doi:10.1016/j.jmb.2009.01.019 J. Mol. Biol. (2009) 386, 1327–1342

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Mechanical Design of the Third FnIII Domain of Tenascin-C

Qing Peng, Shulin Zhuang, Meijia Wang, Yi Cao, Yuanai Khor and Hongbin Li*

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

Received 24 October 2008; received in revised form 9 January 2009; accepted 14 January 2009
Available online 22 January 2009

By combining single-molecule atomic force microscopy (AFM), proline mutagenesis and steered molecular dynamics (SMD) simulations, we investigated the mechanical unfolding dynamics and mechanical design of the third fibronectin type III domain of tenascin-C (TNfn3) in detail. We found that the mechanical stability of TNfn3 is similar to that of other constituting FnIII domains of tenascin-C, and the unfolding process of TNfn3 is an apparent two-state process. By employing proline mutagenesis to block the formation of backbone hydrogen bonds and introduce structural disruption in β sheet, we revealed that in addition to the important roles played by hydrophobic core packing, backbone hydrogen bonds in β hairpins are also responsible for the overall mechanical stability of TNfn3. Furthermore, proline mutagenesis revealed that the mechanical design of TNfn3 is robust and the mechanical stability of TNfn3 is very resistant to structural disruptions caused by proline substitutions in β sheets. Proline mutant F88P is one exception, as the proline mutation at position 88 reduced the mechanical stability of TNfn3 significantly and led to unfolding forces of <20 pN. This result suggests that Phe88 is a weak point of the mechanical resistance for TNfn3. We used SMD simulations to understand the molecular details underlying the mechanical unfolding of TNfn3. The comparison between the AFM results and SMD simulations revealed similarities and discrepancies between the two. We compared the mechanical unfolding and design of TNfn3 and its structural homologue, the tenth FnIII domain from fibronectin. These results revealed the complexity underlying the mechanical design of FnIII domains and will serve as a starting point for systematically analyzing the mechanical architecture of other FnIII domains in tenascins-C, and will help to gain a better understanding of some of the complex features observed for the stretching of native tenascin-C.

Keywords: mechanical unfolding; single molecule force spectroscopy; steered molecular dynamics simulation; unfolding pathway; tenascin

Edited by C. R. Matthews

Introduction

It is well recognized that mechanotransduction is a mechanism in which mechanical force acts on cells to serve as a physiological signal that triggers a variety of biological processes.[1–3] In living tissues, mechanical force is transmitted to cells through the extracellular matrix (ECM), which serves as a mechanical scaffold for cells to adhere, migrate and differentiate.[4] ECM is linked to the intracellular cytoskeleton through cell membrane receptors integrins and establishes a mechanical continuum that allows the mechanical force be transmitted as a physiological signal between the interior and exterior of cells. A wide variety of ECM proteins are subject to mechanical tension in biological environments and many of them share a similar tandem modular architecture.[4] Mechanical tension alters the conformational states of such mechanical proteins, and may modulate biological functions *via* force-modulated conformational changes. Therefore, understanding the relationship between the struc-

*Corresponding author. E-mail address: hongbin@chem.ubc.ca.

Abbreviations used: ECM, extracellular matrix; FnIII, fibronectin type III; AFM, atomic force microscopy; TNfn3, the third FnIII domain of tenascin-C; SMD, steered molecular dynamics; WLC, worm-like chain.

ture and mechanical properties of these proteins is of important biological significance. Tenascin-C, a highly conserved oligomeric ECM glycoprotein,[5–8] is one of the model systems for such studies.

Tenascin-C is an extracellular matrix protein with important roles in regulating the cell-matrix interactions.[7] Tenascin-C is a tandem modular protein and consists of a tenascin-assembly domain, a stretch of epidermal growth factor-like repeats, a fibronectin type III (FnIII) domain region that is composed of a series of FnIII domains, and a terminal knob domain that is homologous to the globular domain of fibrinogen. Tenascins are expressed mainly in regions that are subject to heavy tensile load[9] or in tissues that undergo extensive structural re-modeling during processes such as tissue injury and tumorigenesis.[10–12] As tenascins are subject to mechanical stretching forces under physiological conditions, it is possible that the force-induced unfolding/refolding reactions of FnIII domains may be an important part of tenascins dynamics *in vivo*.[13]

Single-molecule atomic force microscopy (AFM) studies have provided insights into the mechanical design and functions of tenascin-C. It was revealed that tenascin-C is an elastic protein that can extend to several times its resting length *via* force-induced unfolding of FnIII domains.[13,14] It has been suggested that the mechanical unfolding of FnIII domains serves as a shock-absorber to prolong the lifetime of the tenascin-ligand bond.[13] Recent studies revealed that some FnIII domains display weakly populated folded microstates in addition to their native states, which may entail a possible mechanism for these FnIII domains to recover their mechanical resistance more rapidly after mechanical unfolding.[15] Since these studies were carried out on native fragments of tenascin-C, the inherent heterogeneity of the constituting FnIII domains in native fragments makes it difficult to assign mechanical features observed on native tenascin fragments to specific FnIII domains, and makes detailed molecular interpretation of the experimental results difficult.

To overcome these difficulties, detailed studies of the mechanical unfolding dynamics of individual FnIII domains become necessary. Using polyproteins made of identical tandem repeats of the protein of interest has become the standard approach to study their mechanical properties using single-molecule AFM.[16,17] Recently, mechanical ϕ value analysis has been done on the third FnIII domain of tenascin-C (TNfn3) *via* single-molecule AFM and molecular dynamics simulation using an implicit water model.[18] However, this study was focused on probing the role of hydrophobic core in the mechanical unfolding of TNfn3. The role of backbone hydrogen bonds and β sheet stability, two important factors for mechanical stability of proteins, were not probed. Here, in order to address the importance of backbone hydrogen bonds as well as β sheet stability in the mechanical unfolding of TNfn3, we combine single-molecule AFM, proline mutagenesis and steered molecular dynamics (SMD) simulations to investigate the mechanical unfolding dynamics and mechanical design of TNfn3 in detail. Our results revealed that the mechanical stability of TNfn3 is similar to that of other FnIII domains of tenascin-C,[13,15] and the unfolding process of TNfn3 is an apparent two-state process. To probe the mechanical design of TNfn3, we used proline mutagenesis to block the formation of backbone hydrogen bonds and introduce structural disruption in β sheet in order to affect its mechanical stability and unfolding kinetics. Our results revealed that hydrophobic core packing is not the only factor that is important for determining the mechanical stability of TNfn3, and backbone hydrogen bonds in β hairpins are also responsible for the overall mechanical stability of TNfn3. Furthermore, proline mutagenesis revealed the robust mechanical design of TNfn3, as the mechanical stability of TNfn3 is resistant to disruptive proline mutations in β sheets of TNfn3. We identified that residue Phe88 is a weak point in TNfn3, and a single substitution of Phe88 with proline results in the unfolding of TNfn3 at forces that are lower than the detection limit in AFM. We compared the AFM results with SMD simulations to understand the molecular details underlying the mechanical unfolding of TNfn3. A comparison between the mechanical features of TNfn3 with the tenth FnIII domain from fibronectin revealed the significant differences in mechanical unfolding and design of these two structurally homologous FnIII domains. These results pave the way for systematic analysis of the mechanical architecture of other FnIII domains in tenascin-C and will help us to obtain a better understanding of some of the complex features observed for the stretching of native tenascin-C.

Results

The mechanical unfolding of TNfn3 is an apparent two-state process

TNfn3 is an all-β protein of 90 amino acid residues. It has a typical immunoglobulin-like β-sandwich structure, in which the two β sheets of TNfn3 pack against each other (Fig. 1).[19] The top sheet consists of strands A-B-E and the bottom sheet consists of C-C′-F-G. The two force-bearing terminal β strands A and G are parallel with each other, and the N- and C-termini point in opposite directions. Such an arrangement of the two terminal strands forms a shear topology upon stretching, which is a common feature among proteins that are mechanically stable.[16,17,20–26] It is of note that the two terminal force-bearing strands of TNfn3 are not bonded directly by backbone hydrogen bonds like other elastomeric proteins, such as I27 of the muscle protein titin.[17]

Using protein engineering techniques, we engineered a polyprotein $(TNfn3)_8$, which consists of eight identical tandem repeats of TNfn3. We then

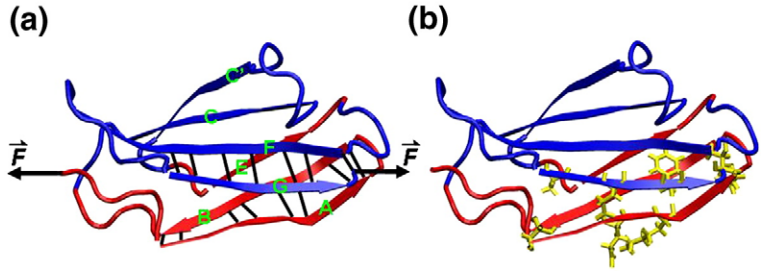

Fig. 1. The three-dimensional structure of the third fibronectin type III domain of tenascin-C (TNfn3). (a) TNfn3 has a typical β-sandwich structure. The two force-bearing β strands are parallel with each other and are pointing in opposite directions. The backbone hydrogen bonds associated with the two force-bearing β strands are indicated by black bars. (b) The location of amino acids that are substituted by proline residues in this work.

used single-molecule AFM to stretch polyprotein $(TNfn3)_8$ and characterize its mechanical unfolding behaviors. Stretching polyprotein $(TNfn3)_8$ resulted in force–extension relationships of the characteristic sawtooth pattern appearance, where the individual sawtooth peaks correspond to the mechanical unfolding of each individual TNfn3 domain in the polyprotein chain. The force–extension relationships of $(TNfn3)_8$ can be well described by the worm-like chain (WLC) model of polymer elasticity,[27] and there is no apparent deviation from the WLC fits (Fig. 2). WLC fits to consecutive unfolding force peaks measure an average contour length increment (ΔL_c) of 29.0±0.8 nm (average±standard deviation). TNfn3 is 90 residues long and the distance between its N-, and C- termini is 3.1 nm in the folded state.[19] The contour length of the unfolded and fully stretched TNfn3 is 32.4 nm (90 residues×0.36 nm/residue). Hence, a complete unraveling of a TNfn3 domain should result in a contour length increment ΔL_c of ~29.3 nm, which is in excellent agreement with the experimentally determined ΔL_c. This result suggests that the mechanical unfolding of TNfn3 corresponds to the complete unfolding of TNfn3 in an apparent two-state fashion, and there is no visible intermediate state along its mechanical unfolding pathway.

