Unfolding of Titin Immunoglobulin Domains by Steered Molecular Dynamics Simulation

Hui Lu,*# Barry Isralewitz,*§ André Krammer,‖ Viola Vogel,‖** and Klaus Schulten*§¶

*Beckman Institute for Advanced Science and Technology, #Department of Nuclear Engineering, §Center for Biophysics and Computational Biology, ¶Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, and Departments of ‖Physics and **Bioengineering, University of Washington, Seattle, Washington 98195 USA

ABSTRACT Titin, a 1-μm-long protein found in striated muscle myofibrils, possesses unique elastic and extensibility properties in its I-band region, which is largely composed of a PEVK region (70% proline, glutamic acid, valine, and lysine residue) and seven-strand β-sandwich immunoglobulin-like (Ig) domains. The behavior of titin as a multistage entropic spring has been shown in atomic force microscope and optical tweezer experiments to partially depend on the reversible unfolding of individual Ig domains. We performed steered molecular dynamics simulations to stretch single titin Ig domains in solution with pulling speeds of 0.5 and 1.0 Å/ps. Resulting force-extension profiles exhibit a single dominant peak for each Ig domain unfolding, consistent with the experimentally observed sequential, as opposed to concerted, unfolding of Ig domains under external stretching forces. This force peak can be attributed to an initial burst of backbone hydrogen bonds, which takes place between antiparallel β-strands A and B and between parallel β-strands A′ and G. Additional features of the simulations, including the position of the force peak and relative unfolding resistance of different Ig domains, can be related to experimental observations.

INTRODUCTION

The giant muscle protein titin, also known as connectin, is a ~30,000 amino acid long filament that spans half of the sarcomere and plays a number of important roles in muscle contraction and elasticity (Labeit et al., 1997; Maruyama, 1997; Kellermayer and Granzier, 1996; Wang et al., 1993). During muscle contraction, titin, which is anchored at the Z-disk and at the M-line, exerts a passive force that keeps sarcomere components uniformly organized. The passive force developed in titin during muscle stretching restores sarcomere length when the muscle is relaxed. Titin is composed of ~300 repeats of two types of domains, fibronectin type III-like (Fn-3) domains and immunoglobulin-like (Ig) domains, and the PEVK (70% proline, glutamic acid, valine, and lysine residue) region (Labeit and Kolmerer, 1995). The Fn-3 domains are located only in the A-band of the molecule, the PEVK region is located in the I-band, and the Ig domains are distributed along the whole length of titin.

The region of titin located in the sarcomere I-band is believed to be responsible for titin's extensibility and passive elasticity (Erickson, 1994; Linke et al., 1996; Granzier et al., 1996; Greaser et al., 1996). The I-band region of titin consists mainly of two tandem regions of Ig domains, separated by the PEVK region. The Ig domains each form β-sandwich structures, but the PEVK region does not hold a stable conformation, because of the charges on its glutamic acid and lysine residues. Among titins expressed in different muscle tissues, both the lengths of PEVK regions and the number of Ig domains composing the tandem regions vary greatly. Lengths of PEVK regions range from 163 residues in cardiac titin to 2174 residues in skeletal titin; the number of tandem Ig domains ranges from 37 in cardiac titin to 90 in skeletal titin (Labeit and Kolmerer, 1995).

To understand titin's function as a molecular spring, recent work has involved direct observations of titin's response to applied force. Atomic force microscope (AFM) (Rief et al., 1997) and optical tweezer (Kellermayer et al., 1997; Tskhovrebova et al., 1997) experiments directly measure the force extension profile of single titin molecules. In an AFM experiment (Rief et al., 1997), a single titin molecule was stretched at constant speed, the force-extension profile showing a sawtooth-like pattern with ~25 nm spacing between the force peaks. The same experiment also involved stretching of an eight Ig domain and a four Ig domain section of the I-band of titin, and convincingly demonstrated that every force peak corresponds to a single Ig domain unfolding.

These experiments, along with data from immunofluorescence electron microscopy experiments (Linke et al., 1996) that track the lengths of different sections of titin during extension, provide a detailed picture of titin's response to forced elongation (see Fig. 1). Under small extensions past the resting length, the tandem Ig domains straighten from a randomly aligned state to a taut aligned state, with each domain itself still folded. At medium extensions, the PEVK region extends from a random or semifolded state to an extended state. At further extensions, the tandem Ig domains unfold one by one, increasing the titin length by ~300 Å for each unfolded domain. Titins expressed in

Received for publication 10 December 1997 and in final form 4 May 1998.

Address reprint requests to Dr. Klaus Schulten, Department of Physics, Beckman Institute 3147, University of Illinois, 405 N. Mathews Ave., Urbana, IL 61801. Tel.: 217-244-1604; Fax: 217-244-6078; E-mail: kschulte@ks.uiuc.edu.

0006-3495/98/08/662/10 $2.00

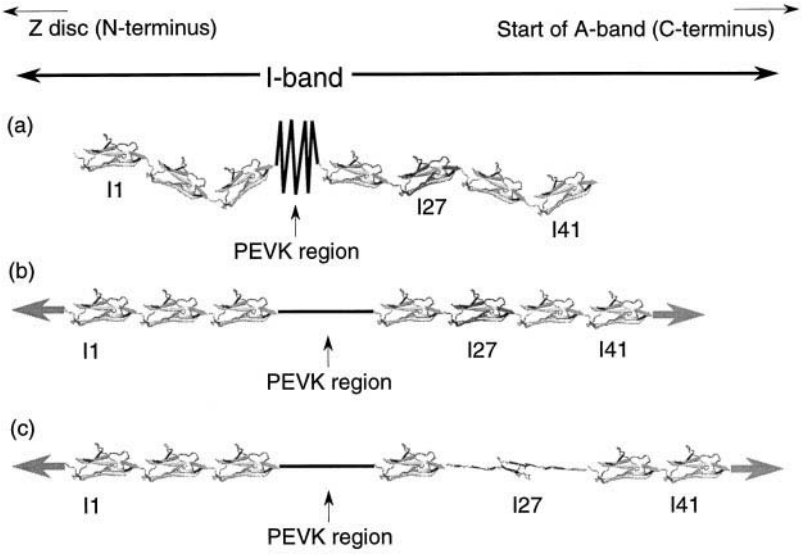

FIGURE 1 Cartoon of titin I-band function. (Actual I-band contains 41 Ig domains (Rief et al., 1997).) Ig domains are explicitly shown, with the PEVK region depicted as a heavy black line. Horizontal gray arrows indicate the external force. (*a*) Titin I-band resting structure. (*b*) Titin I-band with PEVK region extended. (*c*) Titin I-band with Ig domain partially unfolded. This figure was created with VMD (Humphrey et al., 1996).

different muscle types differ in passive forces exerted over physiological extension ranges; these are likely determined by differing lengths of the tandem Ig and titin PEVK domains. Extensions of titin long enough to cause Ig domains to unfold are not sure to occur during normal extensions in skeletal muscle, but are likely to occur around the physiological limits in cardiac titin (Granzier et al., 1997).

Several important features differentiate Ig domains in the I-band region from those in other regions of titin. The Ig domains in titin I-band all share similar, stable folds with high homology (Politou et al., 1994; Pfuhl et al., 1995; Pfuhl and Pastore, 1995; Improta et al., 1996). The domains are made up of a smaller number of amino acids, and contain shorter loops between β-strands, than do other Ig domains. The Ig domains appearing in tandem are directly connected to each other, with no intervening linker sequences between them.

