© Springer 2005

Poly-Ig tandems from I-band titin share extended domain arrangements irrespective of the distinct features of their modular constituents

MARCO MARINO[1], DMITRI I. SVERGUN[2], LAURENT KREPLAK[1], PETER V. KONAREV[2], BOHUMIL MACO[1], DIETMAR LABEIT[3] and OLGA MAYANS[1,*]

[1]*Division of Structural Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056, Basel, Switzerland;* [2]*EMBL, Hamburg Outstation c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany & Institute of Crystallography, Russian Academy of Sciences, Leninsky pr. 59, 117333, Moscow, Russia;* [3]*Institut für Anästhesiologie und Operative Intensivmedizin, Universitätklinikum Mannheim, 68167, Mannheim, Germany*

Abstract

The cellular function of the giant protein titin in striated muscle is a major focus of scientific attention. Particularly, its role in passive mechanics has been extensively investigated. In strong contrast, the structural details of this filament are very poorly understood. To date, only a handful of atomic models from single domain components have become available and data on poly-constructs are limited to scarce SAXS analyses. In this study, we examine the molecular parameters of poly-Ig tandems from I-band titin relevant to muscle elasticity. We revisit conservation patterns in domain and linker sequences of I-band modules and interpret these in the light of available atomic structures of Ig domains from muscle proteins. The emphasis is placed on features expected to affect inter-domain arrangements. We examine the overall conformation of a 6Ig fragment, I65–I70, from the skeletal I-band of soleus titin using SAXS and electron microscopy approaches. The possible effect of highly conserved glutamate groups at the linkers as well as the ionic strength of the medium on the overall molecular parameters of this sample is investigated. Our findings indicate that poly-Ig tandems from I-band titin tend to adopt extended arrangements with low or moderate intrinsic flexibility, independently of the specific features of linkers or component Ig domains across constitutively- and differentially-expressed tandems. Linkers do not appear to operate as free hinges so that lateral association of Ig domains must occur infrequently in samples in solution, even that inter-domain sequences of 4–5 residues length would well accommodate such geometry. It can be expected that this principle is generally applicable to all Ig-tandems from I-band titin.

Introduction

The muscle sarcomere is the most complex biomachinery known to date. Its ultrastructure, mechanics, signaling and even development appear orchestrated to a significant extent by a colossal filament, titin/connectin (Maruyama *et al.*, 1977), housed within its lattice of molecular motors. During the past decade, our understanding of the cellular function of titin has advanced significantly, with a multitude of studies having addressed its role in passive mechanics, stretch-stress biomechanical pathways, myofibrillogenesis, its molecular ruler properties and its involvement in human disease (for recent reviews see Tskhovrebova and Trinick, 2003; Granzier and Labeit, 2004; Miller *et al.*, 2004). In strong contrast, only minimal progress has been made in elucidating the molecular and structural details of this giant protein. The building blocks of titin were initially identified in 1995 (Labeit and Kolmerer, 1995), when sequence data first became available, and were completed in 2001 once the domain composition of the different isoforms was established (Bang *et al.*, 2001). This revealed that most of the 38138 amino acids encoded by the human titin gene fold into immunoglobulin (Ig) and fibronectin-3 (Fn3) domains, which add up to ∼300 modules in the larger isoforms. They form mostly linear tandems distributed along the 1 μm length of the filament. Other important components of titin include the so-called PEVK region in the I-band (rich in proline, glutamate, valine and lysine residues), unique sequences – yet to be characterized – and one single catalytic domain, a kinase, located at the M-line.

Despite its numerous modular components, only a mere handful of atomic structures from domains of titin have been reported to date (Table 1). These 3D-models fall short of accounting for all the representative constituents of the filament. Structural information has also become available for motifs within the PEVK region (Ma *et al.*, 2001; Ma *et al.*, 2003), but the complex conformational dynamics of these unusual sequences is yet to be understood. Structural models from elements of invertebrate elastic filaments are even more meager. While two studies have addressed modules from twitchin (Table 1), no data are currently

* To whom correspondence should be addressed. Phone: +41-61-267-2083; Fax: +41-61-267-2109; E-mail: Olga.Mayans@unibas.ch

Table 1. Atomic structures of titin and twitchin components

Domain	Type	Location	Source	PDB code	Method	Reference
Titin						
M5	Ig	M-line	Human	1NCT/1NCU/1TNM/1TNN	NMR	(Pfuhl *et al.*, 1995)
I91 (I27)	Ig	Constitutive, distal I-band	Human	1TIT/1TIU	NMR	(Improta *et al.*, 1996)
A71	Fn3	A-band (11 domain super-repeat)	Human	1BPV	NMR	(Muhle-Goll *et al.*, 1998)
TK	Kinase	M-line	Human	1TKI	X-ray Cryst.	(Mayans *et al.*, 1998)
I1	Ig	Constitutive, proximal I-band	Human	1G1C	X-ray Cryst.	(Mayans *et al.*, 2001)
Z7:α-actinin	Z-repeat	Z-disc	Human	1H8B	NMR	(Atkinson *et al.*, 2001)
M1	Ig	M-line	Human	2BK8	X-ray Cryst.	(Müller *et al.*, unpublished)
Z1Z2	Ig–Ig	Z-disc	Human	2A38	X-ray Cryst.	(Marino *et al.*, unpublished)
Z1Z2:Telethonin		Z-disc	Human	1YA5	X-ray Cryst.	(Zou *et al.*, unpublished)
Twitchin						
I18	Ig	A-band	*C. elegans*	1WIT/1WIU	NMR	(Fong *et al.*, 1996)
I26	Kinase + Ig	A-band	*C. elegans*	1KOA	X-ray Cryst.	(Kobe *et al.*, 1996)

available on kettins, stretchins or projectins (for a review on these filaments see Bullard *et al.*, 2002). The shortage of structural information on titin affects, in particular, the understanding of its key role in muscle passive mechanics, which is mediated by its elastic I-band. The latter is primarily composed of serially-linked Ig domains and PEVK sequences. Data from immunoelectron microscopy supports a sequential extension model in which Ig-tandems extend at low forces due to a high bending rigidity while the PEVK segment straightens at higher force acting as a permanently unfolded chain (Labeit and Kolmerer, 1995; Gautel and Goulding, 1996; Linke *et al.*, 1996; Trombitas *et al.*, 1998). Titin isoforms vary greatly in their I-band composition, reflecting the specific mechanical properties of filament variants (Bang *et al.*, 2001). Distinct Ig-tandems can be identified in this region according to expression patterns, where two constitutive segments flank the beginning and end of the I-band, and a long central tandem together with the so-termed N2B and N2A regions are differentially expressed elements. The mechanical properties of these domain arrays in response to pulling-forces have been extensively studied (Kellermayer *et al.*, 1997; Rief *et al.*, 1997; Tskhovrebova *et al.*, 1997; Li *et al.*, 2002; Watanabe *et al.*, 2002). Interpretation of stretch-unfolding data in terms of molecular events has used molecular dynamics simulations on available atomic models (Li *et al.*, 2000; Li *et al.*, 2002; Li and Fernandez, 2003). Unfortunately, the fact that structures are currently available for only two Igs from the titin I-band, I1 (Mayans *et al.*, 2001) and I91 (formerly I27) (Improta *et al.*, 1996), has limited, and probably biased, this field of study – highlighting the need for more structural efforts on this filament.

