

This is a repository copy of *Modular proteins from the Drosophila sallimus (sls) gene and their expression in muscles with different extensibility*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/2530/>

Article:

Burkart, C., Qiu, F., Brendel, S. et al. (5 more authors) (2007) Modular proteins from the *Drosophila sallimus (sls)* gene and their expression in muscles with different extensibility. *Journal of Molecular Biology*. pp. 953-969. ISSN 0022-2836

<https://doi.org/10.1016/j.jmb.2007.01.059>

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

promoting access to White Rose research papers



Universities of Leeds, Sheffield and York
<http://eprints.whiterose.ac.uk/>

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/2530/>

Published paper

Burkart, C., Qiu, F., Brendel, S., Benes, V., Hååg, P., Labeit, S., Leonard, K. and Bullard, B. (2007) *Modular proteins from the Drosophila sallimus (sls) gene and their expression in muscles with different extensibility*. *Journal of Molecular Biology*, 367 (4). pp. 953-969.

Accepted Manuscript

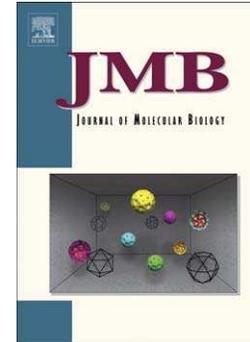
Modular proteins from the *Drosophila sallimus (sls)* gene and their expression in muscles with different extensibility

Christoph Burkart, Feng Qiu, Sigrun Brendel, Vladimir Benes, Petra Hååg, Siegfried Labeit, Kevin Leonard, Belinda Bullard

PII: S0022-2836(07)00112-X
DOI: doi: [10.1016/j.jmb.2007.01.059](https://doi.org/10.1016/j.jmb.2007.01.059)
Reference: YJMBI 59078

To appear in: *Journal of Molecular Biology*

Received date: 30 October 2006
Revised date: 21 January 2007
Accepted date: 23 January 2007



Please cite this article as: Burkart, C., Qiu, F., Brendel, S., Benes, V., Hååg, P., Labeit, S., Leonard, K. & Bullard, B., Modular proteins from the *Drosophila sallimus (sls)* gene and their expression in muscles with different extensibility, *Journal of Molecular Biology* (2007), doi: [10.1016/j.jmb.2007.01.059](https://doi.org/10.1016/j.jmb.2007.01.059)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Modular proteins from the *Drosophila* *sallimus* (*sls*) gene and their expression in muscles with different extensibility

Christoph Burkart^{1,2}, Feng Qiu², Sigrun Brendel², Vladimir Benes², Petra Hååg², Siegfried Labeit¹, Kevin Leonard² and Belinda Bullard^{2,3*}

¹Institut für Anästhesiologie und Operative Intensivmedizin, Universitätsklinikum Mannheim
D-68167 Mannheim, Germany

²EMBL, Meyerhofstrasse 1,
D-69117 Heidelberg, Germany

³Department of Biology, University of York
York YO10 5DD, UK

* Author for correspondence: B. Bullard, Department of Biology, University of York, York
YO10 5DD, UK. Telephone: 44 (0)1904 328823. FAX: 44 (0)1904 328825. Email:
bb500@york.ac.uk

CB and FQ contributed equally to this work.

Present addresses: C. Burkart, The Scripps Research Institute, 10550 Torrey Pines Road, La
Jolla, CA 92037, USA; F. Qiu, CABM, UMDNJ-Robert Wood Medical School, 679 Hoes
Lane, Piscataway, NJ 08854, USA

Abstract

The passive elasticity of the sarcomere in striated muscle is determined by large modular proteins, such as titin in vertebrates. In insects, the function of titin is divided between two shorter proteins, projectin and sallimus (Sls), which are the products of different genes. The *Drosophila sallimus (sls)* gene codes for a protein of 2 MDa. The N-terminal half of the protein is largely made up of immunoglobulin domains and unique sequence; the C-terminal half has two stretches of sequence similar to the elastic PEVK region of titin, and at the end of the molecule there is a region of tandem Ig and fibronectin domains. We have investigated splicing pathways of the *sls* gene and identified isoforms expressed in different muscle types, and at different stages of *Drosophila* development. The 5' half of *sls* codes for zormin and kettin; both proteins contain Ig domains and can be expressed as separate isoforms, or as larger proteins linked to sequence downstream. There are multiple splicing pathways between the kettin region of *sls* and sequence coding for the two PEVK regions. All the resulting protein isoforms have sequence derived from the 3' end of the *sls* gene. Splicing of exons varies at different stages of development. Kettin RNA is predominant in the embryo, and longer transcripts are expressed in larva, pupa and adult. Sls isoforms in the indirect flight muscle (IFM) are zormin, kettin and Sls(700), in which sequence derived from the end of the gene is spliced to kettin RNA. Zormin is in both M-line and Z-disc. Kettin and Sls(700) extend from the Z-disc to the ends of the thick filaments, though, Sls(700) is only in the myofibril core. These shorter isoforms would contribute to the high stiffness of IFM. Other muscles in the thorax and legs have longer Sls isoforms with varying amounts of PEVK sequence; all span the I-band to the ends of the thick filaments. In muscles with longer I-bands, the proportion of PEVK sequence would determine the extensibility of the sarcomere.

Alternative SIs isoforms could regulate the stiffness of the many fibre types in *Drosophila* muscles.

Key words : SIs, kettin, zormin, obscurin, insect muscle, *Drosophila*

Running title: SIs isoforms in *Drosophila* muscles

Abbreviations used: IFM, indirect flight muscle; Ig, immunoglobulin; Fn3, fibronectin 3; *sIs*, *sallimus*; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcriptase polymerase chain reaction; Obs, obscurin; Spec, spectrin; MS-MS, tandem mass spectroscopy; ORF, open reading frame; EST, expressed sequence tag; polyA, polyadenylation; TDT, tergal depressor of the trochanter; Mhc, myosin heavy chain; Mlc1 and 2, myosin light chain 1 and 2.

