Modeling AFM-Induced PEVK Extension and the Reversible Unfolding of Ig/FNIII Domains in Single and Multiple Titin Molecules

Bo Zhang and John Spencer Evans
Laboratory for Chemical Physics, Department of Chemistry, New York University, New York, New York 10010 USA

ABSTRACT Molecular elasticity is associated with a select number of polypeptides and proteins, such as titin, Lustrin A, silk fibroin, and spider silk dragline protein. In the case of titin, the globular (Ig) and non-globular (PEVK) regions act as extensible springs under stretch; however, their unfolding behavior and force extension characteristics are different. Using our time-dependent macroscopic method for simulating AFM-induced titin Ig domain unfolding and refolding, we simulate the extension and relaxation of hypothetical titin chains containing Ig domains and a PEVK region. Two different models are explored: 1) a series-linked WLC expression that treats the PEVK region as a distinct entropic spring, and 2) a summation of *N* single WLC expressions that simulates the extension and release of a discrete number of parallel titin chains containing constant or variable amounts of PEVK. In addition to these simulations, we also modeled the extension of a hypothetical PEVK domain using a linear Hooke's spring model to account for "enthalpic" contributions to PEVK elasticity. We find that the modified WLC simulations feature chain length compensation, Ig domain unfolding/refolding, and force-extension behavior that more closely approximate AFM, laser tweezer, and immunolocalization experimental data. In addition, our simulations reveal the following: 1) PEVK extension overlaps with the onset of Ig domain unfolding, and 2) variations in PEVK content within a titin chain ensemble lead to elastic diversity within that ensemble.

INTRODUCTION

Molecular elasticity is associated with a select number of polypeptides and proteins (Erickson, 1997; Gautel and Goulding, 1996; Politou et al., 1995; Rief et al., 1997; Urry, 1982; Xu and Evans, 1999; Smith et al., 1999; Hayashi and Lewis, 1998; Zhang et al., 2000; Oberhauser et al., 1998). Common to these proteins are unique molecular aspects (e.g., secondary and tertiary structure) which convey elastic properties. Considerable attention has been focused on the muscle protein, titin, which is located within the sarcomere region of skeletal and cardiac muscle (for a review see Linke and Granzier, 1998). In the case of titin, molecular elasticity is conveyed by two structurally distinct domain types. 1) The immunoglobulin (Ig-I subset)/fibronectin type III (FNIII) region (Erickson, 1997; Harpaz and Chothia, 1994; Fraternali and Pastore, 1999; Politou et al., 1995; Labeit et al., 1992; Kellermayer et al., 1997; Rief et al., 1997); and 2) the semistable spring-like Pro-Glu-Val-Lys (PEVK)-repeat domain (Rief et al., 1997; Kellermayer et al., 1997; Linke, 1996). Using experimental techniques, sequential unfolding of titin Ig domains is observed under force, but the refolding phase is nonsequential, and does not initiate until the protein molecule is extensively retracted (Rief et al., 1997, 1998; Linke et al., 1998a; Kellermayer et al., 1997). To help interpret these experimental results, macroscopic simulation methods have been utilized to model the extension and release of the titin chain. The most common method is the entropic spring-based worm-like chain (WLC) model (Flory, 1969). WLC-based simulations have yielded force curves similar to those obtained from AFM, laser tweezer, and immunoelectron microscopy experiments (Rief et al., 1997, 1998; Kellermayer et al., 1997; Linke et al., 1998a, b). Moreover, the WLC model has been adapted for modeling PEVK extension under low force conditions via the use of a linear modulus term (Linke et al., 1998b).

However, the WLC model is not without its shortcomings. First, the titin molecule is composed of a discrete number of globular domains (Ig) and a nonglobular segment (PEVK). Hence, under repetitive force extension and release cycles, we would expect each region to behave with different elastic characteristics. The single WLC spring model treats the entire titin chain as a single spring, and, thus fails to account for the differences in molecular elasticity within the chain. Second, the single WLC model is inadequate for modeling *non-entropic* contributions to PEVK stretch and relaxation. This is clearly observed under conditions of high stretch, where the single WLC model deviates from the experimental data (Linke et al., 1998a, b). It is suspected that these deviations arise from enthalpic contributions (i.e., electrostatic, hydrophobic), which are believed to be responsible for PEVK elastic behavior (Linke et al., 1998a, b). Third, within one-half sarcomere, ~1000–2000 parallel titin molecules (Linke et al., 1998a) will be functioning simultaneously. A multichain simulation may prove to be more informative with regard to the potential synergistic effect of multiple titin molecules undergoing unfolding and refolding (Trombitas et al., 1998a; Helmes et al., 1999).

In the present paper we test specific modifications of the WLC model within the context of simulating AFM-induced

Received for publication 9 June 2000 and in final form 27 November 2000.
Address reprint requests to Dr. John Spencer Evans, Division of Basic Sciences, Laboratory for Chemical Physics, New York University, 345 E. 24th St., Rm. 1007, New York, NY 10010. Tel.: 212-998-9605; Fax: 212-995-4087; E-mail: jse@dave-edmunds.dental.nyu.edu.

0006-3495/01/02/597/09 $2.00

Ig unfolding/refolding and PEVK extension. With the recent development of a time-dependent method for simulating AFM force extension (i.e., standard three-state kinetic protein folding model), we can model the consecutive unfolding and refolding of elastic domains within a single titin protein molecule as a chain propagation reaction (Zhang et al., 1999). We now integrate this kinetic method with two different approaches. This first approach treats the titin chain as two separate entropic springs, i.e., the Ig region and the PEVK segment (Trombitas et al., 1998a, b; Helmes et al., 1999). The second approach examines two kinds of hypothetical multichain titin assemblies: ensembles that have constant PEVK content (mimicking the single isoform titin found in skeletal muscle sarcomeres) or variable PEVK content (mimicking the multiple isoform titin found in cardiac muscle sarcomeres). In addition to these Ig-specific simulations, we also perform force extension simulations on a hypothetical PEVK domain, using a Hooke's spring term to represent the "enthalpic" contribution to PEVK elasticity. This type of simulation is used to gain further insight into PEVK behavior. In each simulation we determine whether specific modifications to WLC theory can lead to improvements in the modeling of titin unfolding and/or PEVK extension. Furthermore, we utilize these simulations to help define the participation of PEVK within the context of single and multichain titin extension and relaxation. We find that specific modifications of the basic WLC model lead to an improved description of several aspects of the Ig unfolding and refolding process and/or PEVK extension. In addition, the simulation results obtained for each model implicate PEVK domain participation in titin chain length compensation, and as a contributor to the heterogeneity of titin chain unfolding events (i.e., creating elastic diversity) (Freiburg et al., 2000) within an ensemble of cardiac titin chains.

