

Steered molecular dynamics simulation of conformational changes of immunoglobulin domain I27 interpret atomic force microscopy observations

Hui Lu^{a,b}, Klaus Schulten^{a,c,1}

^a Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

^b Department of Nuclear Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

^c Beckman Institute for Advanced Science and Technology, Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

Received 9 February 1999

Abstract

Atomic force microscopy and steered molecular dynamics investigations of the response of so-called mechanical proteins like titin, tenascin or their individual immunoglobulin and fibronectin type III domains have lead to qualitative insights about the relationship between the β sandwich domain architecture and the function of this class of proteins. The proteins, linear segments of up to hundreds of domains, through strain induced shape changes, unfolding and refolding, maintain order and elasticity in cellular systems over a nearly tenfold length scale. In this paper we develop a steered molecular dynamics description of the response of the titin immunoglobulin domain I27 at the onset of domain unfolding in quantitative agreement with AFM observations. We show that if forces stronger than 50 pN are applied to the terminal ends the two hydrogen bonds between the antiparallel A and B β strands break with a concomitant 6–7 Å elongation of the protein. If forces strong enough to unfold the domain are applied, the protein is halted in this initial extension until the set of all six hydrogen bonds connecting strands A' and G break simultaneously. This behavior is accounted for by a barrier separating folded and unfolded states, the shape of which is consistent with AFM and chemical denaturation data. We also demonstrate that steered molecular dynamics simulations which induce unfolding through slow pulling (speed 0.1 Å/ps) predict unfolding forces that are within a factor of two within force values extrapolated from AFM observations. © 1999 Elsevier Science B.V. All rights reserved.

1. Introduction

Single molecule techniques, atomic force microscopy (AFM) and optical tweezers, have been applied recently to study the conformational response of proteins to stretching. This has opened a

new avenue for investigating the structure function relationship of proteins, in particular, of the so-called mechanical proteins, i.e., cellular proteins evolved to sustain and control mechanical strain. Atomic force microscopy has measured the response of the muscle protein titin [1,2] and of the extracellular matrix protein tenascin [3] when their terminal ends are pulled apart. The optical tweezers technique has been used likewise to stretch single titin molecules [4,5].

¹ E-mail: kschulte@ks.uiuc.edu

Titin, also known as connectin, is a 30,000 amino acid protein that spans half of the muscle sarcomere and plays important roles in (skeletal and smooth) muscle extension and elasticity [6,7] as well as in controlling the shapes of chromosomes in the cell's nucleus [8]. Titin is the largest known protein and consists of about 300 immunoglobulin-like (Ig) and fibronectin type III (FnIII) domains [9]. When muscle is stretched, titin extends while holding the sarcomere together and provides a passive force. Under extreme conditions like in overstretched muscle, the Ig domains in the titin I-band will unfold to provide the necessary extensions. When forces are released, the unfolded Ig domains refold quickly [2].

Tenascin is an extracellular matrix protein that mediates the attachment of cells to a substrate and the movement of cells on the substrate [10]. The FnIII domains in tenascin act in tandem, e.g., as a tandem of fifteen FnIII domains, endow the tenascin-ligand connection with elasticity and allow the connection to persist over a wide range of extensions [3].

A key property of titin and tenascin entails their protection against strain-induced domain unfolding. AFM experiments have demonstrated that rather strong forces of the order of 100 pN need to be exerted before Ig and FnIII domains get ruptured and unfolded. Proteins which do not have to sustain mechanical strain physiologically have been found to exhibit little resistance against unfolding through stretching forces. This has been demonstrated through AFM experiments on spectrin [11] and on genetically engineered polyproteins consisting of several calmodulin units (J. Fernandez et al., manuscript in preparation).

Atomic force microscopy permits accurate measurement of the forces exerted on a single protein molecule at any specified extension. Respective experiments, when stretching the multi-domain proteins titin and tenascin, record a force-extension profile with a sawtooth pattern [2,3]. AFM force-extension profiles of engineered polyproteins consisting of only a single type of Ig domain reveal a similar sawtooth pattern [12].

The key characteristics of sawtooth AFM force-extension profiles of the multi-domain proteins titin and tenascin are the spacing between the force peaks and the height of the force peaks [2,1,3]. The spacing

of the force peaks matches the length of the completely extended polypeptide chain of one Ig or FnIII domain, proving that when these multi-domain proteins are stretched their domains unfold one by one. The high values of the force peaks (100–300 pN) imply that the Ig and FnIII domains are designed to withstand significant stretching forces. At the extensions where the force peaks occur, the domains lose their resistance and at longer extension the forces drop dramatically. The peak values of the force depend on the type of domains being pulled and on the pulling speed adopted in an experiment. For a pulling speed of 1 $\mu\text{m/s}$, unfolding of titin Ig domains requires about 200 pN while unfolding of tenascin FnIII domains requires only about 140 pN. Unfolding of the helix protein spectrin and of calmodulin units at the same speed requires much weaker forces; in the latter case no distinctive force peak can be discerned. In case of Ig and FnIII domains the force peak values increase logarithmically with the pulling velocity [2,3].

Naturally, one would like to explain the stated observations in terms of the structural properties of the proteins. Currently, only one experimental structure of titin I-band immunoglobulin (Ig) domains is available, the 27th Ig domain (I27) [13]. Fig. 1 demonstrates that I27 forms a β sandwich. AFM experiments of I27 stretching do not resolve atomic level detail of a single molecule's conformational change during the unfolding. Computer simulation of stretching proteins represented by a lattice model [14] also lacked atomic level interactions. Molecular dynamics simulations, in principle, can complement the AFM observations in providing a detailed picture of stretching and unfolding.

