THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

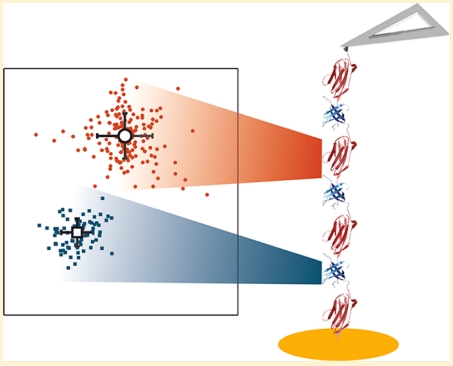
pubs.acs.org/JPCB

Single-Molecule Force Spectroscopy Identifies a Small Cold Shock Protein as Being Mechanically Robust

Toni Hoffmann,[†,‡,§] Katarzyna M. Tych,[†,‡,§] David J. Brockwell,[†] and Lorna Dougan*[,†,‡]

[†]Astbury Centre for Structural Molecular Biology and [‡]School of Physics and Astronomy, University of Leeds, Leeds, LS2 9JT, United Kingdom

ABSTRACT: Single-molecule force spectroscopy has emerged as a powerful approach to examine the stability and dynamics of single proteins. We have completed force extension experiments on the small cold shock protein B from *Thermotoga maritima*, using a specially constructed chimeric polyprotein. The protein's simple topology, which is distinct from the mechanically well-characterized β-grasp and immunoglobulin (Ig)-like folds, in addition to the wide range of structural homologues resulting from its ancient origin, provides an attractive model protein for single-molecule force spectroscopy studies. We have determined that the protein has mechanical stability, unfolding at greater than 70 pN at a pulling velocity of 100 nm s^{-1}. We reveal features of the unfolding energy landscape by measuring the dependence of the mechanical stability on pulling velocity, in combination with Monte Carlo simulations. We show that the cold shock protein has mechanically robust, yet malleable, features that may be important in providing the protein with stability and flexibility to function over a range of environmental conditions. These results provide insights into the relationship between the secondary structure and topology of a protein and its mechanical strength. This lays the foundation for the investigation of the effects of changes in environmental conditions on the mechanical and dynamic properties of cold shock proteins.



INTRODUCTION

Single-molecule force spectroscopy is a powerful tool to study the conformational dynamics and stability of proteins.[1−5] By completing force−extension experiments using an atomic force microscope (AFM), a protein can be extended and unfolded at a constant velocity, yielding information on the mechanical stability of the protein in terms of the force (F_U) required to unfold it. Such studies have revealed rich information on the mechanical stability of proteins, the presence of intermediate states, and the importance of intramolecular interactions in determining protein mechanical stability.[6] This technique is rapidly advancing, and the number of natural and designed proteins studied in experiments, combined with those characterized by computational modeling, provide a growing data set for the detailed analysis of the mechanical stability of proteins.[7−9]

With respect to their mechanical robustness, proteins can be ranked according to their secondary structure content and arrangement, where α-helical proteins generally exhibit lower mechanical stability than those with a high percentage of β-sheet content. The importance of the arrangement of the secondary structure in relation to the direction of the pulling force has also been demonstrated, where the shearing apart of two β-strands requires a greater force than a sequential "unzipping".[10,11] Further studies have examined side chain packing and long-range interactions in topologically similar proteins,[12] hydrophobic packing in the hydrophobic core of a protein,[13] solvent accessibility of hydrogen bonds,[14] and bond patterns as well as the identification of "strong" and "weak" sequence motifs in protein families.[7,14−18]

The family of cold shock proteins (Csps) belongs to a subset of the OB (oligonucleotide/oligosaccharide-binding) class of folds, a protein fold that is found in all three kingdoms of life.[19] The OB-fold is a β-barrel structure formed by five antiparallel β-strands arranged in two β-sheets that in most cases is capped by an α-helix.[19] This α-helix is missing in some cold shock protein domains, which simply contain a Greek-key β-barrel structure. To form the β-barrel, the β-sheets twist and coil to form closed structures in which the first strand is hydrogen-bonded to the last.

The existence of Csps in a broad range of organisms, including the most primitive diverging lineages, suggests they may have ancient origin and were present in the earliest forms of life.[20−23] While the expression of many Csps is upregulated by a sudden decrease in temperature, as part of a cold shock response, most cold shock proteins are present under ambient environmental conditions and therefore may have an important role in response to other forms of stress.[21−24] In their proposed function as RNA chaperones, they are involved in the regulation of a number of biological functions by binding to single-stranded RNA and DNA to regulate ribosome translation, mRNA decay, and termination of transcription.[24]

