doi:10.1016/j.jmb.2008.04.018 J. Mol. Biol. (2008) 379, 871–880

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Configurational Entropy Modulates the Mechanical Stability of Protein GB1

Hongbin Li*, Hui-Chuan Wang, Yi Cao, Deepak Sharma and Meijia Wang

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

Received 18 December 2007; received in revised form 3 April 2008; accepted 8 April 2008
Available online 11 April 2008

Configurational entropy plays important roles in defining the thermodynamic stability as well as the folding/unfolding kinetics of proteins. Here we combine single-molecule atomic force microscopy and protein engineering techniques to directly examine the role of configurational entropy in the mechanical unfolding kinetics and mechanical stability of proteins. We used a small protein, GB1, as a model system and constructed four mutants that elongate loop 2 of GB1 by 2, 5, 24 and 46 flexible residues, respectively. These loop elongation mutants fold properly as determined by far-UV circular dichroism spectroscopy, suggesting that loop 2 is well tolerant of loop insertions without affecting GB1's native structure. Our single-molecule atomic force microscopy results reveal that loop elongation decreases the mechanical stability of GB1 and accelerates the mechanical unfolding kinetics. These results can be explained by the loss of configurational entropy upon closing an unstructured flexible loop using classical polymer theory, highlighting the important role of loop regions in the mechanical unfolding of proteins. This study not only demonstrates a general approach to investigating the structural deformation of the loop regions in mechanical unfolding transition state, but also provides the foundation to use configurational entropy as an effective means to modulate the mechanical stability of proteins, which is of critical importance towards engineering artificial elastomeric proteins with tailored nanomechanical properties.

Keywords: configurational entropy; single molecule atomic force microscopy; mechanical unfolding; mechanical stability; force spectroscopy

Edited by K. Kuwajima

Introduction

Loop regions play important roles in protein structure and functions: they not only serve as connecting units in proteins to connect secondary structural elements, α helices and β strands, but can also participate in forming binding sites and enzyme active sites.[1] Moreover, the length of flexible loops can also have significant influence on the thermodynamic stability and folding/unfolding kinetics of proteins. Such influences are directly associated with the loss of configurational entropy upon closing an unstructured loop region in proteins.[2] Hence, the effect of configurational entropy on the thermodynamic stability as well as chemical folding/unfolding kinetics of proteins has been studied in detail, using ensemble methods, by varying the length of flexible loops. It was found that increasing the loop length of a protein can reduce its thermodynamic stability, and often slows down folding reactions while accelerating the unfolding reactions.[3–7]

The recent development of single-molecule force spectroscopy techniques has opened opportunities to investigate the folding and unfolding dynamics of proteins along a well-defined reaction coordinate defined by the stretching force.[8–12] Protein engineering has been used extensively in combination with single-molecule atomic force microscopy (AFM) technique to investigate the molecular determinants of the mechanical stability of proteins.[9,13] Despite detailed studies on the effect of loop length (or configurational entropy) on the chemical folding and unfolding of proteins, the role of loop regions in protein mechanics and mechanical folding/unfolding dynamics has remained largely unexplored. Although loop insertion strategies have been used in single-protein mechanics studies,[14,15] their use has been largely focused on locating the structural motif

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

Abbreviations used: AFM, atomic force microscopy; wt, wild type.

that provides mechanical resistance or mapping the location of unfolding intermediate state. Using a small protein, GB1, the B1 IgG binding domain of protein G from *Streptococcus*, as a model system, here we utilized single-molecule AFM to directly probe the effect of loop length on the mechanical stability and mechanical unfolding kinetics of GB1. GB1 is a small protein of 56 amino acid residues[16] and its mechanical unfolding and folding dynamics have been investigated in detail using single-molecule AFM.[17–19] Due to its excellent mechanical properties and ease of manipulation, GB1 has become a new paradigm for single-molecule force spectroscopy studies on proteins. GB1 is an α/β protein with a four-strand β sheet packed against an α helix[16] (Fig. 1). There are two loops connecting the α helix to β strands 2 and 3, respectively. Using protein engineering techniques, we inserted various numbers of amino acid residues (2, 5, 24 and 46 residues) into loop 2 of GB1 to lengthen this flexible loop and used single-molecule AFM to investigate the effect of increasing loop length on the mechanical stability and mechanical unfolding kinetics of GB1. We found that increasing the length of the second loop of GB1 decreases the mechanical stability of GB1 and accelerates the mechanical unfolding of GB1. These results can be explained by the loss of configurational entropy upon closing an unstructured flexible loop using classical polymer theory, highlighting the important role of loop regions in the mechanical unfolding of proteins. This study not only demonstrated a general approach to investigating the structural deformation of the loop regions in the mechanical unfolding transition state, but also paved the way to using configurational entropy as an effective means to modulate the mechanical stability of proteins, which is of critical importance towards engineering artificial elastomeric proteins with tailored nanomechanical properties.

Results

The second loop of GB1 is tolerant of loop elongation

There are two loops in GB1 connecting the α helix to β strands 2 and 3, respectively. As a proof of principle, here we use the second loop of GB1 to investigate the effect of loop length on the mechanical stability and mechanical unfolding kinetics of GB1. Loop 2 connects the α helix to β strand 3 and is composed of five residues (residues 37–41, Fig. 1). Using protein engineering techniques, we inserted various numbers of flexible amino acid residues between residues 39 and 40 to obtain the loop insertion mutants GB1-L2, GB1-L5, GB1-L24 and GB1-L46, which elongate loop 2 by 2, 5, 24 and 46 residues, respectively (Fig. 1). For GB1-L2, a two-residue linker, LG, was inserted in loop 2; for GB1-L5, a flexible linker sequence, GGGLG, was inserted in loop 2; while for GB1-L24 and GB1-L46, flexible linkers of GSA repeats were inserted. In the Gly-Ser-Ala linkers, the smaller residues glycine and alanine provide flexibility to avoid formation of residual structure in the loop region, while serine residues increase the solubility and reduce aggregation. Gly-Ser-Ala repeat has been used as a flexible linker to construct a single-chain polypeptide analogue of HIV gp120/CD4 receptor complex.[20] The structural integrity of these loop insertion mutants was investigated using far-ultraviolet circular dichroism (far-UV CD) spectroscopy. CD spectroscopy is an excellent tool for dissecting secondary structures of proteins, as α helix, β sheet and random coil give rise to distinct CD spectra. The CD spectrum of α helix is characterized by two negative bands at ~222 and 208 nm, while the CD spectrum for β sheets has a negative band at ~215 nm. In contrast, the CD spectrum of a random coil is characterized by a negative band at ~195 nm. As shown in Fig. 2, the CD spectrum of wild-type (wt) GB1 shows two broad negative peaks with minima at 208 and 222 nm, consistent with the compact structure of GB1 with both α helix and β sheet features. Upon the insertion of various numbers of residues into loop 2, two minima are still clearly visible in the CD spectra of the four loop insertion mutants, suggesting that the compact structures with both α and β features are well retained in the mutants. In the CD spectra, the minima at 222 nm for the mutants remain unchanged, although their relative amplitude decreased compared with that of wt GB1. Conversely, the minima at 208 nm are gradually shifted towards lower wavelengths with increasing number of residues being inserted into loop 2. Since