Steered molecular dynamics (SMD) simulations [15] are ideally suited for this purpose, adding external forces to conventional force fields, the forces imitating the effects of AFM cantilevers on protein domains. SMD simulations have been applied successfully to ligand binding and unbinding processes [16–19] and, indeed, also to the unfolding of titin's Ig domain I27 [20], of the FnIII₁₀ domain of fibronectin [21], and of ten other domains [22]. Such simulations need to prove their accuracy through comparison with the AFM data. However, due to limits in computer power, SMD simulations presently cover only nanosecond processes requiring forces

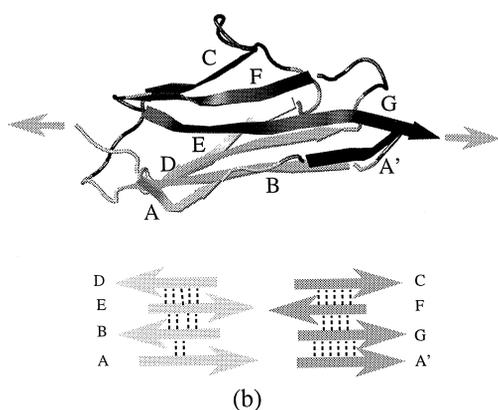


Fig. 1. Titin immunoglobulin domain I27. I27 is a β sandwich domain with four β strands on each sheet. (a) NMR structure of I27. The two sheets are drawn in grey (β strands A, B, D, and E) and black (β strands A', G, F, and C); arrows show the directions of the stretching force applied to the termini. (b) Schematic presentation of I27's secondary structure and hydrogen bonding pattern. (Protein structure shown in this and in the following figures were created with VMD [36].)

stronger than those in AFM experiments, which unfold domains on a millisecond timescale. Nevertheless, simulations of the unfolding of individual I27 and FnIII₁₀ domains reproduced a single dominant force peak as recorded in observed force-extension profiles and revealed how this force peak arises: stretching of these protein domains at their two terminal ends is halted by a set of hydrogen bonds which need to be broken concertedly before the domains unfold; the peak force is the force sufficient to break the hydrogen bonds simultaneously. In case of I27, the hydrogen bonds broken are those between β strands A' and G and between A and B. Subsequently, the remaining interstrand hydrogen bonds can be broken individually in a zipper-like fashion, with the result that further stretching and unfolding requires only weak forces [20–22].

The results of the SMD simulations described so far have been found only in qualitative agreement with AFM observations [20,21] and discrepancies in quantitative detail still cast doubt on the stated explanation of the force peak. In fact, the peak values of the force in the SMD simulations were about ten-fold higher than observed values [20], a deficit, which had been attributed to the much larger pulling speeds

in the simulations. Naturally, one wishes to test the supposition that slower stretching in SMD simulations leads to weaker peak forces, does not change the hydrogen bond breaking scenario described in Refs. [20–22], and reaches in the limit of millisecond stretching the peak forces observed in AFM experiments.

In order to gain more insight into the mechanism that protects mechanical proteins against stretching and in order to provide more opportunity for comparing simulations and observations one seeks a more refined analysis of AFM observations. Can AFM experiments applying new force schedules resolve more detailed information on the initial phase of unfolding and can SMD simulations reproduce such observations?

In this paper we report an important advance in this respect. We will show that a decrease of the pulling speed in SMD simulations by a factor of ten does indeed reduce strongly the peak values of forces encountered during unfolding of the I27 domain without altering the qualitative involvement of hydrogen bond breaking. We will also study the behavior of the I27 domain when stretched by a weak force between 50 pN and 200 pN and demonstrate that for forces exceeding 50 pN the domains are prestretched through breaking of the two hydrogen bonds between β strands A' and G with a resulting geometrical extension of the domain in good agreement with observation [23]. Finally, we show that stationary forces between 500 pN and 1000 pN yield an unfolding behavior that can be interpreted through the theory of mean first passage times [24–26] in terms of an energy barrier separating folded and unfolded states, the (extension) width and height of which agrees well with AFM observation and chemical denaturation data [12]. Eighteen simulations of I27 stretched with a stationary force of 750 pN are shown to lead to a distribution of first passage times that agrees qualitatively with the predictions based on solving of the Smoluchowski diffusion equation [26]. Altogether, this paper demonstrates that a combination of AFM observations and SMD simulation provides a powerful tool for the study of mechanical proteins.

In the following we will first describe the methodological aspects of SMD simulations, present the results of I27 stretching for different pulling speeds

as well as different stationary pulling forces and interpret the SMD trajectories.

2. Methods

The energy-minimized average NMR structure of I27 [13], the 27th immunoglobulin domain of the I-band of cardiac titin, had been obtained from the Brookhaven protein data bank (entry 1TIT). This domain adopts the typical I-frame immunoglobulin superfamily fold [27], consisting of two β sheets packing against each other, as shown in Fig. 1, with each sheet containing four strands. The first sheet comprises strands A, B, E, and D, the second sheet comprises strands A', G, F, and C. All adjacent β strands in both sheets are anti-parallel 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. The structure is stabilized by hydrophobic core interactions between the two β sheets and by hydrogen bonds between β strands.

The simulations were carried out using the CHARMM22 force field [28]. The non-bonded Coulomb and vdW interactions were calculated with a cut-off using a switching function starting at a distance of 10 Å and reaching zero at 13 Å. The TIP3P water model was employed for the solvent [29].

The experimental structure of I27 was solvated before the SMD simulations. I27 was surrounded by a sphere of water molecules of 31 Å radius, which covered the molecular surface of the domain by at least four shells of water molecules. The water molecules within 2.6 Å of the domain surface or within the volume occupied by the domain were deleted. About 3,300 water molecules were kept in the solvated I27 with the resulting protein–water system consisting of 11,400 atoms.

The system described was gradually heated over 10 ps to 300 K, equilibrated with a thermal bath at 300K for another 10 ps. This procedure was followed by 100 ps of free dynamics without constraints and heat bath. From this resulted a stable solvated structure, with a temperature fluctuation of 5 K and a backbone root-mean-square-deviation from the NMR structure [13] of 1 Å. The equilibration and following SMD simulations were performed with the

programs NAMD [30] and XPLOR [31], assuming an integration time step of 1 fs and a uniform dielectric constant of 1.

SMD simulations with constant velocity stretching were carried out by fixing the C_{α} atoms of the first residue (Leu1 in I27), and by applying external forces to the C_{α} atom of the last residue (Leu89 in I27). The latter forces were implemented by restraining the C_{α} atom of the last residue harmonically to a restraint point and moving the restraint point with a constant velocity v in the desired direction. The procedure adopted is equivalent to attaching one end of a harmonic spring to the C_{α} atom of the last residue and pulling on the other end of the spring. This is similar to the procedure performed on titin and tenascin in AFM experiments [2,3], except that the pulling speeds adopted in the simulations are about six orders of magnitude higher than those in the experiment.

