Single Molecule Force Spectroscopy Reveals a Weakly Populated Microstate of the FnIII Domains of Tenascin

Yi Cao and Hongbin Li*

Department of Chemistry, The University of British Columbia, Vancouver, BC, Canada V6T 1Z1

The native states of proteins exist as an ensemble of conformationally similar microstates. The fluctuations among different microstates are of great importance for the functions and structural stability of proteins. Here, we demonstrate that single molecule atomic force microscopy (AFM) can be used to directly probe the existence of multiple folded microstates. We used the AFM to repeatedly stretch and relax a recombinant tenascin fragment TNnALL to allow the fibronectin type III (FnIII) domains to undergo repeated unfolding/refolding cycles. In addition to the native state, we discovered that some FnIII domains can refold from the unfolded state into a previously unrecognized microstate, N* state. This novel state is conformationally similar to the native state, but mechanically less stable. The native state unfolds at \( \sim 120 \) pN, while the N* state unfolds at \( \sim 50 \) pN. These two distinct populations of microstates constitute the ensemble of the folded states for some FnIII domains. An unfolded FnIII domain can fold into either one of the two microstates via two distinct folding routes. These results reveal the dynamic and heterogeneous picture of the folded ensemble for some FnIII domains of tenascin, which may carry important implications for the mechanical functions of tenascins in vivo.

© 2006 Elsevier Ltd. All rights reserved.

*Corresponding author

Keywords: tenascin; FnIII domains; mechanical unfolding; scanning probe microscopy; single-molecule force spectroscopy

Introduction

Proteins exist as a dynamic ensemble of conformationally similar folded microstates in solution. The fluctuation and inter-conversion among different microstates can occur at different time scales spanning from hundreds of microseconds to seconds,\(^1\)–\(^5\) suggesting the existence of high energetic and less populated microstates. This dynamic nature of proteins is essential for protein functions, as recent experimental data shows that the fluctuations among different microstates play important roles in determining catalytic function of enzymes, ligand binding affinities, and structural stability of proteins.\(^1\)–\(^2\),\(^6\)–\(^9\) This dynamic picture also contrasts the average structures of proteins determined by X-ray crystallography or NMR and the average physiochemical properties of proteins determined using ensemble measurements. Despite the importance, it is challenging to experimentally probe the existence and distribution of these heterogeneous microstates, as these microstates resemble each other in conformation and the high energetic microstates are not populated and amount to just a small fraction of the whole population.\(^4\),\(^7\)–\(^11\) It is of great importance to expand the tool box and exploit new methods to monitor the distribution of folded microstates and the fluctuation among them. Here we demonstrate that single molecule atomic force microscopy (AFM) can be applied to directly probe the existence of different mechanical conformational microstates for folded proteins and the fluctuation between these microstates. Unlike the methods mentioned above, which directly monitor the structural difference between different microstates through nuclear magnetic resonance relaxation or single molecule electron transfer, single molecule AFM probes the existence of different microstates by their distinct mechanical stability. Mechanical stability is an intrinsic property of proteins,\(^12\)–\(^17\) and is sensitive to structural perturbation to the folded structures.\(^14\),\(^16\),\(^19\)
example, it has been shown that single point mutation can dramatically alter the mechanical stability of a protein, although the mutant protein remained folded and conformationally resembled the wild-type.14,18,20 Here we use the recombinant human tenascin-C fragment TNfnALL,21 which consists of all the 15 fibronectin type III domains (FnIII) of tenascin-C, as a model system to directly demonstrate the heterogeneity of the native ensemble and the existence of the distinct high energetic folded microstates.

Tenascin-C is an extracellular matrix protein and plays important roles in regulating the cell–matrix interactions.22 Tenascin-C is typically expressed in tissues that are subject to heavy tensile load and is believed to provide elasticity and mechanical strength for mesenchymal tissues.23 Under normal physiological conditions, tenascins are subject to mechanical stretching forces and the force-induced unfolding/refolding reactions may be an important part of tenascins dynamics in vivo.24 Single molecule AFM experiments can mimic the stressful environment of tenascins under physiological conditions and may provide structural and dynamic information that cannot be obtained otherwise. Previous single molecule AFM experiments on tenascins revealed that tenasin is an elastic protein that can extend to several times its resting length via force-induced unfolding of FnIII domains.24,25 It has been suggested that the mechanical unfolding of FnIII domains serves as a shock-absorber to prolong the lifetime of tenasin–ligand bond.24 These studies serve as a foundation for us to use single molecule AFM to probe the existence of microstates in the folded ensemble of FnIII domains.

