Stability and Folding Rates of Domains Spanning the Large A-Band Super-Repeat of Titin

Jared G. Head,* Ahmed Houmeida,† Peter J. Knight,‡ Anthony R. Clarke,* John Trinick,‡ and R. Leo Brady*
*Department of Biochemistry, University of Bristol, Bristol BS8 1TD, United Kingdom; †Department of Biology, University of Nouakchott, Nouakchott, Mauritania; ‡School of Biomedical Sciences, University of Leeds, Leeds, LS2 9JT, United Kingdom

ABSTRACT Titin is a very large (>3 MDa) protein found in striated muscle where it is believed to participate in myogenesis and passive tension. A prominent feature in the A-band portion of titin is the presence of an 11-domain super-repeat of immunoglobulin superfamily and fibronectin-type-III-like domains. Seven overlapping constructs from human cardiac titin, each consisting of two or three domains and together spanning the entire 11-domain super-repeat, have been expressed in *Escherichia coli*. Fluorescence unfolding experiments and circular dichroism spectroscopy have been used to measure folding stabilities for each of the constructs and to assign unfolding rates for each super-repeat domain. Immunoglobulin superfamily domains were found to fold correctly only in the presence of their C-terminal fibronectin type II domain, suggesting close and possibly rigid association between these units. The domain stabilities, which range from 8.6 to 42 kJ mol^{-1} under physiological conditions, correlate with previously reported mechanical forces required to unfold titin domains. Individual domains vary greatly in their rates of unfolding, with a range of unfolding rate constants between 2.6×10^{-6} and 1.2 s^{-1}. This variation in folding behavior is likely to be an important determinant in ensuring independent folding of domains in multi-domain proteins such as titin.

INTRODUCTION

Titin (also known as connectin) remains the largest protein known, with a polypeptide molecular mass of up to 3.7 MDa in human striated muscle and a length of ~1.3 μm (Maruyama et al., 1977; Wang et al., 1979). Comprising ~300 predominantly β-barrel domains and a number of regions of unknown structure, a single titin molecule stretches half the sarcomere in striated muscle where it is believed to play a dual role in both sarcomere assembly and in the subsequent production of force by the working system. Titin has been proposed to act as a molecular ruler, measuring out the length of the thick filament (Whiting et al., 1989) and may also control Z-disc assembly (Sorimachi et al., 1997; Young et al., 1998). Unlike in vitro where myosin readily self-assembles into filaments of varying length, in vertebrate muscle each filament comprises precisely 294 myosin molecules. Spanning the entire length of the thick filament, titin seems an obvious candidate for the regulation of thick filament assembly. It is present early in myogenesis (Schultheiss et al., 1990) before the muscle fiber acquires its striated nature, and hence would have the opportunity to regulate filament assembly. Further, A-band titin is constructed from a series of repeating domain motifs, forming a super-repeat with an expected periodicity of ~44 nm, which correlates with the 42.9-nm physical repeat formed by myosin and other proteins in the thick filament.

The titin gene is ~300 kb in size and is alternatively spliced to produce cardiac and skeletal titin, which are distinguished by a different number of tandem β-barrel repeats in the I-band (between 37 and 90, respectively). Human cardiac titin consists of 29,926 residues with a calculated molecular mass of 2993 kDa and is expressed from an open reading frame DNA sequence of 81 kb. There are 244 recognizable β-barrel domains of which 112 have been assigned to the immunoglobulin superfamily (Ig) and 132 to the fibronectin type III superfamily (Fn). These classes of domains are common in muscle proteins, having also been observed in twitchin, C-protein, X-protein, and telokin. A notable feature in titin is the presence of super-repeats of these domains. These super-repeats comprise motifs of the β-barrel domains repeated a number of times along a length of the molecule, and the homology between a specific domain and its corresponding domain at the same position in other repeats is higher than between the domain and its neighbors within the super-repeat. This suggests that duplication of the super-repeats is a more recent evolutionary event than the duplication of the individual domains (Kenny et al., 1999).

There are two types of A-band super-repeat in titin. Toward the M-line of the A-band are 11 copies of an 11-domain super-repeat in which the domains are arranged Ig-Fn-Fn-Ig-Fn-Fn-Fn-Ig-Fn-Fn-Fn. Further toward the Z-disk is a seven-domain super-repeat (Ig-Fn-Fn-Ig-Fn-Fn-Fn) that is repeated six times. The existence of the 11-domain super-repeat with the observed periodicity of the thick filament, 42.9 nm, reinforces the notion of close association with the filament and hence a role in filament assembly.

Scission of titin results in the loss of muscle passive tension or resistance to extension, which is essential to

Received for publication 21 February 2001 and in final form 15 May 2001.
Address reprint requests to Dr. R. Leo Brady, University of Bristol, Department of Biochemistry, School of Medical Sciences, University Walk, Bristol BS8 ITD, UK. Tel.: 44-117-9287436; Fax: 44-117-9288274; E-mail: L.Brady@bris.ac.uk.