Introduction

Muscles in the *Drosophila* thorax are adapted for different functions. The largest are the dorsal longitudinal and dorso-ventral indirect flight muscles (IFMs), which provide the power for flight. Smaller muscles control the amplitude and frequency of the wingbeat, and affect the manoeuvrability of the fly.^{1,2} Other muscles in the thorax move the legs. The IFMs produce the upstroke and downstroke of the wings by distorting the thorax, rather than acting directly on the base of the wings. The muscles are activated by periodic stretches and oscillatory contraction results from alternate activation of opposing muscles, together with the resonant properties of the whole thorax. The IFMs can respond rapidly to stretch because they are stiff. The sarcomeres have short I-bands, which change very little in length during the contractile cycle. Other muscles in the thorax and legs have longer I-bands and are expected to be more easily extensible. The relative stiffness of the different types of muscle in the thorax is crucial for flight and for the control of flight.¹

Resting elasticity of striated muscle sarcomeres is largely determined by connecting filaments which link the Z-disc with the ends of the thick filaments.^{3,4} The ease with which the sarcomere can be extended depends on the compliance of the proteins in the filaments. In vertebrate muscles, titin molecules reach from the Z-disc, across half the sarcomere to the M-line. The N-terminal part of the molecule is in the I-band, and can be extended a variable amount, depending on the isoform present in a particular muscle. The C-terminal part of titin is in the A-band and is bound to the thick filaments. In invertebrates, the function of titin is divided between two smaller molecules: the *Drosophila* proteins are projectin and SIs; these are the products of different genes and occur in multiple isoforms.⁵⁻⁸ Projectin is associated with thick filaments in the A-band; most of the molecule is made up of regular repeats containing both immunoglobulin-like (Ig) and fibronectin-like (Fn3) domains, similar to

those in A-band titin. Also like titin, projectin has a kinase domain near the C-terminus. In the N-terminal region of projectin, there are tandem Ig domains, followed by sequence similar to the extensible PEVK sequence in titin, and more tandem Ig domains, a pattern that is similar to that of I-band titin.⁹ In IFM, the N-terminal part of projectin extends from the Z-disc across the short I-band, and is a component of connecting filaments; the Ig-Fn3 modules are associated with the end of the thick filament.^{6,8}

Sls is derived from the *Drosophila salinimus* (*sls*) gene in the chromosomal region 62C2 to 62D1 (FlyBase). The complete sequence, obtained from the *Drosophila* genome, is predicted to code for a protein of about 2.3 MDa. The domain structure of the molecule (previously called D-titin) has been assembled from the genome, together with stretches of cDNA sequence.^{10,11} The molecule has two distinct parts: in the N-terminal half, there are Ig domains and sequence of no defined structure (unique sequence); in the C-terminal half there are two long stretches of PEVK-like sequence, separated by tandem Ig domains and unique sequence; both PEVK regions have motifs of repeating residues that are longer than those in vertebrate titin.^{5,12} At the end of the molecule there are more tandem Igs and a few Fn3 domains. There are no repeating Ig-Fn3 modules like those in the A-band region of projectin and titin, and there is no kinase domain.

Sls can be detected in the early embryo^{10,11,13,14} and is necessary for myoblast fusion.^{10,11,14} cDNA probes from different regions of the *sls* gene showed the same expression pattern throughout embryogenesis,¹⁰ suggesting a full-length Sls is expressed in the embryo. It is not known which isoforms are expressed during later stages of development from embryo to larva and pupa.

Kettin is the most abundant isoform of Sls expressed in adult flies, where it was first identified as an Ig-containing modular protein.¹⁵ Kettin is 527 kDa and the entire sequence is in the N-terminal half of Sls. There are 35 Ig domains, the majority of which are separated by linker sequences.^{13,14} Kettin is bound to actin and is oriented in the sarcomere with the N-

terminus in the Z-disc.¹⁶ In *Drosophila* IFM, the molecule spans the I-band and the C-terminus is attached to the end of the thick filament.^{16,17} Most of the high passive stiffness of IFM is due to kettin.¹⁷

We have investigated splicing pathways in *sls* and find there are a large number of possible isoforms. The principal isoforms in IFM and other thoracic muscles have been identified and the layout of the molecule in the sarcomere has been determined. In addition, we have investigated the expression of SIs isoforms at different stages of *Drosophila* development. The results are consistent with a model in which differential splicing of SIs controls myofibril stiffness in invertebrates in a manner similar to the control of myofibril stiffness by titin in vertebrates.

Results

Sequence of the 5' region of the *Drosophila sls* gene

Initial sequencing of the 5' region of *sls* upstream of kettin by rapid amplification of cDNA ends (RACE) was carried out using primers to the main open reading frames (ORFs) and expressed sequence tags (ESTs). The results are shown in Figure 1. A polyadenylation (polyA) site at the end of zormin was found in the cDNA sequence following the EST **GH18167.3**. The sequence has been deposited in the EMBL database as 'zormin' with accession numbers **AJ544075** and **DME544075**. This sequence corresponds to the smaller of the two zormin isoforms, which lacks ORF 6 (see below).

Full molecular characterisation of the *Drosophila sls* gene

We have used a PCR-based approach to characterize isoforms and screen for splicing pathways between ORFs in the *sls* gene. The gene is in the 62C2 to 62D1 chromosomal region. Initially, 27 ORFs larger than 400 bp were predicted by using DNASTar and the *Drosophila* genome sequence. The main ORFs, are shown in Figure 2. Based on the

predicted ORFs, 26 sense and 26 antisense oligonucleotides were designed to test every possible exon junction by reverse transcription PCR (RT-PCR). There are potentially 144 different isoforms of SIs that could be expressed from the splicing pathways identified. Sequencing the amplified PCR products showed that splice sites follow the conventional pattern;¹⁸ pathways shown in Figure 2 code in-frame.

