linker loses its helicity, the two-repeat structure develops a noticeable kink or bend that is indicative of a stretch-splay coupling with localized unwinding of the linker (Figure 2*i*). In contrast, when the linker stays helical, the stretched protein structure remains relatively straight (Figure 2*ii*).

The extension rate is known to generally influence unfolding results,^{7,29,30} including results of SMD simulations,^{29–32} and numerous efforts have addressed these effects. Although the rates used in SMD are normally four to five orders of magnitude

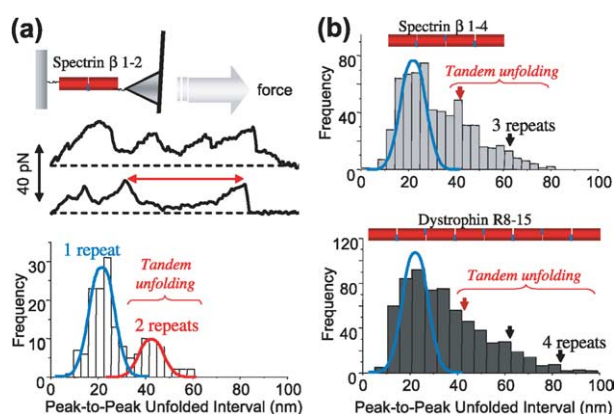


Figure 1. Single molecule experiments showing forced unfolding of serial spectrin repeats, with both single repeat and tandem repeat pathways of unfolding. (a) AFM forcibly extends a two-repeat construct, giving sawtooth patterns of forced unfolding. Tandem unfolding occurs as both repeats unfold at once, giving unfolding curves with twice the unfolding length. (b) Four repeat spectrin or eight repeat dystrophin constructs exhibit similar behavior, except more tandem repeat unfolding is seen.

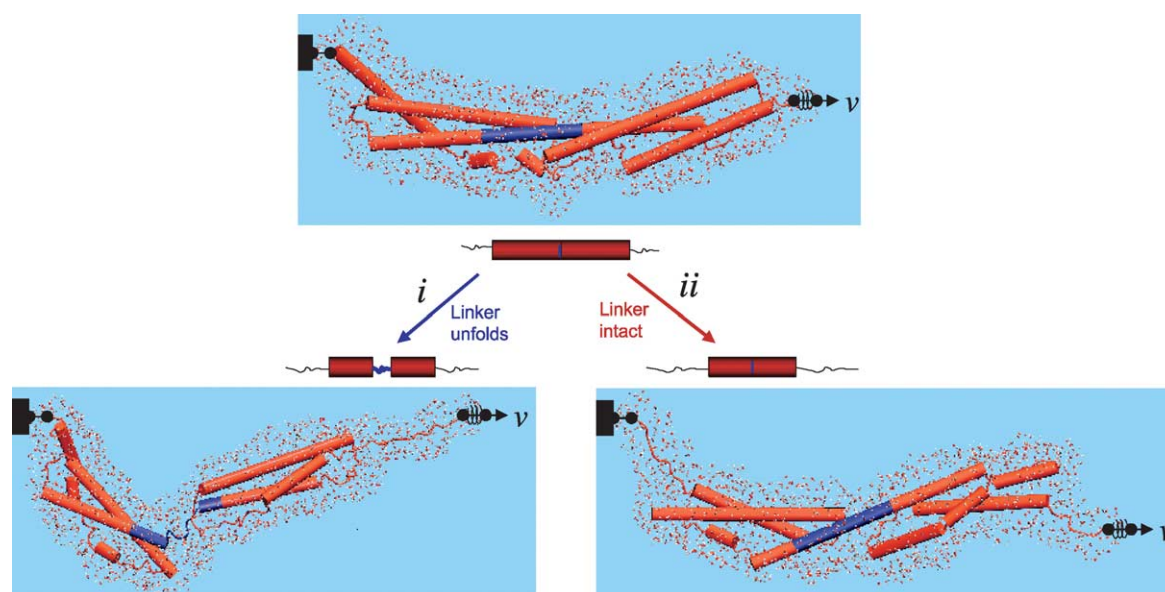


Figure 2. Two initial unfolding pathways for serial spectrin repeats extended under force. An intact or disrupted linker distinguishes the two pathways. Five layers of hydration are shown to emphasize that the simulations are done in explicit water (linker and repeats are shown in blue and red, respectively).

higher than those used in experiments, SMD simulations have given many insights into the molecular mechanisms and energetics of different processes, such as protein folding/unfolding pathways,³³ ligand binding,²⁹ and the passage of small molecules through membrane channels.³⁴

Two pulling rates that differ by an order of magnitude were therefore used to probe the unfolding pathways here. The two pulling rates of 0.5 nm/ns and 5 nm/ns listed in Table 1 yield different frequencies for unfolding the linker between tandem repeats. This is particularly important to show that the linker between repeats 1–2 can either remain intact or unfold, depending on the stresses. Common to both pathways at the slowest rates of pulling of 0.5 nm/ns, the earliest stages of extension show helix unwinding of both the N and C-terminal residues. The translationally immobilized but torsionally free N terminus unwinds while the C terminus is pulled upon. This opposite end effect demonstrates stress propagation through the folded system. At the highest pulling rates, however, the applied force localizes: the C-terminal end (closest to the spring) unwinds but the linker also breaks while the bulk of the protein remains intact.

