Simulated Refolding of Stretched Titin Immunoglobulin Domains

Mu Gao,*† Hui Lu,† and Klaus Schulten*†
*Department of Physics and †Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801 USA

ABSTRACT Steered molecular dynamics (SMD) is used to investigate forced unfolding and spontaneous refolding of immunoglobulin I27, a domain of the muscle protein titin. Previous SMD simulations revealed the events leading to stretch-induced unfolding of I27, the rupture of hydrogen bonds bridging β-strands A and B, and those bridging β-strands A′ and G, the latter rupture occurring at an extension of ~15 Å and preceding the complete unfolding. Simulations are now used to study the refolding of partially unfolded I27 domains. The results reveal that stretched domains with ruptured interstrand hydrogen bonds shrink along the extension direction. Two types of refolding patterns are recognized: for separated β-strands A′ and G, in most simulations five of the six hydrogen bonds between A′ and G stably reformed in 2 ns, whereas for separated β-strands A and B hydrogen bonds seldom reformed in eight 2-ns simulations. The mechanical stability of the partially refolded intermediates has been tested by re-stretching.

INTRODUCTION

Titin, a giant protein (~3000 kDa, ~1 μm long) of muscle sarcomere, is primarily composed of ~300 modules in two motif types, immunoglobulin (Ig) and fibronectin type III (FnIII) domains (Maruyama, 1997; Wang, 1996). It is the longest covalently linked protein known and is coded by the longest gene identified in the human genome (Consortium, 2001). Among the various functional roles that titin plays, providing passive elasticity under tension is important for muscle extensibility and has attracted extensive studies (Erickson, 1994; Labeit and Kolmerer, 1995; Granzier et al., 1996; Labeit et al., 1997; Linke et al., 1999). It has been well established that the sarcomere I-band of titin, consisting of mainly poly-Ig repeats and a PEVK domain (rich in proline, glutamic acid, valine, and lysine residues), contributes to titin extensibility under tension (Labeit and Kolmerer, 1995; Granzier et al., 1996; Erickson, 1997; Linke et al., 1998). Due to a large number of charged residues in its structure, the PEVK domain of I-band titin cannot form stable tertiary structure folds and, hence, easily elongates to develop passive tension under small forces (Labeit and Kolmerer, 1995). Tandem Ig domains, the other main elastic elements in I-band titin, were also speculated to unfold upon stretch in physiological conditions, and in particular under overstretched conditions (Erickson, 1997; Granzier et al., 1996; Linke et al., 1999; Minajeva et al., 2001). Recently a third elastic element, a sequence insertion, I-band isoform N2-B, has been found to account for the passive elasticity of cardiac muscle (Linke et al., 1999). The element, however, seems to be unique to the N2-B isoform.

To study the mechanical properties of Ig-like domains under extreme conditions, atomic force microscopy (AFM) and optical tweezers have been utilized to reversibly unfold titin Ig and FnIII domains (Rief et al., 1997; Carrion-Vazquez et al., 1999; Kellermayer et al., 1997; Tskhovrebova et al., 1997). To exclude heterogeneous effects, polyproteins composed of identical I27 or I28 modules were genetically engineered and unraveled in AFM experiments (Oberhauser et al., 1999; Marszalek et al., 1999; Li et al., 2000; Fisher et al., 2000). Analysis of the sawtooth force-extension profiles revealed that the unfolding of domains occurs in two steps: forces above 50 pN extend the domains by ~6 Å and forces above 150–200 pN extend the domains further, leading then to a complete unfolding of the domains.

The experimental observations required an interpretation in terms of the architecture of the I27 domain, a double β-sheet structure shown in Fig. 1 *a*. For this purpose, steered molecular dynamics (SMD) simulations have been conducted to provide a detailed atomic level picture of forced unfolding of individual, solvated Ig or FnIII domains (Lu et al., 1998; Lu and Schulten, 1999a, 2000; Marszalek et al., 1999; Krammer et al., 1999; Craig et al., 2001, Paci and Karplus, 1999, 2000). The simulations of I27 suggested that the two steps of unfolding indicated in AFM experiments correspond to two sequential events of interstrand hydrogen bond rupture. The two sets of hydrogen bonds are the three hydrogen bonds connecting β-strands A and B and the six bridging between β-strands A′ and G. The first set of hydrogen bonds breaks at weak forces of 50–150 pN and results in a 4–7 Å extension of the domain (Fig. 1 *b*), which explains the minor event of ~6 Å deviation from the position of the major force peak in the extension-force profile recorded in AFM experiments (Marszalek et al., 1999). The second set breaks at stronger forces of 100–300 pN and thereby initiates the complete unfolding (Fig. 1 *c*). The sampling on SMD simulations also provides a match to AFM and chemical denaturation experiments on the amount of energy required to cross the main potential barrier (Car-

Received for publication 28 March 2001 and in final form 5 July 2001.

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

H. Lu's present address is Laboratory of Computational Genomics, Danforth Plant Science Center, St. Louis, MO 63141.

© 2001 by the Biophysical Society

0006-3495/01/10/2268/10 $2.00

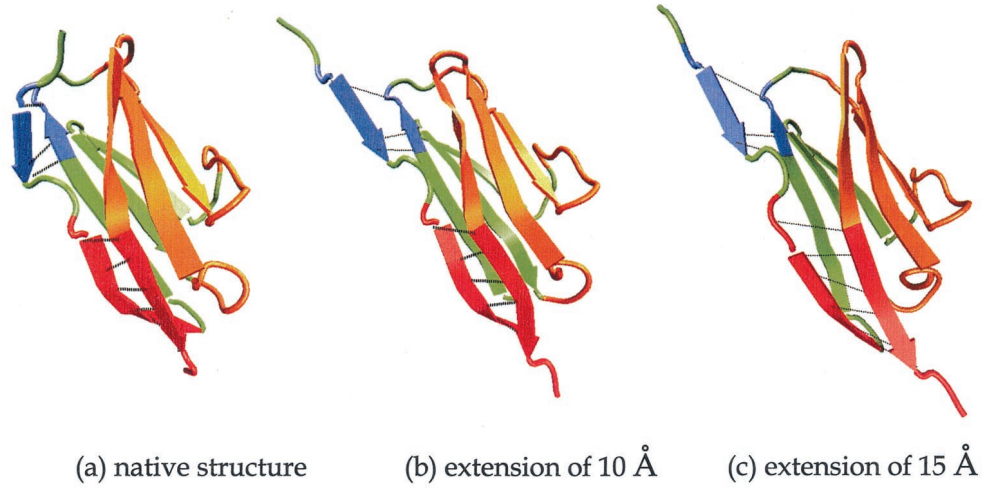

FIGURE 1 Representative snapshots of the 750 pN constant force unfolding scenario of titin immunoglobulin domain I27: (*a*) native structure; (*b*) intermediate at extension of 10 Å; (*c*) intermediate at extension of 15 Å. In snapshot (*b*) the hydrogen bonds between strands A and B have been ruptured. In snapshot (*c*) the interstrand hydrogen bonds between A and B and A′ and G have been ruptured. Green (sheet ABED) and orange (sheet A′GFC) denote the two β-sheets of I27, except that the residues involved in hydrogen bonds between β-strands A and B are shown in blue, and residues involved in hydrogen bonds between β-strands A′ and G are shown in red. Hydrogen bonds are represented as thick black dotted lines, and broken hydrogen bonds as thin lines.