The RT-PCR results showed that some predicted ORFs could not be joined to any other ORF, and were therefore considered not to be transcribed. Instead, many unpredicted small ORFs of 50 to 250 bp were found to code in-frame, especially in the zornin region, and downstream of ORF 12. Taking these results, together with the DNA sequence of zornin (**AJ544075**) and kettin,^{13,14} we were able to determine the genetic structure of the *sIs* gene (Figure 2). As suggested before,^{10,11} we have found that kettin and SIs are two possible splice variants derived from the same genetic locus. Zornin is another splice variant. The translated zornin sequence includes three spectrin-like domains in the region of *sIs* ORFs 2 and 3; and nine Ig domains, mainly in *sIs* 4 and 5. The 5' region of kettin RNA is coded by a short sequence of 172 bp, which is 2 kb upstream of ORF 9.^{13,14} Joining up the RT-PCR sequences spanning the kettin region gives 35 Ig domains, in agreement with cDNA sequence data already published.^{13,14} Zornin RNA can be expressed as proteins of 388 or 324 kDa (*sIs* 1 to 7, with or without *sIs* 6) and kettin RNA as a 527 kDa protein; both can be expressed as larger proteins produced by splicing to ORFs downstream.

To search for potential polyA sites, the 26 sense primers were tested in combination with poly-dT primers by RACE. PolyA sites were found at the end of kettin RNA and after *sIs* 25. Despite the polyA site after *sIs* 7 at the end of zornin RNA, RT-PCR and subsequent sequencing of products, showed that there is an in-frame splicing pathway between *sIs* 7 and 8. The region coding for the final Ig domain of zornin is spliced out of the product. Similarly, *sIs* 13 in the kettin region can be spliced to *sIs* 14, even though a polyA site was found after *sIs* 13 by RACE. Both these polyA sites may lead to internal termination or be

skipped. In human titin, a similar internal termination leads to a 700 kDa isoform called Novex-3-titin.¹⁹ *s/s* 12 can also be spliced to ORFs downstream of kettin RNA, skipping the region coding for the four Ig domains at the end of kettin. All SIs isoforms, except zornin, include kettin at the N-terminus. Strikingly, all isoforms derived from the *s/s* gene, (except zornin and kettin) also include sequence derived from *s/s* 22 to 25 at the C-terminus (Figure 2).

The layout of domains in the protein corresponding to the full-length SIs is in general agreement with that proposed previously.^{10,11} We find that the region of *s/s* downstream of kettin RNA (from *s/s*14 to 21) has many possible alternatively spliced pathways (Figure 2). Seven Ig domains derived from *s/s*14 to 16 are followed by a long stretch of unique sequence. There are two regions that are predicted to be extensible because there is a high proportion of PEVK sequence. In PEVK 1 (derived from *s/s* 14), 58% of the sequence is PEVK; and in PEVK 2 (derived from *s/s*19 to 21), 53% is PEVK overall, although the sequence derived from the small *s/s* 20 is only 44% PEVK. In comparison, the domain in vertebrate skeletal muscle titin has 70 % PEVK residues. The sequence near the end of SIs includes an SH3 domain immediately after sequence derived from *s/s* 21, followed by eleven Ig domains and five Fn3 domains. Interestingly, the many splicing possibilities mean that selected regions of the sequence can be skipped; for example: the Ig domains derived from *s/s* 15; part, or all, of the unique sequence derived from *s/s*16 to18; PEVK 1 (*s/s*14); part, or all, of PEVK 2 (*s/s*19 or *s/s*19 to 21). In addition, *s/s* 20, coding for sequence with a lower PEVK content, can be skipped to increase the potential extensibility of PEVK 2.

The protein corresponding to the full-length SIs encoded by 25 ORFs larger than 400 bp and 23 ORFs in the range of 50 to 250 bp contains three spectrin-like domains; 63 Ig domains (53 Ig domains excluding zornin); one SH3 domain; five Fn3 domains; and two regions with a high proportion of PEVK residues. The number of Ig domains is in agreement with previously determined sequence, excluding the Ig domains in the zornin region.¹¹

Sequence upstream of kettin RNA, and the corresponding Ig domains in that region were described previously, though not identified as a separate isoform.¹⁰ The molecular weight of the whole SIs protein, consisting of 19,712 residues, was calculated to be 2296 kDa.

Semi-quantitative RT-PCR of developmental stages

Differential expression of zornin and kettin RNA, and other spliced products from the *sIs* gene in embryo, larva, pupa and adult were investigated using a panel of cDNAs prepared from the different stages (Figure 3). The RT-PCR results show that *sIs* 9 to10 in the kettin region are transcribed in the embryo at 12 h, and after 24 h, the level of RNA remains high. This is in agreement with *in situ* hybridization and antibody labelling, where kettin expression was detected in the early embryo from stage 11 (7 h).^{10,11,13,14} Expression of the regions spanning *sIs*13 to14 and 18 to19 was detected from larval stage1 onwards; the level of transcription was quite high at larval stage1 and much lower at larval stage 2 and early stage 3. These amplified regions are only present in the larger isoforms of SIs, and semi-quantitative RT-PCR results showed that, during larval and pupal development, less of these isoforms are expressed than the kettin region alone. Previously, expression of a region of *sIs* downstream of kettin was detected in the embryo by *in situ* hybridization at stage 13 (10 h), showing that transcripts of longer isoforms are present in the early embryo,¹⁰ although even at 24 h they are below the level of detection by the RT-PCR method used here. The expression of ORFs close to the 3' end of *sIs* (*sIs* 22 to 23) was low in the embryo and in the late larval stage, but otherwise similar to that of *sIs* 9 to10 in kettin. This is expected because all SIs isoforms, except zornin and kettin, include sequence derived from both *sIs* 9 to12 and *sIs* 22 to 25.