Furthermore, understanding the dynamics of I-band tandems during muscle mechanics intrinsically requires knowledge of the supra-arrangement of domains along the filament. At this effect, data from NMR and small-angle X-ray scattering (SAXS) have been combined with homology modeling to examine the overall conformation of a double construct comprising I91–I92 (formerly I27–I28) from the distal I-band (Improta *et al.*, 1998). In the same study, the SAXS analysis was extended to a 4Ig construct spanning I91–I94 (I27–I30). Results showed that these fragments preferentially adopt extended conformations in solution, characterized by small domain interfaces. Motions proved to be restricted around the linker sequences assayed, suggesting limited molecular flexibility. A complementary report on I10–I11, I91–I92 and I92–I93 used thermodynamic, cross-linking and complexation approaches to analyze interactions between consecutive Ig within a tandem (Politou *et al.*, 1996). The study concluded that only weak interactions occur between modules, which was later confirmed by work on a set of related constructs derived from the I91–I96 region (Scott *et al.*, 2002). Taken together, available data suggest that a "beads-on-a-string" model of organization with limited local conformational freedom is applicable to Ig-tandems within the constitutive I-band.

In the current study, we revisit conservation patterns in domain and linker sequences of I-band modules of titin and interpret these in the light of available atomic structures for Ig domains from muscle proteins. The emphasis is placed on those features expected to affect the structural arrangement of domains in the filament. Given that studies reported to date have focused on the analysis of fragments from the constitutive I-band sections, we examine here a 6Ig construct from the differentially expressed I-band of soleus titin using SAXS and electron microscopy approaches. The possible role of a highly conserved linker group in defining the overall molecular shape of this poly-Ig is investigated by characterization of a triple mutated variant. Results are discussed in the context of I-band architecture.

Materials and methods

Cloning, protein expression and purification

Domains I65–I70 of soleus titin from rabbit (residues 7946–8511, human sequence at TrEMBL Q8WZ42) were cloned into pETM-11, a modified version of pET-9d including an N-terminal His$_6$-tag and a TEV (tobacco etch virus) protease cleavage site prior to the inserted gene. A triple mutated variant, I65–I70$^{3E/A}$, with residue exchanges E8133A, E8230A and E8324A, was produced by site directed mutagenesis using Quik-Change (Stratagene).

For overexpression of both wild-type I65-I70 and I65-I70$^{3E/A}$, *E. coli* Rosetta (DE3) was grown in LB medium supplemented with 34 μg/ml chloramphenicol and 25 μg/ml kanamycin up to an OD$_{600}$ of 0.6 and at 30°C. Induction used 1 mM IPTG and growth was continued at 20°C overnight. Cells were harvested by centrifugation. Bacterial pellet was resuspended in 50 mM Tris–HCl pH 7.5, 50 mM NaCl and sonicated in the presence of 1 mg/ml lysozyme and 2 μg/ml DNase. The homogenate was clarified by centrifugation at 20000 g for 1 h and 4°C. Affinity purification used a HiTrap column (Amersham Biosciences) equilibrated in lysis buffer. Tag removal was by incubation with TEV protease (0.4 mg/ml) overnight at 4°C during dialysis against 50 mM Tris–HCl pH 7.5, 150 mM NaCl, 1 mM DTT. Subtractive purification for protease (His$_6$-tagged at C-terminus) and non-digested protein was carried out on a subsequent chelating step. The eluted protein fractions were further purified to homogeneity by ion exchange chromatography using a Resource Q column and gel filtration on a Superdex 200 HiLoad 16/60 column (both from Amersham Biosciences) in 50 mM Tris–HCl pH 7.5. Protein samples were then concentrated as required and stored at 4°C.

Throughout this work, His$_6$-tags were removed from all samples, at exception of wild-type I65-I70 used in SAXS.

Small angle X-ray scattering

Synchrotron radiation X-ray scattering data from solutions of wild-type I65–I70 and its mutated variant I65–I70$^{3E/A}$ were collected on the X33 camera at EMBL (storage ring DORIS-III, DESY) using a linear proportional gas detector. The scattering patterns were recorded at a sample-detector distance of 3.6 m covering the range of momentum transfer $0.1 < s < 2.4$ nm^{-1} ($s = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle and $\lambda = 0.15$ nm is the X-ray wavelength). Protein solutions were measured in 50 mM Tris–HCl pH 7.5 at a temperature of 15.0 ± 0.2°C. In addition, wild type I65–I70 was also measured at high ionic strength in buffer as above supplemented with 1 M NaCl. All samples were recorded at least at two protein concentrations in the range from 4.0 to 25 mg/ml. To detect possible radiation damage, scattering data were collected in 10 suc-cessive 1-minute frames. Individual frames were averaged after normalization to the intensity of the incident beam, corrected for the detector response and the scattering of the buffer subtracted. Difference curves were scaled for the solute concentration and extrapolated to infinite dilution following standard procedures (Feigin and Svergun, 1987). All data manipulations were performed using the program package PRIMUS (Konarev *et al.*, 2003). The maximum particle dimension, D_{max}, was estimated using the orthogonal expansion program ORTOGNOM (Svergun, 1993). The forward scattering $I(0)$ and the radii of gyration R_g were evaluated using the Guinier approximation, which assumes that at very small angles ($s < 1.3/R_g$) the intensity is represented as $I(s) = I(0) \exp(-(sR_g)^2/3)$. These parameters were also computed from the entire scattering patterns using the indirect transform package GNOM (Svergun, 1992), which provides as well the distance distribution function $p(r)$ of the particle. The molecular mass (MM) of the solute was evaluated by comparison of the forward scattering with that from reference solutions of bovine serum albumin (MM = 66 kDa).