The experimentally solved I27 domain belongs to one of the two tandem Ig sections in titin I-band. This domain adopts the typical I-frame immunoglobulin superfamily fold (Politou et al., 1995; Harpaz and Chothia, 1994), consisting of two β-sheets packing against each other (Fig. 2 *a*), with each sheet containing four strands. The first sheet comprises strands A, B, E, and D, the second sheet A′, G, F, and C (Fig. 2 *b*). All adjacent β-strands in both sheets are antiparallel to each other, except for the parallel pair A′ and G. The β-strands A and A′ belong to different sheets, but are part of the N-terminal strand, with a β-bulge turn at Glu^5 and Lys^6 (Chan et al., 1993). The structure is stabilized by hydrophobic core interactions between the two β-sheets and by the hydrogen bonds between β-strands.

Experiment has not yet resolved the details of unfolding of a single Ig domain, or explained the structure-function relationship of Ig domains. Computer modeling can help to reveal this relationship through an atomic-level description of force-induced titin domain unfolding.

During the last several years, molecular dynamics simulations of unfolding solvated proteins have been reported (e.g., in Karplus and Sali, 1995). These studies utilized a variety of techniques to decrease the computation time needed to observe unfolding, such as highly elevated temperatures at constant volume (e.g., Li and Daggett, 1996; Tirado-Rives et al., 1997), or elevated pressure and application of radial forces (Hunenberger et al., 1995). The unnatural conditions employed may have influenced the unfolding pathway explored by simulations. In this study, the forces applied are intended to accelerate the unfolding of Ig domains, while following the actual forced unfolding pathway in functional muscle and in experiments.

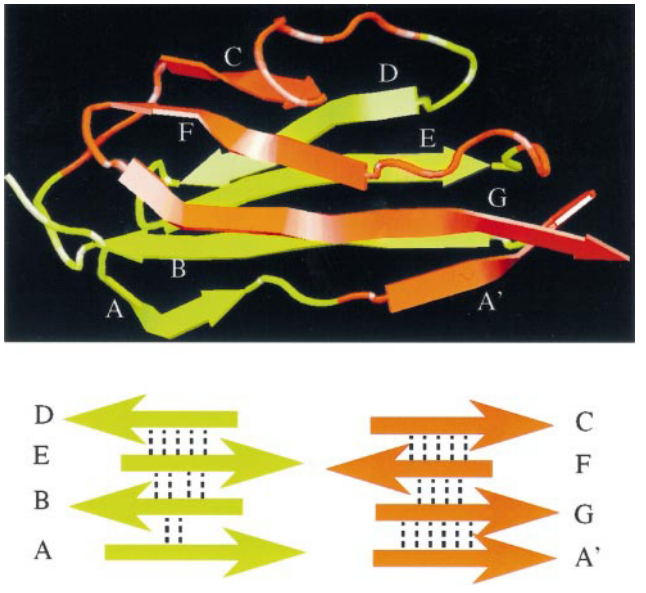

FIGURE 2 (*a*) Secondary structure of the cardiac titin I27 immunoglobulin domain from NMR data (Improta et al., 1996). β-sheets are colored differently, with sheet A B E D in green and sheet A′ G F C in orange. (*b*) Schematic view of all β-sheets and backbone hydrogen bonds (*dotted lines*) between adjacent β-strands. This figure was created with VMD (Humphrey et al., 1996).

The steered molecular dynamics (SMD) technique recently introduced (Izrailev et al., 1997, 1998; Evans and Ritchie, 1997; Balsera et al., 1997; Grubmüller et al., 1996; Leech et al., 1996) is perfectly suited to such study. SMD methods, which involve the application of external forces to molecules in molecular dynamics simulations, have already been used to describe several ligand-protein interactions, e.g., biotin-avidin (Izrailev et al., 1997), retinal and bacterioopsin (Isralewitz et al., 1997), phosphate and actin (Wriggers and Schulten, manuscript in preparation), and lipid and phospholipid A_2 (Stepaniants et al., 1997). Here the SMD method is extended to examine the unfolding of an entire protein domain, an unfolding that is caused in vivo and in vitro by external forces, albeit on a longer time scale than that covered by SMD.

METHODS

The molecular dynamics simulation of stretching titin Ig subunits was carried out with the programs XPLOR (Brünger, 1992) and NAMD (Nelson et al., 1996), with the CHARMM19 force field. The simulations started from an experimentally solved Ig domain, I27 of the I-band of cardiac titin, using the energy-minimized average NMR structure (Improta et al., 1996), deposited as entry 1TIT in the Brookhaven Protein Data Bank (Bernstein et al., 1977).

To identify important conserved unfolding features among Ig domains, SMD simulations performed on Ig domains I27 and I28 were compared.

All Ig domains in the I-band have highly homologous amino acid sequences. Currently I27 is the only I-band Ig with an experimentally solved structure and hence was selected for investigation. However, other Ig domain structures can be built by means of homology modeling. This approach was applied to I28, the Ig domain immediately adjacent to I27, which shows 47% homology and 25% identity with I27 (Fig. 3). Adopting the alignment suggested by Improta et al. (1996), homology modeling of I28 was performed based on the I27 backbone structure, followed by minimization to yield a sufficiently stable structure for MD simulation. From this resulted an I28 structure with backbone atom coordinate RMSD of 1.8 Å relative to equilibrated I27.

The structure of I27 was solvated with the TIP3P model for water (Jorgensen et al., 1983) before SMD simulations. A cube was constructed from 64 identical templates of water molecules provided by AMBER 4.1 (Pearlman et al., 1995). Each template consisted of 216 Monte Carlo-equilibrated water molecules. The Ig domain was placed in the center of the water box, and all water molecules more than 31 Å from the center of the protein, within 2.6 Å of the protein surface, or within the volume occupied by the protein were deleted. The resulting structure, I27 surrounded by a water bubble, has the protein surface covered everywhere by at least four shells of water molecules.

The water-protein system was gradually heated over 10 ps to 300 K, then equilibrated with a thermal bath at 300 K for another 10 ps. During this process the water molecules composing the outer 3-Å shell of the system were harmonically restrained to their original positions to maintain the shape of the water bubble. This was followed by 30 ps of free dynamics without constraints or heat bath, but with harmonic restraints of the outer water molecules still maintained. The free dynamics run exhibited a temperature fluctuation of 5 K and an RMSD of 1 Å from the backbone of the NMR structure (Improta et al., 1996). Finally, 2 ps of free dynamics was performed, with no harmonic restraints set. The same procedure was also applied to solvate the I28 domain.

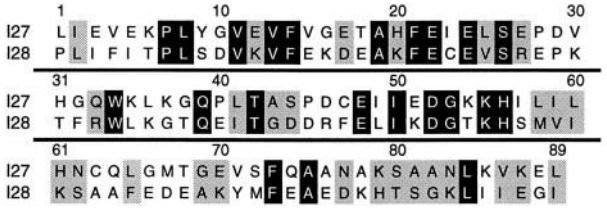

FIGURE 3 Alignment of cardiac titin Ig domains I27 and I28. In each column, gray shading indicates homology, inverse text indicates identity.

In the simulations performed, all atoms, including waters and all hydrogens, were modeled explicitly, with each system containing ~11,000 atoms. The simulations were performed with a time step of 1 fs, a uniform dielectric constant of 1, and a cut-off of Coulomb forces with a switching function starting at a distance of 10 Å and reaching zero at 13 Å.