Received: October 22, 2012
Revised: December 17, 2012
Published: January 8, 2013

ACS Publications © 2013 American Chemical Society

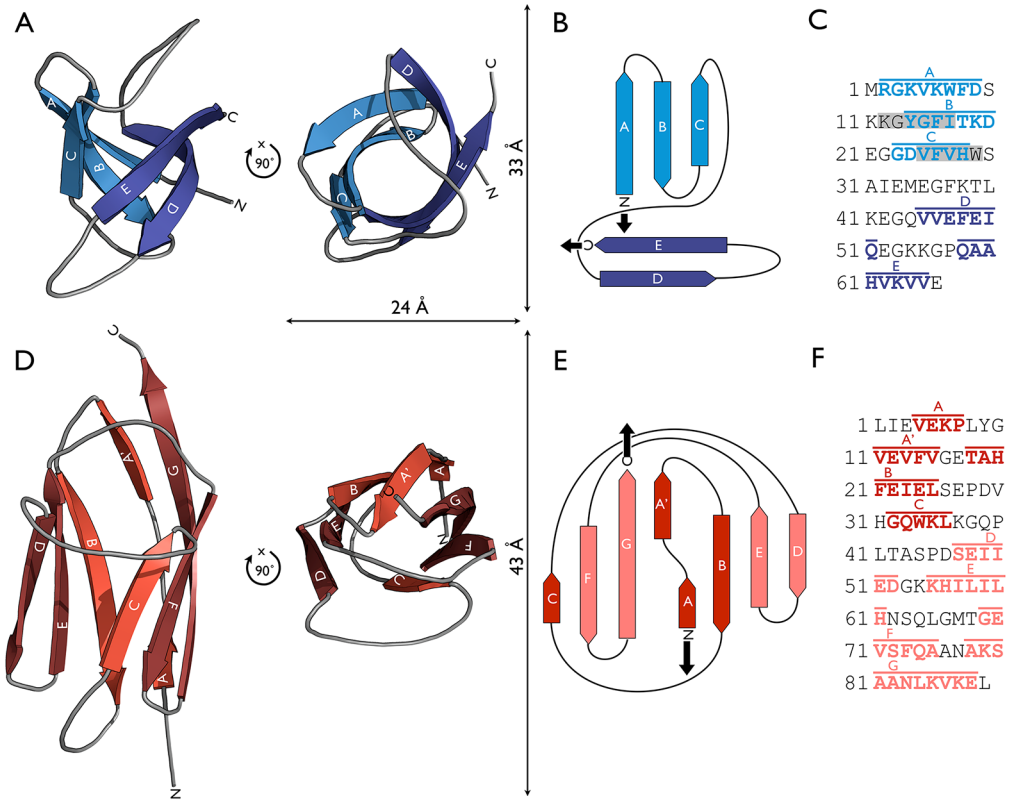


Figure 1. Three-dimensional structure, topology, and sequence of the cold shock protein (Csp) (A, B, C) and I27 (D, E, F). (A, D) The three-dimensional structures of the proteins (Csp PDB accession number: 1G6P[27] and I27 PDB accession number: 1TIT[52]) are visualized as a ribbon diagram that shows β-strands as arrows. Loops and strands have been smoothed for simplicity. In Csp, five antiparallel β-strands (A to F) are arranged in two β-sheets [ABC (light blue) and DE (dark blue)] to form a Greek-key β-barrel. (B, E) The topologies of Csp and I27 are shown, with black arrows indicating the direction of the applied force when concatenated into a polyprotein. (C, F) The linear amino acid sequence of the 66-residue-long peptide chain of Csp and the 89-residue-long peptide chain of I27. The nucleotide-binding motifs in Csp are box shaded in gray. The coloring of β-strand residues and their respective overlined labels follows the same scheme as described above.

Given the prevalence of Csps in all branches of life and the detailed structural characterization of OB-fold proteins, we were interested to determine whether these proteins possess mechanical stability. Does an OB-fold confer mechanical robustness? We were motivated by studies over the past decade, which have pinpointed the interactions and structural features that are responsible for the mechanical stability of proteins.[6] Previous studies of OB-fold proteins have so far been limited to staphylococcal nuclease from *Staphylococcus aureus*,[25] which was found to have a relatively low unfolding force (F_U = 26.3 ± 0.5 pN at a pulling velocity of 200 nm s^{-1}), despite its rich β-strand content. The authors suggest that this mechanical weakness could be attributed to the α-helical segments found near the C-terminus of the protein.

To begin our studies on this interesting protein family, we selected cold shock protein B from *Thermatoga maritima*, which we will refer to in the following as Csp. This small monomeric protein (M_w = 7.5 kDa, Figure 1) has been well characterized structurally and thermodynamically,[22,26−31] making it a good model OB-fold protein for our studies. Figure 1A and B shows the three-dimensional structure and topology of Csp, illustrating the arrangement of the five antiparallel β-strands that form the β-barrel lacking an N-terminal α-helix.

We have completed single-molecule force spectroscopy experiments on Csp, using a specially constructed chimeric polyprotein comprising Csp and I27 domains. The design of a chimeric polyprotein enables the unambiguous identification of the mechanical properties of Csp, due to the presence of the mechanically well-characterized I27 domain[12] (Figures 1D−F). We identify a clear mechanical fingerprint for Csp and determine the mechanical stability of this β-barrel protein. Moreover, measuring the dependence of the unfolding force on pulling velocity, in combination with Monte Carlo simulations, gives us access to further important parameters of the unfolding energy landscape. By establishing a basic set of parameters that describe the energy landscape of Csp, we have laid the foundation for future experiments on the folding/unfolding dynamics of cold shock proteins and other OB-fold proteins.

METHODS

Protein Engineering. DNA manipulations followed standard protocols[32] using enzymes obtained from NEB (MA, USA) and were confirmed by sequencing (Eurofins MWG Operon, Ebersberg, Germany). The cold shock protein (Csp) from *Thermotoga maritima* (UniProtKB/Swiss-Prot accession number: O54310) was reverse translated to optimize codon usage for *Escherichia coli*. ATG start and TGA stop codons were omitted, and *Sac*I and *Apa*I restriction sites were added to the 5′ and 3′ end of the DNA, respectively. The designed DNA sequence was then synthesized and cloned into a *pCR2.1* vector by Eurofins MWG Operon. To generate a chimeric polyprotein, Csp was subcloned into the bacterial expression vector *pET3a-(I27)*$_7$, derived from a pET3a-(I27-ProtL)$_3$-I27 construct,[12] using the added restriction sites *Sac*I

dx.doi.org/10.1021/jp310442s | J. Phys. Chem. B 2013, 117, 1819−1826

and *Apa*I. This replaces the fourth I27 insert in the $(I27)_7$ polymeric gene construct with Csp. The procedure results in a construct that encodes a His-tagged $(I27)_3$-Csp-$(I27)_3$ polyprotein. Then, using PCR, Csp inserts flanked by specific restriction sites were generated. In subsequent cloning steps, these Csp inserts successively replaced the second and sixth *I27* genes in $(I27)_3$-Csp-I27 to generate a final *Csp* triple-insertion construct: $(I27\text{-}Csp)_3$-I27.

Protein Expression and Purification. The constructed expression vector was transformed into *E. coli* BLR [DE3] pLysS (Novagen, Nottingham, UK). Ten liters of LB medium (10 × 1 L in 2 L flasks, supplemented with 25 μg/mL chloramphenicol and 100 μg/mL carbenicillin) was inoculated from a 100 mL overnight starter culture (LB medium as above, at 37 °C in an incubator shaker, Stuart Scientific, UK) to obtain a starting cell density of OD_{600} = 0.05. The cultures were then incubated at 37 °C in an incubator shaker (Sanyo Gallenkamp PLC, UK). Protein expression was induced at an OD_{600} = 0.5−0.6 with 1 mM (final concentration) isopropyl β-D-thiogalactopyranoside (IPTG). The cells were harvested by centrifugation 3 h later. After resuspension in lysis buffer (20 mM Tris, 300 mM NaCl, 0.025% (w/v) sodium azide, 0.15% (v/v) triton X-100, 5 mM imidazole, pH 8), the cells were mechanically disrupted under high-flow pressure (TS series Cell Disrupter, Constant Systems Ltd., Warwick, UK) to release the expressed polyprotein. The histidine-tagged polyprotein was isolated from the cleared lysate using a histidine-binding Ni-NTA affinity chromatography resin (Ni sepharose HP, GE Healthcare, Sweden). The eluted protein was dialyzed into distilled deionized water and freeze-dried. The protein was further purified by size-exclusion chromatography. Here, 5 mL of the protein resuspended in 63 mM sodium phosphate, pH 7.4, was separated by application to a 320 mL HiLoad 26/60 Superdex 75 column (GE Healthcare, Sweden). Fractions containing purified polyprotein were pooled and dialyzed against distilled deionized water before being aliquoted and stored at −20 °C.