0006-3495/01/09/1570/10 $2.00

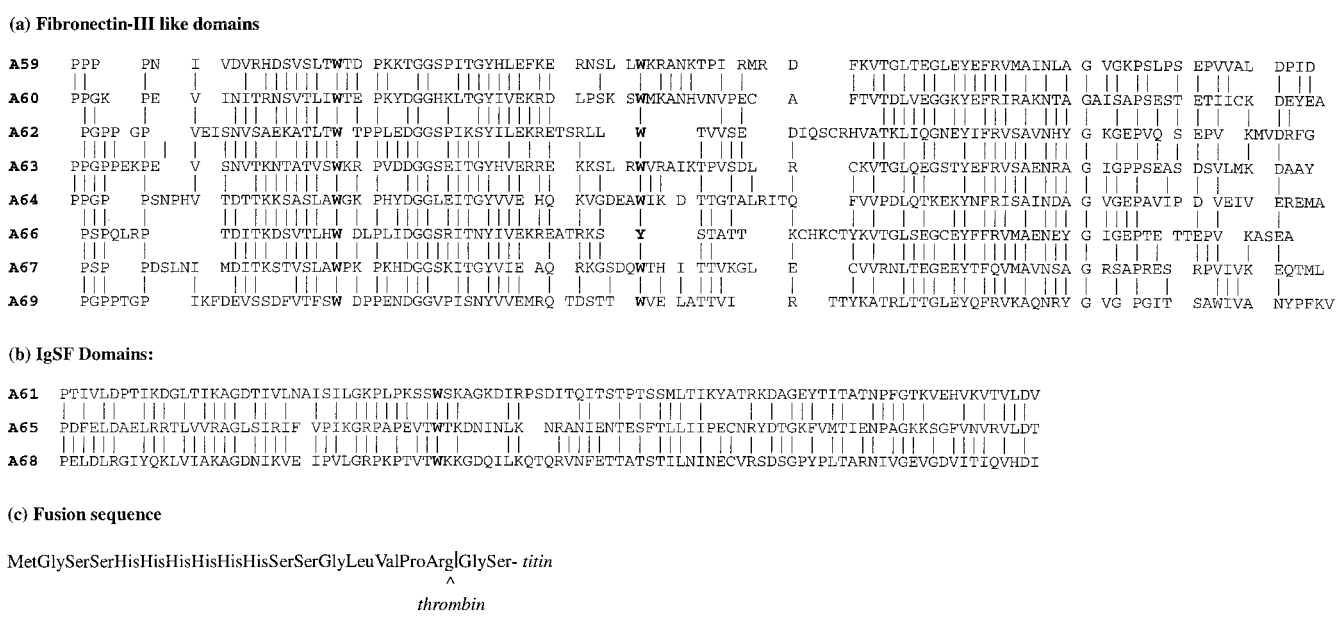

FIGURE 1 Alignment of the 11-domain polypeptide sequences from the A-band super-repeat region of human cardiac titin. The domains are grouped by homology: (*a*) fibronectin type III domains; (*b*) immunoglobulin superfamily domains. The black lines connect similar residues. The allocation of sequences to each domain as shown in this figure was used in the design of the constructs shown schematically in Fig. 2. (C) Sequence of fusion peptide added at the N-terminus when expressed in pET-15b. Thrombin cleavage removes the peptide to the left of the vertical bar.

prevent thick and thin filaments from sliding apart. Additionally, as the force produced by the thick filament is proportional to the number of cross-bridges it is able to form with the neighboring thin filaments, it has been proposed that titin acts to keep the thick filament centered in the middle of the sarcomere (Horowits and Podolsky, 1987). Elasticity of titin molecules is essential to this role.

The molecular mechanism of elasticity involves first straightening of the I-band part of the molecule, followed by unfolding of the polypeptide. Unfolding occurs first in a small region of unknown structure, the PEVK region, perhaps followed by Ig domains. The mechanism has been studied in situ by monitoring epitope movement and in single molecules by atomic force microscopy (AFM) and optical tweezers (Soteriouet al., 1993a,b; Erickson, 1994; Gautel and Goulding, 1996; Rief et al., 1997; Tskhovrebova et al., 1997; Kellermayer et al., 1997). Under extreme stretches the A-band part of the molecule may detach from the thick filament and also unfold.

In this study we have focused on the role of the titin A-band 11-domain super-repeat. We report the expression in bacteria of a series of recombinant constructs comprising overlapping fragments from human cardiac titin extending from domain A59 (using the nomenclature of Labeit and Kolmerer, 1995) in the second 11-domain super-repeat in the A-band to domain A69 in the third super-repeat. From folding studies we have determined the unfolding constants for each of the domains within the super-repeat. These values correlate with previous unfolding rate constants reported by others for similar domains from titin (Politou et al., 1994, 1995; King, 1994; Plaxco et al., 1997). We show that in this sequential series of domains there is significant variation in unfolding rates between neighboring domains, a feature that is likely to be important to minimize misfolding events in this multi-domain protein. Binding studies have also enabled us to define the binding epitope for the A-band-specific CH11 monoclonal antibody, which aids the interpretation of previous electron microscopy observations suggesting in vivo titin lies in an extended conformation along the length of the thick filament.

MATERIALS AND METHODS

Preparation of titin constructs

Domain boundaries were allocated from the amino acid sequence on the basis of several repeating motifs: a proline-rich (PGPP) motif was used to define the N-terminus of the fibronectin domains (Labeit et al., 1990), and the immunoglobulin domains were defined by a PXXXL (where X represents any amino acid) motif at the N-terminus and VLD motif at the C-terminus. The allocation of domain boundaries is shown in Fig. 1.

Each expressed fragment comprised either two or three domains, being designed such that each domain in the super-repeat was expressed at least once, with sufficient overlap to ensure a binding site could be attributed unambiguously to a single domain. The seven fragments expressed are illustrated schematically in Fig. 2.

Two cDNA fragments coding for an 11-domain section of human cardiac titin within the 11-domain super-repeat area of the A-band (between domains A59 and A69; nomenclature of Labeit and Kolmerer, 1995) were kindly donated by Siegfried Labeit (EMBL, Heidelberg, Germany). Polymerase chain reaction (PCR) primers were designed to incorporate a *Nde*I site (CATATG) upstream from, and in the same reading frame as, the coding DNA for each construct, and a *Bam*HI site (GGATCC) was in-

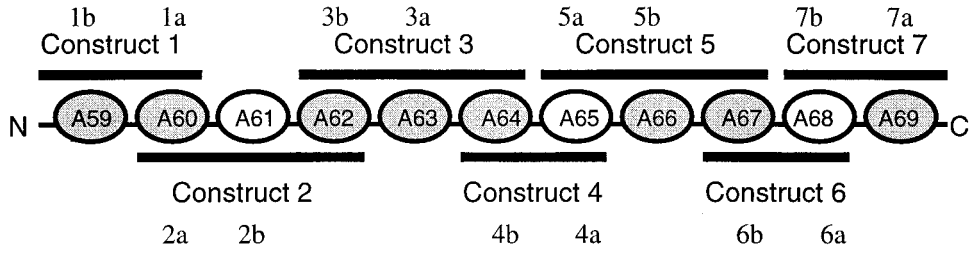

FIGURE 2 Truncated titin A-band constructs. Schematic showing the division of the 11-domain super-repeat region from the A-band of human cardiac titin. Fibronectin type III domains are shaded in gray, and immunoglobulin superfamily domains are unshaded. The black bars represent the expressed constructs, which include each of the domains adjacent to these bars. The amino and carboxyl ends of the super-repeat segment are labeled N and C, respectively. The folding event assignments have been annotated above or below the respective domains.

cluded downstream from the stop codon. Amplified segments of coding DNA were ligated directly into the pGEM-T holding vector (Novagen, Madison, WI) using the cloning host cell line XLI-Blue (New England Biolabs, Beverly, MA) and then inserted into the expression plasmid pET15b (Novagen), which leads to expression with a 6-histidine tag connected by a thrombin-susceptible sequence to the N-terminus side of the cloned protein. Recombinant proteins were expressed after induction with 1 mM isopropyl-β-D-thiogalactopyranoside using the BL21(DE3) (Novagen) host cell line. Pelleting experiments showed that the expressed protein was present in roughly equal amounts in both cell lysate and in insoluble inclusion bodies, and protein recovery was optimized by resuspending frozen cell pellets in 6 M guanidinium hydrochloride (GuHCl), 0.5 M NaCl, 50 mM Tris-HCl, pH 7.9, and passing over nickel-chelating Sepharose (Pharmacia, Uppsala, Sweden) equilibrated with the same buffer. Bound proteins were refolded on the column by introduction of the same buffer without the GuHCl denaturant and eluted with 50 mM EDTA in 0.5 M NaCl, 50 mM Tris-HCl, pH 7.9. The EDTA was then removed by dialysis against 0.5 M NaCl, 50 mM Tris-HCl, pH 7.9. Gel-filtration analysis showed there to be no significant aggregation between proteins refolded on the column in both the presence and absence of mercaptoethanol. His tags were cleaved by the addition of thrombin activated by 1.5 mM $CaCl_2$. Thrombin was removed by a benzamidine Sepharose column, and the cleaved His tag portions were removed by ultra-filtration.

