Journal of Muscle Research and Cell Motility **23**: 435–447, 2002.

Varieties of elastic protein in invertebrate muscles

BELINDA BULLARD[1,*], WOLFGANG A. LINKE[2] and KEVIN LEONARD[1]
[1]*European Molecular Biology Laboratory, D-69012 Heidelberg;* [2]*Institute of Physiology and Pathophysiology, University of Heidelberg, D-69120 Heidelberg, Germany*

Abstract

Elastic proteins in the muscles of a nematode (*Caenorhabditis elegans*), three insects (*Drosophila melanogaster*, *Anopheles gambiae*, *Bombyx mori*) and a crustacean (*Procambus clarkii*) were compared. The sequences of thick filament proteins, twitchin in the worm and projectin in the insects, have repeating modules with fibronectin-like (Fn) and immunoglobulin-like (Ig) domains conserved between species. Projectin has additional tandem Igs and an elastic PEVK domain near the N-terminus. All the species have a second elastic protein we have called SLS protein after the *Drosophila* gene, *sallimus*. SLS protein is in the I-band. The N-terminal region has the sequence of kettin which is a spliced product of the gene composed of Ig-linker modules binding to actin. Downstream of kettin, SLS protein has two PEVK domains, unique sequence, tandem Igs, and Fn domains at the end. PEVK domains have repeating sequences: some are long and highly conserved and would have varying elasticity appropriate to different muscles. Insect indirect flight muscle (IFM) has short I-bands and electron micrographs of *Lethocerus* IFM show fine filaments branching from the end of thick filaments to join thin filaments before they enter the Z-disc. Projectin and kettin are in this region and the contribution of these to the high passive stiffness of *Drosophila* IFM myofibrils was measured from the force response to length oscillations. Kettin is attached both to actin near the Z-disc and to the end of thick filaments, and extraction of actin or digestion of kettin leads to rapid decrease in stiffness; residual tension is attributable to projectin. The wormlike chain model for polymer elasticity fitted the force-extension curve of IFM myofibrils and the number of predicted Igs in the chain is consistent with the tandem Igs in *Drosophila* SLS protein. We conclude that passive tension is due to kettin and projectin, either separate or linked in series.

Introduction

> But, as the world, harmoniously confused:
> Where order in variety we see,
> And where, though all things differ, all agree.
>
> Alexander Pope
> '*Windsor Forest*' (1711)

Invertebrate animals encompass a wide range of species with muscles of differing function. So far, the elastic proteins in the muscles have been studied in only two invertebrate phyla: arthropods (insects and crustacea) and nematodes. But even this limited selection reveals the diversity in the architecture of proteins with an elastic function. The arthropods are the most varied group in the animal kingdom and are the subject of much current work on the relationship between the evolution of genes and diversity in morphology (Averof, 1997; Akam, 2000). Among arthropods, insects represent the greatest number of species, a success that is largely due to the development of flight. Here we will compare the elastic proteins of the invertebrates for which the sequence is known, and give a more detailed account of the function of these proteins in insect flight muscle.

*To whom the correspondence should be addressed: Tel.: +49-6221-387313; Fax: +49-6221-387306; E-mail. bullard@embl-heidelberg.de

The narrow focus of recent work on vertebrate titin, especially the human cardiac isoform, has been at the expense of a broader picture of possible ways in which modular proteins may control muscle elasticity. The invertebrate elastic proteins have some similarity in structure to vertebrate titin, and the large body of information on titin makes this a useful model from which to predict the function of the invertebrate proteins. However, the muscles in many invertebrates have functions not found in vertebrates, and the elastic properties of the proteins will be peculiar to these muscles.

Recent re-appraisal of evolutionary relationships, based on the fossil record and ribosomal RNA sequences, places nematodes and arthropods in the same superclade of Ecdysozoa or moulting animals, and annelids and molluscs in the Lophotrochozoa (Aguinaldo *et al.*, 1997; de Rosa *et al.*, 1999; Conway Morris, 2000). In previous classifications, arthropods and annelids were in the same clade based on their segmentation, and nematodes were considered more primitive because simpler. Insects and crustaceans are also now thought to be more closely related than they were previously (Averof and Cohen, 1997). However, the new clade of Ecdysozoa is controversial (Blair *et al.*, 2002) and the evolutionary relationship of invertebrates is by no means settled. New classifications may be

tested by an analysis of protein families such as the elastic proteins in muscle, though there are few examples at present.

The genome sequences of *Caenorhabditis elegans* and *Drosophila melanogaster* have been completed, and more recently that for a mosquito, *Anopheles gambiae*. In addition there is partial genomic DNA sequence for other invertebrates such as the silkworm, *Bombyx mori*. The four species for which the complete sequence of some modular elastic proteins is known: *C. elegans*, *D. melanogaster*, *A. gambiae* and *Procambus clarkii* (crayfish) belong to groups that are thought to be genetically similar, despite the obvious differences in phenotype between worm, insects and crustacean. As we shall see, there are clear homologies in the elastic muscle proteins of the four species, but parts of the sequences diverge to accommodate widely differing functions of the muscles. The modular design of the proteins is particularly suited to stepwise modification not requiring evolution of new genes to change function.

Mechanical properties of the muscles

A current challenge is to determine the relationship between the passive stiffness of different muscle types and the properties of elastic proteins in the sarcomere. Invertebrates have more than one type of modular elastic protein which compounds the problem. Whereas the elasticity of vertebrate striated muscle can be considered in terms of titin, often in different isoforms, predictions of the elasticity of invertebrate muscles must take account of both projectin and a second modular protein known variously as kettin, D-titin and I-connectin.

In the case of *C. elegans*, nothing is known about the elastic properties of the body wall muscle, beyond what can be deduced from the movements of the worm and the structure of the muscle. Fibres are obliquely striated and have dense bodies instead of Z-discs; the thick filaments are 10 μm long, so the sarcomere length (the distance between dense bodies) is at least this length. Thin filaments are ~6 μm which suggests the muscle is capable of developing tension over a wide range of sacromere lengths (Waterston, 1988). The dense bodies are linked to the cuticle so that contraction of sarcomeres causes the cuticle to shorten; coordination of local contractions produces waves along the worm (Waterston, 1988). Crayfish claw fibres are striated and have long sarcomeres (8–13 μm) with long I-bands. Fibres are easily extensible when relaxed and do not develop appreciable passive tension until stretched to about 140% rest length (Fukuzawa *et al*., 2001). The wide range of sarcomere lengths over which *C. elegans* body wall fibres and crayfish claw fibres are thought to operate is in contrast to the nearly isometric contraction of *Drosophila* flight muscle.

Insect striated muscles may be synchronous or asynchronous. In synchronous muscles, there is a direct correspondence between nervous impulses to the muscle and contractions. Non-flight muscles are synchronous and the structure of the sarcomere is typical of striated muscles with an I-band of varying width, depending on the function of the muscle. Insects with lower wingbeat frequencies, such as dragonflies, butterflies and locusts have synchronous flight muscle. Small insects with high wingbeat frequencies, such as flies, mosquitos, bees and wasps, and some larger insects, including the waterbug *Lethocerus*, have asynchronous flight muscle. In these muscles, oscillatory contractions are produced by delayed activation in response to stretch, combined with the resonant properties of the thorax (Pringle, 1978). The upstroke and downstroke of the wings result from distortions of the thorax rather than direct action of the muscles on the base of the wing, hence the muscles are known as indirect flight muscles (IFM). The sarcomeres in the IFMs are 3.4 μm long in *Drosophila*; there is almost no I-band, and the fibres change in length by only 3.3% during flight (Chan and Dickinson, 1996). Rapid oscillatory contraction requires that the sarcomeres are stiff so that strain is transmitted to the thick and thin filaments. The proteins responsible for the high stiffness will be described. The contrasting sarcomere structures of an IFM and a synchronous muscle in *Drosophila* thorax are shown in Figure 1.

Drosophila has become the natural choice for investigating insect elastic proteins, especially since the completion of the genome sequence (Adams *et al*., 2000). The small size of the muscle makes mechanical measurements on fibres difficult but two groups have been successful in getting the muscle to do oscillatory work (Peckham *et al*., 1990; Kreuz *et al*., 1996). The disadvantage of whole fibre preparations for estimating the elasticity of the sarcomere is that other components have to be taken into account, such as collagen and surrounding mitochondria, particularly in the case of insect flight muscle where mitochondria are packed tightly between myofibrils. In addition, neighbouring myofibrils may be linked through connections between Z-discs. The measurement of tension in isolated myofibrils gives a direct estimate of elasticity without these complications and in the case of *Drosophila*, there is the added advantage that myofibrils are easier to manipulate than fibres (Kulke *et al*., 2001).