The forces experienced by the pulled atom are

$$F = k(vt - x). \quad (1)$$

Here x is the displacement of the pulled atom from its original position, v is the pulling velocity, vt is the position of the restraining point, and k is the spring constant. The pulling direction was chosen along the vector from the C_{α} atom of the first residue to the C_{α} atom of the last residue. We chose for k a value of $10 k_B T / \text{Å}^2$ with $T = 300$ K; this choice implies spatial (thermal) fluctuations of the constrained C_{α} atom of $\delta x = \sqrt{k_B T / k} = 0.38$ Å. To realize a movement of the restraint point with nearly constant velocity, the position of the restraint point was changed every 100 fs by $v\Delta t$. The reader should note that the restraining point, not the C-termini of the domain, was forced to move at constant velocity.

SMD simulations of constant force stretching were implemented by fixing the N-terminus of the domain I27 and by applying a constant force to the C-terminus along the direction connecting the initial positions of N-terminus and C-terminus.

The atomic coordinates of the whole system were recorded every picosecond. For constant velocity stretching, the elongation $d(t)$, defined as the increase of the end-to-end distance between the termini from that of the native fold was monitored along with the force $F(t)$. For the analysis presented below, often the force was plotted as a function of

extension d . The $(F(t), d(t))$ graphs will be referred to as the force-extension profile. In case of the constant force stretching simulation, the elongation $d(t)$ was recorded and plotted as $(d(t), t)$, which will be referred to as the extension curve.

Three pulling speeds have been adopted in our constant velocity stretching simulations, namely, 1.0 Å/ps, 0.5 Å/ps, and 0.1 Å/ps. The respective simulations will be referred to as SMD-(1.0 Å/ps), SMD-(0.5 Å/ps) and SMD-(0.1 Å/ps). Eight values of a stationary force have been adopted for stretching I27: four weaker forces, 50, 100, 150, 200 pN and four stronger forces, 500, 750, 900, and 1000 pN; the respective simulations will be denoted by SMD-(50 pN), ... SMD-(1000 pN).

3. Results

Fig. 2 compares the force-extension profiles for simulations SMD-(1.0 Å/ps), SMD-(0.5 Å/ps) and SMD-(0.1 Å/ps). The three profiles are qualitatively similar: at an extension of about 15 Å arises the dominant force peak that, for the three individual profiles, is about 2 to 3 times higher than forces of other extensions; at extensions larger than 20 Å the domain exhibits little resistance against stretching as

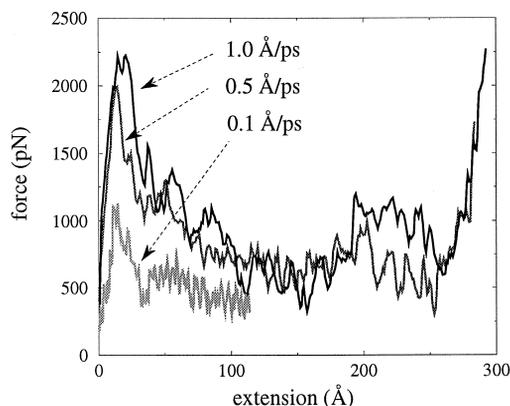


Fig. 2. Force extension curves of I27 stretching for simulations SMD-(1.0 Å/ps), SMD-(0.5 Å/ps) and SMD-(0.1 Å/ps). The simulation times for pulling velocities 1.0 Å/ps, 0.5 Å/ps, and 0.1 Å/ps are 0.3 ns, 0.6 ns and 1.2 ns, respectively. Simulation of SMD-(0.1 Å/ps) had been stopped at an extension of 120 Å, whereas simulations SMD-(1.0 Å/ps) and SMD-(0.5 Å/ps) had been stopped at about 300 Å.

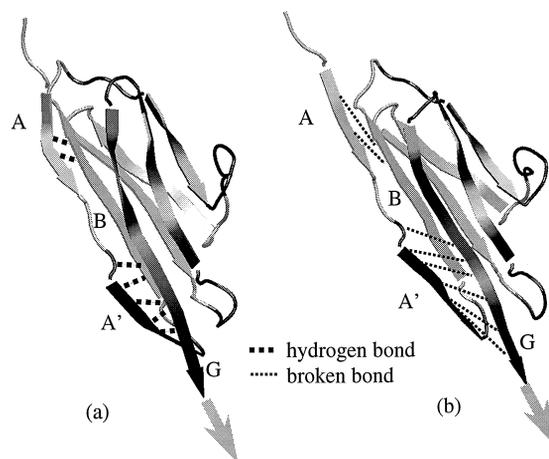


Fig. 3. Two snapshots of I27 unfolding. The domain is drawn in cartoon representation and key hydrogen bonds between strands A–B, and between strands A'–G are shown as dotted lines. (a) Snapshot of I27 before the force peak at 10 Å extension. The hydrogen bonds are all maintained. (b) Snapshot of I27 after the force peak at 17 Å extension. The hydrogen bonds between strands A–B, and between strands A'–G are broken, initiating unfolding.

is evident from the fact that only relatively weak forces are needed to increase extension; at an extension corresponding to a fully stretched polypeptide strand, i.e., about 270 Å, the force required to stretch the now completely unfolded domain increases again as expected. The force-extension profiles in Fig. 2 show also that the lower the pulling speed, the lower the forces needed to extend the domain. The decrease of the speed from 1.0 Å/ps to 0.5 Å/ps to 0.1 Å/ps reduces the peak force from 2200 pN to 2000 pN to 1200 pN, respectively.

An analysis of the trajectories corresponding to the profiles in Fig. 2 reveals that at the extension of the maximum force eight hydrogen bonds break concurrently. Fig. 3 presents snapshots of the domain directly before and after the force peak in which the extensions measure 10 Å and 17 Å. Despite the relatively small change in extension all of the eight hydrogen bonds between β strands A' and G as well as A and B are broken.

The Ig domain has been shown in a recent AFM experiment to experience prior to unfolding an elongation of 7 Å when stretching forces applied exceed 100 pN. In order to derive insights from this AFM

observation, the simulations SMD-(50 pN), SMD-(100 pN), SMD-(150 pN), and SMD-(200 pN) were carried out for 1 ns periods. The comparison between experiment and simulation results as well as the corresponding implication for Ig's function in titin has been presented in Ref. [23], here we will briefly summarize simulation results of that work. During the simulations the 50–200 pN forces did not suffice to unfold the Ig domain within a nanosecond. Fig. 4a presents the extension curves $d(t)$ resulting from these simulations which all reached stationary values within about 200 ps. The final extension resulting in simulation SMD-(50 pN) is about 2 Å, whereas

simulations SMD-(100 pN), SMD-(150 pN), and SMD-(200 pN) all reached larger values around 6–7 Å. Analysis of the SMD trajectories revealed that the difference between a short extension for the 50 pN case and larger extensions for stronger forces arises from a breaking of the two hydrogen bonds between β strands A and B. This breaking of two hydrogen bonds has been suggested as the origin of the domain elongation observed in AFM above 100 pN [23].