Molecular interactions between titin and the CH11 antibody

Interactions between the titin constructs and the monoclonal antibody CH11 (Whiting et al., 1989) were investigated using an enzyme-linked immunosorbent assay. Titin was kindly prepared by Roger Starr and was purified as previously described (Soteriou et al., 1993a). Target proteins in phosphate-buffered saline were absorbed onto plates. Wells were washed with 140 mM NaCl, 2.6 mM KCl, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4, 0.1% Tween, pH 7.5, and then blocked using 1% bovine serum albumin in the wash buffer. After addition of the primary antibody and further washing, bound antibodies were detected by the addition of a secondary anti-mouse-horseradish peroxidase conjugate. Addition of a color reagent (100 mg/ml tetramethylbenzidine in 24.3 mM citric acid, 51.4 mM Na_2 HPO_4, pH 5.0, which was stopped by 2.5 M sulfuric acid) enabled quantitation of binding by measurement of absorbance at 414 nm using a Lab Systems Multiscan Plus plate reader.

Folding characterization of titin constructs

Circular dichroism (CD) spectra were recorded under O_2-free N_2 in a Jobin Yuon CD6 spectrometer. Spectra from protein samples at a number of different concentrations ranging between 1 μM and 30 μM in 5 mM NaH_2PO_4/Na_2HPO_4 buffer (pH 7.5) were recorded and combined by averaging.

Fluorescence spectra under equilibrium conditions were measured using a Perkin-Elmer LS 50B luminescence spectrometer with 1-cm pathlength and emission and excitation slits set between 5 and 10 nm depending on the intensity of signal obtained. Tryptophan fluorophores were excited at 290 nm and emission spectra collected between 300 and 450 nm. Constructs were equilibrated at a range of GuHCl concentrations from 0 to 3 M activity. The protein concentration was between 0.87 and 5.2 mM depending on the construct used, and solutions were buffered using 50 mM Tris at pH 7.5. As construct folding and unfolding occurred over a relatively long timescale, after the addition of GuHCl, samples were allowed to reach an equilibrium overnight. Equilibrium constants were calculated by obtaining the best fit of the unfolding data to the equation:

$$\alpha = \frac{1}{1 + (K_w\, e^{\mathrm{meqD}})},$$

where α is the fraction of unfolded protein, K_w is the equilibrium constant of unfolding at 0 M GuHCl, m_{eq} is the dependence of free energy of unfolding on GuHCl, and D is the activity of GuHCl. The free energy of folding, ΔG_f, was then calculated:

$$\Delta G_f = RT \ln K_w$$

Observed data were fitted to equilibrium curves using the nonlinear least-squares regression software Grafit.

Unfolding transients were also measured by tryptophan fluorescence, using the same fluorimeter to study the slower part of the unfolding transients (>50 s) and an Applied Photophysics DX17-MV stopped-flow fluorimeter to examine that part of the transient that occurred in less than 50 s. Titin fragments were mixed rapidly with varying amounts of GuHCl in 50 mM Tris buffer (pH 7.5) to give GuHCl concentrations ranging between 1.58 and 3.33 M. The excess varied between 10- and 50-fold in fluorimeter experiments and was 10-fold in stopped-flow experiments, and the final protein concentration was between 0.24 and 3.0 mM depending on the construct. The rate constant of unfolding at a specific concentration was calculated by fitting experimental data to an exponential decay equation, where the number of exponentials used was varied until a good fit, as determined by a random distribution of residuals, was obtained. When the log of the observed rate is plotted against GuHCl activity a straight line is obtained that can be extrapolated to the y axis to calculate the rate of unfolding at 0 M GuHCl, k_u, and whose gradient, m_u, gives the dependence of rate of unfolding on GuHCl. Transient data were fitted to exponential curves using the nonlinear least-squares fitting algorithm in Grafit, which was also used to calculate thermodynamic parameters.

RESULTS

DNA sequencing from both 3′ and 5′ termini verified that the appropriate section of titin encoding DNA had been inserted in the correct orientation for each of the seven constructs. Additionally, the plasmids were amplified directly from colonies of transformed BL21(DE3) cells using PCR with a range of primers, and in each case only the correct pair of primers amplified DNA, with the size of the amplified DNA corresponding to the expected size of the insert. All seven expressed titin fragments were efficiently purified using the nickel affinity column to greater than 95% purity as measured by staining of SDS-polyacrylamide gel electrophoresis gels, which also confirmed the expected sizes of each construct. Histidine tags were enzymatically removed from all constructs before the folding studies described below.

Molecular interactions between titin and the CH11 antibody

Fig. 3 *A* shows the binding of the monoclonal antibody CH11 to whole titin. The interaction of each of the titin constructs with CH11 is shown in Fig. 3 *B*, which demonstrates clearly that only construct 3 displays binding activity. When the overlap between constructs 2, 3, and 4 (Fig. 2) and the absence of CH11 binding to constructs 2 and 4 is considered, it can be concluded that the CH11 epitope is located on the mid-domain of construct 3, domain A63 by the nomenclature of Labeit and Kolmerer (1995). Although there is some evidence that the neighboring domains are not well folded in some of the constructs (see later), as many antibodies can bind their epitope when it is unfolded, it seems unlikely this would significantly distort the binding data.