Here, we use single molecule AFM techniques to demonstrate that some FnIII domains of tenascin exist as two distinct populations of microstates in their folded ensembles: the first population unfolds at 120 pN (the native state); while the second one unfolds at only ∼50 pN. The second population is conformationally similar to the native state, but mechanically less stable. It represents a previously unrecognized high energetic folded microstate in the folded ensemble of FnIII domains. This study represents a direct illustration of the heterogeneity among the folded FnIII domains, and revealed the multiple folding routes that may be taken by the same protein. The multiple folding pathways could effectively limit the time that tenascins spend in the unfolded state and help tenasin regain its mechanical stability rapidly, which may have important biological implications for the function of tenascin-C.

Results

The FnIII domains of tenascin have similar mechanical stability

We used single molecule AFM to stretch TNfnALL, which consists of all the 15 FnIII domains of tenascin-C, and measure its force-extension relationships. The stretching of TNfnALL results in force-extension relationships of characteristic saw-tooth pattern appearance, in which the force peaks correspond to the mechanical unfolding of the individual FnIII domains (Figure 1(a)). The last peak in the force-extension curves corresponds to the extension of the unfolded polyprotein chain and its subsequent detachment of the protein from either the substrate or AFM tip. Since the AFM tip picks up molecules randomly along the contour of intact TNfnALL proteins, fragments of TNfnALL were picked up and stretched between the AFM tip and substrate, resulting in force-extension curves with varying number of unfolding force peaks (from 3 to 13).

These pulling experiments are referred to as single pulling experiments, to distinguish them from the repeated stretching-relaxation experiments shown below. The force-extension relationships of TNfnALL can be well described by the worm-like chain (WLC) model of polymer elasticity26,27 and no unfolding intermediate state was detected. WLC fits (red lines, Figure 1(a)) to the consecutive unfolding force peaks measure an average contour length increment (∆Lc) of 29.3(±0.9) nm (n = 1640) for FnIII domains upon their mechanical unfolding (Figure 1(b)). Contour length increment ∆Lc is an inherent structural parameter of a protein. Together with the length difference, Ldifference between Lcontour and ∆Lc measures the distance between the N and C termini prior to the unfolding event, e.g. at the transition state, and can be used to infer information about the transition state during mechanical unfolding. FNIII domains consist of 88 to 92 amino acid residues. The contour length Lcontour of a fully extended FnIII is about 32.4 nm (∼90 aa × 0.36 nm/aa) long, giving rise to an Ldifference of 3.1 nm. Ldifference is very close to the distance between the N and C termini in the folded state (LNC) for the third FnIII domain of tenascin (TNfn3), which is 3.1 nm measured from the crystal structure of TNfn3.30 This result indicates that the unfolding transition state for FnIII domains is just a slightly stretched form of the native state.

The unfolding force peaks in the force-extension curves are of similar amplitudes around 120 pN. The unfolding force histogram shown in Figure 1(c) revealed a narrow distribution of unfolding forces with an average of 117(±19) pN (n = 1640, pulling speed: 400 nm/s), in good agreement with previous AFM measurements.24,25,31 The width of the unfolding force histogram is much narrower than what one would expect for a heterogeneous tandem modular protein, but comparable with that of homopolyproteins,12,20,31 suggesting that the 15 FnIII domains of tenascin unfold at similar forces despite the diverse sequences among them. Consistent with this picture, recent single molecule AFM studies indeed showed that TNfn3 unfolds at ∼120 pN.20 Hence, as an approximation, it is possible to describe...
the average mechanical unfolding of the 15 FnIII domains of tenasin using the same set of kinetic parameters: the unfolding rate constant at zero force, $\alpha_0$, and the distance between the folded state and the transition state, $\Delta x_u$. We modeled the mechanical unfolding of FnIII domains by a simple two-state process and used Monte Carlo simulations to estimate the average $\alpha_0$ and $\Delta x_u$ directly from unfolding force histogram. We found that the unfolding force histogram of TNfnALL was reproduced adequately by Monte Carlo simulations using a $\alpha_0$ of $5 \times 10^{-3} \text{s}^{-1}$ and a $\Delta x_u$ of 0.31 nm. We also measured the dependence of unfolding forces on the pulling speed and the data is shown in Supplementary Data as Figure S1. We found that the speed dependence of the unfolding force for the native states of FnIII domains can be described adequately using the same set of $\Delta x_u$ and $\alpha_0$ as we determined by fitting the unfolding force histogram at a single pulling speed (Figure 1(c)). It is worth noting that such analysis represents an averaged description of the unfolding behaviors of the 15 FnIII domains of tenasin. For more detailed analysis of the unfolding behavior of individual FnIII domains, it is necessary to investigate the mechanical unfolding of individual FnIII domains using polyprotein composed of identical tandem repeats of the same FnIII domain.