Elastic proteins of invertebrates

Twitchin

Early work on the genes involved in muscle function in *C. elegans* led to the identification of the first large muscle protein with a modular structure made up of immunoglobulin-like (Ig) and fibronectin-like (Fn) domains. Worms with mutations in the *unc-22* gene showed characteristic twitching of the body wall muscles (Waterston *et al*., 1980) and genetic evidence suggested an interaction between the gene product, twitchin, and

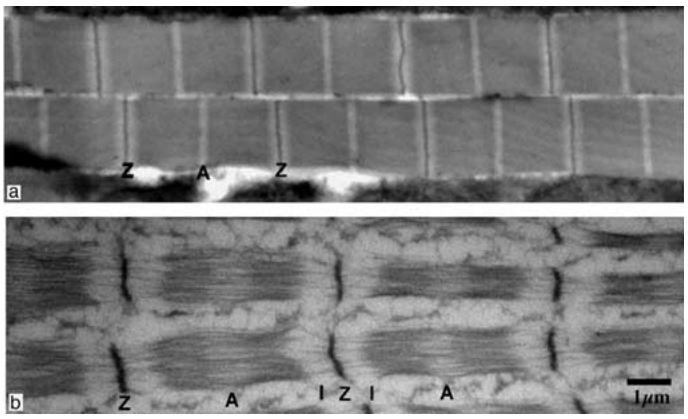

Fig. 1. Variation in the structure of *Drosophila* muscle fibres. Electron micrographs of cryo-sections: (a) IFM showing sarcomeres in which thick filaments extend nearly the whole length of the sarcomere. (b) A synchronous thoracic muscle with less well ordered, longer sarcomeres with longer I-bands. Z, Z-disc; A, A-band; I, I-band. Scale bar 1 μm for a and b.

myosin. Twitchin is associated with thick filaments in the obliquely striated body wall muscle (Moerman *et al*., 1988). A large part of the 753 kDa molecule has a repeating structure consisting of two Fn and one Ig domain. The pattern of modules varies towards the C-terminus and is followed by a protein kinase domain with associated auto-inhibitory sequence (Benian *et al*., 1989, 1993 and Figure 2a). At each end of the molecule there are Ig domains separated by unique sequence. Interestingly, a nearly normal phenotype results when up to six Fn–Ig modules are deleted from the protein, so in the case of *C. elegans*, the exact length of the molecule is not critical (Kiff *et al*., 1988). Twitchin is unlikely to have an effect on the passive stiffness of nematode muscle because it is confined to the thick filaments. However, it is of interest here because of the similarity between twitchin and projectin. Twitchin is also associated with thick filaments in the catch muscle of the mollusc, *Mytilus edulis* (Vibert *et al*., 1993), but although sequence is available for part of the molecule (Funabara *et al*., 2001), there is not enough to be useful in a comparison of sequences.

Projectin

Projectin is a protein of 800–1000 kDa associated with thick filaments in many invertebrate muscles. The sequence of *Drosophila* genomic and cDNA codes for a core region with a repeating pattern of Fn and Ig domains like twitchin but with four extra modules (Ayme-Southgate *et al*., 1991; Fyrberg *et al*., 1992; Daley *et al*., 1998 and Figure 2a). The variation in the pattern of Fn–Ig modules towards the C-terminus is identical to the one seen in the twitchin sequence. There is also a kinase domain near the C-terminus, which has homology to the kinase domain in twitchin. Recently, the projectin sequence has been extended towards the N-terminus using the completed *Drosophila* genome sequence, and here the protein diverges from nematode twitchin (Southgate and Ayme-Southgate, 2001 and Figure 2a). Alternative splicing in this region produces isoforms of differing size. Some isoforms have sequence predicted to be elastic because they are homologous to an extensible region in vertebrate titin which has a high proportion of proline (P), glutamic acid (E), valine (V), and lysine (K) (Labeit and Kolmerer, 1995). At the N-terminus, all isoforms have eight Ig domains upstream of a variable length PEVK region which is followed by five or six Ig domains. There are 13 exons in the PEVK region of the gene which show a rich pattern of alternative splicing reminiscent of the numerous splice forms in the PEVK sequence of vertebrate titin (Bang *et al*., 2001). Thus the sequence of the longer isoforms of projectin has the same general pattern seen in vertebrate skeletal muscle titin: tandem Ig domains either side of a variable length PEVK region, followed by Fn–Ig modules with a kinase domain near the C-terminus. There is a predicted projectin sequence in the *Anopheles* genome which corresponds to a slightly shorter molecule than the full length *Drosophila* projectin: there is one less Fn–Ig module, shorter tandem Ig regions at the N-terminus, and shorter PEVK sequence (Figure 2a). The conservation in the repeating pattern of Fn–Ig modules and the kinase domain in nematode twitchin and the two insect projectins suggest an essential function for this part of the molecule, whereas the variability at the N-terminus is likely to be an adaptation to the mechanical requirements of the different insects. The crayfish claw muscle has a projectin-like protein associated with the thick filament but the sequence is not known (Hu *et al*., 1990).

The position of projectin in the sarcomere differs in different types of insect muscle. In synchronous muscles of *Lethocerus* and *Drosophila*, there are projectin molecules across the whole A-band, with the exception of a gap in the centre of the sarcomere at the H-zone; while in IFM, projectin is confined to the end of the A-band and extends across the short I-band to the Z-disc (Lakey *et al*., 1990; Saide *et al*., 1990; Vigoreaux *et al*., 1991). The distance spanned by projectin along the end of the thick filament up to the Z-disc is about 300 nm, which is approximately the length of the molecule (Nave and