METHODOLOGY

In this section we will briefly summarize our kinetic, concentration-based protein domain unfolding and refolding model (see Zhang et al., 1999, for a full description). Following this, we will describe the WLC modifications that are incorporated into this model and the simulation parameters and conditions utilized for this current study.

Overview

The basic premise of our three-state protein domain folding and unfolding model is the following: *in order for a given stretched, unfolded globular domain to refold to its original length, the domain has to overcome the energy barrier induced by the extension force.* Thus, refolding of a single protein domain can be viewed as a kinetic two-step process:

$$\mathrm{U} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{T} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \mathrm{F} \tag{1}$$

The states involved are the unfolded (U), intermediate or transition (T), and fully folded (F), where k_1 and k_2 are the forward rate constants for the U $\rightarrow$ T and T $\rightarrow$ F refolding transitions, and k_{-1} and k_{-2} are the reverse rate constants for the T $\rightarrow$ U and F $\rightarrow$ T unfolding transitions, respectively. If we consider a single titin molecule to be composed of N protein domains, we can express the overall refolding kinetics of the entire protein as:

$$\mathrm{F}_0 \underset{k_{10}}{\overset{k_{01}}{\rightleftharpoons}} \mathrm{F}_1 \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} \mathrm{F}_2 \cdots \underset{k_{\mathrm{n,n-1}}}{\overset{k_{\mathrm{n-1,n}}}{\rightleftharpoons}} \cdots \mathrm{F}_\mathrm{N} \tag{2}$$

where $\mathrm{F_n}$ represents the species that have n of N domains in the folded state, and $k_{\mathrm{n-1,n}}$, $k_{\mathrm{n,n-1}}$ denote the rate constants for the forward and reverse transitions involving $\mathrm{F_{n-1}}$ to $\mathrm{F_n}$, respectively. From this, we constructed an expression that describes the time-dependent concentration of each species that undergoes refolding. To do this, we make a simplifying assumption that each species unfolds and refolds independently of one another, i.e., a non-cooperative case. We can obtain the concentration distribution of all species, i, for a period of time, t, which corresponds to a specific extension during the simulation of the AFM force-extension experiments. From the concentration distribution, we obtain the average number of folded domains, N:

$$N(t) = \sum i \cdot C_\mathrm{i}^\mathrm{t} \tag{3}$$

and the average force, F:

$$F(t) = \sum f_\mathrm{i} \cdot C_\mathrm{i}^\mathrm{t} \tag{4}$$

We then integrate this concept with the WLC model. The standard WLC model describes a molecular chain as a deformable continuum or rod of a given persistence length, A, which is a measure of the molecule's stiffness. The relationship between the end-to-end length (z) and the external force (f) are given by (Flory, 1969; Kellermayer et al., 1997; Rief et al., 1997; Oberhauser et al., 1998; Zhang et al., 1999):

$$f = \frac{kT}{A} \cdot \left[\frac{1}{4\left(1 - \frac{z}{L}\right)^2} - \frac{1}{4} + \frac{z}{L} \right] \tag{5}$$

where k is the Boltzmann constant, T is temperature, and L is contour length (i.e., the length of a fully extended WLC chain). To integrate this model within the time- and concentration-dependent unfolding and refolding kinetic scheme, we assume a constant pulling/releasing speed of the AFM tip. Thus, each cycle will have a duration, T,

$$\mathrm{T} = 2\left[\frac{L_\mathrm{max} - L_\mathrm{min}}{V}\right] \tag{6}$$

where L_max and L_min are the maximum and minimum contour lengths, respectively, and V is the pulling-releasing velocity. Using a discrete time scheme to approximate the unfolding-refolding simulation (Zhang et al., 1999), the cycle period is partitioned into an M grid lattice, with $\Delta t = T/M$. At a lattice point we have $t_\mathrm{m} = m^*\Delta t$, and $L_\mathrm{m} = L_\mathrm{min} + t_\mathrm{m}^* V$. The values of K and λ are calculated accordingly; as are $C_\mathrm{i}(t)$ for $t_\mathrm{m} < t < t_\mathrm{m+1}$; $C_\mathrm{i}(t)$ at $t = t_\mathrm{m+1}$; and for C_i^0 for $t_\mathrm{m+1}$ to $t_\mathrm{m+2}$. At $t = 0$, all the domains are assumed to exist in the fully folded state.

"Dual spring" (series linked) WLC model

To model the Ig region and PEVK domain within the titin chain, we make the assumption that the elastic properties and force-extension behavior for each region are different from one another. This approach, i.e., defining the titin chain to be a sequential spring composed of two individual springs (PEVK + Ig) linked in series, was used to examine the relationship between contour length and extension (Trombitas et al., 1998a, b; Helmes et al., 1999). In this report we will use the series-linked or "dual spring"

WLC model for a different purpose: to examine the influence of PEVK elasticity on Ig unfolding and refolding cycles within a single titin chain.

For the dual spring model, the overall end-to-end length, z, of the single titin chain can be determined by summing the end-to-end lengths for the globular (z_1) and non-globular (z_2) domain components. In a dual spring model, the force is considered to be identical throughout the whole chain (Trombitas et al., 1998a, b; Helmes et al., 1999). The dual spring WLC equation can be written as:

$$\frac{1}{A_1}\cdot\left[\frac{1}{4\left(1-\frac{z_1}{L_1}\right)^2}-\frac{1}{4}+\frac{z_1}{L_1}\right]=\frac{1}{A_2}\cdot\left[\frac{1}{4\left(1-\frac{z_2}{L_2}\right)^2}-\frac{1}{4}+\frac{z_2}{L_2}\right] \tag{7}$$

The variables in this expression have the same meaning as in the single WLC model (Zhang et al., 1999), i.e., A_1, A_2 are the persistence lengths for each spring and L_1, L_2 are the contour lengths for each domain type. For the dual spring simulation a hypothetical titin polypeptide chain was constructed using 30 Ig-like domains (contour length = 930 nm) and a PEVK domain content corresponding to 25% of the total residues in the human titin primary sequence (Gautel and Goulding, 1996; Zhang et al., 1999; EMBL data library, human skeletal titin, accession X90569). This percentage value represents an *approximation* of the PEVK sequence for our simulations. The persistence lengths for the Ig and PEVK domains were taken to be 15 nm (Higuchi et al., 1993) and 0.5 nm (Linke et al., 1998b), respectively. For each Ig domain, the fully unfolded and folded lengths are 31 and 5 nm, respectively. This simulation utilized a pulling/release velocity of 100 nm/s.