SMD simulations were carried out by fixing one terminus of the domain, and applying external forces to the other terminus. The forces were applied by restraining the pulled end harmonically to a restraint point and moving the restraint point with a constant velocity v in the desired direction. The procedure is equivalent to attaching one end of a harmonic spring to the end of the domain and pulling on the other end of the spring, and is similar to the procedure performed upon whole titin and Ig repeats in AFM experiments (Rief et al., 1997), except that the pulling speeds adopted in the simulations are six to eight orders of magnitude higher than those in the experiments. Force-induced unfolding processes with several choices of pulling positions were simulated. For each domain (I27, I28), two pulling speeds were applied, 0.5 Å/ps and 1 Å/ps. For this purpose one end (C_α of Leu^1 in I27, Pro^1 in I28) of a domain was fixed, and the other end (C_α of Glu^{88} in I27, Leu^{89} in I28) was pulled.

The forces experienced by the C_α atom of a pulled residue are

$$F = k(vt - x). \tag{1}$$

Here x is the displacement of the pulled atom from its original position, and k is the spring constant. The pulling direction was chosen along the vector from fixed atom to pulled atom. To ensure that the direction of pulling did not affect the unfolding process, an I27 unfolding simulation was conducted by fixing C_α^{88} and pulling on C_α^1, and another was conducted by pulling on C_α^1 and C_α^{88} in opposite directions along the line connecting these atoms.

The SMD simulations presented are referred to below in the form (terminal pulled)-(Ig domain number)-(pulling speed in Å/ps). For example, C-28-1.0 represents a simulation pulling on the C-terminal of I28 at a speed of 1.0 Å/ps. The value of k (see Eq. 1) was set at 10 $k_BT/Å^2$, corresponding to a spatial (thermal) fluctuation of the constrained C_α atom of $\delta x = \sqrt{k_BT/k} = 0.32$ Å at $T = 300$ K. The simulations with the values of k and v chosen here correspond to pulling with a stiff spring in the drift regime (Evans and Ritchie, 1997; Izrailev et al., 1997; Balsera et al., 1997). To realize a movement of the restraint point with nearly constant velocity, the position of the restraint point was changed every 100 fs by $v\Delta t$. For the force-extension curve, as shown in Fig. 5, the force was calculated according to Eq. 1.

RESULTS

Six SMD simulations of titin unfolding were performed (listed in Table 1). All simulations began with an equilibrated folded structure and were stopped when a fully extended polypeptide was produced. In the following, the simulation chiefly discussed is C-27-0.5, the extension of I27 with pulling speed 0.5 Å/ps, which was stopped at an extension of 285 Å, after 578 ps of simulation. The features of this simulation are seen in all other simulations. Initially the C-terminal and the N-terminal strands extend gradually while loops BC and CD, which both connect the two β-sheets, also become extended (Fig. 4 *a*). At this time, the two β-sheets slide away from each other, but each maintains

TABLE 1 Characterization of SMD simulations

Simulation	Ig domain	Pulled AA (terminus)	Fixed AA	Speed	Force peak (pN)	Extension at force peak (Å)
C-27-0.5	I27	Glu^{88} (C)	Leu^{1}	0.5	2040	14.6
C-27-1.0	I27	Glu^{88} (C)	Leu^{1}	1.0	2440	18.8
N-27-1.0	I27	Leu^{1} (N)	Glu^{88}	1.0	2360	16.3
NC-27-1.0	I27	Leu^{1} and Glu^{88} (N, C)	None	1.0	2318	18.8
N-28-0.5	I28	Leu^{89} (C)	Pro^{1}	0.5	2082	16.7
N-28-1.0	I28	Leu^{89} (C)	Pro^{1}	1.0	2554	19.8

AA, Amino acid.

a stable structure as well as its intrasheet backbone hydrogen bonds. As the pulling process continues, reaching an extension of the domain of ~14 Å, the structure within each sheet begins to break: in the A′GFC sheet, strands A′ and G slide past each other; in the ABED sheet, A and B slide past each other (Fig. 4 *b*). This motion marks the beginning of the Ig domain collapse, after which the domain gradually unfolds (Fig. 4 *c*) and the strands unravel one by one, with strands E and D the last to unfold. At an extension of 260 Å, the domains are completely straightened and unfolded (Fig. 4 *d*).

Fig. 5 shows the force-extension profile for the C-27-0.5 simulation. Initially, the force applied by the external spring increases until the extension reaches ~14 Å, where the