Electron microscopy

For glycerol spraying/low-angle rotary shadowing, 20 μl of protein sample at 0.1 mg/ml in 50 mM Tris pH7.5, 30% [v/v] glycerol was sprayed onto freshly cleaved mica at room temperature and rotary metal-shadowed in a BAF 511 M freeze-etch apparatus (Balzers) with platinum/carbon at an elevation angle of 3–5°. Images of the metal replica were recorded on an EM912 Omega EFTEM (LEO Electron microscopy, Oberkochen, Germany) at an acceleration voltage of 120 kV, equipped with a slow-scan CCD camera with a 2-MHz read-out, 16-bit information depth, 1024×1024-pixel CCD chip, pixel size 19 μm, and a P43 phosphorus scintillator (Proscan, Scheuring, Germany). End-to-end molecular distances were measured using the ImageJ software (a version of NIH-image).

Results

Conserved features in Ig domains from the I-band

Sequence alignments of Ig domains from the I-band of titin have been previously reported (Improta *et al.*, 1996; Witt *et al.*, 1998). They expose the high levels of conservation in these modules and reveal a key set of features which define domain and linker subclasses in these tandems. Based on those data and for the current study, relevant indicators will be taken as (i) the presence of a PPxf motif at the domain N-terminus in strand A (where P is proline, x any residue and f a hydrophobic group – usually a phenylalanine); (ii) the presence of a lengthened FG β-hairpin housing a conserved NxxG motif at its turn (where N is asparagine, x any residue and G glycine); (iii) the conservation of

a negatively charged group, often a glutamate, in linker sequences; (iv) length of the linkers. A graphical representation of the occurrence of these features in modules of the I-band is shown in Figure 1. Accordingly, two main classes of Ig domains and interconnecting linkers can be detected, which partly correlate with the constitutive or differentially expressed character of the tandems.

A first group, hereby termed "N-conserved type", includes all domains from the differentially expressed I-band as well as flanking domains within constitutive regions (I1, I14–I15, I84–I85, I105) and I25 from the N2B element. In this class, the presence of a long FG β-hairpin correlates strongly with the prevalence of a PP motif at the N-terminus within strand A, so that 63% of I-band domains with an extended hairpin also contain a full PP motif and 90% contain at least one N-terminal proline. In addition, these two features moderately correlate with conservation in length and sequence at the BC turn. While an extended FG β-hairpin favors a longer version of the latter, conservation of the PP motif in strand A appears to promote proline-rich sequences in that turn, usually as PPh or PxP signatures (where h is any hydrophobic residue). Furthermore, the three features occur simultaneously

in this Ig group with detectable frequency, accounting for about one third of the domains. The molecular basis of this correlation can be deduced from the available atomic structures, since most of the models of Ig domains from muscle proteins fall in this category, namely telokin (Holden et al., 1992), I1 (Mayans et al., 2001) and M5 (Pfuhl et al., 1995) from titin, I18 (Fong et al., 1996) and I26 (Kobe et al., 1996) from twitchin. Models of I1, I18 and telokin show one proline rest in strand A, the NxxG motif in the FG turn and a PxP sequence in the BC connection, while slightly degenerated patterns are exhibited by I26, which contains a PPh BC motif, and M5 which lacks proline groups in strand A. These models reveal that the mentioned attributes cluster tightly at the N-terminal loop region of the Ig architecture (Figure 2). Consistently, the proline residue from strand A (corresponding to the second position in the PP motif) stacks against the second proline of the PxP sequence in the BC β-turn. All prolines are in *trans* conformation, with exception of the first proline group of the PxP motif, which adopts a *cis* disposition structurally relevant for the BC turn. Similarly, the asparagine residue in the NxxG sequence of the FG β-turn is central to the architecture of the cluster. Although the interactions mediated by this residue

Fig. 1. Conservation of sequence motifs in Ig tandems from I-band titin: Ig domains and linker segments are represented as filled boxes (as indicated). For simplicity, the PEVK region has been excluded. The presence of an extended FG β-hairpin is given as stars, where red shows conservation of the NxxG motif and black represents any other sequence. The conserved E/D group is displayed on top of the corresponding linker boxes, where blue refers to the C-terminus of strand G and red indicates one additional position ($i+1$). Linkers are classified as short if containing less than three residues and long otherwise. Following classical conventions, "linker" sequences are defined as those inter-domain residues which could not be reliably predicted to form part of β-strands. In the differentially-expressed tandem a two-residue linker corresponds to a zero inter-domain separation (see main text). The conservation of a PP motif in strand A is displayed, as well as proline-rich sequences at the BC loop. Modules for which an atomic structure has been elucidated are marked by a thick bar. Segments for which domain arrangement has been analyzed by SAXS are boxed. Super-repeats of 6–10 Ig domains within the central tandem are marked with a capped-bar (●●).

do not appear entirely conserved, it consistently holds together groups from the FG and BC turns and, in certain cases, from strand A. Remarkably, according to sequence alignment data (Fong *et al.*, 1996), the features described here can also be expected to characterize Ig domains from twitchin.

The second group of I-band domains, hereby termed "N-variable", exhibits a broader heterogeneity in its features. However, it can be defined by shortened BC and FG elements as well as by a minimal number of prolines involved in interactions at the N-terminal loop cluster (Figure 2). This group is loosely represented by the atomic structure of I91 (I27; Improta *et al.*, 1996) and comprises the large majority of constitutive domains.