Table 1. Linker unfolding frequency for serial spectrin repeats at two different pulling rates

Pulling rate (nm/ns)	Linker unfolding frequency (%)	
	Repeats 1–2	Repeats 2–3
5.0	66 (3 runs)	100 (3 runs)
0.5	0 (4 runs)	100 (5 runs)

Breaking the linker helix

For each run in which the linker unfolds, the distance between two chosen residues on either side of the unfolding region is determined as a function of time (Figure 3(a)). For all of the runs, a sudden increase in this distance corresponds to nucleated unfolding of the linker. The discontinuity occurs at different times in the simulations and between different linker residues, although it will be shown below that the unfolding is nonetheless restricted to a small segment of about ten residues in the linker. Figure 3(b) shows the fractional helical content of the linker (see Methods) averaged over all runs. The discontinuity is blurred slightly by ensemble averaging, but the plot clearly summarizes the fact that the linker unfolds under an applied force with loss of about 25% of its helicity in the 2.5 ns trajectory. Conversely, the linker that remains intact conserves its helicity throughout the entire simulation. These two pathways are illustrated in Figure 2 and are clearly related to Figure 1. Loss of linker helicity under force limits a helix–coil transition of adjacent domains and thus yields single-repeat unfolding. An intact linker allows the cooperative transition of sequential domains that sometimes extends beyond just one additional repeat.

Increased hydration of the linker

All past simulation studies of spectrin repeats have been limited to single repeats and to implicit solvent.^{24,28} The present study is the first to include explicit water as well as tandem spectrin repeats. Figure 2 shows five layers of protein hydration. This hydration will clearly influence surface exposure as well as solvent–peptide dynamics and pathways.

In the case when the linker unfolds, two proximal

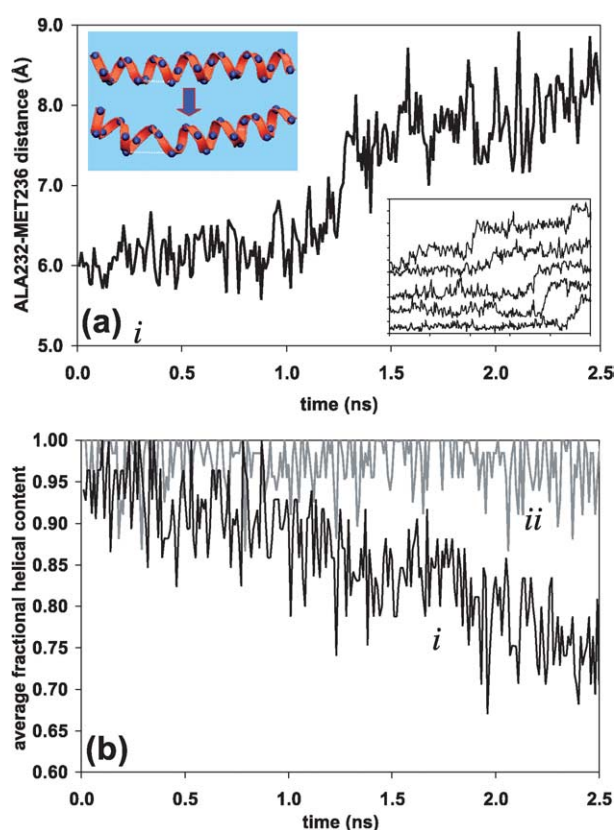


Figure 3. Stretch-induced twist and unfolding of the helical linker. (a) The distance between the α -carbon atoms of linker residues 232 and 236 (marked in the upper inset) is shown as a function of time for a single run of tandem repeats 2–3. The sudden increase in distance at a time between 1.25 ns and 1.5 ns corresponds to unfolding of the linker. The lower inset shows that such a sudden increase is characteristic of linker unfolding for all of the runs, although it occurs at different times and between different linker residues (curves were offset for clarity). (b) Fractional helical content as a function of time, averaged over all runs for each of the two linkers studied. The linker connecting repeats 1 and 2 retains its α -helical character (*ii*), while the linker connecting repeats 2 and 3 (*i*) loses an average of 25% of its helicity in a time period of 2.5 ns at a pulling rate of 5 Å/ns.

loops from each of the repeats that normally sequester and “protect” the linker from solvent are seen to lift away under the applied force as shown in Figure 4(a)*i*. In contrast, for the pathway where the linker remains intact, Figure 4(a)*ii* shows that the linker remains sequestered from solvation throughout the simulations.

The effect shown in Figure 4(a) is quantified in Figure 4(b), which shows the change in linker surface area exposed to solvent water as a function of time, after averaging over all the runs for the two systems studied (see Methods, equation (1)). The linker connecting repeats 2–3 shows five- to sixfold greater exposure than the linker connecting repeats 1–2. The contact between the linker and the sequestering loops is primarily between hydrophobic residues; this loss of tertiary structure

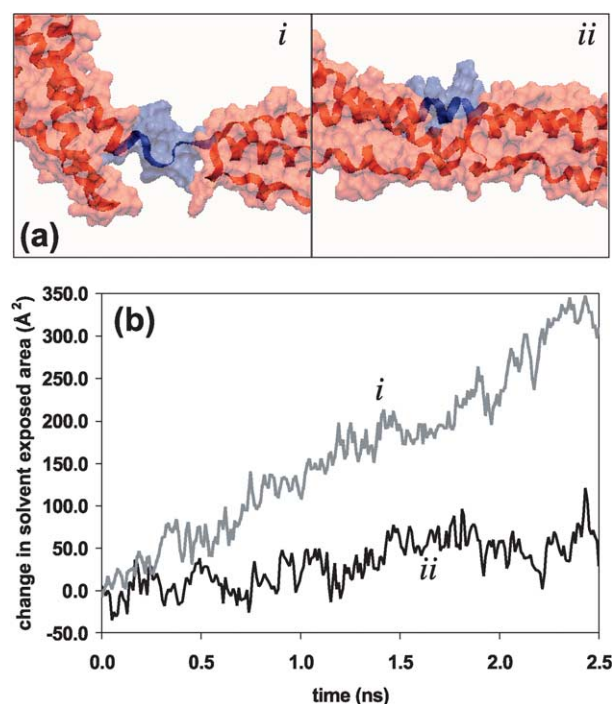


Figure 4. Destabilization of the linker with solvent exposure. (a) Solvated surface near the linker showing either unfolding of the linker region (*i*: repeats 2–3) or an intact linker region (*ii*: repeats 1–2) (repeats are identified with red and linkers with blue). (b) Average change in solvent-exposed area (referenced to $t=0$ ns) for the linker as a function of time. The reduction in shielding from the solvent by the protecting loops destabilizes the linker.