rion-Vazquez et al., 1999; Lu and Schulten, 1999b). Moreover, the SMD simulations revealed that water molecules fluctuate around the backbone atoms during barrier crossing and attack interstrand hydrogen bonds, thereby assisting the unfolding of I27 domains (Lu and Schulten, 2000). Recently, a lattice model (Klimov and Thirumalai, 1999) and an off-lattice model (Klimov and Thirumalai, 2000) had been used to simplify the forced unfolding simulations of I27. Combined with dynamic force spectroscopy theory (Evans and Ritchie, 1997, 1999) the results from these simulations, so-called steered Langevin dynamics simulations, gave good predictions of rupture forces measured in AFM experiments on I27.

If Ig-like domains need to unfold to fulfil their physiological role, they must be able to do it reversibly. The reversible unfolding has been demonstrated indeed in both AFM and optical tweezer experiments (Rief et al., 1997; Carrion-Vazquez et al., 1999; Kellermayer et al., 1997; Tskhovrebova et al., 1997). In AFM experiments, it takes no less than 1 s for an I27 module to completely refold (Carrion-Vazquez et al., 1999). A faster refolding rate of $\sim$20 s^{-1} of human fibronectin FnIII domains was reported in chemical experiments (Plaxco et al., 1996, 1997), and a refolding rate of $\sim$42 s^{-1} of FnIII domains of tenascin was estimated in AFM experiments (Oberhauser et al., 1998). These refolding rates are, however, much faster than spontaneous unfolding rates, which typically range from 10^{-3} to 10^{-5} s^{-1} (Fong et al., 1996; Clarke et al., 1997; Rief et al., 1997; Carrion-Vazquez et al., 1999). Due to the limited resolution of current experimental instruments, we know little about the characteristic of the refolding pathway of Ig-like domains at the atomic level. With MD tools that complement the experimental techniques one can characterize the refolding pathway, especially the reformation of interstrand hydrogen bonds.

Starting from partially unfolded structures of the $FnIII_9$ domain of fibronectin at different stages of unfolding, Paci and Karplus (1999) have simulated the refolding by using an implicit solvent model. Their simulations, however, showed an unnaturally fast refolding process that lasted $<$50 ps, which is even shorter than the time they spent on simulated stretching; refoldings also ended in non-native secondary structures.

In this paper we will simulate the refolding process starting from different stages of partially unfolded I27 intermediates by using an explicit solvent model. The energy of two sets of hydrogen bonds, which involve the minor and major unfolding events characterized above, will be analyzed to characterize the refolding process. The simulations correspond to the final stages of the I27 folding process and the formation of the key force-bearing element of the domain that secures the integrity of the protein under external forces. We will demonstrate that water-protein interactions play an important role during refolding. To test the mechanical elasticity of the refolded proteins, external forces are applied to the resulting structures from two of the refolding simulations.

METHODS

We started our simulations on model immunoglobulin I27 domains that had been partially unfolded in previous studies through the application of

TABLE 1 Sequences of simulations utilized in this study

Step 1	Unfolding	Force (pN)	Initial Structure	Time (ps)	Resulting Structure
	$SMD(nat)_1$	750	nativeI27	200	s_1
→	$SMD(nat)_2$	750	nativeI27	299	s_2
stretch	$SMD(nat)_1$	750	nativeI27	311	s_3
	$SMD(nat)_2$	750	nativeI27	1000	s_4
	$SMD(nat)_2$	750	nativeI27	1010	s_5
	$SMD(nat)_1$	750	nativeI27	1067	s_6
	$SMD(nat)_1$	750	nativeI27	1072	s_7
	$SMD(nat)_2$	750	nativeI27	1085	s_8
Step 2	**Refolding**	**Force**	**Initial Structure**	**Time (ns)**	**Resulting Structure**
	$R(s_1)$	—	s_1	2	s'_1
	$R(s_2)$	—	s_2	2	s'_2
relax	$R(s_3)$	—	s_3	2	s'_3
	$R(s_4)$	—	s_4	2	s'_4
	$R(s_5)$	—	s_5	2	s'_5
	$R(s_6)$	—	s_6	2	s'_6
	$R(s_7)$	—	s_7	2	s'_7
	$R(s_8)$	—	s_8	2	s'_8
Step 3	**Re-stretching**	**Force (pN)**	**Initial Structure**	**Time (ns)**	
→	$SMD(s'_3)$	100	s'_3	1	
re-stretch	$SMD(s'_4)$	750	s'_4	1	

Two independent SMD unfolding simulations reported in Lu and Schulten (2000), named $SMD(nat)_1$ and $SMD(nat)_2$, were started from the native I27 structure by applying forces of 750 pN. At various time steps the unfolding simulations were stopped and the intermediate protein structures, named s_k ($k = 1 - 8$), were selected. These structures were allowed to relax for 2 ns without application of external forces, resulting in structures named s'_k ($k = 1 - 8$). The corresponding simulations described in this paper are denoted as $R(s_k)$ ($k = 1 - 8$). Finally, two SMD simulations, named $SMD(s'_3)$ and $SMD(s'_4)$, were carried out on partially refolded structures s'_3 and s'_4 by applying forces of 100 pN and 750 pN, respectively, to test the mechanical elasticity of s'_3 and s'_4.

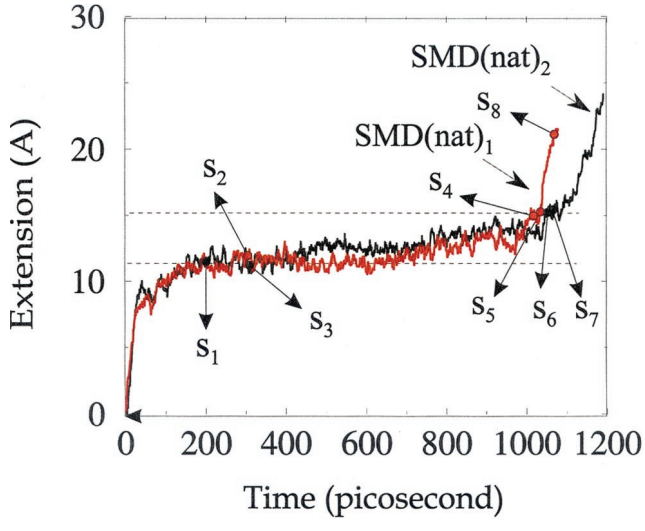

FIGURE 2 Extension-time profiles from simulations $SMD(nat)_1$ and $SMD(nat)_2$, adapted from Lu and Schulten (2000). Eight frames s_k ($k = 1 - 8$) were selected from two trajectories as the initial structures for the refolding simulations $R(s_k)$ ($k = 1 - 8$), as indicated by circles and arrows. Three of the structures, s_1 (200 ps), s_2 (299 ps), and s_3 (311 ps), have hydrogen bonds between β-strands A and B broken (Fig. 1 *b*), while the other five structures, s_4 (1000 ps), s_5 (1010 ps), s_6 (1067 ps), s_7 (1072 ps), and s_8 (1085 ps) have all the hydrogen bonds between β-strands A′ and G and between β-strands A and B broken (Fig. 1 *c*).

