doi:10.1016/j.jmb.2005.07.064 J. Mol. Biol. (2005) 352, 795–806

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Molecular Extensibility of Mini-dystrophins and a Dystrophin Rod Construct

Nishant Bhasin†, Richard Law†, George Liao, Daniel Safer Jennifer Ellmer, Bohdana M. Discher, H. Lee Sweeney and Dennis E. Discher*

Pennsylvania Muscle Institute and Graduate Groups in Physics and Cell & Molecular Biology, University of Pennsylvania, Philadelphia, PA 19104 USA

Muscular dystrophies arise with various mutations in dystrophin, implicating this protein in force transmission in normal muscle. With 24 three-helix, spectrin repeats interspersed with proline-rich hinges, dystrophin's large size is an impediment to gene therapy, prompting the construction of mini-dystrophins. Results thus far in dystrophic mice suggest that at least one hinge between repeats is necessary though not sufficient for palliative effect. One such mini-dystrophin is studied here in forced extension at the single molecule level. Δ2331 consists of repeats (R) and hinges (H) H1-R1-2∼H3∼R22-24-H4 linked by native (−) and non-native (∼) sequence. This is compared to its core fragment R2∼H3∼R22 as well as an eight-repeat rod fragment middle (RFM: R8-15). We show by atomic force microscopy that all repeats extend and unfold at forces comparable to those that a few myosin molecules can generate. The hinge regions most often extend and transmit force while limiting tandem repeat unfolding. From 23–42 °C, the dystrophin constructs also appear less temperature-sensitive in unfolding compared to a well-studied βI-spectrin construct. The results thus reveal new modes of dystrophin flexibility that may prove central to functions of both dystrophin and mini-dystrophins.

*Corresponding author

Keywords: dystrophin; spectrin; AFM; protein folding; hinge

Introduction

Dystrophin localizes to the plasma membrane in contractile myocytes, and is one of the largest known spectrin-repeat, actin-binding proteins. Dystrophin is also encoded by the largest gene in the genome and because of mutations that cause muscular dystrophy (MD), considerable effort is being spent on the engineering of functionally equivalent "minimal" dystrophins.[1–10] Dystrophin functions as a monomer in linking cortical actin and the contractile apparatus to the dystroglycan complex that mediates adhesion (Figure 1(a)).[11,12] While the nature of dystrophin's flexibility and extensibility in relation to its structure is unclear, the mini-dystrophins tested thus far in mice do seem to indicate that a simple concatenation of spectrin repeats is functionally inadequate.

Dystrophin has a central rod domain with 24 spectrin repeats[13,14] that accounts for most of its folded contour length of ∼150 nm. Interspersed between dystrophin domains are four proline-rich, proteolysis-sensitive hinge regions, designated H1–H4,[14] that appear relatively unique among spectrin family proteins. Although there are no crystal structures yet for dystrophin repeats, all of the recent tandem-repeat structures for related α-actinin and spectrins reveal continuous helical linkers between repeats (Figure 1(a) inset).[15–18] Interestingly, for the linker-hinge H3 with distributed proline residues, secondary structure predictions as recently applied to spectrin by McDonald and co-workers,[19] suggest that H3 is purely helical between R2 and R22. Perhaps H3 is a proline-kinked helix as seen in the middle B-helix of several spectrin crystal structures.[17] This putative helicity of H3 is examined here with force and structural probes.

Naturally expressed mini-dystrophins give rise to mild forms of MD and provide further rationale

† R.L. and N.B. contributed equally to this work.

Abbreviations used: aa, amino acids; AFM, atomic force microscopy; RFM, rod fragment middle; MD, muscular dystrophy.

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

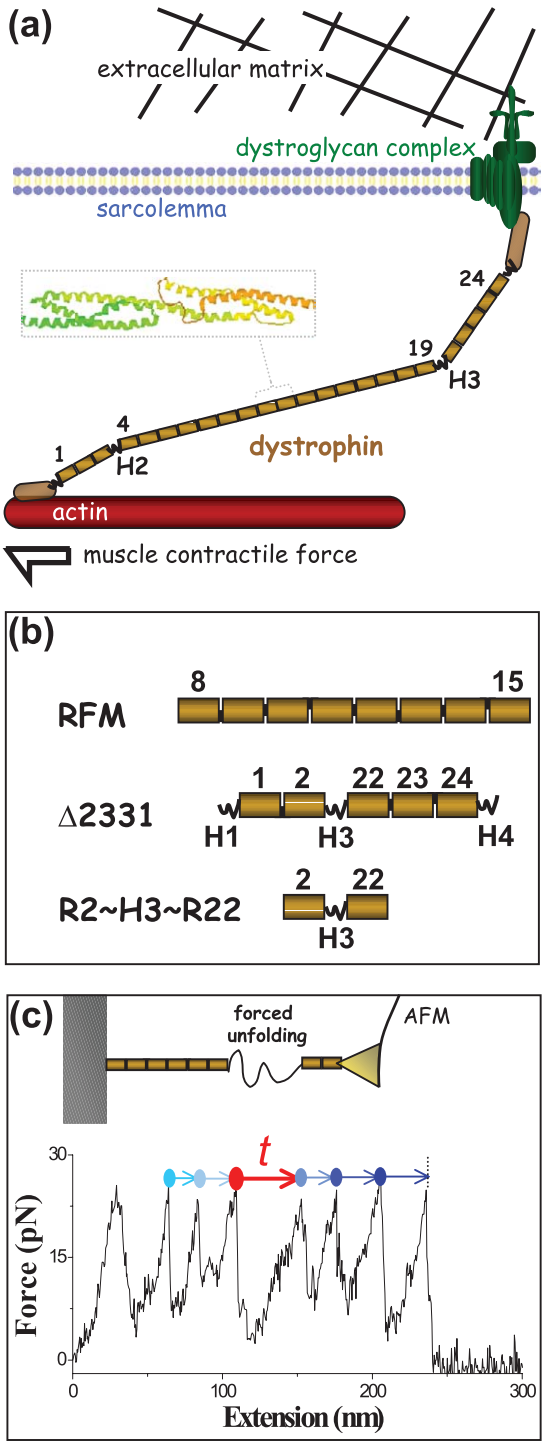

Figure 1. Schematic of dystrophin constructs investigated by AFM force spectroscopy. (a) In myocytes, dystrophin with 24 spectrin repeats and several hinges (H) links the cell membrane to cortical actin. Boxed inset shows tandem repeats 2–3 from α-actinin,[15] in which the triple helix repeats are linked by a continuous helix. (b) Dystrophin RFM consists of repeats R8–15; Δ2331 is a mini-dystrophin with five repeats and three hinges; and the R2∼H3∼R22 construct is the central structure of Δ2331 with a single hinge (see Table 1). (c) Extensible unfolding by AFM. Representative sawtooth-shaped force-extension pattern for RFM. One peak-to-peak extension is especially long and is denoted with a *t* for tandem repeat unfolding.

for the engineering of mini-dystrophins.[20] From gene therapy studies of dystrophin-deficient mice, a critical rod length of at least four repeats appears to be required to functionally bridge the cytoskeleton and the plasma membrane,[4,6] although a more sequence-specific mechanical role for a shortened rod also seems crucial.[4] Additionally, all of the mini-dystrophin constructs that have shown some success contain one or more proline-rich hinge regions both within the shortened rod and at its ends. This includes a mini-dystrophin composed of five repeats and three hinges, Δ2331 (H1-R1-2∼H3∼R22-24-H4; note that ∼ denotes a non-native linkage), with domains added at either end to bind actin and dystroglycan. While this construct has proven at least partially palliative in dystrophin-deficient mice following viral gene transfer,[7] similar constructs have had more limited effect (e.g. R1-2∼R22-24) and imply hinges are key. From proteolysis and electron microscopy, the hinge regions appear to be highly accessible.[14] We therefore seek to understand how these hinge regions influence the molecular extensibility and stability of dystrophin.