(contrast Figure 4(a)*i* and (a)*ii*) exposes these hydrophobic residues to water. This destabilizes the linker, contributing to the loss of its helicity in accord with prevailing views of protein folding.^{35–37}

Residue-by-residue, the trends in exposed area discussed above are also apparent. Figure 5(a) shows the change in water exposure for each residue in the linker at different times of a single simulation (see equation (2)). The increase in exposure of the linker to the solvent is clearly not uniform over the linker’s length. The linker breaks as it divides into two regions separated by a sharp interface of increased solvation. One of the regions remains shielded from the solvent at all times while the other becomes increasingly exposed to the solvent as the protein is extended at a constant velocity. The position of the interface in the linker identifies the first place at which the linker unfolds. Two key residues identified with the unfolding of the linker in Figure 3(a) (residues 232 and 236) are located on either side of the interface and are indicated with boxes in Figure 5(a). Notice that residues Ala232 and Leu233 are hydrophobic and subject the linker to opposing forces due, on the one hand, to the force being applied to the protein, which causes the loops to separate from the linker, and, on the other hand, to the effective force which resists solvent exposure of hydrophobic residues.

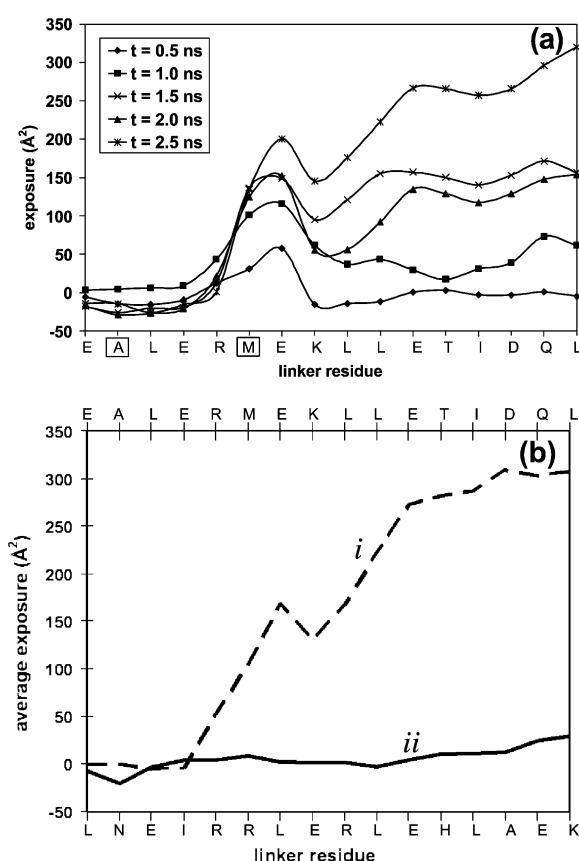


Figure 5. Exposure to water along the linker sequence. (a) Solvation profile at different times for the same run shown in Figure 3(a). A sharp interface develops, separating a protected and an increasingly exposed region of the linker. It is at this interface that the unfolding of the linker begins. Residues 232 and 236 in Figure 3(a) are shown in boxes here and are located across the interface. (b) Average exposure for the final configuration over all low extension rate runs for the two systems studied (see the text for details). The linker connecting repeats 2–3 has an exposure that is about one order of magnitude greater than the linker connecting repeats 1–2. Linker sequences for repeats 1–2 and 2–3 are shown on the bottom and top axis, respectively.

This localized stress eventually destabilizes the peptide, disrupting key hydrogen bonds within the helix. For the case of repeats 1–2, no such stress develops because the linker remains uniformly protected by the loop regions over the entire length of the simulation. This is quantified in Figure 5(b) where the average final exposure over all low extension rate runs is shown on a residue by residue basis for both systems. The maximal difference in solvent exposure over the length of the linker is an order of magnitude greater for repeats 2–3 compared to repeats 1–2.

Supplemental movies demonstrate, with a simple, average “strain” metric and a suitable color-code, that local deformations within the tandem repeat 2–3 are greatest in both the linker and adjacent loop regions (see Methods and

Supplementary Data). Quantifying the local deformations at the residue level has proven successful in probing the mechanical properties of proteins,³⁸ here, the local deformations are calculated at each residue’s α -carbon atom from the mean change in distance to the nearest ten α -carbon atoms in the starting structure. These deformations are converted to directional strains by dividing out the initial mean distance and resolving into parallel or perpendicular components with respect to the pulling direction. Whereas repeats 1–2 develop little internal strain except for end effects, the movies for repeats 2–3 convey clearly the linker unwinding upon the loss of native contacts. A large perpendicular strain in the linker arises as the protecting loops displace away from the pulling direction (see Figure 6 snapshots from movie strainperp2-3.mpg). Shown as spheres in Figure 6 are the two residues highlighted in Figures 3(a) and 5(a). The strain analysis is consistent with the picture emerging from these earlier Figures in showing that residue 232 remains unchanged (green, no strain) throughout the simulation while the strain on residue 236 increases. In addition, since most of the extension occurs in the parallel direction, a high parallel strain is observed for one of the loops that initially sequesters the linker (see strainpar2-3.mpg). The movies thus highlight strain localization to the linker, and suggest analogies to a widening crack that forms as the sequestering loops pull away and expose the linker to water.

For helical peptides, approximately one water molecule is found in the solvation shell of each carbonyl oxygen atom, whereas for a well-hydrated, unfolded configuration of a peptide backbone, an average of about two water molecules coordinate each carbonyl oxygen atom.³⁹ The number of water oxygen atoms in the first hydration shell of the linker’s backbone carbonyl oxygen atoms (see Methods) was therefore tracked as a change in hydration for all of the linker residues (Figure 7(a)). Figure 7(b) shows the final change in hydration for all runs of the system that exhibit linker unfolding. The inset snapshots show the change in backbone-localized water molecules near the break where the linker initially unfolds. For the system that does not exhibit linker unfolding, hydration remains constant throughout the entire simulation.