750 picoNewton (pN) constant forces (Lu and Schulten, 2000). The pre-stretched proteins were allowed to relax spontaneously for 2 ns and were then stretched by applying constant-force SMD protocols to them. The simulations carried out are listed in Table 1.

Eight I27 unfolding intermediates s_k ($k = 1, 2, \ldots, 8$) were selected from two independent SMD unfolding simulation trajectories, $SMD(nat)_1$ and $SMD(nat)_2$ (Lu and Schulten, 2000), as shown in Fig. 2. These intermediates can be classified as two sets: AB intermediates s_k ($k = 1, 2, 3$), named after the disruption of the interstrand hydrogen bonds only between β-strands A and B during forced unfolding, and A′G intermediate s_k ($k = 4–8$), where all six hydrogen bonds between β-strand A′ and G were ruptured. The AB intermediates s_1 and s_2 were selected at 200 ps and 299 ps from trajectories of $SMD(nat)_1$ and $SMD(nat)_2$, respectively. Intermediate s_3 was selected after the hydrogen bonds between β-strands A and B had been broken for ~110 ps in $SMD(nat)_1$. The intermediates s_4 to s_8 were selected at time steps no less than 1000 ps from $SMD(nat)_1$ and $SMD(nat)_2$ after all six hydrogen bonds between β-strand A′ and G were broken. Each of the simulations describing the refolding of the intermediates s_k, referred to below as $R(s_k)$, was carried out for 2 ns. In these simulations no external forces were applied, permitting the domains to refold. To start these simulations we have used restart files to initiate the simulations for $R(s_3)$, $R(s_4)$, and $R(s_8)$ with actual atomic velocities, whereas in the other five refolding simulations the velocities of atoms were randomly assigned according to Maxwellian distributions for a temperature of 300 K.

Simulations $R(s_k)$ lead to resulting structures denoted as s'_k. For s'_3 and s'_4 a second simulation was carried out, referred to as $SMD(s'_3)$ and $SMD(s'_4)$, in which constant forces of 100 pN and 750 pN were applied to the C_α atom of the C-terminus residue of the I27 modules, the C_α atom of the N-terminus being fixed, and the force being directed along the vector connecting the initial positions of the N-terminus to the C-terminus.

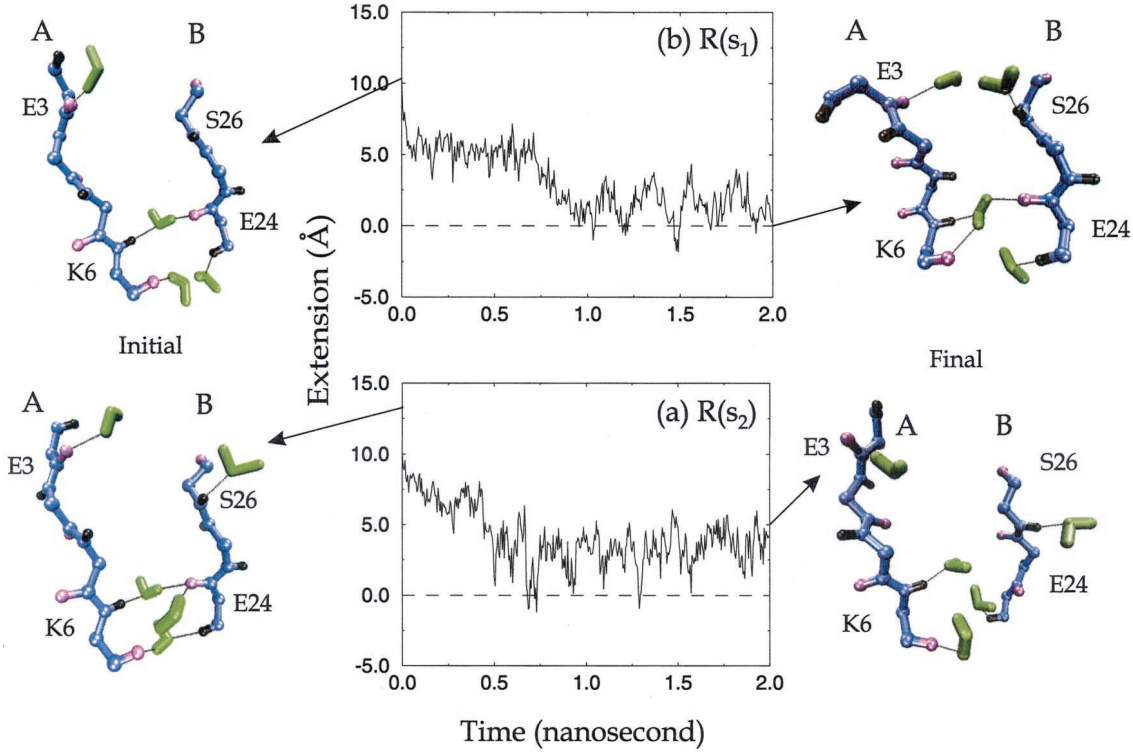

FIGURE 3 Extension-time profiles from (*a*) simulation $R(s_2)$ and (*b*) simulation $R(s_1)$, together with snapshots of β-strands A and B at the initial and final time steps. Both curves show a fast decrease in extension, but slow reformation of hydrogen bonds between β-strands A and B. In all snapshots only the backbone atoms and the hydrogen atoms involved in the backbone hydrogen bonds are shown. Oxygen atoms, hydrogen atoms, water molecules, and other backbone atoms are colored in purple, black, green, and blue, respectively.

Before the original SMD simulations were initiated, the native protein was solvated by a sphere of TIP3 water molecules (Jorgensen et al., 1983) of 31 Å radius and equilibrated for 1 ns. The molecular surface of the domain was covered at this point by at least four shells of water molecules.

All molecular dynamics simulations were carried out using the software packages NAMD (Kalé et al., 1999) and X-PLOR (Brünger, 1992) with the CHARMm22 force field (MacKerell, Jr. et al., 1998). An integration time step of 1 fs and a uniform dielectric constant of 1 was chosen. A switching function for calculating non-bonded Coulomb and van der Waals interactions was used, switching interactions off between 10 Å and 13 Å. The atomic coordinates of the entire system were recorded every picosecond.