Multiple parallel chain WLC simulation

This simulation describes the elastic behavior of a finite number of parallel titin chains that are stretched between a common AFM tip and a given surface. In a single chain WLC simulation, a typical saw-tooth force extension pattern is observed, with each force peak reflecting the unfolding of a single Ig domain within the titin chain (Zhang et al., 1999; Rief et al., 1997, 1998). However, titin molecule stoichiometry in situ is substantially more complex: within a single half-sarcomere, 1000–2000 parallel titin molecules are believed to exist (Freiburg et al., 2000; Cazorla et al., 2000). Thus, more than one titin molecule will undergo extension at the same time.

To emulate this scenario, we wish to determine a series of AFM-induced force extension curves for a discrete number of parallel skeletal or cardiac titin chains (defined as N) that are stretched between a single AFM tip and a common surface. We will consider two types of multichain simulations. 1) Single isoform: the chain ensemble consists of single isoform titin molecules, similar to what one would find in a skeletal muscle sarcomere (Trombitas et al., 1998b; Kellermayer et al., 1997; Labeit et al., 1992). Here, we consider each chain to possess a constant number of Ig domains (i.e., 30, contour length = 930) and PEVK content (25% of the total residues) (Gautel and Goulding, 1996; Zhang et al., 1999, EMBL data library, human cardiac titin, accession X90568, skeletal titin, accession X90569). 2) Multiple isoform: this hypothetical titin chain model represents the cardiac titin molecule, which is expressed as several isoforms (e.g., N2B and N2BA) with PEVK content differences existing within these isoforms (Freiburg et al., 2000). In order to compare this model to the foregoing one, and for simplicity, we model these isoform sequence variations as follows: we maintain a constant number of Ig domains per molecule, but we vary the PEVK content. This treatment leads to contour length variations within each titin chain; the maximum difference would be no more than 31 nm, which is the length of an Ig domain.

Based upon these two multichain models we can model the extension and recovery of a parallel titin chain ensemble in the following way. First, apply a single WLC expression (Eq. 5) to each titin molecule. Next, obtain the overall force-extension relationship via summation of the individual WLC expressions. For single isoform simulations, we utilize the same hypothetical titin chain as per our previous report (Zhang et al., 1999). For multiple isoform simulations, contour length differences are generated via variation in the number of residues in the PEVK domain. In other words, the PEVK domain length difference between two adjacent chains is:

$$\text{In-phase}=\frac{(L_{\text{Ig domain}})}{N}$$
$$\text{Out-of-phase}=\frac{(L_{\text{Ig domain}})}{2N} \tag{8}$$

Note that as the value of N increases, the phase differences will decrease. However, the phase difference cannot be smaller than the contour length of a single residue, which is ~0.4 nm, or the maximum spacing length of an amino acid. So, when the contour length exceeds 0.4 nm, Eq. 8 is utilized; otherwise, a value of 0.4 nm is utilized for in-phase chains and 0.2 nm for out-of-phase chains. Single and multiple isoform multiple chain simulations involved the following titin chain ensembles: 3, 5, 15, 31, 45, 61, and 91.

Modeling "enthalpic" contributions to PEVK elasticity using Hooke's spring model

We now wish to consider a simulation that focuses on the behavior of a hypothetical PEVK domain. If one considers that enthalpic properties (i.e., hydrophobic, electrostatic interactions) contribute to the elastic properties of the PEVK domain, then one can treat the enthalpic effect as a small perturbation, e.g., adding a nonlinear stretch modulus term (f/K_o) to the single WLC model (Linke et al., 1998b). However, the value f/K_o is difficult to obtain precisely. We are interested in addressing the contribution of PEVK domain to the force extension process in a slightly different way. Let us consider a hypothetical chain that contains only a PEVK region (no domain number assigned). This chain can be described by a WLC model to which a *linear* spring term (i.e., the Hooke spring model) is added:

$$f=\frac{kT}{A}\cdot\left[\frac{1}{4\left(1-\frac{z}{L}\right)^2}-\frac{1}{4}+\frac{z}{L}\right]+k'\cdot z \tag{9}$$

Here, the modulus term, $k'\cdot z$, represents the enthalpic contributions, z is the end-to-end distance of the entropic spring, and k' is Hooke's constant. The advantages of using the Hooke's spring over the stretch modulus term are twofold. First, the Hooke model is a classic description of a linear spring and has been utilized for describing elasticity (Vinckier and Semenza, 1998). Thus, Hooke's spring model is used to represent the "enthalpic" spring characteristics of the PEVK domain. Second, for any small perturbation term like $k'\cdot z$, we have the advantage of approximating the polynomial form to the first order-term via Taylor expansion. Thus, the Hooke spring modulus term can be evaluated directly. Using Eq. 9, we perform a force-pulling simulation of this hypothetical chain with no relaxation phase, over the extension region of 0 to 100 nm (i.e., the low stretch regime where PEVK domain extension reportedly occurs). For comparison, 1) we perform separate, parallel AFM pulling simulations using the standard Hooke's law and the single WLC model; and 2) perform a curve-fitting of the WLC + Hooke's spring simulation using the single WLC model. All parameters in Eq. 9 are pseudo-generated, so the units will be normalized to kT. Thus, the contour length, persistence length, and Hooke's constant are 100 kT/1 pN, 1 kT/1 pN, and 1 kT/1 pN, respectively.

RESULTS

General parameters utilized in this series of simulations are identical to those reported in our previous work (Zhang et al., 1999), except as noted in the Methods section. A temperature of 298 K was utilized for all simulations.