Net increases in backbone hydration are restricted to residues involved in the initial linker unfolding, regardless of ultimate solvent exposure. In other words, backbone hydration increases for only the small number of key residues in the linker compared to the number of residues that become exposed to the solvent (compare Figures 5 and 7). This may seem paradoxical, but even when the linker’s surface is exposed, the peptide’s side-chains protect the backbone from hydration. Ghosh *et al.*³⁹ showed that some side-chains are more effective than others at shielding the backbone from water. Not only must the loop-protected helix become exposed to the solvent, but it must also

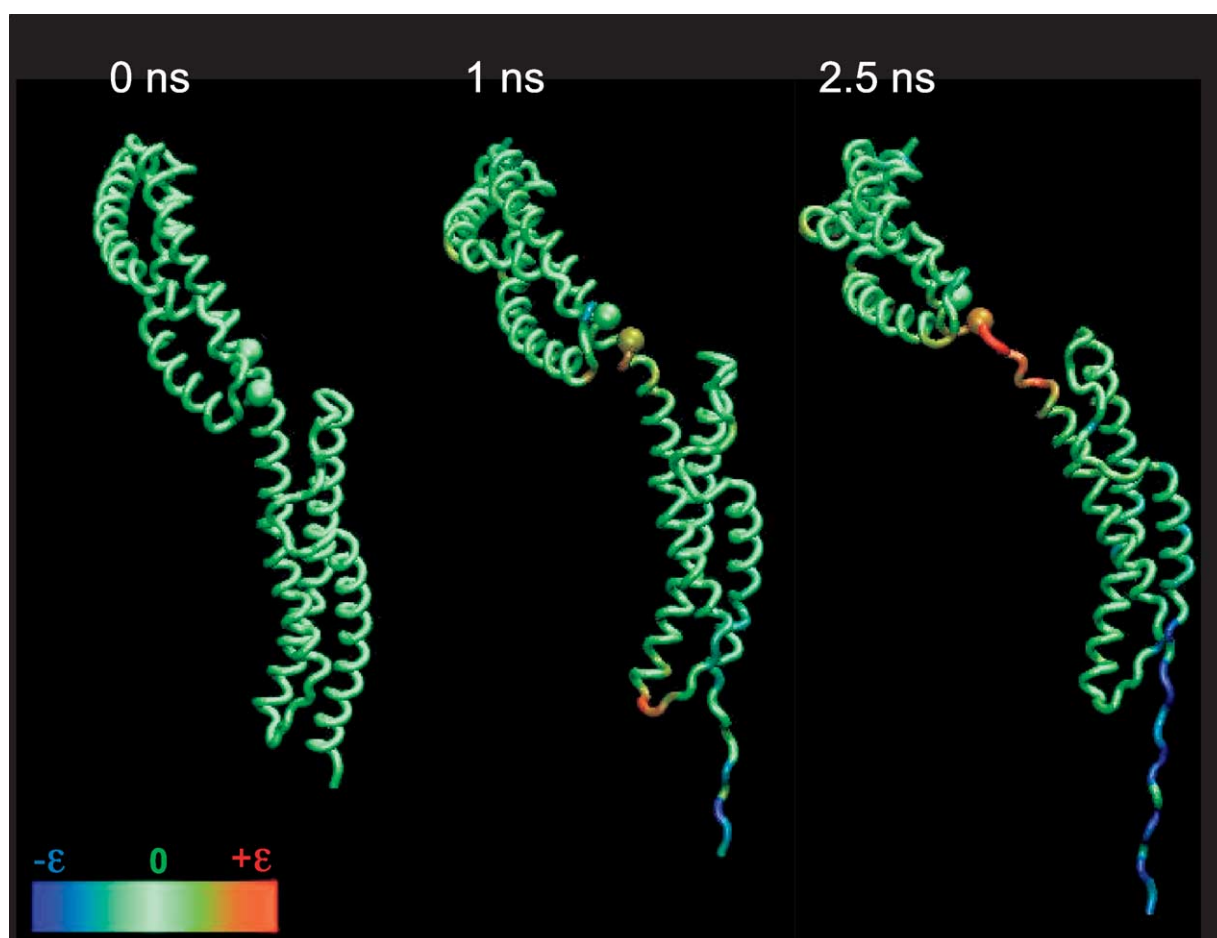


Figure 6. Snapshots from the movie *strainerp2-3.mpg* (see Supplementary Data) showing the strain flow perpendicular to the pulling direction in tandem repeats 2–3 of α -actinin when pulled at 0.5 nm/ns. A high strain value in the linker is a consequence of the separation of the protecting loops, which move away from each other in the perpendicular direction. The two residues highlighted in Figures 3(a) and 5(a) are shown here as spheres. No strain develops at residue 232, while the strain on residue 236 increases with time. Green coloration represents zero strain, blue negative strain (compression), and red positive strain (extension).

unfold in order for water to enter and hydrogen bond to carbonyl oxygen atoms of the backbone. Since water can hydrate the backbone only after the helix has made the transition to the unfolded state, we conclude that backbone hydration is a consequence rather than a cause of the unfolding. From Figure 7(b) we see that all of the unfolding events are constrained to a small region of about ten residues. This observation leads to a more refined definition of the linker compared to conventional definitions based on visual inspection of helical repeat boundaries.