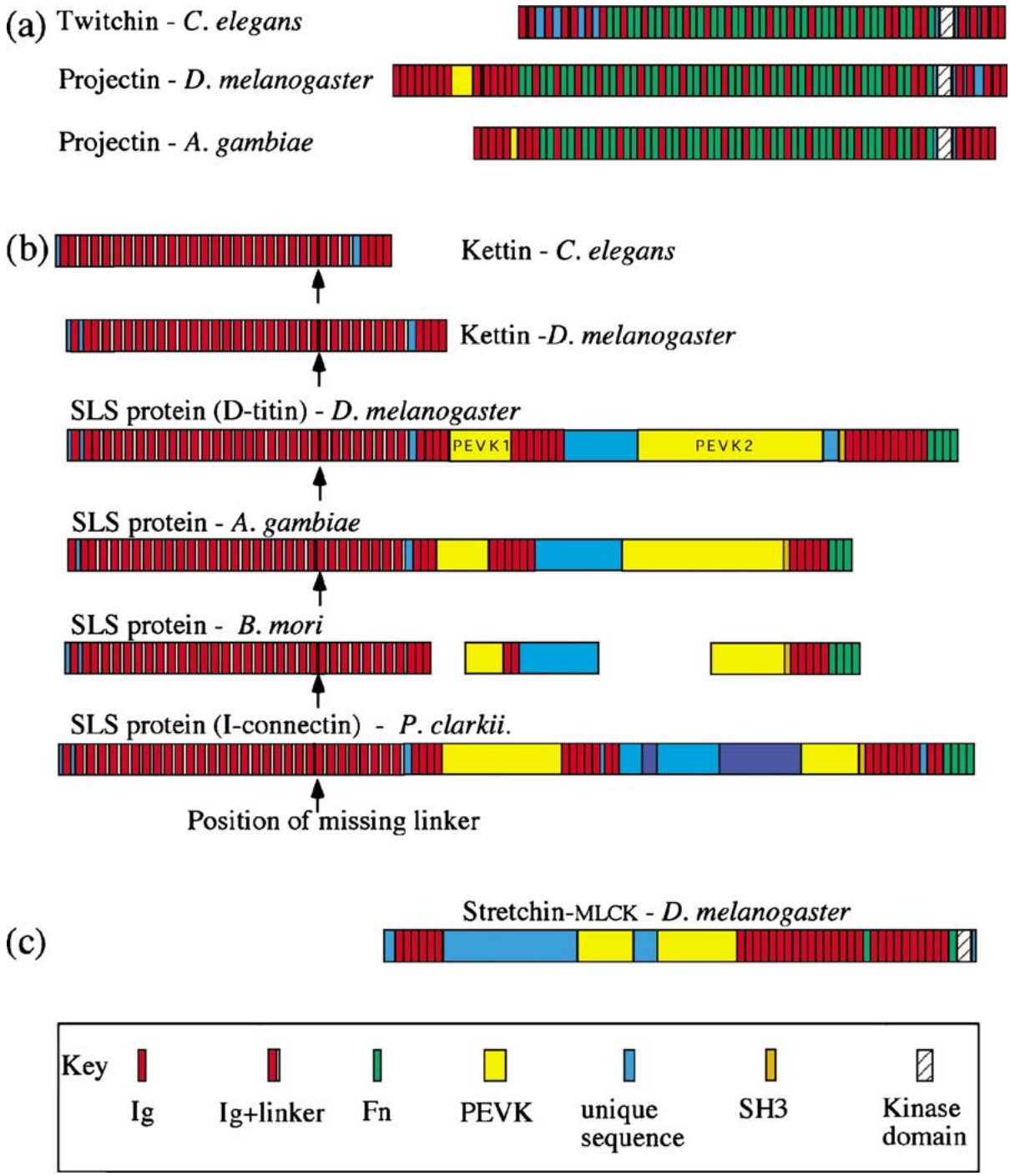

Fig. 2. Modular structure of elastic proteins in invertebrate muscles. (a) Twitchin and projectin. These proteins are associated with thick, myosin-containing filaments and have characteristic Fn–Ig modules and a kinase domain at the C-terminus; the pattern of Fn–Ig modules changes towards the C-terminus. Insect projectins have an N-terminal extension with tandem Igs and PEVK. (b) SLS protein in different species. The N-terminal region of the gene is kettin which is associated with thin filaments. The *C. elegans* genome has no further sequence downstream. In *D. melanogaster* (probably also in the other insects) and *P. clarkii*, kettin is expressed as an isoform of SLS protein. All kettins have a missing linker at the same position in the Ig-linker region. The complete sequence predicted from the genome is given, though there are several isoforms in *D. melanogaster* and probably all the species. Downstream of kettin, SLS protein has two PEVK domains, tandem Igs, unique sequence, and Fns at the C-terminus. Repeats within the unique region of the *P. clarkii* sequence are dark blue. All the insects and the crayfish have a conserved SH3 domain near the C-terminus. (c) Stretchin is a protein predicted from the *D. melanogaster* genome sequence. It has some similarity to the C-terminal region of SLS proteins, except for the kinase domain. The length of the domains in the figure is proportional to the number of amino acids, except for the Ig and Fn domains which are smaller than shown. Sequences are compiled from FlyBase, Wormpep and the EMBL and NCBI data bases. The *Bombyx mori* sequence was assembled from three EMBL data base entries (Y. Koike et al., unpublished data): AB079865 (Bmkettin) which is a homologue of kettin, AB079866 (Bmtitin) which corresponds to the central region of SLS protein and AB079867 which corresponds to the C-terminal region of SLS protein.

Weber, 1990; Bullard and Leonard, 1996). By analogy with vertebrate titin, it is usually assumed that the N-terminus of projectin is in the Z-disc and the C-terminal region, which has the Fn–Ig modules and the kinase domain, is in the A-band, though this has not been established by labelling with antibodies. The *Drosophila* projectin sequence does not have the Z-repeats near the N-terminus that are characteristic of this region of titin, and it is possible that projectin, unlike titin, does not enter the Z-disc. There are uncertainties about the position of different isoforms of projectin in the sarcomere of different muscle types. In *Drosophila*, the IFM isoform is smaller than the isoforms in leg and jump muscle (Vigoreaux *et al*., 1991), probably because it has less PEVK sequence. There are many muscle types in the insect thorax and it is possible that IFMs are not the only muscles in which projectin spans the I-band; some projectin isoforms with longer PEVK sequence may link thick filaments to Z-disc in fibres with longer sarcomeres.

Kettin

Kettin is an unusual member of the Ig-containing modular proteins in that it is bound to the thin filament along most of the length of the molecule. The *Drosophila* protein is 540 kDa and the complete cDNA sequence codes for a protein of this size. The sequence consists predominately of Ig domains separated by linkers of 35 residues, with the exception of a module two-thirds of the way along the molecule, where the linker is missing. At the N- and C-termini there are consecutive Ig domains with regions of unique sequence. An important difference from projectin is that there are no Fn domains (Hakeda *et al*., 2000; Kolmerer *et al*., 2000 and Figure 2b). The *Anopheles* and *Bombyx* genomic sequences predict these insects have a kettin similar in size and sequence to the *Drosophila* protein (Figure 2b). Kettin has been isolated from the crayfish, *P. clarkii*, (Maki *et al*., 1995) and the cDNA sequence codes for a kettin remarkably similar to the insect proteins (Fukuzawa *et al*., 2001 and Figure 2b). The *C. elegans* genome also has kettin sequence but in this case, there are four fewer Ig-linker modules downstream of the missing linker (Hakeda *et al*., 2000; Kolmerer *et al*., 2000 and Figure 2b). The similarity in the pattern of modules in different kettin sequences is striking: the missing linker is common to all and, in the case of the arthropods, the number of Ig-linker modules is the same.

The uniformity in the primary structure of kettin and the invariance in different species are explained by its position in the sarcomere. The orientation and layout of the molecule have been determined in the IFM of *Lethocerus* and *Drosophila* by labelling fibres with antibodies to different parts of the sequence (van Straaten *et al*., 1999). The N-terminus is in the Z-disc and the molecule extends along the thin filament for about 100 nm in *Drosophila* IFM (40 nm within the Z-disc and 60 nm outside). Kettin binds to actin with high affinity and the stoichiometry between kettin Ig-linker modules and actin monomers is close to one-to-one. The most likely model for kettin on the thin filament, based on the stoichiometry, is one in which the molecule follows the helix of actin and Ig domains, separated by linker sequences, are bound to actin monomers (van Straaten *et al*., 1999).

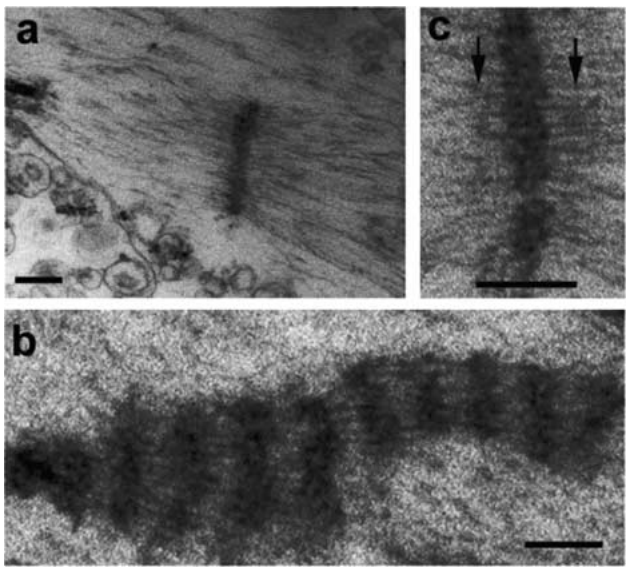

Fig. 3. Kettin in thin filaments. Cryo-sections of IFM in the *Drosophila* mutant Mhc^7 which has no thick filaments in the IFM. Thin filaments emerging from Z-discs vary in length. (a) Low magnification view of Z-disc and thin filaments. (b) Tiger tails formed by regular repeats of Z-discs are commonly seen in sections. The minimum distance from the edge of one Z-disc to the next corresponds to the maximum extent of kettin along an actin filament. (c) Either side of the Z-disc, the thin filaments appear straighter for a distance corresponding to the length of filament spanned by kettin (arrows). Scale bars 200 nm.

In crayfish claw muscle, kettin is close to the borders of the Z-disc and stays in the same position when sarcomeres are stretched to 12 μm (Fukuzawa *et al*., 2001). Since the length of the kettin sequence (Figure 2b) is independent of the length of the I-band, as illustrated by the extremes in insect IFM and crayfish claw muscle, the molecule probably acts to stabilize actin filaments in the region of the Z-disc rather than along the whole length of the I-band. This is supported by the appearance of thin filaments in electron micrographs of a *Drosophila* mutant, Mhc^7 (previously called *Ifm*(*2*)2), which has no thick filaments in the IFM. The length of thin filaments between Z-discs varies, but the minimum distance is about 60 nm, forming characteristic 'tiger-tails' (Figure 3). This suggests that actin is stabilised by kettin where the filaments emerge from the Z-disc. In other sarcomeres, filaments appear straighter and more ordered for a similar distance outside the Z-disc. Kettin promotes the anti-parallel association of actin filaments (van Straaten *et al*., 1999) and this may be the driving force behind the formation of tiger-tails, and possibly also behind the initial formation of

sarcomeres in normal muscle. In *C. elegans* body wall muscle, kettin is associated with dense bodies, the homologues of Z-discs, and here the precise length of the molecule appears to be less critical (Kolmerer *et al.*, 2000 and Figure 2b). The significance of the missing linker in all kettins remains a mystery, though it would be expected to disturb the regular binding of Ig domains to actin.

SLS protein

A gene containing Ig and PEVK domains was identified in *Drosophila* by Andrew and colleagues (Machado *et al.*, 1998) and different regions of the sequence were shown to be expressed in IFM. The gene was called *D-titin* because of the large size of the protein in embyos and the homology to the I-band region of vertebrate titin. It was suggested that, in addition to functioning as an elastic protein in muscle, an isoform of D-titin associated with chromosomes is responsible for their elasticity (Machado *et al.*, 1998; Machado and Andrew, 2000). The *Drosophila* genome sequence shows that D-titin has the whole kettin sequence near the N-terminus (Machado and Andrew, 2000; Zhang *et al.*, 2000 and Figure 2b) and kettin is now known to be an alternatively spliced product. The gene is called *sallimus* (*sls*) in FlyBase and is the locus of *sls* mutations (Kennison and Tamkun, 1988); sallimus is the Finnish word for 'fate' or 'destiny'. As we have seen, the module pattern of projectin is similar to the pattern in vertebrate titin. D-titin is rather different. The whole molecule is predicted to be 1.8–1.9 MDa but there are several alternatively spliced isoforms and it is not clear if the entire sequence is expressed as a single large molecule in muscle. Unlike vertebrate titin, the molecule has no Fn–Ig super repeats and the only Fn domains are the four at the C-terminus (Figure 2b). The lack of titin-like myosin binding modules immediately suggests that the molecule does not extend along the A-band region of the sarcomere (Machado and Andrew, 2000); D-titin is therefore unlikely to be a template regulating thick filament length, as is thought to be the case for vertebrate titin (Trinick, 1994). Surprisingly, nearly half the sequence downstream of kettin is PEVK-like and occurs in two domains. Tandem Igs and unique sequence make up the rest of the C-terminal half of the molecule. Close to the C-terminus there is an SH3 domain, likely to be the site of a proline-rich ligand.