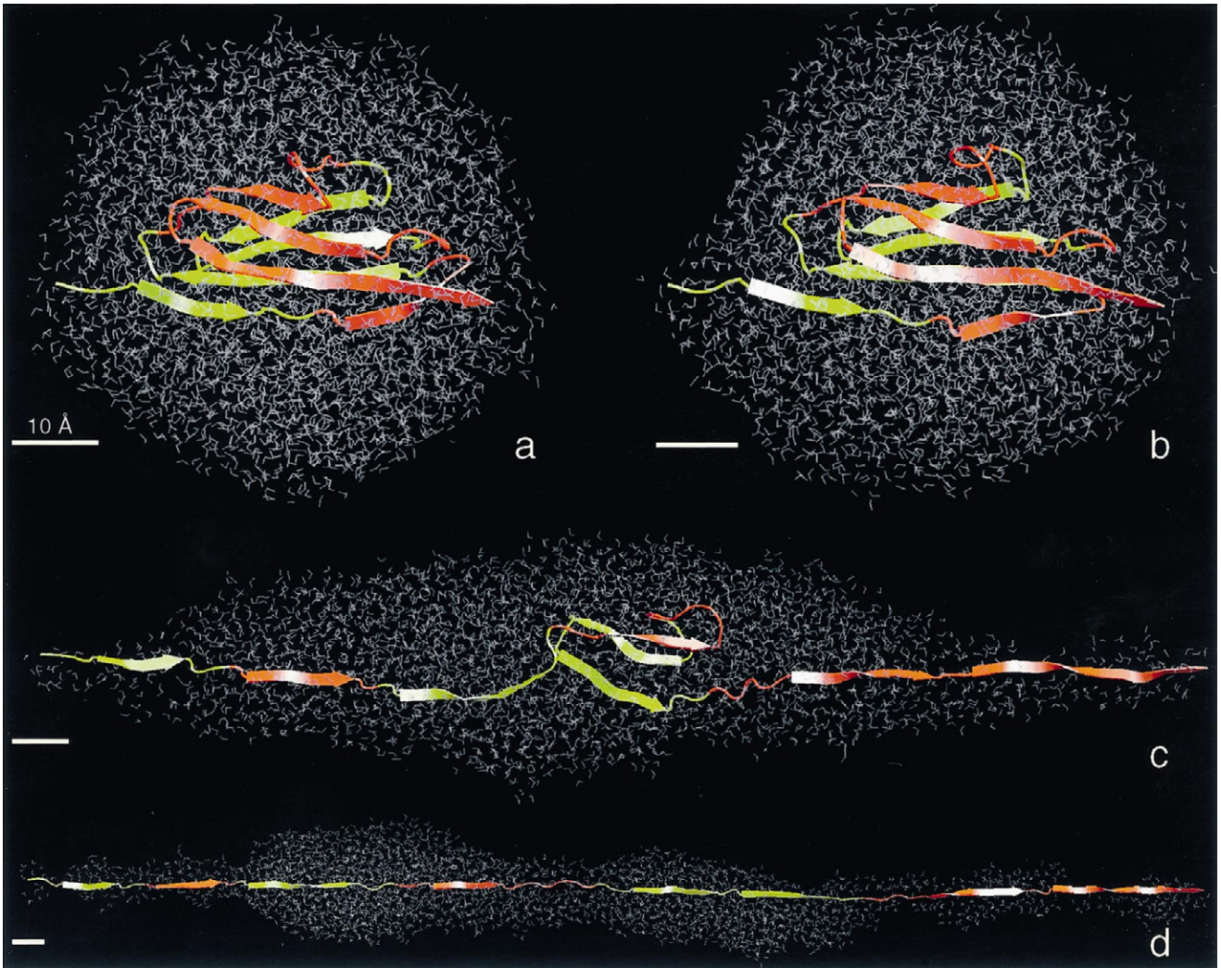

FIGURE 4 The intermediate stages of pulling simulations. The protein domain I27 (residues 1–88) is drawn in cartoon representation with the two β-sheets presented in different colors, and water molecules are drawn in line representation. (*a*) Region I, preburst, at extension 10 Å. (*b*) Region II, immediately after the major burst, at extension 17 Å. (*c*) Region III, postburst, at extension 150 Å. (*d*) Region IV, fully extended domain, at extension 285 Å. The bar at the lower left corner of each figure represents 10 Å. This figure was created with VMD (Humphrey et al., 1996).

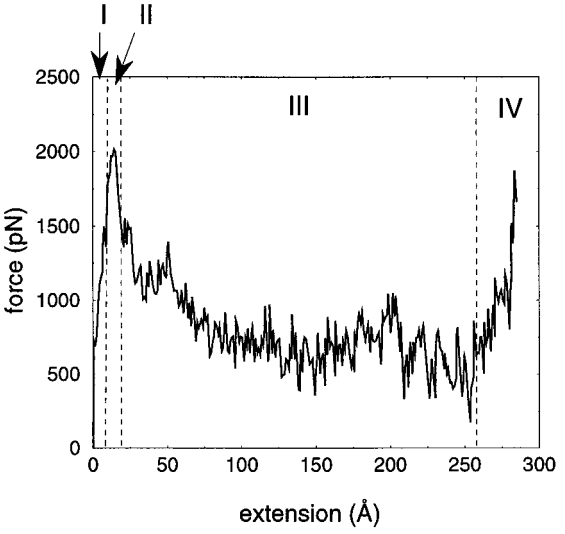

FIGURE 5 Force extension profile of SMD simulations for I27 with a pulling speed of 0.5 Å/ps. The extension domain is divided into four regions: I, preburst; II, major burst; III, postburst; IV, pulling of fully extended chain.

force peaks at ~2000 pN. The force then drops rapidly to 1500 pN within 3 Å and drops further by an additional 300 pN at an extension of ~22 Å. With further stretching, the force drops gradually, reaching 700 pN at 60 Å, then remains relatively constant until an extension of 260 Å, where the force begins to increase again. Beyond 260 Å, the fully unfolded polypeptide is being stretched.

The unfolding process can be divided into four regions: I) preburst, from 0 to 10 Å, during which the protein maintains β-sheet structure and the external force remains smaller than 1500 pN; II) main burst, from 10 to 17 Å, during which the secondary structure begins to break down; III) postburst, from 17 to 260 Å, during which the protein unravels; IV) fully extended (>260 Å), during which the polypeptide chain approaches its maximum length. Simulation C-27-1.0, with the faster pulling speed of 1 Å/ps, results in a similar series of processes and a similar force-extension profile, with a slightly higher peak value of 2440 pN during the main burst.

Other simulations showed similar features of the unfolding process and force profiles with only small variations in force peak value and degree of extension at the force peak. The results are summarized in Table 1.

Observation of hydrogen bond participants revealed that I27 unfolding resistance under external forces can be attributed to inter-β-strand, intra-β-sheet backbone hydrogen bond breaking. Table 2 presents the backbone hydrogen bonds from different β-strands and their involvement in the unfolding process. The distances between hydrogen bond participants of different β-strands during the unfolding are shown in Fig. 6. In region I, all hydrogen bonds remain stable, i.e., the domain maintains its β-sandwich structure. At a slightly larger extension, the hydrogen bonds between strands A and B and between strands A′ and G break nearly simultaneously, the rupture coinciding with the peak of the unfolding force shown in Fig. 5. Fig. 7 compares the hydrogen bonding in two pairs of β-strands, strands A′ and G (involved in the initial burst) and strands G and F (not involved in the initial burst), during the C-27-0.5 simulation. In the initial structure, sheet structures exhibit stable hydrogen bonds. The hydrogen bonds are maintained well in the preburst region (I); however, in the postburst region (III), all hydrogen bonds between A′ and G are broken, whereas those between F and G are maintained.

Once the A′/G and A/B hydrogen bonds are broken, the protein unfolds rapidly; this later unfolding is relatively smooth. The third stage of unfolding is dominated by frictional forces and involves changing the water bubble shape, as well as breaking individual inter-β-strand hydrogen bonds, i.e., backbone hydrogen bonds break one by one instead of in clusters, as in the main burst region. After the main burst, when the domain begins to unfold, the hydrophobic core region gradually becomes exposed to water and destabilizes.

Analysis of the energy between hydrogen bonding atoms provides a complementary view of the unfolding process. Fig. 8 presents individual interaction energies of the hydro-

TABLE 2 Backbone hydrogen bonding pairs of I27

β-strand partners	Connected residues (atoms)	Extension at breakage (Å)
A′G	Tyr^{9}(O)—Asn^{83}(H)	14
	Val^{11}(H)—Asn^{83}(O)	15
	Val^{11}(O)—Lys^{85}(H)	14
	Val^{13}(H)—Lys^{85}(O)	15
	Val^{13}(O)—Lys^{87}(H)	14
	Val^{15}(H)—Lys^{87}(O)	14
AB	Lys^{6}(H)—Glu^{24}(O)	9
	Lys^{6}(O)—Glu^{24}(H)	13
FG	Gly^{69}(H)—Leu^{84}(O)	43
	Gly^{69}(O)—Leu^{84}(H)	41
	Val^{71}(H)—Ala^{82}(O)	37
	Val^{71}(O)—Ala^{82}(H)	35
CF	Gln^{33}(H)—Gln^{74}(O)	43
	Gln^{33}(O)—Gln^{74}(H)	57
	Lys^{35}(H)—Ser^{72}(O)	77
	Lys^{35}(O)—Ser^{72}(H)	106
BE	Ala^{19}(H)—Leu^{60}(O)	48
	Ala^{19}(O)—Leu^{60}(H)	66
	Phe^{21}(H)—Leu^{58}(O)	72
	Phe^{21}(O)—Leu^{58}(H)	102
	Ile^{23}(H)—His^{56}(O)	132
	Ile^{23}(O)—His^{56}(H)	141
	Leu^{25}(H)—Lys^{54}(O)	136
DE	Glu^{48}(H)—Ile^{59}(O)	145
	Glu^{48}(O)—Ile^{59}(H)	202
	Ile^{50}(H)—Ile^{57}(O)	202
	Ile^{50}(O)—Ile^{57}(H)	109
	Asp^{52}(H)—Lys^{55}(O)	104

Hydrogen bond pairs with residue and participant atoms listed. O is backbone oxygen and H is the backbone hydrogen connected to N. Pairs were selected if the NH-O distance was stably less than 3.5 Å and the NH-O angle was smaller than 45° in the equilibrated I27 structure. The extension point of a breakage was recorded as the point when the H-O interaction energy exceeded −5.0 kcal/mol.