Two splicing events in the *sIs* gene were investigated to check if the splice isoforms might be differentially expressed during development. An isoform resulting from skipping *sIs*13 to15 (*sIs*12 to16) was only expressed in pupal and adult stages; another isoform resulting from skipping most of the ORFs coding for the PEVK regions (*sIs*15 to 21), was

only expressed in the adult. Both splicing events could be detected weakly after 34 cycles of PCR.

The expression pattern of zormin RNA is similar to that of kettin: transcripts of *sIs* 3 to 4 were detected in the embryo from 12 h, and in subsequent embryonic, larval, pupal and adult stages. Zormin transcripts were detected at low levels and 34 cycles of PCR were needed to amplify cDNA, compared to 30 cycles for other consecutive ORFs in *sIs*.

Expression of zormin RNA is higher in the pupa than in other stages. Primers linking ORFs of zormin and kettin (*sIs* 8 to 9) yield a PCR product at the pupal stage, after 34 cycles of PCR.

High molecular weight proteins in *Drosophila* thoracic muscles

The major proteins of about 250 to 2000 kDa in whole *Drosophila* thoraces and in IFM and legs were analysed by SDS-PAGE with 2.5% acrylamide gels. Some of the proteins have already been identified.^{15,17,20-22} The thorax contains muscles not directly involved in flight, as well as the IFMs. Muscles that do not perform oscillatory contractions, are enriched in thoraces from which IFMs have been removed (Figure 4). Both IFMs and other muscles in the thorax and legs have kettin, and IFMs have an additional larger isoform of about 700 kDa (SIs700), which is not present in the other muscles.¹⁵ Two proteins of about 1000 kDa are seen in the gel: the smaller one is in IFM and the larger one is the predominant isoform in leg muscle and other non-IFMs in the thorax. There are minor components of 1500 kDa and above, which are more abundant in the muscles remaining after IFMs have been removed. Proteins of about 350 and 450 kDa are in IFM but were not seen in gels of other muscles. Some samples of IFM have more 450 kDa protein than others and two preparations are shown in Figure 4.

Immunoblotting can detect proteins not seen in gels stained with Coomassie blue. Isoforms of SIs, many of them minor, were identified in immunoblots of IFM or whole thoraces incubated with antibodies to different regions of the molecules (Figure 5 and Figure

6). Antibodies were raised to recombinant proteins made from constructs derived from the ORFs shown in Figure 2. The splice isoforms, identified by RT-PCR, were assigned to gel bands, based on the size of the isoforms, splicing pathways, and the antibody epitopes in the ORFs. IFM (Figure 5) has two isoforms of zormin predicted to be 324 and 388 kDa. Kettin is the major SIs isoform and the more minor SIs(700) isoform, reacts with kettin antibodies. Surprisingly, SIs(700) includes the Ig and Fn3 domains derived from *s/s* 22 to 25, near the end of the molecule, (antibody to B2 reacts with SIs(700)). SIs(700) does not have the Ig domains at the end of kettin, derived from *s/s* 13 (antibody to KIg34 does not react with SIs(700)). SIs(700) is predicted to be 724 kDa. Larger SIs isoforms were not detected in blots of IFM; therefore, if present, they are in lesser amounts than in other muscles of the thorax (see below).

Muscles in the thorax (Figure 6a) have the two isoforms of zormin seen in IFM. The predominant SIs isoform in the thorax is kettin, and in addition to SIs(700), there is an isoform of 1000 kDa which includes kettin sequence. SIs(1000), like SIs(700), has the Ig and Fn3 domains derived from *s/s* 22 to 25, but not the Ig domains from *s/s* 13 at the end of kettin. There are three SIs isoforms of between 1500 and 2000 kDa. These are less abundant than other SIs isoforms, and are seen best in a gel of thorax from which IFM has been removed (Figure 4). Immunoblots of heavily loaded gels (Figure 6b) show that the three isoforms contain kettin sequence, and sequence from many of the ORFs downstream of kettin: PEVK-like sequence from *s/s* 14,19 and 21; the five Ig domains from the region of *s/s* 15; and unique sequence from *s/s* 16,17 and 18. The largest of the three isoforms includes sequence from *s/s* 18, and because of its size, is predicted to include sequence from *s/s* 16; sequence derived from these ORFs is not present in the other two isoforms. The smallest of the three isoforms also lacks sequence from *s/s* 19.

Identification of proteins by mass spectroscopy

In order to confirm the presence of protein sequence derived from *sls* ORFs, samples were cut out of the bands in a Coomassie-blue stained gel (similar to the one in Figure 4) and analysed by tandem mass spectroscopy (MS-MS). Peptides are listed in Table 1.

Unexpectedly, SIs isoforms were detected in only two of the gel bands analysed. A 250 kDa band above myosin had five peptides found in the sequence of stretchin-Mlck.²³ A band of about 350 kDa had two, and one of 450 kDa had five, peptides in the sequence of a *Drosophila* protein (**CG 30171**) similar to vertebrate obscurin. The *Drosophila* protein is also similar to *Caenorhabditis elegans* UNC-89.^{24, 25} A minor gel band of 700 kDa had five peptides in the SIs sequence: three are from *sls* 9,10 and 12 of kettin, one is from *sls* 16, and one is from *sls* 24 near the C-terminus of SIs. This is consistent with the proposed splicing pathway for SIs(700), which includes *sls* 9 to 12, *sls* 16, and *sls* 22 to 25 (Figure 5). Two bands of about 1000 kDa contained projectin peptides: the smaller isoform was identified by five peptides and the larger by one peptide. IFM contained only the smaller isoform, while leg muscle contained predominately the larger isoform. Thorax, which has IFM and other muscles, had both isoforms. This is in agreement with previous work by Vigoreaux and co-authors, who found that IFM projectin is smaller than projectin in other *Drosophila* muscles.²⁰ A minor 1500 kDa gel band had a peptide from *sls* 12 in kettin and the same peptide from *sls* 24 as SIs(700). This identifies the protein as an SIs isoform.