Simulation of unfolding and refolding events in a single titin molecule as a function of AFM tip extension: comparison of single and dual spring WLC models

We compare the single WLC and dual-spring WLC models with regard to unfolding and refolding of individual Ig domains during the AFM-induced extension and release cycle (Fig. 1). As shown in Fig. 1, the areas traced by both curves are nonequivalent; the dual spring simulation unfolding + refolding pathway is reduced, compared to that obtained for the single WLC simulation. Moreover, if we examine the interval before the first Ig domain unfolding event, we note that the single WLC simulation has a lag period that is four times longer than that observed in the dual spring simulation. These differences can be explained as follows. In the single WLC simulation, the PEVK domain is modeled as an entropic spring that is indistinguishable in behavior from the Ig domains. Thus, at low stretch, there is the presence of a lag period that reflects chain length compensation in the hypothetical titin molecule (Gautel and Goulding, 1996). This region would correspond to the monotonic segment of the AFM force extension profile, which supposedly represents PEVK extension (Rief et al., 1997). However, in the dual spring simulation, the PEVK domain is modeled as an independent spring with different force characteristics. As a result, we observe a decrease in chain length compensation *before* the unfolding of the Ig segment. Interestingly, the first 4–5 Ig domains experience prolonged unfolding transitions in the dual spring simulation, compared to their counterparts in the single spring simulation (Fig. 1; note the presence of lag periods within unfolding transitions in the extension regime of 200–400 nm). This suggests that the PEVK segment is undergoing extension during the early phases of Ig unfolding and extension. This result is supported by the experiments of Trombitas and co-workers (1998b) which suggest that PEVK is extensible within the moderate to high extension force range where Ig domains are observed to unfold. For extension lengths >400 nm, the lag phase periods are observed to diminish within successive Ig unfolding events (Fig. 1).

Another interesting trend was observed in the initiation of the unfolding and refolding processes. From previous single WLC-based simulations, we know that the extent of Ig refolding is directly related to the PEVK content (Zhang et al., 1999). As shown in Fig. 1, for the single WLC model, unfolding commenced at 380 nm extension and the refolding process commenced at 450 nm extension. However, for the dual spring model, unfolding commenced at 220 nm extension and the refolding process initiated at 570 nm extension. Given that the initiation of the refolding process

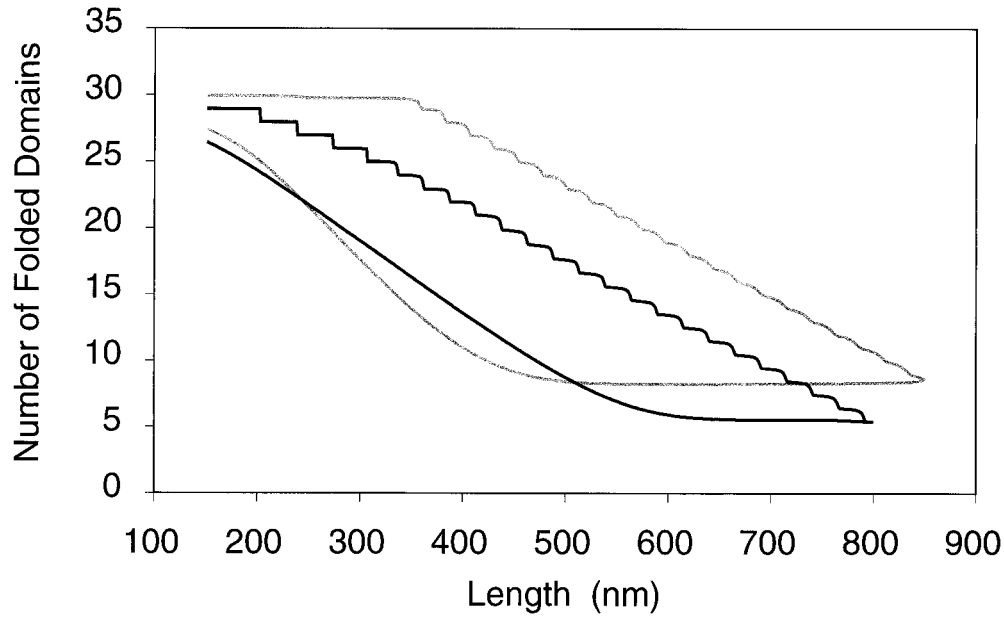

FIGURE 1 Simulation of unfolding and refolding events in a single titin molecule as a function of AFM tip extension. *Gray curve*: single WLC model; *black curve*: dual-sequential WLC model. For comparison with experimental data, see the following references and the original figures noted therein: Rief et al., 1997 (Figs. 1, 3), 1998 (Fig. 4); Oberhauser et al., 1998 (Fig. 1); Kellermayer et al., 1997 (Fig. 3 *A*).

was experimentally observed at half of the titin chain contour length (i.e., 560 nm), (Rief et al., 1997, 1998; Kellermayer et al., 1997), we find that the dual spring WLC model provides a better description of the experimentally probed refolding process. Thus, by incorporating a "PEVK-like" spring within the entropic Ig-containing titin chain, particular aspects of the titin chain unfolding process (i.e., chain length compensation, refolding initiation) approach the experimentally observed behavior.

Comparison of AFM force extension simulations using the single and multiple chain titin WLC models

This type of simulation can provide information on how chain number and, in the case of the cardiac isoform simulation, how sequence length variation (i.e., PEVK content) affect the force extension saw-tooth pattern, the force peak-to-peak distances, and the general domain unfolding and refolding cycle itself. The skeletal isoform force extension simulations, for $N = 1$, 3, 5, 15, 31, 45, 61, and 91, are shown in Fig. 2. In general, as N increases, the areas encompassed by the unfolding-refolding curves increase, and the AFM tip force required for chain extension increases with chain number. This result is not unexpected, given a linear summation of WLC expressions. In comparison, the cardiac isoform extension simulations (Fig. 3, with expansion plots of the low stretch regime presented in Fig. 4) are more complex, viz: 1) note the presence of distinct overlapping or superimposed force peaks in multichain simulations where $N = 3$, 5; this phenomenon is not observed in the single chain simulation (Figs. 3 and 4). Since each chain within the 3- or 5-chain ensemble possesses a slightly different PEVK content, a different extension force will be required to unfold a given Ig domain within each chain. This leads to a heterogeneous force extension response or "elastic diversity" (Freiburg et al., 2000) within the hypothetical cardiac titin multichain ensemble. This heterogeneity is manifest by the presence of superimposed Ig unfolding transitions in each force curve. It should be noted that as chain extension progresses, the superimposed peaks decrease in intensity. 2) For $N \geq 15$, we observe the presence of broader unfolding transitions, particularly within the extension regime of 150–250 nm. These broader transitions are observed to diminish as the cardiac titin chain extension progresses. 3) Simulations involving $N \geq 15$ chains also feature asymmetric profiles or irregularities that are not observed in a single chain simulation. Since the value of N also reflects PEVK content differences, it is plausible to suggest that variations in length compensation will also occur for each cardiac titin chain. Thus, for each titin chain, we would expect a corresponding variation in the initiation and conclusion of each Ig domain unfolding event, resulting in the appearance of a broad force peak on the simulation plot.