Inter-domain linker sequences also exhibit certain recurrent features. In particular, the presence of a conserved, negatively charged group (commonly a
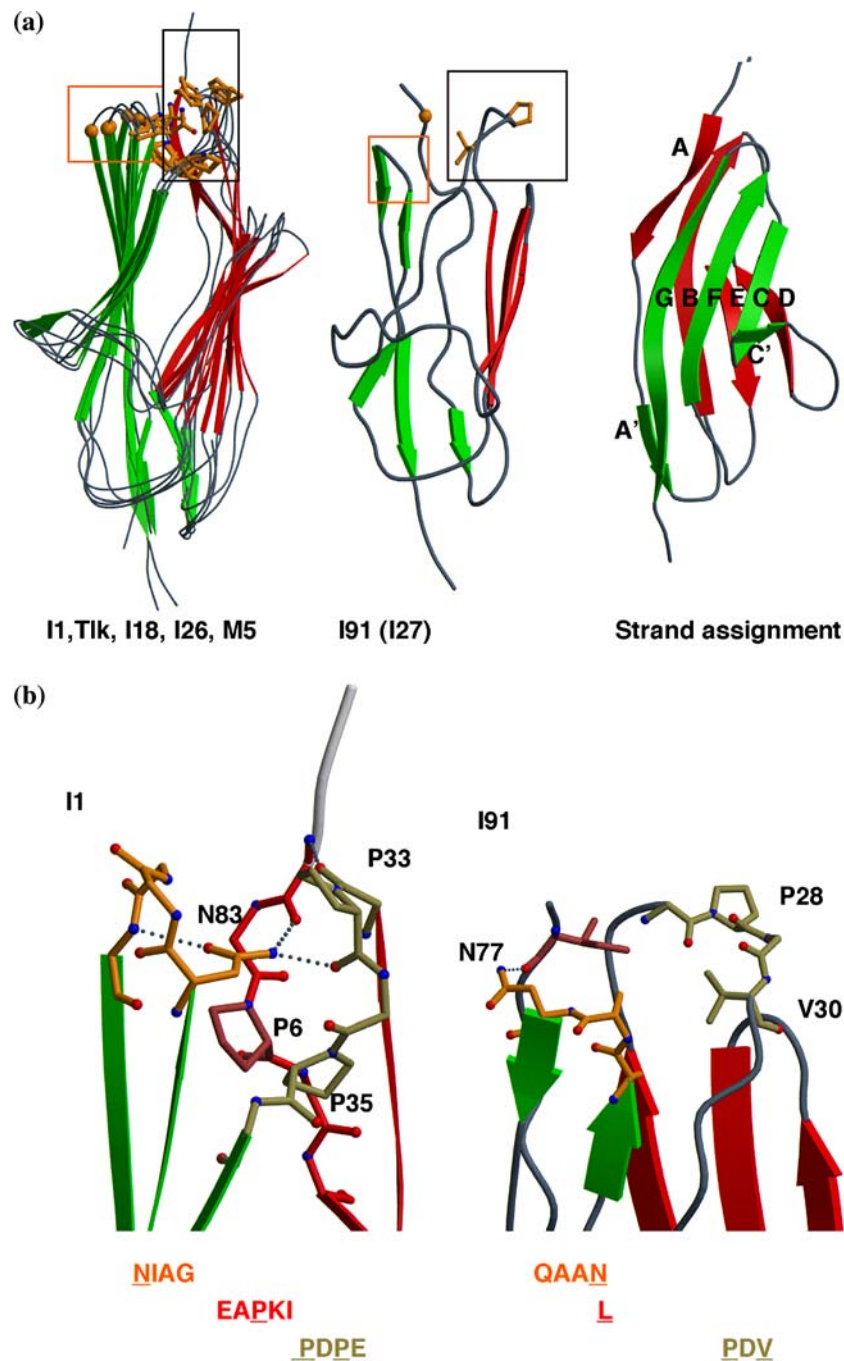
Fig. 2. Atomic structures of Ig domains from muscle proteins: (a) Ribbon representation of Ig models grouped according to the architecture of their N-terminal loop cluster (secondary structure definition as in Mayans *et al.*, 2001). Structures correspond to 1G1C (I1; Mayans *et al.*, 2001), 1NCT (M5; Pfuhl *et al.*, 1996), 1TIT (I91; Improta *et al.*, 1996), 1WIT (I18; Fong *et al.*, 1996), 1KOA (I26; Kobe *et al.*, 1996) and 1FHG (Telokin; Holden *et al.*, 1992). The BC and FG loops are boxed in black and orange, respectively. A ribbon model of I1 is included displaying strand assignment. (b) Stabilizing interactions at the N-terminal loop region. The FG β-turn is colored in yellow, the BC loop in khaki and strand A in red. Residue sequences for these elements are given, where underlined groups correspond to those labeled (numbering as in the respective PDB entries).

glutamate, E) which is present in 82% of all I-band domains. Interestingly, in domains from the first constitutive tandem (I1–I15), this group represents the C-terminal end of strand G (i), while in all other sequences it is shifted to a $i+1$ position. For ''N-conserved'' domains of the differential tandem linked by short sequences, it can be predicted that such shift results in the E residue becoming the first amino acid of strand A in the next Ig domain within the array. This leads to a periodic EPPxf N-terminal motif in these domains. The high conservation of the E groups suggests their involvement in intra- or inter-domain stabilizing interactions. However, this is not supported by structural data available to date. The structure of telokin includes a C-terminal glutamate, located as in the first Ig tandem, while in the structure of I1 this is in N-terminal position, placed as in the EPPxf motif. In none of the models, are these E groups found to establish intra-domain interactions. In I1, the conserved group, E4, hydrogen bonds to residues from the FG and BC turns via its main chain, proving that it is integral part of the N-terminal Ig architecture. A lack of interactions for an E linker residue has also been observed in the structure of an Ig-doublet, Z1Z2, recently elucidated in our laboratory (Marino *et al.*, unpublished). Thus, current data indicate that these E groups are unlikely to determine stability of individual domains. However, given the limited amount of structural models currently available, this possibility cannot be entirely excluded. More importantly, it is yet to be resolved whether such conserved residues located at the domain boundaries could be involved in inter-domain contacts.

Strikingly, modules within the central, spliced tandem show a prominent regular layout of linker features, where groups of domains are connected by short sequences (commonly 2 residues) comprising the conserved glutamate residues and are interspersed by longer linkers housing aspartate rests (Figure 1). The reiteration of the pattern correlates with the previously reported existence of 6–10 Ig super-repeats in this tandem (Gautel, 1996; Kenny *et al.*, 1999). Modules with short connecting sequences result in a highly conserved vxEPPxf linker motif (where v and f indicate prevalence of valine and phenylalanine residues). Analysis of secondary structure elements in the high-resolution models of I1 and telokin, indicate that all residues within this motif are integral part of their respective Ig domains. So that it corresponds to a zero linker length between two consecutive domains in stretched conformation. It is, therefore, likely that Ig domains within the differential I-band section might exhibit very little freedom of motions within the tandem. Taken together, the linker repeats and the conserved loop clusters in Ig domains (expected to constitute domain-domain interfaces; Improta *et al.*, 1998) allow speculating, tentatively, on a sustained regularity of inter-domain arrangement within the skeletal tandem. Unfortunately, the scarcity of structural data for ''N-variable'' domains of titin

and the larger sequence variations in this group, hampers reliable estimations to be made for constitutive I-band regions. A possible local high-order in the fine structure of the filament, if existed, would have deep consequences for its mechanical dynamics.