In contrast to the experimental finding that the FnIII domains of tenasin share similar mechanical stability, FnIII domains from fibronectin were observed to have vastly different mechanical stability. Such contrast illustrates the unique mechanical architecture of tenasin and fibronectin, in which homologous FnIII domains are major components, and reveals the complexity of the relationship between protein structure and mechanical stability.

Repetitive stretching-relaxation of TNfnALL reveals unfolding events with anomalously low unfolding forces.

It has been demonstrated that, by limiting the extension such that the protein molecule does not detach from either the AFM tip or substrate, one can stretch and relax the same protein molecule repeatedly to allow the protein to undergo many unfolding/refolding cycles. Previously, such experiments measured a double-exponential folding kinetics of tenasin FnIII domains, and captured rare dimeric misfolding events involving two neighboring FnIII domains in TNfnALL at the single molecule level. Applying similar techniques to TNfnALL, we reproduced the double-exponential folding kinetics of FnIII domains and also observed the above mentioned dimeric misfolding events (see Figure S2 in Supplementary Data). And more importantly, we discovered that some unfolding events of FnIII domains, during repeated stretching-relaxation cycles, occurred at anomalously low unfolding forces that have not been reported in the literature. Figure 2(a) shows a series of extension traces during a typical repeated stretching-relaxation experiment. In trace i, the protein was first extended and five equally spaced force peaks were observed, which correspond to the mechanical unfolding of five FnIII domains. Then the unfolded protein was relaxed to almost zero extension (trace ii), resulting in elastic behavior that is typical of a simple entropic spring-like polymer chain. Between the consecutive stretching-relaxation cycles, the protein was relaxed at zero extension increment, persistence length of 0.39 nm and an average contour length of $29.3(\pm 0.9)$ nm for $\Delta L_c$ of 29.4 nm. The last peak of each curve corresponds to the detachment of TNfnALL molecule from either the tip or substrate.

Figure 1. Force-extension relationships of TNfnALL measured using single molecule AFM. (a) Typical force-extension curves for individual TNfnALL protein, which is composed of 15 different FnIII domains. Since the TNfnALL molecule is picked up by the AFM tip randomly along its contour length, different numbers of FnIII domains are picked up and stretched during individual trials, resulting in the force-extension curves with different number of unfolding force peaks. The force-extension curves show characteristic saw-tooth patterns, in which each individual saw-tooth peak corresponds to the mechanical unfolding of individual FnIII domains. The unfolding force peaks are of similar amplitude of $\sim 120$ pN. The force-extension curves can be well-described by the WLC model of polymer elasticity. Red lines are WLC fits to the experimental data with a persistence length of 0.39 nm and an average contour length increment, $\Delta L_c$, of 29.4 nm. The last peak of each curve corresponds to the detachment of TNfnALL molecule from either the tip or substrate. (b) Histogram of $\Delta L_c$ for the unfolding of the FnIII domains. Gaussian fit to the experimental data shows an average value of $29.3(\pm 0.9)$ nm for $\Delta L_c$. (c) Histogram of unfolding forces for TNfnALL at a pulling speed of 400 nm/s. The distribution of unfolding forces can be well described by the Monte Carlo simulation (red line) using an unfolding rate constant $\alpha_0$ of $5 \times 10^{-3} \text{s}^{-1}$ and an unfolding distance, $\Delta x_u$, of 0.31 nm.
A Weakly Populated Microstate of the FnIII Domains

