Chapter 1

COMPUTER MODELING OF FORCE-INDUCED TITIN DOMAIN UNFOLDING

Hui Lu[1], André Krammer[2], Barry Isralewitz[1],
Viola Vogel[2] and Klaus Schulten[1]

[1] *Beckman Institute for Advanced Science and Technology*
University of Illinois at Urbana-Champaign,
Urbana, Illinois 61801

[2] *Departments of Physics and Bioengineering*
University of Washington,
Seattle, Washington 98195

Abstract

Titin, a 1 μm long protein found in striated muscle myofibrils, possesses unique elastic and extensibility properties, and is largely composed of a PEVK region and β-sandwich immunoglobulin (Ig) and fibronectin type III (FnIII) domains. The extensibility behavior of titin has been shown in atomic force microscope and optical tweezer experiments to partially depend on the reversible unfolding of individual Ig and FnIII domains. We performed steered molecular dynamics simulations to stretch single titin Ig domains in solution with pulling speeds of 0.1 – 1.0 Å/ps, and FnIII domains with a pulling speed of 0.5 Å/ps. Resulting force-extension profiles exhibit a single dominant peak for each domain unfolding, consistent with the experimentally observed sequential, as opposed to concerted, unfolding of Ig and FnIII domains under external stretching forces. The force peaks can be attributed to an initial burst of a set of backbone hydrogen bonds connected to the domains' terminal β-strands. Constant force stretching simulations, applying 500 – 1000 pN of force, were performed on Ig domains. The resulting domain extensions are halted at an initial extension of 10 Å until the set of all six hydrogen bonds connecting terminal β-strands 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.

Introduction

The giant muscle protein titin, also known as connectin, is a roughly 30,000 amino acid long filament which spans half of the sarcomere and plays a number of important roles in muscle contraction and elasticity [1, 2, 3, 4], as well as controlling chromosome shape in the cell nucleus [5]. During muscle contraction, titin, which is anchored at the Z-disk and at the M-line, exerts a passive force which 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 about 300 repeats of two types of domains, immunoglobulin (Ig) domains and fibronectin type III (FnIII) domains, and the PEVK (70% proline, glutamic acid, valine and lysine residue) region [6]. The FnIII domains are located only in the A-band of the molecule, the PEVK region is located in the I-band, while 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 extensibility and passive elasticity [7, 8, 9, 10]. 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 due to the charges on its glutamic acid and lysine residues. When muscle is stretched, the PEVK region unfolds and enlongates. Under extreme conditions such as in overstretched muscle, the Ig domains in the titin I-band will unfold to provide the necessary extension. At further extension, the A-band of titin will unbind from myosin and contribute additional Ig and FnIII domains to the available pool of extendible domains. When forces are released, the unfolded Ig domains refold quickly [11].

Titin domains have been observed, using single molecule techniques, atomic force microscopy (AFM) [11] and optical tweezer experiments [12, 13], to possess protection against strain-induced domain unfolding. The AFM experiments in particular have demonstrated that rather strong forces, on the order of 100 pN, must be exerted before Ig and FnIII domains rupture and unfold. We refer to proteins which are designed to respond to force application under physiological conditions as *mechanical proteins*. Other proteins, which do not encounter mechanical strain under physiological conditions, have been found to exhibit little resistance against strain-induced unfolding, as has been demonstrated through AFM experiments on the helical protein spectrin [14].

The AFM force-extension profiles of multimers of Ig domains (from titin) and FnIII (from titin, and also from the protein tenascin) display a regularly repeating sawtooth pattern [11, 15, 16]. The spacing between any two 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. 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 μm/s, AFM unfolding of titin Ig domains requires about 200 pN while unfolding of tenascin FnIII domains requires only about 140 pN [11, 16].

We would like to explain one-by-one domain unfolding in titin 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) [17]. Since no experimental structures are available for titin FnIII domains, we use a homologous FnIII domain, FnIII_{10} of fibronectin[18]. Figure 1.1 demonstrates that I27 and FnIII_{10} both form β sandwiches. AFM experiments studying Ig and FnIII stretching do not resolve atomic-level detail of the domains' conformational changes during the unfolding. In this paper we summarize the results of molecular dynamics simulations we have presented elsewhere [19, 20, 21, 22] to provide an overall view of how SMD complements titin AFM observations, and to produce a detailed picture of titin domain stretching and unfolding.

Molecular dynamics (MD) simulations describe molecules as a set of atoms with known coordinates, masses, charges, and bond types [24]. Motion is simulated by advancing through a series of time steps, producing atomic-level detail of molecular movements. At each time step, the forces on every atom from all other atoms due to Coulombic, van der Waals and bonded interactions are calculated. The positions of each atom are updated for their resulting movement during one time step, then the forces are calculated again. Since the fastest oscillations in the system have a period of about 10^{-14} s; a time step of about 1/10 this size, 10^{-15} s (1 femtosecond), is often used. The CHARMM energy function [25] is used in this work to calculate forces on atoms; it has the form