A similar sequence can be identified in the genome of *Anopheles* and in the partial genomic sequence of *Bombyx* (Figure 2b); cDNA from crayfish claw muscle has sequence generally similar to those of the three insects (Fukuzawa *et al.*, 2001 and Figure 2b). The protein needs a name that can be used for all species; D-titin is not appropriate. We will use the general term SLS protein for all species in the tradition of calling proteins after the *Drosophila* gene, and in order not to confuse the protein with vertebrate titin. The crayfish protein has been called I-connectin (Fukuzawa *et al.*, 2001), but again this suggests it is titin (connectin)-like. All the *sls* genes have the sequence of kettin at the N-terminus and, as mentioned above, kettin is a spliced product of the gene which has been identified as a distinct protein in *Lethocerus*, *Drosophila* and crayfish. The *C. elegans* genome has kettin sequence, but unlike other *sls* genes, there is no further sequence downstream. The crayfish is the only species for which the entire cDNA from the *sls* gene has been sequenced and there are some unusual repeats in the region of unique sequence in the C-terminal half of the molecule which are not found in the insect sequences (Fukuzawa *et al.*, 2001 and Figure 2b). The SH3 domain near the C-terminus is conserved in all the SLS proteins, suggesting an essential function. In view of the common overall pattern in the sequence of the genes sequenced so far, we propose the name SLS protein for the gene product in all invertebrates; SLS protein would include the spliced product, kettin.

Several different isoforms of the *Drosophila* SLS protein produced by splicing together different ORFs of the gene, have been identified from the cDNA and by labelling thoracic muscles with antibodies to different regions of the sequence (Bullard *et al.*, 2002). The predominant isoforms in IFM are 540 kDa (kettin), and a less abundant 700 kDa isoform in which the extreme C-terminal region of the gene, including the tandem Igs (Figure 2b), is spliced to the end of kettin. These isoforms are present within the same myofibril (K. Leonard *et al.*, unpublished data). In *Lethocerus*, the molecular weight of the major IFM isoform is 700 kDa (Lakey *et al.*, 1993). Leg muscles, and other non-flight muscles in the *Drosophila* thorax that have sarcomeres with longer I-bands, have larger SLS isoforms in addition to kettin. Antibodies to different regions of the sequence downstream of kettin label the I-band or the edge of the A-band in many different muscle types. Therefore, even the longer isoforms do not extend far into the A-band, as was predicted from the lack of titin-like Fn–Ig repeats in the sequence. The Fn domains at the C-terminus probably bind to the end of the thick filament and may be spliced to isoforms of varying length.

In contrast to *Drosophila* IFM and many of the non-flight muscles, the crayfish claw muscle has only one isoform of SLS protein (I-connectin) in addition to kettin (Fukuzawa *et al.*, 2001). The 1.9 kDa molecule extends to as much as 3.5 μm in stretched sarcomeres of the claw muscle, which is equivalent to the length of the whole sarcomere of *Drosophila* IFM. This is achieved by sequential elongation of different domains downstream of kettin (Fukuzawa *et al.*, 2001). The C-terminus of crayfish SLS protein, which has Fn repeats, is attached to the end of the A-band and the I-band can be stretched until the elastic part of the molecule reaches its contour length. It seems likely that other, shorter, isoforms of SLS protein will be found in crayfish muscles that shorten less than the claw muscle.

These two arthropod examples, insect and crustacean, demonstrate the adaptability of SLS protein. By expres-

sion of different modules, the elasticity is tuned to a stiff muscle like IFM in which the sarcomere is stretched 100 nm or to the extensible claw muscle where the sarcomere is stretched 5 μm. IFM has the additional sophistication of having more than one isoform within a sarcomere.

The genomic sequence of *C. elegans* might be expected to have sequence downstream of kettin but this is not present. In view of the long sarcomeres in the obliquely striated muscle and the potential for considerable extension, it seemed likely that the extensible region of SLS protein would be needed to protect the sarcomeres from excessive stress, as in crayfish claw muscle. However, the linkage of dense bodies to the cuticle may prevent the sarcomeres stretching beyond the extent allowed by the compliance of the cuticle. Thus SLS protein expression is adapted to the needs of muscles attached to the hard exoskeleton of arthropods or to the flexible cuticle of nematodes.

Stretchin-MLCK

A transcription unit in the *Drosophila* genome has been identified which codes for a 926 kDa protein, stretchin-MLCK, predicted to have tandem Ig domains, PEVK and unique sequence, with a kinase domain near the C-terminus (Champagne *et al.*, 2000 and Figure 2c). As well as the full length transcript, an isoform called stretchin, lacking the C-terminal part of the molecule (including the kinase domain) and other isoforms

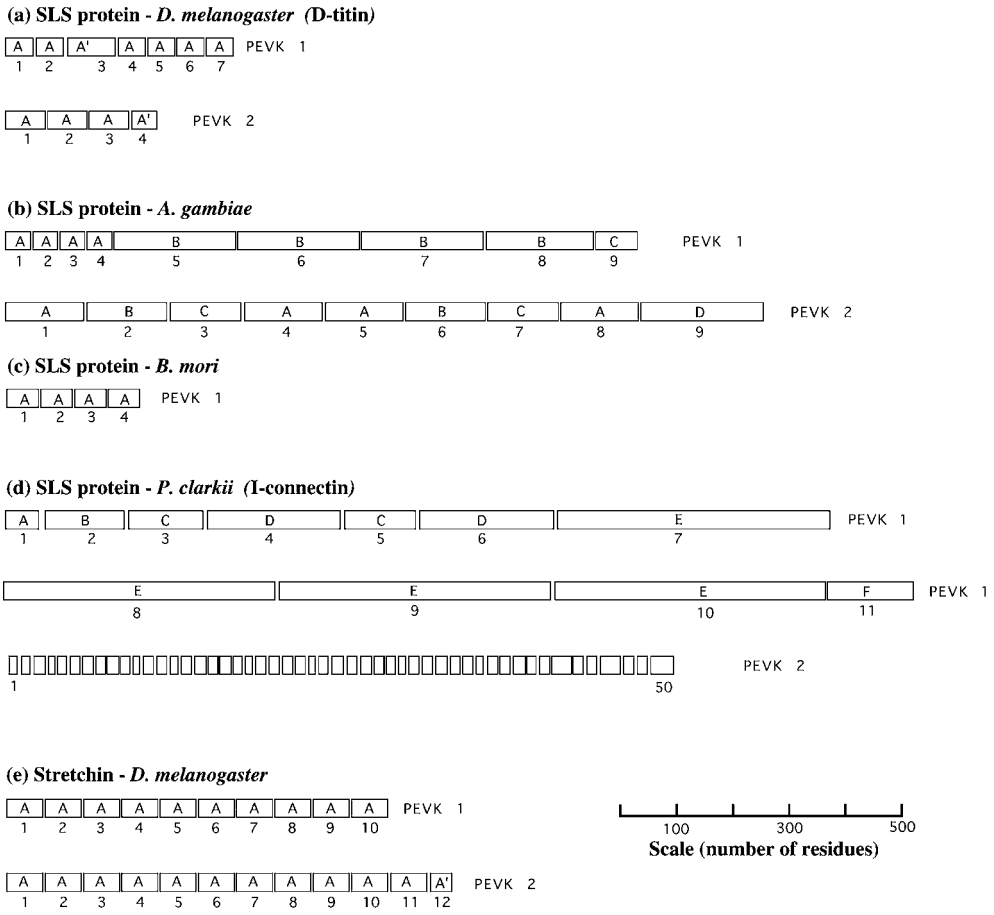

Fig. 4. Repeats in the PEVK sequence of invertebrate elastic proteins. In each case, the sequences are divided into PEVK 1 and 2 regions as shown in Figure 2. Domains are numbered and labelled A, B, C etc. to denote different types of repeat within each group. CLUSTAL X alignments for the PEVK repeats are given on the web site www.krembl.org/SLS (a) SLS protein – *D. melanogaster* (D-titin). PEVK 1 has 6 homologous tandem repeats of approximately 48 residues and a longer one (A′) of 91 residues. PEVK 2 has three highly homologous (>90% identity) repeats of 71 residues and a shorter one (A′) of 45 residues. (b) SLS protein – *A. gambiae*. PEVK 1 has 4 short homologous repeats of 45 residues that are similar to the repeats in PEVK 1 of *D. melanogaster*. These are followed by four longer repeats and a final short repeat. The B-type 216 residue repeats, numbered 5, 6 and 7, have 96, 98 and 98% identity to each other. PEVK 2 has eight repeats of 130–140 residues followed by a ninth 200 residue repeat. There is a pattern of similarity so that repeat 1 is more homologous to 4, 5 and 8, repeat 2 to repeat 6 and repeat 3 to repeat 7, giving the sequence ABCAABCA. (c) SLS protein – *B. mori*. This sequence is incomplete so there are only data for PEVK 1. This has four short repeats (56 residues) that are highly homologous (>90% identity). (d) SLS protein – *P. clarkii* (I-connectin) PEVK 1 has a complex pattern of repeats extending over nearly 2600 residues; six different types of repeat are present. C and D occur twice, E occurs four times. The E repeats are highly homologous despite being about 470 residues in length. E8, E9 and E10 are more than 99% identical at both the protein and DNA level. PEVK 2 has a large number (approximately 50) of short homologous repeats which range in length from 11 to 40 residues. (e) Stretchin – *D. melanogaster*. PEVK 1 and 2 are made up of 21 homologous 64 residue repeats, 10 in PEVK 1 and 11 in PEVK 2; A′ is a shorter repeat. The repeats in PEVK 1 have 90% identity and those in PEVK 2 have 84%.

thought to be myosin light chain kinases (MLCKs) are predicted products of the gene. The only evidence that any of the transcripts is in muscle is that cDNA from the second tandem Ig region codes for partial sequence in a 225 kDa protein, A(225), the antibody to which labels the A-band of *Drosophila* IFM (Patel and Saide, 2001).