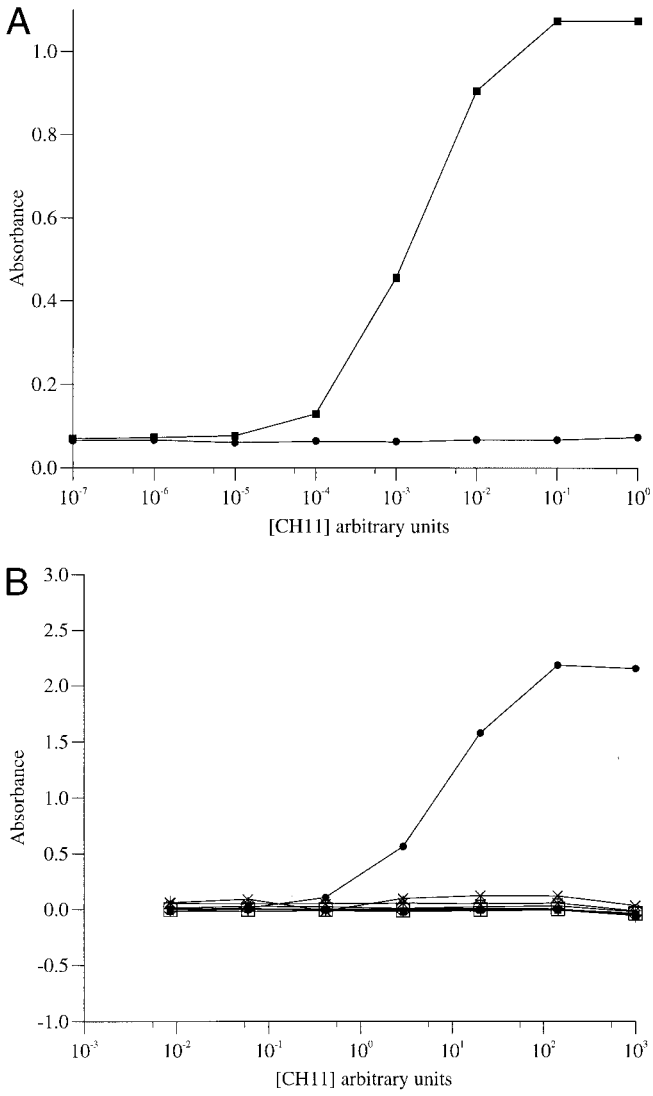

FIGURE 3 Molecular interactions of titin constructs with the CH11 monoclonal antibody. (*A*) Graph showing binding of the CH11 monoclonal antibody to titin (■) and BSA (●). Titin (0.2 mg/ml) and BSA (0.2 mg/ml) were first coated on the plate, CH11 was the primary antibody, and the secondary antibody was a 1 in 2000 dilution of goat anti-mouse Ig/horseradish peroxidase conjugate. The vertical axis represents absorbance at 414 nm; the horizontal axis is concentration of CH11 (arbitrary units representing dilutions of stock solution). (*B*) Graph showing binding of CH11 to titin constructs. Plates were first coated with the constructs (5 μg/ml), after which conditions were as described in *A* above. Only construct 3 (●) is seen to bind CH11. The vertical axis represents absorbance at 414 nm (corrected for background absorbance in the absence of CH11); the horizontal axis is concentration of CH11 (arbitrary units).

Folding characterization of titin constructs

The β-sheet proteins exhibit a CD spectrum with a minimum at ~216 nm, whereas random coiled proteins typically show a minimum of greater magnitude at ~195 nm (Johnson, 1990). Table 1 shows the wavelength of the minimum for each construct. Fig. 4 shows representative CD spectra for constructs 3 and 6.

Emission fluorescence spectra were recorded for each of the constructs under equilibrium conditions, and a typical series of spectra (from construct 1) is shown in Fig. 5 *A*. The degree of unfolding on addition of GuHCl, which was accompanied by a redshift in the fluorescence, was calculated by integrating the emission intensity between 325 and 330 nm, as shown for the representative fitted curve for construct 1 in Fig. 5 *B*. The values of K_w, m_{eq}, and ΔG_f derived from these plots are listed in Table 1.

It was hoped that the equilibrium studies could be deconvoluted to show the number of unfolding events in each construct, such that the degree of folding of each construct could be deduced. However, although two distinct events were visible for constructs 2 and 6, for the remainder of the constructs the equilibria of the individual events proved too close together for them to be separated usefully. Therefore, the constructs other than 2 and 6 were fitted with single unfolding curves, giving an average equilibrium constant for the superimposed events. Constructs 2 and 6 were fitted

TABLE 1 Equilibrium unfolding of titin constructs

Construct number	K_w	m_{eq}	ΔG_f (kJ mol^{-1})	CD minima (nm)
1	5900 ± 1100	−6.2 ± 0.1	−21 ± 0.4	213
2(i)	35 ± 7	−5.6 ± 0.3	−8.6 ± 0.5	202
2(ii)	$2.7 \times 10^6 \pm 1.0 \times 10^6$	−10.0 ± 0.2	−36 ± 1	
3	1400.0 ± 400	−5.9 ± 0.3	−17 ± 0.8	200
4	170 ± 30	−4.4 ± 0.1	−12 ± 0.4	212
5	1300 ± 500	−4.4 ± 0.2	−17 ± 1	208
6(i)	1700 ± 500	−9.3 ± 0.4	−18 ± 0.8	214
6(ii)	$8.1 \times 10^7 \pm 7.2 \times 10^7$	−10.7 ± 0.7	−42 ± 3	
7	190 ± 40	−4.3 ± 0.2	−13 ± 0.5	212

Results of equilibrium unfolding experiments showing the equilibrium constant of unfolding at 0 M GuHCl (K_w), the dependence of free energy of unfolding on GuHCl (m_{eq}), and the calculated free energy of folding (ΔG_f). Wavelengths of CD minima for each titin construct are also shown in the final column.

to double unfolding curves, but it should be noted that the increase in number of parameters to be fitted has led to a decrease in the accuracy of these parameters as is demonstrated by the unusually high m_{eq} values obtained, and particularly in the large error in K_w for event 6b.

As the contributions of the individual domains could not be distinguished from the equilibrium data it was necessary to undertake kinetic unfolding experiments. For each of the constructs two clear unfolding transients could be seen on analysis of the kinetic data. For constructs with three domains, attempts to fit data to three exponentials produced no significant improvement in correlation, whereas these data were readily interpreted by two-exponential fitting. Representative transients (construct 1) are shown in Fig. 6. Extrapolation from the observed rates was used to derive an unfolding rate at 0 M GuHCl activity (k_u), and these values are summarized in Table 2.

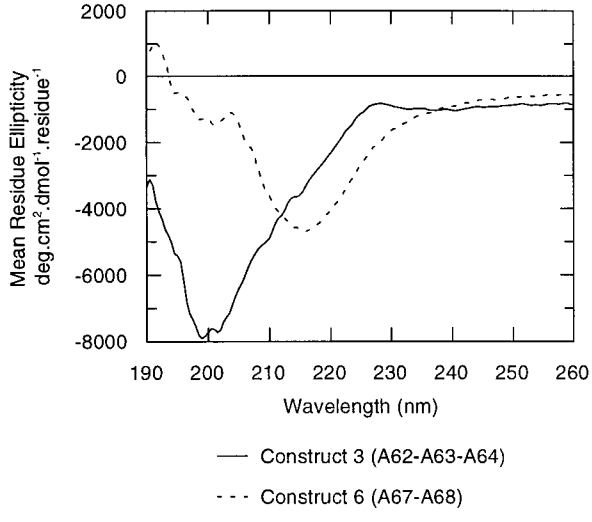

FIGURE 4 Representative CD spectra of titin constructs: construct 3 (——) and construct 6 (- - - -). In each case the protein samples were 2 μM in 5 mM phosphate buffer, pH 7.5; path length = 1 cm.