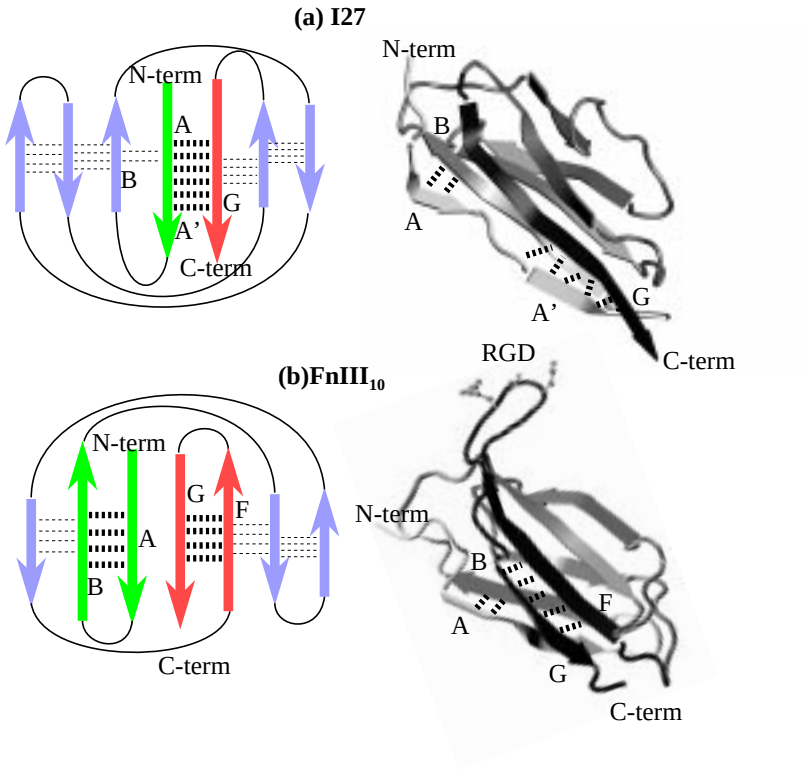

Figure 1.1 Topology schematics and structures. The schematic representations on the left depict the topology of interstrand hydrogen bonding networks of (a) I27 and (b) FnIII_{10}. On the right are shown the corresponding proteins domains rendered with VMD [23] with the essential interstrand hydrogen bonds highlighted as in the schematic representations.

$$V = \underbrace{\sum_{bonds} \frac{1}{2} k_b (r - r_0)^2}_{V_{bond}} + \underbrace{\sum_{angles} \frac{1}{2} k_\theta (\theta - \theta_0)^2}_{V_{angle}} + \underbrace{\sum_{dihedrals} k_\phi (1 + \cos(n_\phi \phi + \delta_\phi))}_{V_{dihedral}} + \underbrace{\sum_{i \neq j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right]}_{V_{van-der-Waals}} + \underbrace{\sum_{i \neq j} \frac{q_i q_j}{\epsilon r_{ij}}}_{V_{Coulomb}}.$$

V_{bond} describes the oscillations about the equilibrium bond lengths. V_{angle} and $V_{dihedral}$ describe the other bonded interactions, oscillations of three atoms about an equilibrium angle and torsional rotations of four atoms about a central bond, respectively. The partial charges q_i and the energy parameters (e.g. k_b, σ_{ij}) are determined experimentally or by quantum chemical calculations. A set of these parameters is referred to as a *force field*.

Although many biological events take place on the millisecond to second time scale, the femtosecond (fs) time step taken in MD, combined with limits in available computational power, limit the time scales of simulations to the nanosecond (10^{-9} s = ns) range for systems with 10,000 to 100,000 atoms, such as the systems described in the paper.

Steered Molecular Dynamics (SMD), reviewed in [26], is a novel approach to the study of the dynamics of binding/unbinding events in bimolecular systems and of their elastic properties. The simulations reveal the details of molecular interactions in the course of ligand unbinding [27, 28, 29, 30, 31, 32] or protein unfolding [19, 20, 21, 22], thereby providing important information about the molecular mechanisms underlying these processes. The advantage of SMD over conventional molecular dynamics is the ability to induce relatively large conformational changes in molecules on the ns time scales accessible to computation.

SMD simulations add external forces to conventional force fields, with the external forces imitating, for example, the effects of AFM cantilevers on protein domains. SMD simulations are ideally suited for the purpose of studying processes in which force is applied to molecules in actual systems, such as in force-induced protein stretching. We will show how SMD stretching trajectories, which are limited to the ns time scale, can provide insights into experimentally observed ms time scale stretching events. The problem of bridging the time scale gap has been discussed in detail in [33].

Methods

In this work, we apply forces to systems with two SMD protocols: constant-velocity moving restraints and constant force restraints.

SMD using constant-velocity moving restraints simulates perturbing a protein with a moving AFM cantilever. At least one atom is held fixed during the simulation, which prevents the protein from simply translating in space when external forces are applied, and which corresponds to attaching the protein to a fixed substrate in the AFM experiments. At least one other atom of the protein is restrained to a point in space (restraint point) by an external, e.g., harmonic, potential. The restraint point is then shifted in a chosen direction [34, 28, 30, 35, 31, 32], forcing the restrained atom to move from its initial position in the protein. Assuming a single reaction coordinate x, and an external potential $U = K(x - x_0 - vt)^2/2$, where K is the stiffness of the restraint, and x_0 is the initial position of the restraint point moving with a constant velocity v, the external force exerted on the system can be expressed as

$$F = K(x_0 + vt - x)\,. \tag{1.2}$$

This force corresponds to a molecule being pulled by a harmonic spring of stiffness K with one end attached to the restrained atom and the other end moving with velocity v. The force applied in constant force SMD, however, retains the same magnitude and direction regardless of the position of the restrained atom.