Start-stop positions or phasings of the repeats also appear to be of paramount importance to mini-dystrophin function.[21] Recent antisense-based therapies of exon skipping[22] undoubtedly disrupt phasing and are very likely to affect folding, since dystrophin's repeats are typically encoded by more than one exon. Moreover, not all three-helix spectrin repeats are functionally interchangeable: α-actinin's four repeat rigid rod domain, for example, shows little success in replacing dystrophin's repeats. α-Actinin is also well known to dimerize antiparallel, which stiffens the molecule and severely limits extension.[23]

Successes as well as failures in gene therapy together with basic questions of structure-function motivate our single molecule studies here of three dystrophin and mini-dystrophin fragments (Figure 1(b) and (c)). Dystrophin's rod fragment middle (RFM) consists of repeats R8–15[21] and serves here as a hinge-less, native

Table 1. Primary structure properties of dystrophin constructs

Domain	Number of aa	Contour, l_c (nm)
Hinge 1	84	31.0
R1	111	41.0
R2	108	39.9
R8	123	45.4
R9	104	38.4
R10	96	35.4
R11	105	38.7
R12	108	39.9
R13	102	37.6
R14	96	35.4
R15	99	36.5
Hinge 3	47	17.3
R22	116	42.8
R23	129	47.6
R24	109	40.2
Hinge 4	72	26.6
Avg for R	108	41.0

Domain contour lengths, l_c, have been calculated with a peptide length of 0.37 nm.

construct with a mean number of amino acid (aa) residues per repeat of 104 (Table 1) that is close to the canonical 106 aa for spectrin repeats. The representative mini-dystrophin Δ2331 construct (i.e. H1-R1-2∼H3∼R22-24-H4) has the basic engineered hinge region R2∼H3∼R22 in the middle with a 47 aa long hinge engineered between two spectrin repeats (of near-typical length). In thermal denaturation, the RFM construct had been shown in the past to be more stable than a construct similar to Δ2331.[21] We extend such studies here to the core construct R2∼H3∼R22 and probe these molecules under force by atomic force microscopy (AFM), showing that the short dystrophin construct R2∼H3∼R22 has a stability similar to the RFM.

Results

Extension of the dystrophin RFM by AFM shows the often-seen sawtooth patterns for forced extension and unfolding (Figure 1(c)). First described for the muscle protein titin,[24,25] and soon thereafter for spectrin[23,26–28] and other membrane-associated proteins,[29] each peak in the sawtooth (blue or red dot in Figure 1(c)) corresponds to the force to unfold one or more repeats. Here, the spacing is >20 nm, which is more than three times the folded length of a spectrin repeat. The peak-to-peak length represents the extension of an unfolded domain before the next domain is stressed to unfold. Thousands of such pulling events provide statistical fingerprints of mechanical unfolding for each protein construct.

Extra long intervals between force peaks indicate tandem repeat unfolding processes (red *t* in Figure 1(c)) and are among the recent surprises in the unfolding of various spectrin family proteins.[26,28] For the given sawtooth pattern, force peaks marked with blue dots indicate the point at which a single repeat unfolds and the blue arrows indicate the forced extension of that repeat. The red dot indicates the point at which tandem repeats unfold and the red arrow is the forced extension of the repeats. The influence of dystrophin hinges on the single or tandem repeat unfolding processes is a primary focus below.

Extensible unfolding of dystrophin: peak counts, lengths, and forces

The sawtooth patterns for the three dystrophin constructs here have a countable number of peaks, N_{pk}, that are seen to range up to the number of spectrin repeats plus two (Figure 2(a), bottom histograms). In buffer alone, contacts of the AFM tip with the surface yield curves with $N_{pk}=0$ or 1. With protein on the surface, unfolding has occurred whenever $N_{pk} \geq 3$; if a molecule attaches and detaches without unfolding ($N_{pk}=2$), and if no molecule attaches to the tip ($N_{pk}<2$). Force spectrograms with zero to two peaks account for >80% of

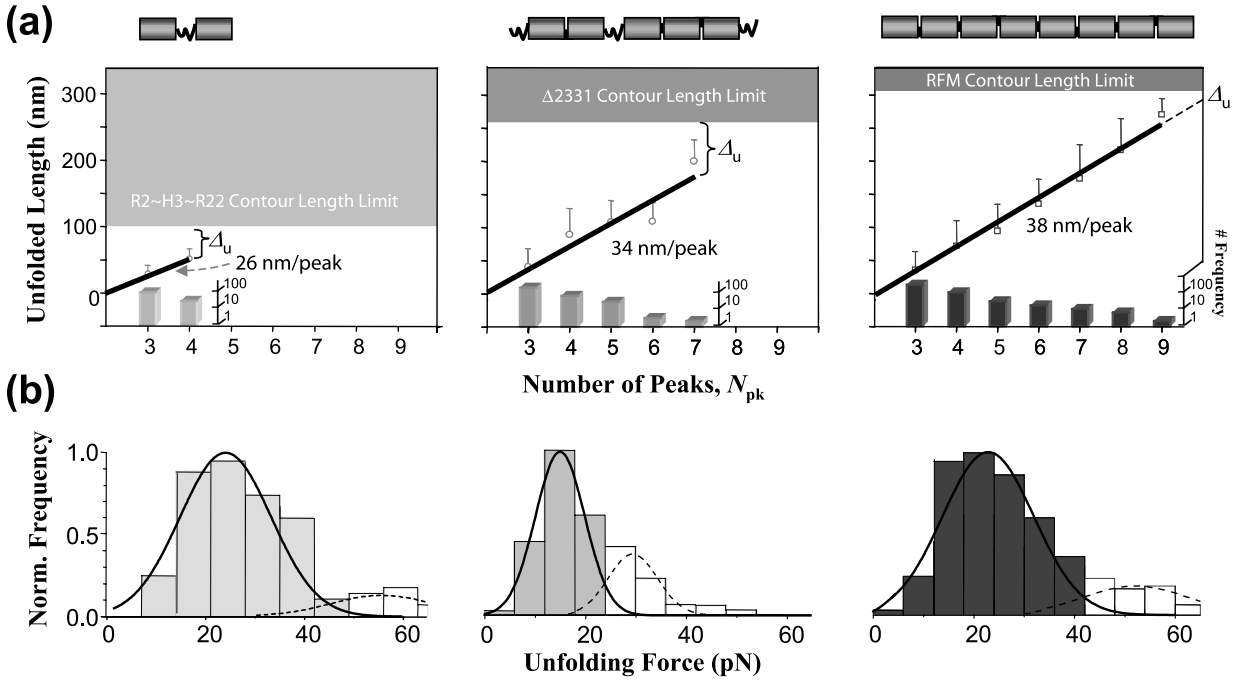

Figure 2. Dystrophin constructs' unfolding length statistics and force distributions. (a) After 8000 contacts of tip to surface, the total unfolding length (avg. ±S.D.) beyond the second peak is obtained for each sawtooth, which is categorized by N_{pk}. The lower bar graph shows frequency distributions of N_{pk}. The upper gray regions indicate the extension limits calculated from the number of amino acid residues (Table 1). The slope of the best-fit line through all of the data (including zero total unfolding length at $N_{pk}=2$) is the average distance between peaks and accounts for both single and tandem repeat unfolding. (b) Histograms of peak unfolding forces. Each bimodal histogram is fitted with two Gaussians of the same width but means that differ by a factor of 2. The major peak corresponds to single chains and the minor peak to pulling on two chains or a loop.

the data, which is consistent with working in the single molecule limit.[26]

It is no surprise that the greater the number of peaks observed, the longer the total extension (Figure 2(a)). However, sawtooths with very large N_{pk} are rare. The length limits calculated as contour lengths from the number of residues (gray regions in Figure 2(a); see Table 1) are approached only for extension curves at the max(N_{pk}). The difference, Δ_u, between the maximum unfolded length and the contour length tends to increase with the total length of hinge (Table 2), suggesting that the hinges are generally (but not always) being extended before unfolding of the repeats. In a sawtooth such as in Figure 1(c), for example, any hinge regions would typically have been extended before the first blue dot where the first domain unfolds and therefore does not account for maximum unfolding length. It thus seems clear that the hinges readily extend without a distinct force peak or domain signature.