The amplitude of the unfolding force peaks varies around ~120 pN. A histogram of the unfolding forces compiled from ~4000 unfolding events at a pulling speed of 400 nm/s measured an average unfolding force of 125±14 pN (n=4198, Fig. 3a), which is in agreement with previous measurements on a similar TNfn3 polyprotein.[18] The measured mechanical stability of TNfn3 is similar to the average mechanical unfolding force of all the 15 FnIII domains measured from a recombinant fragment of tenascin-C containing all the 15 FnIII domains, consistent with the previous conclusion that all the FnIII domains of tenascin-C have similar levels of mechanical stability.[13,28]

It was shown that extending the C-terminus of TNfn3 by two residues can increase the thermodynamic stability of TNfn3 significantly.[29] However, we found that such an extension of TNfn3 does not affect the mechanical stability of TNfn3 in any way (data not shown), indicating that the two additional residues at the C-terminus of TNfn3 are already detached from the folded TNfn3 before TNfn3 reaches the mechanical unfolding transition state.

The mechanical unfolding of TNfn3 is characterized by a long unfolding distance from the native state to the transition state

To further characterize the mechanical unfolding of TNfn3 in detail, we measured the pulling speed-dependence of the unfolding forces of TNfn3 by stretching $(TNfn3)_8$ at different pulling speeds. Similar to other elastomeric proteins, the mechanical unfolding of TNfn3 is a non-equilibrium process and its unfolding force depends on the pulling speed: the higher the pulling speed is, the bigger the unfolding force (Fig. 3b). It is of note that, as compared with Ig domains from the muscle protein titin,[17,30,31] the pulling speed-dependence of the unfolding force of TNfn3 is relatively weak: the unfolding force of TNfn3 increases from 109 pN at a pulling speed of 50 nm/s to 154 pN at a pulling speed of 2700 nm/s. To estimate the spontaneous unfolding rate constant at zero force (α_0) and the unfolding distance (Δx_u) between the folded state and the transition state

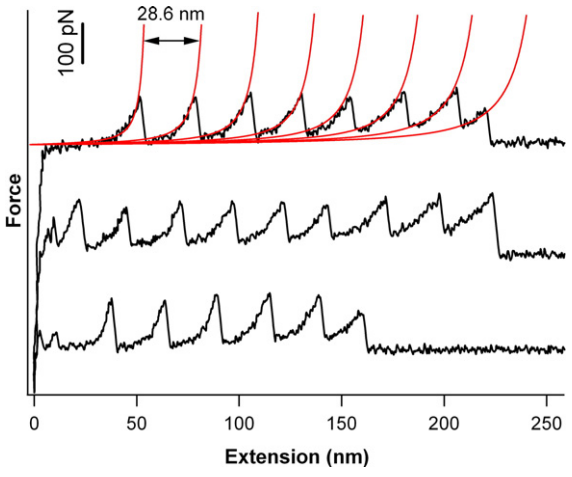

Fig. 2. Typical force-extension curves of polyprotein $(TNfn3)_8$. Stretching polyprotein $(TNfn3)_8$ results in force-extension curves with the characteristic saw-tooth pattern. The equally spaced force peaks result from the mechanical unfolding of the individual TNfn3 domains in the polyprotein chain. The last peak in force-extension curves corresponds to the detachment of the protein from either the AFM tip or the substrate. WLC fits (thin lines) to the consecutive unfolding force peaks measure a contour length increment ΔL_c of ~29.0±0.8 nm.

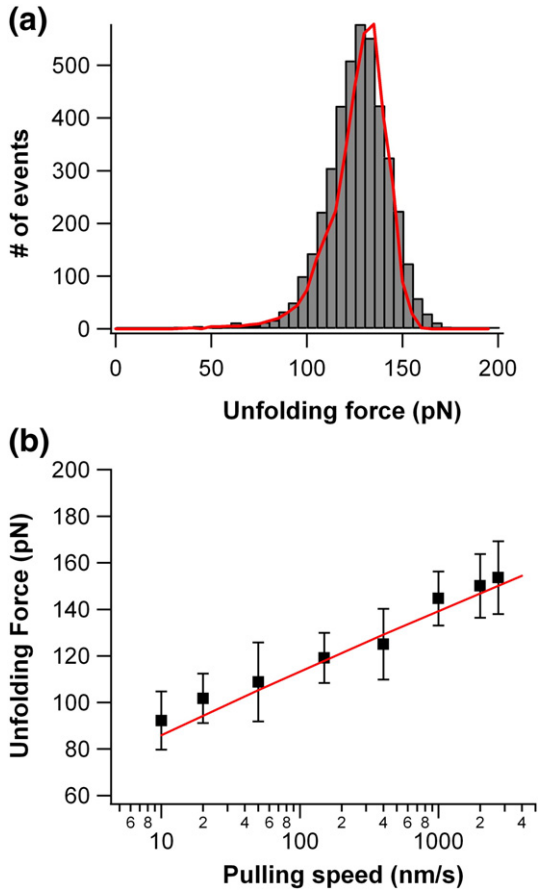

Fig. 3. Unfolding force of TNfn3 and its dependence on the pulling speed. (a) Histogram of unfolding forces for TNfn3. The unfolding force histogram spans a range of ~100 pN (60 pN~160 pN) with an average value of 125± 14 pN (n=4198). The red line corresponds to Monte Carlo simulation of the mechanical unfolding of TNfn3 using α_0 of 1.5×10^{-4} s^{-1} and Δx_u of 0.42 nm. The pulling speed is 400 nm/s. (b) The pulling speed dependence of the unfolding forces of TNfn3 (symbols). The pulling speed dependence of the unfolding forces of TNfn3 can be reproduced adequately by Monte Carlo simulations using α_0 of 1.5×10^{-4} s^{-1} and Δx_u of 0.42 nm (red line).

along the reaction coordinate, two important parameters characterizing the mechanical unfolding free energy diagram, we carried out Monte Carlo simulations to reproduce the force-extension relationships of $(TNfn3)_8$. In the Monte Carlo simulation, we assumed that the unfolding of TNfn3 is a two-state process and the force-dependent unfolding rate constant follows the classical Bell model $\alpha(F) = a_0 \exp\left(\frac{F\Delta x_u}{k_B T}\right)$, where k_B is the Boltzmann constant and T is absolute temperature (in K). We found that both the unfolding force histogram (Fig. 3a) and the pulling speed-dependence of the unfolding forces (Fig. 3b) are well described using an α_0 of 1.5×10^{-4} s^{-1} and a Δx_u of 0.42 nm. This result suggests that the mechanical resistance to unfolding is distributed over a distance of 0.42 nm. This unfolding distance is notably longer than that for other typical elastomeric proteins, such as I27,[17] ubiquitin,[22] and GB1 domains,[23] suggesting that the mechanical resistance of TNfn3 is distributed along a longer distance, which is in contrast with the highly localized mechanical resistance for other elastomeric proteins. The molecular basis for the observed long unfolding distance Δx_u will be addressed in Discussion. The measured α_0 and Δx_u are comparable to those measured in previous studies on tenascin-C fragment and TNfn3 polyprotein.[13,18]

The SMD simulations of the mechanical unfolding of TNfn3

The single-molecule AFM results suggest that the mechanical unfolding of TNfn3 is an apparent two-state process. In order to understand the molecular events leading to the mechanical unfolding of TNfn3, we carried out SMD simulations of the mechanical unfolding of TNfn3. In contrast to previous molecular dynamics simulation work using an implicit water model,[18] we used explicit water model TIP3P in our simulation to explicitly address the potential role of the solvent water molecules during the mechanical unfolding process of TNfn3. This strategy has been used extensively to simulate the mechanical unfolding of proteins, including FnIII domains from fibronectin and TNfn3.[32,33]

In our SMD simulations, we simulated the mechanical unfolding of TNfn3 using both constant velocity and constant force protocols. Starting from 1 ns and 1.5 ns equilibrated conformations, TNfn3 was pulled at a constant force (500 pN) or at a constant velocity (0.05 Å/ps). In total, we performed 16 constant force SMD and 17 constant velocity SMD simulations with a total simulation time of 57 ns. Both constant velocity and constant force trajectories revealed similar features of the mechanical unfolding of TNfn3.

In constant force SMD simulations, TNfn3 was stretched at a constant force of 500 pN from its N- and C-termini, and the distance between the two termini (R_{NC}) was monitored as a function of time. Representative R_{NC} *versus* time curves from constant force SMD simulations are shown in Fig. 4a. The presence of multiple plateaus in R_{NC} *versus* time profiles is clearly visible, indicating the presence of stable intermediates populated along the unfolding trajectories. From the native state with R_{NC} of ~34 Å, TNfn3 elongates by ~6 Å *via* straightening the disordered N-terminal end of the protein and enters into a stable intermediate state I1, in which the tertiary structure of TNfn3 remains largely intact. It is of note that the two backbone hydrogen bonds between Ser6 and Phe23 are relatively weak and break in state I1 during most of the trajectories. The rupture of the backbone hydrogen bonds between Ser6 and Phe23 leads to the detachment of the first six residues of strand A from the folded structure of TNfn3. In the second stage, TNfn3 elongates further by ~10 Å to reach the intermediate state I2. During this process, the two β sheets, one containing strands A–B –E (colored red) and the

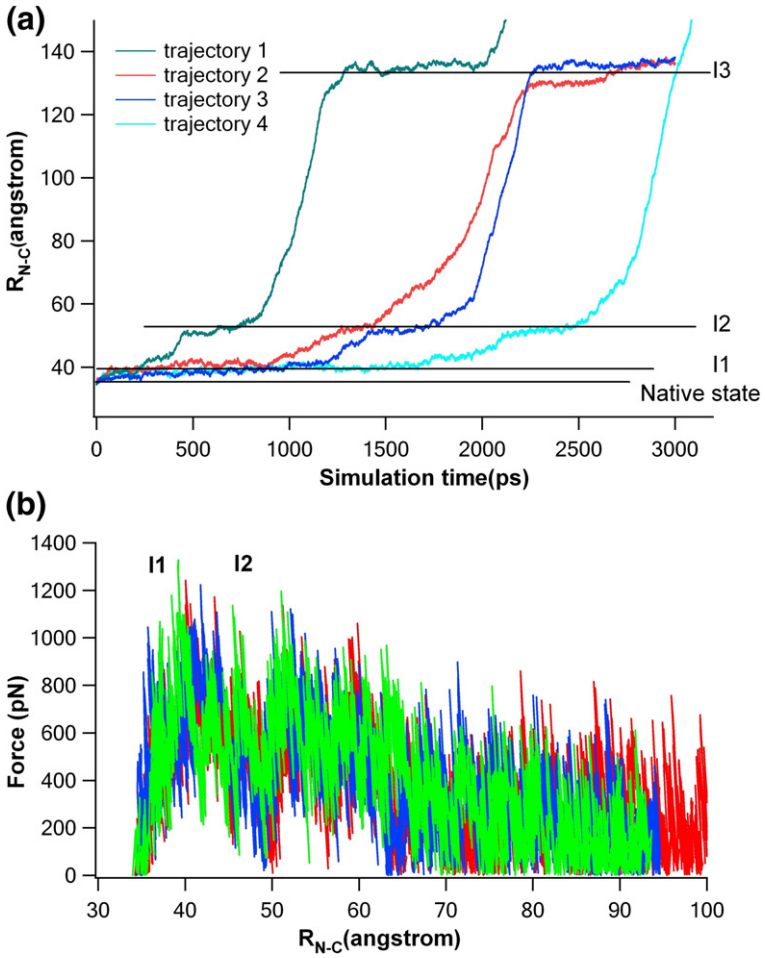

Fig. 4. Constant force and constant velocity SMD simulations of the mechanical unfolding of TNfn3. (a) Representative R_{NC} *versus* time profiles from constant force SMD simulations at a stretching force of 500 pN. The presence of three plateaus indicated kinetic intermediates I1, I2 and I3. Trajectories 1–3 correspond to the A-strand separates first pathway, and trajectory 4 corresponds to the A-G strands separate simultaneously pathway. (b) Representative *Force versus* R_{NC} curves from constant velocity SMD simulations. The pulling velocity is 0.05 Å/ps. The first force peak occurs at an R_{NC} of ~40 Å, corresponding to the transition from I1 to I2; the second peak occurs at an R_{NC} of ~50 Å and corresponds to the unraveling of intermediate I2.