Position of SIs and obscurin in IFM and other muscles

Immunofluorescence microscopy was used to determine the region of the sarcomere labelled by antibodies to zormin B1 and obscurin, compared to labelling by kettin antibodies. IFM myofibrils were incubated with rabbit antibodies to the N- and C-terminal regions of kettin (K1g3 and K1g34) and the pattern of labelling was compared with that of a rat monoclonal antibody to the middle of kettin (K1g16). Antibody to zormin B1 labelled both the Z-disc and the M-line. Both K1g3 and K1g34 antibodies labelled close to the Z-disc, but separation of the

epitopes could not be resolved by immunofluorescence; there was no sign of any label on the M-line (Figure 7a). In contrast, antibody to obscurin labelled only the M-line. Thus, zornin is in both Z-disc and M-line, and obscurin is only in the M-line, at least in adult IFMs.

In order to determine the position of SIs(700), a minor isoform in IFM, myofibrils were labelled with antibody to the C-terminal region of SIs (B2), which reacts with SIs(700) (Figure 5). Surprisingly, B2 antibody only labelled the core of the myofibril in the region of the Z-disc, while KIg16 antibody labelled across the diameter of the Z-disc. This was seen by conventional fluorescence microscopy, but is clearer in confocal images (Figure 7b). Transverse images of the myofibril show B2 in the core. The distance between B2 and the Z-disc is not resolved by immunofluorescence. SIs(700) includes kettin sequence and KIg16 antibody reacts with this isoform (Figure 5); it is therefore not possible to determine whether the core of the IFM myofibril contains only SIs(700) or both SIs(700) and kettin. For comparison, non-IFM myofibrils were labelled with antibody to B2. In these myofibrils, both B2 and kettin antibodies labelled across the diameter of the myofibril (Figure 7c). B2 was distributed further along the I-band than kettin. The longer SIs isoforms in sarcomeres with longer I-bands would place B2 further from the Z-disc than B2 in IFM.

The position of SIs and obscurin in the sarcomere of IFM and other muscles was determined at higher resolution by immuno-electron microscopy. Antibody to zornin B1 labelled sections of IFM in the middle of the H-zone, in the position of the M-line (Figure 8). There was some diffuse labelling in the Z-disc region but this was not as clear as the M-line labelling. The distribution of gold label across the sarcomere is shown in the histogram in Supplementary Figure S2. The pattern of labelling differs from that seen in fluorescence micrographs, where the Z-disc labelled more strongly than the M-line (Figure 7a). Antibodies to KIg3, KIg16 and KIg34 labelled progressively further from the centre of the Z-disc, and antibody to KIg34 was close to the ends of the thick filaments¹⁷(Figure 8). IFM sections were not labelled by antibodies to regions of SIs downstream of kettin that contained PEVK

sequence (not shown). This agrees with the immunoblots of IFM, in which no labelling by these antibodies was detected.

Antibody to B2 labelled a discrete region of the IFM sarcomere either side of the Z-disc; the label did not extend across the myofibril diameter (Figure 8). This is in agreement with the core labelling seen in fluorescence micrographs (Figure 7b). Cryosections of several specimens of IFM showed a similar distribution of gold label, and the core region containing Sls(700) was estimated to be about 6 to 7% of the total cross-sectional area of the myofibril. Electron micrographs show that the B2 region of Sls(700) is further from the Z-disc than the end of kettin (Klg34), as expected of a longer molecule; B2 also appears to be associated with the ends of thick filaments. Antibody to obscurin labelled the M-line only, as seen in fluorescence micrographs; the distribution of gold was narrower than seen in the M-line labelling with antibody to B1.

Cryosections of thoraces oriented to maximize the content of non-IFM fibres, and cryosections of legs, were labelled with antibodies to Sls and obscurin. A selection of labelled fibres is shown in Figure 9. Antibody to zormin B1 labelled the Z-disc of leg and other non-IFM thoracic muscle fibres; there was no label in the M-line region. Thus, zormin can be in the Z-disc, as in leg muscle, or in the M-line and Z-disc, as in IFM (hence zormin: Z or M). Labelling of non-IFM fibres by antibodies to kettin was less regular than in IFM, but generally similar: Klg3 is in the Z-disc, Klg16 at the edge, and Klg34 outside the Z-disc. Because there is an appreciable I-band, the Ig domains at the end of kettin (Klg34) do not reach the thick filaments. Antibodies to the C-terminal half of Sls (B4 and B5) label irregularly, some way out from the Z-disc. The spread of gold particles is probably due to variable extension of the PEVK sequence. In the most extended molecules, B4 and B5 are at the ends of the thick filaments. Antibody to the B2 region of Sls was spread over the ends of the thick filaments and, unlike the labelling of IFM, was distributed across the diameter of the myofibril. The labelling pattern shows the Sls molecule extends from the Z-disc, across the I-

band to the thick filament. Antibody to obscurin labelled a broad region in the middle of the sarcomere in non-IFM thoracic muscle fibres, and did not label the Z-disc (Figure 9).

Therefore, obscurin is in the same region of the sarcomere in IFM and non-IFM fibres, unlike zormin, which can be in the M-line or the Z-disc.

The tergal depressor of the trochanter (TDT, or jump muscle) is in the thorax and moves the middle leg of the fly. Myofibrils in the TDT have a regular structure and short I-bands. Electron micrographs of labelled cryosections of this muscle are shown in Supplementary data (Figure S1). The labelling pattern was similar to that of other non-IFM fibres: zormin is in the Z-disc, kettin extends some way outside the Z-disc, B4 and B5 sequences are in the I-band, and B2 sequence is at the ends of the thick filaments. The position of different regions of SIs in isoforms present in the IFM and non-IFM sarcomere, and the position of obscurin, are summarized in Figure 10.