One other interesting phenomenon was noted in the cardiac titin simulations. As shown in Figs. 3 and 4, for $N = 15$, 31, 45, 61, 91, the in-phase force curves simulations feature abrupt unfolding transitions with larger force amplitudes, whereas the out-of-phase curves feature more gradual unfolding transitions with smaller force amplitudes. In Fig. 5 we plot the average force peak ratio (i.e., OP/IP, where the

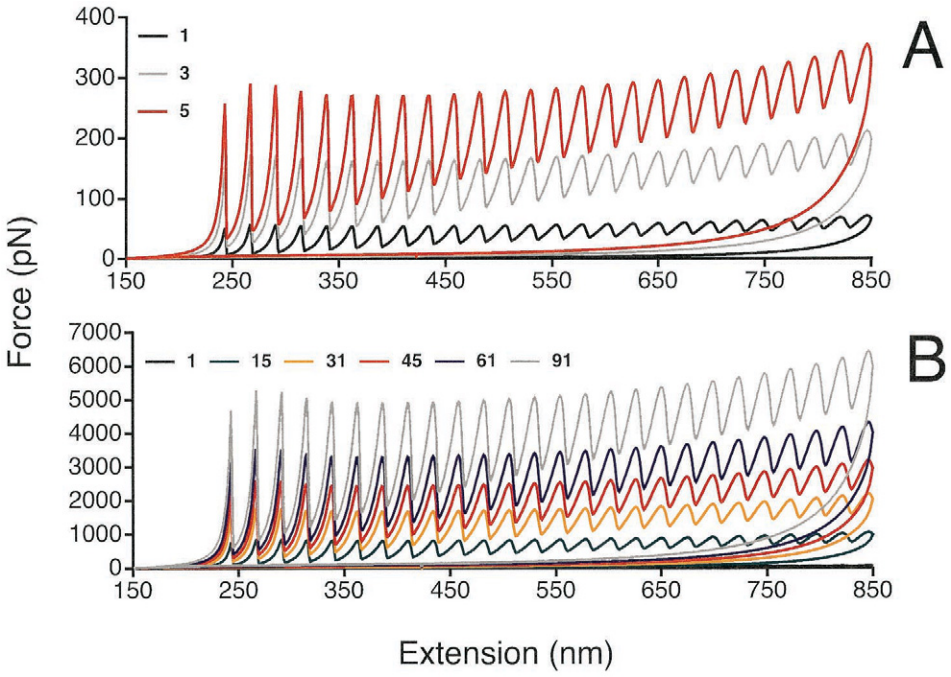

FIGURE 2 Comparison of AFM force extension simulations using the single and multichain skeletal titin WLC models. (*A*) $N = 1$, 3, 5; (*B*) $N = 1$, 15, 31, 45, 61, 91. Each curve is color-coded according to the number of titin chains in each simulation run (see legend within figure). For comparison with experimental data, see the following references and the original figures noted therein: Rief et al., 1997 (Figs. 1, 3), 1998 (Fig. 4); Oberhauser et al., 1998 (Fig. 1); Kellermayer et al., 1997 (Fig. 3 *A*); Smith et al., 1999 (Figs. 2, 3); Zhang et al., 2000 (Figs. 1, 7, 8).

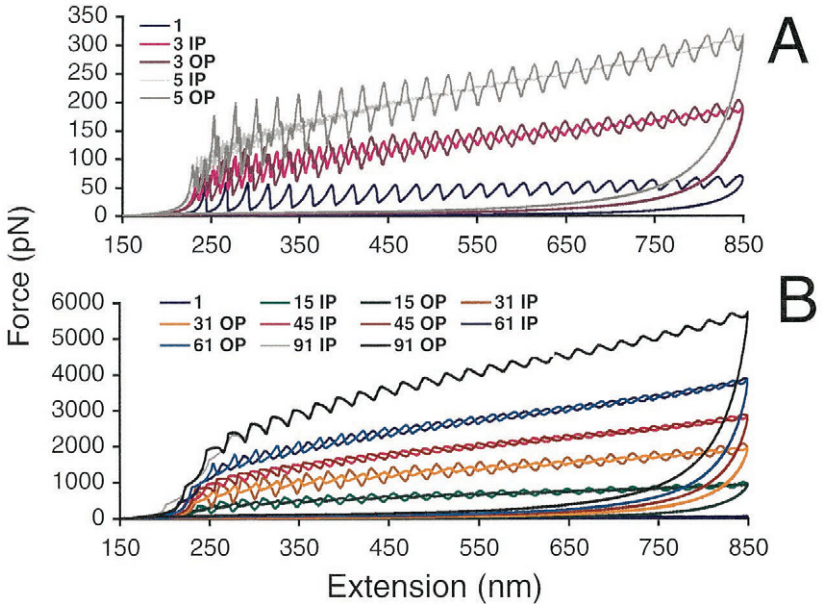

FIGURE 3 Comparison of AFM force extension simulations using the single and multichain cardiac titin WLC models. (*A*) $N = 1, 3, 5$; (*B*) $N = 1, 15, 31, 45, 61, 91$. Each curve is color-coded according to the number of titin chains in each simulation run (see legend within figure). IP = "in-phase"; OP = "out-of-phase." For comparison with experimental data, see the following references and the original figures noted therein: Rief et al., 1997 (Figs. 1, 3), 1998 (Fig. 4); Oberhauser et al., 1998 (Fig. 1); Kellermayer et al., 1997 (Fig. 3 *A*); Smith et al., 1999 (Figs. 2, 3); Zhang et al., 2000 (Figs. 1, 7, 8).

