The Effect of Core Destabilization on the Mechanical Resistance of I27

David J. Brockwell,* Godfrey S. Beddard,† John Clarkson,‡ Rebecca C. Zinober,‡ Anthony W. Blake,* John Trinick,§ Peter D. Olmsted,‡ D. Alastair Smith,‡ and Sheena E. Radford*

*School of Biochemistry and Molecular Biology, ‡Department of Physics and Astronomy, †School of Chemistry, and §School of Biomolecular Sciences, University of Leeds, United Kingdom

ABSTRACT It is still unclear whether mechanical unfolding probes the same pathways as chemical denaturation. To address this point, we have constructed a concatamer of five mutant I27 domains (denoted $(I27)_5$*) and used it for mechanical unfolding studies. This protein consists of four copies of the mutant C47S, C63S I27 and a single copy of C63S I27. These mutations severely destabilize I27 ($\Delta\Delta G_{UN}$ = 8.7 and 17.9 kJ mol^{-1} for C63S I27 and C47S, C63S I27, respectively). Both mutations maintain the hydrogen bond network between the A′ and G strands postulated to be the major region of mechanical resistance for I27. Measuring the speed dependence of the force required to unfold $(I27)_5$* in triplicate using the atomic force microscope allowed a reliable assessment of the intrinsic unfolding rate constant of the protein to be obtained (2.0×10^{-3} s^{-1}). The rate constant of unfolding measured by chemical denaturation is over fivefold faster (1.1×10^{-2} s^{-1}), suggesting that these techniques probe different unfolding pathways. Also, by comparing the parameters obtained from the mechanical unfolding of a wild-type I27 concatamer with that of $(I27)_5$*, we show that although the observed forces are considerably lower, core destabilization has little effect on determining the mechanical sensitivity of this domain.

INTRODUCTION

Since its conception in 1986 (Binnig et al., 1986), the atomic force microscope (AFM) has been used for a wide variety of imaging applications in material science, chemistry, and biology. More recently, the imaging capability of AFM has been complimented by the development of force mode AFM (Burnham and Colton, 1989). By using a cantilever of known stiffness (spring constant) the force applied by the cantilever can be calculated by measurement of its deflection. This allows picoNewton force sensitivity coupled with Ångstrom distance resolution. This technique has been used for the direct measurement of the binding forces of complimentary strands of DNA (Lee et al., 1994), the binding energy of receptor:ligand complexes (Florin et al., 1994), and over larger distances, to measure conformational changes in organic polymers (Rief et al., 1997a).

The effect of applied force upon the energy landscape of protein domains is of particular interest in the context of protein folding and unfolding. The use of force as a protein denaturant was initially attempted by differential chemical derivitization of the tip and substrate (Mitsui et al., 1996). However, these results were difficult to interpret due to the complication of surface effects and identification of true unfolding events. These problems were obviated by the use of the giant muscle protein titin. This modular protein (3 MDa) consists mainly of ~300 immunoglobulin (I-set) and fibronectin type III domains and has been suggested to act as a molecular spring, responsible for the passive tension generated by extended muscle and for maintaining myosin filaments in the middle of the sarcomere (Trinick, 1996). By using optical tweezers (Kellermayer et al., 1997, Tskhovrebova et al., 1997) or AFM (Rief et al., 1997b), it was shown that individual domains in the titin polymer unfold in an all-or-none manner. When measured by AFM (extension rates ~10–10000 nms^{-1}) these domains unfolded at a force of the order of hundreds of picoNewtons, producing a characteristic "saw-tooth" pattern of force versus extension. The method of mechanical unfolding has also been applied to several other naturally occurring modular "beads on a string" proteins: tenascin (Oberhauser et al., 1998), spectrin (Rief et al., 1999), fibronectin (Oberdörfer et al., 2000), and abalone shell protein (Smith et al., 1999). In a similar approach, the sequential abstraction of individual bacteriorhodopsin α-helices from the membrane of *Halobacterium salinarium* has been observed (Oesterhelt et al., 2000). These studies have been important in that they have shown, for example, that the predicted pulling speed dependence of the unbinding force of ligand:receptors (Merkel et al., 1999) is applicable to forced protein unfolding.

The amount of information that can be extracted from mechanical unfolding of natural protein polymers is limited by the heterogeneity of the sample, coupled with the fact that the point of tip and substrate attachment is unknown. Molecular biology has enabled the construction of artificial polyproteins comprising 8 to 12 copies of a single domain joined by amino-acid linkers (Carrion-Vazquez et al., 1999a), or by disulphide bridges (Yang et al., 2000). The 27th immunoglobulin domain of the I band of human cardiac titin (Fig. 1) has become a paradigm for mechanical unfolding, being used as a model system for both experiment (Carrion-Vazquez et al., 1999a) and theoretical studies (Lu et al., 1998). First, experiments on this homogeneous system showed unambiguously that each saw-tooth in the

Submitted November 2, 2001, and accepted for publication March 13, 2002.

Address reprint requests to Sheena E. Radford, School of Biochemistry and Molecular Biology, University of Leeds, Leeds, LS2 9JT UK. Tel.: 44-113-343-3170; Fax: 44-113-343-3167; E-mail: s.e.radford@leeds.ac.uk.

0006-3495/02/07/458/15 $2.00

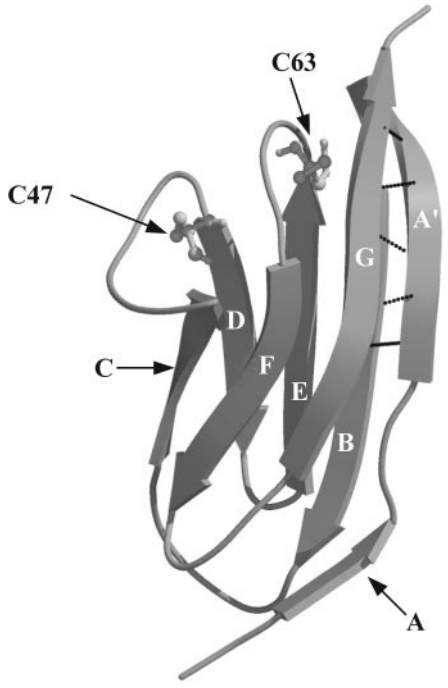

FIGURE 1 Nuclear magnetic resonance solution structure of monomeric I27. C47 and C63 are shown in ball and stick representation, and the hydrogen bonds between the A′ and G strands are shown as dashed lines. The figure was drawn using Molscript (Kraulis, 1991 and Raster3D (Merritt and Murphy, 1994) using the co-ordinates from the protein data base file 1TIT (Improta et al., 1996). Individual β-strands are labeled *A* to *G*.

force distance curve relates to a single-domain unfolding event. Second, the extension rate dependence of the unfolding force (which had been observed previously, Merkel et al., 1999) and the time dependence of the probability of folding allowed the height of the unfolding and folding energy barriers (ΔG_u and ΔG_f) and the placement of the mechanical unfolding transition state on the reaction coordinate (distance) to be determined. Third, and most importantly, the intrinsic unfolding rate constants obtained by classical chemical denaturation and mechanical unfolding were reported to be very similar (4.9×10^{-4} s^{-1} and 3.3×10^{-4} s^{-1}, respectively (Carrion-Vazquez et al., 1999a)). As well as a similarity in barrier height, a similar transition-state placement on the reaction coordinate was also observed (~10% from the native state), suggesting that mechanical and chemical denaturation probe the same unfolding process, at least for this domain.

Steered molecular dynamics simulations (Lu et al., 1998; Lu and Schulten, 2000) have suggested that the occurrence of large unfolding forces in I27 results from the rupture of six hydrogen bonds between the A′ and G strands, which need to be broken before the rest of the protein can be exposed to the force (Fig. 1). Recent mechanical unfolding experiments using proline mutagenesis and loop insertions have supported the suggestion that the A′-G interface acts as a mechanical clamp that resists the applied force (Li et al., 2000a; Carrion-Vazquez et al., 1999b). The dependence of mechanical stability upon the presence of specific, highly localized hydrogen bond "clamps" and their geometry relative to the applied force, is clearly at odds with the proposition that the chemical and mechanical unfolding pathways for this domain are identical. Indeed, chemical denaturation experiments have shown that although the A′ and G strands are disrupted in the transition state for unfolding, other regions of the protein are also significantly perturbed (Fowler and Clarke, 2001). To determine whether chemical and mechanical unfolding of I27 monitor similar or distinct processes, we have constructed a concatamer of five mutated I27 domains. Unique restriction sites, enabling facile incorporation of domains for future mechanical unfolding studies, link the domains. This protein consists of four copies of the double C47S, C63S mutant and a single copy of C63S I27 as the central domain. (The single copy of C63S was incorporated as the central domain as part of ongoing studies that require the labeling of this domain with fluorophores for single molecule fluorescence measurements during mechanical unfolding.) Both of these mutations maintain the hydrogen bond network between the A′ and G strands and, on the basis of the mechanical clamp model, would not be expected to affect the observed mechanically induced unfolding forces, irrespective of their effect on the thermodynamic and kinetic stability of the domains measured chemically.

MATERIALS AND METHODS

All polymerase chain reaction (PCR) protocols were performed using Vent polymerase (New England Biolabs, Hitchen, Herts., UK) under empirically derived melting temperatures and Mg^{2+} concentrations. Restriction endonucleases were purchased from New England Biolabs Ltd (UK). All other enzymes were from Promega (Southampton, Hants, UK).

Engineering of monomeric C63S I27 and C47S, C63S I27

A His-tag modified pET8c vector containing the wild-type I27 gene was obtained from M. Gautel (EMBL, Heidelberg, Germany (Improta et al., 1998)). The single mutant C63S I27 was constructed by Megaprimer PCR (Barik, 1996) using pET8cI27 as template with forward and reverse primers containing the *Xho*I and *Mlu*I endonuclease restriction sites found in the modified pET8c vector. After isolation the PCR product was digested with *Xho*I and *Mlu*I and directly ligated into the modified pET8c vector, which had been predigested with *Xho*I and *Mlu*I and dephosphorylated with calf intestinal alkaline phosphatase. The presence of the mutation was verified by automated sequencing. The double mutant C47S, C63S I27 was generated by a second round of Megaprimer PCR.