Binding of zormin to thick filaments and actin

Zormin was seen in the Z-disc and in the M-line in fluorescence and electron micrograph images of IFM, suggesting the protein might bind both to actin and to thick filaments. Intact thick filaments were isolated from a mutant *Drosophila* in which the IFMs have no thin filaments. The recombinant B1 fragment of zormin (consisting of three Ig domains) was co-sedimented with the thick filaments and with actin (Figure 11); therefore these three Igs are capable of binding to both thick and thin filaments. Overlay dot-blot assays (not shown) confirmed B1 binding to thick filaments and to actin; B1 did not bind to tropomyosin, showing that there was no general non-specific binding.

Discussion

The aims of this investigation were to identify isoforms of SIs in IFM and other muscles; and to find out where different regions of the molecule are in the sarcomere, and how this might

be related to the function of the muscle. SIs isoforms are produced from a single gene. We have shown that a wide variety of splice isoforms are generated from this single locus. Although not all of the different transcripts are necessarily translated, we have confirmed the expression of some of the major isoforms at the protein level. We have identified isoforms in flight and non-flight muscles in the *Drosophila* thorax, and found which are expressed at different developmental stages from early embryo to larva, pupa and adult.

The 5' half of *sIs* produces two discrete proteins with no sequence in common: zormin and kettin. Zormin has three spectrin-like domains, which may function as spacers, like the repeats in α -actinin.²⁶ There are two major isoforms of zormin in IFM and other muscles. Labelling IFM fibres with antibody to zormin, showed the protein is present both in the M-line and in the Z-disc by immunofluorescence; the strong labelling at the M-line and more diffuse distribution at the Z-disc, seen in electron micrographs, may be due to more labile binding to the Z-disc. In leg and other non-IFM fibres, including the TDT, zormin is only in the Z-disc. This dual localisation is confirmed by the finding that Ig domains in the middle of the molecule are capable of binding both to thick filaments and to actin. M-line and Z-disc may have different isoforms of zormin, but the function is not known.

Kettin is the major SIs isoform expressed in all *Drosophila* muscles. The Ig-linker domains that make up most of the molecule are bound to actin and reinforce the thin filament in the region of the Z-disc; the molecule extends along thin filaments for about 100 nm outside the Z-disc^{16,27}. In IFM, the N-terminus of the molecule is in the Z-disc, the middle is at the edge of the Z-disc, and because the I-band is short, the tandem Igs at the C-terminus reach the end of the thick filaments^{16, 17} (Figure 8). In leg and other non-IFM fibres, the orientation of kettin is the same, but the C-terminus is in the I-band because the molecule is not long enough to reach the thick filaments (Figure 9). The shorter I-band of the TDT means the end of kettin is closer to the end of the thick filaments. Zormin and kettin transcripts can be spliced together (Figure 2), but no isoform containing both zormin and kettin sequence

was detected in immunoblots of IFM or muscles in the whole thorax (Figures 5 and 6). If the two were linked, kettin would be expected to be in the M-line with zormin in IFM and this is not observed (Figure 7a).

Kettin can be linked to the C-terminal half of SIs to give longer isoforms; all of which have kettin at the N-terminus. The SIs(700) isoform is specific to the IFM, and is unlike other longer SIs isoforms in having no PEVK sequence. This would make it less easily extensible. The final tandem Ig domains of kettin are skipped in SIs(700) and replaced with tandem Igs and Fn3 domains from the B2 region at the end of SIs. This results from two splicing events that cut out PEVK and most of the unique sequence (Figure 2). SIs(700) is in the core of the IFM myofibril and the C-terminal region is at the end of the thick filaments (Figures 7b and 8). IFM therefore has two SIs isoforms that reach the ends of thick filaments: kettin linked to thick filaments by the four C-terminal Ig domains,¹⁷ and SIs(700) linked by the 16 Ig and Fn3 domains from the end of SIs. The net difference in the lengths of the molecules would be about 50 nm, if each Ig and Fn3 domain is 4 nm.²⁸ The SIs isoforms in IFM are estimated to be responsible for 70% of the passive stiffness,¹⁷ but it is not clear what the effect of having a longer molecule at the core of the myofibril would be on the overall stiffness.

The many splicing pathways in the 3' half of *sIs* produce isoforms of varying length, with varying amounts of extensible PEVK sequence. All longer isoforms have the domains from the B2 region at the C-terminus. Muscles in the thorax, which include flight control muscles and leg muscles, have a variety of larger isoforms that are not present in IFM. We have not established which isoforms are in particular muscles, but it is likely that fibres with longer sarcomeres and wider I-bands have larger isoforms. In leg muscle fibres, the somewhat scattered position of different regions of SIs in the I-band is probably due to different degrees of extension in the PEVK sequence (Figure 9). In the TDT fibres, the I-bands are narrower, and sequence in the C-terminal half of SIs is less extended. In all non-IFM fibres, the B2 region of SIs is at the ends of the thick filaments, and is distributed across

the whole diameter of the myofibril, which would produce uniform elasticity. One relatively abundant isoform, Sls(1000), has only a short region of PEVK sequence, and would be less extensible than the larger isoforms; Sls(1000) is like Sls(700) in that sequence coding for the four tandem Ig domains at the end of kettin is spliced out of the *sIs* transcripts. The larger isoforms all have the full sequence of kettin and would therefore have two potential thick-filament binding sites. However, the relatively wide I-band in the sarcomere of leg muscle fibres means the end of kettin is distant from the thick filaments (Figure 9). The SH3 domain near the end of SIs would be expected to bind a ligand having poly-proline sequence, but this has not been identified as yet. The domain is in all isoforms, except kettin and SIs(700), both of which are in IFM.