Slopes in Figure 2(a) give the unfolding length per peak for each construct and increase by almost 50% from the shortest to longest construct. These best-fit slopes average both single and tandem unfolding events and begin to indicate that tandem events are far more frequent in the two tandem repeat constructs than in the short construct where a single proline-rich hinge separates two helical repeats.

Mean forces to unfold the dystrophin domains are statistically similar among the various constructs and range from 15 pN for Δ2331 to 21 pN for R2∼H3∼R22 and 23 pN for RFM (Figure 2(b)). These are all within the range of forces previously measured for spectrins;[23,26,27] and, because of thermal noise levels with these cantilevers of $(k_C\ k_BT)^{1/2} =$ 6–7 pN, current methods preclude any claim that these forces are statistically different. While the force distributions appear bimodal when fit by Gaussians, minor peaks at high force differ from the main peaks by ∼twofold in force, which indicates that the minor peaks are just a small number of two-chain events where parallel chains are forced to unfold simultaneously. Importantly, forces to unfold tandem repeats and single repeats appear the same: forced unfolding of one repeat propagates rapidly to the next without any need for a higher force. Aside from the small number of two-chain events, the results show single molecule extension and unfolding.

Table 2. Hinge lengths for dystrophin constructs and experimentally determined difference between extended length and total contour length

Dystrophin construct	Total hinge length (nm)	Δ_u (nm)
RFM	0	4
R2-H3-R22	17.7	48
Δ2331	75.3	90

Construct lengths and tandem repeat unfolding

A maximum of seven peaks per force-extension curve for the five repeat spectrin construct (Δ2331; number of repeats, $r=5$) is consistent with the first and the last peaks representing surface and tip desorption processes. Thus $\max(N_{pk})=r+2$. Each repeat in this construct thus contributes no more than one peak, consistent with past trends for monomeric constructs of serial spectrin repeats (Figure 3). The same arithmetic applies to the $r=2$ dystrophin construct R2∼H3∼R22 with $\max(N_{pk})=4$. The eight-repeat RFM dystrophin shows a maximum of nine peaks in sawtooth unfolding patterns after the standard number of contacts, and although this is one peak less than the expected $(r+2)=10$, it is probably another indication of a very high probability of tandem repeat unfolding. Importantly, all of the max(N_{pk}) results above imply that each repeat unfolds in a single all-or-none fashion. If there were a significant frequency of two-stage unfolding (i.e. two peaks per repeat), then max(N_{pk}) would tend to be twice or more the number of repeats in each construct.

While pre-extension of the hinge in the short R2∼H3∼R22 construct predominates (based on the large Δ_u), and should limit coupled unfolding of the two repeats (see Discussion), tandem repeat unfolding clearly occurs at finite frequency based on the bimodal peak-to-peak length histogram of Figure 4. The minor Gaussian occurs at nearly twice the length (41 nm) of the major Gaussian (21 nm) and has a similar width. Consistent with results

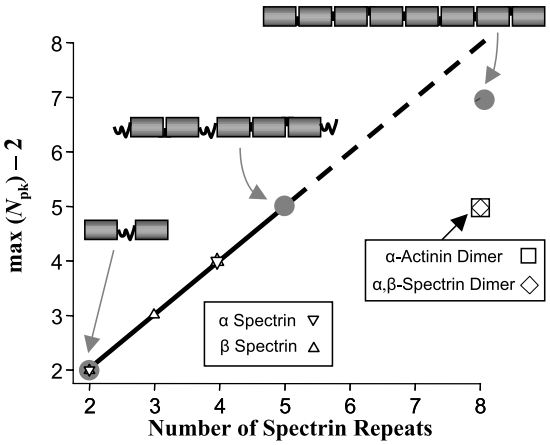

Figure 3. The maximum number of peaks in the sawtooth patterns scales linearly with the number of spectrin repeats in tandem. The quantity max(N_{pk}) −2 is subtracted from the first and last peaks for surface desorption and final desorption, respectively. For the two small dystrophin constructs as well as monomeric α-spectrin and β-spectrin, the unit slope indicates that the maximum contribution to the sawtooths is one peak per repeat, ruling out partial unfolding of repeats. Dystrophin RFM, with its eight repeats in series and extensive tandem repeat unfolding (see Figure 4), falls just below the trendline. The eight-repeat lateral dimers of α,β-spectrin as well as α-actinin fall well below the trendline.

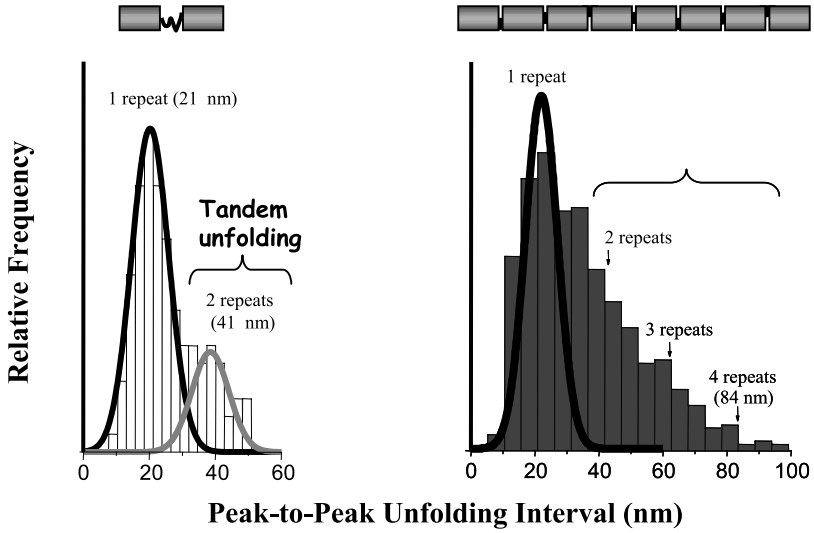

Figure 4. Histograms of peak-to-peak unfolding lengths for dystrophin constructs. Histograms for both R2∼H3∼R22 and RFM constructs were fit with Gaussians for single repeat unfolding and, for R2-H3-R22, tandem unfolding. The major peaks (black) for R2∼H3∼R22 and RFM reflect single domain unfolding and are respective fits of sums of two and eight Gaussians that include effects of different contour lengths for dissimilar domains (see Table 1). For R2-H3-R22, the minor peak (gray line) at ∼twice the major peak value reflects simultaneous unfolding of tandem repeats and was fit with a single Gaussian. For RFM, the bracket spans tandem unfolding events involving two, three, and four repeats.

above, the major peak at 21 nm in Figure 4 is less than the 26 nm slope of Figure 2(a) and is therefore also indicative of tandem repeat unfolding. The difference suggests that tandem unfolding is a minor pathway with a frequency of ∼25%, which is the same small, but finite frequency obtained from the amplitudes of the Gaussians.