Construction of $(I27)_5$* concatamer

The concatamer was constructed using a PCR-generated cassette strategy. The design of the concatamer is shown in Fig. 2 *a*. Each I27 domain was

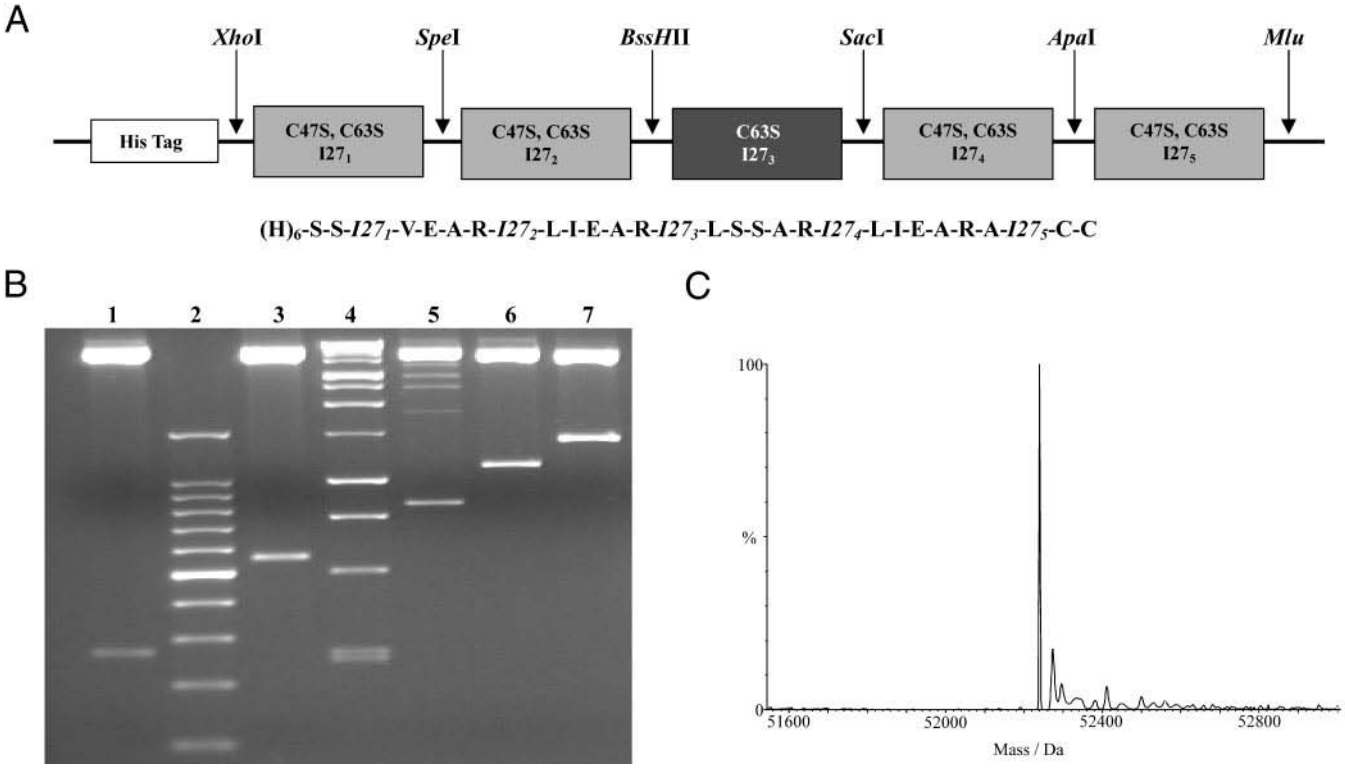

FIGURE 2 (*a*) Cassette design of $(I27)_5$* at the DNA level. Each cassette is defined by a unique pair of restriction endonuclease sites that are shown above the construct. The amino acid composition of the interdomain linkers are shown below the construct. The hexa-histidine tag facilitates purification in two steps. Two C-terminal cysteines allow attachment of the protein to the gold surface. The central domain $(I27)_3$ is a single mutant (C63S), whereas the others are double mutants (C47S, C63S). (*b*) Restriction digest of $(I27)_3$* with different pairs of restriction enzymes. (*Lane 1*) Domain 1 excised by *Xho*I and *Spe*I; (*lane 2*) 100-bp markers; (*lane 3*) domains 1 + 2 excised using *Xho*I and *Bss*HII; (*lane 4*) 1-kb markers; (*lane 5*) domains 1 through 3 excised using *Xho*I and *Sac*I; (*lane 6*) domains 1 through 4 excised using *Xho*I and *Apa*I; (*lane 7*) domains 1 to 5 excised using *Xho*I and *Mlu*I. (*c*) Electrospray ionization mass spectrum of purified $(I27)_5$*. The measured mass was 52,235.5 ± 1.9 Da in excellent agreement with the expected mass (52,235.2 Da).

regarded as comprising leucine 1-leucine 89 (from the original structure determination (Improta et al., 1996)). Linkers consisting of four to six amino acids were inserted between domains to decrease inter-domain interactions. The sequences of the linkers were designed to be as similar as possible to the natural I26-I27 and I27-I28 linkers (linker choice was constrained by restriction site sequence). $I27_1$ was generated by PCR using C47S, C63S I27 as the template. The forward primer coded for a *Xho*I restriction site and the reverse primer added an artificial multiple cloning site (MCS) coding for the following restriction endonuclease sites: *Spe*I, *Bss*HII, *Sac*I, *Apa*I, and *Mlu*I. The blunt ended PCR product was A-tailed with *Taq* DNA polymerase and, after purification, ligated into a pGEM-T vector (pGEM-T vector system) using the manufacturers protocol. Plasmid DNA from colonies bearing inserted DNA was isolated and the DNA sequence verified giving pGEM $I27_1$+MCS.

The rest of the double mutant cassettes ($I27_2$, $I27_4$, and $I27_5$) were created using PCR with pGEM $I27_1$+MCS as template and forward and reverse primers to give the correct pairs of unique restriction sites. Again, after A-tailing and purification these PCR products were ligated into the pGEM-T shuttle vector for automated DNA sequencing. The single mutant cassette, $I27_3$ (C63S I27), was generated in an analogous manner using pET8cI27 C63S as DNA template.

The concatamer was assembled as follows. $I27_1$+MCS was cut from pGEM $I27_1$+MCS using *Xho*I and *Mlu*I and ligated into the original His-tag modified pET8c vector, which had been predigested with the same restriction endonucleases. Cassettes were then formed by digestion of the pGEM shuttle containing the desired gene ($I27_2$, $I27_3$, $I27_4$, or $I27_5$) with its respective pair of restriction enzymes. The cassettes were then ligated into pET8cI27_1+MCS, which had also been digested with the same pair of enzymes. After the first round of ligation all steps were performed in *Escherichia coli* Sure2 cells (Stratagene, Amsterdam, The Netherlands). The final product contained five I27 genes and is denoted $p(I27)_5$*. The presence of the full-length concatamer in the plasmid was assessed by a series of restriction digests that removed one to five genes from the vector (Fig. 2 *b*).

Protein over-expression/purification

Over-expression and purification was essentially identical for both monomeric and polymeric I27. For over-expression, both pI27 C63S and pI27 C47S, C63S were transformed into *E. coli* BL21[DE3]pLysS (Novagen, Nottingham, UK) while the $p(I27)_5$* construct was transformed into *E. coli BLR*[DE3]pLysS (Novagen). An overnight starter culture (50 mL) (LB medium with 25 μg mL^{-1} chloramphenicol and 100 μg mL^{-1} ampicillin) of the appropriate construct was used to inoculate each of 10 × 1 L Luria Bertani medium with supplements as above and incubated in a shaking incubator at 37°C. Protein expression was induced at an OD_{600} = 0.7 to 0.8 with 1 mM isopropyl β-D-thiogalactopyranoside (final concentration). The cells were harvested by centrifugation 4 h after induction.

The protein was isolated using histidine Ni-NTA affinity chromatography resin (Qiagen, Crawley, W. Sussex, UK) following the manufacturers protocol. After elution, the protein was dialyzed into distilled deionized water and freeze dried. The protein was further purified by size-exclusion chromatography. Briefly, 5 mL of protein solution (5 mg mL^{-1}, 50 mM

Na_2HPO_4/NaH_2PO_4, pH 7.6) was applied to a preequilibrated 320-mL Superdex 75 column (26-mm diameter, Amersham Biotech, Little Chalfont, Bucks., UK) and eluted at a flow rate of 1.5 mL min^{-1}. Fractions containing the concatamer were pooled and extensively dialyzed against deionized distilled water. The protein was freeze dried in appropriate aliquots (0.05 mg for AFM studies and 5 mg for all other experiments) and stored at −20°C. The yield of pure (>95%) concatamer was typically 10 mg L^{-1}. Yields of the monomeric domains were typically 10 mg L^{-1}. ESI-MS demonstrated that the mass of the concatamer (52, 235.5 ± 1.9 Da) was in excellent agreement with the mass estimated by its amino acid sequence (52, 235.2 Da) (Fig. 2 *c*). Similarly, the monomeric domains were of the expected mass (10, 894.6 ± 0.4 Da for C63S I27 and 10, 878.2 ± 0.8 Da for C47S, C63S I27).

Equilibrium denaturation

Protein samples (10 μM) in different concentrations of guanidine hydrochloride (GnHCl) were prepared from stock solutions of 20 mM Na_2HPO_4/NaH_2PO_4, pH 7.3, 1 mM EDTA, and 2 mM dithiothreitol containing either 0 M or 8 M GnHCl (ICN Biomedicals Inc., Basingstoke, Hants, UK). These solutions were mixed in different ratios to give a 0 to 8 M range of denaturant concentrations with 0.1- or 0.2-M increments. The samples were centrifuged, vortex mixed, and then allowed to equilibrate overnight in a circulating water bath at 25°C. Fluorescence emission spectra or intensity versus time traces were acquired on a PTI Quantmaster C-61 spectrofluorimeter at 25°C. Denaturation was followed by measuring the intensity of tryptophan fluorescence at 315 nm after excitation at 280 nm. After signal averaging, the intensity was plotted as a function of denaturant concentration and the data fitted to a two-state transition as described previously (Ferguson et al., 1999). In the case of $(I27)_5$* the data were fitted manually to the sum of two independent two-state transitions, the first of which accounts for 80% of the signal (C47S, C63S I27). The *m* values and fluorescence intensities of the native and denatured states for each domain were assumed to be identical. For presentation purposes the raw data were then converted to fraction population of native molecules (Santoro and Bolen, 1988).

Kinetic analysis of monomeric mutant proteins

Kinetic folding experiments were performed using an Applied Photophysics SX.18MV stopped-flow fluorimeter. The temperature was internally regulated using an external probe placed near the cuvette and maintained at 25°C using a Neslab RTE-300 circulating water bath. Tryptophan fluorescence was excited at 280 nm with a 10-nm bandwidth, and the emitted fluorescence was monitored at >320 nm.

All refolding experiments were performed in 20 mM Na_2HPO_4/NaH_2PO_4, pH 7.3, 1 mM EDTA, and 2 mM dithiothreitol. Refolding experiments were performed by dissolving protein (~50 μM) into buffer containing 3.6 M GnHCl and diluting this 1:10 into solutions containing various denaturant concentrations to give final GnHCl concentrations down to 0.33 M. For each GnHCl concentration used, several kinetic transients were averaged and fitted to a double exponential equation using the manufacturers software. The minor phase (~20% of the amplitude) was attributed to proline isomerization and was ignored in further analysis. Unfolding experiments were performed by manual mix. Protein (~50 μM) in native buffer (20 mM Na_2HPO_4/NaH_2PO_4, pH 7.3, 1 mM EDTA, and 2 mM dithiothreitol, or phosphate-buffered saline (PBS) containing 1 mM EDTA and 2 mM dithiothreitol) was diluted 1:9 into solutions containing various GnHCl concentrations. The decrease in fluorescence at 315 nm (excitation 280 nm) was monitored under the same conditions as described above in a 1-cm path length cuvette for 600 s. Kinetic transients were fitted to a 3-parameter single or 5-parameter double exponential equation using Sigmaplot (SPSS Inc., Woking, Surrey, UK) for monomeric and polymeric proteins, respectively. Identical results were obtained in both buffer systems.