Isoforms of a protein similar to vertebrate obscurin and *C. elegans* UNC-89 were identified in IFM by mass spectroscopy. The larger of two isoforms predicted from the genome sequence in Flybase (with six additional Ig domains)²⁴ is about 421 kDa. This probably corresponds to the 450 kDa isoform identified here. There are 21 Ig domains, two Fn3 domains, and two kinase domains near the C-terminus; but the *Drosophila* protein has none of the signalling domains found in vertebrate obscurin and UNC-89. Vertebrate obscurin is in the Z-disc in the early stages of myofibrillogenesis in cardiac cells, and in the M-line at later stages.²⁹ The protein may also be in Z-disc and M-line simultaneously in skeletal fibres, where it is associated with the periphery of the myofibril.³⁰ In contrast, *Drosophila* obscurin is in the M-line of the sarcomere at all stages of development, from embryo and larva to adult (not shown), and immunolabelling of cryosections cut through the interior of the myofibril shows the protein is present throughout the M-line. *Drosophila* obscurin may be a structural protein necessary for A-band assembly, but the relationship to zomin, also in the M-line of IFM, is not known.

During development, *sIs* transcripts are differentially expressed at different stages. Kettin RNA is the major transcript in the embryo. The preponderance of kettin RNA, and

sequence derived from the B2 region of *sIs* in early larval, and pupal stages, when new muscles are being assembled, is consistent with the finding that the majority of isoforms begin and end with these sequences. Larger transcripts, with sequence downstream of kettin RNA, are mostly formed by assembling consecutive ORFs. Transcripts that skip sequence coding for the last four Ig domains of kettin (producing SIs(700) and SIs(1000)), and others in which large stretches of sequence coding for PEVK sequence are missing, are only detected in pupal and adult stages; the corresponding SIs protein isoforms may only be in the adult thorax. Similarly, splicing between zormin and kettin transcripts, which could also produce larger isoforms, was only faintly detected in the pupa, suggesting these isoforms are rare. Zormin RNA is expressed during embryogenesis at 12 h, though the strongest expression is in the pupal stage. Recently zormin has been shown to interact with Rols7. Rols7 is found in fusing myoblasts in the embryo, and in Z-discs during larval stages, and also at the site of attachment of the muscles to the epidermis.³¹ It has been suggested that Rols7, zormin and α -actinin form a complex during assembly of the Z-disc. Although kettin also appears in the early embryo, and is necessary for myoblast fusion,^{10,11,14} it is not part of the complex of Rols7 with zormin and α -actinin.³¹ Therefore, two isoforms of SIs: zormin and kettin, have different functions in the early development of muscles.

The mechanics of flight and a non-flight muscle in *Drosophila* have been compared. Stretch-activation of IFM and the TDT and the stiffness of both types of muscle fibre in relaxing conditions were measured.³² The TDT responded very little to stretch, and the resting stiffness of the fibres was only 14% that of IFM fibres. Although, like IFM, the TDT has a short I-band, SIs in this muscle has extensible sequence found in some of the longer isoforms (Figure S1, Supplementary data), which would make the fibres more compliant than IFM fibres. The TDT initiates flight by straightening the middle leg and making the fly jump. During this sudden action, the I-band in an extended muscle would shorten rapidly; longer

Sls isoforms with elastic elements would give the I-band the optimum compliance for rapid shortening and extension.

In all types of muscle fibres, Sls is firmly anchored to the thin filaments near the Z-disc by kettin, and anchored to the ends of thick filaments by sequence at the end of the molecule. Intermediate sequence, varying in length and extensibility, is in connecting filaments. This arrangement can accommodate variation in I-band length in different muscles, and the need for different extensibility and elasticity of the sarcomere. In IFM, a link between the end of kettin and thick filaments, and the presence of projectin in connecting filaments, would increase the stiffness of the sarcomere.^{8,17} For connecting filaments to be effective in determining the passive elasticity of the sarcomere, the thick filaments must be relatively inextensible. In IFM, flightin on the outside of the thick filament, and paramyosin in the core are needed to maintain high passive stiffness. The elastic modulus was reduced in oscillating fibres lacking flightin,³³ and in fibres in which phosphorylation sites of paramyosin were mutated;³⁴ passive stiffness of myofibrils was also reduced in the paramyosin mutants.³⁵

Recently it has been shown that the Ig domains in kettin³⁶ and projectin³⁷ can refold under high forces. Unfolding in a few domains would protect thick and thin filaments in the IFM sarcomere from being damaged during stretching, and the domains would refold while fibres were still under tension. It is also possible that kettin and projectin could function as folding-based springs during oscillatory contractions.^{36,37}

In conclusion, the *Drosophila* *sls* gene produces many isoforms that have essential functions in different processes, from myoblast fusion in the embryo to regulation of fibre stiffness in the adult. *Drosophila* has proved to be a good system in which to study the function of Sls, because of the variety of muscle types. We have found that the length of Sls isoforms and the proportion of elastic sequence are greater in muscles with long sarcomeres, which are more easily extensible. The function of kettin and the longer Sk isoforms may differ in muscles with highly extensible sarcomeres. Kettin does not span the I-band in these

muscles, but acts to reinforce the thin filament near the Z-disc. The longer SIs isoforms, which do span the I-band, are likely to be largely responsible for the elastic properties. IFM is a special case in which kettin and SIs(700) both span the short I-band, resulting in the high stiffness essential for oscillatory contraction. The power output of flight muscles depends on the resonant property of the thorax, and this is regulated by control muscles.² The relative elasticity of flight and control muscles will be an important element in determining the power produced during oscillatory contraction of flight muscles.

Materials and Methods

Fly stocks

Wildtype *Drosophila melanogaster* were Oregon-R strain. A mutant lacking IFM thin filaments was *Act88F^{KM88}*.

Sequencing of zormin cDNA

Drosophila cDNA for RACE reactions was prepared from adult mRNA using a Marathon cDNA amplification kit from Clontech. Identification of the 5' end (N-terminal sequencing) was carried out using the Invitrogen GeneRacer kit; all other RACE amplification reactions were done with the Clontech Marathon cDNA kit. RACE primers used for zormin cDNA sequencing are given in Table S1 (Supplementary data).