UV Fluorescence Emission Spectroscopy. Purified and lyophilized proteins were resuspended in 63 mM sodium phosphate buffer (pH 7.4) in the presence and absence of 6 M guanidinium hydrochloride (GdHCl). Concentrations were determined by measuring the absorbance at 280 nm using extinction coefficients calculated by the *ProtParam* algorithm (http://web.expasy.org/protparam/[33]). Fluorescence spectroscopy was performed on a Photon Technology International QM-1 spectrofluorimeter (PTI, UK) at 20 °C using a 1 mL quartz cuvette (Hellma, UK). The fluorescence emission spectra of monomeric I27 and CSP and the $(I27\text{-}Csp)_3$-I27 polyprotein were recorded between 305 and 375 nm after excitation with 280 nm wavelength UV light.

Single-Molecule AFM Experiments. Single-molecule force spectroscopy experiments were performed using a custom-built atomic force microscope similar in design to that described previously.[34] Silicon nitride cantilevers (MLCT) were obtained from Veeco (Santa Barbara, CA). The spring constant of each cantilever was found in buffer by applying the equipartition theorem[35] and was typically found to be 32 (±1.8) pN nm^{-1}. Protein (0.1 mg) was reconstituted to 0.5 mg mL^{-1} in sterile sodium phosphate buffer (63 mM, pH 7.4) and centrifuged for 5 min at 14 500*g* (Espresso Personal Microcentrifuge, Thermo Scientific, UK). The protein solution (40 μL) was applied onto a coverslip with a freshly stripped gold surface that resulted in the immobilization of polyproteins via covalent attachment between the sulphydryl groups of two cysteine residues at the C-terminus of the polyprotein and the gold surface. After incubation for 15−30 min, the surface was rinsed with buffer. Mechanical unfolding experiments were performed at pulling velocities of 100, 200, 400, 1000, and 2000 nm s^{-1} at a room temperature of 23 °C over a distance of 400 nm. Three data sets, each containing at least 27 unfolding events, were accumulated at each pulling velocity using a fresh sample and a new cantilever for each data set.

Analysis of Single-Molecule Force Spectra. Single-molecule force spectra were analyzed using in-house software written for Igor PRO (Version 6.0, Wavemetrics, Lake Oswego, OR). Data were filtered to include traces where only one polyprotein chain was observed to unfold, characterized by the presence of seven or less unfolding events, which had a minimum of two I27 unfolding events. Any force−extension traces displaying nonspecific unbinding events at high force or other sources of noise were rejected. Model-free analysis was carried out by measuring the peak unfolding force of each protein unfolding event and the interpeak distance between unfolding events. This is defined as the distance from one unfolding peak to the same force value on the following sawtooth curve.[12] To obtain an estimate of the number of amino acids that are released from a condensed native state to an extended unfolded state upon each unfolding event, the data were analyzed by fitting a worm-like-chain (WLC) model for polymer elasticity[36] to the rising edge of each sawtooth.

Monte Carlo Simulations. Monte Carlo simulations were performed using a two-state model for unfolding, for both I27 and Csp, using software developed previously for heteropolyproteins.[12] Briefly, the simulations were used to generate histograms of unfolding forces for both protein domains. At each pulling velocity, the values of k_U, the unfolding rate constant at zero force, and Δx_U, the distance between the folded state and the transition state, were optimized sequentially until the values of median force and standard deviation of the peak unfolding force histograms matched those obtained experimentally, and the Chi squared value of the simulated peak unfolding forces compared with the experimentally obtained forces was equal to 1. The pair of k_U and Δx_U values which gave the best fit to each pulling velocity was thus obtained. The average values of k_U and Δx_U over all velocities were then used to give a global fit to the data.

RESULTS AND DISCUSSION

Construction of a Chimeric Polyprotein to Identify the Mechanical Fingerprint of Cold Shock Protein. To study the mechanical unfolding properties of Csp using AFM force spectroscopy, we have followed a polyprotein-based approach, which has been successfully employed previously.[12,34,37−39] We have used protein engineering techniques to assemble a chimeric polyprotein composed of repeats of I27-Csp. This engineered chimera, consisting of four I27 domains interdigitated with three Csp domains ($(I27\text{-}Csp)_3$-I27, Figure 3B), yields AFM mechanical unfolding data that contain the uncertain mechanical unfolding profile of Csp together with the well-characterized unfolding fingerprint for I27 domains.[12] This construct thus allows the use of I27 unfolding events as an easily recognizable marker of the unfolding of the polyprotein chain.

To confirm that the Csp domains are stable and remain folded in the context of a polyprotein, we analyzed the purified polyprotein (see Methods) by fluorescence emission spectros-

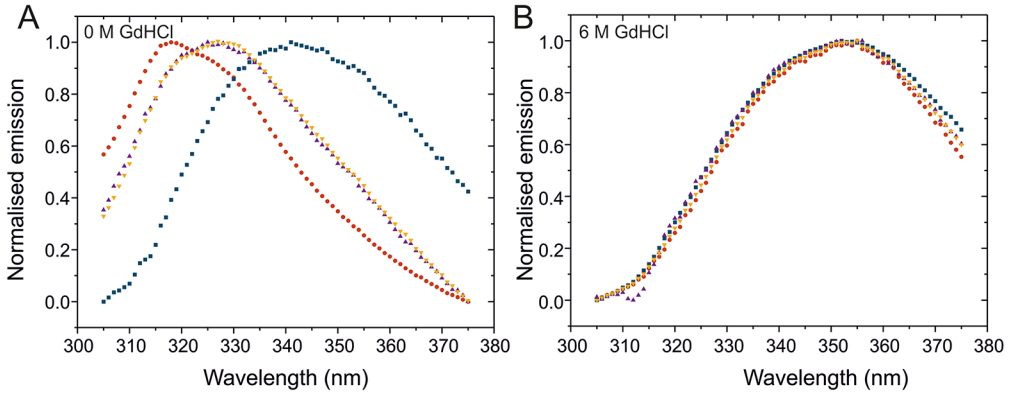


Figure 2. Normalized fluorescence emission spectra of I27 monomers (red circles), Csp monomers (blue squares), $(I27\text{-}Csp)_3$-I27 polyprotein (violet upward pointing triangles), and a simulated $(I27\text{-}Csp)_3$-I27 signal (the sum of 4/7 and 3/7 of the emission spectra of monomeric I27 and Csp, respectively, yellow downward pointing triangles) in the absence (A) and presence of 6 M GdHCl.