other containing strands C′–C–F–G (colored blue), rotate relative to each other and both align with the pulling force (Fig. 5), leading to the so-called aligned β sandwich intermediate I2. This alignment resulted in partial solvation of the periphery of the hydrophobic core. In intermediate I2, the remaining backbone hydrogen bonds in strands A-B between Glu9–Trp21, Lys11–Leu19, Asp12–Leu19, and Thr14–Thr17 remained intact, so do the backbone hydrogen bonds in hairpin F-G.

Immediately following the transition from I1 to I2, TNfn3 elongated further following two distinct unfolding pathways (Fig. 5). The first pathway is characterized by the presence of the partially unfolded intermediate state I3 with R_{NC} of ~130 Å. Along this pathway (the "A-strand separates first" pathway), strand A of TNfn3 separates first from the folded structure, followed by the subsequent unraveling of strands B and E. However, the overall structure of the β sheet containing strands C′-C-F-G remains largely intact. This unfolding pathway was observed in ~70% of the constant force SMD simulation trajectories (11 out of 16) and three representative trajectories of this type are shown in Fig. 4a (trajectories 1–3). The second pathway is characterized by simultaneous detachment of strands A and G from the folded structure and is termed "A-G strands separate simultaneously". After the simultaneous detachment of strands A and G, TNfn3 unfolds readily without significant barrier and hence intermediate state I3 is not detected in this pathway. This type of pathway was observed in ~30% of the constant force SMD simulations (5 out of 16) and trajectory 4 in Fig. 4a is one example.

The dwell time of a given state along the unfolding trajectory is a measure of the stability of the given state. Constant force simulations revealed that the dwell time of intermediate I1 is, on average, longer than that of I2 and I3, as well as that of the native state, suggesting that intermediate state I1 is the most stable one at a force of 500 pN. This result suggests that under a stretching force of 500 pN, the native state of TNfn3 is transformed rapidly into intermediate state I1, and the mechanical unfolding of TNfn3 is not directly from its native state. Thus, the intermediate I1 is the pseudo-native state for the mechanical unfolding of TNfn3. It is important to point out that the dwell time of individual intermediate states does vary from trajectory to trajectory.

To reveal a more detailed picture of the transition from mechanical unfolding of TNfn3, we monitored the breakage of backbone hydrogen bonds in hairpins A-B and F-G by calculating hydrogen bond energies along the SMD trajectories. Five backbone hydrogen bonds between strands A and B,Glu9(H)–Thr21(O), Glu9(O)–Thr21(H), Lys11(H)–Leu19(O), Asp12(O)–Leu19(H), Thr14(H)–Thr17(O) and

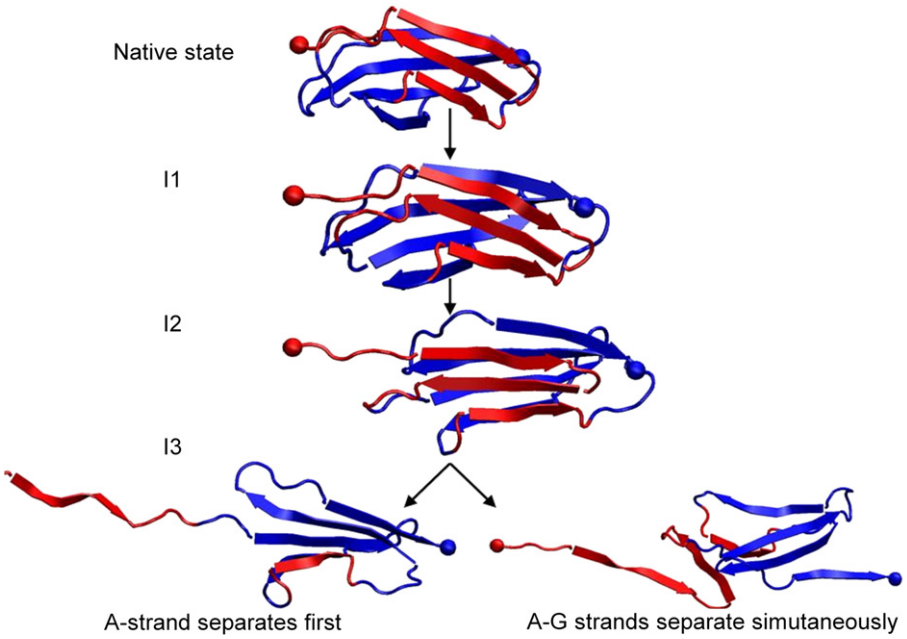

Fig. 5. Snapshots of TNfn3 during its simulated mechanical unfolding. For all the snapshots, the N terminus (red ball) was fixed and the C terminus (blue ball) was pulled during these simulations. From the native state, TNfn3 elongates by straightening the N-terminus and enters into the so-called twist intermediate state I1; then the two β sheets rotate relative to each other and align with the stretching force vector, leading to the so-called aligned intermediate state I2. After I2, TNfn3 unfolds *via* two distinct pathways: the first is by separating A strand first from the folded structure, followed by subsequent unfolding of B and E strands, leading to a partially unfolded intermediate state I3; the second pathway is *via* simultaneous unfolding and detachment of A and G strands from the folded structure. After this event, TNfn3 unfolds readily and the existence of intermediate state I3 is not detected. For the snapshot of I3, the N terminus is artificially shortened to fit into the figure.

another five backbone hydrogen bonds between strands F and G, Phe88(H)–Tyr68(O), Glu86(O)–Val70(H), Glu86(H)–Val70(O), Ala84(O)–Leu72(H), Ala 84(H)–Thr72(O) were selected for the hydrogen bond energy calculation. Figure 6 shows the hydrogen bond energy as a function of time for these ten backbone hydrogen bonds for the trajectories 2 and 4 shown in Fig. 4a. At the beginning of the unfolding trajectories, these backbone hydrogen bonds are stable and the hydrogen bond energies fluctuate between approximately −5 kcal/mol and −3 kcal/mol with an average value of approximately −4.3 kcal/mol. Upon further stretching, β strands will begin to separate and the strength of its hydrogen bonds become weaker; accordingly, the hydrogen bond energies gradually increase to zero, at which point the hydrogen bonds are already broken. In the "A-strand separates first" pathway (Fig. 6a and b), after ~1.8 ns, the hydrogen bond energy of Glu9(H)–Thr21(O), Glu9(O)–Thr21(H), Lys11(H)–Leu19(O), Asp12(O)–Leu19(H), Thr14(H)–Thr17(O)increase to zero fairly rapidly accompanying the transition from intermediate state I2 to I3, indicating that these five backbone hydrogen bonds in strands A and B break concurrently during this process. In contrast, the hydrogen bonds in strands F and G (Phe88(H) – Tyr68(O), Glu86(O) –Val70(H), Glu86(H) –Val70(O), Ala84(O) – Leu72(H), Ala 84(H) – Thr72(O)) remain steady during this process. After 2.6 ns, three hydrogen bonds, Phe88(H) – Tyr68(O), Glu86(O) –Val70(H) and Glu86(H) –Val70(O), start to break.

In the "A-G strands separate simultaneously" pathway, the hydrogen bond energy for Glu9(H) – Thr21(O), Glu9(O) –Thr21(H), Lys11(H)–Leu19(O), Asp12(O) – Leu19(H), Thr14(H) – Thr17(O), Phe88 (H) – Tyr68(O), Glu86(O) –Val70(H), Glu86(H) – Val70(O), Ala84(O) – Leu72(H), and Ala 84(H) – Thr72(O) increases simultaneously after 2.7 ns, indicating that these ten backbone hydrogen bonds break concurrently during the "A-G strands separate simultaneously" pathway (Fig. 6c and d).