The analysis of molecular structures and hydrogen bond energies were conducted using X-PLOR and VMD (Humphrey et al., 1996). The extension of the protein is defined as the change of the end-to-end distance between the two terminal carbon atoms ($L1(C_\alpha)$ and $L89(C_\alpha)$) at the end of the 1-ns equilibration, i.e., before the first round of SMD simulations. An explicit hydrogen-bonding energy term was used in the trajectory analysis, with parameters adopted from param11.pro in XPLOR. The minimum energy of the nitrogen-hydrogen-oxygen (N-H-O) hydrogen bond is -3.5 kcal/mol, occurring when the O-H distance is 1.9 Å and the N-H-O angle is 180°. A hydrogen bond was considered broken when the O-H distance was longer than 3 Å, or when the N-H-O angle was $<90°$, situations that correspond to an H-bond energy above -1 kcal/mol.

RESULTS

Eight refolding simulations $R(s_1)$–$R(s_8)$ have been performed, and two constant force SMD simulations $SMD(s'_3)$ and $SMD(s'_4)$ with stretching forces of 100 pN and 750 pN have been carried out, as described in Methods.

Refolding of intermediates s_1, s_2, and s_3

Fig. 3 presents the extension-time profiles resulting from refolding simulations $R(s_1)$ and $R(s_2)$ together with snapshots of backbone hydrogen bonds between β-strands A and B at the initial and final time steps of the simulations. In simulation $R(s_1)$ (shown in Fig. 3 *b* the extension decreases within 1 ns from 11 Å to <1 Å, which is very close to the native fully folded case, then fluctuates around 2 Å in the remaining time of the simulation. Although the extension of the protein shrinks dramatically, water molecules remain intercalated between β-strands A and B, preventing backbone hydrogen bonds from reforming. All three hydrogen bonds are still broken and backbone atoms form hydrogen bonds with nearby water molecules at the end of the simulations. The simulation results of $R(s_2)$ shown in Fig. 3 *a* are similar to those shown in Fig. 3 *b*. The extension of the I27 domain is seen to diminish from 13 Å to 1 Å in 500 ps, then to fluctuate around 4 Å for the remaining simulation period. Nevertheless, simulations $R(s_2)$ did not lead to a stable

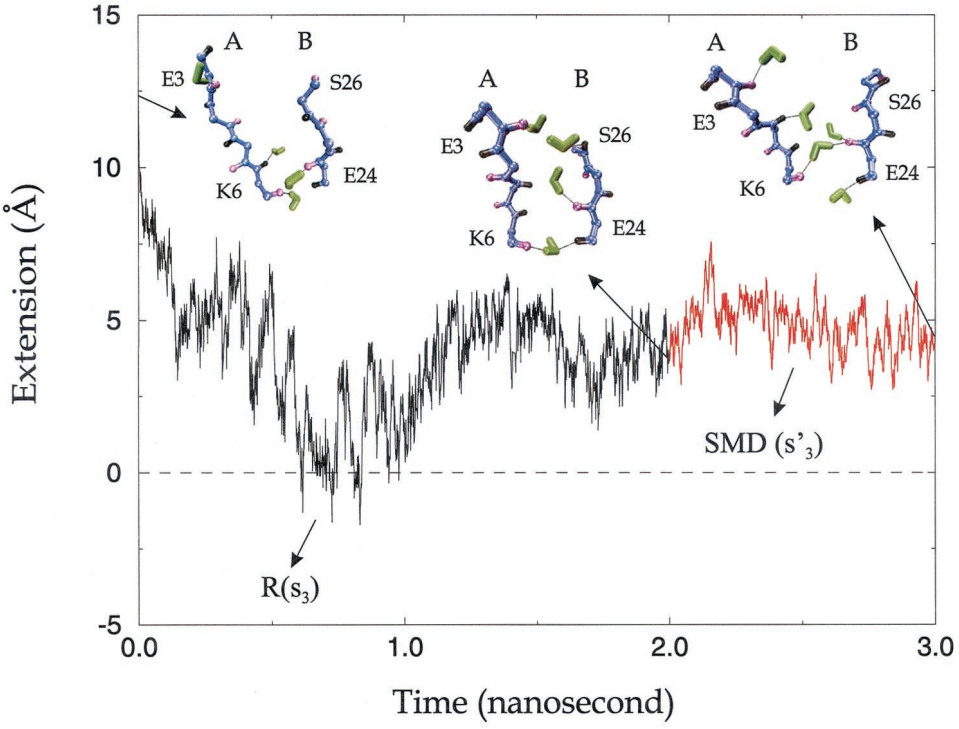

FIGURE 4 Time development of extension of a prestretched I27 domain in simulations R(s_3) (*black*) and SMD(s'_3) (*red*), together with snapshots of backbone hydrogen bonds between β-strand A and B at various stages of the simulations. The left snapshot shows the β-strands A and B of the initial structure of refolding s_3. After 2-ns relaxation (*center snapshot*), the extension of the protein has decreased from ~12.5 Å to ~4 Å. The protein was then stretched by applying a constant force of 100 pN, and reached an extension of ~5 Å (*right snapshot*). Oxygen atoms, hydrogen atoms, water molecules, and other backbone atoms are colored in purple, black, green, and blue, respectively.

reformation of backbone hydrogen bonds between β-strands A and B in 2 ns.

To further investigate the unfolding properties of the partially refolded domains resulting from R(s_3), the s'_3 structure was subjected to a 100 pN constant force in SMD simulation SMD(s'_3). The resulting extension versus time curves are shown in Fig. 4, covering both simulations R(s_3) and SMD(s'_3). The refolding resulting in R(s_3) resembles the curve shown in Fig. 3 for R(s_1) and R(s_2). At the end of simulation R(s_3), the extension of the protein decreased by ~8 Å and fluctuated around a value of 4 Å. Although the backbone hydrogen and oxygen atoms came closer in a tendency toward reformation of the interstrand hydrogen bonds (see center snapshot), water molecules still intercalated between strands A and B prevented hydrogen bonds from reforming. Because those hydrogen bonds did not reform, the protein in state s'_3 was expected to be more easily extendible than in the case of the native structure. In fact, during simulation SMD(s'_3) in which a force of 100 pN was applied, the extension reached a peak of 7 Å in 300 ps and later fluctuated around 5 Å. The β-strands A and B separated further (as seen in the right snapshot) to provide a transitional extension before the main unfolding event.

Fig. 5 shows the hydrogen bond energy versus time for all hydrogen bonds between β-strands A and B. As one can see, the energy of the hydrogen bond between E3(O) and S26(H) is always higher than −1 kcal/mol, i.e., this bond remains in a ruptured state. Hydrogen bonds K6(H)-E24(O) and K6(O)-E24(H) show energy fluctuations around −1 kcal/mol, i.e., they are stronger than the E3(O)-S26(H) bond, yet these bonds cannot be considered to have stably reformed.