Mechanical unfolding

All mechanical unfolding experiments were performed using a commercially available mechanical force probe (MFP-SA, Asylum Research Inc., Santa Barbara, CA). Coated unsharpened microlevers (MLCT-AUNM) were obtained from ThermoMicroscopes (Cambridge, UK). The spring constant of each cantilever was calculated under PBS using the thermal method (Florin et al., 1995) and was typically found to be ~51 ± 5 pN nm^{-1}.

Protein (0.05 mg) was reconstituted to 0.1 mg mL^{-1} in sterile PBS and centrifuged (11,600 × *g*, MSE, MicroCentaur). Typically 45 μL of PBS was dropped onto a recently cleaved template stripped gold surface. Protein solution (15 μL) was then added and the two solutions allowed to mix. At this protein concentration the probability of attaching a molecule to the tip is relatively low (typically 4%). However, under these conditions ~50% of the traces result in the attachment of a single molecule and four or more clear unfolding peaks.

Mechanical unfolding experiments were performed at pulling speeds varying from 70 nms^{-1} to 4000 nms^{-1} at a room temperature of 23 ± 1°C over a distance of 400 nm for $(I27)_5$* and 600 nm for $(I27)_8$ (kindly provided by Jane Clarke, University of Cambridge, UK). Each mechanical unfolding experiment at each speed was performed three times with a fresh sample, on a different day and using a new cantilever.

Monte Carlo simulations

A two-state model was used to simulate the forced extension of the I27 constructs (Rief et al., 1998). Each domain of the molecule was initially assumed to be in the lowest energy state and therefore folded. The folding and unfolding rate constants at applied force *F* were calculated using $\alpha_{i,F} = \alpha^0_{i,F} \exp(\pm F_x x_i / k_B T)$ in which $i = f$ or u for the folding and unfolding events, respectively, and the negative sign is associated with folding. (Note: so as to differentiate between the intrinsic rate constants for chemical and forced unfolding, the notation $k^0_{u,GnHCl}$ or $\alpha^0_{u,F}$, respectively, is used). The constants x_f and x_u represent the distance from the folded and unfolded well to the barrier, respectively (this reaction co-ordinate is assumed to be parallel to the stretch axis). The protein was extended with speeds from 10 to 10,000 nms^{-1} and with different values of x_f, x_u, and $\alpha^0_{f,F}$ and $\alpha^0_{u,F}$, which are the rate constants for folding and unfolding, respectively, in the absence of applied force. The worm-like chain was used to calculate the force applied at any extension *x* as

$$F = \frac{k_B T}{p}\left(\frac{1}{4(1 - x/L)^2} - \frac{1}{4} + \frac{x}{L}\right)$$

in which *p* is the persistence length, and *L* the total contour length calculated as $L = zL_f + (n - z)L_u$ for *z* folded domains of length L_f and $n - z$ unfolded domains each of length L_u. At each extension the probability of folding, unfolding, or extending the chain is calculated. If unfolding (folding) occurs the chain length, *L,* is increased (decreased), as described above, the cantilever extension incremented, and the probability of folding, unfolding, or extending the protein recalculated. The sequence of domain unfolding is random. As a consequence the first domain to unfold, corresponding to the first pulling event, can be any one of those in the construct and not necessarily the first or last in the chain. The procedure is continued until all domains are unfolded. The whole calculation was then repeated 10,000 times. From a histogram of the unfolding forces the mode and mean values of the force were obtained. Because of the exponential dependence of the rate on force, very small increments of extension, or equivalently time, must be used and typical extension steps were 0.005 nm, typically over a length of 200 nm. For each set of parameters, a graph of *F* versus

log (pulling speed) was constructed and compared with the data. The parameters were varied until the simulated speed dependence of the unfolding force matched the best-fit line of the experimental data.

RESULTS

Effect of mutations on the equilibrium and kinetic stability of monomeric C63S and C47S, C63S I27

The mutation(s) in each I27 domain replaced cysteine with serine at position 63 (for C63 I27) or at positions 47 and 63 (for C47, C63 I27) (Fig. 1). By inspection of the structure of I27 (protein data base file 1TIT (Improta et al., 1996)), it is found that C47 is in strand D and makes contacts with residues in the D and E strands. C63 is in the loop between strands E and F and contacts residues in the A′ (V13), E, and G (L84, V86) strands (Fig. 1). The amide hydrogen of V13 forms a hydrogen bond with the oxygen of L85 in the putative hydrogen bond clamp region of the A′G interface (Lu and Schulten, 2000). While the side-chains of V13 and C63 make contact (within ~3 Å), a cysteine to serine mutation is a conservative one, and these contacts would be expected to be retained. Further, the sulfur atom of C63 is distant from the hydrogen bond donor of the V13-L85 interaction (7 Å). It should be noted that the I27 studied previously (Carrion-Vazquez et al., 1999a) is in fact a pseudo wild type. This differs from the published amino acid sequence of I27 (which we have used as our wild type) at two positions: T42A and A78T. Both of these mutations are distant from both the A′ and G strands and make no contacts with any of the residues involved in the clamp region.

Fig. 3 *a* shows the equilibrium denaturation curve obtained for both monomeric I27 mutants, as well as for $(I27)_5$*. The curve expected for the wild-type protein under identical conditions (determined using published parameters (Carrion-Vazquez et al., 1999a)) is shown for comparison. The data show that both mutations significantly destabilize I27 relative to the wild-type protein ($\Delta\Delta G_{UN}$ = 8.7 and 17.9 kJ mol^{-1} for the single and double mutant, respectively; Table 1). The stability of each domain is not altered significantly by its tethering in $(I27)_5$* (Table 1). The effect of these mutations on the rate constants for folding and unfolding of each protein is shown in Fig. 3 *b*. Both mutations markedly increase the intrinsic chemical unfolding rate constant, $k^0_{u,GnHCl}$, of each monomeric domain (from 4.9 × 10^{-4} s^{-1} for the wild-type domain (Carrion-Vazquez et al., 1999a), to 7.6 × 10^{-3} s^{-1} and 10.6 × 10^{-3} s^{-1} for C63S I27 and C47S, C63S I27, respectively). They also decrease the rate of folding (Table 1). The position of the transition state is not altered significantly by the mutations (β_T = 0.9, 0.95 ± 0.05 and 0.94 ± 0.10 for wild-type I27 (Carrion-Vazquez et al., 1999a), C63S I27 and C47S, C63S I27, respectively (Table 1)). The kinetic intermediate previously reported for wild-type I27 (Carrion-Vazquez et al., 1999a) is not detected during refolding of the mutant domains, presumably because it is also significantly destabilized, such that it can no longer be observed. As a consequence, the equilibrium and kinetic data were fitted to a two-state transition and the resulting parameters are comparable (Table 1).

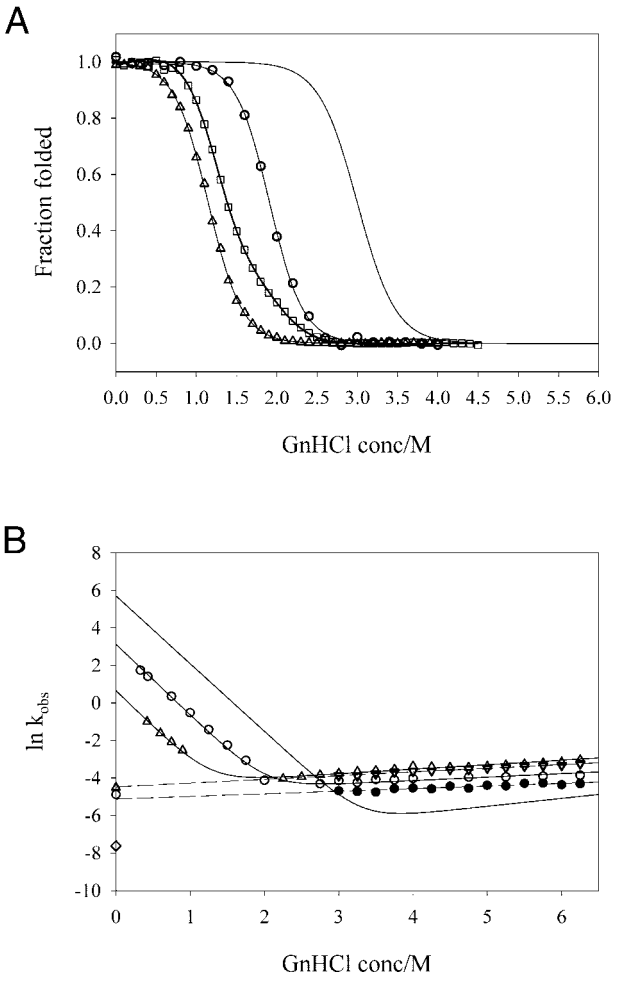

FIGURE 3 (*a*) Equilibrium denaturation of I27 monomers and the $(I27)_5$* concatamer followed by tryptophan emission fluorescence spectroscopy. The data have been converted to fraction of folded molecules for ease of comparison (△, C47S, C63S I27; ○, C63S I27; □, $(I27)_5$*; solid line (no symbols), I27 wild-type data simulated from published parameters (Carrion-Vazquez et al., 1999a). Solid lines to the data for C63S and C47S, C63S are the fits to a two state model. Solid line to the data for $(I27)_5$* is a fit to a two state model for each of two species (C63S I27 and C47S, C63S I27) as described in Materials and Methods. The resulting parameters are shown in Table 1. (*b*) Rate profile of I27 monomers and the $(I27)_5$* concatamer. △, Monomeric C47S, C63S I27; ○, monomeric C63S I27. Solid lines are fits to a two state model (fit parameters are listed in Table 1). The data for wild-type I27 (solid line, no symbols) were simulated using published data (Carrion-Vazquez et al., 1999a). (*Inverted triangles*) Observed rate constant for the faster phase of $(I27)_5$* unfolding. ●, Observed rate constant for the slower phase of $(I27)_5$* unfolding. Dashed lines show the best linear least squares fit for the fast and slow phases of unfolding of the concatamer. The open triangle, circle, and diamond on the ordinate are the extrapolated k^0_u for C47S, C63S I27, C63S I27, and wild-type I27, respectively.