Constant velocity SMD simulations

We also carried out constant velocity SMD simulations of the mechanical unfolding of TNfn3 (17 trajectories in total). Figure 4b shows representative force *versus* R_{NC} curves. It is evident that there are multiple force peaks along the unfolding pathway: the first force peak occurs at an R_{NC} of ~40 Å, which corresponds to the transition from I1 to I2; the second peak occurs at an R_{NC} of ~50 Å and corresponds to the unraveling of intermediate state I2. In all the trajectories, the amplitudes of the first and second peak show slight difference: in 13 trajectories, the first unfolding force peak is higher than the second one (~1300 pN *versus* ~1050 pN), while in four trajectories, the second unfolding force peak in higher than the first one (~1200 pN *versus* 1100 pN). This

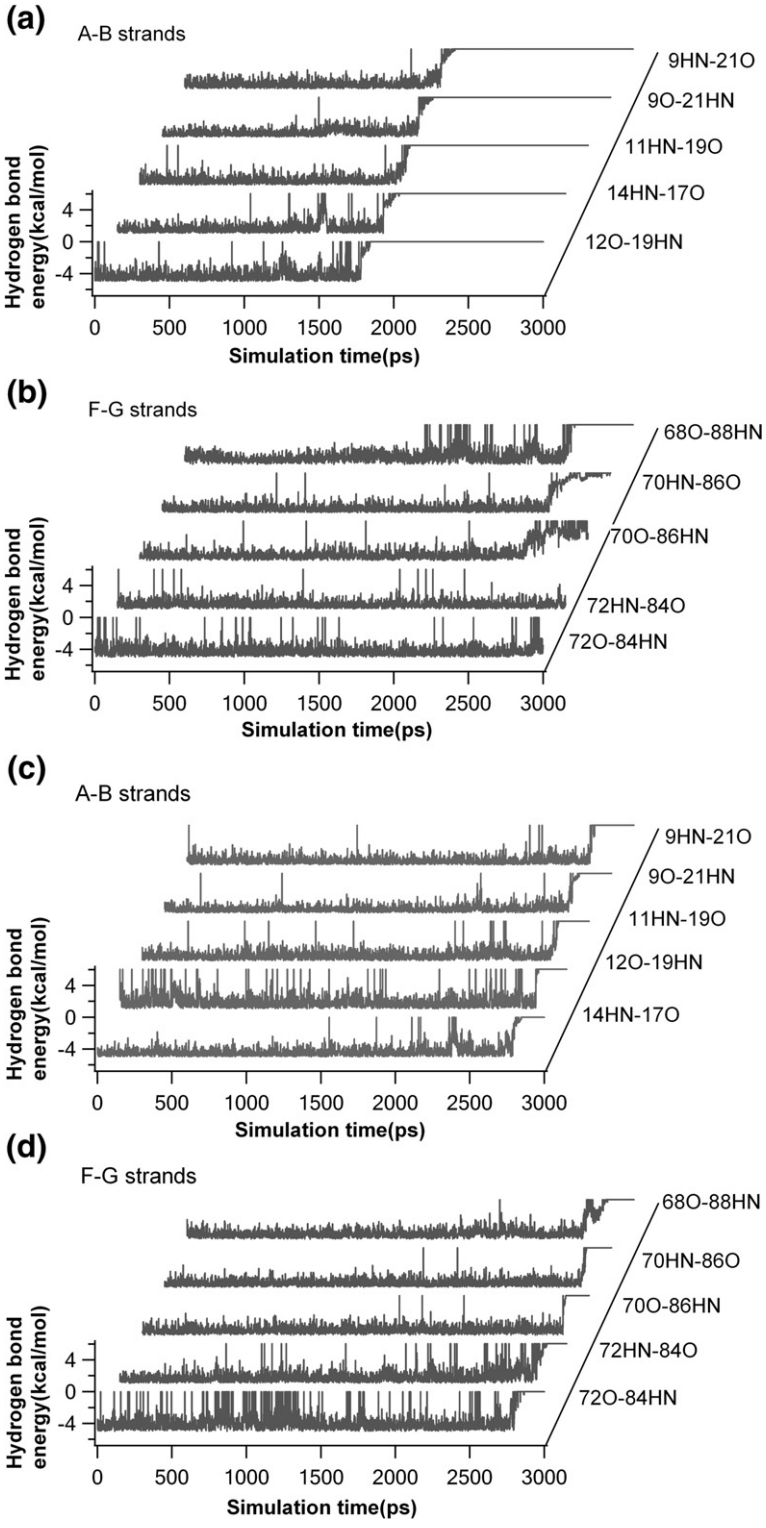

Fig. 6. Profiles of hydrogen bond energy of inter-strand hydrogen bond in A-B and F-G strands *versus* time in two representative SMD unfolding trajectories of TNfn3. (a and b) The energy change of the hydrogen bonds as a function of time in the trajectory following the A-strand separates first pathway. The energy was calculated from trajectory 2 in Fig. 4a. (c and d) The energy change of the hydrogen bonds as a function of time in the trajectory following the A-G strands separate simultaneously pathway. The energy was calculated from trajectory 4 in Fig. 4a. The sudden increase in hydrogen bond energy indicates the breaking of hydrogen bonds. In both unfolding pathways, the hydrogen bonds in A-B and F-G break during the unfolding of the intermediate state I2.

observation is consistent with the observed variation in dwell time of the intermediates.

Using site-directed mutagenesis to probe the nature of unfolding transition state observed in single-molecule AFM

SMD simulations on TNfn3 revealed that the mechanical unfolding of TNfn3 proceeds *via* several intermediate states. However, single-molecule AFM experiments indicate that the mechanical unfolding of TNfn3 is an apparent two-state process. The discrepancy between SMD simulations and AFM experiments suggests that some of the intermediate states observed in SMD simulations do not populate on the time-scale of single-molecule AFM experiments. Considering the large R_{NC} of the intermediate I3, we can easily rule out the possibility that the intermediate I3 is a stable intermediate state. Previous SMD simulations and mechanical

ϕ-value analysis suggested that the transition from the twist intermediate I1 to the aligned intermediate I2 is most likely to be the rate-limiting step for the mechanical unfolding observed in AFM, and that the energy barrier for the transition from intermediate I2 to I3 is too low to be observed experimentally.[18,33] Our SMD simulation results reveal that in majority of the unfolding trajectories, the plateau of intermediate state I1 is the longest in constant force SMD and the unfolding force peak for the transition from I1 to I2 is the highest in constant velocity SMD, supporting the view that the transition from I1 to I2 is the rate-limiting step, which corresponds to the mechanical unfolding force peak observed in single-molecule AFM experiments.

If this view is correct, it means that the rupture events of the backbone hydrogen bonds (Fig. 5) occur after the rate-limiting step. Thus, it is tempting to conclude that the hydrogen bonds are not critical for the mechanical stability of TNfn3. For example, a previous single-molecule AFM study on TNfn3 singled out the importance of hydrophobic interactions to the mechanical unfolding of TNfn3.[18,34] However, the backbone hydrogen bonds are important in protecting the hydrophobic core from the attack by water molecules, especially the backbone hydrogen bonds in hairpins A-B and F-G. It can be imagined that destabilization of strands A-B or F-G by deleting backbone hydrogen bonds and introduction of a bulge would facilitate the attack of the hydrophobic core by water molecules and lead to reduced mechanical stability. To further explore such scenarios, we used proline mutagenesis to selectively disrupt strands A-B and F-G to directly probe the role of backbone hydrogen bonds and β sheet stability on the mechanical stability of TNfn3.

The design of proline mutants of TNfn3

It is well known that proline substitution in a β sheet region can block the formation of backbone hydrogen bonds, cause a bulge in the β strand and disrupt hydrophobic packing.[35] These combined disruptive effects by proline substitution result in the discontinuity of the β strand and lead to the selective disruption of local β sheet structure. Such a relatively large structural perturbation is ideal for probing the mechanical unfolding pathway of proteins, as such a large perturbation of the protein structure generally cannot be compensated easily by the structural rearrangement of the protein. Therefore, the structural perturbation caused by proline substitution can be located easily and its effect on

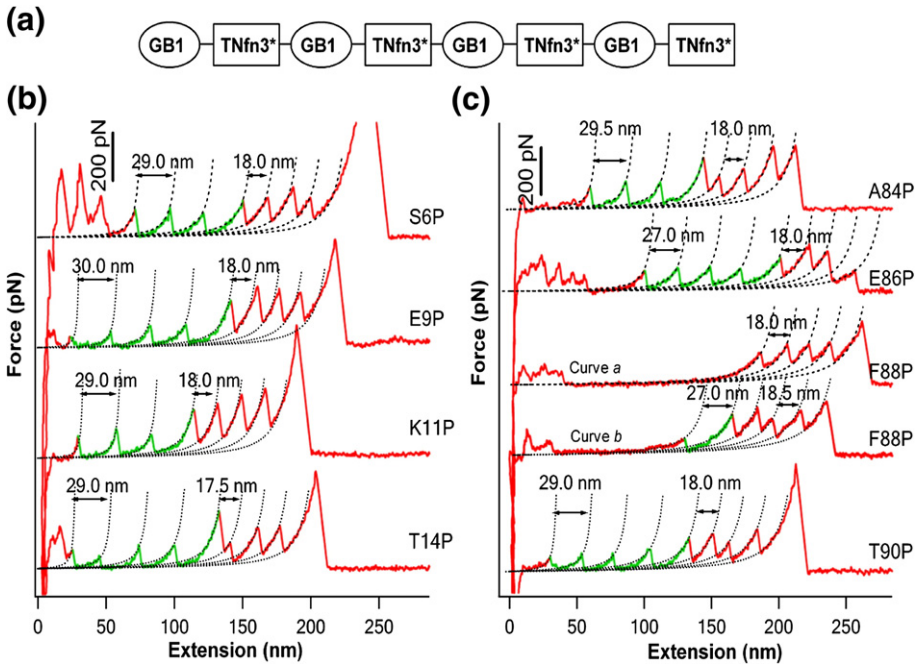

Fig. 7. Mechanical unfolding of proline mutants of TNfn3. (a) An illustration of the polyprotein chimera (GB1-TNfn3*)$_4$. TNfn3* denotes the proline mutant of TNfn3. (b and c) Typical force-extension curves of the polyprotein chimera (GB1-TNfn3*)$_4$ for each proline mutant. b, The force-extension curves of mutants in which residues in the A strand were substituted by proline. (c) The force-extension curves of mutants involving proline mutations in the G-strand. The mechanical unfolding events of the well characterized GB1 domains (colored red) occurred at ~180 pN with $\Delta L_c$of ~18 nm and served as fingerprints to identify single-molecule stretching events. Except for F88P, all the proline mutants show clear mechanical unfolding events with ΔL_c of ~29 nm (colored green), and their unfolding forces are in the range 90~120 pN, which is lower than that of wt TNfn3. Dotted lines are WLC fits to the experimental data. In contrast to other proline mutants, the unfolding of F88P does not result in clear mechanical unfolding events. Instead, long featureless spacer is typically observed before the mechanical unfolding events of GB1 domains (curve *a*), suggesting that F88P domains unfold at forces that are below our detection limit (~20 pN). Only a small fraction of F88P domains show clear unfolding events, as the one shown in curve *b*.

mechanical unfolding pathway can be identified readily.[36,37]

Since breaking the hydrogen bond between S6 and F23 occurs in intermediate state I1 and is the first event during the mechanical unfolding process of TNfn3, we engineered the S6P mutant to specifically probe the mechanical unfolding intermediate state I1. To probe the effect of backbone hydrogen bonds in hairpin A-B during the mechanical unfolding, we engineered mutants E9P, K11P and T14P. Similarly, we introduced proline mutations in strand G. To investigate the importance of the C-terminus on the mechanical stability of TNfn3, we engineered mutant T90P. We engineered proline mutants A84P, E86P and F88P to further probe the importance of strand G in the mechanical unfolding of TNfn3. The locations of these proline probes in TNfn3 are highlighted in Fig. 1b.