copy. The spectra of monomeric I27 and Csp (Figure 2A) have profiles and emission maxima identical to those reported previously.[30,40] The spectrum of the chimeric polyprotein (Figure 2A, violet triangles) resembles an intermediate spectrum of both monomers (Figure 2A, yellow triangles) and is distinct to the denatured state spectra for either the monomeric domains or the polyprotein (Figure 2B). The results indicate that, under the chosen experimental conditions for single-molecule force spectroscopy AFM studies, both Csp and I27 domains are stable and folded.

Single-Molecule Force Spectroscopy Reveals That the Cold Shock Protein Has Mechanical Stability. We use a single-molecule force spectroscopy AFM approach to mechanically unfold the polyprotein construct. A simplified schematic of the AFM setup is shown in Figure 3B. Stretching $(I27\text{-}Csp)_3$-I27 at a constant pulling speed of 200 nm s^{-1} results in a sawtooth-like unfolding pattern, an example of which is shown in Figure 3A. Each peak in the sawtooth pattern corresponds to the unfolding of a single domain of either Csp or I27. While the polyprotein chain unfolds stochastically, the order in which each domain type will unfold is governed by the number of each domain type and the unfolding energy landscape.[5,10] From the force–extension traces we measure the interpeak distance of each unfolding event (x_{p2p}), defined as the distance from one unfolding peak to the same force value on the following curve, and the peak unfolding force (F_U). We measure x_{p2p} and F_U for a number of force–extension traces and construct histograms for both x_{p2p} and F_U. The x_{p2p} frequency histogram (Figure 3C) displays a bimodal distribution centered at 19.0 and 23.7 nm. The value of 23.7 nm is in accord with previously published values of x_{p2p} for I27[12] allowing the assignment of the lower value to the unfolding of Csp. While useful for rapidly identifying the protein domain that unfolded at each event, x_{p2p} cannot be used to estimate the increase in contour length upon unfolding of a domain (ΔL_C), which yields information on the size of the mechanically resistant core of each domain type. Consequently, we used the worm-like chain model to measure ΔL_C, yielding a distribution with two values: 23.5 and 28.0 nm. The value of ΔL_C = 28.0 nm is in accord with previously published values for I27[12] suggesting that the ΔL_C for Csp is 23.5 nm. The contour length increment for Csp can be predicted using the following estimation: each of the 64 amino acids that make up the force-resistant structural portion of Csp contributes about 0.4 nm[41] to the unfolded length of the polyprotein, yielding 25.6 nm. The initial distance between the

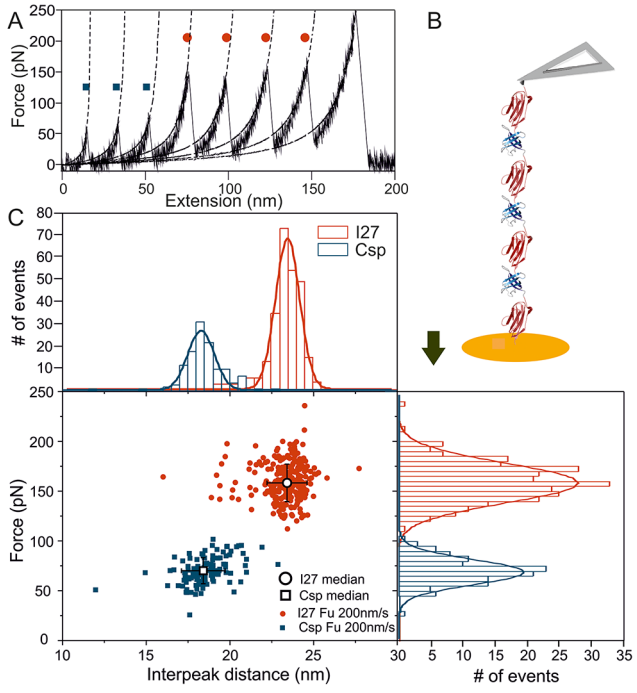


Figure 3. Mechanical unfolding of polyprotein $(I27\text{-}Csp)_3$-I27. (A) Sawtooth force–extension profile that results from the mechanical unfolding of the polyprotein (B) at 200 nm s^{-1}. Dashed lines show worm like chain fits to the data, with a persistence length of 0.4 nm. (B) Schematic showing a single $(I27\text{-}Csp)_3$-I27 molecule (Csp and I27, blue and red, respectively) attached to a gold surface (bottom) and the tip of an AFM cantilever (top). (C) The scatter plots of Csp- and I27-specific unfolding forces and interpeak distances for 79 Csp unfolding events (blue squares) and 123 I27 unfolding events (red circles) are shown in combination with their respective distribution histograms. Error bars on the median values for each data set (open symbols) indicate the standard deviation. Gaussian fits to histograms for each data set are used to obtain a measure of the force and interpeak distance distribution widths.