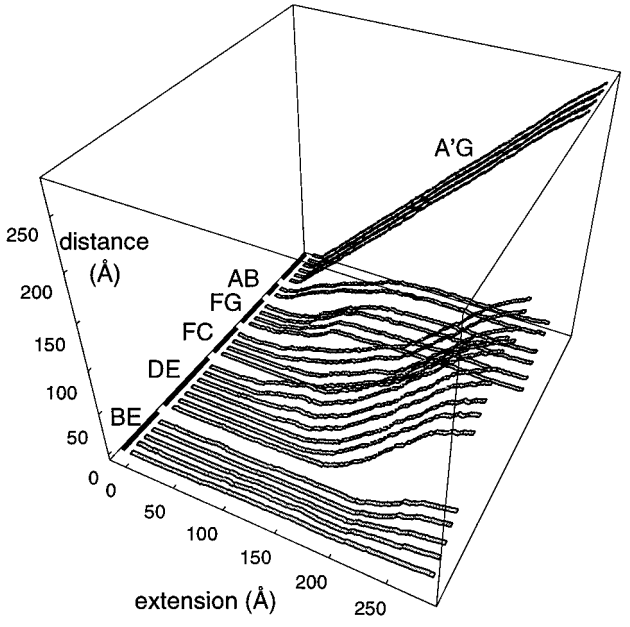

FIGURE 6 Individual distances versus extension for interstrand hydrogen bond pairs in simulation C-27-0.5. Bond pair distances are grouped by β-strand pairs, ordered from far to near according to the ordering in Table 2.

gen bond pairs in I27. One can discern features similar to those in Fig. 6, depicting the hydrogen bond distances. Fig. 9 shows the total interaction energies of atoms involved in hydrogen bonding. One can recognize a sharp increase occurring around an extension of 14 Å. Comparison with the force-extension peak in Fig. 5 strongly suggests that the backbone hydrogen bonds make up the most important component of the force profile recorded.

We have also monitored the solvent-accessible surface area of hydrophobic and backbone residues during the SMD unfolding process, as shown in Fig. 10. The accessibility of hydrophobic residues increases from 2000 Å^2 to ~7000 Å^2. The increase is rather uniform during the first 80-Å extension and, in particular, does not reflect the burst event at an extension of 14 Å. The accessibility of backbone residues, however, reflects unfolding events more closely. The curve shows a plateau in the preburst region I, from 0–10 Å extension, which implies that the backbone structure is stable in this region. At larger extensions (10–80 Å), the accessible surface exhibits a linear increase; a second linear increase with smaller slope arises for 80–285 Å extension; the reduced rate of solvent exposure is likely due to the fact that the inner faces of the β-sandwich have become completely exposed at 80 Å extension.

DISCUSSION

Applied force experiments (Rief et al., 1997; Kellermayer et al., 1997; Tskhovrebova et al., 1997) have elucidated the chief design requirements for titin I-band Ig domains under extreme stretch conditions: they must unravel one by one, and must increase the length of titin at each unraveling event by a set amount without affecting the stability of those domains that still remain folded. The aim of the present paper is to explain, on the basis of atomic-level simulations, how the architecture of titin's Ig domains controls extension.

At small extensions of titin, when the link regions are pulled taut to form a straight chain, all Ig domains are at their resting contour length of ~40 Å. This regime, most prevalent in muscle action, has not been covered by our

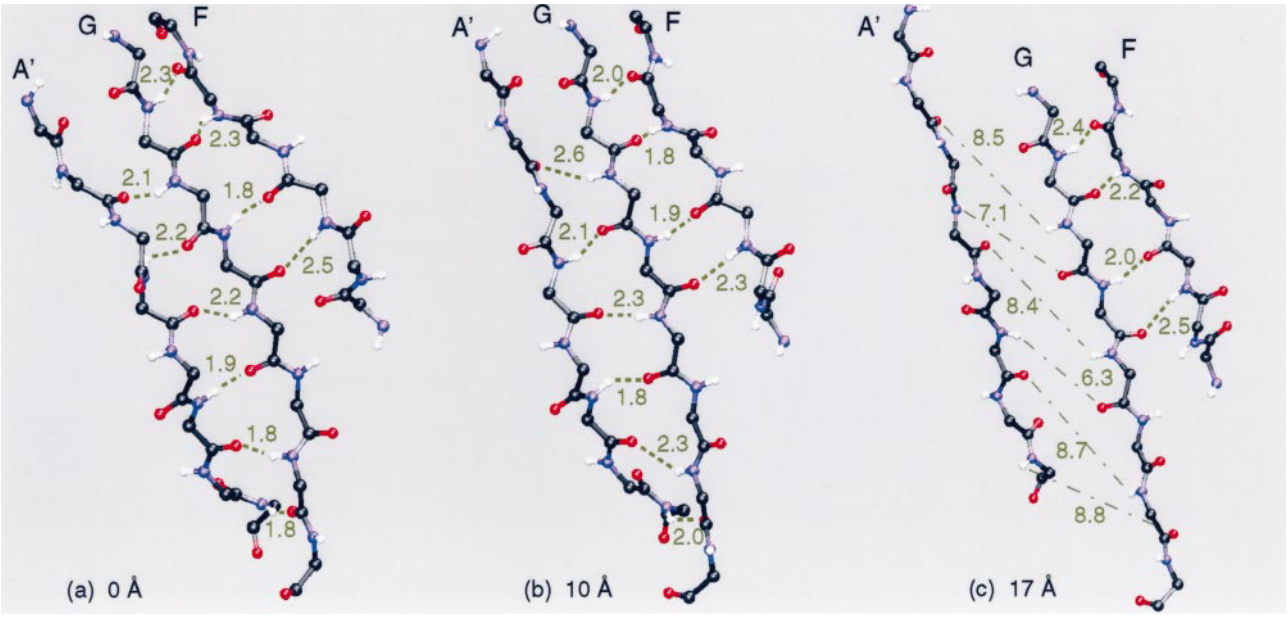

FIGURE 7 Hydrogen bonding structure of β-strands A′, G, and F of I27. (*a*) At 0-Å extension (initial structure). (*b*) At 10-Å extension (preburst structure). (*c*) At 17-Å extension (postburst structure). Hydrogen bonds are drawn as dotted lines; broken hydrogen bonds are drawn as dashed lines.

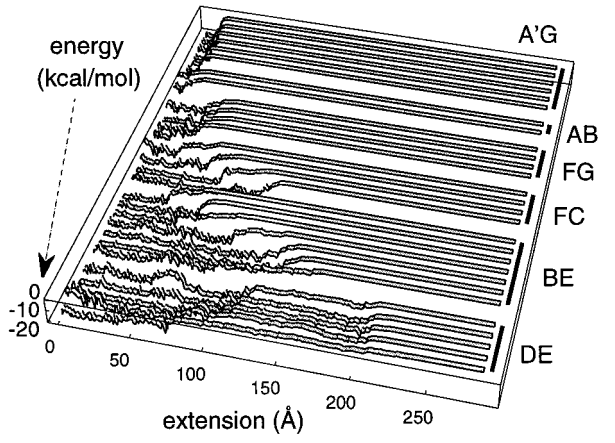

FIGURE 8 Individual interaction energy versus extension for interstrand hydrogen bond pairs in simulation C-27-0.5. Energies are grouped by β-strand pairs, ordered from far to near according to the ordering in Table 2.