Forces of unfolding

Forced unfolding in SMD is typically carried out on single-domain proteins and up to full unfolding of the contour, but this can be done only at rates of unfolding that are far higher than experiments. The slower the pulling and the larger the protein, the more relevant the simulation is to experiment.^{7,29–32}

The focus of the work here on tandem repeat systems has been on the inter-repeat helical linker and only in the very initial stages of unfolding: within 3 nm of end-to-end extension of a tandem repeat. At a more advanced stage, the extension of previously unfolded domains appears to fit a worm-like chain model,²³ which is a generic model that omits any detail on either the primary structure of the protein or residual secondary structure. These two stages are presented in Figure 8; the force-extension data for the initial stage are shown in Figure 8(a), and its location in the full force-extension profile is indicated in Figure 8(b). The force-extension profile from SMD is nearly flat, which is in good accordance with its identification within the schematic sawtooth of Figure 8(b). The data in Figure 8(a) are an overlay of all runs for both repeats 1–2 and repeats 2–3, and the similarity in the data between these two systems indicates the independence of the force on the unfolding pathway of the protein, which is consistent with experiment.⁹

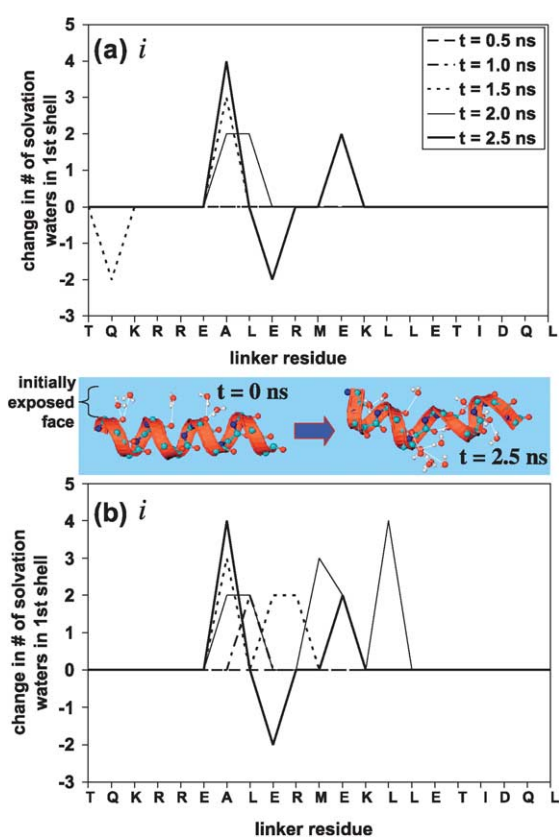


Figure 7. Change in the number of water oxygen atoms within the first hydration shell (of 3.6 Å) for the carbonyl oxygen of each residue in the linker (relative to $t=0$ ns). (a) Change in hydration at different times for a single run. Increased hydration occurs predominantly on the same face of the unfolding linker (see color picture), which further identifies the “unfolding interface” along the linker (see Figure 5(a)). (b) Final change in hydration for all runs of repeats 2–3; data for repeats 1–2 are not shown because there is no change in hydration from the initial configurations. Unfolding events in the linker connecting repeats 2–3 are concentrated in a region of about ten residues (231–240). Changes less than or equal to one are suppressed for clarity.

Discussion

What is it about linkers between domains such as spectrin repeats that makes them stable some of the time but almost as often unfoldable under force (Figures 1 and 2)? SMD results here suggest that peptide–water interactions are prominent in unfolding processes, with hydration playing a critical role both before and after unfolding. Before unfolding, highly localized changes in linker exposure to solvent contribute to unfolding by placing an additional stress on the helical linker (Figure 3). Water infiltration is the likely cause for this destabilization (Figures 4 and 5). All of the residues that are on the face of the linker that is initially protected by the key proximal loop but later exposed (Figure 4(a)i), are hydrophobic (residues 232, 233, 236, 239, 240 and 243). Additionally, the

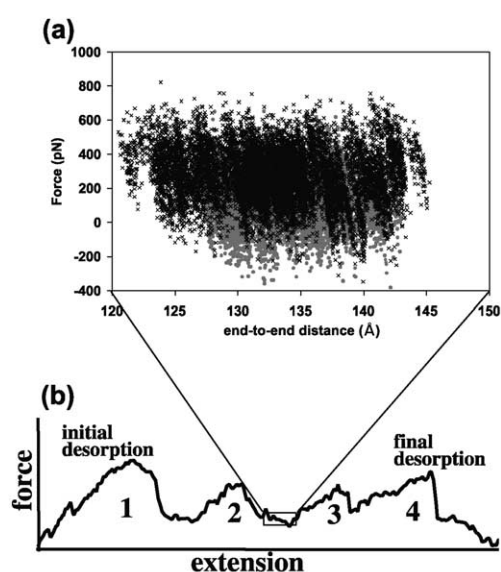


Figure 8. Profiles of force *versus* extension. (a) The data for all runs (● repeats 1–2, × repeats 2–3) and (b) a typical profile obtained from an AFM experiment. The range of extensions and forces covered in this work are very small relative to those covered in AFM experiments because of computational limitations. The simulation probes conformations of the protein that are representative of the marked region in (b), therefore the force is observed to remain nearly constant throughout the simulation. Note that the forces are similar for repeats 1–2 and repeats 2–3, indicating that the force is independent of the unfolding pathway.

residue sitting next to the sharp interface in Figure 5 but on the side that remains more protected is hydrophobic. The interfacial stress that develops as the hydrophobic residues are exposed to water appears relieved by backbone hydration. This is indeed quantified by the increased hydration of the clearly unfolded backbone (Figure 7). This scenario is supported by our observation that unfolding and hydration occur precisely at the point where the linker loses contact from the splaying, protective loops. Based on a recent crystal structure of tandem repeats 8–9 from human erythrocyte β -spectrin, MacDonald and co-workers have described similar and important roles for water in destabilization of individual repeats.²¹

As mentioned in Results, residue side-chains play an important role in determining water access to the backbone after unfolding. This result agrees with findings on β -sheet domains subjected to forced unfolding in SMD simulations.⁴⁰ Specifically, by performing simulations on FnIII modules from the cell adhesion protein fibronectin, which have natural variations in their amino acid sequences, the mechanical stability was found to be tuned by substitutions of just a few key amino acid residues. These alter the access of water molecules to hydrogen bonds that break early in the unfolding pathway. There is a clear ordering of events, which is emphasized both in this work and in the fibronectin work; namely, unfolding of the backbone

followed by stabilization of the unfolded structure through the formation of hydrogen bonds to water molecules.