Example, in trace iii we observed four equally spaced unfolding force peaks, indicating that one of the five domains remained unfolded while the other four managed to refold. We observed that most unfolding force peaks had amplitudes of ∼120 pN, similar to the average unfolding force observed during single pulling experiments, suggesting that these FnIII domains folded back to their native states and acquired similar mechanical stability. However, during repeated stretching-relaxation cycles, we also observed that some unfolding force peaks occurred at much lower forces (∼50 pN), as those indicated by arrows. These events occur predominantly as the first unfolding force peak in a given force-extension curve, followed by unfolding force peaks occurring at normal forces (∼120 pN), suggesting that only one FnIII domain (out of the FnIII domains that are being stretched between the AFM tip and substrate) unfolded at the anomalously low unfolding force. These anomalously low unfolding force peaks were rarely observed in single pulling experiments (63 out of total 1640 events), but were frequent occurrences in repeated stretching-relaxation experiments. Using repeated stretching-relaxation protocols, we examined a total of 59 molecules. Since the protein is picked up randomly along its contour by the AFM tip, all the 59 molecules were fragments of the intact TNfNAll protein and contained between three and nine FnIII domains (on average five domains/molecule). We observed that 56 molecules clearly showed unfolding events with anomalously low unfolding forces (<70 pN), as those shown in Figure 2(a). Out of a total 3625 unfolding events, 379 unfolding events occurred at lower forces (<70 pN).

From the repeated stretching-relaxation of the same molecule, we observed that the unfolding of the same FnIII domain can occur at either regular force ∼120 pN or lower force at ∼50 pN. For example, in Figure 2(a) the first unfolding events in traces i and iv occurred at ∼120 pN, while those in traces v and vi occurred at ∼50 pN. In all these traces, additional unfolding events at ∼120 pN were followed. It seems that, during refolding, only one FnIII domain out of the many FnIII domains being unfolded can fold into two distinct states: one is mechanically stable and unfolds at 120 pN; the other one is mechanically labile and unfolds at ∼50 pN.

Since the first unfolding event in force-extension curves of TNfNAll during repeated stretching-relaxation cycles can occur at either regular or lower forces, we compiled an unfolding force histogram (shown in Figure 2(b)) for the first unfolding events for all the 56 molecules that were studied and showed anomalously low force unfolding events. As expected, the histogram shows a pronounced bimodal distribution in unfolding forces, with the first peak centered at 50(±20) pN and the second at 112(±19) pN. The latter one is very close to the mechanical stability of the FnIII domains in the native state (117(±19) pN) measured in the single pulling experiments. This result clearly

extension for a few seconds (typically ∼5 s) to ensure that at least 80% of the unfolded FnIII domains can refold back to their folded states, according to previously measured folding kinetics.24,31 When stretched again, the refolded domains unfold again to produce the characteristic saw-tooth like force-extension relationships. For

Figure 2. Repeated stretching-relaxation of TNfNAll reveals that some FnIII domains fold into a novel folded state. (a) During a typical repeated stretching-relaxation experiment, a TNfNAll molecule containing five FnIII domains was first picked up and extended by AFM to unfold all the five FnIII domains (trace i). Then the protein was relaxed to almost zero extension to allow the unfolded domains to refold (trace ii). After a waiting period (1–10 s), the protein was stretched again. Four unfolding force peaks were observed (trace iii), indicating that four out of five FnIII domains refolded. We repeatedly stretched and relaxed the same molecule many times, resulting in the stretching traces as shown as traces iv to vii (for clarity, relaxation traces are not shown). We observed that most of unfolding force peaks occur at ∼120 pN, similar to the average unfolding force observed in single pulling experiments, indicating that FnIII domains refold into their native states. We also observed that some unfolding events occurred at much lower values, ∼50 pN, as indicated by red arrows. These lower unfolding force events always precede those occurring at 120 pN, indicative of mechanical unfolding hierarchy. (b) Histogram of unfolding forces for the first peaks in unfolding traces of TNfNAll during stretching-relaxation cycles at a pulling speed of 400 nm/s. A bimodal distribution of unfolding forces is observed with the first peak centered at 50(±20) pN and the second at 112(±19) pN (Gaussian fits, red lines).
indicates that some FnIII domains can fold into a distinct state that shows marked difference from the native state in mechanical stability.