The energy-minimized average NMR structure of I27 [17], the 27th immunoglobulin domain of the I-band of cardiac titin, was obtained from the Brookhaven protein data bank entry 1TIT, was employed in this modeling study. This domain adopts the typical I-frame immunoglobulin superfamily fold [36], consisting of two β sheets packing against each other, as shown in Figure 1.1, 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.

To simulate a solvent environment, the I27 structure 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 [19] consisting of 11,400 atoms.

The MD simulations were performed using the CHARMM22 force field [37] with the programs NAMD [38] and XPLOR [39], assuming an integration time step of 1 fs and a uniform dielectric constant of 1. 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 [40].

SMD simulations with constant velocity stretching were carried out on I27 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 [11, 16], except that the pulling speeds adopted in the simulations are about six orders of magnitude higher than those in the experiment. The value of K (see Eq. 1.2) was set to 10 $k_B T/\text{Å}^2$, corresponding to spatial (thermal) fluctuations of the constrained C_α atom with variance $\delta x = \sqrt{k_B T/k} = 0.32$ Å at $T = 300\ K$.

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 the N-terminus to the C-terminus.

The atomic coordinates of the entire 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 is 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 I27-(1.0 Å/ps), I27-(0.5 Å/ps) and I27-(0.1 Å/ps). Four force values have been adopted for stretching I27 under constant force conditions: 500, 750, 900, and 1000 pN; the respective simulations will be denoted by I27-(500 pN), ..., I27-(1000 pN).

The system setup, solvation, equilibration, SMD simulations, and data recording performed for FnIII domains [20] are similar to those performed for I27. The X-ray crystallographic structure of $FnIII_{10}$ was obtained from the tetramer $FnIII_{7-10}$ (PDB code 1fnf) [18]. All atoms, including hydrogens, were described explicitly. In the simulations, the N-terminal C_α atom (Leu^1) of $FnIII_{10}$ was constrained to a fixed point, while the C-terminal C_α atom (Thr^{94}) was attached to a constant velocity (0.5 Å/ps) moving SMD restraint.

Results

The force profiles for the I27 extensions are depicted in Figure 1.2a, which compare the simulations I27-(1.0 Å/ps), I27-(0.5 Å/ps) and I27-(0.1 Å/ps), as presented already in [22]. The three profiles are qualitatively similar: at an extension of about 15 Å a dominant force peak arises that, in each individual profile, is about 2 to 3 times higher than forces at other extensions. At extensions larger than 20 Å the domain exhibits little resistance against stretching as 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., at about 270 Å, the force required to stretch the now completely unfolded domain increases again as expected. The force-extension profiles in Figure 1.2 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. All the simulated force-extension curves share the main qualitative feature with those observed in AFM stretching experiments: a single dominant force peak.

An analysis of the trajectories corresponding to the profiles in Figure 1.2 reveals that at the extension of the maximum force eight hydrogen bonds break concurrently, as pointed out in [19], as well as in [22]. Figure 1.3 presents snapshots of the domain directly before and after the force peak in which the extensions measure 10 Å and 17 Å, during which 8 hydrogen bonds break concurrently: the 2 hydrogen bonds between strands A and B break during 10–14 Å of extension, and the 6 hydrogen bonds between strands A' and G break during the 15–17 Å extension. During the following 190 Å of extension, the remaining 22 intra-β-strand hydrogen bonds break individually.

Simulations stretching I27 with constant force characterize the unfolding barrier responsible for the force peak described above. Simulations I27-(500 pN), I27-(750 pN), I27-(900 pN), and I27-(1000 pN) explore the influence of the stretching forces on the time between the start of

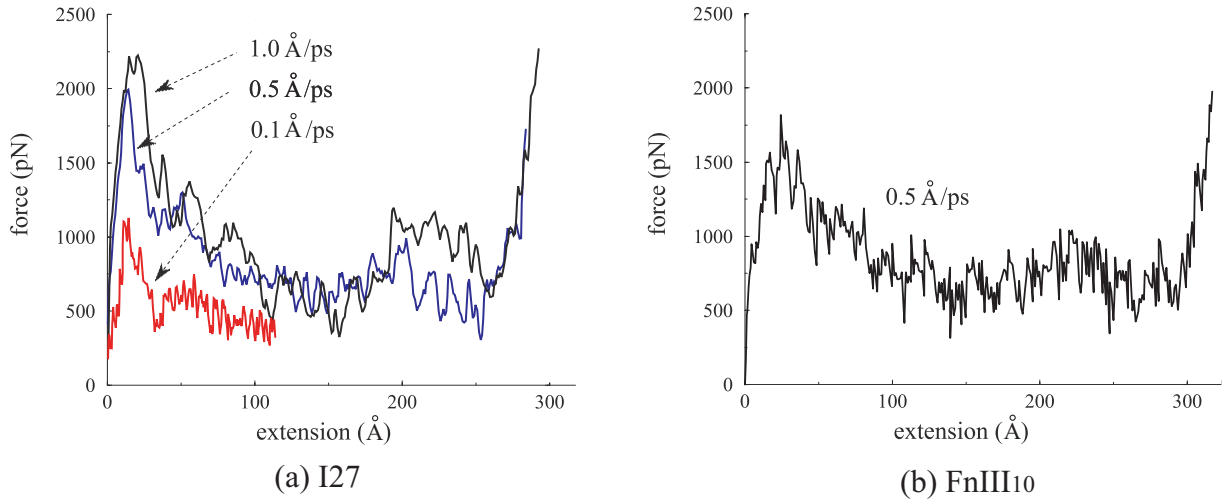

Figure 1.2 Force extension profiles of I27 and $FnIII_{10}$. (a) Simulation results of I27-(1.0 Å/ps), I27-(0.5 Å/ps) and I27-(0.1 Å/ps). The simulation times for pulling velocities 1.0 Å/ps, 0.5 Å/ps, and 0.1 Å/ps are 0.3 ns, 0.6 ns and 1.2 ns, respectively. The simulation of I27-(0.1 Å/ps) was stopped at an extension of 120 Å, whereas simulations I27-(1.0 Å/ps) and I27-(0.5 Å/ps) were stopped at about 300 Å. (b) Simulation results of $FnIII_{10}$ extension. The simulation time is 0.3 ns, with a velocity of 0.5 Å/ps.