Hydrogen bond interactions between the backbone and water molecules stabilize the unfolded state of the backbone compared to vacuum or low dielectric mediums.^{39,41} Water makes it relatively easier for the protein to move repeatedly between folded and unfolded states, which highlights the physiological function of spectrin family proteins in repeated extension/compression cycles.^{13,14}

AFM results on various spectrin family proteins in Figure 1 show that tandem unfolding (unfolding of two repeats at the same time) occurs with relatively high frequency and at essentially no extra cost in free energy.⁹ On the basis of the SMD results here, the two main pathways observed in AFM (single-repeat and tandem-repeat unfolding) relate to whether the linker unfolds or stays intact. For all trajectories at the slower pulling rate, the linker between repeats 1–2 never unfolded while the linker between repeats 2–3 always unfolded. It would be tempting to relate this to primary structure of the linker and secondary structure predictions, but MacDonald and co-workers have suggested that such predictions are perhaps better correlated with unfolding temperature in solution than with linker helicity in crystal structures.^{21,22} Secondary structure predictions have a basis in energetics and structure rather than energy gradients and protein extensibility under force. Broadening structure predictions to somehow include the effects of forced extension would be useful to better understand spectrin family proteins, since they are mechanically important proteins that exhibit a large natural variability in their linker sequences. Insights from crystals are important but come from only a very small subset of this protein family and under fixed, static constraints. The main goal of this study was to provide an explanation for the results obtained in the AFM experiments and we believe this goal has been achieved. We understand that, because of the limitations of both experimental and computational methods, a direct relationship cannot yet be established between the results and the behavior of spectrin family proteins in their native environment. However, we are convinced that this work provides an important piece of information about the mechanical properties of these proteins, which complements existing understanding (such as the cited work of MacDonald) and is an additional step towards reaching a complete description of their structure and function in the human body.

Atomistic simulations, if not structural predictors that incorporate mechanical function, could also be exploited in functional designs of hinges for mini-dystrophins in gene therapy. Native dystrophin has 24 spectrin repeats with several proline-rich hinge regions that likely limit α -helix formation;⁴² these hinges have recently been shown to be critical to success in gene therapy.⁴³ Linker sequence is clearly not the only determinant of stability, since we have demonstrated here by SMD that intra-domain loops in particular play a critical role in steric stability

through linker sequestration. The full molecular detail of the entire protein afforded by simulation thus allows insight into key structural and solvation mechanisms at play.

Methods

AFM experiments

Forced extension unfolding was done on purified recombinant α -spectrin repeats 1–2 and 1–4, as described elsewhere.⁹ A recombinant dystrophin construct consisting of spectrin repeats 8–15 in the middle of this protein's "rod" domain was purified as described elsewhere⁴⁴ and studied under the same conditions as the spectrin constructs.

Simulations

Tandem repeats 1–2 and repeats 2–3 of α -actinin were obtained (as residues 1–241 and 117–364, respectively) from the B-chain of the dimerized α -actinin rod domain pdb 1HCI.¹² The very high level of homology of these α -actinin repeats with the N-terminal repeats of human erythroid β -spectrin has been noted.⁴⁵ Both of these two-repeat segments were hydrated in a box of explicit water (TIP3P) for total system sizes exceeding 73,000 atoms. Each system was equilibrated at 300 K for a total time of 2.0 ns, from which uncorrelated configurations were taken as seeds for SMD simulation.

All simulations were carried out with the program NAMD⁴⁶ employing the CHARMM27 force field.⁴⁷ A time-step of 2 fs and a non-bonded force cutoff function starting at 10 Å and vanishing at 12 Å were used. The Shake algorithm was used to fix bonds with hydrogen, and full electrostatic interactions were incorporated with the Particle Mesh Ewald (PME) algorithm. For temperature and pressure control, the Langevin dynamics option of NAMD was used.

During the SMD simulations,^{27,33} the N terminus was immobilized and an external force was applied through a harmonic spring coupled to the C terminus of the protein. The force was applied by displacing at chosen rates the unattached end of the spring with the stretching direction chosen along the vector pointing from the N to the C terminus. The spring constant was 4840 pN/nm. The total SMD simulation time was ≥ 2.5 ns, and configurations were saved every 10 ps.

Simulation analysis

Distances between residues as well as the end-to-end distance are defined between the relevant C α backbone atoms. The linkers are nominally identified here as 21 amino acid residues in length: residues 117–137 for repeats 1–2 and residues 226–246 for repeats 2–3. Fractional helical content for each linker was calculated using the Lifson–Roig model,⁴⁸ i.e. a residue i is marked helical if, and only if, its backbone dihedral angle pair (ϕ_i , ψ_i), as well as its adjacent residue angle pairs (ϕ_{i-1} , ψ_{i-1}) and (ϕ_{i+1} , ψ_{i+1}), lie in the helical region $\phi = -65(\pm 35)^\circ$ and $\psi = -37(\pm 30)^\circ$. Profiles of the fractional helical content of the linker were calculated as a function of time and then averaged over all runs. The solvent-exposed area for the linker, A_{exp} , was calculated as:

$$A_{\text{exp}} = 1/2(A_{\text{protein}} + A_{\text{linker}} - A_{\text{repeat}}) \quad (1)$$

where A_{protein} is the molecular surface area for the entire

segment, A_{linker} is the same quantity for the linker residues only, and A_{repeat} is the area for all the residues in the protein not forming the linker; namely, the residues of the repeats. Molecular surface areas were obtained with the program MSMS.⁴⁹ The exposure, defined as:

$$\text{exposure}(t) = A_{\text{exp}}(t) - A_{\text{exp}}(t = 0) \quad (2)$$

was then determined for each residue in the linker at different times.

Hydration of the backbone was quantified by counting water oxygen atoms in the first solvation shell (within 3.6 Å) of backbone carbonyl oxygen atoms. The instantaneous hydration was thus measured for each residue at different times during the simulation. As done in calculating $\text{exposure}(t)$, initial values were subtracted from the instantaneous values in order to gauge the change in hydration.