SAXS characterization of a poly-Ig fragment from the skeletal I-band

In order to investigate the influence of specific domain subclasses and linker attributes on overall domain arrangements in I-band poly-Ig fragments, we have characterized a 6Ig construct from the skeletal tandem of soleus titin, comprising domains I65–I70, using SAXS. The mechanical properties of this fragment have been previously characterized via AFM (Watanabe *et al.*, 2002). While the features of its component Ig modules are constant (Figure 1), the segment includes a melange of linker sequences of different length and composition (Table 2), extending over 2–6 residues and varying in the presence of conserved glutamate residues. The X-ray scattering pattern and computed distance distribution function, $p(r)$ (which represents the probability distribution of the interatomic vectors inside the molecule) are shown in Figure 3. The parameters computed from the scattering data are given in Table 2. The estimated molecular mass (MM) of the solute (62 kDa) was in agreement with that expected for a monomeric state of the I65–I70 sample (63.8 kDa). The R_g value of 6.3 nm indicates an asymmetric and elongated molecular shape. The observed D_{max} (maximum linear dimension), calculated as 23.0 nm, suggests a fairly extended arrangement of all Ig domains in the construct when considering the average length of one Ig as 4.4 nm. Interestingly, this implies that neither linker length nor composition impose particular restraints in inter-domain arrangement that would make specific domain pairs deviate strongly from an average hinge opening.

Results from the current study are well in agreement with previous data on constructs from the distal I-band as well as the Z-disc. SAXS measurements on I91–I92 (I27–I28), I91–I94 (I27–I30) (Improta *et al.*, 1998) and Z1Z2 (Zou *et al.*, 2003) yielded D_{max} values for these samples consistent with a domain disposition predominantly extended along the axial length of the molecule (Table 2). This, in spite of structural differences in Ig domains and linker length and composition in these constructs. The structural relation between adjacent domains in poly-Ig constructs from titin can be further confirmed from SAXS data by considering the local maxima in the respective $p(r)$ functions. The $p(r)$ distributions obtained in the current study as well as those previously reported (Improta *et al.*, 1998; Zou *et al.*, 2003) exhibit a primary local maximum at approximately 2.0 nm, which derives from all inter-atomic distances within a single Ig module, as well as a secondary maximum at

Table 2. X-ray scattering parameters

Sample	Linker Sequence[a]	R_g (nm)	D_{max} (nm)	Linear Length (nm)	MM (kDa)
wt I65–I70	I65–I66: **V**IKARKL**PP**S**F** I66–I67: **I**LSEHEV**PPFF** I67–I68: **V**Q....EPP**R**F I68–I69: **V**K....E**PP**I**F** I69–I70: **L**K....A**PP**R**F**	6.3 ± 0.2	23 ± 1	26.4	58 ± 6
I65–I70³E/A	I65–I66: **V**IKARKL**PP**S**F** I66–I67: **I**LS**A**HEV**PPFF** I67–I68: **V**Q....**A**PP**R**F I68–I69: **V**K....**A**PP I**F** I69–I70: **L**K....A**PP**R**F**	6.4 ± 0.2	23 ± 1	i.d.	i.d.
wt I65–I70 (1M NaCl)		7.8 ± 0.4	27 ± 2	i.d.	68 ± 7
I91–I92[b]	I27–I28: **V**KEL...**P**LI**F**	2.8 ± 0.03	9 ± 0.3	8.8	–
I91–I94[b]	I27–I28: **V**KEL...**P**LI **F** I28–I29: **I**EGI...RLK**F** I29–I30: **V**LEG...DPY**F**	4.5 ± 0.22	15 ± 0.5	17.6	–
Z1Z2[c] 8.8	Z1–Z2: **V**KAE.TA**PP**N**F**	3.20 ± 0.04	11.5 ± 0.5		23 ± 2

[a]Residues in bold correspond to conserved groups demarking the end and beginning of consecutive Ig domains (linker definition as in Figure 1). Underlined glutamate residues are those conserved in position with respect to the differentially expressed I-band. Enhanced characters in sample I65–I70³E/A indicate residue exchanges; [b](Improta *et al.*, 1998); [c](Zou *et al.*, 2003).

~4.4–4.8 nm, which represents the separation between the centres of mass of two consecutive Ig domains. This distance between adjacent modules is close to the average length of an Ig domain, indicating extended domain arrangements common to the different constructs.

Since data on I65–I70 indicated that linker residue composition had little effect on the overall molecular conformation of this sample, we tested this further by assaying a mutated variant, I65–I70³E/A, where three glutamate residues at the linker regions had been exchanged for alanine groups (Table 2). Scattering data (Figure 3, Table 2) coincided within experimental error with those of the wild-type, revealing no detectable alteration of the overall molecular conformation in solution. Although it cannot be excluded that changes might occur at atomic level beyond the resolution limit of the SAXS technique, it can be reliably deduced that the bending stiffness of the hinges is not significantly altered. Furthermore, since it has been questioned how media conditions might affect the latter (Tskhovrebova and Trinick, 2001), we measured the overall molecular parameters of wild-type I65–I70 in buffer solution supplemented with 1 M NaCl (Figure 3, Table 2). In this case, the overall particle parameters were clearly larger than those of the wild-type and $p(r)$ functions displayed a significantly broader distribution. This may be attributed to partial, but detectable, aggregation of the protein, which revealed itself at very low scattering angles. Apart from this, the scattering in high salt did not show a significant difference from that of the wild-type indicating that I65–I70 retains its elongated shape in high ionic strength media. The absence of molecular collapse suggests that interdomain flexibility is not greatly increased by this factor either.