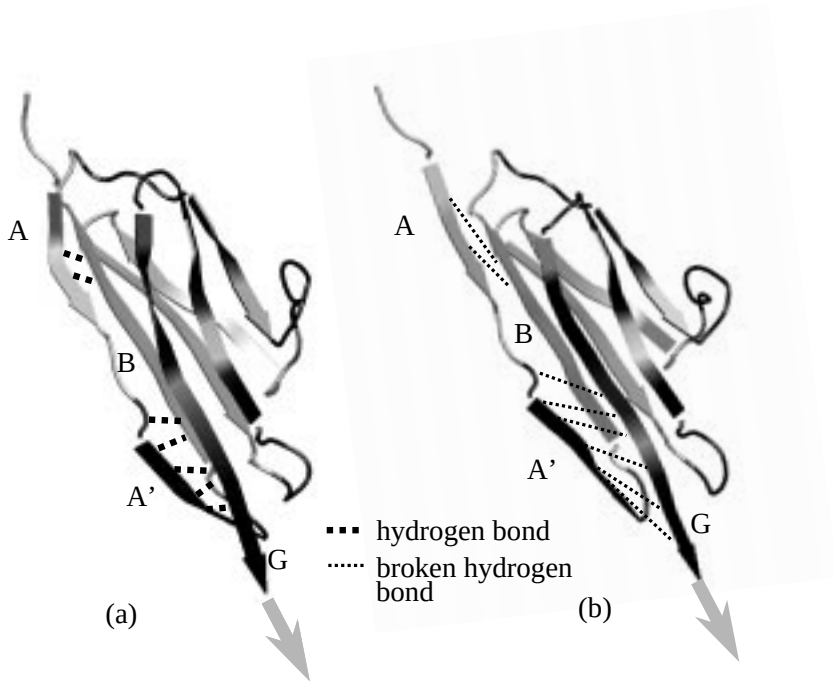

Figure 1.3 Two snapshots of I27 unfolding. The domain is drawn in cartoon representation and key hydrogen bonds between strands A-B, and between strands A'-G are shown as dotted lines. (a) Snapshot of I27 before the force peak at 10 Å extension. The hydrogen bonds are all maintained. (b) Snapshot of 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.

force application and the beginning of domain unfolding. An analysis of the four simulations revealed that an initial 10 Å extension takes place which is due to a straightening of the polypeptide chain near its terminal ends and due to a breaking of the two hydrogen bonds between β strands A and B. This extension, shown in Figure 1.4, and presented originally in [22], continues until a plateau region (which is longer for weaker forces) is reached, at which point the six A'– G interstrand hydrogen bonds come under mechanical strain. The domain extensions fluctuate around a constant value during the plateau period, until the A'–G bonds are broken, after which extension rapidly increases. Snapshots of I27 structure at extensions smaller and greater than the plateau (dotted lines in figure 1.4) reveal the breaking event as described; the structural changes are roughly indentical to those observed before and after the force peak seen in constant velocity stretching (figure 1.3). During the rapid post-plateau extension increase, the domain unravels, exhibiting relatively little resistance to extension.

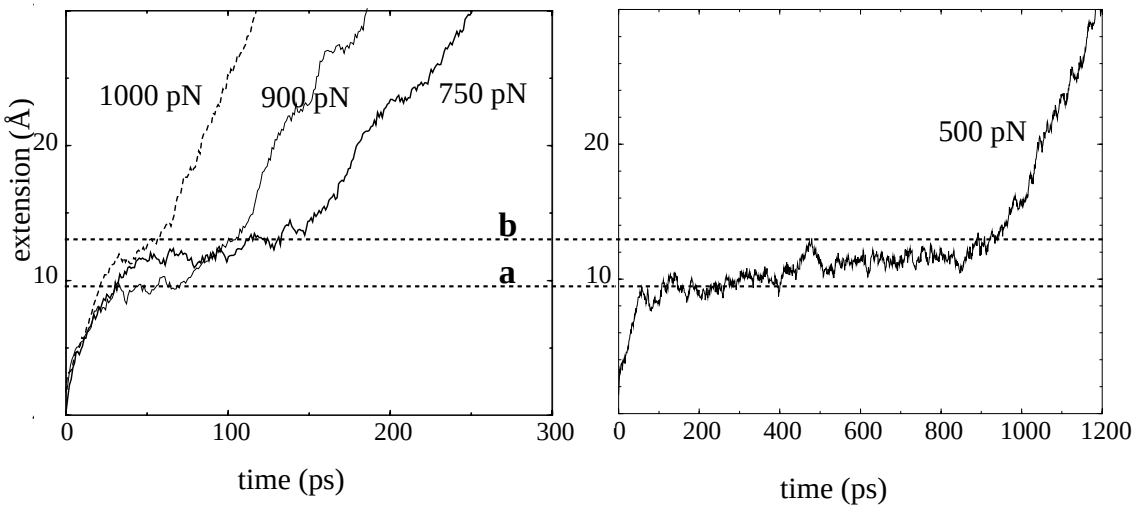

Figure 1.4 Extension curves of I27 unfolding by means of constant forces of 500, 750, 900 and 1000 pN strength. In all cases the domain extended quickly by 10 Å, then remained approximately at constant length for long periods of time until the system unfolded (appearing as linear extension with time). Conformations at dotted lines a and b correspond to conformations shown in figure 1.3a and 1.3b respectively (see text).