N and C termini of a single folded Csp domain is 1.4 nm. Thus, the expected increase in contour length of Csp upon its forced unfolding is 25.6 − 1.4 = 24.2 nm, which is close to the value obtained from the worm-like chain analysis. Similarly for the I27 protein, each of the 81 amino acids that make up the force-resistant structural portion of I27[5] contributes about 0.4 nm[41]

dx.doi.org/10.1021/jp310442s | J. Phys. Chem. B 2013, 117, 1819–1826

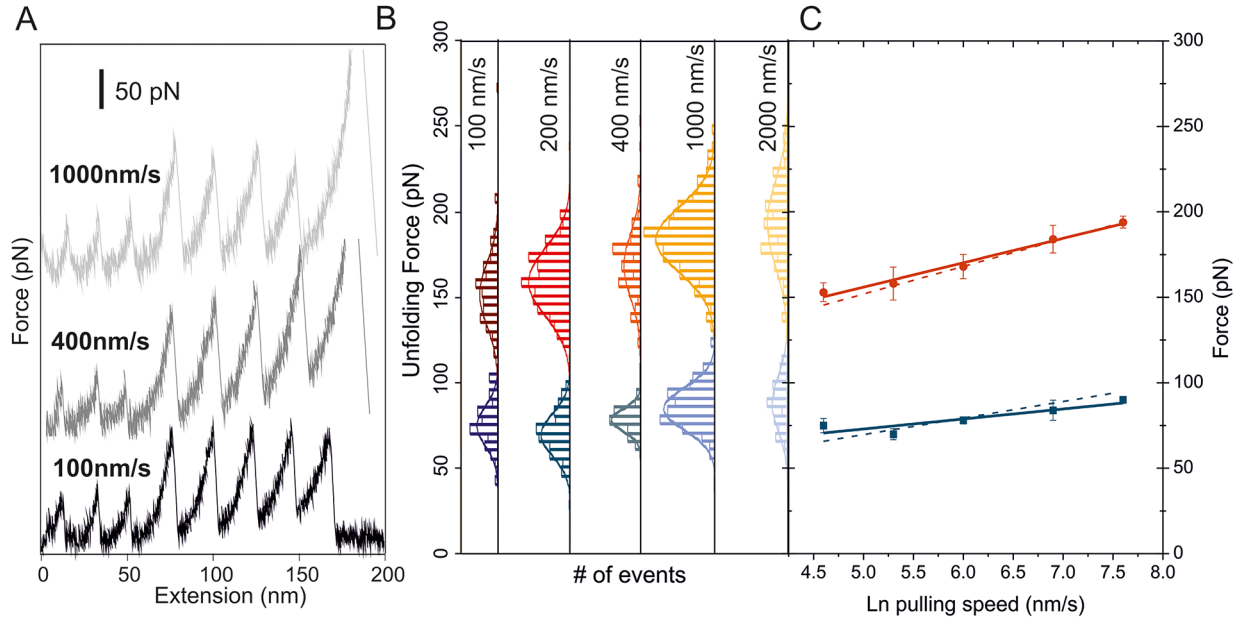


Figure 4. (A) Example sawtooth force−extension traces at three different pulling speeds. (B, C) The dependency of the unfolding force on pulling speed provides the data for Monte Carlo simulations to obtain the unfolding rate constant, k_U, and the distance between native state and transition state Δx_U, where (B) the force distributions are shown as histograms with Gaussian fits used to measure the distribution widths (solid lines), and where (C) the unfolding force for both protein domains increases with the natural logarithm of the pulling speed. Error bars indicate the standard deviation between the three experiments at each pulling speed (Table 1). The solid line is a best fit to the experimental data. Monte Carlo fits to the experimental data are shown as dotted lines.

Table 1. Summary of Mechanical Unfolding Data for $(I27\text{-}Csp)_3$-I27

speed [nm s^{-1}]	# Csp	# I27	median unfolding force CSP [pN] (±SD)	average [pN] (±SE)	median unfolding force I27 [pN] (±SD)	average [pN] (±SE)
100	32	33	72 (±13)	75 (±4)	153 (±26)	153 (±5)
	29	48	74 (±13)		159 (±17)	
	11	25	80 (±13)		148 (±19)	
200	42	62	68 (±15)	70 (±3)	169 (±20)	158 (±10)
	15	24	74 (±10)		154 (±16)	
	22	37	69 (±11)		151 (±17)	
400	12	29	78 (±8)	78 (±2)	176 (±18)	168 (±7)
	8	19	76 (±8)		162 (±18)	
	12	26	80 (±11)		167 (±26)	
1000	43	78	82 (±13)	84 (±6)	177 (±20)	184 (±8)
	42	61	91 (±12)		193 (±22)	
	53	75	80 (±12)		183 (±20)	
2000	12	18	90 (±16)	90 (±1)	194 (±27)	194 (±4)
	19	18	89 (±17)		198 (±23)	
	20	38	90 (±14)		191 (±25)	

to the unfolded length of the polyprotein, yielding 32.4 nm. The initial distance between the N and C termini of a single folded I27 domain is 2.8 nm. Thus, the expected increase in contour length of I27 upon its forced unfolding is 32.4−2.8 = 29.6 nm, which is close to the value obtained from the worm-like chain analysis.

The spacing between each event thus allows each to be assigned to a protein domain type, allowing mechanical unfolding data to be obtained for Csp, with I27 providing a robust mechanical marker protein. The F_U frequency distribution also displays a bimodal distribution centered at 70 and 158 pN at a pulling speed of 200 nm s^{-1} (Figure 3C). The different values of F_U reflect the different forces required to unfold the different protein structures, I27 and Csp. Previous studies on a chimeric polyprotein containing I27 found an F_U of 170 pN[12] at a slightly faster pulling velocity of 270 nm s^{-1}, in agreement with the values we measure. Therefore, we can identify the distribution centered on 78 pN as representing the Csp. The ability to assign unfolding forces to specific domains can be clearly illustrated by examining a scatter plot (Figure 3C) that combines the data for x_{p2p} and F_U and reveals two clear populations of events at a pulling speed of 200 nm s^{-1}. This approach has thus permitted the identification of a clear mechanical fingerprint for Csp, which can be differentiated from the I27 protein. Csp unfolds at a lower force (70 pN versus 158 pN) and has a shorter x_{p2p} and ΔL_C (19 nm vs 24 nm and 23.5 nm vs 28.0 nm, respectively).

Having obtained a clear mechanical fingerprint for Csp, we followed the same procedure to obtain force−extension traces at four other pulling velocities: 100, 400, 1000, and 2000 nm

dx.doi.org/10.1021/jp310442s | J. Phys. Chem. B 2013, 117, 1819−1826

s^{-1} (Figure 4). At each pulling velocity we measured F_U for each unfolding peak and constructed a histogram of F_U (Figure 4B). We fitted this with a Gaussian distribution to obtain a median value of F_U. A summary of the resulting F_U for each of the triplicate experiments completed (see Methods) at each pulling speed is given in Table 1. To examine the pulling speed dependence of F_U for each protein, we plotted the natural logarithm of pulling speed against the median F_U for the two populations observed in the F_U histograms. Figure 4C shows F_U for Csp (blue squares) and I27 (red circles) and demonstrates that as the pulling speed increases the force needed to unfold the protein increases, with the slope of this dependence being steeper for I27 than for Csp.