Allocation of unfolding events to individual domains

A notable feature of the transient unfolding experiments is the observation of two separate unfolding events for each construct (Table 2). One interpretation is that in each of the constructs there are only two domains correctly folded. Thus, each of the four double-domain constructs numbers 1, 4, 6, and 7, are all correctly folded, whereas each of the triple-domain constructs, namely, 2, 3, and 5, have two domains folded and one domain poorly folded. This interpretation is supported by the CD spectroscopy where, in general, a CD minimum of 216 nm would be expected for a protein comprised exclusively of β-sheets, and a minimum of 195 nm for random coil, or unfolded, segments of protein. Using this approximation, a significant content of β-sheet (which we interpret as correct folding) is expected in all of the two domain constructs, for which CD minima above 210 nm are observed (Table 1). In contrast, the minima of 208 nm for construct 5 and especially 202 nm and 200 nm for constructs 2 and 3 are indicative of a larger degree of random coil present in these constructs. This is consistent with our interpretation of the transient unfolding data that in each of these three constructs one domain is poorly folded. It is not, however, possible to reliably quantitate the degree of unfolding from these data. From the rate data it can be seen that the unfolded domains in the trimers are folded in the dimers, which fortuitously simplifies the assignation of domain unfolding rates.

Because of the overlap between neighboring constructs, the unfolding of some of the domains can be observed in two different constructs. Assuming that the unfolding kinetics of one domain are not affected by neighboring domains, this enables the event to be allocated to a specific domain in the super-repeat, and by a process of elimination it is also possible to deduce the locations of the events not seen as overlaps. In the allocation of data to individual domains, one domain in construct 6 that becomes more fluorescent on unfolding is a useful starting point. This domain (6b) un-

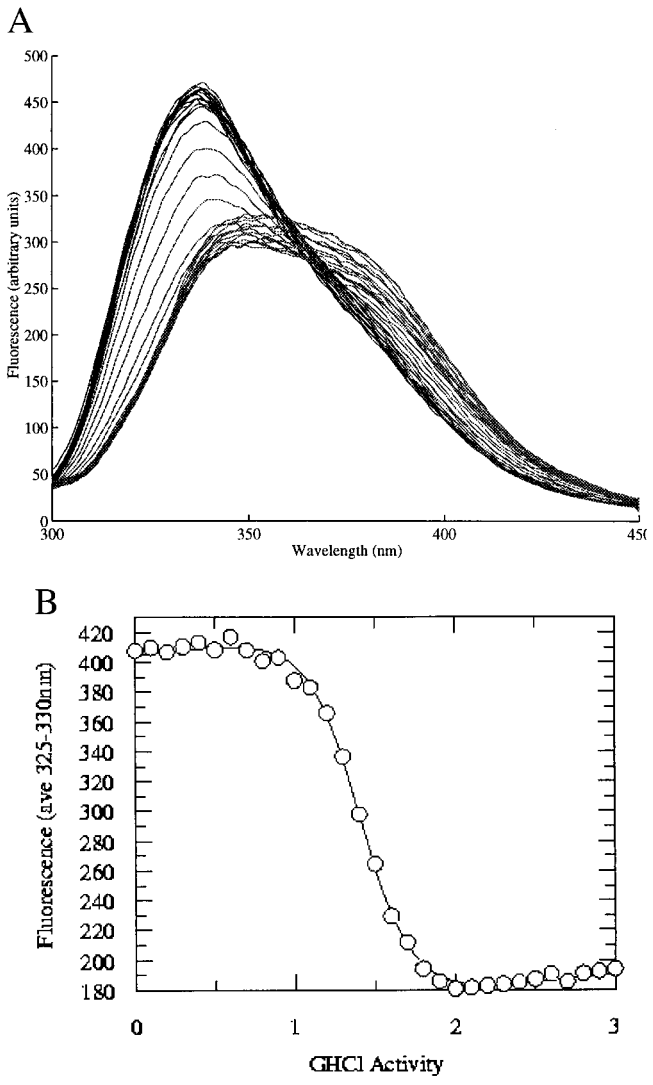

FIGURE 5 Equilibrium unfolding data for construct 1. (*A*) Fluorescence emission spectra obtained as described in the text. The protein concentration was 2.9 mM, and λ_{ex} was 285 nm. Slits were at 5 nm for both excitation and emission. (*B*) Graph showing the degree of unfolding as calculated by averaging of emission spectra between 325 nm and 330 nm. The fitted curve has been modeled as a single unfolding curve and corresponds to $K_w = 5870$, $m_{eq} = -6.19$.

folds faster than either of the unfolding events observed for construct 7, whereas event 6a matches 7b. This enables allocation of these unfolding events to the last three domains in Fig. 2 (A67 = 6b, A68 = 6a and 7b, and A69 = 7a). Neither of the unfolding events observed for construct 5 overlap with construct 6, but 5a correlates with 4a, leading to the conclusion that A66 = 5b, A65 = 5a and 4a, and A64 = 4b. As neither of the observed events for construct 3 match 4b (domain A64) it must be this domain in construct 3 that is unfolded (A62 and A63 = 3a and/or 3b). Construct 2 is more ambiguous, with 2a matching either 1a or 3b. If 1a and 2a are matched, construct 2 has its C-terminal domain unfolded; if 3b and 1a are equated then it