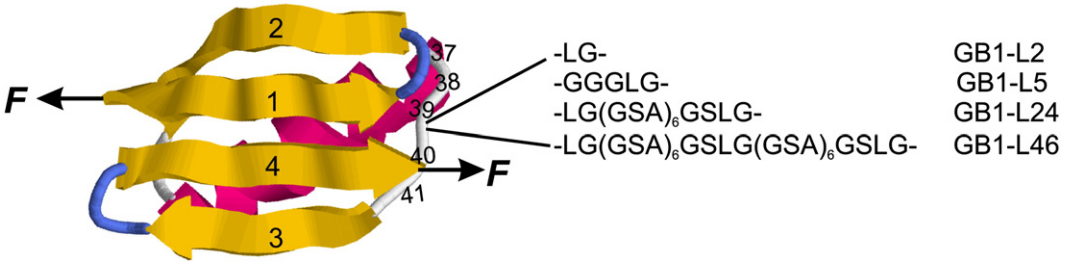

Fig. 1. Loop insertions in GB1. GB1 has a typical α/β structure with a four-strand β sheet packed against an α helix. There are two loops connecting the α helix to β strands 2 and 3, respectively. Loop 2 is composed of five residues (residues 37–41) and connects the α helix to β strand 3. Flexible linker sequences are inserted between residues 39 and 40. The sequences of the inserted residues are also shown.

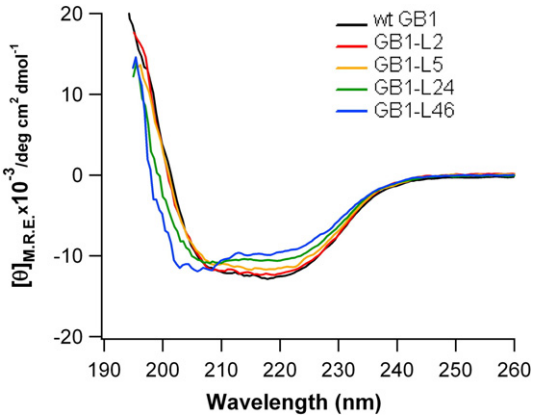

Fig. 2. Far-UV CD spectra of the loop insertion mutants indicate that loop 2 is tolerant of loop insertion and GB1 mutants are properly folded. The CD spectrum of wt GB1 exhibits broad bands with minima at 208 and 222 nm, consistent with the α/β structure of GB1. These features are also clearly visible in the CD spectra of the loop insertion mutants, suggestive of compact structures with α helix and β sheet features. Since the CD spectrum of a random coil is characterized by a negative band at ~195 nm, the increasing content of random coil sequences in GB1 by loop insertion leads to the observed shift of the minima at 208 nm towards lower wavelength.

the CD spectrum of random coil is characterized by a negative band at ~195 nm, the shift of the minima at 208 nm towards lower wavelength is indicative of the increase of random coil sequences in loop insertion mutants, consistent with the gradual increase of random coil sequences being inserted into loop 2. These results indicate that loop 2 is well tolerant of loop insertions without affecting GB1's native structure, even if the insertion is 46 residues long, which is almost the length of the whole wt GB1. These results are consistent with the finding that GB1 fragments, which were split in the middle of loop 2, can reconstitute its three-dimensional structure by complexation.[21,22] Therefore, these mutants of GB1 with long loop insertion provide good model systems for detailed investigations of the effect of loop length on the mechanical unfolding of proteins.

Features of the mechanical unfolding of loop insertion mutants

To investigate the mechanical properties of loop insertion mutants of GB1, we constructed four polyproteins, $(GB1\text{-}L2)_8$, $(GB1\text{-}L5)_8$, $(GB1\text{-}L24)_8$ and $(GB1\text{-}L46)_8$, which are composed of eight identical tandem repeats of the mutant protein under investigation. Stretching polyproteins of loop insertion mutants resulted in force–extension curves (Fig. 3) with characteristic sawtooth pattern appearance, where the individual unfolding force peaks correspond to the mechanical unfolding of the folded loop mutant GB1 domains. Compared with the force–extension curves of wt GB1, it is evident that the spacing between the consecutive unfolding events of the loop insertion mutants increased, while the unfolding forces decreased as the loop length increased. For example, the contour length increment (ΔL_c) of wt GB1, measured by fitting the Wormlike-chain model of polymer elasticity[23] to the consecutive unfolding events, is ~18 nm and the unfolding force is ~180 pN[17] (Fig. 3). In comparison, insertion of 24 residues into loop 2 results in a ΔL_c of ~27 nm, and the unfolding force decreased to ~80 pN. Mutant GB1-L24 is 80 residues long and