From the speed dependence of the unfolding force, we can access basic features of the energy landscape of unfolding. The unfolding of a protein under an external force can be described as a lowering of the free energy barrier between the folded and unfolded state of the protein. This increases the likelihood of thermal fluctuations, leading to a transition from the folded to the unfolded state. For a two-state unfolding process, this reduction in the energy barrier is dependent on the magnitude of the applied force and the distance between the free energy barrier and the folded state energy well, as described by the Bell model[42]

$$k(F) = A \exp\left[\frac{-(\Delta G - F\Delta x_U)}{k_B T}\right] \quad (1)$$

where $k(F)$ is the force-dependent rate constant; F is the applied force; A is the attempt frequency; Δx_U is the distance from the folded state to the transition state; k_B is Boltzmann's constant; and T is the temperature.[42] The value of Δx_U is related to the transition state position, the distance between the folded state and the transition barrier on the unfolding pathway. A movement of the transition state toward the unfolded state would result in an increased Δx_U. ΔG_U is the height of the activation energy barrier to unfolding and can be related to the unfolding rate at zero force k_U by the following expression

$$k_U = A \exp\left[\frac{-\Delta G_U}{k_B T}\right] \quad (2)$$

Single-molecule AFM experiments in combination with Monte Carlo (MC) simulations provide access to Δx_U and k_U, allowing features of the underlying energy landscape of proteins to be uncovered.[9,12,25,43] MC simulations of the stretching of the $(I27\text{-}Csp)_3$-I27 constructs were completed to extract k_U and Δx_U for each experimental data set. The MC technique assumes that the polyproteins used contain three Csp and four I27 domains and that these unfold via a two-state all-or-none process, governed by the unfolding rate constant, k_U, and the distance from the native state to the transition state along the measured reaction coordinate, Δx_U. This enables the distribution of forces and number of protein unfolding events in a polyprotein chain to be simulated, producing unfolding force histograms similar to those generated experimentally, for a specified set of parameters. This is completed over the experimental range of pulling speeds, to obtain the pulling speed dependence of the unfolding force F_U (dashed lines, Figure 4C). The resulting values of Δx_U = 0.28 nm and k_U = 0.0015 s^{-1} obtained from the MC simulations for I27 are in agreement with the previously reported values in the literature.[12] For Csp, a higher value of Δx_U = 0.49 nm and k_U = 0.015 s^{-1} was found. Assuming a value of the prefactor[44] $A = 1 \times 10^6$ s^{-1} this gives an activation energy barrier height of 10.7 kcal/mol (18.0 $k_B T$) for Csp, which is slightly smaller than that of 12.0 kcal/mol (20.3 $k_B T$) for I27 (eq 2).

The schematic energy landscape in Figure 5 shows the calculated Δx_U and ΔG_U for the I27 and Csp protein. While the height of the activation energy barrier to unfolding is lower in the Csp, the distance to the unfolding transition state is larger, suggesting that Csp is easier to unfold but less sensitive or more malleable to changes in the pulling velocity. This may imply that a similar amount of work is spread over a larger distance in Csp compared to I27 and that structural changes occur at lower forces.[38,45]

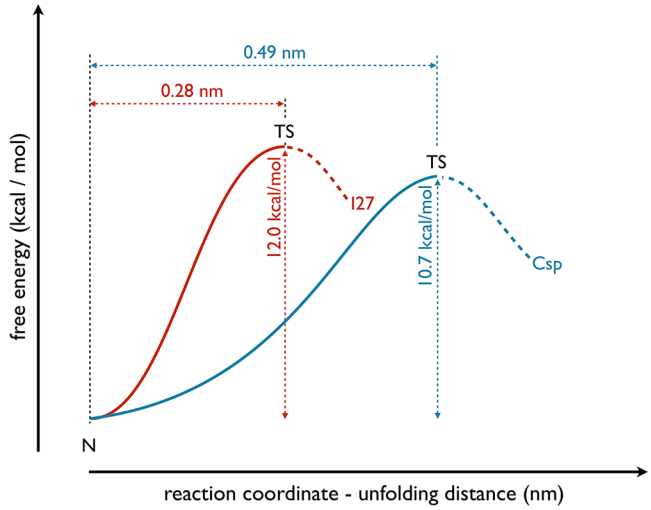


Figure 5. Schematic depiction of the unfolding barriers of I27 and Csp as determined in this study, illustrating the greater distance between native state (N) and transition state (TS) of Csp (blue line) compared to I27 (red line). Lines are dashed after the transition state as this region of the energy landscape was not mapped in our experiments.

CONCLUSIONS

Single-molecule force spectroscopy experiments using a chimeric polyprotein, $(I27\text{-}Csp)_3$-I27, have revealed two populations of mechanical unfolding events. One population corresponds to the previously well-characterized I27 protein. The other reveals the mechanical fingerprint for the Csp, which appears to be an all-or-none unfolding event at the pulling speeds used. The results of our single molecule studies indicate that the structure of Csp can be disrupted by mechanical forces of ~75 pN. While this force is lower than that measured for the I27 protein, it is relatively large compared to other proteins outside the Ig-like and β-grasp families. In addition to lower unfolding forces, MC simulations reveal that the barrier to unfolding of Csp has a relatively long distance to the unfolding state (0.49 nm). This distance is greater than that observed for other all-β-strand proteins previously examined using AFM, including I27 (values are typically in the range 0.1−0.4[5,46−48]). A longer distance to the unfolding transition state is suggestive of a protein conformation that is less sensitive to the pulling force and is more malleable or flexible.[38] Such large values of Δx_U have been reported previously, in particular in the special case of the protein Per-ARNT-Sim, a signal transduction protein, which has an exceptionally large Δx_U of 2.0 nm. This is

thought to be important for the flexibility of the protein during its function.[45] Whether the mechanical properties of Csp discerned here are of functional relevance is, as yet, unclear. Csp is thought to act as a RNA chaperone to facilitate translation by preventing the formation of secondary RNA structures,[49] and the higher flexibility (or native ensemble) may enhance complex formation with single-stranded RNA/DNA molecules. The relatively high mechanical stability (high F_U), alongside relatively easy deformability (high Δx_U), allows the protein to remain in the native energy well under the application of force, which may endow the protein with a robustness to carry out its function over a range of environmental conditions.