Refolding of intermediate s_4 to s_8

Fig. 6 presents the extension versus time profile from the contiguous simulations R(s_4) and SMD(s'_4), together with snapshots of backbone hydrogen bonds between β-strands A and B and β-strands A′-G. During simulation R(s_4) the domain gradually became more compact, with its extension decreasing from over 15 Å to 0 Å, the native folded length. At the end of the 2-ns refolding simulation, one hydrogen bond between A-B, K6(O)-E24(H) (see bottom right snapshot), stably reformed; five of the six hydrogen bonds connecting A′ and G stably reformed, except the bond between Y9(O) and N83(H) (see upper right snapshot). There is no water molecule observed within 3 Å of the hydrogen and oxygen atoms that pairs these stably reformed hydrogen bonds. In simulation SMD(s'_4), the protein, under

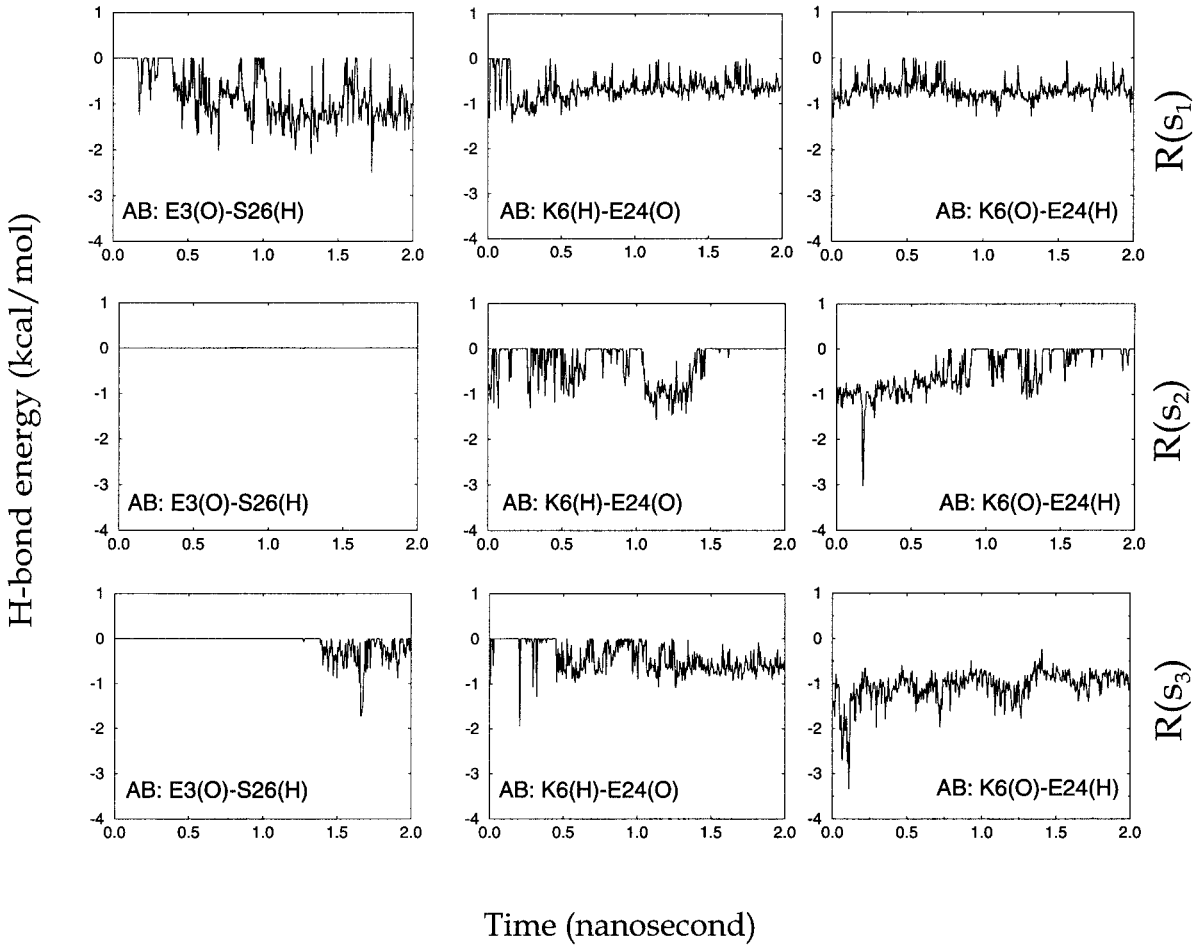

FIGURE 5 Backbone hydrogen bond energies versus time for individual hydrogen bonds between β-strands A and B. *Top row*: simulation R(s_1); *middle row*: simulation R(s_2); *bottom row*: simulation R(s_3).

a force of 750 pN, extends by 10 Å in <10 ps, then fluctuates in the middle plateau region for ~500 ps. After this, the protein crosses the main unfolding energy barrier (at ~2.5 ns). The right snapshot in Fig. 6 was made when the hydrogen bonds between β-strands A′ and G were ruptured again, and water molecules attacked backbone oxygens and hydrogens to form protein-water hydrogen bonds.

The energy of individual interstrand hydrogen bonds was monitored during simulation R(s_4), as shown in Fig. 7. Hydrogen bonds between β-strands A and B displayed a tendency to be stably reformed, except for the bond between E3(O) and S26(H). The bond K6(H)-E24(O) reforms (the energy is <-1 kcal/mol) at ~200 ps, but breaks after 1 ps. It reaches a more favorable low energy configuration at ~1600 ps and stays there for ~200 ps. The bond then breaks again and is found to fluctuate between formed and broken states for the remaining time of the simulation. The other bond between K6 and E24, K6(O)-E24(H), is adjusted to a configuration with a bond energy lower than -3 kcal/mol in <200 ps; it then fluctuates between formed and broken states until it is stabilized. No water molecule forms a hydrogen bond with either K6(O) or E24(H) at the end of simulation R(s_4).

Hydrogen bonds between β-strand A′ and G show a stronger tendency to reform, except for the one between hydrogen bond partners Y9(O) and N83(H), located at the end of the A′-G strands. The four hydrogen bonds, V11(H)-N83(O), V11(O)-K85(H), V13(H)-K85(O), and V13(O)-K87(H) reform in <200 ps and are stable after the reformation. It takes hydrogen bond V15(H)-K87(O) 400 ps to reform. After the 2-ns simulation R(s_4), five of the six hydrogen bonds between A′ and G are reformed.