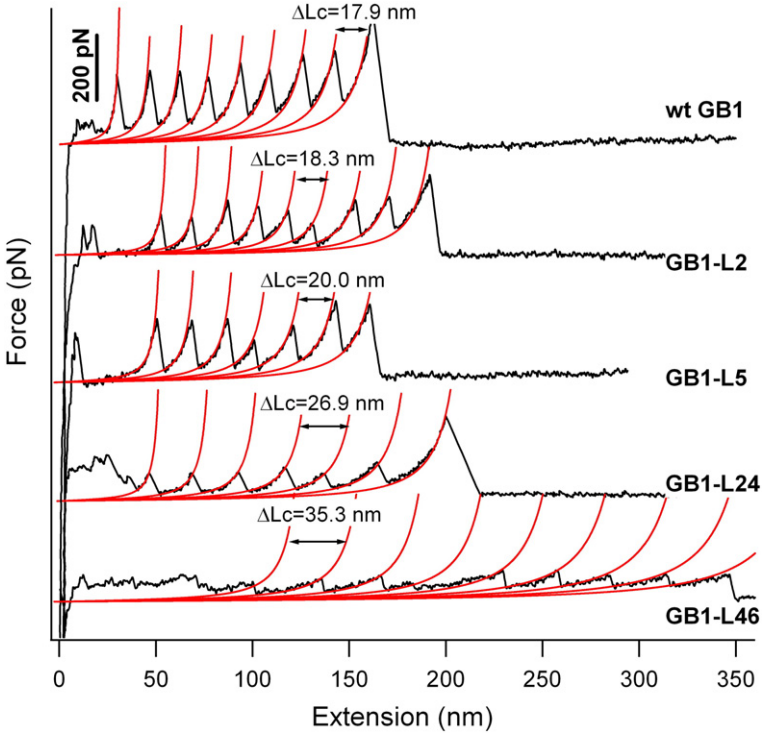

Fig. 3. Force–extension relationships of loop insertion mutants of GB1. Stretching polyproteins of GB1 loop insertion mutants results in characteristic sawtooth-like force–extension curves, in which the individual force peaks correspond to the mechanical unfolding of the individual mutant domains. The unfolding forces of loop insertion mutants decrease as the number of the amino acid residues inserted in loop 2 increases. WLC fits (red lines) to consecutive unfolding force peaks revealed that the contour length increment of loop insertion mutants increases as a function of the inserted number of amino acid residues.

is 28.8 nm long when it is unfolded and fully extended, assuming that the length of a single amino acid residue is 0.36 nm.[24] Since the insertion of 24 residues does not affect the proper folding of GB1, we take the distance of 2.6 nm[16] between the N- and C-termini of wt GB1 as the distance between the N- and C-termini of GB1-L24. Therefore, the complete unfolding of GB1-L24 will result in ΔL_c of 26.2 nm (28.8–2.6 nm), in good agreement with the experimentally observed ΔL_c for GB1-L24. This result corroborates the finding that the mutants are properly folded and the force–extension curves shown in Fig. 3 indeed correspond to the stretching and mechanical unfolding of the loop insertion mutants. Therefore, the mechanical unfolding forces measured from Fig. 3 correspond to the mechanical stability of the loop insertion mutants.

As predicted by molecular dynamics simulation,[25] the force-bearing β strands 1 and 4 of GB1 constitute a "mechanical clamp" to provide mechanical resistance to extending and unraveling of GB1. Disrupting the mechanical clamp will result in the mechanical unfolding of GB1 and exposure of the "hidden" amino acids in the core,[14] giving rise to the observed ΔL_c. If additional amino acids are inserted after the mechanical clamp, mechanical unfolding will result in the increase in ΔL_c. Since loop 2 is located after the mechanical clamp, we should expect to observe an increase in ΔL_c for the loop insertion mutants. Indeed, the ΔL_c of loop insertion mutants increases progressively as a function of the number of amino acids inserted in loop 2 (as shown in Fig. 4a).

Figure 5 shows the relationship between $\Delta\Delta L_c$ ($\Delta L_{c(\text{loop mutant})} - \Delta L_{c(\text{wt})}$) and the number of amino acid residues being inserted into loop 2. It is evident that $\Delta\Delta L_c$ is linearly proportional to the number of inserted residues. Linear regression to the experimental data in Fig. 5 measures a slope of ~0.37 nm per amino acid ($R^2 = 0.997$), indicating that the length of an individual amino acid residue is 0.37 nm, in close agreement with the expected length of a single amino acid in a fully extended polypeptide chain.[24]

Loop elongation decreases the mechanical stability of GB1

In addition to the progressively increasing ΔL_c of GB1 in response to the loop elongation, the mechanical stability of loop insertion mutants also decreases progressively as a function of the loop length. Figure 4b shows the unfolding force histogram for the loop mutants of GB1 as well as the wild-type GB1. It is