TABLE 1 Derived parameters from equilibrium and kinetic folding/unfolding of I27 monomers and $(I27)_5$*

	WT I27*	C63S I27	C47S,C63S I27
Equilibrium			
ΔG_{UN} (kJ mol^{-1})	31.4 ± 1.3	22.7 ± 1.2	13.4 ± 0.3
		24 ± 2.0‡	14.5 ± 1.5‡
m_{un} (kJ/mol.M)	10.5	11.9 ± 0.6	11.7 ± 0.2
		11.7 ± 1.0‡	11.7 ± 1.0‡
Kinetic			
$k^0_{f,GnHCl}$ (s^{-1})	30 ± 2	22.8 ± 3.4	1.9 ± 0.4
m_f (kJ/mol.M)	NA†	9.36 ± 0.35	9.53 ± 0.68
$k^0_{u,GnHCl}$ (s^{-1})	$4.9 \pm 0.6 \times 10^{-4}$	$7.6 \pm 1.4 \times 10^{-3}$	$10.6 \pm 0.7 \times 10^{-3}$
		$5.9 \pm 0.5 \times 10^{-3}$‡	$10.7 \pm 1.0 \times 10^{-3}$‡
m_u (kJ/mol.M)	NA†	0.46 ± 0.11	0.60 ± 0.04
		0.34 ± 0.04‡	0.49 ± 0.02‡
m_{un} (kJ/mol.M)	NA†	9.82 ± 0.37	10.13 ± 0.68
β_T§	0.90	0.95 ± 0.05	0.94 ± 0.10
ΔG_{UN} (kJ-mol^{-1})	NA†	19.9 ± 0.5	12.9 ± 0.5

*Data taken from Carrion-Vazquez et al. (1999a).
†Wild type protein populates a kinetic folding intermediate.
‡Measured in $(I27)_5$*. All other values are those measured for individual monomeric domains.
§$\beta_T = m_f/m_u + m_f$ and gives a measure of transition state placement.

Does incorporation into a polyprotein affect unfolding kinetics?

To assess whether the unfolding rate constants of the monomeric domains are altered when their N and/or C termini are tethered, the unfolding kinetics of $(I27)_5$* were also measured between 3.0 and 6.5 M GnHCl (Fig. 3 *b*). Each transient was well described by a double exponential decay with a short lifetime corresponding to 81 ± 5% of the total amplitude. The rate constant of this phase ($k^0_{u,GnHCl}$ = $10.7 \pm 1 \times 10^{-3}$ s^{-1}) is identical to that of monomeric C47S, C63S I27 ($k^0_{u,GnHCl} = 10.6 \pm 0.7 \times 10^{-3}$ s^{-1}; Table 1) (which comprises four of the five I27 domains in $(I27)_5$*). Assuming that the fluorescence yield of the natively folded mutants in the monomer and polymer are similar, it seems reasonable to assign this lifetime to the unfolding of four C47S, C63S I27 domains in the concatamer. The denaturant dependence of the rate constant of this phase (the *m* value) is also similar to that of monomeric C47S, C63S I27 and both show no sign of curvature at low denaturant concentration. The second kinetic phase detected for the polyprotein has a rate constant of $5.9 \pm 0.5 \times 10^{-3}$ s^{-1}, which contributes ~20% to the total decay. This phase, therefore, is presumably that of the single cysteine mutant domain. The rate constant of this phase is very similar to that observed for monomeric C63S I27 ($7.6 \pm 1.4 \times 10^{-3}$ s^{-1}), confirming the assignment of this kinetic phase to unfolding of the single copy of C63S in $(I27)_5$*. The estimation of the intrinsic unfolding rate constant at zero denaturant concentration ($k^0_{u,GnHCl}$) assumes that there is no curvature in the relationship between the logarithm of the observed rate constant and the denaturant concentration even at low denaturant concentration. However, it is not possible to measure this rate constant directly below the midpoint of the transition. One of the advantages of using a severely destabilized mutant ($\Delta\Delta G_{UN} \approx 18$ kJ mol^{-1} for C47S, C63S I27) is that the extrapolation to zero denaturant is much shorter. As described above, the parameters obtained by equilibrium studies are within error of those obtained by kinetic studies (Table 1). This is diagnostic of two-state folding kinetics. Further, the unfolding kinetics of $(I27)_5$* has been measured between 1.80 and 7.05 M Gdn.HCl in PBS to match the buffer used in the mechanical unfolding studies. Again the unfolding branch of the chevron plot is totally linear and, importantly, analysis of the amplitudes of the kinetic phases reveals no burst phase species (data not shown).

Forced unfolding of $(I27)_5$*

The mechanical unfolding properties of $(I27)_5$* were investigated by extending the molecule via the molecular force probe at 70, 120, 200, 332, 600, 1000, 2100, and 4000 nms^{-1}. Several replicates were obtained at each speed to obtain a satisfactory data set. A force-extension profile is shown for each pulling speed in Fig. 4. As shown previously (Rief et al., 1997b; Carrion-Vazquez et al., 1999a), force extension profiles in which only one molecule is attached between tip and surface result in a series of "sawteeth." These number between one and the total number of domains present in the concatamer (in this case five). The leading edge of the first event corresponds to stretching the fully folded concatamer. The final peak represents the energy required to extend the fully denatured concatamer up to a maximal force whereupon tip or surface desorption occurs (a force of ~1.5 nN would be expected for the rupture of the gold substrate-sulfur bond (Grandbois et al., 1999)). The leading edge of the 3rd to 6th peaks is adequately described by the worm-like chain model for simple polymer elasticity

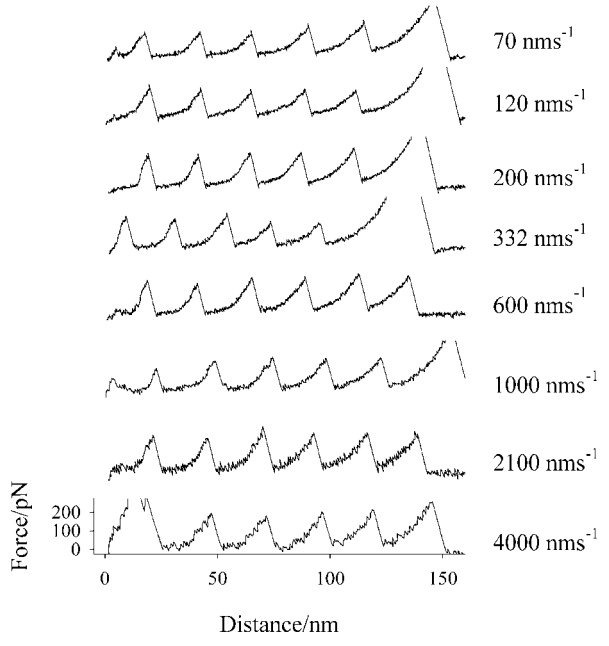

FIGURE 4 Sample force profiles obtained from $(I27)_5$* at different pulling speeds. Note the first peaks in each trace deviate from a typical WLC polymer extension consistent with the population of an unfolding intermediate (Marszalek et al., 1999).

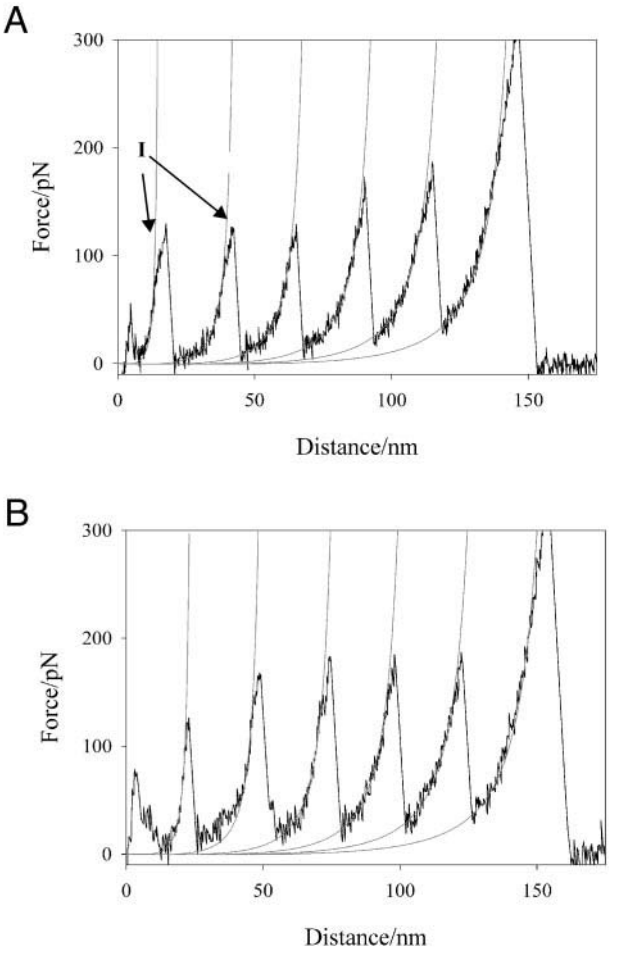

FIGURE 5 Worm-like chain model (*gray lines*) fitted to the rising edge of the unfolding peaks at (*a*) 70 nms^{-1} and (*b*) 1000 nms^{-1} (*black lines*). Both data sets can be fitted satisfactorily using a persistence length of 0.4 nm. This gives an average $\Delta L = 28$ nm. Note the progressively better fit of the WLC to the data at 70 nms^{-1} at increased extensions. This is consistent with the presence of an unfolding intermediate (labeled I) observed by Marszalek et al. (1999) using a wild-type I27 concatamer.

(Fig. 5), despite the fact that this model does not take into account anything more than the elasticity of the unfolded polypeptide chain. The compliance of the folded domains, for example, is ignored. The unfolding intermediate (labeled I in Fig. 5) observed previously in both experiment and simulation (Lu et al., 1998; Marszalek et al., 1999) is clearly visible in the leading edge of the first few unfolding peaks of some of the force-profiles shown here. The first portion of the unfolding peak was best fit using a persistence length (p) of 0.4 nm, which was held constant in subsequent fits. The resultant change in contour length (ΔL) was found to be 28 ± 1 nm. The expected contour length upon unfolding of one complete domain should be equal to the number of "structured" amino acids (i.e., K6-K87, 81 residues) multiplied by the distance between two adjacent C^{α} atoms in a fully extended state (0.34 nm (Yang et al., 2000)), less the initial separation between the two boundary amino acids in the native state (0.23 nm). The increase in contour length estimated in this way is 27.3 nm in good agreement with the experimentally derived value of 28 ± 1 nm.