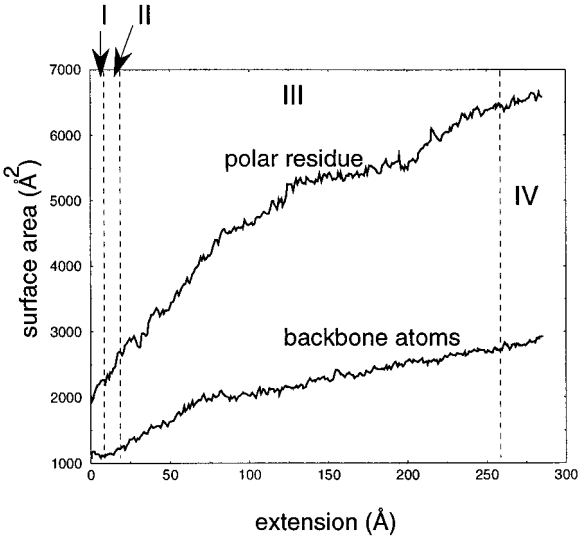

FIGURE 10 Surface accessible area versus extension in simulation C-27-0.5. Apolar residue surface area and protein backbone atom surface area are plotted against extension. The four regions of extension are defined as in Fig. 5.

description. Rather, the stretching of titin for stronger forces has been investigated. Every Ig domain exhibits a preburst increase in contour length of ~10 Å per Ig domain. These length increases could contribute to the extensions exhibited before the sawtooth force pattern observed in AFM experiments (Rief et al., 1997). Such preburst extension could partially account for the observed differences in this experiment between the observed spacing of 250–280 Å between the force peaks and the contour length increase of 280–290 Å required for the wormlike chain (WLC) model to fit the data (Rief et al., 1997).

With further extension, Ig domains continue to lengthen, but only after the force exceeds a given value, which varies slightly among different Ig domains. Our simulations provide an explanation for this bursting behavior. The applied force acts along the direction between the C- and N-termini of the Ig domains. The links that must be ruptured first to initiate the unfolding of the Ig domains are the hydrogen bonds between β-strands A and B and between β-strands A′ and G. Because of the topology of the Ig domain (Fig. 11 *c*), until the A′G bonds break, force cannot be transmitted along the backbone to unravel the rest of the protein. Only when the A′G and AB strands are separated, after all interstrand hydrogen bonds are broken, can the unfolding of an Ig domain continue, involving rupture of the interstrand hydrogen bonds between the remaining β-strands. The difference in the force peak distinguishing the Ig domains can be due to either differences in the strength of the hydrogen bonds between the A and B, and A′ and G strands, or to a different angle between these strands and the C- and N-termini that could lead to differences in the component of the external force actually available for hydrogen bond breaking.

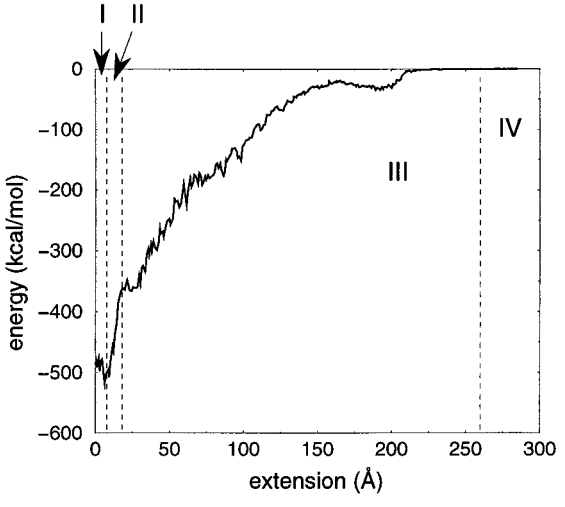

FIGURE 9 Total interaction energy of all interstrand hydrogen bond partners versus extension in simulation C-27-0.5. The four regions are defined as in Fig. 5.

Once the AB and A′G strands have been separated, the hydrophobic core of the Ig domain becomes exposed as the two sheets of the β-sandwich separate, and the rest of the domain can be unraveled with a much lower force than that required to separate the AB and A′G strands. While this occurs, the other I-band Ig domains do not experience a force high enough to separate their A and B, and A′ and G strands. Only when the unfolded domain is fully extended, beyond 260 Å, does the pulling force increase again, causing the next weakest domain to experience a burst and unfold.

Because of the short simulation times imposed by limitations of computer resources, SMD simulations must stretch Ig domains at a speed six to eight orders of magnitude higher than in AFM and optical tweezer experiments. As a result, in comparison to the experimental 30-pN to 250-pN forces recorded during Ig domain unfolding, simulation yields dominant force peaks of ~2000 pN. Extrapo-