Electron microscopy

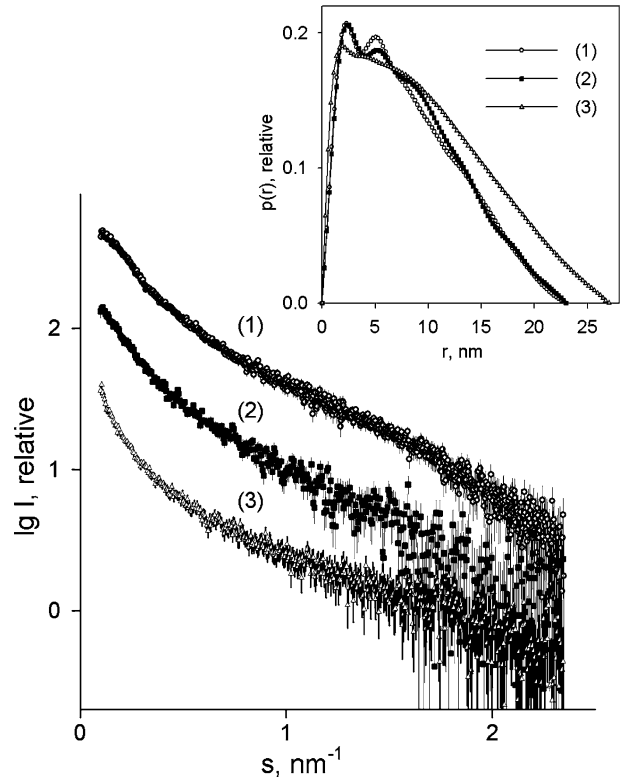Given that the solution scattering technique reports space- and time-averaged molecular parameters for a

Fig. 3. SAXS characterization: X-ray scattering patterns of wild type I65–I70 (1), variant I65–I70³E/A (2) recorded in 50 mM Tris–HCl pH 7.5 and wild-type protein in buffer supplemented with 1 M NaCl (3). The curves are displaced along the logarithmic axis for better visualization. The inset displays distance distribution functions computed from the experimental data using the program GNOM.

sample population in solution, we examined the distribution of conformations adopted by I65–I70 and its mutated variant using rotary shadowing electron microscopy. A typical sample field and end-to-end distance distributions are shown in Figure 4. Microscopy images show I65–I70 molecules as elongated rods of kinked, bowed or straight visual appearance. The end-to-end distance, R, measured for $n = 320$ molecules yielded a mean value $\langle R \rangle$ of 24.6 ± 3.3 nm (value corrected for the contribution of the metal coating estimated as 2 nm), which is well in agreement with D_{max} values derived from SAXS data. An insight into the conformational freedom allowed by inter-domain hinges in this construct can be derived from the experimental mean square $\langle R^2 \rangle_{exp}$, which equals 619 ± 160 nm^2. To analyze further these data, we estimated the contour length of the 6Ig fragment under study to approximate 29–30 nm. This was calculated by considering: (i) the average length of an Ig domain as 4.4 nm; (ii) that two-residue linkers add zero inter-domain distance (as discussed previously); and (iii) that two linkers in this construct have 4 additional residues which, if in extended antiparallel β-strand conformation, will add ∼0.35 Å per residue. Of models available to describe the conformation of biopolymers, the worm-like chain (WLC) model is the most commonly applied to titin filaments analyzed in vitro. The $\langle R^2 \rangle_{WLC}$ for I65–I70 estimated using a contour length of 30 nm and applying the WLC model for a polymer confined in two dimensions (Rivetti et al., 1996), yields values of 578 and 662 nm^2
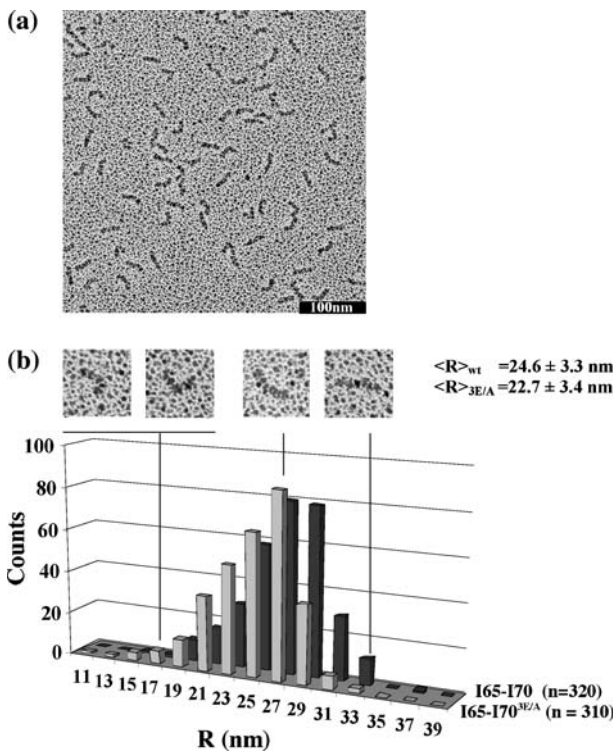

Fig. 4. EM characterization: (a) Typical Sample field of wild-type I65–I70; (b) Histogram of molecular classes for I65–I70 (dark grey) and mutated I65–I70$^{3E/A}$ (light grey) (the histogram contains uncorrected values).

when assuming persistence lengths of 10 and 15 nm (Higuchi et al., 1993; Di Cola et al., 2005), respectively. Our experimental $\langle R^2 \rangle_{exp}$ value falls within these estimations, indicating that the apparent persistence length of this 6Ig construct is within the range calculated for full titin.

It should be noted that the WLC model is suitable for a smoothly bent polymer, in the form of a continuous, homogeneous chain with no weak points or irregularities. It is generally used for polymers where the contour length is much larger than the bond length, so that the latter is considered to be zero and the number of bonds infinite. However, the current 6Ig construct has a limited number of bonds and the bond length is large compared to the contour length. Such a discreet system, with a finite bead-on-a-string architecture and putatively weak bonds between subsequent joints, can be better described by using the freely-jointed chain (FJC) model. This is the simplest model to describe a discreet polymer (Flory, 1969) so that $\langle R^2 \rangle_{FJC} = n * l^2$, where n is the number of bonds and l is the bond length. Since the 6Ig fragment has an estimated contour length of 29–30 nm, the average bond length is 5 nm. Thus, six freely-jointed modules would yield an $\langle R^2 \rangle_{FJC}$ of 150 nm^2, while a straight rod of 29–30 nm length would result in $\langle R^2 \rangle_{straight}$ of 841–900 nm^2. A comparison of the experimentally measured $\langle R^2 \rangle_{exp}$ and these calculated values suggests that I65–I70 is a relatively stiff formation. The analysis of the mutated variant I60–I75$^{3E/A}$ yielded equivalent results ($\langle R \rangle = 22.7 \pm 3.4$ nm; value corrected as above), confirming that glutamate residues at the linker sequences do not determine hinge opening values or flexibility parameters. In this context, a more detailed analysis of $\langle R^2 \rangle_{exp}$ for I65–I70 can be performed assuming that domains are not free to rotate unhindered within the protein chain and that, in fact, they are limited to a cone of motion with angle θ. Then, $\langle R^2 \rangle = C_6 \langle R^2 \rangle_{FJC}$, where C_6 is a function of $\alpha = \cos \theta$ (P.J. Flory, 1969): $C_6 = ((1 + \alpha)/(1 - \alpha)) - (\alpha/3) \times ((1 - \alpha^6)/(1 - \alpha)^2)$. The experimentally determined end-to-end distance distribution leads to $C_6 = 4.13$ and $\theta = 37°$, which corresponds to an averaged opening angle of 143° between any two consecutive Igs. This is in agreement with estimations for the pair I91–I92 (I27–I28) based on NMR and SAXS data, which yielded a hinge opening of ≈155° (Improta et al., 1998). The calculated value gives an insight into the filament architecture and, it being a structural parameter, can be validated experimentally using future atomic structures of multi-Ig fragments.