The same peak-to-peak length histogram for the eight repeat RFM, which lacks any hinges, has a very prominent, long-length tail that suggests that up to four repeats or more can unfold all at once when stressed. This would cause a gain in dystrophin length from a ∼25 nm folded structure to an 84 nm unfolded chain in a single transition (!). Importantly, all such length changes (from 21 nm up to 84 nm) are highly cooperative in unfolding force, in that they occur at the same force of ∼20 pN as per Figure 2(b). Recent studies of spectrin unfolding by AFM have attributed these tandem repeat unfolding events to helix-to-coil transitions that propagate through helical linkers between repeats.[26]

Pathway shifts and softening with temperature

Tandem repeat unfolding events, as a combined percentage (Figure 5), increase with the number of repeats from ∼25% for a two-repeat construct to over 50% for the eight-repeat construct (at 23 °C). Initial studies of the rate-dependence of the ∼25% tandem events for R2∼H3∼R22 from 0.3 nm/ms to 5 nm/ms extension, suggest no more than a weak increase of about 5% over this log range in rate. However, these percentages tend to decrease significantly with increasing temperature up to 37 °C or 42 °C, and the decrease is strong for the shortest construct, down to an almost imperceptible 10–15% at 37 °C (Figure 5). $\beta_{1\text{-}4}$ Spectrin shows a similar but less dramatic decrease in tandem events upon heating.[27]

Although tandem unfolding events are less frequent upon heating, the dystrophin constructs' peak-to-peak length distributions from 23 °C to 42 °C showed no significant changes in the major peak values. This is most likely because the helical linkers between repeats[15–18] tend to be destabilized by heat before the repeats themselves are.[27] Thermal unfolding of the linker limits helix-to-coil tandem unfolding events that otherwise would propagate to adjacent repeats. Thermal unfolding of a highly heat-susceptible linker-hinge in R2∼H3∼R22 thus results in a steep decrease in tandem percentages as compared to RFM and $\beta_{1\text{-}4}$.

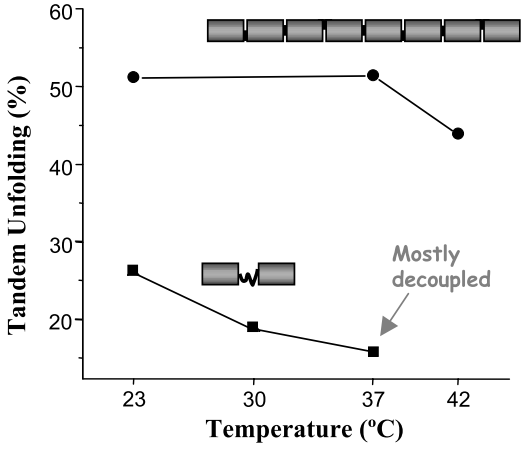

Figure 5. Decrease in fraction of tandem unfolding events with temperature.

Past work on intramolecular hydrogen-bonding[30] and ionic strength effects on the thermal unfolding[31] of spectrin peptides indeed support the conclusion that interdomain linkers are less stable and more susceptible to conformational changes than the coiled-coil triple helical repeats.

Comparisons and insights with circular dichroism (CD)

Thermal unfolding in solution, based on loss of helicity, proved distinct among the different constructs (Figure 6). At 42 °C, the percentages of initial helix are >80% for both dystrophin RFM and R2~H3~R22, and 56% for spectrin $\beta_{1\text{-}4}$. Spectrin $\beta_{1\text{-}4}$ thus loses nearly half of its initial helicity over the temperature range studied here in AFM. Over this same temperature range for the dystrophin constructs, there was also no significant change in unfolding force in AFM (data not shown); at the physiological temperature of 37 °C, mean unfolding forces of 24 pN and 26 pN were measured for RFM and R2~H3~R22, respectively. For $\beta_{1\text{-}4}$ spectrin, in contrast, the unfolding forces decrease by 50% (from 23 °C to 42 °C) with the same non-linear trend for loss of helicity with temperature.

While CD measures of helicity might provide a first indication of stability under force, there are other possible effects. R2~H3~R22 exhibits a considerable drop in helicity from 112% (relative to 4 °C) to 78% helicity (change of 34%) from 23–45 °C (Figure 6(b)). This is similar, if less complete, than the loss of helix and drop in force for $\beta_{1\text{-}4}$ spectrin over this same temperature range. While unfolding forces do not change significantly for R2~H3~R22, tandem repeat percentages drop more than 50% (extrapolating to 45 °C). This suggests that any low-temperature helicity predicted for H3 (see Introduction), which would allow helix-to-coil transitions to propagate from one repeat to the next, is lost upon heating. Fully helical H3 is about 20% of total predicted helicity and

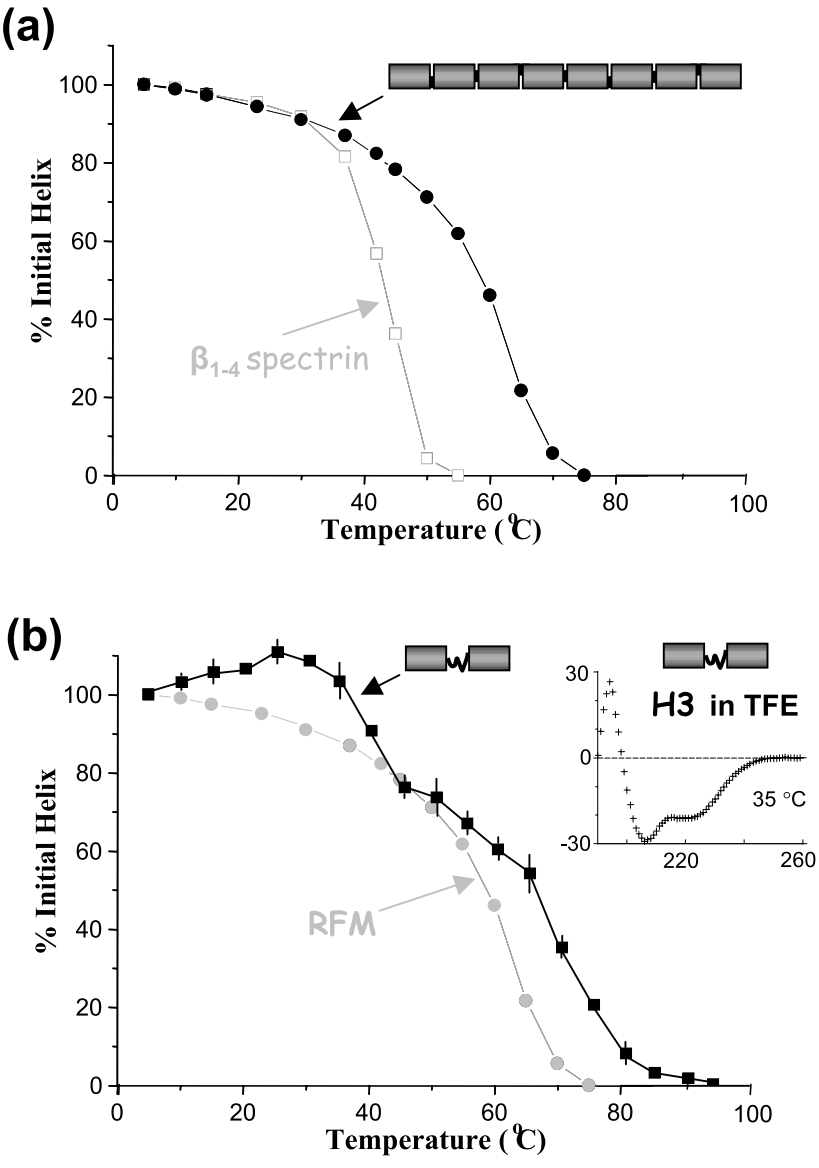

Figure 6. Change in helical content of dystrophin constructs *versus* temperature by CD. (a) Melting curve for RFM *versus* that for $\beta_{1\text{-}4}$ spectrin. (b) Melting curve for R2~H3~R22 (*versus* RFM) is for two separate samples, with error bars indicating the range in helicity. The inset CD spectrum demonstrates helicity induced in H3 construct at 0.16 mg/ml in 40% TFE:60% buffer with *y*-axis corresponding to CD signal as $\theta(10^3 \text{ deg cm}^2 \text{ dmol}^{-1})$ and *x*-axis to wavelength (nm). Increasing the peptide concentration threefold had no significant effect on H3 helicity from 5–90 °C, although helicity decreased almost linearly by 20% over this range.

would be most of the 34% loss of helix. The fact that the remaining majority 78% of R2∼H3∼R22′s helicity is not lost until 80 °C (*versus* 70 °C for RFM) is consistent with these two repeats being very stable.