averages are taken from the first force peak through the last observed force peak). Here we find that OP/IP approaches 1:1 for $3 \leq N \leq 45$, and 1.5:1 for $N > 45$. For $N = 91$, we observe negligible phase and amplitude differences between the force peaks at extensions > 300 nm (Figs. 3–5). However, this is not the case for $N < 91$. These observations can be explained as follows. First, when N is small, the differences in PEVK content are significant for adjacent cardiac titin chains, and thus a force peak phase shift would be observed. However, as N increases, the probability of having two or more non-adjacent chains with similar PEVK content also increases. Thus, we would expect force peak phase shifting until we reach a phase shift of 360° (i.e., at $N = 91$), where phase coherence would occur. This is what we observe in Figs. 3–5.

Overall, our multichain simulations indicate the following: 1) an increase in N leads to an increase in the extension force required to unfold Ig domains and extend the titin molecule. Although there are no experimental data available for AFM extension of multichain skeletal or cardiac titin bundles, we note that observations of increasing force response and force peak broadening have been observed in AFM pulling experiments involving nacre layer matrix consisting of multiple elastic protein chains (Smith et al., 1999) and in AFM pulling experiments involving multiple *Bombyx mori* silk fibroin proteins (Zhang et al., 2000). More-

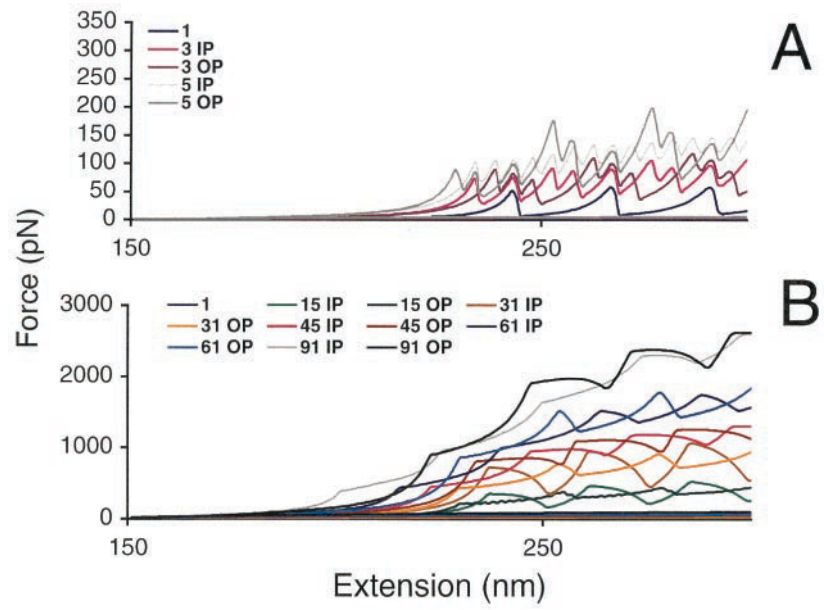

FIGURE 4 Comparison of AFM force extension simulations using the single and multichain cardiac titin WLC models. As per Fig. 3, displaying expansion of the 150–300 nm extension range.

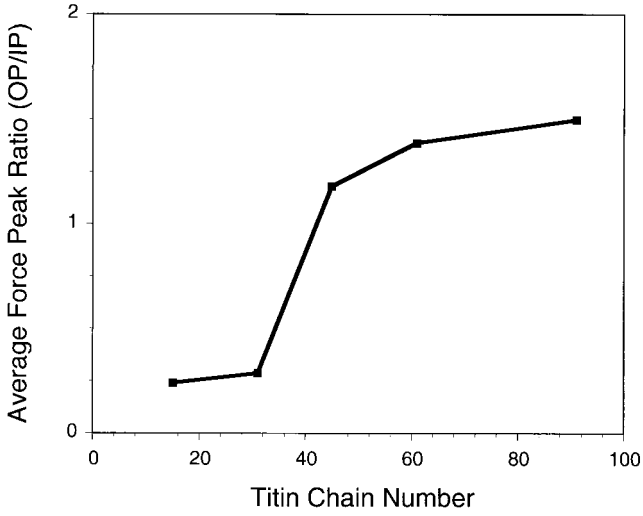

FIGURE 5 Average OP/IP ratio as a function of titin chain number.

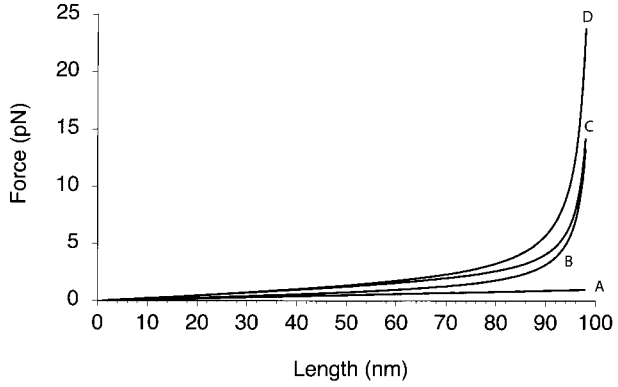

FIGURE 6 Comparison of AFM force extension simulations using the WLC ("entropic") and WLC + Hooke's spring ("enthalpic") model. Plots reflect the extension of a single titin molecule in the low stretch regime (<100 nm extension). Legend to figure: (*A*) Hooke's spring, with k' (Hooke's constant) = 0.01 kT; (*B*) single WLC model, for persistence length = 1; (*C*) single WLC + Hooke's spring; (*D*) single WLC model fitting of WLC + Hooke's spring simulation curve. For reference to experimental data, see Linke et al., 1998b, Fig. 4.

over, using laser tweezer instrumentation, Kellermayer and colleagues (1997) observed increased force resistance for triple titin chain assemblies compared to a single titin chain. Collectively, these results parallel those obtained from our single isoform and multiple isoform simulations (Figs. 2–4). 2) Variations in PEVK content modulate the force required to unfold Ig domains and generate force extension heterogeneity (i.e., create elastic diversity) within an ensemble of cardiac titin chains. With regard to cardiac titin, it has been shown that exon-skipping pathways modulate the fractional extensions of the tandem Ig and PEVK segments, thereby influencing myofibrillar elasticity (Freiburg et al., 2000). Furthermore, the expression of small and large titin isoforms at different ratios has been postulated to be a means of modulating cardiac myocyte stiffness. These findings are in agreement with our multichain multiple isoform simulations (i.e., cardiac titin), which show that extension force variations and elastic diversity within an ensemble can occur when sequence length variations are present [i.e., PEVK length (Figs. 3 and 4)]. Hence, our simulation data are qualitatively consistent with available multichain experimental data.