Repeating structure of PEVK domains

The SLS proteins are notable for the high proportion of PEVK sequence in two separate domains downstream of kettin; stretchin also has two PEVK domains (Figure 2). Projectin, like vertebrate titin, has one PEVK domain. Since this region is likely to be important in determining the elasticity of the proteins, we have analysed the repeating patterns that are seen in all the longer PEVK domains (but not in projectin). Details of the repeats are given in Figure 4. The pattern in both PEVK domains of *Drosophila* SLS protein is relatively simple; PEVK 2 contains repeats identified previously (Machado *et al*., 1998). In contrast, the pattern in the *Anopheles* protein is more complicated and most of the repeats are longer. Repeats in PEVK domains of crayfish SLS protein (I-connectin) were noted by Fuzukawa *et al*. (2001). Further analysis reveals a varied pattern of repeats that extends over nearly 2600 residues in PEVK 1 and shorter repeats of different lengths in PEVK 2. The entire PEVK 2 is short compared with that of the other species (Figure 2b), but there is an extra domain immediately upstream which has another type of repeating structure not present in other SLS proteins (Fukuzawa *et al*., 2001). Stretchin has a simple pattern of repeats in both PEVK domains (Champagne *et al*., 2000).

There is a large variation in the length of sequence making up a repeat, which ranges from around 11–480 residues. The high degree of homology between repeats is remarkable: three of the 480 residue repeats in the crayfish protein are more than 99% identical at both the protein and the DNA level. This suggests recent duplication of sequence within the gene, because genetic drift would be expected to produce some random sequence variation. Such sequence multiplication may be a mechanism for rapidly extending the length of extensible region in these proteins and changing the elasticity. The designation of domains as 'PEVK' is somewhat subjective and there is variation in the proportion of these residues in a domain: for example, the PEVK domain of soleus titin has 36% proline while stretchin PEVK 1 and 2 have 4.6 and 1.5% proline respectively. CLUSTAL X alignments for the PEVK repeats are given on the web site www.krembl.org/SLS

The diversity in the size and repeating pattern of PEVK repeats in invertebrate elastic proteins is in contrast to the makeup of vertebrate titin, where the single PEVK domain has one type of repeat of 28 residues (Greaser, 2001). This difference is probably due to the large range of sarcomere lengths in invertebrate muscles compared to vertebrate muscles, where sarcomere lengths vary little in different muscle types and in general are less than 2.5 μm. Isoforms of widely differing length and elasticity could be expressed from the *sls* gene.

The origin of passive stiffness in IFM

The structure of the IFM sarcomere

The stiffness of IFM is usually attributed to connections between the ends of the thick filaments and the Z-disc, spanning the short I-bands, the so-called C filaments (Reedy, 1971; Trombitas and Tigyi-Sebes, 1974; White, 1983; Trombitas, 2000). These filaments can be seen

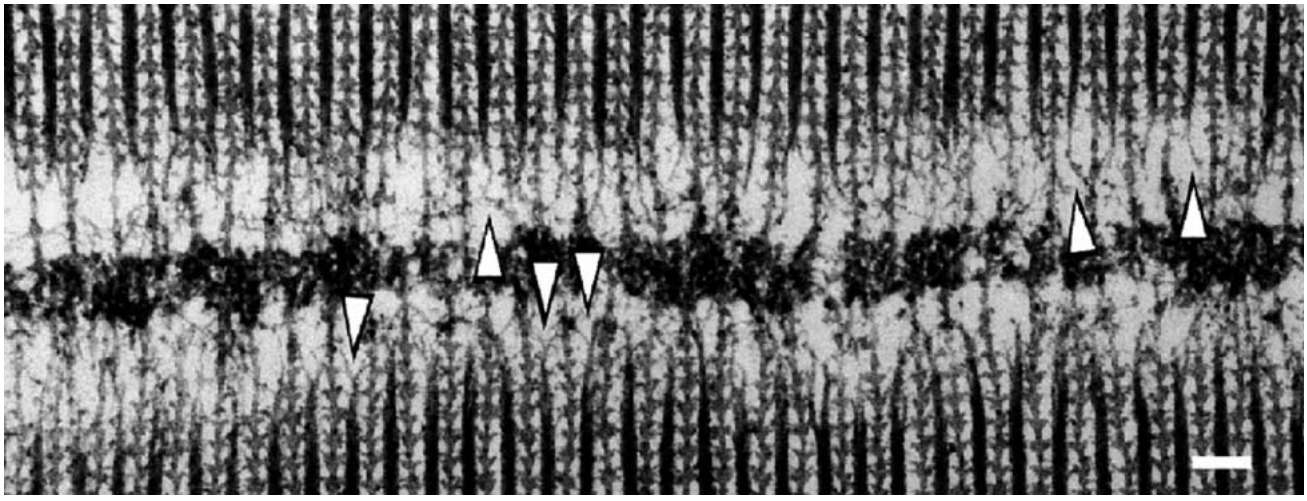

Fig. 5. Elastic filaments in *Lethocerus* IFM. Electron micrograph of a longitudinal section of a *Lethocerus* IFM fibre swollen at low ionic strength in rigor. Arrowheads point to filaments splitting from the ends of the thick filaments and joining actin outside the Z-disc. In less stretched regions, there are dense blobs where the fine filaments join actin. The electron micrograph is by Mary Reedy. Scale bar 100 nm.

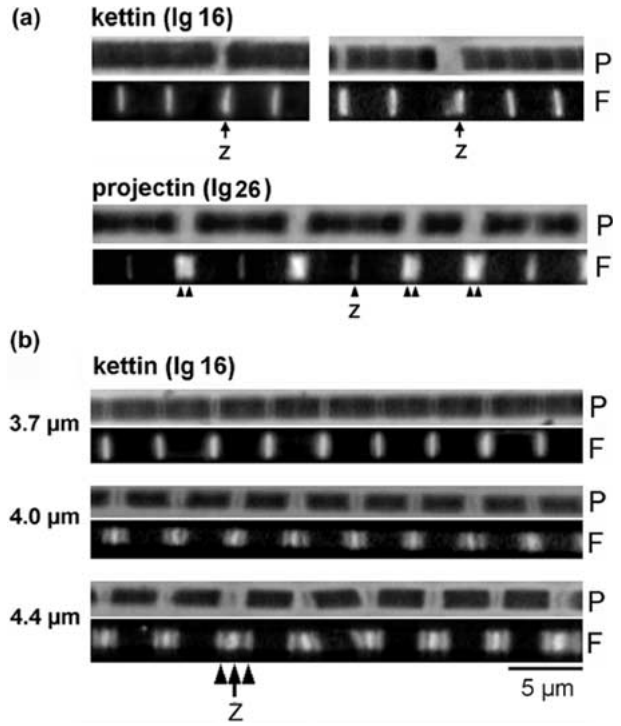

Fig. 6. The position of kettin and projectin in *Drosophila* IFM. (a) Single myofibrils stretched in relaxing conditions, labelled with monoclonal antibody to kettin Ig16 (top) or projectin Ig26 (bottom), followed by fluorescent second antibody. Where sarcomeres are disrupted and thin filaments broken, the kettin epitope is near the Z-disc and the projectin epitope is pulled out from the end of the A-band. (b) Myofibrils from which actin has been extracted with gelsolin at different sarcomere lengths, labelled with anti-kettin Ig16. Both the Z-disc (arrow) and the edge of the A-band (arrowheads) are labelled in stretched myofibrils.

clearly in electron micrographs of fibres stretched in rigor so that thin filaments are broken away from the Z-disc and do not confuse the picture. Projectin is in this region of the IFM sarcomere and was identified as a component of C filaments (Bullard *et al*., 1977; Saide, 1981; Lakey *et al*., 1990). Better preservation of *Lethocerus* IFM during preparation for electron microscopy has meant that the connections between thick filaments and Z-disc can be seen without removing thin filaments. Electron micrographs show strands emerging from thick filaments, which diverge and join the thin filaments before entering the Z-disc (Trombitas, 2000 and Figure 5). There is a dense blob where the strands join the thin filament and this is less pronounced when thick filaments are further away from the Z-disc, as though material were unravelling from it (Figure 5). This arrangement is rather different from the classical picture of C filaments joined directly to the Z-disc and will affect the interpretation of passive stiffness measurements.