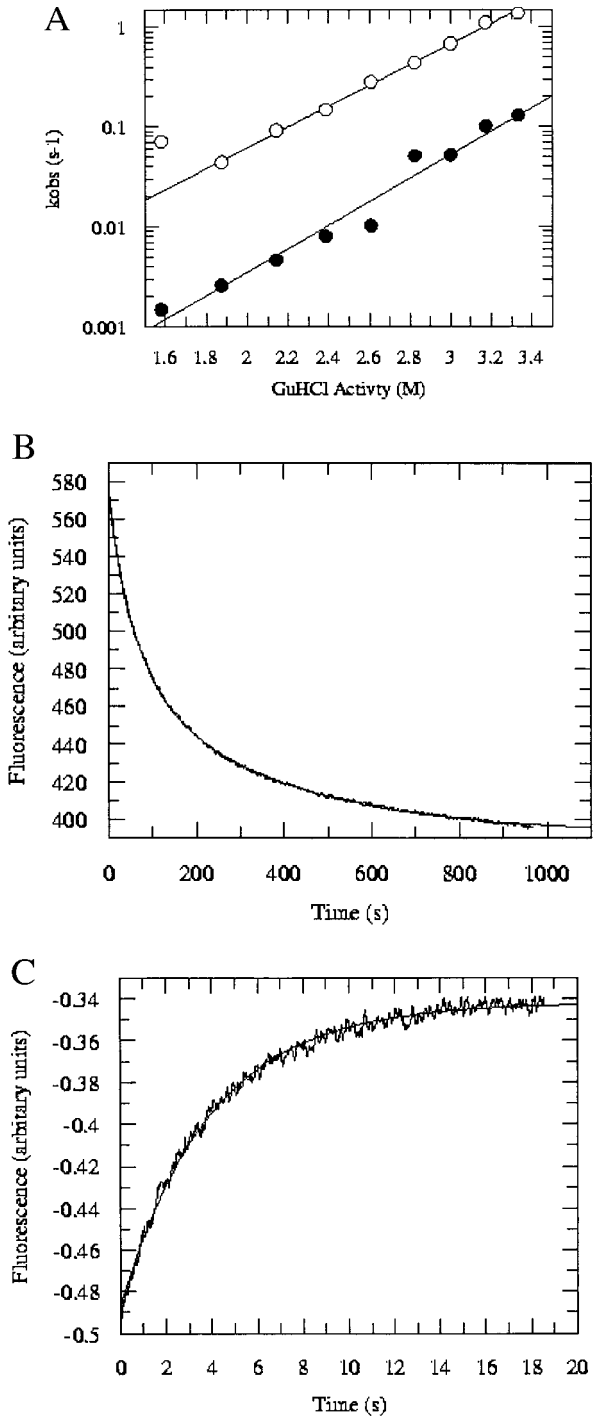

FIGURE 6 Transient unfolding data for construct 1. (*A*) Observed rate constants obtained using a Perkin-Elmer fluorimeter (slow unfolding event, ●) and stopped-flow fluorimeter (fast unfolding event, ○). The data have been fitted by linear regression and give $k_{u1} = 1.7 \times 10^{-5}$, $m_{k1} = 1.2$, $k_{u2} = 5.1 \times 10^{-4}$, and $m_{k2} = 1.0$. (*B*) Representative fluorimeter transient data for construct 1. The final concentration of protein is 0.57 mM, λ_{ex} was 290 nm, λ_{em} was 327 nm, and slits were adjusted to 15 nm for excitation and 10 nm for emission. The data have been fitted to a double-exponential curve. (*C*) Representative stopped-flow transient data for construct 1. The final protein concentration is 2.4 μM, λ_{ex} was 285 nm, and λ_{em} was >360 nm. The data have been fitted to a single-exponential curve, representing the single fast unfolding event.

TABLE 2 Transient unfolding of titin constructs

Construct number	k_u (s^{-1})	m_k
1 a	$(1.7 \pm 0.78) \times 10^{-5}$	1.2 ± 0.08
b	$(5.1 \pm 0.64) \times 10^{-4}$	1.0 ± 0.02
2 a	$(1.8 \pm 0.95) \times 10^{-5}$	1.2 ± 0.1
b	$(6.5 \pm 1.8) \times 10^{-2}$	0.57 ± 0.05
3 a	$(1.7 \pm 0.62) \times 10^{-6}$	1.2 ± 0.06
b	$(2.0 \pm 0.74) \times 10^{-5}$	0.83 ± 0.18
4 a	$(2.0 \pm 0.93) \times 10^{-5}$	0.91 ± 0.54
b	$(7.4 \pm 2.7) \times 10^{-4}$	0.57 ± 0.10
5 a	$(8.8 \pm 5.8) \times 10^{-5}$	1.1 ± 0.05
b	$(1.5 \pm 1.1) \times 10^{-2}$	0.71 ± 0.13
6 a	$(6.9 \pm 4.9) \times 10^{-5}$	0.85 ± 0.15
b	$(6.9 \pm 2.4) \times 10^{-4}$	0.97 ± 0.06
7 a	$(2.6 \pm 1.0) \times 10^{-6}$	1.2 ± 0.07
b	$(2.3 \pm 1.6) \times 10^{-4}$	0.77 ± 0.15

Results of transient unfolding analysis showing the rate of unfolding at 0 M GuHCl (k_u) and dependence of rate of unfolding on GuHCl (m_k) for each unfolding event, labeled a for the slower event of each construct and b for the faster event.

is the N-terminal domain that is incorrectly folded. Because in the case of the other two incompletely folded constructs it is the C-terminal domain that is unfolded, it seems more probable that this is also the case in construct 2. The most plausible allocation of unfolding events is annotated to the domains shown in the schematic in Fig. 2.

DISCUSSION

Molecular interactions of titin

In electron micrographs, the monoclonal antibody CH11 has previously been noted to form a distinctive stripe in the A-band when bound to muscle fibers (Whiting et al., 1989). From the binding data in this study we have localized the CH11 epitope to domain A63. In terms of distances along the titin molecule, as each domain is ~4 nm in length, positioning of the CH11 epitope on domain A63 provides a reference location for the antibody stripe to ±2 nm, a useful measurement when related to an overall molecule length of 1 μm.

Electron microscopy shows that CH11 is bound 10 nm closer to the Z disk than stripe 11 of the 11 non-myosin stripes in the A-band (Whiting et al., 1989). Comparing these results with those of Freiburg and Gautel (1996), where the binding site of C-protein is allocated to the first Ig domain of the super-repeat (A65), suggests that the two binding sites are located two domains apart, or $\sim 2 \times 4$ nm, which is similar to the separation reported in the electron microscopy studies. The CH11 site is also closer to the Z disk as would be expected for an epitope closer toward the N-terminus of titin. Isolated titin molecules have been observed by electron microscopy to adopt a range of conformations, from fully extended to globular structures (e.g., Houmeida et al., 1995), consistent with the individual domains forming a tandem array like beads on a string. However, despite the expectation that titin would adopt a linear, extended conformation within a muscle fiber, this has not been demonstrated experimentally. The correspondence between the calculated linear separation of the CH11 stripes in electron micrographs of muscle filaments and the actual distance observed on the thick filament supports the theory that titin is arrayed straight alongside the thick filament. Titin must therefore adopt an extended, linear conformation, rather than being coiled or bunched, at least in this region of the protein. This extended conformation for titin in vivo is consistent with the notion of titin acting as a molecular ruler in the assembly of muscle filaments.