Although none of three hydrogen bonds between β-strands A and B stably forms in simulations R(s_5), R(s_6), and R(s_7), results of hydrogen bond formation between β-strands A′ and G from these simulations are similar to those from R(s_4). The three hydrogen bonds, V11(H)-N83(O), V11(O)-N83(H), and V13(H)-K85(O), located in the middle of the hydrogen bond cluster between A′-G strands, reform faster than the three H-bonds, V13(O)-K87(H), V15(H)-K87(O), and Y9(O)-N83(H) located at the

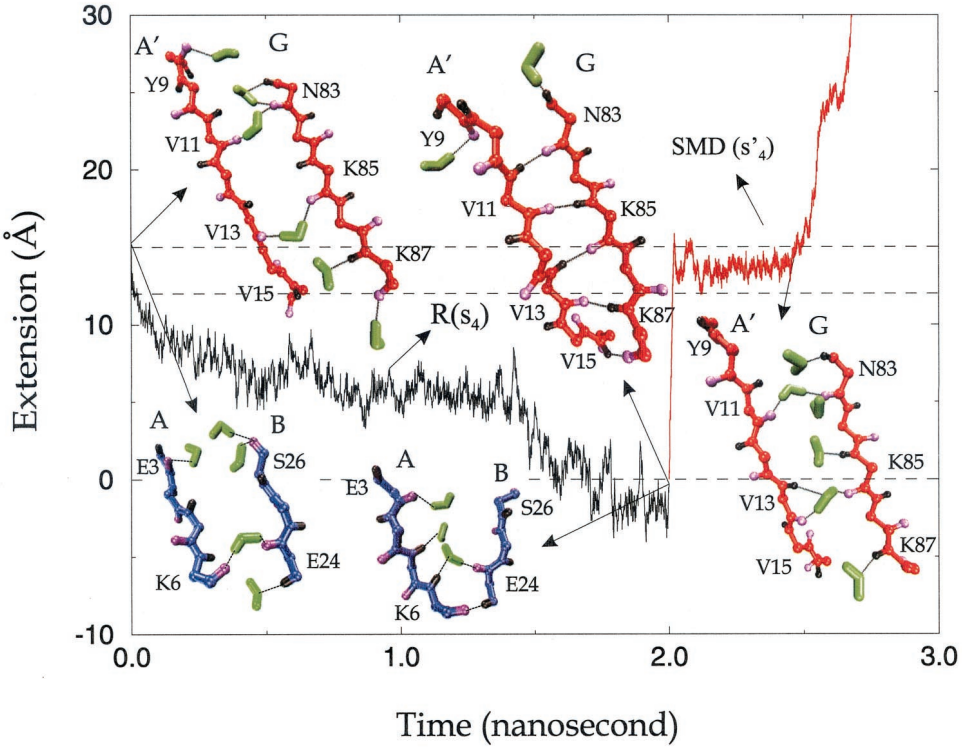

FIGURE 6 Extension versus time profiles from simulations R(s_4) (*black*) and SMD(s'_4) (*red*), together with snapshots of β-strands A and B and β-strands A′ and G at various stages during the simulations. The left snapshots show β-strands A and B (*bottom left*) and β-strands A′-G (*upper left*) of the initial structure s_4. After 2-ns relaxation, the extension of the domain decreased from over 15 Å, to a value of ~0 Å. One hydrogen bond between A and B (*bottom center*) and five hydrogen bonds between A′ and G (*upper right*) have stably reformed. The protein was then stretched by applying a constant force of 750 pN. The bottom right snapshot of β-strands A′ and G was captured when the hydrogen bonds between β-strands A′–G were broken again and water molecules formed hydrogen bonds with backbone oxygens and hydrogens. Oxygen atoms, hydrogen atoms, and water molecules are colored in purple, black, and green. Other backbone atoms in snapshots of A and B and A′ and G are colored in blue and red, respectively.

edges of the hydrogen bond cluster. However, in simulation R(s_8) none of the hydrogen bonds has been found to reform over a 2-ns time period, though the extension of the protein decreased from 23 Å to ~9 Å. Five of the six broken A′-G backbone hydrogen bonds in R(s_8), however, were observed to shorten their O-H distance from 10–13 Å to 7–10 Å, demonstrating a tendency for reformation.

Summary of refolding simulations R(s_1) to R(s_8)

Simulations R(s_1) to R(s_3) were performed to simulate the refolding of A-B β-strands. The simulations R(s_4) to R(s_8) simulated the reformation of hydrogen bonds between both A-B strands and A′-G strands. In Table 2 we have summarized the reformation of hydrogen bonds and overall extension from all eight refolding simulations. We consider a hydrogen bond stably reformed if the bond energy has been <-1 kcal/mol for over 500 ps during the second half of the simulation period, i.e., the period from 1 ns to 2 ns. The three hydrogen bonds between β-strands A and B have seldom been stably reformed, except for one case in R(s_4), as shown in Fig. 7. However, in four of five simulations of A′-G reformation most of the backbone hydrogen bonds characterizing this pair of β-strands do reform quickly in the first half of the 2-ns simulation period, and then stay in an energetically favorable configuration (<-1 kcal/mol) for over 500 ps for the rest of time, especially those located in the middle of the hydrogen bond cluster.

In all simulations except R(s_8) the domain length shrinks to a near-native level. In R(s_4), where β-strands A′-G reform stably and one hydrogen bond in β-strand A-B reforms, the domain refolds and fluctuates around its native folded length. Six simulations for which β-strands A-B become not stably bonded reveal an overall extension that is longer by 2 to 5 Å than the native fold. In R(s_8), where both β-strands A′-G and β-strands A-B do not reform, the extension is 9 Å longer than for the native fold. This observation indicates that the extension of the protein can serve well as the reaction coordinate in the unfolding process.

DISCUSSION

Starting from the partially stretched I27 domain, the refolding simulations showed that the reformation of β-sheet AB