Movies accompanying this manuscript can be found on the JMB website and show, in color code, how the forced extension applied at either end of both tandem repeats affects the distances between neighboring α -carbon atoms during the simulations. This form of strain can be directly related to the loss of native contacts in the protein during unfolding. The strain was calculated for the entire molecule on a residue-by-residue basis as the deformation observed in the group of ten α -carbon atoms closest to the reference atom at time zero. The criterion used to choose the ten neighbors is based solely on the initial spatial distance and not on the basis of primary structure connectivity. This group of ten atoms is followed throughout the simulation and the change in their positions relative to the reference atom forms the basis for the strain calculation. Specifically, the mean of the distances between each of the ten atoms and the reference atom was resolved into components parallel with and perpendicular to the forced-extension direction, and the strain was computed, at each time-step, as:

$$\varepsilon_i(t) = \frac{L_i(t) - L_i(0)}{L_i(0)} \quad (3)$$

where i is the residue index and L is either the perpendicular or the parallel component of the distances defined above.

The calculated values, $\varepsilon_i(t)$, were used to color code the molecule and were rendered into movies (from which Figure 6 was extracted) to visualize the strain flow through the protein. Four such movies are included in the Supplementary Data. The files *strainpar1-2.mpg* and *strainpar2-3.mpg* show the parallel strain for simulations on repeats 1–2 and 2–3, respectively, while *strainperp1-2.mpg* and *strainperp2-3.mpg* present the perpendicular strain results for the two constructs. The color scale is based on the blue-green-red gradation, where green represents zero deformation, blue represents a negative deformation (compression), and red represents a positive deformation (extension). It should be clarified that even though the biggest values of strain occur at the ends of the molecule, these are end-effects as discussed in Results.

Acknowledgements

This work was supported, in part, by the NSF IGERT program (grant DGE-0221664), the Center for Science and Engineering of Nanoscale Systems (SENS), and the National Institutes of Health (NIH).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2005.03.086

References

- Andreeva, A., Howorth, D., Brenner, S. E., Hubbard, T. J. P., Chothia, C. & Murzin, A. G. (2004). SCOP database in 2004: refinements integrate structure and sequence family data. *Nucl. Acids Res.* **32**, D226–D229.
- Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M. & Gaub, H. E. (1997). Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science*, **276**, 1109–1112.
- Kellermayer, M. S. Z., Smith, S. B., Granzier, H. L. & Bustamante, C. (1997). Folding–unfolding transitions in single titin molecules characterized with laser tweezers. *Science*, **276**, 1112–1116.
- Fernandez, J. M. & Li, H. B. (2004). Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science*, **303**, 1674–1678.
- Oberhauser, A. F., Marszalek, P. E., Erickson, H. P. & Fernandez, J. M. (1998). The molecular elasticity of the extracellular matrix protein tenascin. *Nature*, **393**, 181–185.
- Gao, M., Craig, D., Lequin, O., Campbell, I. D., Vogel, V. & Schulten, K. (2003). Structure and functional significance of mechanically unfolded fibronectin type III1 intermediates. *Proc. Natl Acad. Sci. USA*, **100**, 14784–14789.
- Evans, E. & Ritchie, K. (1997). Dynamic strength of molecular adhesion bonds. *Biophys. J.* **72**, 1541–1555.
- Scott, K. A., Steward, A., Fowler, S. B. & Clarke, J. (2002). Titin; a multidomain protein that behaves as the sum of its parts. *J. Mol. Biol.* **315**, 819–829.
- 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.
- Speicher, D. W. & Marchesi, V. T. (1984). Erythrocyte spectrin is comprised of many homologous triple helical segments. *Nature*, **311**, 177–180.
- Koenig, M., Monaco, A. P. & Kunkel, L. M. (1988). The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. *Cell*, **53**, 219–228.
- Ylanne, J., Scheffzek, K., Young, P. & Saraste, M. (2001). Crystal structure of the alpha-actinin rod reveals an extensive torsional twist. *Structure*, **9**, 597–604.
- Mohandas, N. & Evans, E. (1994). Mechanical-properties of the red-cell membrane in relation to molecular-structure and genetic-defects. *Annu. Rev. Biophys. Biomol. Struct.* **23**, 787–818.
- An, X. L., Lecomte, M. C., Chasis, J. A., Mohandas, N. & Gratzer, W. (2002). Shear-response of the spectrin dimer–tetramer equilibrium in the red blood cell membrane. *J. Biol. Chem.* **277**, 31796–31800.
- Lee, J. C. M. & Discher, D. E. (2001). Deformation-enhanced fluctuations in the red cell skeleton with theoretical relations to elasticity, connectivity, and spectrin unfolding. *Biophys. J.* **81**, 3178–3192.
- Yan, Y., Winograd, E., Viel, A., Cronin, T., Harrison, S. C. & Branton, D. (1993). Crystal-structure of the repetitive segments of spectrin. *Science*, **262**, 2027–2030.
- Pascual, J., Pfuhl, M., Walther, D., Saraste, M. &