Folding integrity of expressed proteins

Our results from both the CD spectroscopy of the constructs and the transient unfolding experiments indicated a single domain in each of the three-domain constructs did not fold correctly. As each of the three incompletely folded domains was inferred to be located at the C-terminus of their respective three-domain constructs (Fig. 2), one explanation might be the incorrect allocation of domain C-terminal boundaries. However, the NMR structure of the A71 Fn domain (equivalent to A60 in constructs 1 and 2), which was reported after the design of these constructs (Goll et al., 1998), confirms that our assignment of the PGPP motif for the N-terminus of the Fn domains is appropriate. A more likely explanation is that the folding of several domains (A62, A64, and A67) is dependent on the presence of their neighboring C-terminal domains, or even other molecules. If the requirement is provided by the adjacent titin domain, this implies a close association between these pairs, namely A62-A63, A64-A65, and A67-A68, perhaps similar to the rigid associations observed between the D1D2 and D3D4 IgSF pairs of CD4 (Wang et al., 1990; Ryu et al., 1990; Brady et al., 1993). This would be expected to have a considerable impact on the relative rigidity of these sections of the A-band part of titin. The B/C and F/G loops of the β-sandwich structure are expected to contact neighboring domains if domain pairs are closely associated in tandem arrays. Modeling analysis of I-band domains (Fraternali and Pastore, 1999) has been used to distinguish between flexibility of this region and other sections of the titin molecule. Our analysis suggests inter-domain flexibility is likely to vary across the 11-domain super-repeat, perhaps influencing the overall conformation of this region.

Implications of unfolding rate constants

The unfolding and re-folding of β-sandwich domains similar to those found in the A-band have been proposed to play a role in the extreme elasticity observed for titin (Soteriou et al., 1993a; Erickson, 1994; Gautel and Goulding, 1996; Rief

TABLE 3 Stability of titin and similar domains

Type	Domain	ΔG_u (kJ mol^{-1})
Titin Ig domains	Ab1	13.8*
	Ab2	17.5
	M11	18.0
	I11	12.6†
	I27	30.8
	I28	16.2
	I29	10.7
	I30	18.5
	M5	42.6
Average titin domain		41‡
		13§
CD2	1	26.4†
Fn III	^{9}FnIII	5.0¶
	^{10}FnIII	25.6
Titin constructs	1	21
	2a	36
	2b	8.6
	3	17
	4	12
	5	17
	6a	18
	6b	42
	7	13

Comparison of calculated free energy of unfolding (G_u) values obtained for titin constructs (from Table 1) with published values for comparable domains.
*Politou et al. (1994).
†Politou et al. (1995).
‡Soteriou et al. (1993).
§King (1994).
¶Plaxco et al. (1997).

et al., 1997; Tskhovrebova et al., 1997; Kellermayer et al., 1997). Although the current consensus is that I-band rather than A-band domains are more likely to be directly associated with the in vivo elasticity of titin, mechanical unfolding of A-band domains in response to tension has been demonstrated using recombinant constructs (Tskhovrebova et al., 1997), and the similarity of domain structure in both regions suggests a common response toward applied tension should the molecule become detached from the thick filament under extreme stress. The values obtained in this study for free energies of unfolding of the A-band constructs are of the same order as those reported for similar molecules (Table 3). The range of values obtained (8–42 kJ/mol) seems fairly large for such structurally similar domains, but perhaps, if unfolding of these domains mimics similar events elsewhere along the titin molecule, it reflects a mechanism for achieving elasticity over a range of tensions. The rates of unfolding, and consequently refolding, of titin domains vary significantly, but in general are relatively slow (Table 2). These rates are of similar order to the previously reported values of 3×10^{-5} s^{-1} for the unfolding of a titin IgSF domain obtained using AFM (Schultheiss et al., 1990) and 2.8×10^{-4} s^{-1} for a twitchin IgSF domain (Fong et al., 1996). This is surprising when it would seem to be more useful for titin to be able to recover its elasticity as quickly as possible, but it is consistent with the hysteresis exhibited by stretched muscle fiber where after extension the elastic ability takes time to recover. This result also corresponds to the findings of Rief et al. (1997), which show only partial refolding of titin domains within 1 s after mechanical unfolding.

The amount of mechanical work corresponding to the derived values for free energies of unfolding can be calculated. Because energy = force × distance the ΔG_u can be translated into a force required to unfold the domain. The nature of the unfolding event means that instead of a constant force being applied over the course of the extension from completely folded to completely unfolded, there is an energy barrier at the start of the extension where the secondary and tertiary structure of the domain are disrupted, so the bulk of the work is done in the first short distance of extension. The unfolding potential is calculated by Rief et al. (1997) to have a width, $\Delta x = 0.3$ nm, for a representative titin Ig domain. The force required to extend a single domain can be therefore be estimated from the equation:

$$F = \frac{\Delta G_f}{\Delta x \times N_a},$$

which gives a range of forces from 44 pN for ΔG_u of 8 kJ/mol to 232 pN for ΔG_u of 42 kJ/mol. These values agree well with the range obtained by AFM of 150–300 pN. A wide variety of forces required to unfold the individual domains in titin would seem to be useful so that it remains elastic over a broad range of stresses, and it is possible that the varying values seen in this super-repeat are representative of the molecule as a whole.

A striking feature of the unfolding rates for the titin domains is the significant differences observed in most instances between one domain and both of its neighbors (Fig. 7). Protein β-sandwich domains are known to be prone to misfolding to intertwined structures when similar domains are folded in close proximity, as demonstrated for the N-terminal domain of CD2 (Murray et al., 1995). It has been suggested that the inter-sheet disulfide helps to minimize these events in extracellular IgSF domains (Murray et al., 1995), but this strategy is not available to intracellular domains such as those in titin. Gene duplication seems very likely to have played a role in the evolution of molecules such as titin with its unusually large number of sequential and structurally related domains. Intertwining of neighboring β-sheet structure needs to be avoided in both the synthesis and, if regular folding/unfolding events are part of the activity cycle, function of titin. Misfolding events of this type could be minimized by effectively refolding each domain in isolation. The folding studies reported here suggest two ways in which this might be achieved. First, the variation in the energy required to unfold individual domain types, although not highlighting any particular domain as