Phenotypic effects of proline mutations on strand A of TNfn3

To investigate the phenotypic effects of proline mutations in the region of strand A, we constructed the four proline mutants S6P, E9P, K11P, and T14P. In order to characterize unambiguously the mechanical unfolding of proline mutants using single-molecule AFM, we constructed heteropolyprotein $(GB1\text{-}ProlineMutant)_4$, in which TNfn3 mutants alternate with GB1 domains (Fig. 7a). In the polyprotein chimera, the well-characterized GB1 domains serve as fingerprints for identifying single-molecule stretching events and discerning the signatures of the mechanical unfolding of TNfn3 proline mutants.[38–40] The mechanical unfolding of GB1 is characterized by contour length increment ΔL_c of ~18 nm and unfolding force of ~180 pN at a pulling speed of ~400 nm/s.[15,23] Typical force-extension curves of the four heteropolyproteins involving the proline mutations in strand A are shown in Fig. 7b. Since TNfn3 alternates with GB1 domains in the heteropolyprotein, if we observed *N* unfolding events of GB1 in a given force-extension curve, we are certain that the force-extension curve must contain the signature of the stretching and unfolding of at least *N–1* TNfn3 mutant domains. Indeed, in the force-extension curves shown in Fig. 7b, the GB1 unfolding events with ΔL_c of ~18 nm (in red) are preceded by the low force unfolding events at ~100-120 pN (colored green). It is evident that

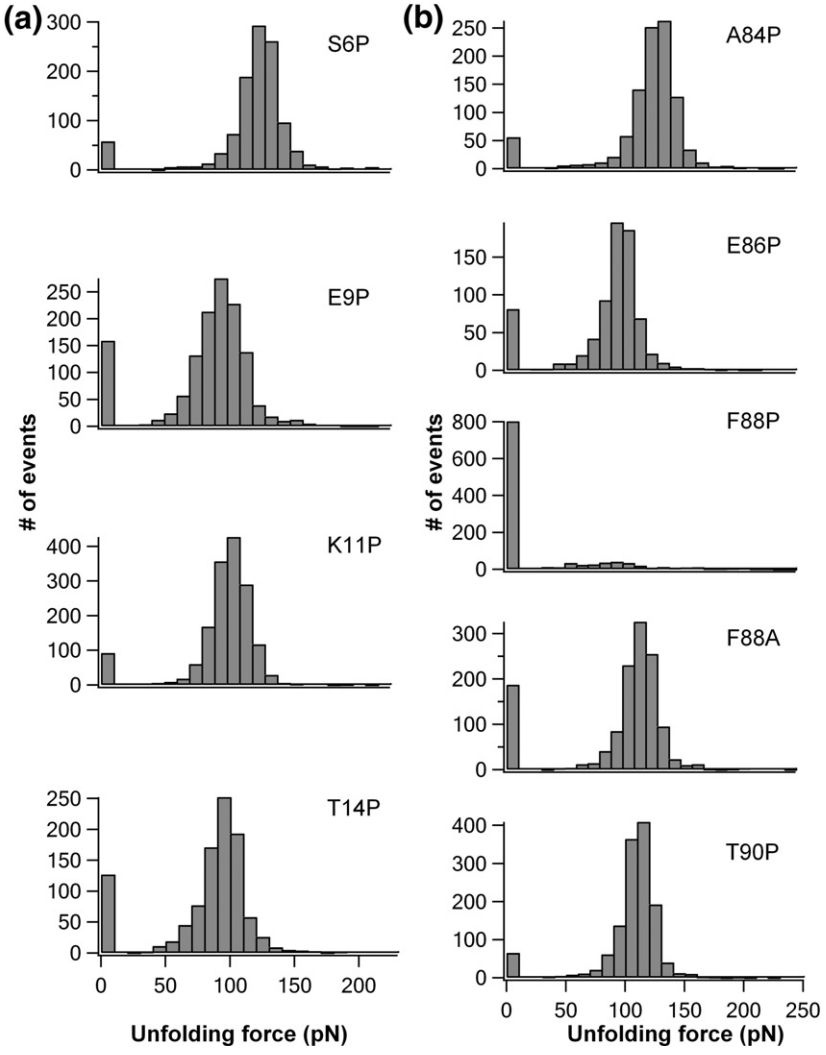

Fig. 8. The unfolding force histograms of TNfn3 proline mutants as well as F88A. All the proline mutants, except F88P, show well defined mechanical unfolding forces in the range 90 pN~120 pN. Most of F88P unfold at forces below 20 pN, and a small population of F88P show clear unfolding events at significant forces.

these low force unfolding events correspond to the mechanical unfolding of the proline mutants (S6P, E9P, K11P, and T14P) in their respective heteropolyprotein. Indeed, WLC fits to these unfolding events measure ΔL_c of ~28 nm, corroborating that these events indeed correspond to the complete mechanical unfolding of TNfn3 proline mutants. The average ΔL_c is 28.9 nm, 28.8 nm, 28.6 nm, and 28.8 nm for S6P, E9P, K11P, and T14P, respectively, which is identical with that of wild type TNfn3 within the resolution of our experiments.

It is important to note that in some force-extension curves, it appears that the unfolding events of some TNfn3 mutant domains are "missing". For example, in the curve shown in Fig. 1S (Supplementary Data), there are four GB1 unfolding events but only one E9P unfolding event. The absence of at least two additional E9P domains in this particular curve suggests that at least two additional E9P domains unfold at forces of less than 20 pN. This is a general feature for all the proline mutants investigated here, suggesting that a minute population of TNfn3 mutants are already "unfolded" before the stretching and there might be conformational heterogeneity in the native conformation of the proline mutants. Nonetheless, the origin of this observation will be investigated in detail elsewhere.

The unfolding force histograms of S6P, E9P, K11P, and T14P are shown in Fig. 8a. It is noticeable that the average unfolding force of S6P is ~127 pN, almost identical with that of wild type TNfn3, despite the significant local structure disruption. This result is in good agreement with the SMD simulation result. SMD simulations showed that the first six residues are detached from the rest of TNfn3 in the twist intermediate state I1. Since the first six residues are already disordered in the pseudo native state of TNfn3, disrupting the interactions in this region should not have any effect on the mechanical unfolding kinetics as well as mechanical unfolding force. Indeed, our results on S6P are consistent with this picture.

In contrast, mutations E9P, K11Pm and T14P reduce the average unfolding force of TNfn3 by an average of ~30 pN (see Table 1), reflecting the effect of structural disruption on the mechanical stability of TNfn3. Compared with the destabilization effect of more subtle alanine mutations in the same region, for example L8A,[18] proline mutations in the A strand causes slightly bigger destabilization effects. Considering that E9P, K11P, and T14P are mutations from polar or charged residues to proline, these results suggest that in addition to the hydrophobic interaction, the overall stability of the β hairpin plays important role in determining the mechanical stability of TNfn3.

Phenotypic effects of proline mutations on strand G of TNfn3: F88 is the Achilles heel of TNfn3

Using similar strategies, we constructed four proline mutants in the strand G region to weaken hairpin F-G to investigate their effects on the mechanical unfolding of TNfn3. Fig. 7c shows the typical force-extension curves of the heteropolyproteins containing proline mutants. The force-extension curves of mutant A84P, E86P, and T90P show clear mechanical unfolding events with ΔL_c of ~28 nm (colored green), corresponding to the complete mechanical unfolding of these three proline mutants. The unfolding force histograms for these three proline mutants are shown in Fig. 8b. It is evident that the average unfolding force for these three proline mutants is slightly lower than that for wild type TNfn3 by 10 – 20 pN, indicating that such a large perturbation to the β sheet structure of TNfn3 in strand G has a very mild effect on the mechanical stability of TNfn3.

Compared with A84P, E86P, and T90P, mutant F88P exhibits the strongest phenotypic effect in its mechanical unfolding behavior. In contrast to other proline mutants, the force-extension curves of $(GB1\text{-}F88P)_4$ are characterized by a long featureless spacer followed by the GB1 unfolding events (colored red). Clear unfolding events of ΔL_c of ~28 nm were absent from the vast majority of the force-extension curves of $(GB1\text{-}F88P)_4$ (Fig. 7c, force-extension curve a). Since GB1 alternates with F88P in the heteropolyprotein, the force-extension curves should contain roughly the same number of the stretching and unfolding events of the GB1 and F88P domains. Therefore, the long featureless spacer must correspond to the stretching and subsequent unfolding of F88P domains, suggesting that F88P domains unfold

Table 1. Unfolding force and kinetic parameters for the mechanical unfolding of TNfn3 and its mutants

Strand	Mutant	Unfolding force (pN)[a]	n	ΔL_c (nm)	α_0 (s^{-1})	Δx_u (nm)
	Wt	125±14	4198	29.0±0.8	1.5×10^{-4}	0.42
A	S6P	127±22	1053	28.9±1.2	1.0×10^{-4}	0.44
	E9P	96±21	1176	28.8±1.4	1.0×10^{-2}	0.41
	K11P	103±16	1499	28.6±1.2	5×10^{-3}	0.42
	T14P	94±18	879	28.8±1.4	8×10^{-3}	0.42
G	A84P	118±20	959	28.7±1.0	7×10^{-4}	0.44
	E86P	98±20	678	28.0±1.2	4×10^{-2}	0.44
	F88P	<20 pN	–	–	–	–
	F88A	115±23	1121	28.9±1.1	1.5×10^{-3}	0.42
	T90P	110±17	1283	28.0±1.4	7×10^{-4}	0.44

[a] The data are given as average±standard deviation, and n indicates the number of observations. All of these unfolding forces were measured at a pulling speed of 400 nm/s.

at forces < 20 pN. This result indicates that mutation F88P causes significant destabilization on TNfn3, such that the mechanical unfolding of TNfn3 occurs at very low forces. Occasionally, we observed that a small number of F88P domains can unfold at significant forces. For example, one F88P domain in the force-extension curve b shown in Fig. 7c unfolds at ~68 pN with ΔL_c of 27 nm, while the other three F88P domains in the same chimera polyprotein did not generate any unfolding event. Unfolding force histogram of such rare unfolding events for F88P (Fig. 8b) measures an average unfolding force of 104±49 pN (n = 320). Such high unfolding forces are unlikely, due to the fluctuations in unfolding forces from those that occurred at <20 pN. Instead, the difference in unfolding forces suggests that there are two distinct populations of F88P that have different mechanical stability: the majority of F88P is mechanically weak and unfolds at forces <20 pN, and a small percentage of F88P can unfold at forces of ~100 pN. This observation is suggestive of heterogeneity of the native state of F88P, which deserves further experimental work in the future to fully characterize the origin of such a conformational heterogeneity of the native states of F88P.