- Nilges, M. (1997). Solution structure of the spectrin repeat: a left-handed antiparallel triple-helical coiled-coil. *J. Mol. Biol.* **273**, 740–751.
18. Grum, V. L., Li, D. N., MacDonald, R. I. & Mondragon, A. (1999). Structures of two repeats of spectrin suggest models of flexibility. *Cell*, **98**, 523–535.
19. Kliche, W., Fujita-Becker, S., Kollmar, M., Manstein, D. J. & Kull, F. J. (2001). Structure of a genetically engineered molecular motor. *EMBO J.* **20**, 40–46.
20. Djinovic-Carugo, K., Young, P., Gautel, M. & Saraste, M. (1999). Structure of the alpha-actinin rod: molecular basis for cross-linking of actin filaments. *Cell*, **98**, 537–546.
21. Kusunoki, H., Macdonald, R. I. & Mondragon, A. (2004). Structural insights into the stability and flexibility of unusual erythroid spectrin repeats. *Structure*, **12**, 645–656.
22. MacDonald, R. I. & Cummings, J. A. (2004). Stabilities of folding of clustered, two-repeat fragments of spectrin reveal a potential hinge in the human erythroid spectrin tetramer. *Proc. Natl Acad. Sci. USA*, **101**, 1502–1507.
23. Rief, M., Pascual, J., Saraste, M. & Gaub, H. E. (1999). Single molecule force spectroscopy of spectrin repeats: low unfolding forces in helix bundles. *J. Mol. Biol.* **286**, 553–561.
24. Altmann, S. M., Grunberg, R. G., Lenne, P. F., Ylanne, J., Raae, A., Herbert, K. *et al.* (2002). Pathways and intermediates in forced unfolding of spectrin repeats. *Structure*, **10**, 1085–1096.
25. Menhart, N., Mitchell, T., Lusitani, D., Topouzian, N. & Fung, L. W. M. (1996). Peptides with more than one 106-amino acid sequence motif are needed to mimic the structural stability of spectrin. *J. Biol. Chem.* **271**, 30410–30416.
26. Law, R., Liao, G., Harper, S., Yang, G. L., Speicher, D. W. & Discher, D. E. (2003). Pathway shifts and thermal softening in temperature-coupled forced unfolding of spectrin domains. *Biophys. J.* **85**, 3286–3293.
27. Izrailev, S., Stepaniants, S., Isralewitz, B., Kosztin, D., Lu, H., Molnar, F. *et al.* (1998). In *Steered Molecular Dynamics* (Deufhard, J., Hermans, J., Leimkuhler, B., Mark, A. E., Reich, S. & Skeel, R. D., eds), vol. 4, Springer, Berlin.
28. Paci, E. & Karplus, M. (2000). Unfolding proteins by external forces and temperature: the importance of topology and energetics. *Proc. Natl Acad. Sci. USA*, **97**, 6521–6526.
29. Izrailev, S., Stepaniants, S., Balsera, M., Oono, Y. & Schulten, K. (1997). Molecular dynamics study of unbinding of the avidin–biotin complex. *Biophys. J.* **72**, 1568–1581.
30. Balsera, M., Stepaniants, S., Izrailev, S., Oono, Y. & Schulten, K. (1997). Reconstructing potential energy functions from simulated force-induced unbinding processes. *Biophys. J.* **73**, 1281–1287.
31. Heymann, B. & Grubmuller, H. (1999). AN02/DNP-hapten unbinding forces studied by molecular dynamics atomic force microscopy simulations. *Chem. Phys. Letters*, **303**, 1–9.
32. Li, P. C. & Makarov, D. E. (2003). Theoretical studies of the mechanical unfolding of the muscle protein titin: bridging the time-scale gap between simulation and experiment. *J. Chem. Phys.* **119**, 9260–9268.
33. Lu, H. & Schulten, K. (1999). Steered molecular dynamics simulations of force-induced protein domain unfolding. *Proteins: Struct. Funct. Genet.* **35**, 453–463.
34. Jensen, M. O., Park, S., Tajkhorshid, E. & Schulten, K. (2002). Energetics of glycerol conduction through aquaglyceroporin GlpF. *Proc. Natl Acad. Sci. USA*, **99**, 6731–6736.
35. Plotkin, S. S. & Onuchic, J. N. (2002). Understanding protein folding with energy landscape theory - Part I: Basic concepts. *Quart. Rev. Biophys.* **35**, 111–167.
36. Onuchic, J. N. & Wolynes, P. G. (2004). Theory of protein folding. *Curr. Opin. Struct. Biol.* **14**, 70–75.
37. Daggett, V. & Fersht, A. (2003). The present view of the mechanism of protein folding. *Nature Rev. Mol. Cell Biol.* **4**, 497–502.
38. Navizet, I., Cailliez, F. & Lavery, R. (2004). Probing protein mechanics: residue-level properties and their use in defining domains. *Biophys. J.* **87**, 1426–1435.
39. Ghosh, T., Garde, S. & Garcia, A. E. (2003). Role of backbone hydration and salt-bridge formation in stability of alpha-helix in solution. *Biophys. J.* **85**, 3187–3193.
40. Craig, D., Gao, M., Schulten, K. & Vogel, V. (2004). Tuning the mechanical stability of fibronectin type III modules through sequence variations. *Structure*, **12**, 21–30.
41. Lu, H. & Schulten, K. (2000). The key event in force-induced unfolding of titin's immunoglobulin domains. *Biophys. J.* **79**, 51–65.
42. Zarrinpar, A. & Lim, W. A. (2000). Converging on proline: the mechanism of WW domain peptide recognition. *Nature Struct. Biol.* **7**, 611–613.
43. Harper, S. Q., Crawford, R. W., DelloRusso, C. & Chamberlain, J. S. (2002). Spectrin-like repeats from dystrophin and alpha-actinin-2 are not functionally interchangeable. *Hum. Mol. Genet.* **11**, 1807–1815.
44. Kahana, E., Flood, G. & Gratzer, W. B. (1997). Physical properties of dystrophin rod domain. *Cell Motil. Cytoskeleton*, **36**, 246–252.
45. Ursitti, J. A., Kotula, L., DeSilva, T. M., Curtis, P. J. & Speicher, D. W. (1996). Mapping the human erythrocyte beta-spectrin dimer initiation site using recombinant peptides and correlation of its phasing with the alpha-actinin dimer site. *J. Biol. Chem.* **271**, 6636–6644.
46. Kale, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N. *et al.* (1999). NAMD2: Greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **151**, 283–312.
47. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J. *et al.* (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, **102**, 3586–3616.
48. Lifson, S. (1961). Theory of helix–coil transition in polypeptides. *J. Chem. Phys.* **34**, 1963–1974.
49. Sanner, M. F., Olson, A. J. & Spehner, J. C. (1996). Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers*, **38**, 305–320.

Edited by B. Honig

(Received 7 October 2004; received in revised form 14 March 2005; accepted 31 March 2005)