Discussion

Although the mechanical properties of poly-Ig tandems from I-band titin have been extensively investigated, their structure and dynamics remain poorly characterized. To date, only two atomic models from

I-band domains are available: I1 (Mayans *et al.*, 2001) and I91 (I27) (Improta *et al.*, 1996) (Table 1). Similarly, the study of the arrangement of domains in I-band tandems has been limited to the constitutive segment I91–I94 (I27–I30) using SAXS and NMR in combination (Improta *et al.*, 1998) (Table 2). Whether data on these domains are representative of general I-band features is yet to be determined. Ig components from titin I-band belong to different subclasses (Improta *et al.*, 1996; Witt *et al.*, 1998), characterized by variations in size and composition of a N-terminal loop cluster where strand A and β-turns FG and BC pack together (Figure 1). These features partly correlate with the constitutive or differentially expressed character of the tandems. Conservation at this cluster can be understood by analysis of atomic structures available for Ig domains of muscle proteins. Interestingly, distributed along the entire I-band and irrespective of module type, a preference for glutamate groups at the domain interfaces is observed. Structural data currently available do not support an involvement of such groups in intra-domain stabilization, although more atomic models will be required before this possibility can be discarded. The current study assayed their possible influence in inter-domain conformation and dynamics by examination of a triply mutated variant lacking the conserved groups. No differences between wild-type and mutant were detected, although the coarseness of the techniques here employed do not allow excluding possible fine changes in filament structure or dynamics. It can, nevertheless, be concluded that these residues are not a key determinant of either molecular stability or inter-domain conformation. Thus, the basis for the conservation of these surface-exposed groups remains unknown.

Given that linker length and N-terminal loop structure are believed to determine domain interfaces and flexibility within Ig-tandems (Improta *et al.*, 1998), we explored the overall molecular conformation of a 6Ig fragment from the skeletal tandem of soleus titin. No data have been previously reported on the skeletal fraction of the I-band. The fragment, I65-I70, has conserved Ig-domain characteristics but linker sequences which diverge in length and composition (Table 2). Our data, derived from SAXS and rotary-shadowing EM approaches and supported by a structure-based sequence analysis of components, indicate that this segment adopts an extended conformation in solution and exhibits certain stiffness. Remarkably, our findings agree closely with studies previously reported on the constitutive I-band fragment I91–I94 (Improta *et al.*, 1998) and a doublet from the Z-disc, Z1–Z2 (Zou *et al.*, 2003), regardless of contrasted differences in domain and linker composition (Table 2). A detectable alteration of molecular parameters was not achieved either by mutation of conserved glutamate linker groups or by increase of the ionic strength of the medium, even if the several linkers in this construct would permit varied Ig arrangements to occur. The molecular

basis of the conformational stability of this construct are unknown, although this is unlikely to result from the conservation of the N-terminal EPPxf motif in strand A since these residues are embedded within the Ig fold and are not architectural part of the hinges. Thus, it can be concluded that poly-Ig tandems from I-band titin tend to adopt extended arrangements with low or moderate intrinsic flexibility, independently of the specific features of linker or component Ig domains across constitutively- and differentially-expressed tandems. Linkers do not appear to operate as free hinges leading to frequent lateral association of Ig domains, although connections of about 4 residues length would well accommodate such geometry.

Although the current study does not reveal details of the fine structure of the filament or inter-domain dynamics, based on measured mean square $\langle R^2 \rangle$ values, we speculate that average hinge openings between any two consecutive Ig will approximate 140°–150°. In any case, it should be borne in mind that poly-Ig segments in solution are bound to exhibit certain conformational heterogeneity where extended module-pairs might be subject to varying degrees of bending, twisting and tilting, according to the relative stiffness of the individual hinges. Relative motions such as the sharp kinks seen in EM images might be enhanced or induced by the surface adsorption procedure used in this technique and might only occur rarely in dilute solutions of the sample, as indicated by SAXS. Our results and structure-based deductions are consistent with recent microscopy data obtained from isolated titin filaments (for some recent reports see Tskhovrebova and Trinick, 2001; Kellermayer *et al.*, 2003), which hint at possible local substructure in the filament and a certain torsional stiffness. Taken together, sequence conservation patterns and biophysical data suggest that poly-Ig fragments from I-band titin might build well-defined formations. Knowledge on the dynamic conformation of titin components at the atomic level will be essential to unveil the structural and mechanical intricacy of the filament.

Acknowledgements

References

Atkinson RA, Joseph C, Kelly G, Muskett FW, Frenkiel TA, Nietlispach D and Pastore A (2001) Ca2+-independent binding of an EF-hand domain to a novel motif in the alpha-actinin-titin complex. *Nat Struct Biol* **8**: 853–857.

Bang ML, Centner T, Fornoff F, Geach AJ, Gotthardt M, McNabb M, Witt CC, Labeit D, Gregorio CC, Granzier H and Labeit S (2001) The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. *Circ Res* **89:** 1065–1072.

Bullard B, Linke WA and Leonard K (2002) Varieties of elastic protein in invertebrate muscles. *J Muscle Res Cell Motil* **23:** 435–447.

Di Cola E, Waigh T, Trinick J, Tskhovrebova L, Houmeida A, Pyckhout-Hintzen W and Dewhurst (2005) Persistence Length of Titin from Rabbit Skeletal Muscles Measured with Scattering and Microrheology Techniques. *Biophys J.* **88:** 4095–4106.

Feigin LA and Svergun DI (1987) *Structure Analysis by Small-Angle X-ray and Neutron Scattering*. Plenum Press, New York.

Flory PJ (1969) *Statistical Mechanics of Chain Molecules*. Interscience publishers, New York.

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

Gao M, Lu H and Schulten K (2001) Simulated refolding of stretched titin immunoglobulin domains. *Biophys J* **81:** 2268–2277.

Gao M, Wilmanns M and Schulten K (2002) Steered molecular dynamics studies of titin I1 domain unfolding. *Biophys J* **83:** 3435–3445.