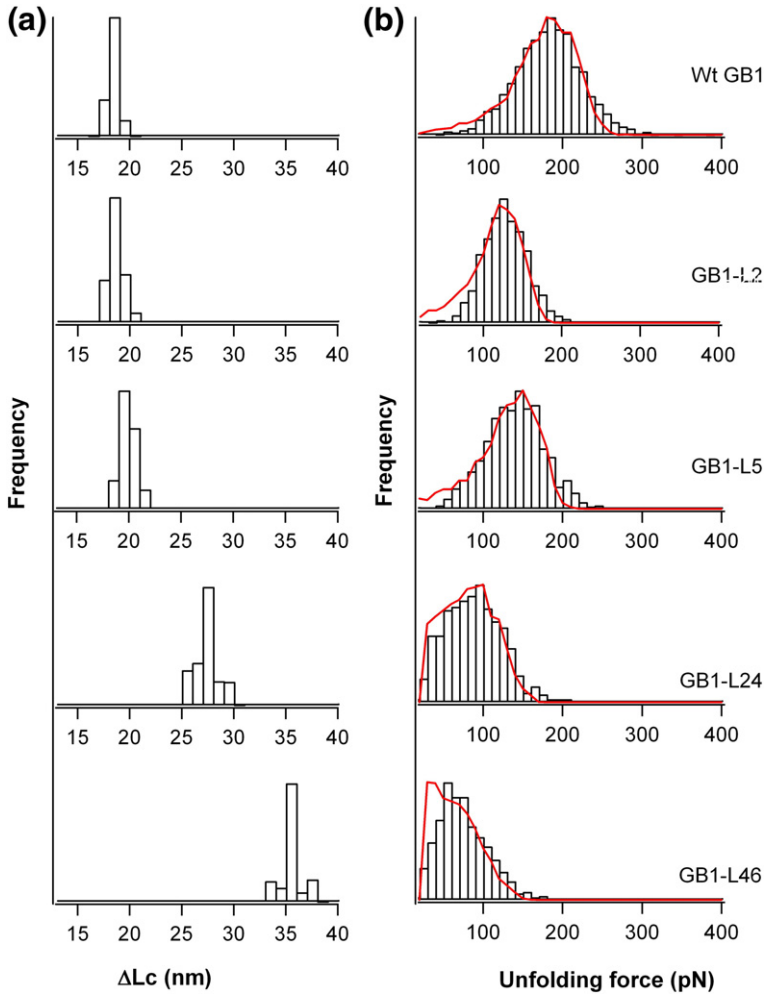

Fig. 4. Histograms for ΔL_c and the mechanical unfolding force of loop insertion mutants. (a) Histograms for ΔL_c of loop insertion mutants. The average ΔL_c is 18.3±0.5 nm (n=470) for wt GB1, 18.5±0.7 nm (n=415) for GB1-L2, 19.8±0.7 nm (n=307) for GB1-L5, 26.9±1.1 nm (n=796) for GB1-L24 and 35.0±1.0 nm (n=595) for GB1-L46. It is evident that ΔL_c is increased from ~18 nm for wt GB1 to ~35 nm for GB1-L46. (b) Histograms of the mechanical unfolding forces for loop insertion mutant. The average unfolding force is 177±40 pN (n=2593) for wt GB1, 121±28 pN (n=1483) for GB1-L2, 134±36 pN (n=1209) for GB1-L5, 80±35 pN (n=659) for GB1-L24 and 63±31 pN (n=595) for GB1-L46. Red lines are Monte Carlo fits to the experimental data using the following unfolding distance Δx_u and unfolding rate constant α_0 values: 0.17 nm and 0.039 s^{-1} for wt GB1; 0.22 nm and 0.075 s^{-1} for GB1-L2; 0.19 nm and 0.090 s^{-1} for GB1-L5; 0.19 nm and 0.40 s^{-1} for GB1-L24; and 0.19 nm and 0.62 s^{-1} for GB1-L46. It is of note that the detection limit measured for the unfolding force in our AFM is ~20 pN.

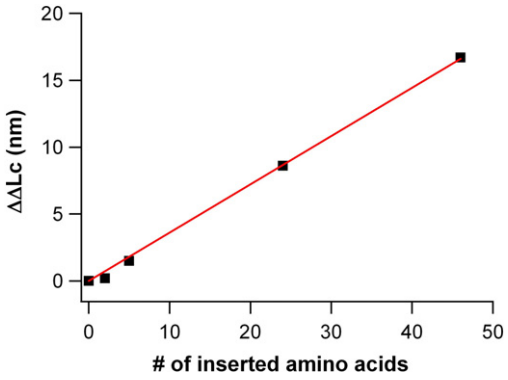

Fig. 5. $\Delta\Delta L_c$ is linearly proportional to the number of inserted amino acid residues in loop2. A linear regression measures a slope of 0.37 ± 0.01 nm per amino acid ($R^2 = 0.997$), in good agreement with the predicted value of the length of a single amino acid residue.

evident that the unfolding forces of loop mutants of GB1 showed significant decrease: the unfolding force decreases from 180 pN for the wt-GB1to ~60 pN for the GB1-L46 mutant. Figure 6a depicts the dependency of the unfolding forces on the length of loop 2. The mechanical stability decreases in a nonlinear fashion as a function of the inserted loop length, and the decreasing trend tends to attenuate when the loop length increases further. It is of note that mutant GB1-L2 is somewhat special, as its unfolding force is clearly lower than that for the mutant GB1-L5. These results strongly indicate that the loop length can have a profound influence on the mechanical stability of proteins.

The unfolding force of a protein is determined by two characteristic parameters of the free energy diagram for the mechanical unfolding reaction: the height of the free energy barrier ΔG_{T-N} for mechanical unfolding and the unfolding distance Δx_u between the folded state (N) and unfolding transition state (T). From the unfolding force histograms it is possible to estimate the spontaneous unfolding rate constant at zero force (α_0), which is directly related to ΔG_{T-N}, and unfolding distance Δx_u. It is evident that the width of the unfolding force histogram for GB1-L2 is much narrower than that for wt-GB1, while the width for the other three loop mutants remains largely unchanged as compared with that for wt-GB1. Since the width of the unfolding force histogram is directly related to Δx_u,[26] this result indicates that Δx_u of the loop mutants GB1-L5, L24 and L46 is similar to that of the wt-GB1, implying that the mechanical unfolding pathway of the loop mutants of GB1 remains largely unaltered. However, there is a change in the Δx_u for GB1-L2, suggesting that the unfolding pathway has been altered by the insertion of linker sequence LG. To estimate α_0 and Δx_u accurately, we carried out pulling experiments on these GB1 mutant proteins at different pulling speeds. As shown in Fig. 7, the unfolding forces of GB1 loop insertion mutants depend on the pulling speeds: the higher the pulling speed is, the higher the force required to unfold the proteins. It is of note that the slope of the pulling speed dependence of the unfolding force of GB1-L2 is clearly different from that of wt GB1 as well as other loop insertion mutants, suggesting a change in Δx_u for GB1-L2, consistent with the conclusion drawn from the unfolding force histograms. Using standard Monte Carlo simulation procedures,[27,28] we reproduced force–extension curves of GB1 mutants. By fitting the unfolding force distribution (Fig. 4b) and the speed dependence of the unfolding force (Fig. 7) simultaneously, we have estimated α_0 and Δx_u for individual mutant proteins. For loop mutants GB1-L5, L24 and L46, Δx_u was estimated to be 0.19 nm, which is slightly bigger than that of wt GB1.[17] For loop mutant GB1-L2, Δx_u was estimated to be ~0.22 nm, which is significantly bigger than that of wt GB1. Δx_u measured for wt GB1 as well as their loop insertion mutants are quite small, suggesting that the mechanical unfolding transition state is highly native-like. It is worth noting that the Monte Carlo simulation method we used here is based on the Bell–Evans model for force-induced unfolding[26,29]; thus, the values we measured here for

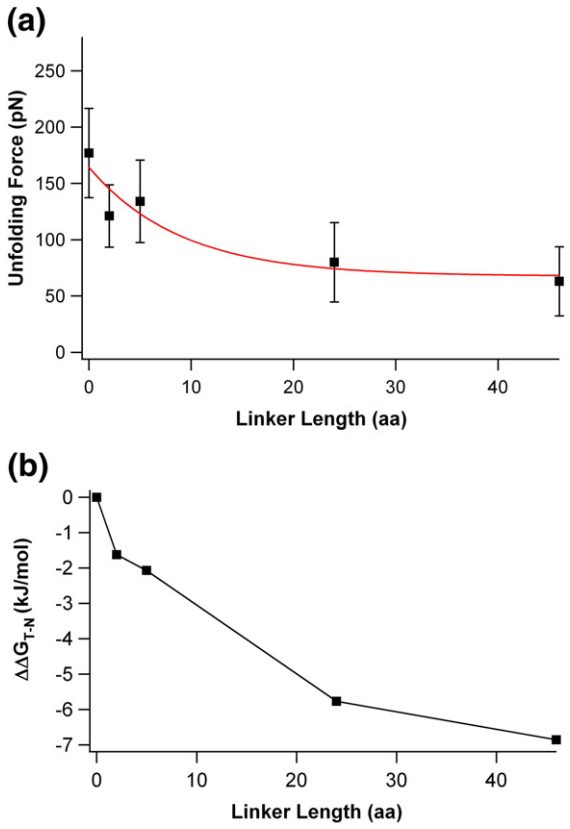

Fig. 6. Mechanical stability (a) and free energy barrier for mechanical unfolding (b) show nonlinear dependence on the number of amino acid residues inserted in loop 2. The effect of mechanical destabilization and acceleration of unfolding kinetics tends to attenuate with further increase of the number of amino acid residues in loop 2. Red line in (a) is an exponential fit to the experimental data to illustrate the trend of how the unfolding force changes as a function of linker length.