Simulations SMD-(500 pN), SMD-(750 pN), SMD-(900 pN), and SMD-(1000 pN) explored the influence of the stretching forces on the time which passes until the domain begins to unfold. In AFM

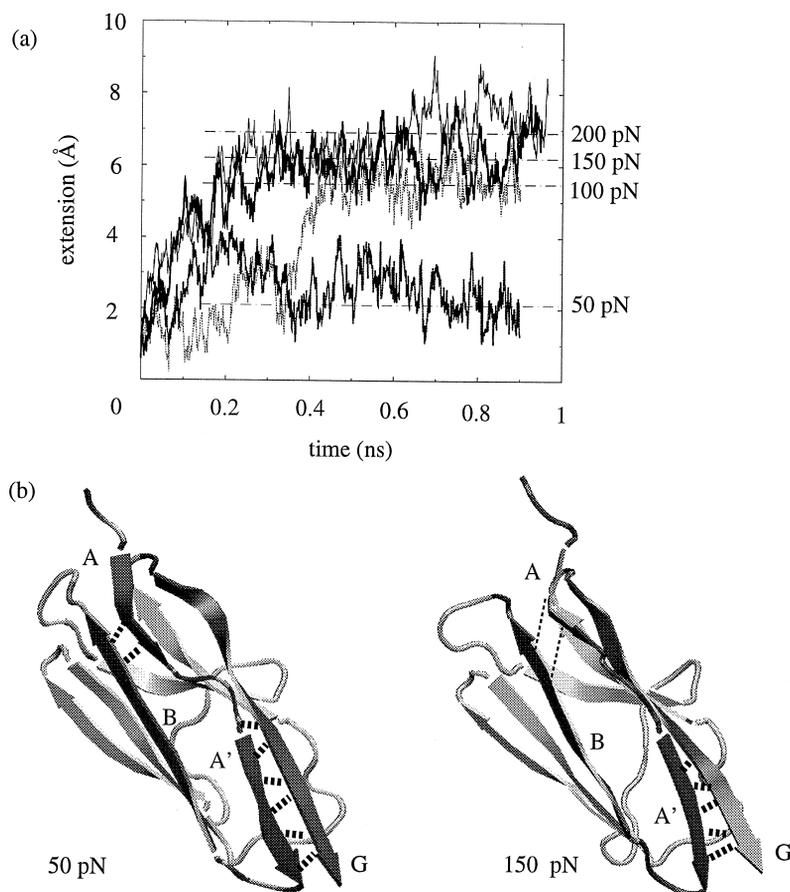


Fig. 4. I27 stretching by means of stationary forces of 50, 100, 150 and 200 pN. The data are from [23]. (a) Extension versus time plot. The dashed horizontal lines indicate the average extensions (from 0.5 ns to 1 ns) for the various forces applied to the domain. (b) Snapshots of I27 at 0.8 ns for stationary forces of 50 and 150 pN. The domain is drawn in cartoon representation with β strands A, A', B, and G marked and shown in black, other strands being shown in grey. Hydrogen bonds between strands A–B as well as between strands A'–G are presented as dotted lines. One can recognize that A–B bonds remained intact in case of the 50 pN snapshot, but became extended, i.e., broken, in case of the 150 pN snapshot.

observations that have applied forces of about 200 pN, unfolding times are about 1 ms. In the stated SMD simulations the domain exhibits a rapid extension until the $d(t)$ curves reach a plateau value, the plateau being more pronounced for the weaker forces. This behavior is shown in Fig. 5. During the plateau period the domain's extensions fluctuate around a constant value, but at one point exhibit a strong increase that extends into a nearly linear (with time) increase to very large extensions. The latter increase is a clear signature of unfolding. One can measure

from the $d(t)$ curves the times when the plateau value for d is reached for the first time and the time when unfolding begins. The difference in time will be denoted by τ_{barrier} . One can attribute τ_{barrier} to a barrier crossing process, the barrier arising from the six hydrogen bonds between β strands A' and G.

An analysis of simulations SMD-(500 pN), SMD-(750 pN), SMD-(900 pN), and SMD-(1000 pN) showed that the initial extension of the domain is due to a straightening of the polypeptide chain near its terminal ends and due to a breaking of the two