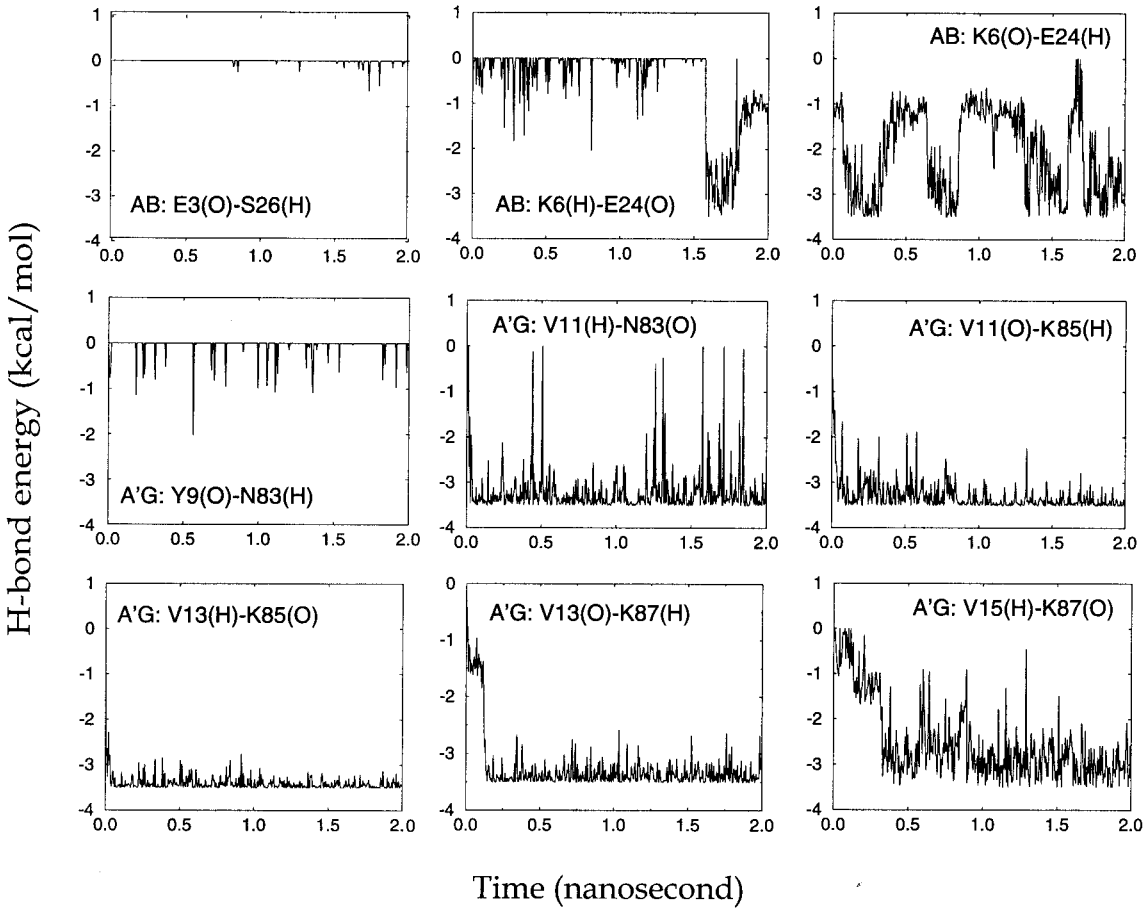

FIGURE 7 Fluctuation of interstrand hydrogen bonds between β-strands A and B, and β-strands A′ and G during simulation R(s_4). Hydrogen bonds between A′ and G reform in 500 ps, except the one between Y9 and N83, which never reforms during 2-ns simulation. Hydrogen bonds between β-strands A and B have a tendency to reform, except the one between E3 and S26, but remain unstable and are easily disrupted during the simulations.

and β-sheet A′G exhibit different characteristics. Comparing the current study with the 1-ns equilibration described in Lu and Schulten (2000), it is interesting to notice that the more stable a hydrogen bond appeared during equilibration, the easier this hydrogen bond reforms stably during refolding. In both equilibration and refolding simulations, the hydrogen bonds between the same pair of β-strands showed a certain cooperativeness. The larger number of hydrogen

TABLE 2 Hydrogen bonding and overall extension after refolding

Refolding Simulations	AB H-bonds		A′G H-bonds		Extension (Å)	
	Reformed	Stabilized	Reformed	Stabilized	Initial	Final
R(s_1)	3	0	–	–	11	2
R(s_2)	2	0	–	–	13	4
R(s_3)	3	0	–	–	12	4
R(s_4)	2	1	6	5	15	0
R(s_5)	2	0	6	5	16	5
R(s_6)	1	0	5	5	16	3
R(s_7)	2	0	5	4	17	5
R(s_8)	0	0	0	0	23	9

A summary of the number of reformed hydrogen bonds between A and B strands and A′-G strands and the extension of the I27 domain in eight refolding simulations R(s_k) ($k = 1 - 8$), respectively. A hydrogen bond is defined as reformed if the value of its hydrogen bond energy has ever been <-1 kcal/mol during the 2-ns refolding simulation, and it is counted as a stably reformed hydrogen bond if the energy of the bond has been <-1 kcal/mol for over 500 ps during the second half of the simulation period, i.e., from 1 ns to 2 ns. The final extension is calculated by averaging the extension of the protein over the last 500 ps in a simulation.

bonds involved in β-sheet A′G than those in β-sheet AB may determine why β-sheet A′G reforms much faster than β-sheet AB.

AFM experiments revealed a small transitional extension of I27 tandem repeats before the main unfolding event (Marszalek et al., 1999). This humplike extension of 6.6 Å per module was suggested to correspond to the unfolding of hydrogen bonds between β-strands A and B, and further analysis implied that an average time period of 40 ms must pass before I27 can fully refold from this intermediate. Our refolding simulations of intermediates with broken interstrand hydrogen bonds between β-strands A and B demonstrated a decrease in the extension of I27, from a configuration of early forced unfolding to close to its native fully folded length, in <1 ns, as shown in Figs. 3 and 4. The energy analysis of hydrogen bonds shown in Figs. 5 and 7 seldom exhibits a stable reformation of the bonds between β-strands A and B, except that the K6(O)-E24(H) in R(s_4) stably reforms, yet the tendency for refolding in terms of overall domain length was clearly demonstrated within the nanosecond time scale of the simulations. The rare reformation of hydrogen bonds between β-strands A and B may explain the low probability of "hump" events in the post-initial force peaks of AFM experiments, because hydrogen bonds bridging β-strands A and B of unfolded domains might remain ruptured during the unfolding of the first module and the following unfolding events. In SMD(s'_3) we applied a constant force of 100 pN to the partially refolded intermediate, showing that the domain elongates from 4 Å to 5 Å in <200 ps. As suggested in Marszalek et al. (1999), the disruption of hydrogen bonds between A and B corresponds to the transitional extension of ~6 Å. The extension of the I27 domain in the re-stretching SMD simulation shows that the separation between β-strands A and B increases the extension again easily, which is in good agreement with the observations in AFM experiments. The current work provides new evidence to support our previously proposed mechanism that hydrogen bonds between β-strands A and B control the unfolding intermediate of I27.

The refolding simulations in this study extend previous studies (Lu and Schulten, 2000) confirming again that water plays a major role in the reformation of backbone hydrogen bonds. As one can see in snapshots shown in Figs. 3 and 4 that characterize refolding of the AB intermediates, water molecules compete for hydrogen bonds with backbone atoms on β-strands A and B, hindering the reformation of hydrogen bonds between backbone atoms themselves. This behavior, however, cannot be accounted for by the implicit solvent models that Paci and Karplus (1999, 2000) used.

In contrast to the refolding of β-sheet AB, the refolding of β-sheet A′G occurs very fast, reforming hydrogen bonds between β-strands A′ and G shortly after their concurrent disruption. As an example, one can see from Fig. 6 of simulation R(s_4) that water molecules in the initial snapshot of A′-G have been driven out after 2 ns relaxation and the extension is reduced to the native folded length of the domain. Further details of the hydrogen bond energy analysis in Fig. 7 show that, except for the bond between Y9(O) and N83(H), the five other interstrand hydrogen bonds between β-strands A′ and G reform in <400 ps and remain in an energetically favorable configuration thereafter. Because the length of hydrogen bond Y9(O)-N83(H) had been demonstrated to be more flexible (fluctuating between 2 and 5 Å during the 1-ns equilibration reported in Lu and Schulten (2000)) than that of the other five hydrogen bonds between β-strands A′ and G, it is not surprising that the reformation of hydrogen bond Y9(O)-N83(H) requires more time. The fluctuation of the energies of the other five hydrogen bonds shown in Fig. 7 closely resembles the fluctuation observed during the equilibration of the native I27 domain (Lu and Schulten, 2000) and reflects stable bonds. When the I27 domain is stretched further (beyond 17 Å), it is more difficult for A′G hydrogen bonds to reform during the 2-ns simulation period. As shown in Table 2, simulation R(s_8) was started at an extension of 23 Å, and none of the hydrogen bonds between β-strands A′ and G is subsequently reformed, though the distance between hydrogen bond partners decreases ~3 Å.