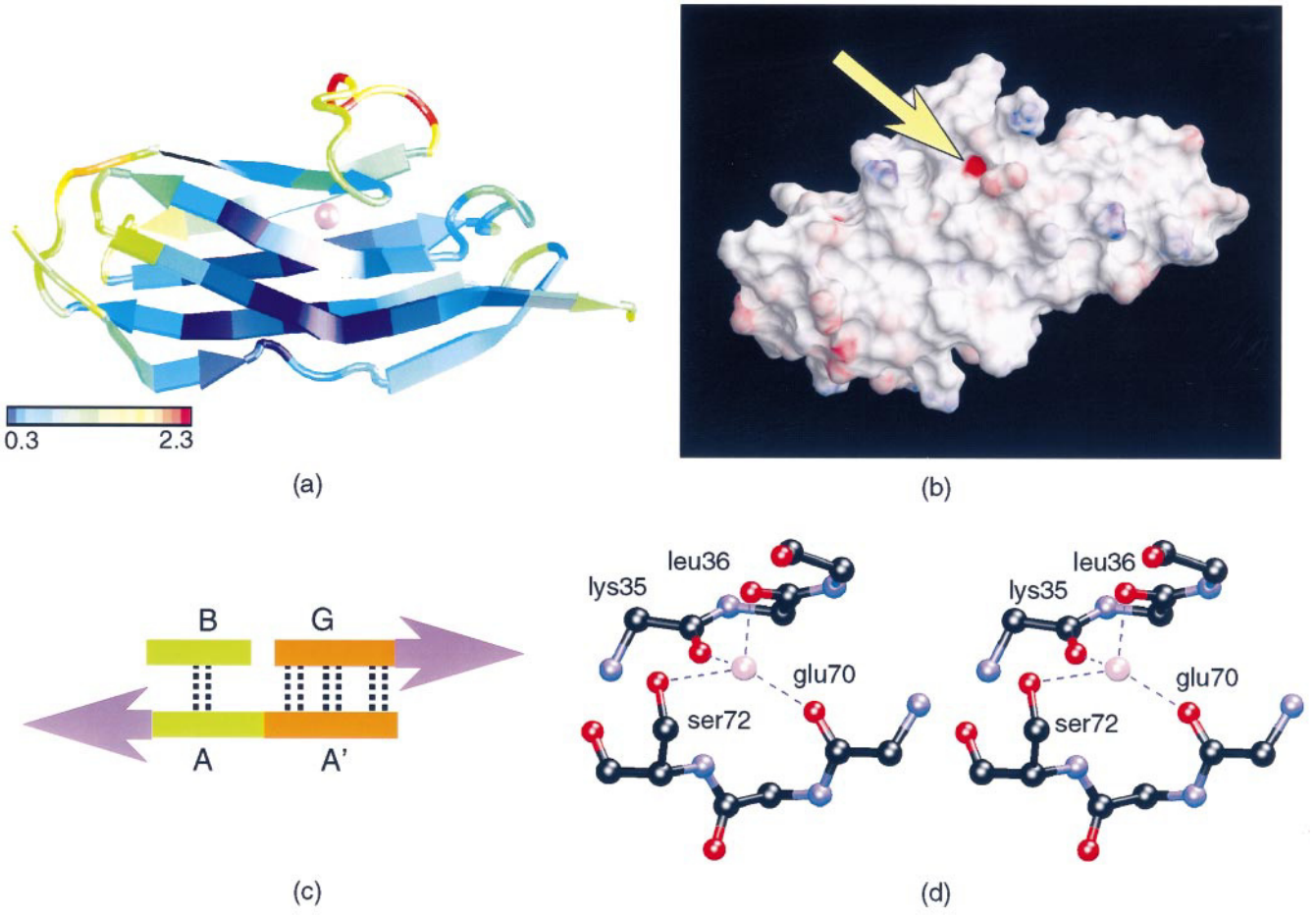

FIGURE 11 Titin topology, metal binding, and structure. (*a*) A predicted calcium-binding site (*pink sphere*) placed on a titin cartoon colored according to NMR temperature factors. The calcium-binding site was predicted by evaluating the bond valence $\nu = (R/R_1)^{-N}$ on a three-dimensional lattice, where R is the distance from the cation binding site to the ligating atom. For Ca^{2+} ions, the parameters were chosen to be $R_1 = 1.909$ Å and $N = 5.4$ (Brown and Wu, 1976), with grid spacing set to 0.1 Å. In the search for binding sites, only oxygen atoms were considered to contribute to ν. The suggested binding site has a valence of $\nu = 1.601$, which is well above the lower cutoff of $\nu \geq 1.4$ for identification of binding sites as suggested by Nayal and Cera (1994). The binding site is solvent accessible; any additional ligating water molecule would yield a trigonal bipyramidal coordination of Ca^{2+}. An octahedral coordination is less likely to occur, because Ca^{2+} is embedded in a surface dent that is only accessible to one water molecule. (*b*) Electrostatic potential surface of I27 (visualization with SURF (Varshney et al., 1994) and PSSHOW (Swanson, 1996)). The yellow arrow points to the predicted binding site. Under physiological conditions, mutation of Ser^{72} into Ala leads to an electrostatic potential difference of $|(-10.53) - (-3.50)| = 7.03\ kT/e$ at the predicted binding site. (*c*) Cartoon representation of topology of strands A, A′, B, and G of I27. Purple arrows indicate sites of force application; orange and green strand colorings indicate β-sheet membership, as in Fig. 2. (*d*) Stereo view of the Ca^{2+}-binding site (*pink sphere*).

lating from the results of Rief et al. (1997), the force required to unfold a domain for the pulling speed of the current SMD simulation should be around 500 pN, but the simulations required a peak force four times larger. The difference is due to the fact that the SMD simulations operate in a regime where forces are large enough that all energy barriers are eliminated, whereas AFM experiments operate in a regime in which barriers still exist, but are low enough to allow thermally activated barrier crossing within the millisecond experimental time scale (Izrailev et al., 1997). The strong forces and rapid motion needed in the SMD case imply that much irreversible work is performed on the system, leading to larger peak force values. In fact, we monitored temperature increases of ~18–20 K in SMD simulations with pulling speeds of 0.5 Å/ps.

Ig domains have been seen to unfold spontaneously without the application of external forces, the average time for this process measuring ~40 min with an activation energy of 3–7 kcal/mol (Politou et al., 1994). The force-induced unfolding pathway of Ig domains, with their C- and N-termini moving apart diametrically, is likely different from the pathway for spontaneous unfolding. It is highly desirable to understand the difference in the pathways of forced and spontaneous unfolding of Ig domains, in particular, because the refolding of unfolded Ig domains is likely to be a reversal of spontaneous unfolding.

Information about the Ig spontaneous unfolding pathways may be provided by temperature factors derived from NMR structure data. The temperature factors are often indicative of the thermal flexibility of protein domains. Using temper-

ature factors from I27 NMR data (Improta et al., 1996), loops are seen to be more flexible than sheets, and the B-C and C-D loops are significantly more flexible than the other loops (Fig. 11 *a*). β-strand C is the only strand in I27 with flexible loops at both ends, and has only one side hydrogen bonded to another β-strand. Fluctuations of strand C could cause the initial exposure of the hydrophobic core during spontaneous unfolding. Strands A′, G, A, and B, according to the NMR results, are stable and are not likely to be the first sites of the Ig domain to unfold. These strands, however, are the first to break in the SMD simulations.

The important role of Ca^{2+} in muscle suggests searching titin Ig domains for Ca^{2+} binding sites that may control spontaneous unfolding and refolding. For this purpose we employed the program VALE (Nayal and Cera, 1994). A binding site is predicted in the region between β-strands C and F near the proline-containing CD loop (Fig. 11 *a*). A stereo view of the predicted binding is shown in Fig. 11 *d*. Three oxygens (Lys^{35}, Leu^{36}, and Glu^{70}) belonging to the protein backbone ligate the ion, and a fourth ligating oxygen belongs to the hydroxyl group of Ser^{72}, a site with 60% conservation of Ser and Thr residues (Improta et al., 1996) among cardiac titin I-band Ig domains. Further evidence for a Ca^{2+}-binding site can be derived from the electrostatic potential shown in Fig. 11 *b*, produced by DELPHI (Honig and Nicholls, 1995), which exhibits a minimum electrostatic potential at a deep dent in the surface.

The putative Ca^{2+}-binding site could play a role in Ig domain stabilization and refolding. The location of the Ca^{2+} ion suggests that Ca^{2+} could stabilize the β-strand C, locking it into place next to the β-strand F (Fig. 11 *a*). The putative Ca^{2+} binding site is distant from the strands A, B, A′, and G, i.e., Ca^{2+} should not be a significant factor for forced unfolding.

A main shortcoming of the present investigation is the overestimation of the maximum unfolding force. To repair this deficit, simulated pulling needs to be slowed down, and thus simulation times extended. With improved computational resources, simulation times of 100 ns for properly solvated Ig domains with 11,000 atoms may soon be achieved. It is also desirable to devise a method to discount the irreversible work performed during rapid unfolding, such as the method suggested by Balsera et al. (1997).

Even with the shortcomings mentioned above, the present work demonstrates the power of the SMD method. For a rather limited computational expense, the method reveals key features of the unfolding process of an Ig domain under external forces that are not yet amenable to direct observation. Future simulations will include pulling multiple, connected Ig domains, and should clearly demonstrate sequential unfolding. Non-I-band Ig domains will also be studied to determine differences in response to force application. The opportunities for modeling studies are matched by advances of experimental methods. A recently completed tenascin Fn-3 domain stretching experiment by AFM (Oberhauser et al., 1998) showed sawtooth peaks similar to those observed in the titin stretching experiments; SMD simulations to reveal the unfolding process of the Fn-3 domain are ongoing. SMD will also be complemented in the future by kinetic modeling of strained β-sheets, generalizing the statistical mechanical model of β-hairpin folding and unfolding suggested on the basis of experimental observations by Munoz et al. (1997).

The authors thank A. Balaeff, M. Balsera, S. Izrailev, X. Hu, and T. Lybrand for fruitful discussions.

This work was supported by the National Institutes of Health (NIH PHS 5 P41 RR05969), by the National Science Foundation (NSF BIR 94-23827 EQ, NSF/GCAG BIR 93-18159, MCA93S028), and by the Roy J. Carver Charitable Trust. BI was partially supported by a GAANN Fellowship from the U.S. Department of Education. AK and VV further acknowledge support from NIH (GM 49063, First Award to VV).