Elastic proteins and IFM stiffness

An estimate can be made of the contribution of the two elastic proteins, projectin and SLS protein, to the passive stiffness of single IFM myofibrils. The main isoforms of SLS protein in *Drosophila* IFM are kettin and kettin spliced to sequence from the C-terminal region of the gene. There are also minor amounts of an isoform containing PEVK sequence (Kulke *et al*., 2001). The main isoform of projectin in *Drosophila* IFM is about 900 kDa (Kulke *et al*., 2001).

The position of kettin and projectin in stretched myofibrils can be followed with fluorescently labelled antibody, which shows that kettin remains close to the Z-disc in disrupted sarcomeres where thin filaments are broken (Figure 6a). Projectin is exposed and labels more strongly at the edge of the A-band in stretched sarcomeres, suggesting the labelled epitope, Ig26, is pulled from the A-band by part of the molecule which is still attached to the Z-disc (Kulke *et al*., 2001 and Figure 6a). When kettin is freed from thin filaments by extracting actin with gelsolin, the sarcomeres extend uniformly and while some of the kettin Ig16 epitope remains near the Z-disc, some is on the end of the A-band and moves with it on increasing stretch (Figure 6b). Therefore, the C-terminal part of kettin is attached directly or indirectly to the A-band. The position of projectin is not affected by removing actin, and projectin appears to remain intact when the extracted myofibrils are stretched. Similar results are obtained when kettin is selectively cleaved by µ-calpain. Myofibrils extend easily because thin filaments can be pulled from the overlap region when not restrained by kettin, and the Ig16 epitope is detected at the edge of the A-band after it is cleaved from the rest of the molecule (Kulke *et al*., 2001). Thus, kettin restricts sarcomere extension, and when kettin is rendered inoperable, projectin bears the remaining stress.

The force developed by relaxed IFM myofibrils at increasing sarcomere length is a measure of passive stiffness. Figure 7a shows that single *Drosophila* IFM myofibrils are much stiffer than vertebrate skeletal or cardiac myofibrils (Kulke *et al*., 2001). The myofibrils give after an extension of only about 5% (Figure 7a), which was suggested to be due to breakage of actin and/or loosening up the connections between elastic filaments and their anchorage points at the A-band edge or Z-disc. Alternatively, it is possible that an elastic element unfolds. Ig domains could unfold *in situ* at high physiological sarcomere lengths (Li *et al*., 2002).

A different way to look at stiffness is to oscillate the single myofibril using a micromotor and measure the force response at the force-transducer end (Figure 7b). The average peak-to-valley amplitude of force oscillations indicates myofibrillar stiffness. The passive stiffness of single *Drosophila* IFM myofibrils under control relaxing conditions (SL = 3.7 µm) decreased slowly during a 30-min experiment (Figure 7b, open circles and fit), indicating normal force relaxation. In contrast, stiffness dropped immediately and greatly upon exposure of myofibrils to a Ca^{2+}-independent gelsolin fragment to extract actin (Figure 7b, squares and fit curve). A similar magnitude of stiffness decrease was seen when myofibrils were treated with µ-calpain to specifically degrade kettin (Kulke *et al*., 2001). At time

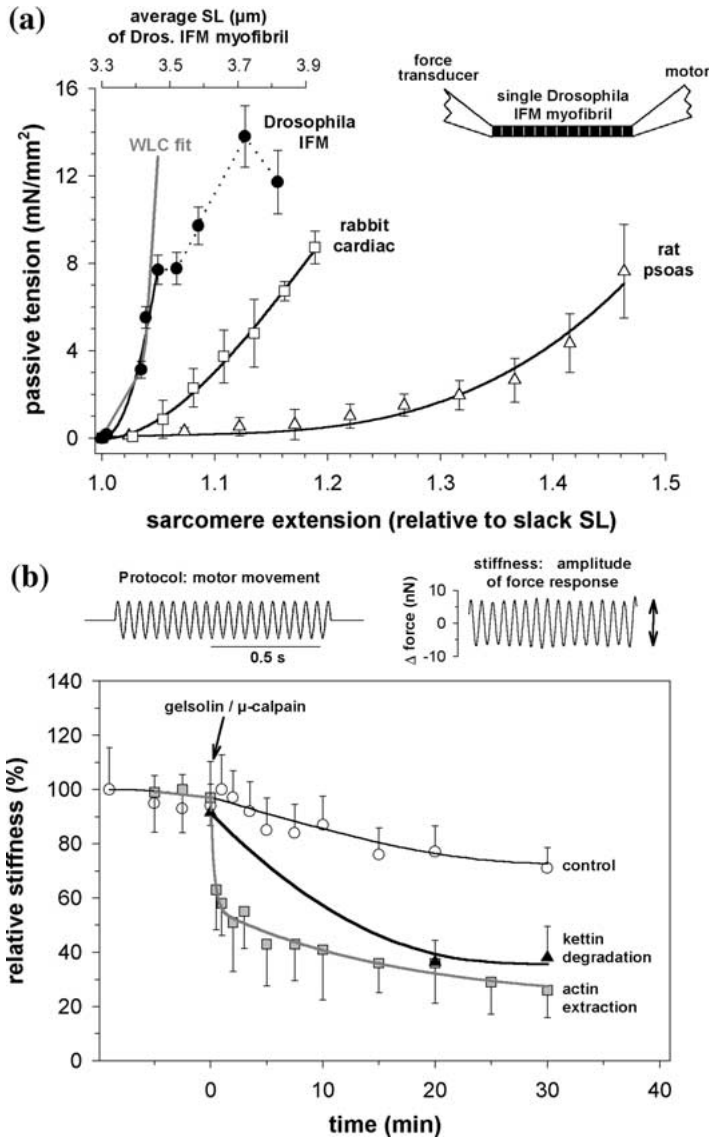

Fig. 7. Force measurements on single, non-activated, myofibrils. (a) Passive-tension *versus* sarcomere-extension relationship of *Drosophila* IFM myofibrils ($n = 24$), in comparison with that of rabbit heart ($n = 17$) and rat psoas ($n = 14$) myofibrils. Pooled data points (mean ± SEM) represent quasi steady-state tension recorded after decay of most of the viscous/viscoelastic force. Black solid curves are third-order polynomial fits. The tension-SL curve of *Drosophila* IFM myofibrils shows several steps at extensions >5% (dotted line). The grey solid curve ('WLC fit') is a fit to the steep initial portion of the *Drosophila* IFM data, calculated using the WLC model of entropic elasticity (Eq. (1)) (parameters: A, 10 nm; L, 126 ± 0.5 nm; z of elastic polymer at slack sarcomere length, 30 nm; scaling factor to obtain force of single molecule, 2.5×10^9/mm^2 cross-sectional area; for further details, see text). (b) Force response of single *Drosophila* IFM myofibrils to 20 Hz sinusoidal motor oscillations (taken as an indicator of stiffness) under different experimental conditions. Stiffness was measured at SL = 3.7 µm (pooled data points are mean ± SD, $n = 3$ myofibrils, for each condition). Fit curves are simple exponential decay functions.

point 30 min, average stiffness of control myofibrils was 71.0 ± 7.7%, statistically significantly higher ($P < 0.001$ in Student's t-test) than that of kettin-extracted (38.0 ± 11.5%) or actin-depleted (26.0 ± 10.1%) myofibrils; differences between the latter two types of specimen were statistically insignificant. Thus, both actin extraction and kettin proteolysis result in a large drop in myofibrillar stiffness. The 'residual' stiffness may be due to the remaining projectin molecules. In *Lethocerus* IFM, which has a 700 kDa isoform of SLS protein (kettin), the myofibrils are less stiff than in *Drosophila* IFM (Kulke *et al.*, 2001); removal of actin with gelsolin also results in a drop in passive stiffness in *Lethocerus* IFM fibres (Granzier and Wang, 1993).

Wormlike-chain behaviour of IFM elastic proteins

We wanted to know whether the force-extension curve of single *Drosophila* IFM myofibrils can be approximated by the wormlike-chain (WLC) model of entropic polymer elasticity, which has been successfully applied to explain the elasticity of titin molecules in vertebrate muscle sarcomeres (e.g. Linke *et al.*, 1998). This model

states (Bustamante *et al*., 1994; Marko and Siggia, 1995) that external force (f) is related to fractional extension (z/L) of a polymer chain by

$$f = \left(\frac{k_B T}{A}\right)\left[\frac{1}{4(1 - z/L)^2} - \frac{1}{4} + \frac{z}{L}\right] \quad (1)$$

where A is the persistence length, k_B is the Boltzmann constant, T is absolute temperature (~300 K in our experiments), z is end-to-end length (extension), and L is the chain's contour length. An entropic chain undergoes thermally-induced bending movements that tend to shorten its end-to-end length. Stretching the chain reduces its conformational entropy and thus, requires an external force. Entropic compliance results when $L \gg A$, due to the numerous configurations the polymer may adopt. The larger the persistence length (at a given contour length), the stiffer the polymer and the smaller the external forces required for stretching out (straightening) the molecule.