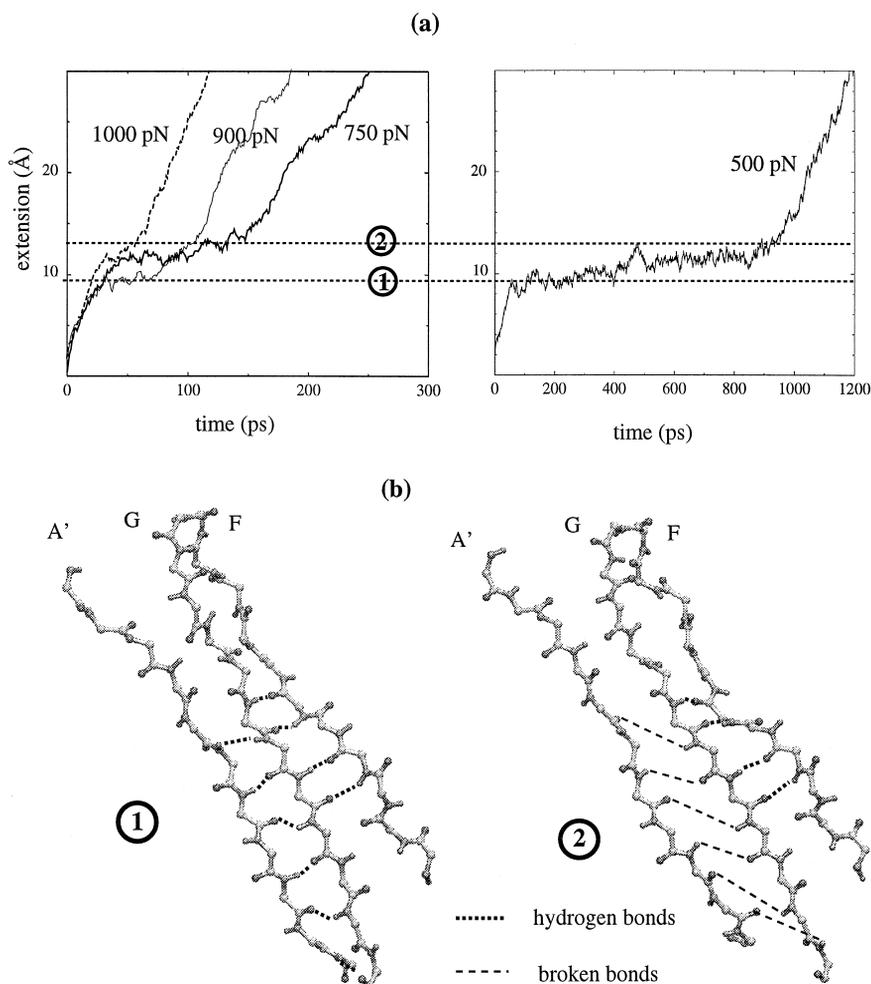


Fig. 5. I27 unfolding by means of stationary forces of 500, 750, 900 and 1000 pN strength. (a) Extension versus time plots. In all cases the domain extended quickly by 10 Å, then remained approximately at constant length for long periods of time until the system unfolded (appearing as linear extension with time). (b) Snapshots of β strands A', G, and F before and after the plateau region seen in the extension curve in (a). One can recognize that at the beginning of the plateau region all hydrogen bonds are maintained, and that at the end of the plateau region the hydrogen bonds between strands A' and G are broken.

hydrogen bonds between β strands A and B. The plateau region is reached when the six hydrogen bonds come under mechanical strain. When these bonds are broken the domains begin to extend rapidly and to leave the plateau region. The snapshots of I27 before and after the plateau (Fig. 5b) reveal the breaking event as described.

When forces are applied to the domain's termini the motion separating diametrically the termini experiences the stated barrier and comes to a halt in front of it. However, the applied force effectively lowers the barrier such that stronger forces lead to faster barrier crossing ($\tau_{\text{barrier}} = 0.04$ ns for 1000 pN) than weaker forces ($\tau_{\text{barrier}} = 0.9$ ns for 500 pN). In any case, the motion gets "stuck" in front of the barrier and only thermal fluctuations permit the system to overcome the barrier. This scenario can be described as Brownian motion governed by a potential which is the sum of the indigenous barrier and a linear potential accounting for the applied force. For this scenario the mean time to cross the barrier can be evaluated using the expressions for the mean first passage time developed in Refs. [24–26]. By comparing the mean first passage times with the respective times τ_{barrier} for various forces one can estimate the height of the indigenous potential barrier.

Fig. 6 shows the dependence of τ_{barrier} on the applied force. One can apply the theory of mean first passage times to estimate the shape of the indigenous potential barrier assuming for this purpose a simple model for the barrier provided by

$$U(x) = \begin{cases} +\infty & \text{for } x < a, \\ \Delta U(x-a)/(b-a) & \text{for } a \leq x \leq b, \\ -\infty & \text{for } x > b. \end{cases} \quad (2)$$

Here ΔU is the height of the barrier, and $b-a$ the barrier width. The choice of this potential function is dictated by the fact that for this barrier type the mean first passage time can be expressed analytically [17]:

$$\tau(F) = 2\tau_d \delta(F)^{-2} [e^{\delta(F)} - \delta(F) - 1]. \quad (3)$$

We have introduced here $\tau_d = (b-a)^2/2D$ and $\delta(F) = \beta[\Delta U - F(b-a)]$.

Assuming a width of 3 \AA , estimated from the fluctuation of the extension curves in Fig. 5a, a least square fit procedure matching the data points in

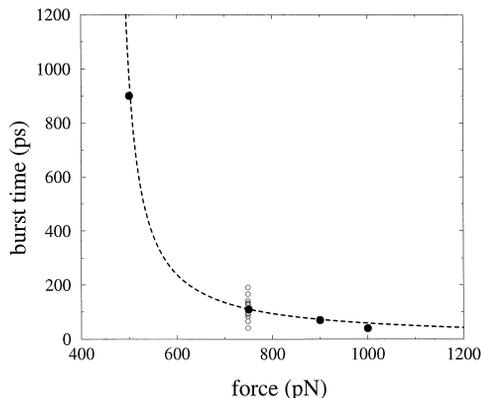


Fig. 6. Time needed for I27 to unfold versus stationary stretching forces. The solid dots represent simulation results and the dotted line is the least square fit of the mean first passage time described by Eq. (3). The little circles at the force of 750 pN demonstrate the first passage time is widely distributed. The actual distribution of passage time from eighteen SMD simulations is presented in Fig. 7.

Fig. 6 to the expression (3) results in the satisfactory match shown also in Fig. 6. This curve corresponds to a barrier height of 1420 pN\AA , i.e., 20.3 kcal/mol .

This analysis of Fig. 6 neglects the possibility that the actual passage (unfolding) times observed in SMD simulations exhibit fluctuations around a mean. It is desirable to estimate the distribution of passage times corresponding to the model potential (2) modified by an applied homogeneous force F . The probability of finding the system at an extension x at time t when it had an extension a at time t_0 , $p(x,t|a,t_0)$, obeys the Smoluchowski diffusion equation

$$\partial_t p(x,t|a,t_0) = \partial_x D \left[\partial_x - \frac{\delta(F)}{b-a} \right] p(x,t|a,t_0) \quad (4)$$

subject to the boundary conditions

$$\left[\partial_x - \delta(F)/(b-a) \right] p(x,t|a,t_0) = 0 \quad \text{at } x = a; \text{ reflective boundary} \quad (5)$$

$$p(x,t|a,t_0) = 0 \quad \text{at } x = b; \text{ absorptive boundary} \quad (6)$$

and initial condition $p(x,t_0|a,t_0) = \delta(x-a)$. The constant $\delta(F)/(b-a)$ describes the drift of the system following the combination of a force due to the intrinsic potential (2) and the applied force F .