As the data set from protein mechanical unfolding experiments has increased, it has become clear that proteins display high mechanical strength when extended from distal ends of β-strands within a β-strand-rich region. Proteins with this type of topology include I27, protein L,[12] GB1,[1] and c7a.[50] In addition to high unfolding forces, these proteins are typically "brittle", with small Δx_U values. The clear relationship between secondary structure topology and unfolding force and the ability to simulate their unfolding has resulted in a detailed understanding of the determinants of mechanical strength for these proteins. For other types of proteins, however, such an understanding is missing due to the lack of experimental data. Here, we have begun to address this by characterizing the mechanical properties of Csp, which can be compared to two similar-sized, β-strand-rich proteins: the tenth fibronectin type III domain from fibronectin (10FNIII) and staphylococcal nuclease (SNase), with the latter being a member of the same OB-fold family.

Csp is similar in size and β-strand content to 10FNIII, which is composed of 94 amino acids. Previous AFM studies have shown that 10FNIII has a F_U of 75 pN at a pulling speed of 600 nm s^{-1},[46] similar to that for Csp. A pulling speed dependence and MC simulation study found a Δx_U = 0.38 nm and a k_U = 0.02 s^{-1}. Again, assuming a prefactor of 1×10^6 s^{-1}, this would result in ΔG_U = 10.49 kcal mol^{-1} (the authors obtain a result of 22.2 kcal mol^{-1}, using the prefactor of 6×10^{12} s^{-1}), suggesting that 10FNIII has a similar activation energy barrier to Csp. However, the smaller value of Δx_U for 10FNIII than for Csp suggests that the latter protein is less sensitive to pulling speed, or more malleable. Interestingly, steered molecular dynamics simulations of 10FNIII[51] suggest that two energy barriers have to be overcome to unravel the protein. First, the two opposing β-sheets have to be untwisted and aligned (the first energy barrier) before adjacent β-strands can break apart (the second energy barrier). The existence of a similar intermediate unfolded state in Csp cannot be excluded with the data reported here. If such an intermediate state occurs, it may be mechanically very weak and not observable with our current methods.

To date, staphylococcal nuclease (SNase) is the only other OB-fold containing protein that has been mechanically unfolded, using a similar AFM force spectroscopy approach.[25] This study found that SNase unfolded in an all-or-none fashion, with an F_U of 26 pN at a pulling speed of 200 nm s^{-1}. The Csp in the present study requires about three times the force to unfold at the same pulling speed. Interestingly, contrary to SNase, Csp lacks a C-terminal α-helical region. This suggests that the applied force acts directly on the β-barrel of the Csp, resulting in a higher mechanical stability. Indeed, the authors of the SNase study suggested that the surprisingly low unfolding force, given the high β-sheet content of SNase, may result from easy detachment of the two C-terminal α-helices of the SNase protein.[25] This action could then expose the β-barrel so that the interactions connecting the β-strands can be unzipped, thus lowering the activation energy barrier height. This is in contrast to the more resilient shear topology found in mechanically stronger proteins, such as the I27 protein, which exhibit a mechanically robust mechanical clamp composed of interactions which must be broken simultaneously by the shearing apart of the β-sheet.

In this study we have shown that a small cold shock protein combines mechanical robustness with flexibility. Its simple topology and ancient origin provide an attractive model protein for further single-molecule force spectroscopy studies. In the future, the further analysis of Csp, its homologues, and designed mutants in combination with varying experimental conditions will help to better understand how this particular family of protein folds achieves its mechanical stability and how structure and stability relate to its function.

AUTHOR INFORMATION

Corresponding Author

*E-mail: l.dougan@leeds.ac.uk.

Author Contributions

[§]These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Dr. Lorna Dougan is supported by a grant from the European Research Council (258259-EXTREME BIOPHYSICS). We would like to thank Robert Schiffrin (University of Leeds) for his help in generating the $(I27\text{-}Csp)_3$-I27 chimeric polyprotein construct. We would like to thank all members of the Dougan group for useful discussions and feedback.

REFERENCES

(1) Cao, Y.; Yoo, T.; Li, H. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11152−11157.

(2) Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, J. M.; Gaub, H. E. *Science* **1997**, *276*, 1109−1112.

(3) Carrion-Vazquez, M.; Oberhauser, A. F.; Fowler, S. B.; Marszalek, P. E.; Broedel, S. E.; Clarke, J.; Fernandez, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3694−3699.

(4) Dougan, L.; Koti, A. S. R.; Genchev, G.; Lu, H.; Fernandez, J. M. *ChemPhysChem* **2008**, *9*, 2836−2847.

(5) Brockwell, D. J.; Beddard, G. S.; Clarkson, J.; Zinober, R. C.; Blake, A. W.; Trinick, J.; Olmsted, P. D.; Smith, D. A.; Radford, S. E. *Biophys. J.* **2002**, *83*, 458−72.

(6) Hoffmann, T.; Dougan, L. *Chem. Soc. Rev.* **2012**, *41*, 4781−4796.

(7) Lu, W.; Negi, S. S.; Oberhauser, A. F.; Braun, W. *Proteins* **2012**, *80*, 1308−15.

(8) Li, H.; Cao, Y. *Acc. Chem. Res.* **2010**, *43*, 1331−41.

(9) Galera-Prat, A.; Gómez-Sicilia, A.; Oberhauser, A. F.; Cieplak, M.; Carrión-Vázquez, M. *Curr. Opin. Struct. Biol.* **2010**, *20*, 63−69.

(10) Brockwell, D. J.; Paci, E.; Zinober, R. C.; Beddard, G. S.; Olmsted, P. D.; Smith, D. A.; Perham, R. N.; Radford, S. E. *Nat. Struct. Biol.* **2003**, *10*, 731−737.

(11) Carrion-Vazquez, M.; Li, H.; Lu, H.; Marszalek, P. E.; Oberhauser, A. F.; Fernandez, J. M. *Nat. Struct. Mol. Biol.* **2003**, *10*, 738−743.

(12) Sadler, D. P.; Petrik, E.; Taniguchi, Y.; Pullen, J. R.; Kawakami, M.; Radford, S. E.; Brockwell, D. J. *J. Mol. Biol.* **2009**, *393*, 237−248.