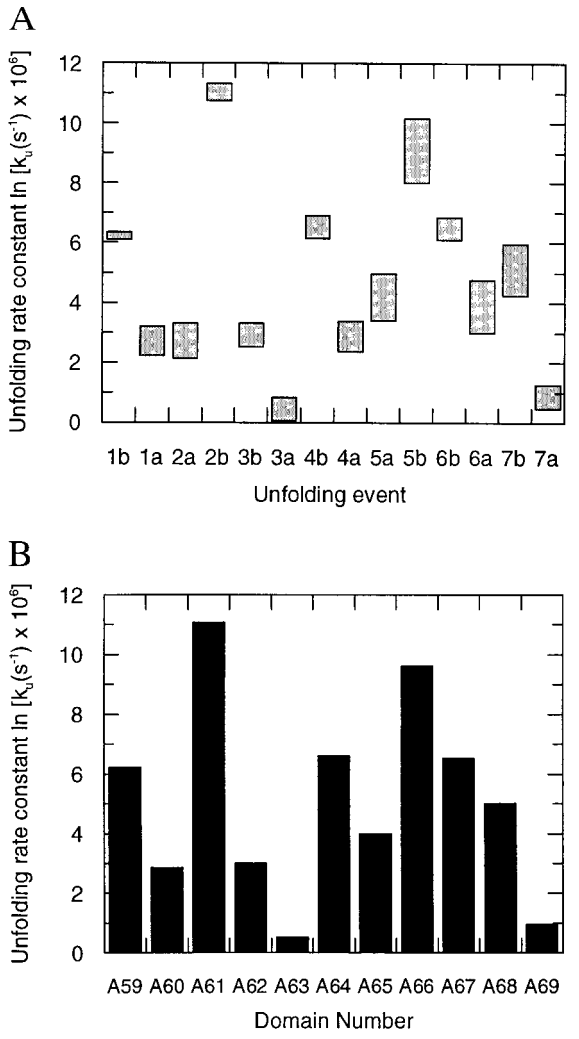

FIGURE 7 Unfolding rate constants for titin A-band domains. (*A*) Graph showing unfolding rate constants (k_u) for all unfolding events. In each case the shaded bars represent the range for each k_u value as reflected in the uncertainty values shown in Table 2. (*B*) Graphical representation showing the variation in the unfolding rate constants (k_u) between the 11 domains of the A-band super-repeat of human cardiac titin.

especially susceptible to unfolding, may stagger the unfolding of specific domains at separate locations, ensuring a correlation with the applied mechanical force. Second, the significant variation in unfolding rates, and hence refolding rates, should maximize the likelihood that neighboring domains are at different stages of the folding process at any one time. There is no correlation between folding stability and Ig or Fn domain classifications within the super-repeat, although the inter-dispersal of these differing topologies throughout the super-repeat may also act to minimize the likelihood of misfolding or intertwining. The attainment of individualized folding behavior for neighboring domains is likely to be an important goal immediately after domain duplication during the evolution of multi-domain proteins, in order to minimize misfolding events.

REFERENCES

Brady, R. L., E. J. Dodson, G. G. Dodson, G. Lange, S. J. Davis, A. F. Williams, and A. N. Barclay. 1993. Crystal structure of domains 3 and 4 of rat CD4: relationship to the NH2-terminal domains. *Science.* 260: 979–983.

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

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

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

Freiburg, A., and M. Gautel. 1996. A molecular map of the interactions between titin and myosin-binding protein-C: implications for sarcomeric assembly in familial hypertrophic cardiomyopathy. *Eur. J. Biochem.* 235:317–323.

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.

Goll, C. M., A. Pastore, and M. Nilges. 1998. The three-dimensional structure of a type I module from titin: a prototype of intracellular fibronectin type III domains. *Structure.* 6:1291–1302.

Horowits, R., and R. J. Podolsky. 1987. The positional stability of thick filaments in activated skeletal muscle depends on sarcomere length: evidence for the role of titin filaments. *J. Cell Biol.* 105:2217–2223.

Houmeida, A., J. Holt, L. Tskhovrebova, and J. Trinick. 1995. Studies of the interaction between titin and myosin. *J. Cell Biol.* 131:1471–1481.

Johnson, W. C. 1990. Protein secondary structure and circular dichroism: a practical guide. *Proteins.* 7:205–214.

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.

Kenny, P. A., E. M. Liston, and D. G. Higgins. 1999. Molecular evolution of immunoglobulin and fibronectin domains in titin and related muscle proteins. *Gene.* 232:11–23.

King, L. 1994. Effects of denaturant and pressure on the intrinsic fluorescence of titin. *Arch. Biochem. Biophys.* 311:251–257.

Labeit, S., D. P. Barlow, M. Gautel, T. Gibson, C. L. Holt, U. Hsieh, K. Francke, J. Leonard, A. Wardale, A. Whiting, and J. Trinick. 1990. A regular pattern of two types of 100-residue motif in the sequence of titin. *Nature.* 345:273–276.

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

Marszalek, P. E., H. Lu, H. B. Li, M. CarrionVazquez, A. F. Oberhauser, K. Schulten, and J. M. Fernandez. 1999. Mechanical unfolding intermediates in titin modules. *Nature.* 402:100–103.

Maruyama, K., S. Matsubara, R. Natori, Y. Nonomura, S. Kimura, K. Ohashi, F. Murakami, S. Handa, and G. Eguchi. 1977. Connectin, an elastic protein of muscle: characterization and function. *J. Biochem. (Tokyo).* 82:317–337.

Murray, A. J., S. J. Lewis, A. N. Barclay, and R. L. Brady. 1995. One sequence, two folds: a metastable structure of CD2. *Proc. Natl. Acad. Sci. U.S.A.* 92:7337–7341.

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

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

Politou, A. S., D. 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.

Ryu, S. E., P. D. Kwong, A. Truneh, T. G. Porter, J. Arthos, M. Rosenberg, X. P. Dai, N. H. Xuong, R. Axel, R. W. Sweet, and W. A. Hendrickson. 1990. Crystal structure of an HIV-binding recombinant fragment of human CD4. *Nature.* 348:419–426.

Schultheiss, J., Z. Lin, M.-H. Lu, J. Murray, D. A. Fischman, K. Weber, T. Masaki, M. Imamura, and H. Holtzer. 1990. Differential distribution of subsets of myofibrillar proteins in cardiac nonstriated and striated myofibrils. *J. Cell Biol.* 110:1159–1172.

Sorimachi, H., A. Freiburg, B. Kolmerer, S. Ishiura, G. Stier, C. C. Gregorio, D. Labeit, W. A. Linke, K. Suzuki, and S. Labeit. 1997. Tissue-specific expression and actinin binding properties of the Z-disc titin: implications for the nature of vertebrate Z-discs. *J. Mol. Biol.* 270:688–695.

Soteriou, A., A. R. Clarke, S. Martin, and J. Trinick. 1993a. Titin folding energy and elasticity. *Proc. R. Soc. Lond. B.* 254:83–86.

Soteriou, A., M. Gamage, and J. Trinick. 1993b. A survey of interactions made by the giant protein titin. *J. Cell Sci.* 104:119–123.

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

Wang, K., J. McClure, and A. Tu. 1979. Titin: major myofibrillar components of striated muscle. *Proc. Natl. Acad. Sci. U.S.A.* 76:3698–3702.

Wang, J. H., Y. W. Yan, T. P. J. Garrett, J. H. Liu, D. W. Rodgers, R. L. Garlick, G. E. Tarr, Y. Husain, E. L. Reinherz, and S. C. Harrison. 1990. Structure of a fragment of human CD4 containing two immunoglobulin-like domains. *Nature.* 348:411–418.

Whiting, A., J. Wardale, and J. Trinick. 1989. Does titin regulate the length of muscle thick filaments? *J. Mol. Biol*. 205:263–268.

Young, P., C. Ferguson, S. Bañuelos, and M. Gautel. 1998. Molecular structure of the sarcomeric Z-disk: two types of titin interactions lead to an asymmetrical sorting of -actinin *EMBO J.* 17:1614–1624.