The applied force F has to be large enough to render $\delta(F)$ negative or otherwise the unfolding would be a thermally activated process which would not complete during the extremely short simulation time. This implies that the force pulls the system away from the boundary at $x = a$ and one may assume, therefore, a boundary at $x \rightarrow -\infty$. For this choice Eqs. (4) and (6) have the solution [32]

$$p(x, t|a, t_0) = p_1(x, t|a, t_0) - p_2(x, t|a, t_0) \quad (7)$$

$$p_1(x, t|a, t_0) = \frac{1}{\sqrt{4\pi Dt}} \exp\left[-\frac{1}{4Dt} \left(x + a + \frac{\delta(F)Dt}{b-a}\right)^2\right], \quad (8)$$

$$p_2(x, t|a, t_0) = \frac{1}{\sqrt{4\pi Dt}} \exp\left[\frac{\delta(F)a}{b-a} - \frac{1}{4Dt} \left(x - a + \frac{\delta(F)Dt}{b-a}\right)^2\right]. \quad (9)$$

The fraction of domains not yet unfolded at time t , $N(t)$, corresponds to the fraction of systems that have not reached the boundary $x = b$ at time t and is

$$N(t) = \frac{1}{2} \operatorname{erfc}\left[\frac{a + \delta(F)Dt/(b-a)}{\sqrt{4Dt}}\right] - \frac{1}{2} \exp\left[\frac{-\delta(F)a}{b-a}\right] \times \operatorname{erfc}\left[\frac{-a + \delta(F)Dt/(b-a)}{\sqrt{4Dt}}\right]. \quad (10)$$

In Fig. 7b we present the rate $-\dot{N}(t)$ corresponding to the expression (10) with a and b chosen as -3 \AA and 0 \AA , respectively; the force F was set to 750 pN . One can recognize in Fig. 7b that the rate is rather narrowly peaked around the value 80 ps , but that a significant spread of passage times is predicted. We have carried out, therefore, 18 trajectories for an external force of 750 pN for which passage (unfolding) times of about 100 ps are expected. The required simulations were computationally feasible,

amounting to a total simulation time of 6 ns . The resulting passage times are shown in Fig. 7a, exhibiting a clearly the distribution close to the predicted one described by Eq. (10).

The fraction of unfolded domains, $N(t)$, assumes an initial value one and has initially a vanishing time derivative since domains require a non-zero time to reach the boundary at b . Accordingly, the simplest functional form that can represent $N(t)$ is the double exponential [26]

$$N(t) = \frac{[t_1 \exp(-t/t_1) - t_2 \exp(-t/t_2)]}{(t_1 - t_2)}, \quad (11)$$

where amplitudes $t_1/(t_1 - t_2)$ and $-t_2/(t_1 - t_2)$ have been chosen to match the initial values $N(0)$ and $\dot{N}(0)$. The relaxation times t_1 and t_2 can be determined through conditions derived by means of the theory of first passage times [26]

$$t_1 + t_2 = \tau_1(a, b), \quad (12)$$

$$t_1^2 + t_1 t_2 + t_2^2 = \tau_2(a, b). \quad (13)$$

The τ_1 and τ_2 have been derived in Schulten et. al [26].:

$$\tau_1(a, b) = \int_a^b dx [D(x) p_{\text{eq}}(x)]^{-1} \int_{-\infty}^x dy p_{\text{eq}}(y), \quad (14)$$

$$\tau_2(a, b) = \int_a^b dx [D(x) p_{\text{eq}}(x)]^{-1} \times \int_{-\infty}^x dy p_{\text{eq}}(y) \tau_1(y, b). \quad (15)$$

Here we have defined

$$p_{\text{eq}} = \exp[-\beta U(x)] / \int_{-\infty}^b dx \exp[-\beta U(x)]. \quad (16)$$

Integrating the above equations using the potential described in Eq. (2) and the numerical values of $\delta(F)$ and D resulting from a match of expression (3) to the simulation data (c.f. Fig. Fig. 6) yields for t_1 and t_2 the values 238 ps and 54 ps , respectively. The resulting $-\dot{N}(t)$, plotted in Fig. 7b, is found to

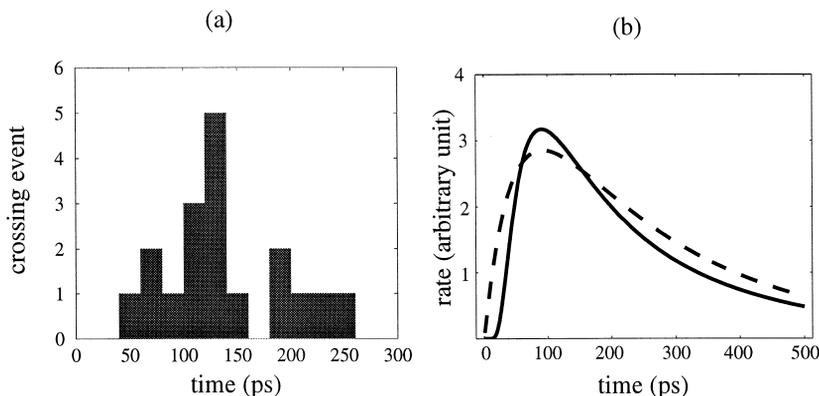


Fig. 7. (a) Barrier crossing time resulting from eighteen SMD simulations with a stationary force of 750 pN. (b) Predicted distributions of crossing (first passage) time. Result from a description using $-N(t)$ with $N(t)$ calculated from Eq. (10) is drawn in solid line, and from Eqs. (11)–(16) dashed line. For $\delta(F)$ and D values of 2 and 0.015, respectively, were chosen assume that correspond to the model of Eq. (3) and simulation data in Fig. 6.

compare favorably with the analytical solution and, hence, also with the simulation

4. Discussion

Mechanical multi-domain proteins like those involving immunoglobulin and fibronectin type III domains constitute a fascinating class of biopolymers with important cellular functions. Single molecule experiments based on AFM that probe the mechanical response of these protein systems provide a unique source of information that becomes truly valuable in combination with steered molecular dynamics simulations. The latter provide atomic level pictures of the conformational processes governing the function of mechanical proteins which, however, need to be verified through comparison with AFM data.