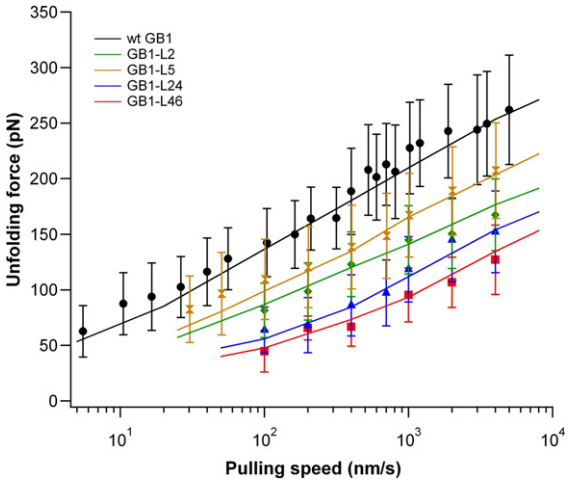

Fig. 7. Speed dependence of the mechanical unfolding forces of wt GB1 and loop insertion mutants. The average unfolding forces of loop insertion mutants and wt GB1 increase with the rising of pulling speed. Error bars indicate the standard deviation of unfolding force. Continuous lines correspond to Monte Carlo fits of the experimental data using the following values of Δx_u and α_0: 0.17 nm and 0.039 s^{-1} for wt GB1; 0.22 nm and 0.075 s^{-1} for GB1-L2; 0.19 nm and 0.090 s^{-1} for GB1-L5; 0.19 nm and 0.40 s^{-1} for GB1-L24; and 0.19 nm and 0.62 s^{-1} for GB1-L46.

α_0 and Δx_u may differ from those estimated using Kramer-theory-based approach.[30]

Having measured the unfolding rate constant α_0 for GB1 and its mutants, we can then examine the effect of loop length on the free energy barrier ΔG_{T-N} of GB1. Since $\alpha_0 = \upsilon \exp(-\Delta G_{T-N}/RT)$, where υ is the prefactor, R is the gas constant and T is the temperature, the difference of the unfolding kinetic energy barrier upon loop elongation, $\Delta\Delta G_{T-N}$, follows the relationship:

$$\Delta\Delta G_{T-N} = \Delta G_{T-N}(mut) - \Delta G_{T-N}(wt) = RT\ln\left(\frac{\alpha_0(wt)}{\alpha_0(mut)}\right)$$

Plotting $\Delta\Delta G_{T-N}$ *versus* loop length allows us to investigate how the mechanical unfolding kinetic barrier changes with the increasing of loop length (Fig. 6b). It is evident that $\Delta\Delta G_{T-N}$ shows a nonlinear dependency on the loop length and the effect of loop length tends to attenuate with increasing number of amino acids in loop 2. Such nonlinear dependency of unfolding free energy barrier is similar to that observed in chemical unfolding of circular permutants of α-spectrin SH3.[3] These results suggest that for mutants GB1-L5, GB1-L24 and GB1-L46, the effect of loop length on the mechanical stability of GB1 is accomplished largely by reducing the free energy barrier for unfolding, rather than shifting the mechanical unfolding pathways. In contrast, the effect of loop length on the mechanical stability of mutant GB1-L2 is more complex and involves the change in both free energy barrier and Δx_u.

Discussion

Configurational entropy plays important roles in mechanical unfolding kinetics of GB1

Previous studies on the mechanical unfolding of proteins were mainly focused on specific interactions, such as backbone hydrogen bonds, in the key region of the protein of interest, and the effect of configurational entropy on the mechanical unfolding of proteins remains largely unexplored. By increasing the length of loop 2 of GB1 systematically, we use single-molecule AFM to demonstrate that loop insertion decreases the mechanical stability and accelerates the mechanical unfolding kinetics of GB1. Formation of loop regions in proteins is directly associated with the loss of configurational entropy.[31] Thus, our results provide the first direct experimental evidence that configurational entropy plays important roles in determining the mechanical unfolding kinetics and mechanical stability of proteins.

The effect of configurational entropy on mechanical unfolding kinetics can be explained using classical polymer theory.[2,31] It can be shown that the change in configurational entropy of closing an unstructured loop of L residues relative to the loop of L_0 residues can be estimated by the following relationship:

$$\Delta\Delta S_{N-U} = \Delta S_{N-U}(L) - \Delta S_{N-U}(L_0) = -C^*R\ln(L/L_0) \quad (1)$$

where L_0 is the length of the reference loop (in the case of GB1, L_0 is 5), L is the length of the loop after loop insertion, C^* is a coefficient that is dependent upon the polymer used[3,32] and R is the gas constant. Therefore, the change in the thermodynamic stability for the two proteins after loop insertion is:

$$\Delta\Delta G_{U-N} = \Delta G_{U-N}(L) - \Delta G_{U-N}(L_0) = -C^*RT\ln(L/L_0) \quad (2)$$

where $\Delta G_{U-N}(L_0)$ is the thermodynamic stability of the protein with a loop of L_0 residues, $\Delta G_{U-N}(L)$ is the thermodynamic stability of the protein after the loop is elongated to L and T is the temperature in Kelvin. Thus, loop insertion will considerably increase the entropic cost of closing an unstructured loop, resulting in decreased thermodynamic stability of the protein. Moreover, if the closing of an unstructured loop is required in the folding transition state, loop insertion will also decelerate the rate of folding reaction.[5]