**ΔLc shows that a great majority of these anomalous unfolding events correspond to the complete unfolding of FnIII domains**

Contour length increment (ΔLc) upon domain unfolding is an inherent structural parameter of a protein and contains valuable information about a protein’s mechanical topology. We used WLC model of polymer elasticity to fit the consecutive unfolding force peaks (Figure 3(a)) to extract the information about ΔLc for these anomalously low force unfolding events. We measured ΔLc for a total of 379 such low force events from 56 different molecules and the histogram is shown in Figure 3(b). It is evident that ΔLc for these low force events shows a broad distribution, spanning a range from ∼17 nm up to ∼70 nm. However, a great majority of these events (295 out of 379, or 78%) showed ΔLc around 29 nm, giving rise to the predominant peak. A Gaussian fit to the experimental data shows an average of 29.3(±1.1) nm (n=295, average ±S. D.), which is identical to the contour length increment for unfolding force peaks occurring at regular unfolding forces ∼120 pN (Figure 1(b)). This can be further elaborated by superimposing two recordings on top of each other (Figure 3(c)). Fitting WLC to experimental data can lead to some uncertainty in estimating ΔLc. In contrast, superposition of two force-extension curves avoids such potential ambiguity and clearly demonstrates that the low force peaks have the same ΔLc as those regular force peaks. This result suggests that the anomalously low force unfolding events correspond to the complete unfolding of FnIII domains that fold into a previously un-recognized, novel folded state. We refer to this state as N* state.

In addition to the main peak at 29.3 nm, 84 unfolding events, which amount to 22%, occurred at different values of ΔLc which clustered at ∼23 nm and ∼60 nm (Figure 3(b)), respectively. Since these events are observed during single molecule stretching-relaxation cycles, we are certain that these events indeed correspond to the mechanical unfolding of FnIII domains. Hence, these events may result from the unfolding of partially folded states or dimeric misfolded states formed by two neighboring domains, as those reported previously for tenascin FnIII domains. In our subsequent analysis, we focused on the events that correspond to the unfolding of FnIII domains that fold into the N* state.

The contour length increment ΔLc can be used to derive structural information on the topology of a protein or intermediate state along the mechanical unfolding pathway, which otherwise cannot be obtained by traditional techniques. Within the resolution of our measurements, ΔLc for the N* state is indistinguishable from that of the native state, strongly suggesting that the overall structure of the N* state is very similar to that of the native state. In addition, L_difference (L_difference = L_contour − ΔLc), a measure of the distance between the N and C termini at the mechanical unfolding transition state, for the N*
state is also identical to that of the native state, suggesting that mechanical resistance is localized at a similar region of the folded structure, despite the fact that the native state and N* state have very different mechanical stability.

**N* state is mechanically less stable than the native state**

The force histogram shown in Figure 2(b) contains ~22% events, which amount to a total of 84 events, that do not correspond to the unfolding of N* states, as their $\Delta L_c$ are not ~29 nm. Hence, we have corrected the unfolding force histogram shown in Figure 2(b) by removing these 22% events, and the corrected histogram is shown in Figure 4. It is evident that this correction has little effect on the average unfolding forces for the two peaks. The width of the unfolding force histogram is directly related to the distance between the folded state and the transition state ($\Delta x_u$) along the mechanical unfolding reaction coordinate. The similar width of the two unfolding force peaks in Figure 2(b) and Figure 4 suggests that the N* state and native state have similar $\Delta x_u$. This is further confirmed by Monte Carlo simulations. Using Monte Carlo simulation and two-stage unfolding, we found that the unfolding force distribution for the N* state can be explained adequately using an unfolding rate constant $\alpha_0$ of 0.15 s$^{-1}$ and $\Delta x_u$ of 0.31 nm (continuous line, Figure 4). The regular force peaks corresponding to the native state can be reproduced using the same unfolding rate constant $\alpha_0$ of $5 \times 10^{-3}$ s$^{-1}$ and $\Delta x_u$ of 0.31 nm as we used for the single pulling experiment. This result strongly indicates that the N* state is conformationally similar to the native state, but mechanically less stable.

**Figure 4.** Corrected unfolding force histogram for the first unfolding force peaks during repeated stretching-relaxation cycles. This histogram was corrected from Figure 2(b) by removing the low force unfolding events that do not correspond to the unfolding of N* state in Figure 2(b). The unfolding forces show a bimodal distribution with the first peak centered at ~50 pN and the second at ~110 pN. The ratio of the two populations is ~1:1. Continuous lines correspond to Monte Carlo simulations of the unfolding of an FnIII domain at a pulling speed of 400 nm/s. The parameters used in Monte Carlo simulations are $\Delta x_u$ of 0.31 nm and $\alpha_0$ of 0.15 s$^{-1}$ for the first peak, and $\Delta x_u$ of 0.31 nm and $\alpha_0$ of $5 \times 10^{-3}$ s$^{-1}$ for the second peak, respectively.