One last curious feature in Figure 6(b) is that the hinge-linked construct shows an initial, unique, cold-denaturation regime of about 12% helix induction from 4–27 °C. Note that the error bars in the plot indicate the range of helicity results obtained with two separate batches of protein, and so this helix induction phase is reproducible. Since it is not seen in either the dystrophin RFM or the β_{1-4} spectrin construct, a feature unique to R2∼H3∼R22 seemed implicated. Initial efforts to express the synthetic tandem R2∼R22 were unsuccessful, and so we expressed the H3 alone (extended by a few residues as Met∼R^{2417}–Q^{2476}) and studied its secondary structure. The predicted structure of H3 in R2∼H3∼R22 was described as helical (100%) in the Introduction based on algorithms previously applied to spectrin,[19] and dependent on the adjacent R2 and R22. As an isolated peptide, H3 is predicted to be mostly coil. CD measurements indeed show H3 is a coil in PBS. However, addition of 40% trifluoroethanol (TFE), which is known to induce helicity preferentially in polypeptides with an underlying disposition for helicity,[32] induces at least 50% helix in the H3 construct (Figure 6(b), inset).

Discussion

More than a dozen mini-dystrophin constructs have been tested in mice by various morphological and physical assessments of muscle.[1–10] Peripheral nuclei provide one of the simplest measures of effectiveness since, in normal muscle, nuclei appear pushed to the periphery by the contractile apparatus whereas dystrophin-deficient muscles show more than 50% of nuclei centrally located in the myotubes. Chamberlain and co-workers reported that only two of their ten constructs suppressed nuclear mislocalization (to <1%), with the smallest effective construct being R1-3∼H2∼R24.[5] A longer construct of R1-3∼H3-R20-24, based on H3 instead of H2, gave similar results, but other combinations with or without H2 and/or H3 still led to a significantly higher number of central nuclei (and often lower contractile forces). Similar to the functional R1-R3-H2∼R24 construct reported by Chamberlain and co-workers, the R1-2∼H3∼R22-24 construct reported by Xiao and co-workers that was tested here was the only one to have an intervening hinge (H3) among their three mini-dystrophin constructs made.[7,8] It was the only one tested to show a positive effect both on muscle force[8] and nuclear localization.[7] We sought here to better understand the force-dependent flexibility of dystrophin constructs, especially the role of the hinge H3, in forced extension. Of course, throughout, it is very important to note that differences between constructs may also reflect differences in individual repeats; for example, some individual repeats have been found to unfold thermally at or below 37 °C, whereas most repeats unfold at or above 40–45 °C.[19]

Single molecule extension of a number of multi-domain proteins have shown that unfolding forces can easily range from almost 300 pN for some titin Ig domains[24,25] down to 15–30 pN for spectrin repeats[23,26,28] as also found here. Such forces are typical for the mid-range of AFM-extension velocities (∼1 nm/ms) and are even lower at slower, physiologically relevant rates of extension. Myosin molecules individually generate peak forces on F-actin of about 6 pN with similar milliseconds kinetics.[33] This suggests that the collective action of several myosin molecules would be sufficient to extensibly unfold dystrophin in muscle. Titin is also believed to unfold in muscle,[34] despite its much higher unfolding forces.

Tandem unfolding through helical linkers and hinges

Because differences in unfolding force among constructs here are small and cannot be considered significant (Figure 2(b)), sustainable forces on repeats are perhaps less revealing of distinctive, linker-related mechanisms here than tandem repeat unfolding processes. While the skipped sawtooths exemplified by Figure 1(c) for the dystrophin RFM might be interpreted as simultaneous unfolding of non-adjacent repeats instead of tandem repeats, tandem repeat unfolding is most clearly implicated by the bimodal length distribution for the two-repeat R2∼H3∼R22 (Figure 4). We had reported a similar bimodal distribution for the two-repeat β_{1-2} spectrin[26] and a multi-modal distribution for β_{1-4} spectrin,[35] which suggest a qualitatively similar coupling between some if not all spectrin repeats. In addition, kinetic Monte Carlo simulations of such AFM data based on equivalence of all repeats for tandem unfolding also capture the global statistics of such experiments.[36] However, the tandem repeat unfolding pathway for β_{1-2} spectrin was shown to occur about as frequently as the single-repeat pathway, which is significantly more frequent than with R2∼H3∼R22. For β-spectrins, tandem repeat unfolding is understood to reflect a helix-to-coil transition that propagates through a helical linker from one repeat to the next. A reduced helical propensity for the proline-rich linker of R2∼H3∼R22 would tend to reduce, but not eliminate, such propagated unfolding. While hinges (at least H3) thus seem to have a primary role as mechanical separators of repeats in forced extension of dystrophin, the occasional coupling between repeats could be of some importance, especially in a distinctive functional role for H3 *versus* other hinges in dystrophin.

Tandem repeat unfolding is much more prevalent in the longer constructs here, based on both the unfolding lengths per peak (slopes in Figure 2(a))

and the peak-to-peak length histograms (Figure 5). Moreover, at the physiological temperature of 37 °C the percentage of tandem repeat unfolding for R2∼H3∼R22 is just 10–15% *versus* >50% for the dystrophin RFM. The proline-rich hinge H3 is thus most likely less helical, on average, than most of the linkers in the RFM construct. Indeed, the R2∼H3∼R22 construct loses ∼10% helix upon heating from 23 °C to 37 °C in the first phase of an apparent two-phase melting curve (Figure 6(b)), whereas the RFM construct loses much less helix over this temperature range. The fact that the unfolding forces change very little while the tandem repeat frequency decreases by nearly half suggests that the repeats are intact and stable but increasingly decoupled. β_{1-4} Spectrin is the example that establishes the rule, in that its melting curve clearly tracks the decrease in unfolding force (Figure 6(b)). Another parallel is that, at low temperatures where β_{1-4} spectrin repeats are thermally stable (≤25 °C), the frequency of tandem repeat unfolding of β_{1-4} spectrin decreases.[27] In other words, helical linkers between repeats tend to stochastically couple the unfolding of tandem repeats and these linkers tend to be more thermally susceptible structures than the repeats they link.

Secondary structure predictions used here, as on spectrins,[19] predict helical linkers between spectrin repeat motifs for all of the tandem repeats studied here (Figure 7). This includes H3 as a helical linker in the engineered construct R2∼H3∼R22, which is a prediction consistent with the finite probability of tandem repeat unfolding in R2∼H3∼R22. H3 helicity is induced, in structure predictions, by the adjacent repeats; likewise, we show that helicity of an H3 construct can be induced in TFE (Figure 6(b) inset). The H3 here contains five proline residues spaced by three or more amino acid residues and is distinct from H2 (not studied here), which has five proline residues in a row and is predicted to be a coil where proline density is high. More generally, recent database analyses confirm that proline