These results indicate that the phenotypic effect of proline substitution in strand G is context-dependent: at the N-terminal end of strand G, proline substitution had little effect on the mechanical stability. Proline substitutions at the C-terminal end of strand G, with the exception of F88P, have only a mild effect on the mechanical stability of TNfn3. F88P has the strongest destabilization effect on the mechanical stability of TNfn3 and seems to be the Achilles heel for the mechanical unfolding of TNfn3. These observations suggest that the C-terminal end of TNfn3 can play important roles in the mechanical unfolding of TNfn3. It is of note that similar context-dependent phenotypic effect have been observed in FNfn10 domain.[37] Comparison between the two FnIII domains is discussed in Discussion.

It is of note that the F88P substitution blocks the formation of backbone hydrogen bond between F88 and Y68 and significantly disrupts the hydrophobic packing interactions of TNfn3 mediated by hydrophobic residue F88. To distinguish the contribution of a hydrogen bond from that of hydrophobic interactions to mechanical stability, we constructed an alanine mutant F88A, which affects only hydrophobic interactions and not the backbone hydrogen bond. Single-molecule AFM experiments on $(GB1\text{-}F88A)_4$ shows that the average unfolding force of F88A is ~115 pN, indicating that mutant F88A is much more stable than F88P. This result suggests that the backbone hydrogen bond between residues 88 and 68 is critical for the mechanical and thermodynamic stability of TNfn3. Proline mutation F88P will block the original backbone hydrogen bond and introduce a bugle at the C-terminal end of G-strand, which is likely to open the flood-gate for water molecules to enter and solvate the hydrophobic core of TNfn3, leading to the significant destabilization of TNfn3.

Proline substitutions do not affect the mechanical unfolding distance of TNfn3

To investigate how the proline substitutions affect the mechanical unfolding distance, we carried out pulling experiments on TNfn3 proline mutants at different pulling speeds. Similar to that of wild type TNfn3, the average unfolding force for TNfn3 proline mutants also exhibits weak dependence on the pulling speeds at which the polyprotein is being stretched and unraveled (Fig. 9). The slope of the pulling speed dependence of the unfolding force for TNfn3 proline mutants is similar to that of the wild type FNfn3, suggesting that proline mutations in strands A and G do not have a significant effect on the mechanical unfolding distance Δx_u of TNfn3. Therefore, the phenotypic effects of proline mutants result from lowering the mechanical unfolding energy barrier. Using Monte Carlo simulations, we found that using the parameters given in Table 1 for α_0 and Δx_u can adequately describe the unfolding force histogram and the pulling speed dependence of the unfolding forces for TNfn3 proline mutants. These results corroborate that the phenotypic effect observed on TNfn3 proline mutants are due to the reduced mechanical unfolding energy barrier, rather than the change in mechanical unfolding distance.

Discussion

Mechanical unfolding of TNfn3: an FnIII domain of a robust mechanical design

Our single-molecule AFM experiments revealed that TNfn3 is a mechanically stable protein, which unfolds at an average force of ~130 pN at a pulling speed of 400 nm/s. The mechanical unfolding of TNfn3 is an apparent two-state process without any observable intermediate state populating along its mechanical unfolding pathway. Using proline mutagenesis, we demonstrated that disrupting the A-B or F-G β hairpin by proline substitution can lead to reduction of the mechanical stability of TNfn3, and the amplitude of reduction of the mechanical stability depends on the location of the proline substitution. Recent work downplayed the role of backbone hydrogen bonds in the mechanical unfolding of FnIII domains and suggested that the hydrophobic packing is critical for the mechanical stability of TNfn3,[18] and this hypothesis was the basis for swapping the hydrophobic cores between different FnIII domains in order to increase the mechanical stability of engineered FnIII domains.[41] Our results with proline mutants clearly demonstrate that hydrophobic packing is not the only important factor in determining the mechanical stability of TNfn3. Backbone hydrogen bonds as well as the structural integrity of force-bearing β hairpins are important structural para-

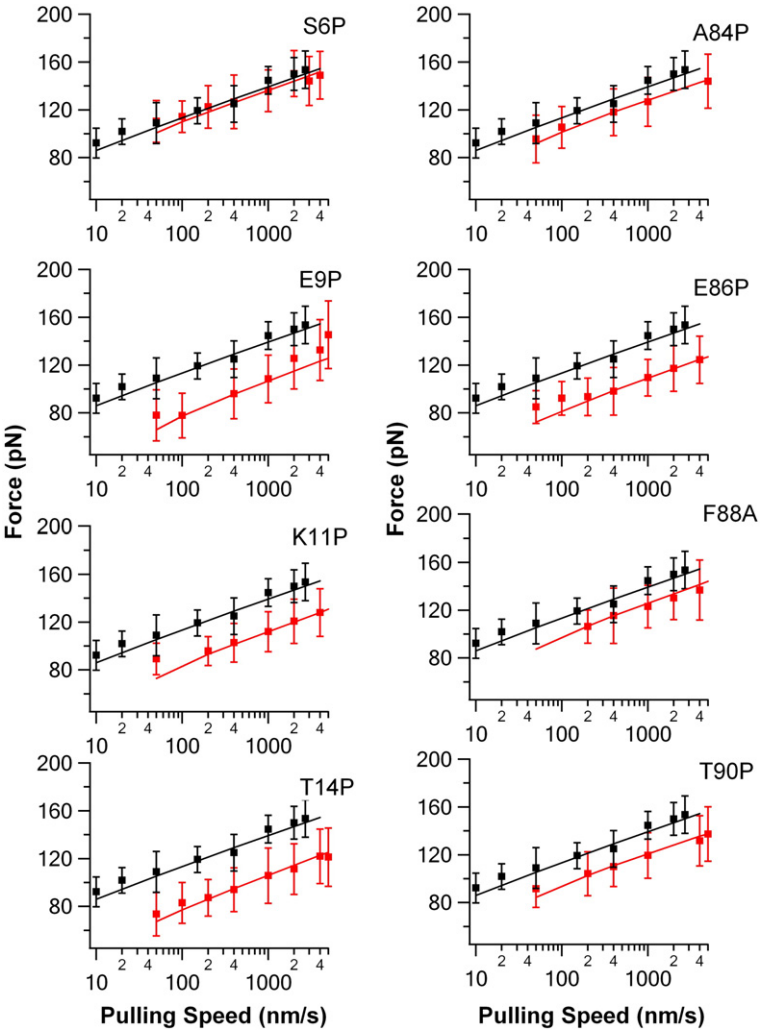

Fig. 9. Pulling speed dependence of the mechanical unfolding force of TNfn3 mutants. For comparison, the data for wt TNfn3 is also shown (black symbols). It is evident that the slope of the pulling speed dependence of the mechanical unfolding force of TNfn3 mutants is similar to that for wt TNfn3, indicating that the unfolding distance Δx_u of TNfn3 mutants is similar to that of wt TNfn3. The continuous lines are Monte Carlo fits using the parameters given in Table 1.

meters in defining the mechanical stability of TNfn3. Hence, the mechanical stability of TNfn3 carries both global and local structural attributes, necessitating the synergetic consideration of both hydrophobic core packing and backbone hydrogen bonds in the force-bearing β hairpins when designing novel FnIII domains with tailored mechanical properties.

Although proline mutagenesis has helped to reveal the importance of both hydrophobic core packing and backbone hydrogen bonds in β hairpins in determining the mechanical stability of TNfn3, it is not possible to estimate the relative contributions of the two types of interactions to the mechanical stability of TNfn3. This is because proline mutation deletes the backbone hydrogen bond, affects hydrophobic interaction, and introduces local structural disruption. Amide to ester mutation would be an ideal approach to evaluate quantitatively the contribution of backbone hydrogen bonds to mechanical stability.[42–44] Despite the current challenge in applying the amide-to-ester mutagenesis technique in protein mechanics, further developments of amide-to-ester mutagenesis technique will make it feasible to address this question in the future.

Our proline mutagenesis results also reveal the robustness of the mechanical design of TNfn3, which has been underestimated by conservative side chain deletion approaches in some AFM studies. Except for residue F88, introducing disruptive proline mutation into the β sheets of TNfn3 does not result in catastrophic effects on its mechanical stability. Instead, a mild reduction in mechanical stability by 10~30 pN was observed for these proline mutants, an effect comparable to that resulting from conservative side chain deletion mutations in the same region of TNfn3.[18] Such mild effects are in contrast with the significant destabilization effect of proline substitution on the native state observed for the tenth FnIII domains of fibronectin.[37] These results highlight the global nature of the mechanical resistance of TNfn3, which is in sharp contrast to the rather local attributes of the mechanical resistance observed for other well-studied elastomeric proteins, such as I27.[17] Since the energy barrier for the transition from twist state I1 to aligned state I2 is likely to be the rate-limiting step for the mechanical unfolding of TNfn3, proline substitution in AB and FG hairpins in TNfn3 not only disrupts β hairpins, but also facilitates the rotation and alignment of the two β

sheets. However, the exact molecular mechanism underlying such mechanical weakening effect on TNfn3 still needs to be elucidated. Moreover, the observation of a single weak point in TNfn3, the Achilles heel F88, is quite surprising. It seems that F88 is the key to protecting the hydrophobic core *via* the combination of backbone hydrogen bonds and hydrophobic interactions. Since the mechano-phenotype is context-dependent, the dramatic phenotype observed for F88P cannot be extended directly to its neighboring residues 87 and 89. Since residues 87 and 89 do not involve the formation of backbone hydrogen bonds, it will be of interest to examine whether proline mutation on such residues leads to a similar dramatic destabilization effect. Furthermore, since all the FnIII domains of tenascin-C share similar mechanical stability,[13,15,28] it will be interesting to check whether the features observed here for the mechanical design of TNfn3 will also apply to other FnIII domains of tenascin-C.

Comparison of the mechanical unfolding of TNfn3 *versus* FNfn10: similar structure but different unfolding behaviors

The mechanical unfolding of the tenth FnIII domain of fibronectin (FNfn10) has been characterized in detail using single-molecule AFM and SMD simulations.[32,33,37,45,46] Despite the highly homologous structures of FNfn10 and TNfn3, their mechanical unfolding behavior shows interesting commonality as well as differences.

SMD simulations show strikingly similar signatures in the mechanical unfolding of these two domains.[32,33,45,46] FNfn10 and TNfn3 follow very similar sequences of molecular events: first, the predetachment of the first few residues of the N-terminal end of strand A leads to a slightly twisted state I1; then the two β sheets rotate and align with each other to enter the aligned state I2; after that, the protein unfolds in two distinct pathways; e. g. strand A separates first and strands A-G separate simultaneously, and reach the partially unfolded intermediate state I3 in the "A-strand separates first" pathway; finally the domain unravels completely. The transition from the twisted state I1 to the aligned state I2 has been indicated as the main unfolding barrier.