**Discussion**

**N* state is unique to a few FnIII domains**

By repeatedly stretching and relaxing the same TNfnALL molecule, we discovered a mechanically less stable state, N* state, for FnIII domains in tenascin. This N* state is a previously unrecognized state, and is conformationally similar to the native state. This finding reveals the dynamic and heterogeneous picture of the ensemble of folded FnIII domains.

However, TNfnALL is a heterogeneous tandem modular protein. Each individual FnIII domain in TNfnALL is different from each of the others, although they do unfold at similar forces when subject to a mechanical stretching force. Is this N* state unique to a particular FnIII domain or a common feature shared by all the 15 FnIII domains in tenascin? From a total of 59 TNfnALL molecules containing three to nine domains studied using the repeated stretching-relaxation protocols, we found that three molecules did not show any unfolding of the N* state (one molecule contained four FnIII domains, and the other two contained three FnIII domains), assuming that any event with an unfolding force below 70 pN corresponds to a novel folded state. The remaining 56 molecules contained between one and three unfolding force peaks that correspond to the unfolding of the N* state. This observation indicates that not all of the 15 FnIII domains are capable of folding into the N* state. Although in one force-extension curve we observed up to three lower force peaks, the vast majority of force-extension curves (270 out of 295) contained only one unfolding event of the N* state, suggesting that only one domain in the stretched fragment is capable of folding into the N* state. Since TNfnALL fragments were picked up randomly along the contour of the intact TNfnALL proteins by the AFM tip, different molecules may contain different stretches of FnIII domains and none of the 56 molecules is the full-length intact TNfnALL protein. Hence, our observations suggest that the novel folding behavior is neither specific to only one particular FnIII domain, nor a common feature for all the 15 FnIII domains. Instead, this novel folding behavior is unique to a few FnIII domains. Indeed, our recent refolding experiments on polypeptide (TNfm3)$_{15}$, which is composed of eight identical tandem repeats of TNfm3, showed that TNfm3 does not show this novel folding behavior. Hence, it is of importance to identify the exact FnIII domains that can fold into the N* state in the future, as these domains will provide ideal model systems for detailed structural and kinetic characterizations of the N* state. To accomplish such goals, expanded experimental efforts are required to examine the folding properties of the remaining FnIII domains individually by engineering polypeptides made of identical tandem repeats of individual FnIII domain.
N* state or stochastic unfolding?

Mechanical unfolding of proteins is a stochastic process and it is possible that a protein will unfold at lower forces with finite probability, raising the question whether the unfolding events of the N* state are merely resulted from the stochastic unfolding of the native state of FnIII domains. Indeed, we observed some unfolding events (63 out of 1640 events) occurring at lower forces (<70 pN) in single pulling experiments, which form the tail for the unfolding force distribution. Such a low force tail in the unfolding force histogram is typical of mechanical unfolding of proteins and reflects the stochastic nature of protein unfolding. However, the frequency of occurrences for low force unfolding events in repeated stretching-relaxation experiments stands in sharp contrast with that observed in single pulling experiments, in disagreement with the possibility that the low force events in repeated stretching-relaxation experiments are merely due to the stochastic unfolding. In single pulling experiments, low force events (<70 pN) are only about 3.8% (63/1640) of the total number of unfolding events; while the low force events in repeated stretching-relaxation experiments amount to 8.1% (295/3625) of all the unfolding events with the regular 29 nm spacing, corroborating our conclusion that the repeated stretching-relaxation experiments expose a new mechanical unfolding hierarchy. Previous single molecule AFM experiments have demonstrated that the mechanical unfolding of different domains in a heterogeneous modular protein follow their mechanical unfolding hierarchy. For example, in single pulling experiments only 65% (41/63) of the low force peaks are shown to be the first one. Therefore, the lower force unfolding events can only be attributed to the unfolding of states that distinctly differ from the native state and have intrinsic lower mechanical stability, corroborating our conclusion that some FnIII domains can fold into a mechanically less stable N* state.