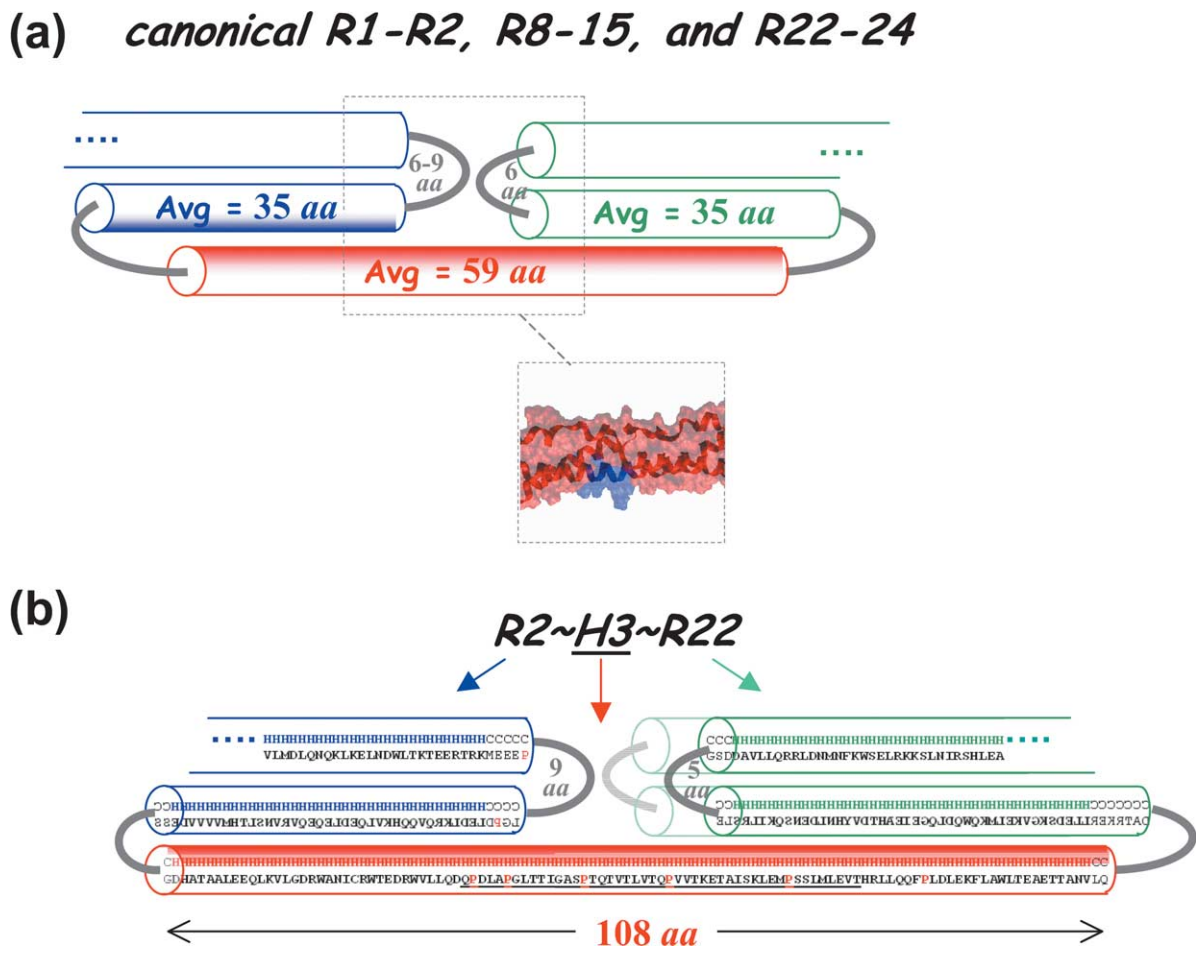

Figure 7. Secondary structure predictions of constructs studied here. The algorithm used was PSIPRED (www.expasy.org) and has recently been applied to spectrin for refined prediction of linkers.[19] (a) Canonical, helix-linked spectrin structures are predicted for all tandem repeats of dystrophin, but the linking helices are much shorter, averaging 59 aa (range of 53–65 aa). The predictions are consistent with those of Koenig and Kunkel.[15] For the 35 aa helices in the repeats, the range is 29–43 aa. The inset, from α-actinin's PDB structure, shows the close-packed structure of adjacent loops that sequester and probably stabilize the linker, in blue. (b) Prediction for the R2∼H3∼R22 construct with hinge-3 (H3) underlined as part of a long, 108 aa helix that links adjacent helical repeats.

residues have an intermediate propensity, among all amino acids, to be in short α-helices.[37] Proline residues are also in turns in the spectrin structures here: in the two turns within each repeat, single proline residues are often found. We thus conclude from predictions as well as AFM data and preliminary CD data that, at least a small fraction of the time, the hinge H3 is helical: perhaps a kinked helix seen in the B-helix of some spectrin repeats.[17]

As cited in the Introduction, a helical linker between tandem repeats is found in all of the crystal structures for spectrin repeat superfamily proteins.[15–18] This seems to apply to dystrophin's H3 linker also, though it appears significantly softer than the "typical" linker between spectrin repeats; the lower stability is probably due to H3's proline residues, which might introduce kinks but are sufficiently distributed along the sequence. The prediction here of a 108 aa helical linker is about 80% longer than the 59 aa (average length) linkers between all other helical repeats in dystrophin, but structural evidence of even longer helices in other muscle proteins is clear and includes the neck region of myosin-V, which is about 150 aa in length. Thus, while these hinges may add flexibility to the folded structure as speculated originally,[14] they also limit extensible unfolding of helical repeats and thereby make mini-dystrophins with strategically placed hinges less extensible whenever domain unfolding occurs.

Our current understanding of how the results here translate to full-length dystrophin or other mini-dystrophins is perhaps best summarized in the analyses in Figure 8, which is based on the results for all three constructs studied (see Materials and Methods). Plotted *versus* the number of repeats, *r*, the exponential fits for both unfolding length (which combines single plus all tandem unfolding pathways as per Figures 2 and 4) and the decaying frequency distributions of N_{pk} (Figure 2), yield an integer decay constant of $r_o = 3$. This indicates that, ~67% of the time, three or fewer spectrin repeats in dystrophin behave as a single mechanical unit in forced extension and unfolding. Although tandem repeat unfolding is frequent even with hinge H3, massive unfolding of more than three spectrin repeats in dystrophin is unlikely. While averaging through hinged and non-hinged constructs, the result provides insight into the extensibility of a full-length, 24-repeat dystrophin with two intervening hinges (Figure 1(a)): very similar sawtooth statistics are expected for dystrophin constructs larger than about 12 repeats, regardless of hinges being present or not. Such studies also provide physical insight into refined designs of mini-dystrophins that might have statistics similar to or distinct from longer constructs. For example, knocking out the proline residues in the hinges or adding more potent helix-breakers could have useful consequences.

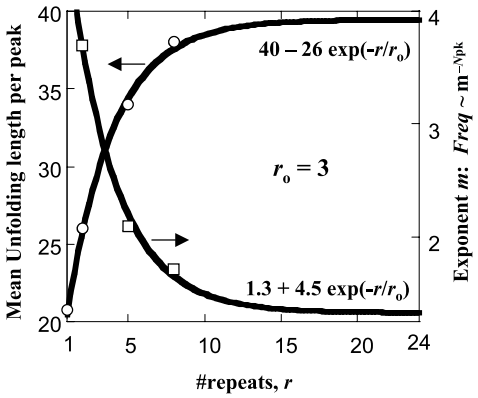

Figure 8. Dependence on the number of repeats, *r*, of two key measures of sawtooth statistics. First, the unfolding lengths per peak of Figure 2, including the single repeat interval in Figure 4, increase with the total number of domains *r* and fit well to the indicated exponential plateau function with the decay constant, $r_o = 3$ (see Materials and Methods). Additionally, the N_{pk} distributions of Figure 2 consistently show a scaled decay with N_{pk} that fit to $\sim m^{-N_{pk}}$,[26] and this factor *m* also decreases as indicated with the total number of domains *r*. Such fits from the current constructs provide simple extrapolations for other possible mini-dystrophins and dystrophin constructs, including full-length ($r = 24$).

Conclusion

The single molecule, ensemble-scale studies here elaborate the thermo-mechanical properties of various dystrophin mini-constructs, including one mini-dystrophin hinged-rod domain used already in gene therapy. The constructs probed reveal aspects of mechanical extension and unfolding forces as well as hinge and temperature effects. Our findings that the hinges will not unfold as three-helix repeats do show that these proline-rich regions lack the character of stable domains. While these hinges have been speculated to add flexibility during the contraction-relaxation of muscle fibers, our results indeed show that they generally facilitate extension when unfolding is minimal. The hinges, or at least H3 as a marginally stable proline-kinked helix, allow only very limited unfolding of tandem repeats. Similarities between dystrophin and spectrin extension include forces and linker mechanisms as well as very limited aspects of temperature dependence.