Analysis of mechanical unfolding data

Empirically it has been found that the effect of pulling speed (v) upon maximal unfolding force (F) is given by $F = a + b \ln(v)$ (Merkel et al., 1999; Carrion-Vazquez et al., 1999a). To accurately quantify this relationship the following criteria were applied to each experiment to select peaks for further analysis: 1) the last peak (~300 pN), which shows protein detachment from the tip, was omitted; 2) only force-extension profiles containing two or more unfolding peaks were included; 3) only experiments in which a single molecule was attached were included; 4) each peak had to have the correct interpeak distance after unfolding (23.1 ± 1.3 nm). The peak force for an individual unfolding event in each data set was measured, allotted to a 10-pN bin, and frequency histograms were then generated. These histograms fit to a Gaussian function (Fig. 6). Because $(I27)_5$* consists of four copies of C47S, C63S I27 and one copy C63S I27, the resulting histograms are a convolution of two frequency distributions with different amplitudes. Monte Carlo simulations (with a data set of $n = 10{,}000$ com-

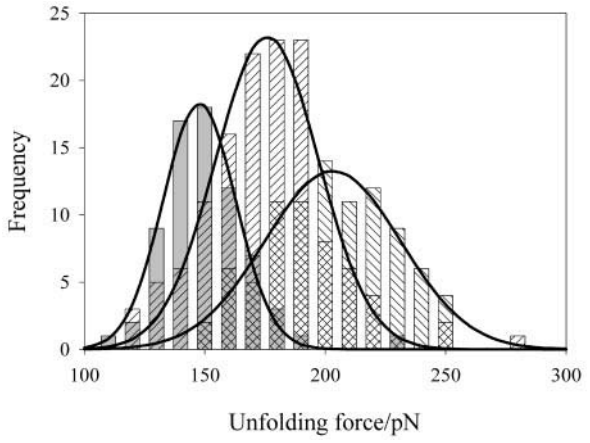

FIGURE 6 Sample histograms of the mechanical unfolding forces of $(I27)_5$* obtained at: 70 nms^{-1} (*solid gray bars*), 600 nms^{-1} (*up diagonals*), and 4000 nms^{-1} (*down diagonals*). Fits of a Gaussian distribution to each histogram are shown in bold lines.

pared with $n = 150$ for experimental data) suggest that the respective $\alpha^0_{u,F}$ values would have to differ by at least an order of magnitude before separate distributions could be resolved. The kinetic analysis of the domains in the concatamer by chemical denaturation showed the difference in intrinsic unfolding rate constants of each domain to be only twofold (Table 1). The maximum of each fitted Gaussian (the mode) was therefore used in further analysis. To assess the validity of this approach the mean of each data set was also calculated and was found to be consistently less than the mode although within experimental error (Table 2). The precision of the experiment was assessed by pulling the concatamer at each speed in triplicate and taking relatively few data points (n = 39–150). Averaging the three modes or means at each pulling speed gives an indication of the reproducibility of the experiments. The standard error of the mean varies from ± 3 to ± 6 pN, which shows that there is little variability between experiments.

Two methods were used to interpret the experimentally determined constants a and b in terms of the fundamental parameters x_u and $\alpha^0_{u,F}$: first using the theory of Bell (1978) and Evans and Ritchie (1997) and second using a Monte-Carlo method. When a force is applied to the N- and C-terminal residues of a protein, the bonds can be stretched to a greater or lesser extent depending upon whether they are in line with the applied force. For those bonds that are stretched, such as the hydrogen bonds shown in Fig. 1, we assume that this causes a lowering of the barrier to bond breaking and reformation, and the extent of lowering is the difference in energy. According to the theory of Bell (1978) and Evans and Ritchie (1997), which assumes a single barrier to bond rupture (or in this case to protein unfolding/refolding), the application of force, F, to a protein decreases this barrier (ΔG_u) by $F.x_u$, in which x_u is a single measure of the "distance" in configuration space from the native geometry to the mechanical unfolding transition state:

$$\alpha_{u,F} = \alpha^0 e^{-(\Delta G - Fx_u)/k_B T} = \alpha^0_{u,F} e^{Fx_u/k_B T} \quad (1)$$

in which α^0 is the pre-exponential factor that represents the rate in the limit of high temperature. This reduction in energy then allows barrier crossing driven by thermal fluctuations (of the order of k_BT) to occur more frequently than would otherwise be the case. It can be seen, therefore, that the mechanical sensitivity of a protein (i.e., the extent to which the unfolding barrier is lowered upon application of force to a protein) depends on the product $F.x_u$. As will be discussed later, proteins studied so far exhibit a broad range of values for F and x_u, which are difficult to interpret taken in isolation. From Eq. 1, it can be seen that unfolding occurring in a short time interval (i.e., at high pulling speeds) must occur over a smaller total barrier, hence larger applied force, (see Eq. 1) than if the force is applied for longer, i.e., slower pulling speed, because other things being equal, fast pulling allows less time for thermal motions to provide energy to surmount the barrier. At the point of bond rupture the maximum in the distribution of forces is related to x_u and the unfolding rate constant at zero applied force, $\alpha^0_{u,F}$ (Bell, 1978; Evans and Ritchie, 1997):

$$F = \frac{k_B T}{x_u} \ln\left(\frac{r x_u}{k_B T \alpha^0_{u,F}}\right)$$

or (2)

$$F = \frac{k_B T}{x_u} \ln\left(\frac{r x_u}{k_B T \alpha^0}\right) + \frac{\Delta G}{x_u}$$

in which r is the force loading rate. At 300 K, $k_BT = 4.1$ pN nm and x_u is in the region 0.1 to 0.4 nm producing forces in the 100 to 300 pN range at the loading rates used. The constant $\alpha^0_{u,F}$ should be modified when more than one (identical) domain is being pulled to become $\alpha^0_{u,F} n!^{1/n}$, since

$$F = \frac{1}{n} \sum_{i=1}^{n} \frac{k_B T}{x_u} \ln\left(\frac{r x_u}{k_B T \alpha^0_{u,F} i}\right).$$

The mean unfolding force measured is given by:

$$\langle F \rangle = \frac{k_B T}{x_u} \ln\left(\frac{k_B T \alpha^0_{u,F}}{r x_u}\right) En_1\left(\frac{k_B T \alpha^0_{u,F}}{r x_u}\right)$$

in which En is the exponential integral of the first type. At large force, which for the parameters used in this paper is >50 pN, then this equation can be approximated as:

$$\langle F \rangle \approx \frac{k_B T}{x_u} \left(\ln\left(\frac{r x_u}{k_B T \alpha^0_{u,F}}\right) - \gamma\right)$$

and so differs from (2) only by Eulers constant, $\gamma = 0.5772$. This behavior is in agreement with the force histograms,

TABLE 2 Summary of mechanical unfolding data for $(I27)_5$*

Speed/nms^{-1}	n	Mean unfolding force/pN (±SD)	Average/pN (±SE)	Mode unfolding force/pN (±SD)	Average/pN (±SE)
$(I27)_5$* MFP I					
77	39	130 (21)		134 (10)	
	50	147 (19)	141 (6)	150 (16)	144 (5)
	71	146 (18)		148 (15)	
120	40	148 (20)		152 (14)	
	68	131 (22)	142 (5)	136 (23)	144 (5)
	51	146 (18)		145 (11)	
200	71	158 (24)		165 (23)	
	52	151 (19)	157 (3)	155 (21)	162 (4)
	80	161 (18)		167 (19)	
332	60	168 (22)		174 (23)	
	111	153 (23)	159 (5)	157 (24)	164 (5)
	87	156 (22)		161 (21)	
600	104	165 (24)		167 (25)	
	90	178 (25)	172 (4)	181 (26)	175 (4)
	131	172 (26)		176 (21)	
1000	117	178 (22)		180 (21)	
	100	167 (26)	173 (3)	171 (23)	177 (3)
	84	174 (25)		180 (27)	
2100	131	196 (26)		201 (26)	
	150	188 (27)	194 (3)	193 (26)	198 (3)
	92	198 (24)		201 (23)	
4000	63	197 (29)		201 (31)	
	85	206 (26)	200 (3)	207 (28)	204 (2)
	91	198 (26)		203 (28)	
$(I27)_5$* MFP II					
600	94	176 (27)		178 (26)	
	105	167 (23)	171 (3)	173 (23)	172 (3)
	86	169 (24)		166 (17)	
2100	98	182 (24)		184 (20)	
	48	202 (28)	193 (6)	206 (27)	196 (6)
	76	196 (26)		199 (24)	
Wild-type $(I27)_8$ MFP I					
200	60	179 (22)		185 (21)	
	45	196 (22)	184 (6)	203 (17)	189 (7)
	35	177 (25)		179 (25)	
1000	105	194 (28)		201 (28)	
	39	214 (24)	206 (6)	218 (25)	212 (5)
	73	209 (25)		217 (20)	

which show that the mean force is slightly less than the mode.

A plot of force versus log pulling speed (Fig. 7 *a*) shows that the speed dependence of the unfolding force for $(I27)_5$* agrees well with the empirical observation $F = a + b \ln(v)$. Comparison of the data obtained here (Fig. 7 *a*; Table 2) and elsewhere (Carrion-Vazquez et al., 1999a) on polymers of wild-type I27 shows that $(I27)_5$* unfolds at significantly lower forces than the wild-type I27 polyprotein. For example, at 10 nms^{-1} and 1000 nms^{-1}, the wild-type I27 concatamer, $(I27)_8$ (Carrion-Vazquez et al., 1999a) unfolds at a force that is 21 and 39 pN higher, respectively, than $(I27)_5$*. The number of domains in the polymer is known to affect both the measured peak and average forces for unfolding (Makarov et al., 2001). Calculation shows that the smaller the number of domains in a polymer, the greater the measured unfolding force. Increasing the number of domains from 1 to 50 reduces the mean unfolding force by ~40 pN, due to the fact that domains unfold independently of one another. The more domains, the greater is the chance of unfolding in a given time interval. Monte Carlo simulations show that, in our system, this effect would cause a 6-pN decrease in unfolding force of pulling an octamer relative to a pentamer, further increasing the difference between the average unfolding forces of $(I27)_5$* and wild-type $(I27)_8$ domains.

Estimation of $\alpha^0_{u,F}$ and x_u by experiment: the assumption of a single domain and a single barrier

The forced unfolding rate constant in terms of the intrinsic forced unfolding rate constant under zero applied load, $\alpha^0_{u,F}$,