Despite the similarity in the unfolding sequences, a significant difference sets the two FnIII domains apart: the presence of stable intermediate state I3 predicted in SMD simulations for FNfn10 was observed experimentally in single-molecule AFM experiments,[37] while the similar intermediate state I3 was not observed for TNfn3 in AFM experiments.

The similarity and difference between the two FnIII domains reside in their response to proline mutation. The phenotypic effect of proline mutations in strand A is catastrophic for FNfn10, as proline mutations in strand A destabilize the protein so dramatically that the transition from the aligned state I2 to I3 is abolished;[37] and the FNfn10 proline mutant unfolds directly from the intermediate state I3. In contrast, the phenotypic effect of proline mutations in strand A in TNfn3 is much milder and the destabilization facilitates the transition from I1 to I2.

Compared with the difference in phenotypic effect in strand A, proline mutations in strand G have strikingly similar effects for FNfn10 and TNfn3. Residue 88 seems to be the weak point for both proteins, as proline substitution at this residue leads to significant destabilization of both proteins: for FNfn10, destabilization leads to the elimination of the transition from I2 to I3, and FNfn10 unfolds directly from intermediate state I3;[37] for TNfn3, since the intermediate state I3 is much less stable, destabilization caused by substitution of residue 88 with proline leads to the complete unraveling of TNfn3 at low forces. Proline substitution at the N-terminal end of strand G shows a very similar minute effect on the mechanical unfolding of both proteins.

Understanding the relationship between sequence variation and mechanical unfolding features of these two proteins will be critical for future efforts to engineer FnIII domains of tailored mechanical properties. A recent study has made an encouraging first step toward such purposes.[41]

Single-molecule AFM *versus* SMD: similarities and discrepancies

Combining protein engineering, single-molecule AFM and SMD simulation techniques, we have characterized the mechanical unfolding pathways of TNfn3 on two vastly different time-scales. Comparing the single-molecule AFM results with SMD predictions, some of the SMD predictions are verified by the single-molecule AFM results, but there are some important discrepancies between the two, and the SMD simulation results cannot fully explain the experimental findings in single-molecule AFM experiments.

Similar to previous SMD simulations,[18,33] our SMD simulations of the mechanical unfolding of TNfn3 using an explicit water model predicted that the first step of the mechanical unfolding process of TNfn3 is the disruption and detachment of the N-terminus (residues 1–6), followed by the alignment of the two β-sheets. Afterwards, the unfolding of the aligned β-sheets is initiated by the rupture of AB β hairpin followed by a mechanical unfolding intermediate state I3 or by the simultaneous separation of strands A-G from TNfn3. The unraveling of the I3 state will lead to a fully unfolded TNfn3. The three on-pathway unfolding intermediate states observed in SMD simulation have different stabilities. Analysis of both the unfolding force (from the constant velocity SMD) and dwell time (from the constant force SMD) of the three intermediate states showed that I1 is the most stable and I2 is the least stable.

Compared with the SMD simulation results, the single-molecule AFM results confirm that the disruption and detachment of the N-terminus of TNfn3 is likely to be the very first step in the mechanical

unfolding of TNfn3, as the S6P mutation does not have any effect on the mechanical unfolding of TNfn3. This agreement confirms that the mechanical unfolding of TNfn3 observed in single-molecule AFM is not the unraveling of TNfn3 from its native state, but from a force-induced pseudo ground state, in which the N-terminal end of strand A is already detached. In addition, SMD simulations predict that the transition from twisted intermediate state I1 to the aligned state I2 is the rate-limiting step for the unfolding of TNfn3. This transition involves the rotation and alignment of the two β sheets, and is accompanied by the extension of the distance between the N- and C-termini by ~10 Å. Therefore, TNfn3 can be deformed over a longer distance before it unfolds. This description provides a plausible molecular level explanation for the observed longer unfolding distance Δx_u for TNfn3. In contrast, unraveling I27-like elastomeric proteins requires simultaneous rupture of multiple hydrogen bonds holding the two force-bearing β strands together. Therefore, I27-like proteins can be deformed only over a shorter distance before they unfold, giving rise to shorter unfolding distances.

Despite the agreements between SMD simulation predictions with single-molecule AFM experiments, some discrepancies exist between the two. In contrast to the SMD prediction of the existence of intermediate I3, the single-molecule AFM results indicated that the mechanical unfolding of TNfn3 is an apparent two-state process, and no intermediate state is observed in single-molecule AFM experiments. Although intermediate I3 was predicted to be mechanically stable in both constant force and constant velocity SMD simulations, intermediate state I3 does not populate on the time-scale of single-molecule AFM experiments. The origin of such a discrepancy remains to be explored. This situation is in sharp contrast to that for FNfn10, for which a very good agreement was reached for the similar intermediate state I3 between single-molecule AFM experiments and SMD simulations.[32,33,37,46] Despite such limitations, it is clear that SMD has offered valuable insights into the molecular mechanism of mechanical unfolding of proteins. Improvements in SMD methodology will continue to further our understanding of the mechanical unfolding and mechanical design of elastomeric proteins at an unprecedented detail.

In summary, we have characterized the mechanical unfolding of TNfn3 using a combination of single-molecule AFM, SMD simulations, and proline mutagenesis techniques. Our results have revealed the robust mechanical design of TNfn3 that confers the resistance of the protein to disruptive mutations such as proline substitution. Moreover, both local and global structural features are important for determining the mechanical resistance of TNfn3. It has become evident that the mechanical unfolding of TNfn3 is much more complex than other typical elastomeric proteins, such as I27,[17,47] due to the structural flexibility and deformability of TNfn3. Therefore, the TNfn3 and FnIII domains, in general, present further challenges for protein engineers to enhance the mechanical stability of FnIII domain in a systematic and rational fashion.

Materials and Methods

Protein engineering

The DNA sequence coding TNfn3, flanked with a 5′ BamHI restriction site and 3′ BglII, HindIII restriction sites, was amplified by the polymerase chain reaction (PCR) from the plasmid TNfnALL encoding all of the 15 FnIII domains. The plasmid TNfnALL was a generous gift from Professor Harold Erickson (Duke Univeristy). Polyprotein $(TNfn3)_8$ was constructed using a consecutive DNA concatamerization method based on the identity of the sticky ends generated by the BamHI and BglII restriction enzymes.[17]

To facilitate the identification of the mechanical unfolding signatures of the proline mutants of TNfn3 using AFM, we constructed heteropolyproteins consisting of alternating GB1 domains and proline mutant TNfn3, where the well-characterized GB1 domains serve as internal fingerprints for identifying single-molecule stretching events. Since GB1 gene carries a 5′ BamHI restriction site and 3′ BglII, KpnI restriction sites, we constructed a new version of TNfn3, which carries a 5′ BamHI restriction site and 3′ BglII, KpnI restriction sites, to facilitate the construction of the heteropolyprotein. Proline mutants were constructed using standard site-directed mutagenesis methods. Genes encoding heteropolyproteins were constructed using protocols similar to those used for constructing polyprotein $(TNfn3)_8$ based on the identity of the sticky ends generated by the BamHI and BglII restriction enzymes.

Polyproteins were over-expressed in strain DH5α and purified from supernatant using Ni^{2+}-affinity chromatography. The polyproteins were kept at 4 °C at a concentration of ~200 μg/mL in PBS.

Single-molecule AFM experiments

Single-molecule AFM experiments were done with a custom-built atomic force microscope, which was constructed as described.[48] All the force-extension measurements were carried out in PBS. In a typical experiment, the polyprotein sample (1 μL) was deposited onto a clean glass coverslip covered by PBS (50 μL), resulting in a 10–20 nm thick layer of protein. The thickness of this protein layer depends upon the amount of protein deposited onto the glass coverslip; i.e., the more protein, the thicker the layer. The thickness of the protein layer may contribute to the apparent contour length of the polyprotein, sometimes resulting in an apparent contour length that is greater than the theoretical contour length of the polyprotein.[38] The spring constant of each individual cantilever (Si_3N_4 cantilevers from Vecco, with a typical spring constant of 40 pN/nm) was calibrated in solution using the equipartition theorem before and after each experiment.

SMD simulation

The crystal structure of TNfn3 (PDB accession code 1TEN) was used as the starting conformation of TNfn3 for simulated equilibration. The protein was solvated in a water box (length 107 Å, width 62 Å, height 53 Å) with a

TIP3P water model.[49] The whole protein–water system contains 33,194 atoms. SMD simulations of the mechanical unfolding of TNfn3 was done with the program NAMD 2.6[50] and with the CHARMM22 force field,[51] as described.[39,52] The initial structure of TNfn3 was equilibrated for 1.5 ns at 300 K. Compared to the crystal structure, the equilibrated structures at 1 ns and 1.5 ns have a backbone RMSD of 0.93 Å and 0.98 Å, respectively, and were used as the starting conformation for the constant force and constant velocity SMD simulations. During the SMD simulations, the N-terminal C^{α} atom was fixed and the C-terminal C^{α} atom was pulled. The pulling speed was 0.05 Å/ps in constant velocity SMD simulations and a pulling force of 500 pN was used in the constant force SMD simulations. The simulation time was 57 ns in total. System setup, structural analysis and calculation of hydrogen bond energies were performed using VMD1.86[53] and hydrogen bond energies were calculated using custom written script kindly provided by Hui Lu and Morten Kallberg.

Monte Carlo simulation

Monte Carlo simulations of the stretching and unfolding of the polyproteins were done as described.[20] The unfolding rate constant at zero force α_0 and the distance of the native state to the transition state Δx_u along the reaction coordinate of the mechanical unfolding reaction were estimated using Monte Carlo simulation procedures in a trial-and-error fashion. Such a procedure is necessary due to the lack of analytical solutions to the unfolding force distribution obtained from force-extension measurement for polyproteins. The accuracy of the fitting parameters is exemplified in Supplementary Data Fig. 2S, which plots the simulated unfolding force histogram and the pulling speed dependence of the unfolding forces using different sets of unfolding rate constant α_0 and unfolding distance Δx_u. Typically, α_0 is accurate within a factor of 3 and the unfolding distance Δx_u is accurate within 0.05 nm.