The effect of configurational entropy on mechanical unfolding kinetics reflects its influence on the free energy barrier for mechanical unfolding. The free energy barrier change for mechanical unfolding upon loop insertion can be written as:

$$\Delta\Delta G_{T-N} = \Delta\Delta G_{T-U} - \Delta\Delta G_{N-U} = C^*RT[\ln(L'/L_0') - \ln(L/L_0)] \quad (3)$$

where L'_0 and L' correspond to the effective loop length of the transition state before and after loop insertion, respectively, and ΔG_{T-N} represents the

free energy barrier for unfolding. Depending on whether the loop formation is complete or not in the transition state compared with that in the native state, the effect of configurational entropy can be different. If the loop is as formed in the mechanical unfolding transition state as in the native state, loop insertion will destabilize the unfolding transition state the same amount as it does on the native state (Fig. 8a), leading to zero for $\Delta\Delta G_{T-N}$. In this case, the loop insertion has no effect on the unfolding free energy barrier. If the loop formation in the transition state is not as complete as in the native state, loop insertion will have a weaker destabilization effect on the unfolding transition state than it does on the native state, leading to a negative $\Delta\Delta G_{T-N}$. In this case, loop insertion will reduce the unfolding free energy barrier and accelerate the unfolding kinetics (Fig. 8b). Our single-molecule AFM results demonstrate that loop insertion accelerates the mechanical unfolding kinetics. These results suggest that the formation of loop 2 in the mechanical unfolding transition state is partially disrupted by stretching force, and thus, loop 2 in the mechanical unfolding transition state is no longer as formed as in the native state of GB1. This result highlights the involvement of loop regions in the mechanical unfolding transition state.

Previous single-molecule AFM experiments and molecular dynamics simulation were largely focused on the rupture of key interactions, such as hydrogen bonds, during mechanical unfolding, which reflect the enthalpic contributions to the mechanical unfolding. Our results reported here suggest that enthalpic factors are not everything and entropic contribution can also play significant roles in determining the mechanical unfolding kinetics. Our result suggests that the deformation of the loop regions during mechanical unfolding of proteins could have been overlooked in previous single-molecule AFM experiments. Hence, our present study demonstrates an effective approach to investigating the importance of loop regions in the mechanical unfolding kinetics and mechanical stability of proteins. For GB1, it will be interesting to compare the roles of the two different loops played on the mechanical unfolding of GB1 in our future endeavors.

Furthermore, our result is somewhat unexpected because of the rather small unfolding distance Δx_u between the native state and the mechanical unfolding state. Our results thus can serve as a benchmark for future molecular dynamics simulations to illustrate, at the molecular level, the role the loop region played in the mechanical unfolding transition state.

Configurational entropy and mechanical stability of proteins

Mechanical stability, i.e., unfolding force, of proteins is determined by two factors: the free energy barrier for unfolding (ΔG_{T-N}) and the distance Δx_u between the native state and mechanical unfolding transition state. Compared with the relatively straightforward explanation of the effect of configurational entropy on the mechanical unfolding kinetics, the effect of configurational entropy on mechanical stability can be more complex, if loop elongation would result in a change in Δx_u for the mechanical unfolding.

For loop insertion mutants GB1-L5, L24 and L46, Δx_u shows only a small change compared with wt GB1. The effect of configurational entropy on their mechanical stability can be largely accounted for by the change in the mechanical unfolding free energy barrier. In contrast, Δx_u for GB1-L2 is significantly larger than that for other loop insertion mutants as well as wt GB1. The effect of configurational entropy on the mechanical stability of GB1-L2 is thus the combined effect of increased unfolding rate constant α_0 and Δx_u. This also explains why the mechanical unfolding force of GB1-L2 is abnormal in compar-

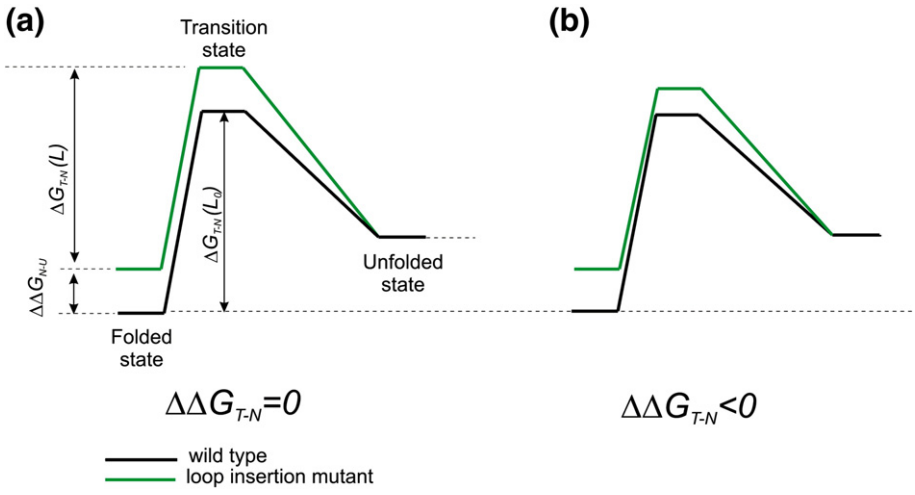

Fig. 8. Schematic free energy diagram explains the effect of loop length on the mechanical unfolding kinetics. If the loop is fully formed in the mechanical unfolding transition state as it does in the folded state, increasing the loop length will result in the same amount of destabilization for both native state and transition state (a). Thus, loop length will not affect the free energy barrier for mechanical unfolding. If the loop is partially deformed in the transition state, loop length will affect the free energy barrier for mechanical unfolding (b).

ison with that of wt GB1 and other loop insertion mutants. However, the molecular mechanism remains unclear on how the insertion of two residues leads to such a big change in the unfolding distance Δx_u for GB1-L2. The two amino acid residues introduced in GB1-L2, resulted from the use of the restriction site of AvaI, are leucine and glycine. Since leucine is a relatively bulky and hydrophobic amino acid, it is possible that the inserted leucine in this short loop of GB1 may interact with other parts of GB1, leading to the change in Δx_u. In the case of other loop insertion mutants, it seems that the impact of leucine is buffered by the further introduction of additional flexible amino acid residues. However, this explanation is just a speculation. Molecular-level interpretation of this effect would require high-resolution structure of this particular mutant and steered molecular dynamics simulation of the mechanical unfolding processes.

Nonetheless, if loop elongation does not change the unfolding distance, the effect of configurational entropy on the mechanical stability can be determined solely by its effect on the mechanical unfolding kinetics. Such results provide the basis for the use of configurational entropy to modulate the mechanical stability of a given protein. Tuning the mechanical stability of proteins is usually done via site-directed mutagenesis approach,[13,33–37] which is largely trial-and-error based. The range of tuning the mechanical stability by site-directed mutagenesis is also rather limited and difficult to predict. In comparison, the loop insertion approach demonstrated here, as well as loop reduction (by engineered disulfide crosslinking[38]), is potentially a much more general one to modulate the mechanical stability of proteins, especially when one considers the ubiquitous existence of loops in proteins. This approach, complementing our recently reported mechanical stability tuning approach by chemical denaturant,[39] potentially can enable the tuning of the mechanical stability of a given protein continuously in a semi-predictable fashion according to the nonlinear dependence of the unfolding forces on the number of amino acid residues inserted. For example, based on our experimental data, it is now possible to predict the mechanical stability of the GB1 mutant with 15 residues being inserted into loop 2. Furthermore, by combining loop insertion and engineered disulfide crosslinking, it will become possible to tune the mechanical stability of protein reversibly by changing the redox potential of the aqueous solution. Such studies will be of critical importance to engineering artificial elastomeric proteins of tailored nanomechanical properties for nanomechanical and biomedical applications.