(13) Ng, S. P.; Billings, K. S.; Ohashi, T.; Allen, M. D.; Best, R. B.; Randles, L. G.; Erickson, H. P.; Clarke, J. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9633−7.
(14) Guzmán, D. L.; Randall, A.; Baldi, P.; Guan, Z. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1989−94.
(15) Balamurali, M. M.; Sharma, D.; Chang, A.; Khor, D.; Chu, R.; Li, H. *Protein Sci.* **2008**, *17*, 1815−26.
(16) Garcia, T. I.; Oberhauser, A. F.; Braun, W. *Proteins* **2009**, *75*, 706−18.
(17) Li, H. B.; Linke, W. A.; Oberhauser, A. F.; Carrion-Vazquez, M.; Kerkviliet, J. G.; Lu, H.; Marszalek, P. E.; Fernandez, J. M. *Nature* **2002**, *418*, 998−1002.
(18) Sharma, D.; Perisic, O.; Peng, Q.; Cao, Y.; Lam, C.; Lu, H.; Li, H. B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9278−9283.
(19) Arcus, V. *Curr. Opin. Struct. Biol.* **2002**, *12*, 794−801.
(20) Graumann, P. L.; Marahiel, M. A. *Trends Biochem. Sci.* **1998**, *23*, 286−290.
(21) Nelson, K. E.; Clayton, R. A.; Gill, S. R.; Gwinn, M. L.; Dodson, R. J.; Haft, D. H.; Hickey, E. K.; Peterson, J. D.; Nelson, W. C.; Ketchum, K. A.; et al. *Nature* **1999**, *399*, 323−329.
(22) Wassenberg, D.; Welker, C.; Jaenicke, R. *J. Mol. Biol.* **1999**, *289*, 187−193.
(23) Nettels, D.; Hoffmann, A.; Schuler, B. *J. Phys. Chem. B* **2008**, *112*, 6137−6146.
(24) Horn, G.; Hofweber, R.; Kremer, W.; Kalbitzer, H. *Cell. Mol. Life Sci.* **2007**, *64*, 1457−1470.
(25) Wang, C.-C.; Tsong, T.-Y.; Hsu, Y.-H.; Marszalek, P. E. *Biophys. J.* **2011**, *100*, 1094−1099.
(26) Jung, A.; Bamann, C.; Kremer, W.; Kalbitzer, H. R.; Brunner, E. *Protein Sci.* **2004**, *13*, 342−350.
(27) Kremer, W.; Schuler, B.; Harrieder, S.; Geyer, M.; Gronwald, W.; Welker, C.; Jaenicke, R.; Kalbitzer, H. R. *Eur. J. Biochem.* **2001**, *268*, 2527−39.
(28) Motono, C.; Gromiha, M. M.; Kumar, S. *Proteins* **2008**, *71*, 655−669.
(29) Phadtare, S.; Hwang, J.; Severinov, K.; Inouye, M. *Genes Cells* **2003**, *8*, 801−810.
(30) Welker, C.; Böhm, G.; Schurig, H.; Jaenicke, R. *Protein Sci.* **1999**, *8*, 394−403.
(31) Perl, D.; Welker, C.; Schindler, T.; Schroder, K.; Marahiel, M. A.; Jaenicke, R.; Schmid, F. X. *Nat. Struct. Mol. Biol.* **1998**, *5*, 229−235.
(32) Sambrook, J.; Russell, D. *Molecular Cloning: A Laboratory Manual*, 3rd Revised ed.; Cold Spring Harbor Laboratory Press: U.S., 2000.
(33) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S. e.; Wilkins, M. R.; Appel, R. D.; Bairoch, A.; Walker, J. M., Ed.; Humana Press: New York, 2005; pp 571−607.
(34) Dougan, L.; Li, J.; Badilla, C. L.; Berne, B. J.; Fernandez, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 12605−12610.
(35) Florin, E. L.; Rief, M.; Lehmann, H.; Ludwig, M.; Dornmair, C.; Moy, V. T.; Gaub, H. E. *Biosens. Bioelectron.* **1995**, *10*, 895−901.
(36) Marko, J. F.; Siggia, E. D. *Macromolecules* **1995**, *28*, 8759−8770.
(37) Steward, A.; Toca-Herrera, J. L.; Clarke, J. *Protein Sci.* **2002**, *11*, 2179−2183.
(38) Schlierf, M.; Rief, M. *J. Mol. Biol.* **2005**, *354*, 497−503.
(39) Peng, Q.; Zhuang, S.; Wang, M.; Cao, Y.; Khor, Y.; Li, H. *J. Mol. Biol.* **2009**, *386*, 1327−1342.
(40) Blake, A. W., PhD Thesis, University of Leeds, 2004.
(41) Carrion-Vazquez, M.; Oberhauser, A. F.; Fisher, T. E.; Marszalek, P. E.; Li, H.; Fernandez, J. M. *Prog. Biophys. Mol. Biol.* **2000**, *74*, 63−91.
(42) Bell, G. I. *Science* **1978**, *200*, 618−627.
(43) Sikora, M.; Sulkowska, J. I.; Witkowski, B. S.; Cieplak, M. *Nucleic Acids Res.* **2011**, *39*, D443−D450.
(44) Li, M. S.; Klimov, D. K.; Thirumalai, D. *Polymer* **2004**, *45*, 573−579.
(45) Gao, X.; Qin, M.; Yin, P.; Liang, J.; Wang, J.; Cao, Y.; Wang, W. *Biophys. J.* **2012**, *102*, 2149−2157.
(46) Oberhauser, A. F.; Badilla-Fernandez, C.; Carrion-Vazquez, M.; Fernandez, J. M. *J. Mol. Biol.* **2002**, *319*, 433−47.
(47) Li, H.; Fernandez, J. M. *J. Mol. Biol.* **2003**, *334*, 75−86.
(48) Oberhauser, A. F.; Marszalek, P. E.; Erickson, H. P.; Fernandez, J. M. *Nature* **1998**, *393*, 181−185.
(49) Schroeder, R.; Barta, A.; Semrad, K. *Nat. Rev. Mol. Cell. Biol.* **2004**, *5*, 908−919.
(50) Valbuena, A.; Oroz, J.; Hervas, R.; Vera, A. M.; Rodriguez, D.; Menendez, M.; Sulkowska, J. I.; Cieplak, M.; Carrion-Vazquez, M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 13791−13796.
(51) Craig, D.; Krammer, A.; Schulten, K.; Vogel, V. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5590−5595.
(52) Improta, S.; Politou, A. S.; Pastore, A. *Structure* **1996**, *4*, 323−337.