We can model crossing the plateau as a barrier crossing process, with $\tau_{barrier}$ denoting the time spent at the plateau. The applied force effectively lowers the barrier such that stronger forces lead to faster barrier crossing ($\tau_{barrier}$ = 0.04 ns for 1000 pN) than weaker forces ($\tau_{barrier}$ = 0.9 ns for 500 pN). In all cases, the motion gets temporarily "stuck" in front of the barrier, which can then only be overcome by thermal fluctuations. 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. The mean time to cross the barrier can be evaluated using the expressions for the mean first passage time developed in [41, 42, 43]. By comparing the mean first passage times with the respective times $\tau_{barrier}$ for various forces one can estimate the height of the indigenous potential barrier.

To estimate the shape of the indigenous potential barrier we assume a simple model for the energy $U(x)$ of the barrier: $U(x) = \Delta U\,(x - a)/(b - a)$ for $a \leq x \leq b$, with a reflective boundary condition at a, and an absorptive boundary condition at b. 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 [27]:

$$\tau(F) = 2\tau_d \delta(F)^{-2}[e^{\delta(F)} - \delta(F) - 1] \tag{1.3}$$

We have introduced here $\tau_d = (b-a)^2/2D$ and $\delta(F) = \beta[\Delta U - F(b-a)]$. Assuming a width of 3 Å, estimated from the fluctuation of the extension curves in Figure 1.4, a least square fit procedure matching the data points in Figure 1.5 to the expression (1.3) results in the satisfactory match shown also in Figure 1.5. This curve corresponds to a barrier height of 1420 pNÅ, i.e., 20.3 kcal/mol.

In the constant velocity moving restraint SMD performed for $FnIII_{10}$, the domain was stretched from its initially compact and folded structure to its fully elongated configuration. The force-extension profile (Figure 1.2b) exhibits a single dominant force peak, following an unfolding process similar to that described above for Ig domain. This 1800 pN force peak, arising at 25 Å extension, implies again a barrier which must be overcome to initiate unfolding. Simulations presented in [20] revealed that, just as for the Ig domain, the respective barrier is caused by the need to simultaneously break several hydrogen bonds early in unfolding. In this case, the force peak is caused by 6 hydrogen bonds breaking during 8 Å of extension, a range longer than the 2 Å extension period required to break the same number of hydrogen bonds at the Ig force peak.

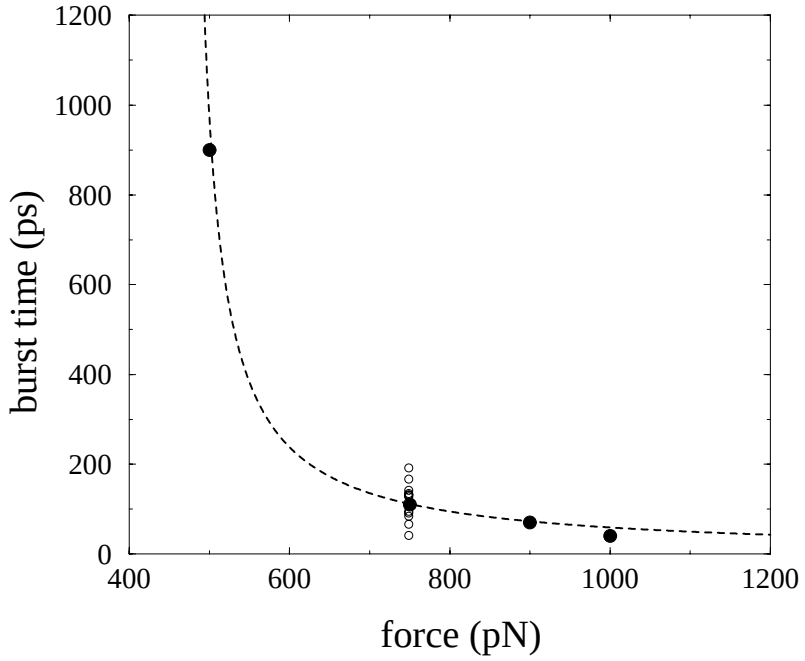

Figure 1.5 Time needed for I27 to unfold vs. constant stretching forces. The solid dots represent simulation results and the dotted line is the least square fit of the mean first passage time described by Eq. (1.3). The small circles at the force of 750 pN demonstrate that the first passage time is widely distributed.

In fibronectin, the $FnIII_{10}$ domain mediates cell adhesion to surfaces via its integrin binding motif, Arg^{78}, Gly^{79}, and Asp^{80} (RGD), which is placed at the apex of the loop connecting β-strands F and G. Fibronectin and other RGD-containing proteins serve as static anchoring sites against which cells can build up tensile strength.