N* state is a weakly populated, kinetically stable state

As discussed before (Figure 2(a)), during repeated stretching-relaxation cycles, the same FnIII domain can unfold at either ~120 pN or ~50 pN, suggesting that an unfolded FnIII domain can fold back to either the native state or the N* state following distinct routes. It is possible that, in one molecule (such as the one shown in Figure 2) undergoing repeated unfolding-refolding cycles, the lower force unfolding events (the unfolding of N* state) observed in different force-extension curves originate from the same FnIII domain. Then, for a given FnIII domain that can fold into the N* state, the ratio between folding into the native state versus N* state is ~1:1 (estimated from the corrected bimodal distribution shown in Figure 4). However, it is also possible that a number of FnIII domains in the same molecule are able to form the N* state with a much lower probability than 1:1. To fully dissect these two possibilities, one has to map the FnIII domains that can fold into the N* state and carry out detailed measurements of the folding behaviors of homopolyproteins made of identical tandem repeats of these identified FnIII domains. Nonetheless, it is clear that a few FnIII domains can assume two distinct folded states: the native state or the N* state, revealing the heterogeneic nature of the folded microstates of some FnIII domains of tenascin.

As shown before, the unfolding of the N* state was only rarely (~3.8%) observed in single pulling experiments, while its unfolding was observed more frequently in refolding experiments (~8.1%). This difference in abundance is most likely originating from the kinetics of conformational conversion between the N* state and the native state. Normally the native state is also the most stable state. Upon reaching thermal equilibrium, the native state will be most populated, giving rise to the low abundance of the N* state in single pulling experiments. However, during stretching-relaxation cycles, the abundance of the N* state is enriched, suggesting that the N* state is kinetically stable and the conversion kinetics between the N* state and the native state is slow. As such, kinetic trapping allows a time window for us to observe the N* state with increased frequency. In our experiments, we changed the waiting time between consecutive stretching-relaxation cycles from 1 s up to 10 s, and we did not observe any pronounced change in the relative population of the two folded states. This result suggests that it is most likely that it takes more than 10 s for the conversion to occur between the native state and N* state.

N* state versus folding intermediate state

Folding intermediate states are frequent occurrences in protein folding studies. Typically, a part of the structure in these folding intermediate states is not formed, resulting in spectroscopic signatures that differ from that of the native state and can be readily recognized in traditional ensemble measurements. However, it is difficult
to capture rare states in ensemble experiments. In this aspect, single molecule techniques have unique advantages. Previous single molecule force spectroscopy experiments revealed various rare partial folded states and folding intermediate states for different proteins by identifying different $\Delta L_c$. However, the N* state we report here is different from these “traditional” intermediate states in a few aspects, although one may well still call the N* state as an intermediate state. The N* state for FnIII domains in tenascin has the same $\Delta L_c$ and $\Delta x_u$ as the native state, indicating that the N* state has an overall structure that closely resembles that of the native state and the mechanical resistance is localized at similar region of the folded structure. The significant difference in the mechanical stability may merely result from minute differences in the bonding patterns of the local structure that are critical for the mechanical stability in these two folded states. For example, it is possible that non-native contact(s), instead of native ones, were formed in the key region that determines the mechanical stability while the rest of the overall structure folded correctly. The N* state reported here shares common features with the recently discovered new folded microstates of enzymes (so-called “excited state” of enzymes) by NMR relaxation experiments and single molecule spectroscopy studies. In these examples, the novel folded microstates are invisible in traditional spectroscopic studies, because the structure of the novel folded microstate only differs from the native state by a few local conformations, such as the loop conformation, which results in energetic difference with the native state. For example, Wright and colleagues showed that a ternary complex formed by dihydrofolate reductase and its substrate showed conformational exchange between a ground state, in which the active site loops adopt a closed conformation, and a weakly populated (4.2%) excited state with the loops in the occluded conformation.

Possible roles of proline residues in the formation of N* state

Proline cis-trans isomerization is a frequent cause for slow folding and folding intermediate state. Sequence analysis of tenascin showed that on average each FnIII domain contains six proline residues (the actual number of proline residues for individual FnIII ranges from 4 to 12), raising the question whether the observed N* state for FnIII domains can result from proline isomerization, which may give rise to the decreasing mechanical stability. Multiple sequence alignment of the constitutive and alternative spliced FnIII domains (Figure S3 in Supplementary Data) shows that three proline residues are conserved in constitutive FnIII domains, while only one proline residue is conserved within alternatively spliced FnIII domains. It is unlikely that these conserved proline residues are responsible for the formation of the N* state. As shown above, we carried out refolding experiments on polyprotein (TNfn3)$_8$ and did not observe the formation of N* state, despite that Fn3 contains five proline residues. This result excludes the role of conserved proline residues in the N* state. However, without detailed three-dimensional structural information of the FnIII domains (especially the information about the configuration of proline residues), we cannot rule out the possibility that there exist some unique proline residues in a few FnIII domains that may indeed play important roles in the formation of N* state. For example, the sixth and seventh FnIII domains (TNfn6 and TNfn7) contain eight and 12 proline residues, respectively. It will be of importance to investigate the potential role in refolding kinetics of the non-conserved proline residues in such proline-rich FnIII domains.