The idea here was to calculate the contour length of the polymer chain that makes up the freely extensible molecular segment in the short I-band of *Drosophila* IFM myofibrils, by taking into consideration the myofibrillar force–extension curve and the persistence-length value for a chain consisting solely of serially linked Ig-domains; the latter is known to be 10 nm (Li *et al*., 2002). Then, the calculated contour length should give a hint at how many Ig-domains could, in theory, make up the extensible segment, since one Ig-domain has a length of approximately 4–4.5 nm (Improta *et al*., 1996; Liversage *et al*., 2001).

Accordingly, we tried to fit the initial data points of the *Drosophila* IFM force–extension curve by using the WLC model (Equation (1)) with the following assumptions:

(1) The polymer chain consists of a certain number of serially linked Ig-domains.
(2) The persistence length of this poly-Ig-domain polymer is 10 nm (Li *et al*., 2002).
(3) The end-to-end length of this polymer (=extension) at slack sarcomere length is between 15 and 30 nm. (These values were assumed; the centre-to-centre distance between thick and thin filaments in *Drosophila* IFM is 28 nm (Irving and Maughan, 2000) and the outside-to-outside distance is 14–15 nm; the extensible segment must span at least this distance.)
(4) The myofibrillar force measured at a given extension is scaled to the single molecule by using scaling factors between 1.25×10^9 and $2.5 \times 10^9/\text{mm}^2$ myofibrillar cross-sectional area (the number of thin filaments per mm^2 cross-sectional area is $1.07 \times 10^9/\text{mm}^2$; the stoichiometry between thin and elastic filaments is not known but likely no more than 1:1 to 1:3).

Using these assumptions, and after scaling up the calculated single-molecule data to the myofibril level (using the same scaling factors as above), we obtain a curve of the kind shown in Figure 7a (grey solid line), which reveals a contour length of 126 ± 0.5 nm. Varying the (scaled-down) force and/or the end-to-end length in the WLC model by a factor of two, did not dramatically change the calculated contour-length value. Similarly, variations in persistence length between 8 and 15 nm did not lead to dramatically different contour-length values. The range of contour lengths obtained, was approximately 100–130 nm. This would suggest that the flexible polymer chain of *Drosophila* IFM myofibrils consists of approximately 22–32 Ig-domains.

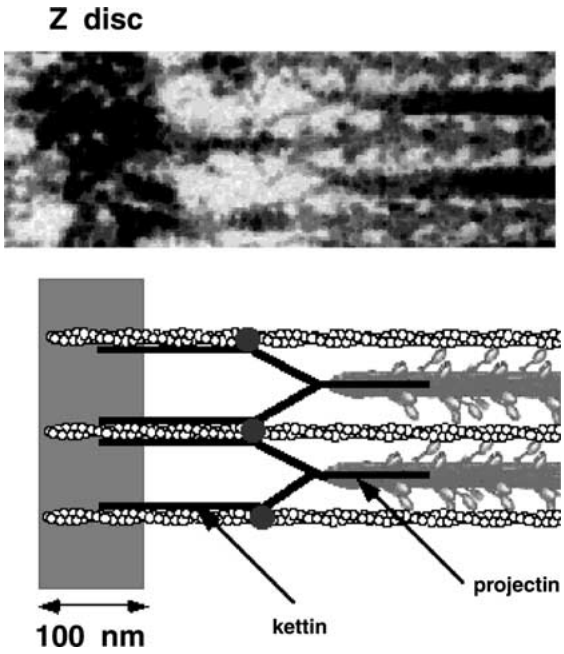

Fig. 8. Model of elastic filaments in *Lethocerus* IFM. Top, part of the electron micrograph in Figure 5. Bottom, suggested position of kettin and projectin. Kettin is bound to actin outside the Z-disc and projectin branches from thick filaments to join the thin filaments. The fine filaments may be made up of both proteins.

Examination of the sequences of SLS protein and projectin (Figure 2) in *Drosophila* shows there are four tandem Ig domains at the end of kettin and a further 18 downstream. If these were spliced together (the C-terminal Igs are known to be spliced to the end of kettin in one IFM isoform), there would be 22 tandem Igs; it is unlikely there is appreciable PEVK because there are no large isoforms of SLS protein in IFM. Thus, there are potentially enough tandem Igs to approximate the contour length of the flexible chain predicted from the force extension curve using the WLC model. There are a maximum of 14 tandem Igs at the N-terminus of projectin, and these may be responsible for the passive stiffness in myofibrils when kettin is broken (Figure 7b). An alternative scenario is that kettin and projectin are somehow linked in series, possibly at the position of the dot-like structures seen in electron micrographs. A possible model based on the micrograph of *Lethocerus* IFM is shown in Figure 8.

Future studies will be necessary to distinguish between these possibilities.

Conclusions

The analysis of sequence domains in the elastic proteins of a nematode and two arthropod groups (insects and crustacea) demonstrates the variety of forms these proteins may take, while being made up of common modules. Closer examination of structure within these modules reveals some are highly conserved, for example those in most of projectin and the whole of kettin; other parts seem freely variable, like the pattern of repeats in PEVK domains. The duplication of long sequence repeats within the PEVK domains is recent and may be a mechanism for increasing the length of the molecule.

Flight muscle needs well regulated stiffness and this is provided by projectin and kettin. The model for entropic polymer elasticity used to predict the elastic properties of titin also applies to IFM, which suggests the basic mechanism is the same in both cases, although IFM has two elastic components and vertebrate striated muscle only has one. IFM is particularly suitable for this study because the muscle has predominately the shorter isoforms of the elastic proteins, without the more complicated arrangement of sequence modules in larger isoforms.

This study of invertebrate proteins puts titin into perspective as an elastic protein adapted to the special requirements of vertebrate muscle. Some of the complexity has been lost and the functions of twitchin or projectin, and SLS protein have been combined in one molecule. Titin appears to have evolved from projectin with the loss of SLS protein in vertebrates. The sequence of twitchin or projectin and SLS protein in the annelids and molluscs, when available, will show if the conserved and highly variable regions that are a feature of the nematode and arthropod proteins, are present in these invertebrates. Comparison of domain sequences in elastic proteins may help to resolve the controversy over the relationship of arthropods to nematodes and annelids.

Acknowledgements

We are very grateful to Mary Reedy for the electron micrograph in Figure 5, and to Toby Gibson for helpful discussion.

Note added in Proof

Recently a *C. elegans* gene coding for an elastic protein of 2.2 MDa has been identified (Flaherty *et al*., J Mol Biol (2002) **323**: 533–549). The protein (called Ce titin) has Ig and Fn domains, a PEVK-like domain and a kinase domain. The domain pattern differs from that of twitchin and SLS protein.

References

Adams MD, *et al*. (2000) The genome sequence of *Drosophila melanogaster*. *Science* **287:** 2185–2195.

Aguinaldo AM, Turbeville JM, Linford LS, Rivera MC, Garey JR, Raff RA and Lake JA (1997) Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387:** 489–493.

Akam M (2000) Arthropods: developmental diversity within a (super) phylum. *Proc Natl Acad Sci USA* **97:** 4438–4441.

Averof M (1997) Arthropod evolution: same Hox genes, different body plans. *Curr Biol* **7:** R634–R636.

Averof M and Cohen SM (1997) Evolutionary origin of insect wings from ancestral gills. *Nature* **385:** 627–630.

Ayme-Southgate A, Vigoreaux J, Benian G and Pardue ML (1991) *Drosophila* has a twitchin/titin-related gene that appears to encode projectin. *Proc Natl Acad Sci USA* **88:** 7973–7977.

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.

Benian GM, Kiff JE, Neckelmann N, Moerman DG and Waterston RH (1989) Sequence of an unusually large protein implicated in regulation of myosin activity in *C. elegans*. *Nature* **342:** 45–50.

Benian GM, L'Hernault SW and Morris ME (1993) Additional sequence complexity in the muscle gene, unc-22, and its encoded protein, twitchin, of *Caenorhabditis elegans*. *Genetics* **134:** 1097–1104.

Blair JE, Ikeo K, Gojobori T and Hedges SB (2002) The evolutionary position of nematodes. *BMC Evol Biol* **2:** 7–13.

Bullard B, Hååg P, Brendel S, Benes V, Qiu F and Leonard K (2002) Mapping kettin and D-titin in the invertebrate sarcomere. *J Mus Res Cell Motil* **22:** 602.

Bullard B, Hammond KS and Luke BM (1977) The site of paramyosin in insect flight muscle and the presence of an unidentified protein between myosin filaments and Z-line. *J Mol Biol* **115:** 417–440.

Bullard B and Leonard K (1996) Modular proteins of insect muscle. *Adv Biophys* **33:** 211–221.

Bustamante C, Marko JF, Siggia ED and Smith S (1994) Entropic elasticity of lambda-phage DNA. *Science* **265:** 1599–1600.