Comparison of PEVK extension simulations using the WLC and WLC-Hooke's spring model

In Fig. 6 we examine the extension of a single PEVK chain over the effective range of PEVK domain elastic extension (0–100 nm). We first generated a force extension curve using the WLC + Hooke's spring model (i.e., "entropic" spring model with "enthalpic" perturbation term). Next, we generated a force extension curve using a Hooke's spring model. Finally, we utilized the single WLC model for fitting to the entropic + enthalpic curve. As shown in Fig. 6, it is evident that the single WLC model provides a good fit to the single WLC + Hooke's simulation force curve at extensions ≤50 nm. However, at extensions >50 nm, the curve fitting is less adequate. This finding is in agreement with experimental and theoretical studies of PEVK elasticity (Linke et al., 1998b), where it was noted that the single WLC + modulus model provided a closer fit to the AFM pulling data as compared to the results obtained using the pure entropic WLC model. Although it is clear that the WLC model cannot fully reproduce experimental AFM pulling data, we do concur with Linke and co-workers that the "entropic + enthalpic" WLC model provides a better description of PEVK domain extension in titin.

DISCUSSION

In this report we explore different modifications of the single WLC expression and their applicability for modeling titin chain Ig unfolding and refolding and PEVK extension in response to force. Using our kinetic-based time-dependent multiple domain unfolding model, we examine the "dual spring" (i.e., PEVK + Ig/FNII) and multichain adaptations of the WLC model and their ability to reproduce experimental force extension data. In addition, we simulate the PEVK domain as an "enthalpic" + "entropic" spring under extension and examine the performance of the PEVK domain under these conditions. Although it is unrealistic to expect any one simulation model, or combination of models, to fully reproduce experimental force extension data, we note that each modification of the WLC expression led to an improvement in our ability to reproduce experimental data (Figs. 1–4, 6). In a sense, these results underline the inadequacies of the *single* WLC model for simulating PEVK extension and Ig unfolding and refolding transitions, and

point to a real need for better simulation methods that capture the essence of AFM, laser tweezer, and other types of force extension experiments.

The results obtained from "dual spring" WLC simulations indicate that the PEVK segment is extensible within the moderate to high extension force range where Ig domains are observed to unfold (Fig. 1), in agreement with the findings of Trombitas and co-workers (1998b). Based upon these observations, we hypothesize that PEVK extension and subsequent chain length compensation may influence the unfolding of the first 4–5 Ig/FNII domains. Another interesting finding is the impact that the PEVK segment has on the titin *refolding process*, or more specifically, the initiation of refolding and the recovery of the titin chain (Fig. 1; note the difference in curve areas). Much discussion has been accorded to the possible involvement of PEVK in Ig unfolding (Linke et al., 1998a, b; Trombitas et al., 1998b). A similar finding was also noted in our earlier simulation study (Zhang et al., 1999). At this time, it is not known precisely how, if at all, the PEVK domain assists with Ig refolding within the titin chain in vivo or in vitro.

It is known that PEVK content differences have been found to exist within the N2B and N2BA cardiac titin isoforms (Freiburg et al., 2000). The question is, how does PEVK content variation affect the force extension behavior of the titin ensemble within a cardiac sarcomere? The multichain multiple isoform simulations (Figs. 3 and 4) reveal that variations in PEVK content (modeled here as phase differences) can lead to the appearance of superimposed or overlapping Ig unfolding transitions for different titin chains. These superimposed unfolding transitions arise from 1) variations in pulling force that are required to unfold a Ig domain within a given titin chain; and/or 2) the extension of PEVK domains under low force before the unfolding of the first Ig domain. Thus, within an ensemble of titin chains, we hypothesize that PEVK content variations may result in individual titin molecules experiencing slightly different percentages of chain extension and Ig unfolding as a function of pulling force. This, in turn, could lead to elastic diversity within the cardiac titin chain ensemble, as proposed elsewhere (Freiburg et al., 2000). With regard to the multichain single and multiple isoform simulations, it is evident that as the number of titin molecules increases, 1) the force required to unfold the titin chain also increases, and 2) force peak broadening effects will also increase (Figs. 2–4). Although no comparable AFM experiments exist for titin proteins, recent AFM studies conducted on abalone shell nacre layer organic matrix proteins revealed the presence of broad, overlapping force saw-tooth peaks (Smith et al., 1999), which possibly reflects the unfolding of different proteins, particularly Lustrin A. Similar force peak characteristics were noted in AFM force spectroscopy experiments involving multiple silk fibroin proteins (Zhang et al., 2000). It is possible that multichain simulation methods may have application to future AFM force extension studies involving titin chain bundles, particularly with regard to evaluating the PEVK content differences that exist within a cardiac titin chain ensemble.

There is evidence from our simulations that PEVK domains play an important role in the initial phase of titin chain extension (i.e., up to 300 nm extension). As shown in Fig. 5, with the addition of a Hooke spring "enthalpic" perturbation term, the unfolding simulation mimics the early events in force extension to a better degree than the single WLC model can. This supports the earlier findings of Linke and co-workers (1998b) that "enthalpic" contributions (i.e., electrostatic, hydrophobic interactions) provide part of the driving force for PEVK elasticity, and underlines the importance of the non-globular PEVK domain in titin extension (Linke et al., 1998a, b; Trombitas et al., 1998a, b; Zhang et al., 1999). However, even this "perturbation" term is not enough to fully describe the initial force curve characteristics of a titin force-extension experiment (Linke et al., 1998a, b; Trombitas et al., 1998a, b). We conclude that macroscopic treatments, such as the WLC model, are insufficient for modeling what are essentially interactions at the atomic and molecular level. Thus, future simulations of PEVK spring-like unfolding and refolding will most likely require mesoscopic or even microscopic approaches (i.e., explicit representation, free energy perturbation molecular dynamics, Monte Carlo, solvation treatments, mean force potentials) to accurately depict the "enthalpic" contributions to force resistance.