Materials and Methods

Protein engineering

The gene encoding the DNA sequence of GB1 is in the pUC19 plasmid. In pUC19-GB1, GB1 is flanked by a 5′ BamHI site and a 3′ BglII site followed by a KpnI site.

In order to facilitate the insertion of various numbers of amino acids to loop 2, we first inserted a nonpalindromic AvaI restriction site (CTC GGG) between codons for the 39th and 40th amino acid residues of wt GB1. The insertion of the AvaI site was carried out using standard protocols of site-directed mutagenesis, resulting in pUC19-GB1(AvaI). The presence of the AvaI restriction site in purified clones was confirmed by restriction digestion followed by DNA sequencing. The insertion of the AvaI restriction site resulted in the mutant GB1-L2, in which a two-residue linker LG was inserted into loop 2 of GB1.

To construct mutant GB1-L5, which incorporates three Gly residues in loop 2, the insert encoding GB1 residues 1–39 was PCR amplified with a 5′ overhang of restriction site BamHI and 3′ overhang of DNA encoding three Gly residues followed by restriction site AvaI. The amplified PCR product was gel purified and digested with restriction enzymes BamHI and AvaI and subcloned into pUC19-GB1(AvaI) digested with BamHI and AvaI restriction enzymes to produce pUC19-GB1-L5. The resulting construct, GB1-L5, contains the insertion of five residues, GGGLG, between the 39th and 40th amino acid residues of wt GB1.

To construct mutants GB1-L24 and GB1-L46, both the Watson and Crick strands for DNA encoding the flexible linker sequence $(GSA)_6GS$ flanked with the AvaI restriction site at both ends were synthesized separately by oligosynthesis (Integrated DNA Technologies). Both strands were mixed to a final concentration of 50 μM in 1 mM Tris–HCl buffer, pH 8.0, and incubated at 50 °C for 10 min. The product was digested with restriction enyzme AvaI and ligated into pUC19-GB1(AvaI) digested with the same enzyme. The ligation mixture was transformed into *Escherichia coli* strain DH5α. It is of note that during the ligation, the insert encoding $(GSA)_6GS$ can self-ligate to produce concatemers of the insert $(GSA)_6GS$, as well as ligate into the digested vector. Transformants were screened for plasmids containing one or two repeats of $(GSA)_6GS$ by digesting purified plasmids with restriction enzymes BamHI and KpnI. Plasmids containing one or two linkers were named as pUC19-GB1-L24 and pUC19-GB1-L46, respectively. The resulting GB1-L24 and GB1-L46 contain the 24-residue insertion $LG(GSA)_6GSLG$ or the 46-residue insertion $LG(GSA)_6GSLG(GSA)_6GSLG$ between the 39th and 40th amino acid residues of wt GB1, respectively. The extra inserted LG are resulted from the AvaI restriction sites.

Polyproteins $(GB1\text{-}L2)_8$, $(GB1\text{-}L5)_8$, $(GB1\text{-}L24)_8$ and $(GB1\text{-}L46)_8$ were constructed using procedures similar to the construction of $(GB1)_8$, which are based on the identity of the sticky ends of BamHI and BglII sites. Polyprotein genes were subcloned into expression vector pQE80L, which contains an N-terminal $(His)_6$ purification tag to facilitate the purification of expressed proteins. The polyprotein was expressed in DH5α strain and purified using Ni-NTA affinity chromatography.

Single-molecule AFM

All the single-molecule AFM measurements were performed with a custom-built atomic force microscope as described.[17] The cantilevers were calibrated in phosphate-buffered saline (PBS) solution using the equipartition theorem with an average error of 10%. In a typical unfolding experiment, 1 μL of polyprotein solution was deposited onto a freshly cleaned glass coverslip containing ~50 μL of PBS and thoroughly mixed. After ~10 min of equilibration, single-molecule AFM pulling experiments were performed. The pulling speed was 400 nm/s

for all the unfolding experiments except as reported otherwise. Monte Carlo simulations were carried using standard procedures.[27] To reflect the noise level of the unfolding force measurements in our single-molecule AFM experiments, the minimal detectable unfolding force in Monte Carlo simulation was set to 30 pN.

Circular dichroism measurements

CD spectra were recorded on a Jasco-J810 spectropolarimeter flushed with nitrogen gas. The spectra were recorded in a 0.2-cm path-length cuvette at a scan rate of 50 nm/min. For each protein sample, an average of three scans is reported. The protein samples were measured in 0.1× PBS at pH 7.4. Data have been corrected for buffer contributions.

Acknowledgements

This work is supported by Natural Sciences and Engineering Research Council of Canada, Canada Research Chairs program and Canada Foundation for Innovation. The CD measurements were performed on a Jasco-J810 spectropolarimeter, which was provided through a grant from the Canada Foundation for Innovation to the Laboratory of Molecular Biophysics, University of British Columbia.