Biological implications

Considering the mechanical function of tenasin, we speculate that the N* state of FnIII domains reported here may carry important biological implications in the mechanotransduction in extracellular matrix. Two distinct folding routes are available for some FnIII domains, resulting in a quasi-parallel folding reaction to the folded ensemble. The apparent folding rate constant will be the sum of the two rate constants along the two distinct folding routes from the unfolded state to the native and N* state, effectively reducing the time that an unfolded FnIII domain could spend in the unfolded state and helping the unfolded FnIII domain to rapidly refold into a mechanically stable state, be it the N* state or native state. Tenascins play important roles in providing elasticity and mechanical strength to tissues and regulating the cell–matrix interactions. For such purposes, the mechanical stability of tenasin is of great importance. Since tenasin is subject to mechanical stretching force in vivo, the unfolding and folding reactions may be part of its natural life cycle. It is required that tenascins regain their mechanical stability rapidly. The complex quasi-parallel folding scheme for some FnIII domains reported here may help fulfill such requirement.

Materials and Methods

Materials

The recombinant TNfnALL fragment comprising 15 FnIII domains from human tenascin-C was a generous gift from Harold Erickson of Duke University.

Single molecule AFM

Single-molecule AFM experiments were carried out on a custom-built atomic force microscope, which was constructed as described. All the force-extension measurements were done in PBS buffer. In a typical experiment,
1 μl of recombinant TNfnALL protein sample (at a concentration of ∼530 μg/ml) was deposited onto a clean glass cover slip cover by ~50 μl PBS buffer, and was allowed to adsorb for ~5 min before pulling experiments proceeded. The spring constant of each individual cantilever (Si3N4 cantilevers from Vecco, with a typical spring constant of 40 pN/nm) was calibrated in each experiment.46,47

During the unfolding experiment, the cantilever was brought into contact with the substrate to pick up a molecule and then pulled away at a constant pulling velocity. The contact force between the cantilever and the substrate was about 600–1000 pN and the pulling speed was 400 nm/s.

Drift is a common problem for AFM pulling experiments, especially for repeated stretching-relaxation experiments. To minimize the effect of drift on our repeated stretching-relaxation measurements, we screened the AFM cantilevers and only used those with minimum drift in our refolding experiments. In doing so, we ensure that the drift in 10 s in our refolding experiments is negligible. During the refolding experiments, we checked the contact point by gently touch the surface every few stretching-relaxation cycles and then adjust the starting point of piezo scan accordingly to ensure that the polyprotein was relaxed to zero extension.

Monte Carlo simulations

The mechanical unfolding of FnIII domains in tenascins can be modeled as an all-or-none process with force-dependent rate constants,35,48,49 in which only the folded and unfolded states will populate during the reaction. The force-dependent unfolding rate constant can be described as: \( \alpha(F) = \alpha_0 \exp(F \Delta \chi_u/k_B T) \), where \( k_B \) is the Boltzmann constant, \( T \) is the temperature in Kelvin, \( \alpha(F) \) is the unfolding rate constant at a stretching force of \( F \), \( \alpha_0 \) is the unfolding rate constant at zero force, and \( \Delta \chi_u \) is the distance between the folded state and the transition state. Monte Carlo simulations were carried out according to published procedures2,3,33 to estimate the unfolding rate constant at zero force (\( \alpha_0 \)) and the distance of the folded state to the transition state (\( \Delta \chi_u \)) along the reaction coordinate of the mechanical unfolding.

Acknowledgements

This work is supported by the Canada Research Chair program and the start-up grant from University of British Columbia. Y. C. is partially supported by the Laird Fellowship. The authors also thank Professor Harold Erickson for the generous gift of TNfnALL samples and Dr Deepak Sharma for his help with the sequence alignment of FnIII domains.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2006.06.014

References

A Weakly Populated Microstate of the FnIII Domains


Edited by C. R. Matthews

(Received 14 March 2006; received in revised form 8 June 2006; accepted 8 June 2006)
Available online 27 June 2006