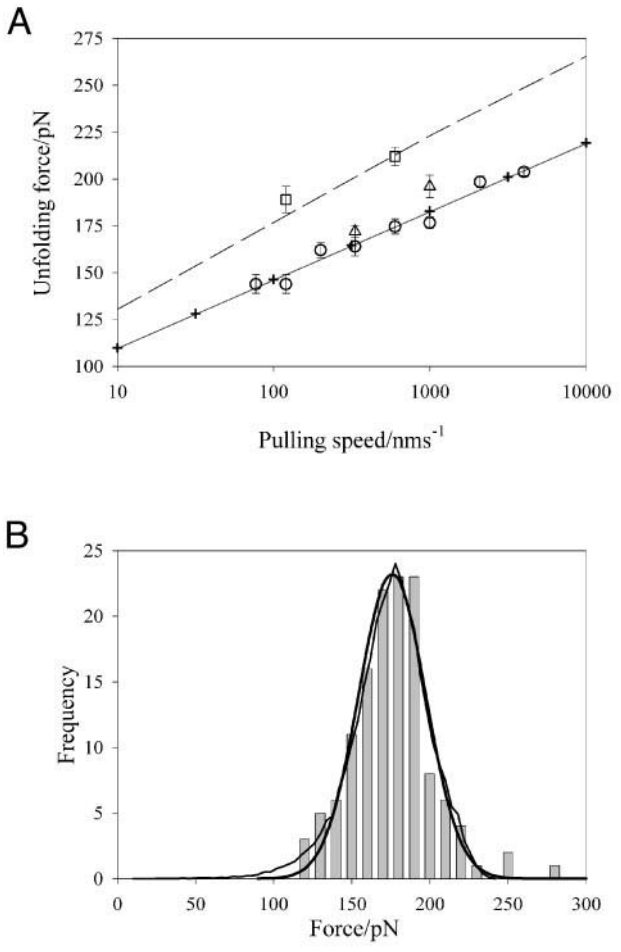

FIGURE 7 (*a*) Graph showing the pulling speed dependence of the observed unfolding force. $(I27)_5$* measured "in-house" (○) unfolds at a lower force when compared with the published data (Carrion-Vazquez et al., 1999a) for wild-type $(I27)_8$ (*dashed line*). To obviate calibration errors as a source of this difference, $(I27)_5$* was unfolded using a different molecular force probe (Dr. J. Clarke, University of Cambridge, UK) at 600 and 2100 nms^{-1} in triplicate (△). A concatamer of wild-type I27 ($(I27)_8$, a kind gift from Dr. J. Clarke, University of Cambridge, UK) was also unfolded at 200 and 1000 nms^{-1} in triplicate "in house" (□). All error bars are standard error of the mean. The solid line is a linear least squares fit through $(I27)_5$* data obtained "in-house." Cross hairs show the best fit Monte Carlo simulation with parameters $\alpha^0_{u,F} = 2.0 \times 10^{-3}$ s^{-1} and x_u = 0.29 nm. (*b*) Experimental force frequency histogram of $(I27)_5$* at 600 nms^{-1} fitted to a Gaussian function (*solid heavy line*). The histogram obtained by Monte Carlo simulation at the same pulling speed is shown as a light line.

is given by Eq. 1. If the unfolding rate of the protein in question is much slower than the time scale on which the protein is being extended, then the observed unfolding rate constant, $k_{u,F}$, is given simply by the pulling speed in each experiment divided by the length of a fully unfolded domain (28 nm), i.e., the number of domains unfolding per second. From Eq. 1 it can be seen that x_u can be estimated from the gradient, (k_BT/x_u), of a plot of force versus ln unfolding rate constant. In this way the parameter x_u was found to be 0.26 ± 0.02 nm for $(I27)_5$* (data taken from Fig. 7 *a*, replot not shown). In comparison, x_u for the wild-type $(I27)_8$ polymer determined in the same manner with data taken from Carrion-Vazquez et al. (1999a) and Fig. 7 *a* is ~0.20 nm, suggesting the mechanical unfolding transition state is significantly displaced in $(I27)_5$*. As reported previously (Carrion-Vazquez et al. 1999a), the small value of x_u indicates that the transition state for mechanical unfolding is close to the native state, in agreement with the transition state placement inferred from chemical denaturation data. However, the relevance of this similarity is difficult to assess due to the differing nature of the reaction co-ordinate. In classical chemical denaturation this co-ordinate is usually the accessible surface area exposed to solvent (Myers et al. 1995). The reaction co-ordinate in mechanical unfolding is simply distance. The physical meaning of this one-dimensional quantity in configurational space is, however, difficult to interpret. Comparison of the transition state placement as measured by the two techniques in structural terms is, therefore, not meaningful. The values for x_u determined for $(I27)_5$* and $(I27)_8$ are both significantly lower than those obtained by simulation (see below). At first glance it would appear that the intercept with the ordinate in the plot of force versus ln unfolding rate constant should yield the intrinsic unfolding rate constant, $\alpha^0_{u,F}$. However, Eq. 1 does not take the number of domains in the polyprotein into account and, as discussed above, the number of domains affects the observed unfolding force and therefore affects the observed unfolding rate constant. This can be corrected for if this is the only effect, but there is also a varying change in length of the whole polymer as it unfolds and these unfolded domains affect the force experienced by the folded domains in a manner that is not accounted for by analytic theory (Eq. 2). Thus, the intercept with the ordinate in this plot is difficult to interpret. To obtain a value of $\alpha^0_{u,F}$ a Monte Carlo (MC) simulation of the data must be performed.

Estimation of $\alpha^0_{u,F}$ and x_u by MC simulation: the effect of multiple domains and an additional barrier

Estimation of the intrinsic unfolding rate constant directly by the linear extrapolation method as described above is complicated by the effect of domain number on the measured unfolding force. However, the parameters $\alpha^0_{u,F}$ and x_u can also be obtained by simulating the effect of displacing the N terminus upon the rate of domain unfolding using a Monte Carlo method (Rief et al., 1998). The force applied to the protein by extending the termini at a certain pulling speed is calculated assuming a worm-like chain (WLC) model for polymer elasticity. We emphasize that the purpose of the simulations is to calculate the critical breaking force, not the detailed shape of the saw-tooth of the force versus extension data. These simulated force histograms can

be fitted to either individual experimental histograms or, more accurately, to a full speed dependence of the measured unfolding force. A sample Monte Carlo generated histogram of unfolding for $(I27)_5$* is shown in Fig. 7 *b*. Both the mode force and the force distribution fit the experimental data well. However, the simulation produces a distribution skewed toward lower force due to there being a probability of barrier crossing at zero force, which is not seen experimentally in the data set, presumably because it is too small (Carrion-Vazquez et al., 1999a; Marszalek et al., 1999). The best fit of the Monte Carlo simulation to the experimental data across the entire range of unfolding speeds is shown in Fig. 7 *a*. This fit was obtained using $\alpha^0_{u,F} = 2.0 \times 10^{-3}\ s^{-1}$ and $x_u = 0.29$ nm (and $\alpha^0_{f,F} = 1.2\ s^{-1}$, $x_f = 150$ nm, $L_u = 28.0$ nm, $L_f = 4.0$ nm, and $p = 0.4$ nm). The parameters obtained by fitting similar simulations to the data for wild type $(I27)_8$ yielded values of $\alpha^0_{u,F} = 3.3 \times 10^{-4}\ s^{-1}$ and $x_u = 0.25$ nm (Fig. 7 *a*; Carrion-Vazquez et al., 1999a). The values obtained for x_u, by both experiment and simulation suggest that for $(I27)_5$*, the transition state has moved closer to the denatured state relative to wild type $(I27)_8$.

It is interesting to note that x_u obtained from MC simulations is always greater than that obtained directly from the plot of force versus ln unfolding rate constant, in contrast to predictions based on single barrier models (Evans and Ritchie, 1997). A possible simple, and general, explanation for this is that the process of unfolding is not a one-step process. Disruption of several hydrogen bonds is necessary for complete unfolding, and the process of unfolding can, in principle, trap intermediate states and elements of secondary structure (Makarov et al., 2001). The simplest theoretical treatment of such effects is to include a second barrier (Evans, 1998) that models a step-wise removal of secondary structure from the domain. We have tested this by calculating the peak forces assuming that the concatamer has two types of barrier; a barrier to unfold initially, and a further secondary barrier to completely unfold the domain. The second barrier is related, along the reaction coordinate, to that for the initial unfolding by $\Delta x = x_0 - x_u$. Using the Evans and Ritchie model (Evans and Ritchie, 1997) modified for two barriers we find that the maximal force F is given by the equation:

$$c(x_u - \Delta x)\exp\left(\frac{F\Delta x}{k_BT}\right) + x_u - \frac{\alpha^0_{u,F}k_BT}{r}\exp\left(\frac{Fx_u}{k_BT}\right) = 0 \quad (3)$$

in which $c > 1$ is the ratio of rate constants for unfolding the folded and stretching the unfolded protein. If we solve Eq. 3 numerically for F using the parameters for $(I27)_5$* from the Monte Carlo calculation ($x_u = 0.29$ nm, $\alpha^0_{u,F} = 2.0 \times 10^{-3}\ s^{-1}$) and then vary c and x_u to fit the force versus pulling speed experimental data, we find that $c = 2$ and $\Delta x = 0.05$ nm. This crude model, we believe, explains why the x_u value calculated from the Monte Carlo calculation is slightly greater than that determined assuming only a single barrier according to Eq. 2.

The intrinsic forced unfolding rate constant for a monomer in $(I27)_5$* obtained by Monte Carlo simulations ($\alpha^0_{u,F} = 2.0 \times 10^{-3}\ s^{-1}$) is five times smaller than that obtained by chemical denaturation ($k^0_{u,GnHCl} \sim 0.01\ s^{-1}$ for both mutants either in monomeric or polymeric form (see above)). The data suggest, therefore, that the mutations affect the chemical and mechanical transition states differently assuming that 1) the dependence of the unfolding rate constant on dentaurant concentration is linear to 0 M GnHCl (see above) and 2) mechanical unfolding involves a single barrier (biased MD simulations suggest two barriers are present (Paci and Karplus, 2000)).

DISCUSSION

Previous data obtained using a polymer of eight wild-type I27 domains (Carrion-Vazquez et al., 1999a) have suggested that mechanical and chemical unfolding probe the same pathway, at least in that the intrinsic rate constant of unfolding (at zero force or in the absence of denaturant) determined by each method is identical for this domain. Although this interpretation is consistent with the data obtained by Carrion-Vazquez et al. (1999a), it is not unequivocal. Recent molecular dynamics simulations on model systems (Klimov and Thirumalai, 2000) and three different protein motifs (Paci and Karplus, 2000) suggest that thermal and mechanical unfolding occur by different pathways. Further evidence is thus required, for example by comparing the effect of mutations on the transition state for unfolding, before the similarity, or otherwise, of the processes can be confirmed. By constructing a concatamer of five mutant I27 domains we have explored the effect of mutation on the transition state for both mechanical and chemical denaturation. We show that, by contrast to the data for wild-type I27, the rate constant for mechanical unfolding of $(I27)_5$*, determined by fitting the AFM data directly using Monte Carlo methods, is more than fivefold slower than the rate constant determined using classical chemical denaturation. Although this difference is small, by obtaining data for mechanical unfolding in triplicate on independent occasions, we show that the difference in rate constants is significantly larger than the errors in the data. Our results suggest, therefore, that chemical and mechanical methods probe different unfolding pathways, at least for this I27 domain. Further mutagenesis studies will now be needed to determine more precisely the similarities or differences in the structural properties of the transition states for chemical and mechanical unfolding. Interestingly, a similar observation has recently been made for the enzyme barnase, a protein that naturally does not have a mechanical function. This protein unfolds slowly chemically ($k^0_{u,GnHCl} \sim 10^{-5}\ s^{-1}$) but, in contrast to I27, unfolds mechanically at a force of only ~65 pN at 600 nms^{-1} (Best et al., 2001).