In the above simulations, dramatic conformational changes of the RGD-loop are observed in the initial stages of the forced unfolding pathway of the module. The loop is initially located at the apex of a hairpin β-turn connecting strands G and F. In the unperturbed native state the RGD-containing loop is located at ~ 12 Å away from the outer surface of the $FnIII_{10}$ module and is thus accessible to cell surface integrins. Optimal binding is obtained for distances between RGD and substrate in the range 11 Å – 30 Å [44]. The RGD hairpin is bent at the C_α atoms of residues Thr^{76} and Ala^{83}, but upon domain extension past the force peak, this bend is straightened. A steep increase in the angle defined by the C_α atoms of residues Asp^{80}, Ala^{83}, and Ser^{84} is observed at 20 Å extension , followed by a movement of the apex of the loop closer to the surface of the domain from 12 Å to 8 Å while the module is extended from 25 Å to 35 Å. The RGD loop is distorted further as the extensions continue. During the first 35 Å of extension of $FnIII_{10}$ we are likely seeing a force-regulated decrease in integrin accessibility of the RGD segment.

Discussion

Multi-domain proteins like titin 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 more 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.

Applied force experiments [11, 12, 13] 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 which still remain folded. 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 about 40 Å. Upon SMD force application, every Ig domain exhibits a pre-burst increase in

contour length of roughly 10 Å per Ig domain. With further extension, Ig domains continue to lengthen, but only after the force exceeds a given value. Our simulations provide an explanation for this bursting behavior. The links which 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. Due to the topology of the Ig domain (Figure 1.1), until the A'G bonds break, force can not be transmitted along the backbone to unravel the rest of the protein. Only when the A'G and AB strands are separated, after breaking of all inter-strand hydrogen bonds, can the unfolding of an Ig domain continue, involving rupture of the inter-strand hydrogen bonds between the remaining β-strands.

The SMD stretching of FnIII proceeds similarly to the Ig stretching. The chief difference between Ig and FnIII domain SMD unfolding is the smaller rupture forces observed for FnIII as compared to Ig; the smaller forces are consistent with observations in three AFM experiments [16, 15, 11]. The AFM force peaks recorded for FnIII unfolding are 70% the size of those recorded for Ig domain unfolding. From analysis of structures and SMD trajectories [20] one can discern structural features, specifically differences in the topology of interstrand hydrogen bonding, likely responsible for the differences in force peak. In the case of Ig (Figure 1.1a) there exist backbone hydrogen bonds directly between N-terminal strand A or A' and C-terminal strand G. In FnIII (Figure 1.1b) backbone hydrogen bonds do not directly connect the termini, they connect the N-terminal strand A with its neighboring strand B, and the C-terminal strand G with its neighboring strand F. In Ig, the breaking of interstrand hydrogen bonds occurs over a very small extension range (roughly the length of a hydrogen bond), whereas in FnIII, the breaking of the interstrand bonds occurs over a wider extension range. In FnIII the terminal segments where force is applied are not directly connected by hydrogen bonds; the FnIII topology of interstrand bonding renders the domain more flexible under stretching, allowing bond breaking over a longer extension range and hence with a reduced rupture force, as seen in Figure 1.2.

Both Ig and FnIII domains exhibit high ($>$ 150 pN) dominant force peaks when stretched in AFM experiments, dominant force peaks are also seen for both domains when stretched with SMD. The domains are both members of a class of proteins, which we call Class I [21], that exhibit resistance to forced unfolding; their fold topologies are such that interstrand hydrogen bonds must break in clusters in order to allow extension of the domain. Other domains (e.g. all-helix domains) have topologies that can be extended while breaking interstrand hydrogen

bonds singly, these do not exhibit dominant force peaks when stretched in SMD simulations[21]; we call these class II domains.

The current study suggests that AFM experiments on proteins that consist of several class II domains should show force-extension profiles different from those performed on proteins consisting of several class I domains. Either no dominant force peak (and thus no sawtooth pattern) or much lower force peak values than those observed for class I domains should be observed. The discernible-but-lower force peak values may arise from hydrophobic effects, which do not play an important role in force-induced unfolding of class I domains. Both possibilities, i.e., non-discernible (J. Fernandez, personal communication) and lower, discernible [14] peaks have already been confirmed by recent AFM experiments.

The main obstacles to relating MD simulations and AFM measurements, namely the time scale discrepancy and the related discrepancy in stretching forces required to induce unfolding, may be overcome through slower pulling speeds in the moving restraint SMD simulations. The slower I27-(0.1 Å/ps) and I27-(0.5 Å/ps) stretching simulations [22] produce the same scenario of clustered hydrogen bond breaking as the faster I27-(1.0 Å/ps) simulation. Figure 1.5 presents the measured AFM forces and their logarithmic extrapolation to faster pulling speeds together with the forces from simulations I27-(1.0 Å/ps), I27-(0.5 Å/ps) and I27-(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.

The results of constant force SMD simulations can be extrapolated to even more closely approach the forces observed in AFM. One may employ the crossing times $\tau_{barrier}$ for simulations I27-(500 pN), I27-(750 pN), I27-(900 pN), I27-(1000 pN) to estimate pulling velocities and extend the set of simulation data shown in Figure 1.5. We assume for this purpose that the forces applied are the peak forces and that the estimated barrier width of 3 Å provides a suitable estimate for the distance traveled. In case of simulation I27-(1000 pN), i.e., for a constant force of 1000 pN, one determines accordingly a velocity of 0.075 Å/ps. The corresponding data point in Figure 1.5 lies indeed close to the data point of simulation I27-(0.1 Å/ps), which exhibited a peak force of 1200 pN. The velocities corresponding to the constant forces 900 pN, 750 pN, and 500 pN are 0.05 Å/ps, 0.027 Å/ps, and 0.003 Å/p s, respectively, which contribute the corresponding data points in Figure 1.5. One can discern from these data points that according to our rough estimate the simulation data approach the measured AFM peak forces when scaled with log(velocity).