In this paper we have shown first that the main obstacle between simulations and AFM measurements, the time scale discrepancy and the related discrepancy in stretching forces required to induce unfolding, may be overcome through slower pulling speeds in the simulations. The results reported have decreased the force discrepancy significantly, retaining the same scenario of hydrogen bond breaking found in the earlier [20] simulations which used faster pulling speeds. Fig. 8 presents the measured AFM forces and their logarithmic extrapolation to

faster pulling speeds together with the forces from simulations SMD-(1.0 Å/ps), SMD-(0.5 Å/ps) and SMD-(0.1 Å/ps). One can discern that the extrapolated AFM forces correspond to a force of about 500 pN at a pulling speed of 0.1 Å/ps whereas the simulated force is twice as large. This result is not satisfactory yet. We will argue below that results of stationary force SMD simulations can be extrapolated even more closely to the forces observed in AFM.

When the immunoglobulin domain is stretched in AFM experiments by relatively weak forces of about 200 pN, the separation of the domain's termini is halted due to a potential energy barrier connected with a set of interstrand hydrogen bonds as described above. Under the experimental, i.e., weak force, conditions it requires thermal fluctuations and a time period of about a 1 ms to overcome this barrier and to begin domain unfolding. Prior to this, i.e., during the period “stuck” in front of the barrier, the proteins experience over the nanosecond time scale, i.e., the period covered in SMD simulations, essentially a stationary force; the latter can be considered constant because of the relatively weak force constants of the AFM cantilevers. Hence, the SMD simulations carried out with stationary constant external forces and reported in this paper should match rather closely AFM experiments. We note that barrier crossing events during the forced unfolding of Ig domains have also been studied by means of Monte Carlo

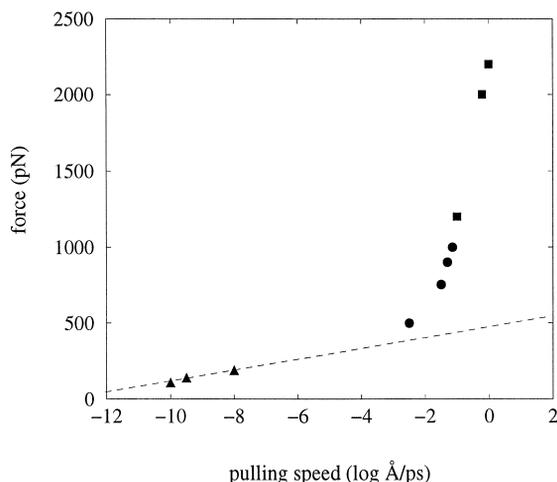


Fig. 8. Force peak value during unfolding versus pulling speed. Triangles represent AFM data and the dotted line is an extrapolation of these data [12]; squares represent results from SMD simulations; circles represent forces estimated (see text) from simulations SMD-(1000 pN), SMD-(900 pN), SMD-(750 pN), and SMD-(500 pN).

simulations [33] and that the results agreed well with AFM data.

We thought to interpret the results of stationary force simulations in terms of a potential energy barrier described by Eq. (2) and have found, in applying the theory of first passage times [24–26], that a barrier of about 3 Å width and 20 kcal/mol height matched the simulation results. These characteristics of the barrier separating the folded and unfolded forms of I27 agree well with AFM observations and chemical denaturation (Chevron plot) data in which a potential barrier of 22 kcal/mol was measured [12].

A further support of our analysis of stationary force SMD simulations can be derived from the distribution of forces measured by AFM experiments on I27 unfolding. This distribution matches closely the distribution predicted in Ref. [17] by means of the theory of mean first passage times for the potential barrier described by Eq. (2). The reason why a rather artificial functional form of the barrier like that given by Eq. (2) describes the actual behavior of stretched Ig so well is likely the fact that the barrier crossing (mean first passage) times $\tau(a \rightarrow b)$ are given by a double quadrature involving only certain

average characteristics of the potential $U(x)$, namely is given by

$$\tau(a \rightarrow b) = \int_a^b dx D^{-1} \exp[U(x)/k_B T] \times \int_{x_{\min}}^x dy \exp[-U(y)/k_B T], \quad (17)$$

where D is an effective diffusion constant and x_{\min} denotes the smallest possible extension of the domain. This expression does not depend sensitively on the detailed shape of the barrier.

One may employ the crossing times τ_{barrier} for simulations SMD-(500 pN), SMD-(750 pN), SMD-(900 pN), SMD-(1000 pN) to estimate pulling velocities and extend the set of simulation data shown in Fig. 8. We assume for this purpose that the forces applied are the peak forces and that the barrier width of 3 Å provides a suitable estimate for the distance traveled. In case of simulation SMD-(1000 pN), i.e., for a stationary force of 1000 pN, one determines accordingly a velocity of 0.075 Å/ps. The corresponding data point in Fig. 8 lies indeed close to the data point of simulation SMD-(0.1 Å/ps) which exhibited a peak force of 1200 pN. The velocities corresponding to the stationary forces 900 pN, 750 pN, and 500 pN are 0.05 Å/ps, 0.027 Å/ps, and 0.003 Å/ps, respectively, which contribute the corresponding data points in Fig. 8. One can discern from these data points that according to our rough estimate the simulation data approach the measured AFM peak forces when linearly scaled with velocity.

Our simulations have shown that the response of immunoglobulin domains to stretching is closely connected to the double β sheet (β sandwich fold of this protein family. There exist three types of responses of I27 to stretching determined by this fold: (i) Weak forces above 50 pN suffice to break two interstrand (A and B) hydrogen bonds and prestretch the domain by about 6 Å[23]. (ii) Stronger forces succeed to break the set of six hydrogen bonds between strands A' and G and, thereby, help the domain to overcome the barrier separating the folded from the unfolded state. (iii) After this barrier is crossed only weak forces are necessary to break the remaining interstrand hydrogen bonds in a zipper-like fashion, i.e., break them “one-by-one”. Proteins with folds not designed for protection against

strain, in particular, stretching of the terminal ends, do not exhibit the second type of response as suggested by the SMD simulations reported in Ref. [22].

Both the first and the second response can be functionally important as has been shown in Ref. [21] for fibronectin type III domains. These proteins are composed similarly to I27 of seven antiparallel β strands (A–G) that are arranged in two sheets. One of these domains, FnIII₁₀, mediates in fibronectin cell adhesion via its integrin-binding RGD peptide sequence located at the apex of the loop between the two C-terminal β strands F and G. The SMD simulations in Ref. [21] revealed that the first response of the domain to forces is an extraction of β strand G while the remaining fold maintains its structural integrity. The separation of β strand G leads to drastic conformational changes of the RGD-loop that results in a reduced accessibility and affinity of the module to cell surface-bound integrins. Further pulling of the module leads to the release of the N-terminal β strand A that initiates then unfolding of the domain. The conformational changes of the RGD-loop in the first response of the module appears to constitute a mechanosensitive control of ligand recognition.