References

1. Branden, C. & Tooze, J. (1999). *Protein Structure* (2nd ed). Garland Publishing, Inc., New York.
2. Chan, H. S. & Dill, K. A. (1988). Intrachain loops in polymers. *J. Chem. Phys.* **90**, 492–509.
3. Viguera, A. R. & Serrano, L. (1997). Loop length, intramolecular diffusion and protein folding. *Nat. Struct. Biol.* **4**, 939–946.
4. Nagi, A. D., Anderson, K. S. & Regan, L. (1999). Using loop length variants to dissect the folding pathway of a four-helix-bundle protein. *J. Mol. Biol.* **286**, 257–265.
5. Grantcharova, V. P., Riddle, D. S. & Baker, D. (2000). Long-range order in the src SH3 folding transition state. *Proc. Natl Acad. Sci. USA*, **97**, 7084–7089.
6. Fersht, A. R. (2000). Transition-state structure as a unifying basis in protein-folding mechanisms: contact order, chain topology, stability, and the extended nucleus mechanism. *Proc. Natl Acad. Sci. USA*, **97**, 1525–1529.
7. Scalley-Kim, M., Minard, P. & Baker, D. (2003). Low free energy cost of very long loop insertions in proteins. *Protein Sci.* **12**, 197–206.
8. Fisher, T. E., Marszalek, P. E. & Fernandez, J. M. (2000). Stretching single molecules into novel conformations using the atomic force microscope. *Nat. Struct. Biol.* **7**, 719–724.
9. Carrion-Vazquez, M., Oberhauser, A., Diez, H., Hervas, R., Oroz, J., Fernandez, J. & Martinez-Martin, A. (2006). Protein nanomechanics as studied by AFM single-molecule force spectroscopy. In *Advanced Techniques in Biophysics* (Arrondo, J. & Alonso, A., eds), pp. 163–236, Springer-Verlag, Berlin, Heidelberg.
10. Fernandez, J. M. & Li, H. (2004). Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science*, **303**, 1674–1678.
11. Zhuang, X. W. & Rief, M. (2003). Single-molecule folding. *Curr. Opin. Struct. Biol.* **13**, 88–97.
12. Brockwell, D. J. (2007). Force denaturation of proteins—an unfolding story. *Curr. Nanosci.* **3**, 3–15.
13. Li, H. (2007). Engineering proteins with tailored nanomechanical properties: a single molecule approach. *Org. Biomol. Chem.* **5**, 3399–3406.
14. Carrion-Vazquez, M., Marszalek, P. E., Oberhauser, A. F. & Fernandez, J. M. (1999). Atomic force microscopy captures length phenotypes in single proteins. *Proc. Natl Acad. Sci. USA*, **96**, 11288–11292.
15. Schwaiger, I., Kardinal, A., Schleicher, M., Noegel, A. A. & Rief, M. (2004). A mechanical unfolding intermediate in an actin-crosslinking protein. *Nat. Struct. Mol. Biol.* **11**, 81–85.
16. Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T. & Clore, G. M. (1991). A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science*, **253**, 657–661.
17. Cao, Y., Lam, C., Wang, M. & Li, H. (2006). Non-mechanical protein can have significant mechanical stability. *Angew. Chem., Int. Ed. Engl.* **45**, 642–645.
18. Cao, Y. & Li, H. (2007). Polyprotein of GB1 is an ideal artificial elastomeric protein. *Nat. Mater.* **6**, 109–114.
19. Cao, Y., Balamurali, M. M., Sharma, D. & Li, H. (2007). A functional single-molecule binding assay via force spectroscopy. *Proc. Natl Acad. Sci. USA*, **104**, 15677–15681.
20. Fouts, T. R., Tuskan, R., Godfrey, K., Reitz, M., Hone, D., Lewis, G. K. & DeVico, A. L. (2000). Expression and characterization of a single-chain polypeptide analogue of the human immunodeficiency virus type 1 gp120–CD4 receptor complex. *J. Virol.* **74**, 11427–11436.
21. Honda, S., Kobayashi, N., Munekata, E. & Uedaira, H. (1999). Fragment reconstitution of a small protein: folding energetics of the reconstituted immunoglobulin binding domain B1 of streptococcal protein G. *Biochemistry*, **38**, 1203–1213.
22. Kobayashi, N., Honda, S. & Munekata, E. (1999). Fragment reconstitution of a small protein: disulfide mutant of a short C-terminal fragment derived from streptococcal protein G. *Biochemistry*, **38**, 3228–3234.
23. Marko, J. F. & Siggia, E. D. (1995). Stretching DNA. *Macromolecules*, **28**, 8759–8770.
24. Eisenberg, D. (2003). The discovery of the alpha-helix and beta-sheet, the principal structural features of proteins. *Proc. Natl Acad. Sci. USA*, **100**, 11207–11210.
25. Li, P. C. & Makarov, D. E. (2004). Ubiquitin-like protein domains show high resistance to mechanical unfolding similar to that of the 127 domain in titin: evidence from simulations. *J. Phys. Chem. B*, **108**, 745–749.
26. Evans, E. (2001). Probing the relation between force–lifetime–and chemistry in single molecular bonds. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 105–128.
27. 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.
28. Rief, M., Fernandez, J. M. & Gaub, H. E. (1998). Elastically coupled two-level systems as a model for biopolymer extensibility. *Phys. Rev. Lett.* **81**, 4764–4767.
29. Bell, G. I. (1978). Models for specific adhesion of cells to cells. *Science*, **200**, 618–627.
30. Schlierf, M. & Rief, M. (2006). Single-molecule

unfolding force distributions reveal a funnel-shaped energy landscape. *Biophys. J.* **90**, L33–L35.
31. Jacobsen, H. & Stockmayer, W. H. (1950). Intramolecular reaction in polycondensations. I. The theory of linear systems. *J. Chem. Phys.* **18**.
32. Pace, C. N., Grimsley, G. R., Thomson, J. A. & Barnett, B. J. (1988). Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J. Biol. Chem.* **263**, 11820–11825.
33. Li, H., Carrion-Vazquez, M., Oberhauser, A. F., Marszalek, P. E. & Fernandez, J. M. (2000). Point mutations alter the mechanical stability of immunoglobulin modules. *Nat. Struct. Biol.* **7**, 1117–1120.
34. Li, H. & Fernandez, J. M. (2003). Mechanical design of the first proximal Ig domain of human cardiac titin revealed by single molecule force spectroscopy. *J. Mol. Biol.* **334**, 75–86.
35. Marszalek, P. E., Lu, H., Li, H., Carrion-Vazquez, M., Oberhauser, A. F., Schulten, K. & Fernandez, J. M. (1999). Mechanical unfolding intermediates in titin modules. *Nature*, **402**, 100–103.
36. Williams, P. M., Fowler, S. B., Best, R. B., Toca-Herrera, J. L., Scott, K. A., Steward, A. & Clarke, J. (2003). Hidden complexity in the mechanical properties of titin. *Nature*, **422**, 446–449.
37. Brockwell, D. J., Beddard, G. S., Clarkson, J., Zinober, R. C., Blake, A. W., Trinick, J. *et al.* (2002). The effect of core destabilization on the mechanical resistance of I27. *Biophys. J.* **83**, 458–472.
38. Ainavarapu, S. R., Brujic, J., Huang, H. H., Wiita, A. P., Lu, H., Li, L. *et al.* (2007). Contour length and refolding rate of a small protein controlled by engineered disulfide bonds. *Biophys. J.* **92**, 225–233.
39. Cao, Y. & Li, H. (2008). How do chemical denaturants affect the mechanical folding and unfolding of proteins? *J. Mol. Biol.* **375**, 316–324.