We also interpreted the results of the constant force simulations in terms of mean first passage time of barrier crossing (Eq. 1.3) and have found, in applying the theory of first passage times [41, 42, 43], that the estimated 3 Å barrier width, combined with the fit 20 kcal/mol barrier height, match the simulation results; they correspond to the same barrier crossing time as observed in the simulations. These characteristics of the barrier separating the folded and unfolded forms of I27 agree well with AFM observations and chemical denaturation (Chevron plot) data [45] in which a potential barrier of 22 kcal/mol was measured.

Recently, several theoretical groups have also published studies on forced unfolding of protein domains, employing approaches and methods different from those applied in this paper. Rohs *et al.* [46] studied the stretching of α helix and β hairpin systems using molecular mechanics; they estimated the magnitude of forces involved in the unfolding of these secondary protein structures. Socci *et al.* [47] studied the relation between force dependence and the reaction coordinate by stretching a lattice model. Evans and Ritchie [48] modeled the Ig domain unfolding as a single bond braking event and approached the problem using Kramers-Smoluchowski theory. Paci and Karplus [49] studied FnIII unfolding by means of biased molecular dynamics using an implicit solvent model to reduce computational effort; the authors suggested that the dominant barrier against unfolding is due to vdW interactions, and not due to hydrogen bonds. Future SMD simulations will help clarify the differences among different theoretical approaches including those described above and will be aimed to provide a means to unite these approaches, while still relating them to experiment.

Our goal in this modeling of forced titin domain unfolding is to obtain an atomic-level view of the process. We approached this by incorporating into simulations two known properties of the system: the experimentally-derived static structure and the experimentally known force-extension curve. Our resulting theoretical model should meet two criteria: first, the model should correspond closely to the experimental (AFM) situation, so as to provide a non-ambiguous check on the model's validity; second, the results should provide information on the process at atomic-level resolution that cannot be obtained from experiment. The SMD method has satisfied both criteria listed above. Force-extension curves produced by simulations can be directly compared to AFM data to check validity. The SMD trajectories account for the unfolding process at atomic-level detail and, hence, reveal structure-function relationships for protein elasticity properties.

Future SMD simulations can be used to help design proteins with unfolding barriers of desired strength. Additionally, modeling force-

sensitive proteins with SMD can provide insights into novel processes such as the RGD loop conformation changes observed for $FnIII_{10}$. SMD can contribute fundamentally to the understanding of structure-function relationships of mechanical proteins.

Acknowledgments

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

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

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

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

[5] C. Machado, C.E. Sunkel, and D.J. Andrew. Human autoantibodies reveal titin as a chromosomal protein. *J. Cell Biol.*, 141:321–333, 1998.

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

[7] Harold Erickson. 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, 1994.

[8] W. Linke, V. Popov, and G. Pollack. Passive and active tension in single cardiac myofibrils. *Biophys. J.*, 67:782–792, 1994.

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

[10] M.L. Greaser, M.G. Sebestyen, J.D. Fritz, and J.A. Wolff. cdna sequence of rabbit cardiac titin/connectin. *Adv. Biophys.*, 33:13–25, 1996.

[11] Matthias Rief, Mathias Gautel, Filipp Oesterhelt, Julio M. Fernandez, and Hermann E. Gaub. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science*, 276:1109–1112, 1997.

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

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

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

[15] Matthias Rief, Mathias Gautel, Alexander Schemmel, and Hermann Gaub. The mechanical stability of immunoglobulin and fibronectin III domains in the muscle protein titin measured by afm. *Biophys. J.*, 75:3008–3014, 1998.

[16] Andres F. Oberhauser, Piotr E. Marszalek, Harold Erickson, and Julio Fernandez. The molecular elasticity of tenascin, an extracellular matrix protein. *Nature*, 393:181–185, 1998.

[17] Sabina Improta, Anastasia Politou, and Annalisa Pastore. Immunoglobulin-like modules from titin I-band: extensible components of muscle elasticity. *Structure*, 4:323–337, 1996.

[18] D. J. Leahy, I. Aukhil, and H. P. Erickson. 2.0 Angstrom crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region. *Cell*, 84:155–164, 1996.

[19] Hui Lu, Barry Isralewitz, Andre Krammer, Viola Vogel, and Klaus Schulten. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys. J.*, 75:662–671, 1998.

[20] André Krammer, Hui Lu, Barry Isralewitz, Klaus Schulten, and Viola Vogel. Forced unfolding of the fibronectin type III module reveals a tensile molecular recognition switch. *Proc. Natl. Acad. Sci. USA*, 96:1351–1356, 1999.

[21] Hui Lu and Klaus Schulten. Steered molecular dynamics simulations of force-induced protein domain unfolding. *PROTEINS: Struc., Func., and Genetics*, 35:453–463, 1999.

[22] Hui Lu and Klaus Schulten. Steered molecular dynamics simulation of conformational changes of immunoglobulin domain I27 interpret atomic force microscopy observations. *Chem. Phys.*, 1999. In press.