REFERENCES

Balsera, M., S. Stepaniants, S. Izrailev, Y. Oono, and K. Schulten. 1997. Reconstructing potential energy functions from simulated force-induced unbinding processes. *Biophys. J.* 73:1281–1287.

Bernstein, F. C., T. F. Koetzle, G. J. Williams, E. F. Meyer, M. D. Brice, J. R. Rogers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.

Brown, I., and K. Wu. 1976. Empirical parameters for calculating cation-oxygen bond valences. *Acta Crystallogr. B.* 32:1957–1959.

Brünger, A. T. 1992. X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR. The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT.

Chan, A., E. Hutchinson, D. Harris, and J. Thornton. 1993. Identification, classification, and analysis of beta-bulges in proteins. *Protein Sci.* 2:1574–1590.

Erickson, H. 1994. Reversible unfolding of fibronectin type III and immunoglobulin domains provides the structural basis for stretch and elasticity of titin and fibronectin. *Proc. Natl. Acad. Sci. USA.* 91:10114–10118.

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

Granzier, H., M. Helmes, and K. Trombitas. 1996. Nonuniform elasticity of titin in cardiac myocytes: a study using immunoelectron microscopy and cellular mechanics. *Biophys. J.* 70:430–442.

Granzier, H., M. Kellermayer, M. Helmes, and K. Trombitas. 1997. Titin elasticity and mechanism of passive force development in rat cardiac myocytes probed by thin-film extraction. *Biophys. J.* 73:2043–2053.

Greaser, M., M. Sebestyen, J. Fritz, and J. Wolff. 1996. cDNA sequence of rabbit cardiac titin/connectin. *Adv. Biophys.* 33:13–25.

Grubmüller, H., B. Heymann, and P. Tavan. 1996. Ligand binding and molecular mechanics calculation of the streptavidin-biotin rupture force. *Science.* 271:997–999.

Harpaz, Y., and C. Chothia. 1994. Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. *J. Mol. Biol.* 238:528–539.

Honig, B., and A. Nicholls. 1995. Classical electrostatics in biology and chemistry. *Science.* 268:1144–1149.

Humphrey, W. F., A. Dalke, and K. Schulten. 1996. VMD—visual molecular dynamics. *J. Mol. Graph.* 14:33–38.

Hunenberger, P. H., A. E. Mark, and W. van Gunsteren. 1995. Computational approaches to study protein unfolding: hen egg white lysozyme as a case study. *Proteins Struct. Funct. Genet.* 21:196–213.

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

Israelewitz, B., S. Izrailev, and K. Schulten. 1997. Binding pathway of retinal to bacterioopsin: a prediction by molecular dynamics simulations. *Biophys. J.* 73:2972–2979.

Izrailev, S., S. Stepaniants, M. Balsera, Y. Oono, and K. Schulten. 1997. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.* 72:1568–1581.

Izrailev, S., S. Stepaniants, B. Israelewitz, D. Kosztin, H. Lu, F. Molnar, W. Wriggers, and K. Schulten. 1998. Steered molecular dynamics. *In* Algorithms for Macromolecular Modelling, Lecture Notes in Computational Science and Engineering. P. Deuflhard, J. Hermans, B. Leimkuhler, A. Mark, R. D. Skeel, and S. Reich, editors. Springer-Verlag, New York (in press).

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

Karplus, M., and A. Sali. 1995. Theoretical studies of protein folding and unfolding. *Curr. Opin. Struct. Biol.* 5:58–73.

Kellermayer, M., and H. Granzier. 1996. Elastic properties of single titin molecules made visible through fluorescent F-actin binding. *Biochem. Biophys. Res. Commun.* 221:491–497.

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

Labeit, S., and B. Kolmerer. 1995. Titins, giant proteins in charge of muscle ultrastructure and elasticity. *Science.* 270:293–296.

Labeit, S., B. Kolmerer, and W. Linke. 1997. The giant protein titin: emerging roles in physiology and pathophysiology. *Circ. Res.* 80: 290–294.

Leech, J., J. Prins, and J. Hermans. 1996. SMD: visual steering of molecular dynamics for protein design. *IEEE Comp. Sci. Eng.* 3:38–45.

Li, A., and V. Daggett. 1996. Identification and characterization of the unfolding transition state of chymotrypsin inhibitor 2 by molecular dynamics simulations. *J. Mol. Biol.* 257:412–429.

Linke, W., M. Ivemeyer, N. Olivieri, B. Kolmerer, J. Ruegg, and S. Lebeit. 1996. Towards a molecular understanding of the elasticity of titin. *J. Mol. Biol.* 261:62–71.

Maruyama, K. 1997. Connectin/titin, a giant elastic protein of muscle. *FASEB J.* 11:341–345.

Munoz, V., P. Thompson, J. Hofrichter, and W. Eaton. 1997. Folding dynamics and mechanism of beta-hairpin formation. *Nature.* 390: 196–199.

Nayal, M., and E. D. Cera. 1994. Predicting Ca^{2+}-binding sites in proteins. *Proc. Natl. Acad. Sci. USA.* 91:817–821.

Nelson, M., W. Humphrey, A. Gursoy, A. Dalke, L. Kalé, R. D. Skeel, and K. Schulten. 1996. NAMD—a parallel, object-oriented molecular dynamics program. *J. Supercomputing Appl.* 10:251–268.

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

Pearlman, D. A., D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham III, D. M. Ferguson, G. L. Seibel, U. C. Singh, P. K. Weiner, and P. A. Kollman. 1995. AMBER4.1. University of California, San Francisco.

Pfuhl, M., M. Gautel, A. Politou, C. Joseph, and A. Pastore. 1995. Secondary structure determination by NMR spectroscopy of an immunoglobulin-like domain from the giant muscle protein titin. *J. Biomol. NMR.* 5:48–58.

Pfuhl, M., and A. Pastore. 1995. Tertiary structure of an immunoglobulin-like domain from the giant muscle protein titin: a new member of the I set. *Structure.* 3:391–401.

Politou, A., M. Gautel, M. Pfuhl, S. Labeit, and A. Pastore. 1994. Immunoglobulin-type domains of titin: same fold, different stability? *Biochemistry.* 33:4730–4737.

Politou, A. S., D. Thomas, and A. Pastore. 1995. The folding and the stability of titin immunoglobulin-like modules, with implications for mechanism of elasticity. *Biophys. J.* 69:2601–2610.

Rief, M., M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub. 1997. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science.* 276:1109–1112.

Stepaniants, S., S. Izrailev, and K. Schulten. 1997. Extraction of lipids from phospholipid membranes by steered molecular dynamics. *J. Mol. Model.* 3:473–475.

Swanson, E. 1996. PSSHOW User's Guide, Version 2.2. University of Washington, Seattle, WA.

Tirado-Rives, J., M. Orozco, and W. Jorgensen. 1997. Molecular dynamics simulations of the unfolding of barnase in water and 8M aqueous urea. *Biochemistry.* 36:7313–7329.

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

Varshney, A., F. P. Brooks, and W. V. Wright. 1994. Linearly scalable computation of smooth molecular surfaces. *IEEE Comp. Graph. Appl.* 14:19–25.

Wang, K., R. McCarter, J. Wright, J. Beverly, and R. Ramirez-Mitchell. 1993. Viscoelasticity of the sarcomere matrix of skeletal muscles. *Biophys. J.* 64:1161–1177.