The relatively fast refolding between β-strands A′ and G may be a mechanical property that leads to the characteristic sawtooth pattern generated in AFM experiments (Rief et al., 1997; Carrion-Vazquez et al., 1999). With this property, even when the patch of hydrogen bonds between β-strands A′ and G of two domains is ruptured concurrently by accident, if the domain remained in near-native extension the hydrogen bonds reformed quickly and contributed to a second force peak in a later unraveling event. In SMD(s'_4) we applied a constant force of 750 pN to the partially refolded structure, showing a 500-ps-long plateau where the domain is trapped in a transitional state. This passage time of 500 ps, spent on crossing the energy barrier, is comparable to what we observed earlier in stretching native I27 domains (~800 ps, as shown in Fig. 2). The close agreement of the mean first passage times indicates good reformation of hydrogen bonds between β-strands A′ and G, reformed at the end of simulation R(s_4), and restoration of the I27's resistance against stretching.

One must note that the fast reformation of β-sheet A′G does not correspond to the refolding of I27 from a random coil or completely extended conformation, which would take milliseconds to seconds. The simulation of the complete folding process in these cases is still impossible due to the limitation of computational power. According to the current understanding of protein folding, a major barrier lies in the formation of native topology. Our starting structures are only partially unfolded I27 domains that already have the correct topology. Thus the refolding simulations carried out in this study correspond to the final step of the protein folding process. Our emphasis here is the atomic level description of the formation of the force-bearing hydrogen

bonds between β-strands A′ and G, which control I27's mechanical behavior.

We thank Barry Isralewitz for preparation of the simulations and critical reading of the manuscript.

This work was supported by the National Institutes of Health (Grant NIH PHS 5 P41 RR05969) and the National Science Foundation (Grants NSF BIR 94-23827EQ, NSF/GCAG BIR 93-18159, MCA93S028).

REFERENCES

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.

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

Clarke, J., S. J. Hamill, and C. M. Johnson. 1997. Folding and stability of a fibronectin type III domain of human tenascin. *J. Mol. Biol.* 270:771–778.

Craig, D., A. Krammer, K. Schulten, and V. Vogel. 2001. Comparison of the early stages of forced unfolding of fibronectin type III modules. *Proc. Natl. Acad. Sci. U.S.A.* 98:5590–5595.

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. U.S.A.* 91: 10114–10118.

Erickson, H. 1997. Stretching single protein modules: titin is a weird spring. *Science.* 276:1090–1093.

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

Evans, E., and K. Ritchie. 1999. Strength of a weak bond connecting flexible polymer chains. *Biophys. J.* 76:2439–2447.

Fisher, T. E., P. E. Marszalek, and J. M. Fernandez. 2000. Stretching single molecules into novel conformations using the atomic force microscope. *Nature Struct. Biol.* 7:719–724.

Fong, S., S. J. Hamill, M. Proctor, S. M. V. Freund, G. M. Benian, C. Clothia, M. Bycroft, and J. Clarke. 1996. Structure and stability of an immunoglobulin superfamily domain from twichin, a muscle protein of the nematode *Caenorhabditis elegans*. *J. Mol. Biol.* 264:624–639.

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.

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

International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. *Nature.* 409:860–921.

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.

Kalé, L., R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. 1999. NAMD2: Greater scalability for parallel molecular dynamics. *J. Comp. Phys.* 151:283–312.

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.

Klimov, D.K., and D. Thirumalai. 1999. Stretching single-domain proteins: phase diagram and kinetics of force-induced unfolding. *Proc. Natl. Acad. Sci. U.S.A.* 96:1306–1315.

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.

Krammer, A., H. Lu, B. Isralewitz, K. Schulten, and V. Vogel. 1999. Forced unfolding of the fibronectin type III module reveals a tensile molecular recognition switch. *Proc. Natl. Acad. Sci. U.S.A.* 96: 1351–1356.

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.

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

Linke, W. A., D. E. Rudy, T. Centner, M. Gautel, C. Witt, S. Labeit, and C. C. Gregorio. 1999. I-band titin in cardiac muscle is a three-element molecular spring and is critical for maintaining thin filament structure. *J. Cell Biol.* 146:631–644.

Linke, W. A., M. R. Stockmeier, M. Ivemeyer, H. Hosser, and P. Mundel. 1998. Nature of PEVK-titin elasticity in skeletal muscle. *Proc. Natl. Acad. Sci. U.S.A.* 95:8052–8057.

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. 1999a. Steered molecular dynamics simulations of force-induced protein domain unfolding. *Proteins.* 35:453–463.

Lu, H., and K. Schulten. 1999b. Steered molecular dynamics simulation of conformational changes of immunoglobulin domain I27 interpret atomic force microscopy observations. *Chem. Phys.* 247:141–153.

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

MacKerell, Jr., A. D., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, I. W. E. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-hydrogen empirical potential for molecular modeling and dynamics studies of proteins using the CHARMM22 force field. *J. Phys. Chem. B.* 102:3586–3616.

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.

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

Minajeva, A., M. Kulke, J. M. Fernandez, and W. A. Linke. 2001. Unfolding of titin domains explains the viscoelastic behavior of skeletal myofibrils. *Biophys. J.* 80:1442–1451.

Oberhauser, A. F., P. E. Marszalek, M. Carrion-Vazquez, and J. Fernandez. 1999. Single protein misfolding events captured by atomic force microscopy. *Nature Struct. Biol.* 6:1025–1028.

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.

Paci, E., and M. Karplus. 1999. Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations. *J. Mol. Biol.* 288:441–459.

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.

Plaxco, K. W., C. Spitzfaden, I. D. Campbell, and C. M. Dobson. 1996. Rapid refolding of a proline rich all β-sheet fibronectin type III domain. *Proc. Natl. Acad. Sci. U.S.A.* 93:10703–10706.

Plaxco, K. W., C. Spitzfaden, I. D. Campbell, and C. M. Dobson. 1997. A comparison of the folding kinetics and thermodynamics of two homologous fibronectin type III modules. *J. Mol. Biol.* 270:763–770.

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.

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.

Wang, K. 1996. Titin/connectin and nebulin: giant protein ruler of muscle structure and function. *Adv. Biophys.* 33:123–134.