Increase in computer power will soon permit simulations in closer agreement with experimental time scales than achieved so far and, thereby, further improve the combination of single molecule techniques and SMD simulations suggested in this paper. The method of reconstructing a potential energy barrier from SMD simulation data adopted in this paper can also be improved through new algorithms. A promising avenue in this respect has been suggested and tested in Ref. [34] involving a reconstruction of potentials of mean force along the direction of pulling in SMD simulations on the basis of the path integral formulation of Brownian dynamics.

The combination of single molecule experiments and MD simulations can be applied to further biomolecular processes. An example is antibody - antigen recognition that is also amenable to AFM measurements [35].

Acknowledgements

This work was supported by the National Institutes of Health (NIH PHS 5 P41 RR05969) and the

National Science Foundation (NSF BIR 94-23827EQ, NSF/GCAG BIR 93-18159, MCA93S028).

References

- [1] M. Rief, M. Gautel, A. Schemmel, H. Gaub, *Biophys. J.* 75 (1998) 3008.
- [2] M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, H.E. Gaub, *Science* 276 (1997) 1109.
- [3] A.F. Oberhauser, P.E. Marszalek, H. Erickson, J. Fernandez, *Nature* 393 (1998) 181.
- [4] M. Kellermayer, S. Smith, H. Granzier, C. Bustamante, *Science* 276 (1997) 1112.
- [5] L. Tskhovrebova, J. Trinick, J.A. Sleep, R.M. Simmons, *Nature* 387 (1997) 308.
- [6] K. Wang, R. McCarter, J. Wright, J. Beverly, R. Ramirez-Mitchell, *Biophys. J.* 64 (1993) 1161.
- [7] K. Maruyama, *FASEB J.* 11 (1997) 341.
- [8] C. Machado, C.E. Sunkel, D.J. Andrew, *J. Cell Biol.* 141 (1998) 321.
- [9] S. Labeit, B. Kolmerer, *Science* 270 (1995) 293.
- [10] H.P. Erickson, *Curr. Opin. Cell Biol.* 5 (1993) 869.
- [11] M. Rief, J. Pascual, M. Saraste, H. Gaub, *J. Mol. Biol.* 286 (1999) 553.
- [12] M. Carrion-Vazquez, A. Oberhauser, S. Fowler, P. Marszalek, S. Broedel, J. Clarke, J. Fernandez, Mechanical and chemical unfolding of a single protein: A comparison, *Proc. Natl. Acad. Sci. USA*, 1999, p. 3694.
- [13] S. Improt, A. Politou, A. Pastore, *Structure* 4 (1996) 323.
- [14] N. Socci, J. Onuchic, P. Wolynes, *Proc. Natl. Acad. Sci. USA* 96 (1999) 2031.
- [15] S. Izrailev, S. Stepaniants, B. Isralewitz, D. Kosztin, Hui Lu, F. Molnar, W. Wriggers, K. Schulten, Steered molecular dynamics, in: P. Deuffhard, J. Hermans, B. Leimkuhler, A.E. Mark, S. Reich, R.D. Skeel (Eds.), *Computational Molecular Dynamics: Challenges, Methods, Ideas*, vol. 4 of *Lecture Notes in Computational Science and Engineering*, Springer, Berlin, 1998, p. 36.
- [16] H. Grubmüller, B. Heymann, P. Tavan, *Science* 271 (1996) 997.
- [17] S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, K. Schulten, *Biophys. J.* 72 (1997) 1568.
- [18] B. Isralewitz, S. Izrailev, K. Schulten, *Biophys. J.* 73 (1997) 2972.
- [19] D. Kosztin, S. Izrailev, K. Schulten, *Biophys. J.* 76 (1999) 188.
- [20] Hui Lu, B. Isralewitz, A. Krammer, V. Vogel, K. Schulten, *Biophys. J.* 75 (1998) 662.
- [21] A. Krammer, Hui Lu, B. Isralewitz, K. Schulten, V. Vogel, Forced unfolding of the fibronectin type III module reveals a tensile molecular recognition switch, *Proc. Natl. Acad. Sci. USA*, vol. 91, 1999, p. 1351.
- [22] Hui Lu, K. Schulten, Steered molecular dynamics simulations of force-induced protein domain unfolding, *PROTEINS: Struc., Func., and Genetics*, 1999, in press.

- [23] J. Fernandez et al., 1999, submitted.
- [24] K. Schulten, Z. Schulten, A. Szabo, *Physica* 100A (1980) 599.
- [25] A. Szabo, K. Schulten, Z. Schulten, *J. Chem. Phys.* 72 (1980) 4350.
- [26] K. Schulten, Z. Schulten, A. Szabo, *J. Chem. Phys.* 74 (1981) 4426.
- [27] A.S. Politou, D. Thomas, A. Pastore, *Biophys. J.* 69 (1995) 2601.
- [28] A.D. MacKerell Jr., 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, M. Karplus, *J. Phys. Chem. B* 102 (1998) 3586.
- [29] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, *J. Chem. Phys.* 79 (1983) 926.
- [30] M. Nelson, W. Humphrey, A. Gursoy, A. Dalke, L. Kalé, R.D. Skeel, K. Schulten, *J. Supercomputing App.* 10 (1996) 251.
- [31] A.T. Brünger, 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, 1992.
- [32] G. Lamm, K. Schulten, *J. Chem. Phys.* 78 (1983) 2713.
- [33] M. Rief, J. Fernandez, H. Gaub, *Phys. Rev. Lett.* 81 (1998) 4764.
- [34] J. Gullingsrud, R. Braun, K. Schulten, Reconstructing potentials of mean force through time series analysis of steered molecular dynamics simulations, *J. Comp. Phys.*, Special Issue on Computational Biophysics, 1998, in press.
- [35] O.H. Willemsen, M.M.E. Snel, K.O. van der Werf, B.G. de Groot, J. Greve, P. Hinterdorfer, H.J. Gruber, H. Schindler, Y. van Kooyk, C.G. Figdor, *Biophys. J.* 75 (1998) 2220.
- [36] W.F. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* 14 (1996) 33.