Champagne MB, Edwards KA, Erickson HP and Kiehart DP (2000) *Drosophila* stretchin-MLCK is a novel member of the Titin/Myosin light chain kinase family. *J Mol Biol* **300:** 759–777.

Chan WP and Dickinson MH (1996) *In vivo* length oscillations of indirect flight muscles in the fruit fly *Drosophila virilis*. *J Exp Biol* **199 (12):** 2767–2774.

Conway Morris S (2000) The Cambrian 'explosion': slow-fuse or megatonnage? *Proc Natl Acad Sci USA* **97:** 4426–4429.

Daley J, Southgate R and Ayme-Southgate A (1998) Structure of the *Drosophila* projectin protein: isoforms and implication for projectin filament assembly. *J Mol Biol* **279:** 201–210.

de Rosa R, Grenier JK, Andreeva T, Cook CE, Adoutte A, Akam M, Carroll SB and Balavoine G (1999) Hox genes in brachiopods and priapulids and protostome evolution. *Nature* **399:** 772–776.

Fukuzawa A, Shimamura J, Takemori S, Kanzawa N, Yamaguchi M, Sun P, Maruyama K and Kimura S (2001) Invertebrate connectin spans as much as 3.5 micron in the giant sarcomeres of crayfish claw muscle. *Embo J* **20:** 4826–4835.

Funabara D, Kinoshita S, Watabe S, Siegman MJ, Butler TM and Hartshorne DJ (2001) Phosphorylation of molluscan twitchin by the cAMP-dependent protein kinase. *Biochemistry* **40:** 2087–2095.

Fyrberg CC, Labeit S, Bullard B, Leonard K and Fyrberg E (1992) *Drosophila* projectin: relatedness to titin and twitchin and correlation with *lethal(4) 102 CDa* and *bent-dominant* mutants. *Proc Roy Soc Lond B Biol Sci* **249:** 33–40.

Granzier HL and Wang K (1993) Interplay between passive tension and strong and weak binding cross-bridges in insect indirect flight muscle. A functional dissection by gelsolin-mediated thin filament removal. *J Gen Physiol* **101:** 235–270.

Greaser M (2001) Identification of new repeating motifs in titin. *Proteins* **43:** 145–149.

Hakeda S, Endo S and Saigo K (2000) Requirements of kettin, a giant muscle protein highly conserved in overall structure in evolution, for normal muscle function, viability, and flight activity of *Drosophila*. *J Cell Biol* **148:** 101–114.

Hu DH, Matsuno A, Terakado K, Matsuura T, Kimura S and Maruyama K (1990) Projectin is an invertebrate connectin (titin): isolation from crayfish claw muscle and localization in crayfish claw muscle and insect flight muscle. *J Muscle Res Cell Motil* **11:** 497–511.

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

Irving TC and Maughan DW (2000) In vivo X-ray diffraction of indirect flight muscle from *Drosophila melanogaster*. *Biophys J* **78:** 2511–2515.

Kennison JA and Tamkun JW (1988) Dosage-dependent modifiers of *polycomb* and *antennapedia* mutations in *Drosophila*. *Proc Natl Acad Sci USA* **85:** 8136–8140.

Kiff JE, Moerman DG, Schriefer LA and Waterston RH (1988) Transposon-induced deletions in *unc-22* of *C. elegans* associated with almost normal gene activity. *Nature* **331:** 631–633.

Kolmerer B, Clayton J, Benes V, Allen T, Ferguson C, Leonard K, Weber U, Knekt M, Ansorge W, Labeit S and Bullard B (2000) Sequence and expression of the kettin gene in *Drosophila melanogaster* and *Caenorhabditis elegans*. *J Mol Biol* **296:** 435–448.

Kreuz AJ, Simcox A and Maughan D (1996) Alterations in flight muscle ultrastructure and function in *Drosophila* tropomyosin mutants. *J Cell Biol* **135:** 673–687.

Kulke M, Neagoe C, Kolmerer B, Minajeva A, Hinssen H, Bullard B and Linke WA (2001) Kettin, a major source of myofibrillar stiffness in *Drosophila* indirect flight muscle. *J Cell Biol* **154:** 1045–1057.

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

Lakey A, Labeit S, Gautel M, Ferguson C, Barlow DP, Leonard K and Bullard B (1993) Kettin, a large modular protein in the Z-disc of insect muscles. *Embo J* **12:** 2863–2871.

Lakey A, Ferguson C, Labeit S, Reedy M, Larkins A, Butcher G, Leonard K and Bullard B (1990) Identification and localization of high molecular weight proteins in insect flight and leg muscle. *Embo J* **9:** 3459–3467.

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 Lond* **418:** 998–1002.

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

Liversage AD, Holmes D, Knight PJ, Tskhovrebova L and Trinick J (2001) Titin and the sarcomere symmetry paradox. *J Mol Biol* **305:** 401–409.

Machado C and Andrew DJ (2000) D-TITIN: a giant protein with dual roles in chromosomes and muscles. *J Cell Biol* **151:** 639–652.

Machado C, Sunkel CE and Andrew DJ (1998) Human autoantibodies reveal titin as a chromosomal protein. *J Cell Biol* **141:** 321–333.

Maki S, Ohtani Y, Kimura S and Maruyama K (1995) Isolation and characterization of a kettin-like protein from crayfish claw muscle. *J Muscle Res Cell Motil* **16:** 579–585.

Marko JF and Siggia ED (1995) Statistical mechanics of supercoiled DNA. *Physical Review. E. Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* **52:** 2912–2938.

Moerman DG, Benian GM, Barstead RJ, Schriefer LA and Waterston RH (1988) Identification and intracellular localization of the *unc-22* gene product of *Caenorhabditis elegans*. *Genes Dev* **2:** 93–105.

Nave R and Weber K (1990) A myofibrillar protein of insect muscle related to vertebrate titin connects Z band and A band: purification and molecular characterization of invertebrate mini-titin. *J Cell Sci* **95:** 535–544.

Patel SR and Saide J (2001) A(225), a novel A-band protein of *Drosophila* indirect flight muscle. *Biophys J* **80:** 71a.

Peckham M, Molloy JE, Sparrow JC and White DC (1990) Physiological properties of the dorsal longitudinal flight muscle and the tergal depressor of the trochanter muscle of *Drosophila melanogaster*. *J Muscle Res Cell Motil* **11:** 203–215.

Pringle J (1978) Stretch activation of muscle: function and mechanism. *Proc Roy Soc Lond* B **201:** 107–130.

Reedy MK (1971) Electron microscope observations concerning the behavior of the cross-bridge in striated muscle. In: Podolsky RJ (ed.) *Contractility of Muscle Cells and Related Processes*. (pp. 229–246) Prentice Hall Inc., Englewood Cliffs, NJ.

Saide JD (1981) Identification of a connecting filament protein in insect fibrillar flight muscle. *J Mol Biol* **153:** 661–679.

Saide JD, Chin-Bow S, Hogan-Sheldon J and Busquets-Turner L (1990) Z-band proteins in the flight muscle and leg muscle of the honeybee. *J Muscle Res Cell Motil* **11:** 125–136.

Southgate R and Ayme-Southgate A (2001) Alternative splicing of an amino-terminal PEVK-like region generates multiple isoforms of *Drosophila* projectin. *J Mol Biol* **313:** 1035–1043.

Trinick J (1994) Titin and nebulin: protein rulers in muscle? *Trends Biochem Sci* **19:** 405–409.

Trombitas K (2000) Connecting filaments: a historical prospective. *Adv Exp Med Biol* **481:** 1–23.

Trombitas K and Tigyi-Sebes A (1974) Direct evidence for connecting C filaments in flight muscle of honey bee. *Acta Biochim Biophys Acad Sci Hung* **9:** 243–253.

van Straaten M, Goulding D, Kolmerer B, Labeit S, Clayton J, Leonard K and Bullard B (1999) Association of kettin with actin in the Z-disc of insect flight muscle. *J Mol Biol* **285:** 1549–1562.

Vibert P, Edelstein SM, Castellani L and Elliott BW, Jr. (1993) Minititins in striated and smooth molluscan muscles: structure, location and immunological crossreactivity. *J Muscle Res Cell Motil* **14:** 598–607.

Vigoreaux JO, Saide JD and Pardue ML (1991) Structurally different *Drosophila* striated muscles utilize distinct variants of Z-band-associated proteins. *J Muscle Res Cell Motil* **12:** 340–354.

Waterston RH, Thomson JN and Brenner S (1980) Mutants with altered muscle structure of *Caenorhabditis elegans*. *Dev Biol* **77:** 271–302.

Waterston RH (1988) Muscle. In: Wood WB (ed.) *The Nematode Caenorhabditis elegans*. (pp. 281–335) Cold Spring Harbor Laboratory Press, New York, NY.

White DC (1983) The elasticity of relaxed insect fibrillar flight muscle. *J Physiol* **343:** 31–57.

Zhang Y, Featherstone D, Davis W, Rushton E and Broadie K (2000) *Drosophila* D-titin is required for myoblast fusion and skeletal muscle striation. *J Cell Sci* **113:** 3103–3115.