The intrinsic unfolding rate constants, $\alpha^0_{u,F}$ determined using Monte Carlo simulations, which are intrinsically those of the monomeric form and not the polyprotein with n domains measured by experiment, are 3.3×10^{-4} s^{-1} (Carrion-Vazquez et al., 1999a) and 2.0×10^{-3} s^{-1} for $(I27)_8$ and $(I27)_5$*, respectively. The barrier to mechanical unfolding, therefore, is reduced upon mutation. Chemical denaturation experiments also show that C47S and C63S unfold more rapidly than the pseudo-wild type ($k^0_{u,GnHCl} = 4.9 \times 10^{-4}$ s^{-1} and 1.1×10^{-2} s^{-1} for wild type and C47S, C63S I27, respectively). The data show, therefore, that the barrier heights for chemical and mechanical unfolding respond differently to mutation $\alpha^0_{u,F}$ increases sixfold upon mutation, whereas $k^0_{u,GnHCl}$ increases 22-fold. This difference cannot simply be explained by changes in the thermodynamic stability of the native protein, because the native state is destabilized by 18 kJ mol^{-1} upon mutation ($\Delta\Delta G_{UN}$), whereas the transition states for unfolding determined mechanically and chemically are destabilized by ~5 and ~8 kJ/mol, respectively.

If proteins have a similar ΔG_u, then the energy $F.x_u$ (Eq. 1) must be of the same order to reduce the barrier height sufficiently to allow crossing by thermal fluctuations at the same rate. The importance of x_u to the observed unfolding forces has been reported previously (Rief et al., 1999). Proteins with the same ΔG_u, therefore, can respond very differently to force, depending on the magnitude of x_u. In proteins with a small x_u the work is done over a very short distance, allowing high forces to be withstood while maintaining structure. In contrast, proteins with a large x_u spread the same work over a greater distance, causing greater structural changes at low force. This observation suggests, therefore, that proteins that have evolved to withstand mechanical stress could generally unfold with a small x_u, irrespective of their topology or mechanical stability. Although the available data are currently too sparse to permit detailed correlations of this kind, it is interesting to note that for both the titin modules, I27 and I28, $x_u = 0.25$ nm (Li et al., 2000b). A similar value (0.3 nm) was obtained for extracellular matrix protein tenascin, a protein that is also exposed naturally to mechanical stress and is partly composed of a series of fibronectin type III domains (Oberhauser et al., 1998). For the helical structural protein, spectrin, $x_u = 1.5$ nm (Rief et al., 1999). By contrast, the x_u for a polymer of the naturally monomeric enzyme, T4 lysozyme, was indirectly calculated to be 0.81 nm (Yang et al., 2000), whereas the enzyme barnase was found to have similar pulling speed dependence to that of $(I27)_8$ (Best et al., 2001). From the available data, therefore, there appears to be no simple correlation between the natural function of a protein and the value of x_u. There may, however, be a correlation between this parameter and secondary structure and/or protein topology.

The emerging data suggest that the type of protein secondary structure, the number and geometry of interstrand hydrogen bonds, and the unusually late transition-state in the folding of I27 and I28 domains optimize their resistance to mechanical force (Li et al., 2000b). Indeed these domains unfold at the largest forces so far recorded for any domain (~200 and 260 pN for the wild-type I27 and I28, respectively (Li et al., 2000b)). By contrast, tenascin (FNIII domains), barnase, T4 lysozyme, the C2 domain of synaptotagmin I and spectrin unfold at 140 pN (Oberhauser et al., 1998), 65 pN (Best et al., 2001), 64 pN (Yang et al., 2000), 60 pN (Carrion-Vazquez et al., 2000), and 30 pN (Rief et al., 1999). Calmodulin unfolded at too small a force to measure (Carrion-Vazquez et al., 2000). The data suggest that proteins with β-sheet secondary structure are mechanically most stable, whereas α-helical proteins are relatively mechanically unstable. Proteins with mixed α/β topologies fall in between these two extremes. Steered molecular dynamics simulations have suggested that the mechanical stability of β-sheet proteins depends critically on the topology of the protein, proteins with parallel N- and C-terminal strands showing the greatest resistance to mechanical unfolding (Lu and Schulten, 1999). This results in force being applied orthogonally to interstrand hydrogen-bonds and requires all of these bonds to be broken simultaneously for significant extension to occur. By contrast, proteins with antiparallel terminal β strands unfold at relatively low forces, possibly because the force is applied parallel to the hydrogen bond and results in the sequential "zipper-like" rupture of these bonds with relatively low force. The most mechanically resistant proteins were found to possess a hydrogen bond network directly between N- and C-terminal strands, as is found for the immunoglobulin domains of titin (Lu and Schulten, 1999).

Comparison of the force required to unfold wild-type and mutant I27 domains clearly shows that the mutant form is less stable to mechanical stress (at all pulling speeds the mutant domains unfold with a lower force than that of the wild-type polymer (Fig. 7 *a*)). However, as well as changing the barrier height for mechanical unfolding, the mutation also affects the parameter, x_u (using Monte Carlo methods x_u wild type = 0.25 nm (Carrion-Vazquez et al., 1999a), x_u mutant = 0.29 nm). As stated above, the term $F.x_u$ equates to the energy required to reduce the barrier height sufficiently to allow crossing by thermal fluctuations. This product can be used, therefore, to assess the energetic effect of a mutation upon the mechanical transition state. At a pulling speed of 600 nms^{-1} the barrier for mechanical unfolding of wild-type I27 domains (for the polymer $(I27)_8$) is reduced by 31 kJ mol^{-1}. Interestingly, whereas the mutant concatamer unfolds at a significantly lower force than its wild-type counterpart ($\Delta F = 37$ pN at 600 nms^{-1}) it also has a higher x_u. As a consequence, the reduction in barrier height for unfolding of the mutant (29 kJ mol^{-1}) is the same as that of the wild-type domain, indicating that the mutations have no effect on the mechanical sensitivity of the protein (however, it should be noted that this analysis is sensitive to the

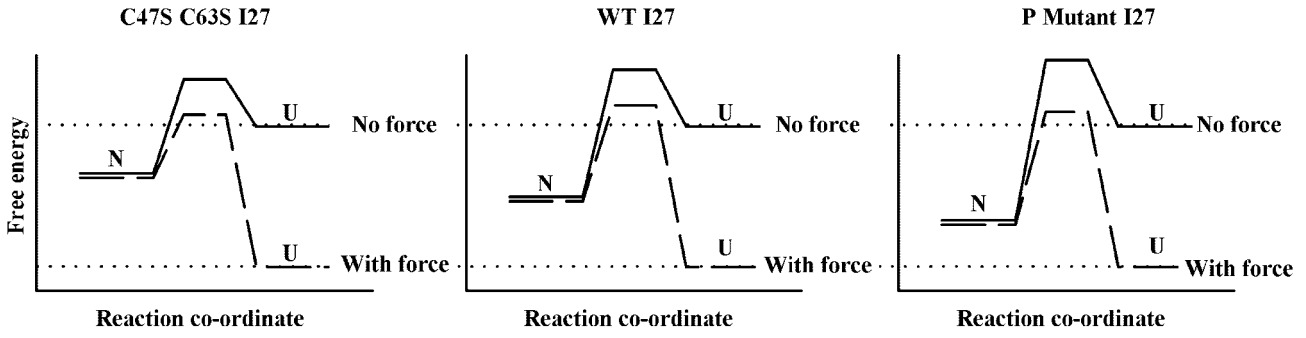

FIGURE 8 Schematic diagram showing effect of mutations and force on the unfolding of C47S, C63S I27, wild-type I27, and a proline mutant (P mutant) in the A′G strand interface. All proteins are referenced to a denatured state (*U*). This state is the fully unfolded state in denaturant for chemical unfolding and the "condensed denatured" state (the denatured state before polymer extension) for mechanical unfolding. The native state (*N*) of C47S, C63S I27 is destabilized relative to wild type, whereas the proline mutation stabilizes *N* due to the destabilization of the unfolded state (assuming that the native state structure is unperturbed by the introduction of a proline). Note: for the proline mutant the thermodymic properties have been inferred from mechanical folding/unfolding studies (Li et al., 2000a). The barrier height for unfolding of C47S, C63S I27 is reduced relative to that of the wild-type protein leading to an increase in $k^0_{u,GnHCl}$. By contrast, the proline mutation decreases $k^0_{u,GnHCl}$. Upon the application of force (shown as a dotted line), the transition state barrier for unfolding the wild-type and core destabilized I27 (C47S, C63S I27) behave similarly upon the application of force. However, the proline mutation causes a greater reduction in barrier height for the same displacement.

accuracy of the value derived for x_u we estimate that x_u varies by less than 10%). By examining the spread of values for x_a and $\alpha^0_{u,F}$ in Monte Carlo simulations that describe the best fit to the experimental data.

What is the origin of the mechanical strength of I27 domains? Previous molecular dynamics simulations have suggested that the six hydrogen bonds linking the A′ and G strands in the N- and C-terminal strands of the I27 domains are the source of their unusually high mechanical strength (Lu et al., 2000). Accordingly, rupture of these strands is thought to result in the facile and cooperative unfolding of the remaining strands in the domain. Interestingly, recent characterization of the conformational properties of the transition state for unfolding of I27 monomers using chemical denaturation and ϕ-value analysis has suggested that the A′ and G strands are also unfolded in the transition state for chemical denaturation (Fowler and Clarke, 2001). However, other regions are also substantially denatured in the transition state for chemical unfolding, most notably the A′-B loop and the E-F loop (the latter of which contains residue 63). By contrast, other regions of the I27 domain were found to be highly structured in the rate-limiting transition state for unfolding, particularly residues toward the center of the B, C, E, and F strands, which form a folding nucleus, similar to that found for the 3rd FNIII domain from human tenascin (Hamill et al., 2000). Our data on mechanical unfolding of I27 domains mutated at residues 47 and 63 are consistent with the predications of Li et al. (2000a), at least in that these residues (which lie in the D strand and the E-F loop) do not play a major role in determining the mechanical sensitivity of the domain. What role, if any, other residues play in mechanical unfolding will have to await further analysis of the structure of the mechanical unfolding transition state using polymers constructed from other mutant domains.

In previous studies, Li et al. (2000a) also used mutagenesis to assess the importance of the A′ and G strands in mechanical resistance by creating a series of proline mutants that ablated hydrogen bonds from this proposed mechanical clamp. Interestingly, whereas three of the mutants (V11P, V13P, and V15P) were less resistant to mechanical unfolding (they unfolded at lower force), a fourth mutant, V9P, was found to unfold at greater force than its wild-type counterpart. Pulling speed dependencies revealed that the transition state for unfolding of the three mechanically destabilized mutants had all moved closer to the denatured state, whereas that of V9P remained similar to that of the wild-type protein, suggesting that the mutations had varied effects on unfolding. These data can now be rationalized using the same argument as applied to the mutant concatamer described above. Thus, the amount of energy by which the transition state is lowered on displacement of its termini ($F.x_u$; in this example at 600 nms^{-1}) shows that all mutants behave in a similar fashion: each mutation causes a lowering of the unfolding barrier by 40 to 48 kJ mol^{-1}, a value significantly greater than that of the wild-type protein (31 kJ mol^{-1} at 600 nms^{-1}) (Fig. 8). Most importantly, the data for the cysteine mutants presented here indicate that the barrier is reduced by only 29 kJ mol^{-1} at the same pulling speed. The data confirm, therefore, that removal of one or more hydrogen bonds from the "mechanical clamp" significantly destabilizes the mechanical transition state. By contrast, mutation of C47 and C63 that lie mostly outside this region (there may be a small contact between the side chains

of V13 and C63) have little effect on the barrier to unfolding upon the displacement of its termini.