Materials and Methods

Protein preparation

The central rod fragment of dystrophin, comprising repeats 8–15, was subcloned into expression vector pMW172 as described.[21] The mini-dystrophin construct Δ2331 was made by PCR, using the construct Δ3990[7] as a template generously donated by Dr X. Xaio. A start codon-embedded NdeI site, an added stop codon and an EcoRI site were engineered into PCR primers such that the PCR product could be subsequently subcloned into expression vector pMW172 to yield the minidystrophin

Δ2331 construct containing hinge 1, repeats 1 and 2, hinge 3, repeats 22–24 and hinge 4, i.e. H1-R1-2∼H3∼R22-24∼H4. In such constructs, we denote any change in native sequence with a ∼. The Δ2331 is essentially the Δ3990 without the N-terminal and CR domains. The R2∼H3∼R22 construct is the core of Δ2331. It corresponds to residues 448–556 (rod domain 2: R2), 2424–2470 (hinge 3: H3), and 2684–2802 (rod domain 22: R22) of human dystrophin (GeneBank NM004006). It was cloned out of Δ3990 and ligated into the expression vector pMW172[31] between BamHI and HindIII sites. Resulting plasmids were used to transform *Escherichia coli* strain BL21 (Star)DE3 (Invitrogen). Bacteria were grown at 37 °C in Luria-Bertani medium with 200 μg/ml of ampicillin, to an $A_{600\,nm}$ of 0.8–1.0, and expression was induced by addition of 1 mM isopropyl-β-D-thiogalactoside. After 3–4 h, cells were harvested by centrifugation at 5000*g* for 10 min and frozen in liquid N_2. The frozen cell pellet (∼10 ml, from 2 l of culture) was thawed for 10 min at 25 °C and lysed by resuspension in 40 ml of Cellytic B (Sigma-Aldrich) containing 1 mM phenylmethylsulfonyl fluoride, 10 μg/ml of leupeptin, and 10 μg/ml of aprotinin. After 15 min at 25 °C, the cell lysate was clarified by centrifugation at 45,000*g* for 15 min; centrifugation and all subsequent steps were performed at 4 °C. The supernatant was loaded onto a DEAE-cellulose column (Whatman DE-52, 2.5 cm × 30 cm) pre-equilibrated with 20 mM Hepes (pH 7.0), 1 mM dithiothreitol, 1 mM sodium azide, and eluted at 35 ml/h with a 325 ml × 325 ml gradient of 0–0.5 M sodium chloride in Hepes buffer; fractions were collected at 12 min intervals. The expressed protein, identified by SDS-PAGE, eluted about halfway through the gradient. Solid ammonium sulfate was added to the pooled peak fractions to a final concentration of 1.6 M, and the material was applied to a 1.5 cm×30 cm column of Butyl ToyoPearl 650 M (Tosoh Bioscience) in 1.6 M ammonium sulfate, 20 mM Hepes (pH 7.0), 1 mM dithiothreitol, 1 mM sodium azide. The column was eluted at 10 ml/h with a descending gradient of 1.6 M–0 M ammonium sulfate in 150 ml. Fractions were collected at 10 min intervals, and fractions containing the purified protein were identified by SDS-PAGE, pooled, and dialyzed against 100 volumes of phosphate-buffered saline, with two changes. The concentration was determined from the $A_{280\,nm}$ using an extinction coefficient of 1.313 at 1 mg/ml, calculated from the amino acid composition.

Stable monomer was ultimately purified by gel-permeation chromatography in phosphate-buffered saline (PBS) and kept on ice for AFM studies. Immediately before use, any protein aggregates were removed by centrifugation at 2 °C for 1 h and monodispersity was verified by dynamic light-scattering.

Dynamic force spectroscopy

An AFM experiment was begun by adsorbing 50 μl of 0.1 mg/ml of protein for 15 min at room temperature onto either freshly cleaved mica or amino-silanized glass coverslips. The surface was then lightly rinsed with PBS and placed, without drying, under the head of the AFM; all measurements were carried out in PBS. Lower protein concentrations generated minimal AFM results; higher protein concentrations showed higher unfolding forces, proving consistent with the results below, indicating that domains in multiple, parallel chains are forced to unfold all at once. Fluorescence imaging of labeled protein demonstrated homogeneous adsorption, and AFM imaging after scratching the surface showed that no more than a monolayer of molecules covered the substrate.[29]

Two AFMs were used with similar results: (i) a Nanoscope-E Multimode AFM (Digital Instruments, Santa Barbara, CA) equipped with a liquid cell; and (ii) an Epi-Force Probe from Asylum Research. Sharpened silicon nitride (Si_3N_4) cantilevers (Park Scientific, Sunnyvale, CA) of nominal spring constant k_C =10 pN/nm were typically used, with equivalent results obtained using 30 pN/nm cantilevers. k_C was measured for each cantilever using the manufacturer's directions at each temperature, and additional calibrations were performed as described.[29] Temperatures in the range used seemed to have little effect on k_C. Experiments were done at imposed displacement rates of 1 nm/ms. For the high-temperature experiments, samples were first studied at 23 °C, and the temperature was then raised. The experiments proved challenging due to bending of the bi-layered cantilever (gold on silicon nitride). The laser had to be realigned after raising the temperature, and only the rare cantilever yielded a high laser sum and low data noise. Temperatures higher than 42 °C caused the cantilever to bend sufficiently that the laser would not deflect off the tip. The desired temperatures were controlled and monitored using the Nanoscope Heater Controller by Digital Instruments. For any one temperature, thousands of surface to tip contacts were collected and analyzed with the aid of a semi-automated, visual analysis program custom written in C++.[29] Since protein unfolding events are stochastic and the experiment intrinsically random in many ways, collecting and analyzing thousands of peaks is necessary to provide an accurate statistical survey of extensible unfolding. Initial results at the beginning of a many hours experiment were similar to those obtained at the end of the experiment. Nearly all of the data was thus analyzed.

As a check of our experimental technique and analysis methods, we have conducted unfolding studies on a Titin-$(I27)_8$ construct (AthenaES, Baltimore, MD). Not only do we obtain mean unfolding forces and mean unfolding lengths within established variations for this well-studied protein, but we also find no significant tandem repeat unfolding (<10–15%).

Circular dichroism measurements

Circular dichroism spectra were measured at a number of temperatures using a 1 mm path-length cell on a Jasco J715 spectropolarimeter in the same buffer (PBS, pH 7.4) as that used in AFM experiments. The instrument was calibrated with d-10-camphorsulfonic acid. Samples were equilibrated at each temperature for 20 min before taking measurements. Each spectrum was the average of three consecutive measurements. The effect of 40% (v/v) TFE on the secondary structure of the peptides was also studied on a construct containing only hinge 3 (H3).

Cumulative analyses of construct lengths

Two analyses of AFM results were conducted in order to identify overall trends *versus* construct length or number of repeats, *r*. With such trends in hand, one might be able to predict single molecule results for other dystrophin constructs or even, perhaps, for full-length dystrophin. First, it is clear that the mean unfolding length per peak in Figure 2 increases with the total number of repeats: 26 nm, 34 nm, 38 nm for r=2, 5, 8, respectively. This is because of the increasing frequency of

tandem repeat unfolding. In addition, a single repeat should also give a forced unfolding of 21 nm for $r=1$ (Figure 4). Plotted *versus* r (Figure 8), the trend with these four points is non-linear and tending to plateau with r, since there is little difference between the two larger dystrophin constructs. We therefore fit the mean unfolding length per peak to the function:

$$f(r) = A + B\exp(-r/r_o)$$

For physical reasons, r_o should be an integer.