Gautel M (1996) The super-repeats of titin/connectin and their interactions: glimpses at sarcomeric assembly. *Adv Biophys* **33:** 27–37.

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

Granzier HL and Labeit S (2004) The giant protein titin: a major player in myocardial mechanics, signaling, and disease. *Circ Res* **94:** 284–295.

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

Holden HM, Ito M, Hartshorne DJ and Rayment I (1992) X-ray structure determination of telokin, the C-terminal domain of myosin light chain kinase, at 2.8 A resolution. *J Mol Biol* **227:** 840–851.

Improta S, Krueger JK, Gautel M, Atkinson RA, Lefevre JF, Moulton S, Trewhella J and Pastore A (1998) The assembly of immunoglobulin-like modules in titin: implications for muscle elasticity. *J Mol Biol* **284:** 761–777.

Improta S, Politou AS and Pastore A (1996) Immunoglobulin-like modules from titin I-band: extensible components of muscle elasticity. *Structure* **4:** 323–337.

Kellermayer MS, Bustamante C and Granzier HL (2003) Mechanics and structure of titin oligomers explored with atomic force microscopy. *Biochim Biophys Acta* **1604:** 105–114.

Kellermayer MS, Smith SB, Granzier HL and Bustamante C (1997) Folding-unfolding transitions in single titin molecules characterized with laser tweezers. *Science* **276:** 1112–1116.

Kenny PA, Liston EM and Higgins DG (1999) Molecular evolution of immunoglobulin and fibronectin domains in titin and related muscle proteins. *Gene* **17:** 11–23.

Kobe B, Heierhorst J, Feil SC, Parker MW, Benian GM, Weiss KR and Kemp BE (1996) Giant protein kinases: domain interactions and structural basis of autoregulation. *EMBO J* **15:** 6810–6821.

Konarev PV, Volkov VV, Sokolova AV, Koch MHJ and Svergun DI (2003) PRIMUS – a Windows-PC based system for small-angle scattering data analysis. *J Appl Crystallogr* **36:** 1277–1282.

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

Li H, Carrion-Vazquez M, Oberhauser AF, Marszalek PE and Fernandez JM (2000) Point mutations alter the mechanical stability of immunoglobulin modules. *Nat Struct Biol* **7:** 1117–1120.

Li H and Fernandez JM (2003) Mechanical design of the first proximal Ig domain of human cardiac titin revealed by single molecule force spectroscopy. *J Mol Biol* **334:** 75–86.

Li H, Linke WA, Oberhauser AF, Carrion-Vazquez M, Kerkvliet JG, Lu H, Marszalek PE and Fernandez JM (2002) Reverse engineering of the giant muscle protein titin. *Nature* **418:** 998–1002.

Linke WA, Ivemeyer M, Olivieri N, Kolmerer B, Ruegg JC and Labeit S (1996) Towards a molecular understanding of the elasticity of titin. *J Mol Biol* **261:** 62–71.

Lu H, Isralewitz B, Krammer A, Vogel V and Schulten K (1998) Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys J* **75:** 662–671.

Ma K, Kan L and Wang K (2001) Polyproline II helix is a key structural motif of the elastic PEVK segment of titin. *Biochemistry* **40:** 3427–3438.

Ma K and Wang K (2003) Malleable conformation of the elastic PEVK segment of titin: non-co-operative interconversion of polyproline II helix, beta-turn and unordered structures. *Biochem J* **374:** 687–695.

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

Mayans O, van der Ven PF, Wilm M, Mues A, Young P, Furst DO, Wilmanns M and Gautel M (1998) Structural basis for activation of the titin kinase domain during myofibrillogenesis. *Nature* **395:** 863–869.

Mayans O, Wuerges J, Canela S, Gautel M and Wilmanns M (2001) Structural evidence for a possible role of reversible disulphide bridge formation in the elasticity of the muscle protein titin. *Structure* **9:** 331–340.

Miller MK, Granzier H, Ehler E and Gregorio CC (2004) The sensitive giant: the role of titin-based stretch sensing complexes in the heart. *Trends Cell Biol* **14:** 119–126.

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

Pfuhl M and Pastore A (1995) Tertiary structure of an immunoglobulin-like domain from the giant muscle protein titin: a new member of the I set. *Structure* **3:** 391–401.

Politou AS, Gautel M, Improta S, Vangelista L and Pastore A (1996) The elastic I-band region of titin is assembled in a "modular" fashion by weakly interacting Ig-like domains. *J Mol Biol* **255:** 604–616.

Rief M, Gautel M, Oesterhelt F, Fernandez JM and Gaub HE (1997) Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276:** 1109–1112.

Rivetti C, Guthold M and Bustamante C (1996) Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis. *J Mol Biol.* **264:** 919–932.

Scott KA, Steward A, Fowler SB and Clarke J (2002) Titin; a multidomain protein that behaves as the sum of its parts. *J Mol Biol* **315:** 819–829.

Svergun DI (1992) Determination of the regularization parameter in indirect transform methods using perceptual criteria. *J Appl Crystallogr* **25:** 495–503.

Svergun DI (1993) A direct indirect method of small-angle scattering data treatment. *J Appl Crystallogr* **26:** 258–267.

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

Tskhovrebova L and Trinick J (2001) Flexibility and extensibility in the titin molecule: analysis of electron microscope data. *J Mol Biol* **310:** 755–771.

Tskhovrebova L and Trinick J (2003) Titin: properties and family relationships. *Nat Rev Mol Cell Biol* **4:** 679–689.

Tskhovrebova L, Trinick J, Sleep JA and Simmons RM (1997) Elasticity and unfolding of single molecules of the giant muscle protein titin. *Nature* **387:** 308–312.

Watanabe K, Muhle-Goll C, Kellermayer MS, Labeit S and Granzier H (2002) Different molecular mechanics displayed by titin's constitutively and differentially expressed tandem Ig segments. *J Struct Biol* **137**: 248–258.

Witt CC, Olivieri N, Centner T, Kolmerer B, Millevoi S, Morell J, Labeit D, Labeit S, Jockusch H and Pastore A (1998) A survey of the primary structure and the interspecies conservation of I-band titin's elastic elements in vertebrates. *J Struct Biol* **122**: 206–215.

Zou P, Gautel M, Geerlof A, Wilmanns M, Koch MH and Svergun DI (2003) Solution scattering suggests cross-linking function of telethonin in the complex with titin. *J Biol Chem* **278**: 2636–2644.