Mechanical resistance appears to be mainly a locally endowed feature: destabilization of the core has little effect upon mechanical sensitivity ($F.x_u$) (although the individual parameters, x_u, $\alpha^0_{u,F}$ and force do vary). By contrast, mutation of residues involved in the local hydrogen bond clamp region results in a more mechanically sensitive phenotype. Further mutagenesis studies are needed to map the transition state for mechanical unfolding of I27 to compare mechanical and chemical unfolding in more detail. The models for mechanical unfolding of proteins suggest that hydrogen bond topology/geometry is the key determinant of mechanical stability, irrespective of function. It is, therefore, important to mechanically unfold proteins with different topologies, as well as those with similar folds but different functions, so that general conclusions about the evolution of mechanical stability of different proteins can be formed.

We thank Andres Oberhauser for the kind gift of the wild-type $(I27)_8$ concatamer and several helpful discussions and Mathias Gautel for the original I27 clone. We also thank Jane Clarke and her laboratory for the use of their Asylum molecular force probe for purposes of comparison, the gift of their wild-type $(I27)_8$ concatamer and many interesting discussions. We thank Keith Ainley for technical support and Stan Gorski for creating Fig. 1. D.J.B and J.C. are funded by the University of Leeds and Biotechnology and Biological Sciences Research Council, R.C.Z. is funded by The Wellcome Trust and A.W.B. is funded by the Engineering and Physical Sciences Research Council. S.E.R. is a BBSRC Professorial Fellow. A.W.B., D.J.B., G.S.B., J.C., P.D.O., S.E.R., D.A.S., and J.T. are members of the Astbury Center for Structural Molecular Biology, which is part of the North of England Structural Biology Center (NESBIC) and is funded by the BBSRC.

REFERENCES

Barik, S. 1996. Site-directed mutagenesis in vitro by megaprimer PCR. *In* Methods in Molecular Biology, Vol. 57: In Vitro Mutagenesis Protocols. M.K. Trower editor. Humana Press Inc., Totowa, NJ. 203–215.

Bell, G. I. 1978. Models for the specific adhesion of cells to cells. *Science.* 200:618–627.

Best, R. B., L. Bin, A. Steward, V. Daggett, and J. Clarke. 2001. Can non-mechanical proteins withstand force: stretching barnase by atomic force microscopy and molecular dynamics simulation. *Biophys. J.* 81: 2344–2356.

Binnig, G., C. F. Quate, and C. Gerber. 1986. Atomic force microscope. *Phys. Rev. Lett.* 56:930–933.

Burnham, N. A., and R. J. Colton. 1989. Measuring the nanomechanical properties and surface forces of materials using an atomic force microscope. *J. Vac. Sci. Technol.* A7:2906–2913.

Carrion-Vazquez, M., P. E. Marszalek, A. F. Oberhauser, and J. M. Fernandez. 1999b. Atomic force microscopy captures length phenotypes in single proteins. *Proc. Natl. Acad. Sci. U. S. A.* 96:11288–11292.

Carrion-Vazquez, M., A. F. Oberhauser, T. E. Fisher, P. E. Marszalek, H. Li, and J. M. Fernandez. 2000. Mechanical design of proteins studied by single-molecule force spectroscopy and protein engineering. *Prog. Biophys. Mol. Biol.* 74:63–91.

Carrion-Vazquez, M., A. F. Oberhauser, S. B. Fowler, P. E. Marszalek, S. E. Broedel, J. Clarke, and J. M. Fernandez. 1999a. Mechanical and chemical unfolding of a single protein: a comparison. *Proc. Natl. Acad. Sci. U. S. A.* 96:3694–3699.

Evans, E. 1998. Energy lanscapes of biomolecular adhesion and receptor anchoring at interfaces explored with dynamic force spectroscopy. *Faraday Discuss.* 111:1–16.

Evans, E., and K. Ritchie. 1997. Dynamic strength of molecular adhesion bonds. *Biophys. J.* 72:1541–1555.

Ferguson, N., A. P. Capaldi, R. James, C. Kleanthous, and S. E. Radford. 1999. Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9. *J. Mol. Biol.* 286: 1597–1608.

Florin, E. L., V. T. Moy, and H. E. Gaub. 1994. Adhesive forces between individual ligand-receptor pairs. *Science.* 264:415–417.

Florin, E. L., M. Rief, H. Lehmann, M. Ludwig, C. Dornmair, V. T. Moy, and H. E. Gaub. 1995. Sensing specific molecular interactions with the atomic force microscope. *Biosens. Bioelectron.* 10:895–901.

Fowler, S. B., and J. Clarke. 2001. Mapping the folding pathway of an immunoglobulin domain: structural detail from phi value analysis and movement of the transition state. *Structure.* 9:1–12.

Grandbois, M., M. Beyer, M. Rief, H. Clausen-Schaumann, and H. E. Gaub. 1999. How strong is a covalent bond? *Science.* 283:1727–1730.

Hamill, S. J., A. Steward, A., and J. Clarke. 2000. The folding of an immunoglobulin-like greek key protein is defined by a common-core nucleus and regions constrained by topology. *J. Mol. Biol.* 297:165–178.

Improta, S., J. K. Krueger, M. Gautel, R. A. Atkinson, J.-F. Lefèvre, S. Moulton, J. Trewhella, and A. Pastore. 1998. The assembly of immunoglobulin-like modules in titin: implications for muscle elasticity. *J. Mol. Biol.* 284:761–777.

Improta, S., A. S. Politou, and A. Pastore. 1996. Immunoglobulin-like modules from titin I-band: extensible components of muscle elasticity. *Structure.* 4:323–337.

Kellermayer, M. S. Z., S. B. Smith, H. L. Granzier, and C. Bustamante. 1997. Folding-unfolding transitions in single titin molecules characterized with laser tweezers. *Science.* 276:1112–1116.

Klimov, D. K., and D. Thirumalai. 2000. Native topology determines force-induced unfolding pathways in globular proteins. *Proc. Natl. Acad. Sci. U. S. A.* 97:7254–7259.

Kraulis, P. J. 1991. Molscript-a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24:946–950.

Lee, G. U., L. A. Chrisey, and R. J. Colton. 1994. Direct measurement of the forces between complimentary strands of DNA. *Science.* 266: 771–773.

Li, H., M. Carrion-Vazquez, A. F. Oberhauser, P. E. Marszalek, and J. M. Fernandez. 2000a. Point mutations alter the mechanical stability of immunoglobulin modules. *Nat. Struct. Biol.* 7:1117–1120.

Li, H., A. F. Oberhauser, S. B. Fowler, J. Clarke, and J. M. Fernandez. 2000b. Atomic force microscopy reveals the mechanical design of a modular protein. *Proc. Natl. Acad. Sci. U. S. A.* 97:6527–6531.

Lu, H., B. Isralewitz, A. Krammer, V. Vogel, and K. Schulten. 1998. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys. J.* 75:662–671.

Lu, H., and K. Schulten. 1999. Steered molecular dynamics simulations of force-induced protein domain unfolding. *Proteins Struct. Funct. Genet.* 35:453–463.

Lu, H., and K. Schulten. 2000. The key event in force-induced unfolding of titin's immunoglobulin domains. *Biophys. J.* 79:51–65.

Makarov, D. E., P. K. Hansma, and H. Metiu. 2001. Kinetic Monte Carlo simulation of titin unfolding. *J. Chem. Phys.* 114:9663–9673.

Marszalek, P. E., H. Lu, H. Li, M. Carrion-Vazquez, A. F. Oberhauser, K. Schulten, and J. M. Fernandez. 1999. Mechanical unfolding intermediates in titin modules. *Nature.* 402:100–103.

Merkel, R., P. Nassoy, A. Leung, K. Ritchie, and E. Evans. 1999. Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature.* 397:50–53.

Merritt, E. A., and M. E. P. Murphy. 1994. Raster3D Version 2.0- a program for photorealistic molecular graphics. *Acta Crystallogr. D.* 50:869–873.

Mitsui, K., M. Hara, and A. Ikai. 1996. Mechanical unfolding of α2−macroglobulin molecules with atomic force microscope. *FEBS Lett.* 385: 29–33.

Myers, J. K., C. N. Pace, and J. M. Scholtz. 1995. Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* 4:2138–2148.

Oberdörfer, Y., H. Fuchs, and A. Janshoff. 2000. Conformational analysis of native fibronectin by means of force spectroscopy. *Langmuir.* 16: 9955–9958.

Oberhauser, A. F., P. E. Marszalek, H. P. Erickson, and J. M. Fernandez. 1998. The molecular elasticity of the extracellular matrix protein tenascin. *Nature.* 393:181–185.

Oesterhelt, F., D. Oesterhelt, M. Pfeiffer, A. Engel, H. E. Gaub, and D. J. Müller. 2000. Unfolding pathways of individual bacteriorhodopsins. *Science.* 288:143–146.

Paci, E., and M. Karplus. 2000. Unfolding proteins by external forces and temperature: the importance of topology and energetics. *Proc. Natl. Acad. Sci. U. S. A.* 97:6521–6526.

Rief, M., J. M. Fernandez, and H. E. Gaub. 1998. Elastically coupled two-level systems as a model for biopolymer extensibility. *Phys. Rev. Lett.* 81:4764–4767.

Rief, M., M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub. 1997b. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science.* 276:1109–1111.

Rief, M., F. Oesterhelt, B. Heymann, and H. E. Gaub. 1997a. Single molecule force spectroscopy on polysaccharides by atomic force spectroscopy. *Science.* 275:1295–1297.

Rief, M., J. Pascual, M. Saraste, and H. E. Gaub. 1999. Single molecule force spectroscopy of spectrin repeats: low unfolding forces in helix bundles. *J. Mol. Biol.* 286:553–561.

Santoro, M., and D. W. Bolen. 1988. Unfolding free energy changes determined by the linear extrapolation method: 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry.* 27:8063–8068.

Smith, B. L., T. E. Schäffer, M. Viani, J. B. Thompson, N. A. Frederick, J. Kindt, A. Belcher, G. D. Stucky, D. E. Morse, and P. K. Hansma. 1999. Molecular mechanistic origin of the origin of the toughness of natural adhesives, fibres and composites. *Nature.* 399:761–763.

Trinick, J. 1996. Cytoskeleton: titin as a scaffold and spring. *Curr. Biol.* 6:258–260.

Tskhovrebova, L., J. Trinick, J. A. Sleep, and R. M. Simmons. 1997. Elasticity and unfolding of single molecules of the giant muscle protein titin. *Nature.* 387:308–312.

Yang, G., C. Cecconi, W. A. Baase, I. R. Vetter, W. A. Breyer, J. A. Haack, B. W. Matthews, F. W. Dahlquist, and C. Bustamante. 2000. Solid-state synthesis and mechanical unfolding of polymers of T4 lysozyme. *Proc. Natl. Acad. Sci. U. S. A.* 97:139–144.