There is a second, more subtle but similar trend with r in the AFM results. First, we have noted before[26] that frequency distributions for the number of unfolding peaks in the sawtooths, N_{pk} (Figure 2) generally decay with N_{pk}, and for spectrin proteins these decays have been shown to fit frequency $\sim m^{-N_{pk}}$. As explained by Law *et al.*,[26] the factor m provides a measure of the m-fold fewer ways of achieving one more unfolded domain (single or multi-repeat domains) spanning the gap between tip and surface. Experiments on a series of β-spectrin constructs as well as kinetic Monte Carlo simulations that fit such results[36] show that $m=m(r)$ and this function also decays exponentially as $f(r)$, albeit with different A and B.

Acknowledgements

We thank S. Harper and D. W. Speicher for invaluable assistance with the protein purification and preparation. This work was supported by NIH-BRP, NSF-PECASE, and MDA grants.

References

1. Chamberlain, J. S. (2002). Gene therapy of muscular dystrophy. *Hum. Mol. Genet.* **11**, 2355–2362.
2. Decrouy, A., Renaud, J-M., Davis, H. L., Lunde, J. A., Dickson, G. & Jasmin, B. J. (1997). Mini-dystrophin gene transfer in mdx4cv diaphragm muscle fibers increases sarcolemmal stability. *Gene Ther.* **4**, 401–408.
3. Fabb, S. A., Wells, D. J., Serpente, P. & Dickson, G. (2002). Adeno-associated virus vector gene transfer and sarcolemmal expression of a 144 kDa micro-dystrophin effectively restores the dystrophin-associated protein complex and inhibits myofibre degeneration in nude/mdx mice. *Hum. Mol. Genet.* **11**, 733–741.
4. 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.
5. Harper, S. Q., Hauser, M. A., DelloRusso, C., Duan, D., Crawford, R. W., Phelps, S. F. *et al.* (2002). Modular flexibility of dystrophin: implications for gene therapy of Duchenne muscular dystrophy. *Nature Med.* **8**, 253–261.
6. Sakamoto, M., Yuasa, K., Yoshimura, M., Yokota, T., Ikemoto, T., Suzuki, M. *et al.* (2002). Micro-dystrophin cDNA ameliorates dystrophic phenotypes when introduced into mdx mice as a transgene. *Biochem. Biophys. Res. Commun.* **293**, 1265–1272.
7. Wang, B., Li, J. & Xiao, X. (2000). From the Cover: Adeno-associated virus vector carrying human minidystrophin genes effectively ameliorates muscular dystrophy in mdx mouse model. *Proc. Natl Acad. Sci. USA*, **97**, 13714–13719.
8. Watchko, J., O'Day, T., Wang, B., Zhou, L., Tang, Y., Li, J. & Xiao, X. (2002). Adeno-associated virus vector-mediated minidystrophin gene therapy improves dystrophic muscle contractile function in mdx mice. *Hum. Gene Ther.* **13**, 1451–1460.
9. Yoshimura, M., Sakamoto, M., Ikemoto, M., Mochizuki, Y., Yuasa, K., Miyagoe-Suzuki, Y. & Takeda, S. i. (2004). AAV vector-mediated micro-dystrophin expression in a relatively small percentage of mdx myofibers improved the mdx phenotype. *Mol. Ther.* **10**, 821–828.
10. Yuasa, K., Miyagoe, Y., Yamamoto, K., Nabeshima, Y. i., Dickson, G. & Takeda, S. i. (1998). Effective restoration of dystrophin-associated proteins *in vivo* by adenovirus-mediated transfer of truncated dystrophin cDNAs. *FEBS Letters*, **425**, 329–336.
11. Amann, K. J., Renley, B. A. & Ervasti, J. M. (1998). A cluster of basic repeats in the dystrophin rod domain binds F-actin through an electrostatic interaction. *J. Biol. Chem.* **273**, 28419–28423.
12. Hoffman, E. P., Brown, R. H. & Kunkel, L. M. (1987). Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell*, **51**, 919–928.
13. Calvert, R., Kahana, E. & Gratzer, W. B. (1996). Stability of the dystrophin rod domain fold: evidence for nested repeating units. *Biophys. J.* **71**, 1605–1610.
14. Koenig, M. & Kunkel, L. M. (1990). Detailed analysis of the repeat domain of dystrophin reveals four potential hinge segments that may confer flexibility. *J. Biol. Chem.* **265**, 4560–4566.
15. 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.
16. 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.
17. Kusunoki, H., MacDonald, R. I. & Mondragon, A. (2004). Structural insights into the stability and flexibility of unusual erythroid spectrin repeats. *Structure*, **12**, 645–656.
18. Kusunoki, H., Minasov, G., MacDonald, R. I. & Mondragon, A. (2004). Independent movement, dimerization and stability of tandem repeats of chicken brain [alpha]-spectrin. *J. Mol. Biol.* **344**, 495–511.
19. 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.
20. England, S. B., Nicholson, L. V. B., Johnson, M. A., Forrest, S. M., Love, D. R., Zubrzycka-Gaarn, E. E. *et al.* (1990). Very mild muscular dystrophy associated with the deletion of 46% of dystrophin. *Nature*, **343**, 180–182.
21. Kahana, E., Flood, G. & Gratzer, W. B. (1997). Physical properties of dystrophin rod domain. *Cell Motil. Cytoskel.* **36**, 246–252.
22. Goyenvalle, A., Vulin, A., Fougerousse, F., Leturcq, F., Kaplan, J.-C., Garcia, L. & Danos, O. (2004). Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. *Science*, **306**, 1796–1799.
23. Law, R., Harper, S., Speicher, D. W. & Discher, D. E.

(2004). Influence of lateral association on forced unfolding of antiparallel spectrin heterodimers. *J. Biol. Chem.* **279**, 16410–16416.
24. Carrion-Vazquez, M., Oberhauser, A. F., Fowler, S. B., Marszalek, P. E., Broedel, S. E., Clarke, J. & Fernandez, J. M. (1999). Mechanical and chemical unfolding of a single protein: a comparison. *Proc. Natl Acad. Sci. USA*, **96**, 3694–3699.
25. 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.
26. 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.
27. Law, R., Liao, G., Harper, S., Yang, G., 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.
28. 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.
29. Carl, P., Kwok, C. H., Manderson, G., Speicher, D. W. & Discher, D. E. (2001). Forced unfolding modulated by disulfide bonds in the Ig domains of a cell adhesion molecule. *Proc. Natl Acad. Sci. USA*, **98**, 1565–1570.
30. Ragone, R. (2001). Hydrogen-bonding classes in proteins and their contribution to the unfolding reaction. *Protein Sci.* **10**, 2075–2082.
31. Lusitani, D., Menhart, N., Keiderling, T. A. & Fung, L. W.-M. (1998). Ionic strength effect on the thermal unfolding of alpha-spectrin peptides. *Biochemistry*, **37**, 16546–16554.
32. Buck, M. (1998). Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Quart. Rev. Biophys.* **31**, 297–355.
33. Ishijima, A., Kojima, H., Higuchi, H., Harada, Y., Funatsu, T. & Yanagida, T. (1996). Multiple- and single-molecule analysis of the actomyosin motor by nanometer-piconewton manipulation with a micro-needle: unitary steps and forces. *Biophys. J.* **70**, 383–400.
34. Li, H., Linke, W. A., Oberhauser, A. F., Carrion-Vazquez, M., Kerkvliet, J. G., Lu, H. *et al*. (2002). Reverse engineering of the giant muscle protein titin. *Nature*, **418**, 998–1002.
35. Ortiz, V., Nielsen, S. O., Klein, M. L. & Discher, D. E. (2005). Unfolding a linker between helical repeats. *J. Mol. Biol.* **349**, 638–647.
36. Law, R., Bhasin, N. & Discher, D. E. (2005). Application of probe microscopy to protein unfolding: adsoprtion and ensemble analyses. *ACS Symp. Series*, **897**, 162–181.
37. Wang, J. & Feng, J.-A. (2003). Exploring the sequence patterns in the {alpha}-helices of proteins. *Protein Eng.* **16**, 799–807.

Edited by C. R. Matthews

(Received 9 February 2005; received in revised form 8 July 2005; accepted 26 July 2005)
Available online 10 August 2005