[23] William F. Humphrey, Andrew Dalke, and Klaus Schulten. VMD – Visual Molecular Dynamics. *J. Mol. Graphics*, 14:33–38, 1996.

[24] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, New York, 1987.

[25] Bernard R. Brooks, Robert E. Bruccoleri, Barry D. Olafson, David J. States, S. Swaminathan, and Martin Karplus. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.*, 4:187–217, 1983.

[26] Sergei Izrailev, Sergey Stepaniants, Barry Isralewitz, Dorina Kosztin, Hui Lu, Ferenc Molnar, Willy Wriggers, and Klaus Schulten. Steered molecular dynamics. In P. Deuflhard, J. Hermans, B. Leimkuhler, A. E. Mark, S. Reich, and R. D. Skeel, editors, *Computational Molecular Dynamics: Challenges, Methods, Ideas*, volume 4 of *Lecture Notes in Computational Science and Engineering*, pages 39–65. Springer-Verlag, Berlin, 1998.

[27] Sergei Izrailev, Sergey Stepaniants, Manel Balsera, Yoshi Oono, and Klaus Schulten. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.*, 72:1568–1581, 1997.

[28] Barry Isralewitz, Sergei Izrailev, and Klaus Schulten. Binding pathway of retinal to bacterio-opsin: A prediction by molecular dynamics simulations. *Biophys. J.*, 73:2972–2979, 1997.

[29] Willy Wriggers and Klaus Schulten. Investigating a back door mechanism of actin phosphate release by steered molecular dynamics. *PROTEINS: Struc., Func., and Genetics*, 35:262–273, 1999.

[30] Sergey Stepaniants, Sergei Izrailev, and Klaus Schulten. Extraction of lipids from phospholipid membranes by steered molecular dynamics. *J. Mol. Model.*, 3:473–475, 1997.

[31] Dorina Kosztin, Sergei Izrailev, and Klaus Schulten. Unbinding of retinoic acid from its receptor studied by steered molecular dynamics. *Biophys. J.*, 76:188–197, 1999.

[32] Ferenc Molnar, Lawrence S. Norris, and Klaus Schulten. Simulated unbinding and binding of fatty acid substrates in the cyclooxygenase site of prostaglandin H_2 synthase-1. *Biophys. J.*, 1999. Submitted.

[33] Justin Gullingsrud, Rosemary Braun, and Klaus Schulten. Reconstructing potentials of mean force through time series analysis of steered molecular dynamics simulations. *J. Comp. Phys.*, 151:190–211, 1999.

[34] Helmut Grubmüller, Berthold Heymann, and Paul Tavan. Ligand binding and molecular mechanics calculation of the streptavidin-biotin rupture force. *Science*, 271:997–999, 1996.

[35] Siewert-Jan Marrink, Oliver Berger, Peter Tieleman, and Fritz Jähnig. Adhesion forces of lipids in a phospholipid membrane stud-

ied by molecular dynamics simulations. *Biophys. J.*, 74:931–943, 1998.

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

[37] 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, and M. Karplus. All-hydrogen empirical potential for molecular modeling and dynamics studies of proteins using the CHARMM22 force field. *J. Phys. Chem. B*, 102:3586–3616, 1998.

[38] Mark Nelson, William Humphrey, Attila Gursoy, Andrew Dalke, Laxmikant Kalé, Robert D. Skeel, and Klaus Schulten. NAMD – A parallel, object-oriented molecular dynamics program. *Int. J. Supercomput. Appl. High Perform. Comput.*, 10:251–268, 1996.

[39] Axel 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.

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

[41] Klaus Schulten, Zan Schulten, and Attila Szabo. Reactions governed by a binomial redistribution process. the Ehrenfest urn problem. *Physica*, 100A:599–614, 1980.

[42] Attila Szabo, Klaus Schulten, and Zan Schulten. First passage time approach to diffusion controlled reactions. *J. Chem. Phys.*, 72:4350–4357, 1980.

[43] Klaus Schulten, Zan Schulten, and Attila Szabo. Dynamics of reactions involving diffusive barrier crossing. *J. Chem. Phys.*, 74:4426–4432, 1981.

[44] J.H. Beer, K.T. Springer, and B.S. Coller. Immobilized arg-gly-asp (RGD) peptides of varying lengths as structural probes of the platelet glycoprotein IIb/IIIa receptor. *Blood*, 79(1):117–128, 1992.

[45] Mariano Carrion-Vazquez, Andres Oberhauser, Susan Fowler, Piotr Marszalek, Sheldon Broedel, Jane Clarke, and Julio Fernandez. Me-

chanical and chemical unfolding of a single protein: A comparison. *Proc. Natl. Acad. Sci. USA*, 96:3694–3699, 1999.

[46] Remo Rohs, Catherine Etchebest, and Richard Lavery. Unraveling proteins: A molecular mechanics study. *Biophys. J.*, 76:2760–2768, 1999.

[47] N. Socci, J. Onuchic, and P. Wolynes. Stretching lattice models of protein folding. *Proc. Natl. Acad. Sci. USA*, 96:2031–2035, 1999.

[48] Evan Evans and Ken Ritchie. Strength of a weak bond connecting flexible polymer chains. *Biophys. J.*, 76:2439–2447, 1999.

[49] Emanuele Paci and Martin Karplus. Forced unfolding of fibronectin type 3 modules: An analysis by biased molecular dynamics simulations. *J. Mol. Biol.*, 288:441–459, 1999.