Finally, we would like to discuss some potential applications of the kinetic protein unfolding/refolding model, and point to some future simulation trends that may improve our ability to interpret experimental AFM force extension data. First, with regard to applications, we suggest that the kinetic protein unfolding/refolding model, when paired with an appropriate spring model, could be used to simulate the force extension behavior of other elastomeric proteins, such as the silk fibroin protein (Zhang et al., 2000) and biomineralization-specific matrix proteins (Smith et al., 1999). In these potential applications one may need to modify the kinetic-based method to accommodate several different domain types and their folding equilibria, as well as the presence of more than one intermediate folded state. In terms of future simulation trends, other macroscopic chain models, such as the broken rodlike chain model (BR) (Muroga, 2000) and the modified freely jointed chain (MFJC) (Zhang et al., 2000), may need to be considered as a starting point for developing an improved simulation of molecular elasticity in proteins. In fact, the MFJC model was recently applied to non-titin AFM force spectroscopy studies with reasonable success (Zhang et al., 2000). In addition to the use of more descriptive entropic spring models, there is also a need to incorporate the enthalpic contributions arising from the PEVK or other similar domains within elastic proteins. In light of these new demands, we are currently

investigating new approaches that address these modeling requirements.

This work was supported by the National Science Foundation (Grants DMR 99-01356, MCB 98-16703) and a Young Investigator Award from the Army Research Office (DAAD19-99-0225). This paper is contribution 13 from the Laboratory for Chemical Physics, New York University.

REFERENCES

Cazorla, O., A. Freiburg, M. Helmes, T. Centner, M. McNabb, Y. Wu, K. Trombitas, S. Labeit, and H. Granzier. 2000. Differential expression of cardiac titin isoforms and modulation of cellular stiffness. *Circ. Res.* 86:59–67.

Erickson, H. P. 1997. Stretching single protein molecules: titin is a weird spring. *Science.* 276:1090–1091.

Flory, P. J. 1969. Statistical Mechanics of Chain Molecules. The Statistical Distribution of Configuration. Wiley Interscience, New York.

Fraternali, F., and A. Pastore. 1999. Modularity and homology: modeling of the type II module family from titin. *J. Mol. Biol.* 290:581–593.

Freiburg, A., K. Trombitas, W. Hell, O. Cazorla, F. Fougerousse, T. Centner, B. Kolmerer, C. Witt, J. S. Beckmann, C. C. Gregorio, H. Granzier, and S. Labeit. 2000. Series of exon-skipping events in the elastic spring region of titin as the structural basis for myofibrillar elastic diversity. *Circ. Res.* 86:1114–1121.

Gautel, M., and D. Goulding. 1996. A molecular map of titin/connectin elasticity reveals two different mechanisms acting in series. *FEBS Lett.* 385:11–14.

Hayashi, C. Y., and R. V. Lewis. 1998. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. *J. Mol. Biol.* 275:773–784.

Helmes, M., K. Trombitas, T. Centner, M. Kellermayer, S. Labeit, W. A. Linke, and H. Granzier. 1999. Mechanically driven contour-length adjustment in rat cardiac titin's unique N2B sequence. Titin is an adjustable spring. *Circ. Res.* 84:1339–1352.

Higuchi, H., Y. Nakauchi, K. Maruyama, and S. Fujime. 1993. Characterization of B-connectin (Titin 2) from striated muscle by dynamic light scattering. *Biophys. J.* 65:1906–1915.

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

Labeit, S., M. Gautel, A. Lakey, and J. Trinick. 1992. Towards a molecular understanding of titin. *EMBO J.* 11:1711–1716.

Linke, W. A. 1996. Towards a molecular understanding of the elasticity of titin. *J. Mol. Biol.* 261:62–71.

Linke, W. A., and H. Granzier. 1998. A spring tale: new facts on titin elasticity. *Biophys. J.* 75:2613–2614.

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

Linke, W. A., M. R. Stockmeier, M. Ivemeyer, H. Hosser, and P. Mundel. 1998a. Characterizing titin's I-band Ig domain region as an entropic spring. *J. Cell Sci.* 111:1567–1574.

Muroga, Y. 2000. Applicability of broken-rodlike chain model to conformational analysis of polypeptide chain. *Biopolymers.* 54:58–63.

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

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

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

Rief, M., M. Gautel, A. Schemmel, and H. E. Gaub. 1998. The mechanical stability of immunoglobulin and fibronectin III domains in the muscle protein titin measured by atomic force microscopy. *Biophys. J.* 75: 3008–3014.

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

Trombitas, K., M. Greaser, G. French, and H. Granzier. 1998b. PEVK extension of human soleus muscle titin revealed by immunolabeling with the anti-titin antibody 9D10. *J. Struct. Biol.* 122:188–196.

Trombitas, K., M. Greaser, S. Labeit, J. P. Jin, M. Kellermeyer, M. Helmes, and H. Granzier. 1998a. Titin extensibility in situ: entropic elasticity of permanently folded and permanently unfolded molecular segments. *J. Cell Biol.* 140:853–859.

Urry, D. W. 1982. Characterization of soluble peptides of elastin by physical techniques. *Methods Enzymol.* 82:673–716.

Vinckier, A., and G. Semenza. 1998. Measuring elasticity of biological materials by atomic force microscopy. *FEBS Lett.* 430:12–16.

Xu, G., and J. S. Evans. 1999. Model peptide studies of sequence repeats derived from the intracrystalline biomineralization protein, SM50. I. GVGGR and GMGGQ repeats. *Biopolymers.* 49:303–312.

Zhang, B., G. Xu, and J. S. Evans. 1999. A kinetic molecular model of the reversible unfolding and refolding of titin under force extension. *Biophys. J.* 77:1306–1315.

Zhang, W., Q. Xu, S. Zou, H. Li, W. Xu, and X. Zhang. 2000. Single molecule force spectroscopy on *Bombyx mori* silk fibroin by atomic force microscopy. *Langmuir.* 16:4305–4308.