Acknowledgements

We thank Prof. Hui Lu and Dr. Gang Feng for their generous help in SMD simulation, and Dr. Mu Gao for stimulating discussion. We also thank Prof. Hui Lu and Morten Kallberg for their generous script and help in hydrogen bond energy calculations. This work is supported by Canadian Institutes of Health Research (CIHR) Operating grant MOP-81225, Michael Smith Foundation for Health Research, University of British Columbia Health Research Resources Office and Canada Research Chairs Program. H. L. is a Michael Smith Foundation for Health Research Career Investigator. Q. P. is supported by a Pacific Century Graduate Scholarship from the Province of British Columbia.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2009.01.019

References

1. Alenghat, F. J. & Ingber, D. E. (2002). Mechanotransduction: all signals point to cytoskeleton, matrix, and integrins. *Sci. STKE*, **2002**, PE6.
2. Kjaer, M. (2004). Role of extracellular matrix in adaptation of tendon and skeletal muscle to mechanical loading. *Physiol. Rev.* **84**, 649–698.
3. Chen, C. S., Tan, J. & Tien, J. (2004). Mechanotransduction at cell-matrix and cell-cell contacts. *Annu. Rev. Biomed. Eng.* **6**, 275–302.
4. Jones, F. S. & Jones, P. L. (2000). The tenascin family of ECM glycoproteins: structure, function, and regulation during embryonic development and tissue remodeling. *Dev. Dyn.* **218**, 235–259.
5. Erickson, H. P. (1994). Reversible unfolding of fibronectin type III and immunoglobulin domains provides the structural basis for stretch and elasticity of titin and fibronectin. *Proc. Natl Acad. Sci. USA*, **91**, 10114–10118.
6. Chiquet-Ehrismann, R. (1995). Tenascins, a growing family of extracellular matrix proteins. *Experientia*, **51**, 853–862.
7. Jones, P. L. & Jones, F. S. (2000). Tenascin-C in development and disease: gene regulation and cell function. *Matrix Biol.* **19**, 581–596.
8. Hsia, H. C. & Schwarzbauer, J. E. (2005). Meet the tenascins: multifunctional and mysterious. *J. Biol. Chem.* **280**, 26641–26644.
9. Kannus, P., Jozsa, L., Jarvinen, T. A., Jarvinen, T. L., Kvist, M., Natri, A. & Jarvinen, M. (1998). Location and distribution of non-collagenous matrix proteins in musculoskeletal tissues of rat. *Histochem. J.* **30**, 799–810.
10. Jarvinen, T. A., Jozsa, L., Kannus, P., Jarvinen, T. L., Hurme, T., Kvist, M. *et al.* (2003). Mechanical loading regulates the expression of tenascin-C in the myotendinous junction and tendon but does not induce de novo synthesis in the skeletal muscle. *J. Cell Sci.* **116**, 857–866.
11. Jarvinen, T. A., Kannus, P., Jarvinen, T. L., Jozsa, L., Kalimo, H. & Jarvinen, M. (2000). Tenascin-C in the pathobiology and healing process of musculoskeletal tissue injury. *Scand. J. Med. Sci. Sports*, **10**, 376–382.
12. Chiquet-Ehrismann, R. & Chiquet, M. (2003). Tenascins: regulation and putative functions during pathological stress. *J. Pathol.* **200**, 488–499.
13. Oberhauser, A. F., Marszalek, P. E., Erickson, H. P. & Fernandez, J. M. (1998). The molecular elasticity of the extracellular matrix protein tenascin. *Nature*, **393**, 181–185.
14. Rief, M., Gautel, M., Schemmel, A. & Gaub, H. E. (1998). The mechanical stability of immunoglobulin and fibronectin III domains in the muscle protein titin measured by atomic force microscopy. *Biophys. J.* **75**, 3008–3014.
15. Cao, Y. & Li, H. (2006). Single molecule force spectroscopy reveals a weakly populated microstate of the FnIII domains of tenascin. *J. Mol. Biol.* **361**, 372–381.
16. Carrion-Vazquez, M., Oberhauser, A. F., Fisher, T. E., Marszalek, P. E., Li, H. & Fernandez, J. M. (2000). Mechanical design of proteins studied by single-molecule force spectroscopy and protein engineering. *Prog. Biophys. Mol. Biol.* **74**, 63–91.
17. 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.

18. Ng, S. P., Rounsevell, R. W., Steward, A., Geierhaas, C. D., Williams, P. M., Paci, E. & Clarke, J. (2005). Mechanical unfolding of TNfn3: the unfolding pathway of a fnIII domain probed by protein engineering, AFM and MD simulation. *J. Mol. Biol.* **350**, 776–789.
19. Leahy, D. J., Hendrickson, W. A., Aukhil, I. & Erickson, H. P. (1992). Structure of a fibronectin type-Iii domain from tenascin phased by MAD analysis of the selenomethionyl protein. *Science*, **258**, 987–991.
20. Li, H. (2008). "Mechanical engineering" of elastomeric proteins: toward designing new protein building blocks for biomaterials. *Adv. Funct. Mater.* **18**, 2643–2657.
21. Schwaiger, I., Kardinal, A., Schleicher, M., Noegel, A. A. & Rief, M. (2004). A mechanical unfolding intermediate in an actin-crosslinking protein. *Nature Struct. Mol. Biol.* **11**, 81–85.
22. Carrion-Vazquez, M., Li, H., Lu, H., Marszalek, P. E., Oberhauser, A. F. & Fernandez, J. M. (2003). The mechanical stability of ubiquitin is linkage dependent. *Nature Struct. Biol.* **10**, 738–743.
23. Cao, Y., Lam, C., Wang, M. & Li, H. (2006). Nonmechanical protein can have significant mechanical stability. *Angew Chem. Int. Ed. Engl.* **45**, 642–645.
24. Oberhauser, A. F., Badilla-Fernandez, C., Carrion-Vazquez, M. & Fernandez, J. M. (2002). The mechanical hierarchies of fibronectin observed with single-molecule AFM. *J. Mol. Biol.* **319**, 433–447.
25. Schoenauer, R., Bertoncini, P., Machaidze, G., Aebi, U., Perriard, J. C., Hegner, M. & Agarkova, I. (2005). Myomesin is a molecular spring with adaptable elasticity. *J. Mol. Biol.* **349**, 367–379.
26. Brockwell, D. J., Beddard, G. S., Paci, E., West, D. K., Olmsted, P. D., Smith, D. A. & Radford, S. E. (2005). Mechanically unfolding the small, topologically simple protein L. *Biophys. J.* **89**, 506–519.
27. Marko, J. F. & Siggia, E. D. (1995). Stretching DNA. *Macromolecules*, **28**, 8759–8770.
28. Wang, M. J., Cao, Y. & Li, H. B. (2006). The unfolding and folding dynamics of TNfnALL probed by single molecule force-ramp spectroscopy. *Polymer*, **47**, 2548–2554.
29. Clarke, J., Hamill, S. J. & Johnson, C. M. (1997). Folding and stability of a fibronectin type III domain of human tenascin. *J. Mol. Biol.* **270**, 771–778.
30. 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.
31. Li, H. B., Oberhauser, A. F., Fowler, S. B., Clarke, J. & Fernandez, J. M. (2000). Atomic force microscopy reveals the mechanical design of a modular protein. *Proc. Natl Acad. Sci. USA*, **97**, 6527–6531.
32. Gao, M., Craig, D., Vogel, V. & Schulten, K. (2002). Identifying unfolding intermediates of FN-III(10) by steered molecular dynamics. *J. Mol. Biol.* **323**, 939–950.
33. Craig, D., Gao, M., Schulten, K. & Vogel, V. (2004). Tuning the mechanical stability of fibronectin type III modules through sequence variations. *Structure*, **12**, 21–30.
34. Billings, K. S., Best, R. B., Rutherford, T. J. & Clarke, J. (2008). Crosstalk between the protein surface and hydrophobic core in a core-swapped fibronectin type III domain. *J. Mol. Biol.* **375**, 560–571.
35. Wood, S. J., Wetzel, R., Martin, J. D. & Hurle, M. R. (1995). Prolines and amyloidogenicity in fragments of the Alzheimer's peptide beta/A4. *Biochemistry*, **34**, 724–730.
36. Li, H., Carrion-Vazquez, M., Oberhauser, A. F., Marszalek, P. E. & Fernandez, J. M. (2000). Point mutations alter the mechanical stability of immunoglobulin modules. *Nature Struct. Biol.* **7**, 1117–11120.
37. Li, L., Huang, H. H., Badilla, C. L. & Fernandez, J. M. (2005). Mechanical unfolding intermediates observed by single-molecule force spectroscopy in a fibronectin type III module. *J. Mol. Biol.* **345**, 817–826.
38. Cao, Y. & Li, H. (2007). Polyprotein of GB1 is an ideal artificial elastomeric protein. *Nature Mater.* **6**, 109–114.
39. Sharma, D., Perisic, O., Peng, Q., Cao, Y., Lam, C., Lu, H. & Li, H. (2007). Single-molecule force spectroscopy reveals a mechanically stable protein fold and the rational tuning of its mechanical stability. *Proc. Natl Acad. Sci. USA*, **104**, 9278–9283.
40. Peng, Q. & Li, H. (2008). Atomic force microscopy reveals parallel mechanical unfolding pathways of T4 lysozyme: evidence for a kinetic partitioning mechanism. *Proc. Natl Acad. Sci. USA*, **105**, 1885–1890.
41. Ng, S. P., Billings, K. S., Ohashi, T., Allen, M. D., Best, R. B., Randles, L. G. *et al.* (2007). Designing an extracellular matrix protein with enhanced mechanical stability. *Proc. Natl Acad. Sci. USA*, **104**, 9633–9637.
42. Blankenship, J. W., Balambika, R. & Dawson, P. E. (2002). Probing backbone hydrogen bonds in the hydrophobic core of GCN4. *Biochemistry*, **41**, 15676–15684.
43. Deechongkit, S., Dawson, P. E. & Kelly, J. W. (2004). Toward assessing the position-dependent contributions of backbone hydrogen bonding to beta-sheet folding thermodynamics employing amide-to-ester perturbations. *J. Am. Chem. Soc.* **126**, 16762–16771.
44. Deechongkit, S., Nguyen, H., Powers, E. T., Dawson, P. E., Gruebele, M. & Kelly, J. W. (2004). Context-dependent contributions of backbone hydrogen bonding to beta-sheet folding energetics. *Nature*, **430**, 101–105.
45. Craig, D., Krammer, A., Schulten, K. & Vogel, V. (2001). Comparison of the early stages of forced unfolding for fibronectin type III modules. *Proc. Natl Acad. Sci. USA*, **98**, 5590–5595.
46. Paci, E. & Karplus, M. (1999). Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations. *J. Mol. Biol.* **288**, 441–459.
47. Lu, H. & Schulten, K. (1999). Steered molecular dynamics simulations of force-induced protein domain unfolding. *Proteins: Struct. Funct. Genet.* **35**, 453–463.
48. Fernandez, J. M. & Li, H. (2004). Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science*, **303**, 1674–1678.
49. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.
50. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E. *et al.* (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802.
51. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J. *et al.* (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, **102**, 3586–3616.
52. Lee, E. H., Hsin, J., Mayans, O. & Schulten, K. (2007). Secondary and tertiary structure elasticity of titin Z1Z2 and a titin chain model. *Biophys. J.* **93**, 1719–1735.
53. Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. 14, 33-8, 27–28.