Sequence analysis and verification of *sIs* splice isoforms using RT-PCR

DNASTar was used to predict ORFs in the 62C2 - 62D1 region of the *Drosophila* genome. In this way, 27 ORFs larger than 400 bp were detected and used for a systematic analysis of splicing by a reverse transcriptase–polymerase chain reaction (RT-PCR) approach. Adult *Drosophila* cDNA was screened for splice isoforms with combinations of 26 sense and 26 reverse primers, in order to check every possible exon junction (see Table S2, Supplementary data, for primer sequences). The synthetic oligonucleotides of ~30 bases in length were

initially synthesized by Sigma (formerly known as Sigma-ARK) and later by Operon Biotechnologies. For the screening PCR, AmpliTaq® (Applied Biosystems) was used. Instead of the PCR buffer supplied, the following buffer was used: 10 x Taq polymerase buffer with 100 mM Taps (pH 8.5), 500 mM KCl, 20 mM MgCl₂, 0.2 % Triton X-100 (Serva). If AmpliTaq® did not yield any PCR product, Combizyme® DNA polymerase (Invitex, Berlin) was used in a second PCR run. The cDNA for the RT-PCR was synthesized using wild-type *Drosophila* total RNA and Superscript II (Invitrogen). An analytical PCR reaction (20µl) typically contained 2 µl 10x Taq polymerase, 1.6 µl dNTP (2.5 mM each), 0.4 µl of each oligonucleotide (50 µM), 0.2 µl DNA template (~2 ng), 0.1 µl AmpliTaq® (5U/µl) and 15.3 µl H₂O. PCR reactions using Combizyme® DNA polymerase were performed as recommended by the supplier. To detect small concentrations of transcripts, the number of cycles was set to 34. The following PCR conditions were used: 20 s at 95°C for denaturation, followed by 2-6 min at 68°C for annealing/extension. For each kb of target cDNA, 2 min at 68°C was estimated to be necessary. Amplified PCR products were cloned using the TOPO TA Cloning® Kit (Invitrogen) and subsequently sequenced by Genecore (EMBL, Heidelberg).

PolyA sites in transcripts from the end of the kettin region, and from the end of the *sls* gene were identified by RACE with *Drosophila* cDNA. Primer FQ29 was used for 5'RACE of kettin cDNA and primer FQ312 for 5'RACE of *sls* (see Supplementary data Table S2). Immunoglobulin, fibronectin, spectrin-like and SH3 domains in the SIs amino acid sequence were identified using the SMART database.

Semi-quantitative RT-PCR from developmental stages

First strand cDNA was synthesized from total RNA prepared from different developmental stages, using StrataScript® (Stratagene). PCR reactions were carried out using *Taq* Polymerase (New England Biolabs). Specific primer pairs were used: for ORFs in the zormin

(*sIs* 3 to 4), kettin (*sIs* 9 to 10), and other *sIs* regions (*sIs* 8 to 9, 13 to 14, 18 to 19 and 22 to 23). Primer pairs were also used to detect spliced transcripts with skipped ORFs (*sIs* 12 to 16 and 15 to 21). Primers for the *Drosophila* housekeeping gene, *RP49*, were used to check that the amount of cDNA was approximately the same for different developmental stages. The PCR products were detected on a 1.5% (w/v)-agarose gel with ethidium bromide staining. The primer sequences used for RT-PCR are listed in Table S2 (Supplementary data).

Cloning and expression of constructs from the *sIs* and *obscurin* genes

Constructs from exons in the *sIs* and *obscurin* genes were cloned and protein expressed in *Escherichia coli*; the position of the constructs in the sequence of *sIs* is shown in Figure 2. Cloning and expression of the following peptides have been described previously: B1, Ig4-Ig6 in zormin (SIg4-SIg6);³⁷ KIg3, Ig3-Ig4 in kettin (KET1);¹³ KIg34, Ig34-Ig35 at the end of kettin (Ig34/35).¹⁷ New constructs were amplified by PCR with *Drosophila* adult cDNA as template. The peptides, and the *sIs* sequence from which they were derived, are as follows: Spec, a spectrin-like sequence from exon 2 of the zormin region; PIg, PEVK sequence from exon 14, with an Ig domain at the C-terminus; B3, four Ig domains from exon 15; B4, sequence derived from splicing between exons 17 and 19; B5, unique sequence from exon 18; Cb1, PEVK sequence from exon 19; B2, three Ig domains from exon 22. A construct from the *Drosophila obscurin* (*obs*) gene was cloned, which coded for the three Ig domains Ig14-Ig16 in the *obscurin* sequence. DNA sequences of constructs were checked before expression. Amino acid sequences at the N- and C-termini of expressed fragments of zormin (GenBank accession number [AJ544075](#)), SIs ([AJ245406](#) and [AF241652](#)) and *obscurin* (FlyBase number [CG30171](#)) are shown in Table S3 (Supplementary data).

Zormin, SIs and *obscurin* cDNA constructs were cloned into a pET9d expression vector with a His₆-tag at the N-terminus of the protein, or a pETM11 vector with a TEV protease cleavage site between the His₆-tag and the protein (both vectors from Novagen). Vectors were transformed in to *E. coli* strain BL21(DE3)pLysS (Spec, PIg, B3, B4 and B2)

or BL21(DE)RP (Cb1 and Obs). Soluble protein was purified from the lysate of sonicated cells on a Ni-NTA agarose column (Qiagen); in some cases, fragments were purified further by ion exchange chromatography on a Mono-Q column (Amersham).

Production of antibodies

Polyclonal antibodies to recombinant proteins Spec, B1, PIg, B3, B4, B5 and B2 were raised in rabbits at EMBL. Specific IgG was affinity-purified on a column of the antigen coupled to CM-Sepharose (Pharmacia).³⁸ Antibodies to Cb1 and Obs were raised in rabbits and IgG was isolated from serum by Biogenes (Berlin). Antibody to KIg16 was a rat monoclonal antibody (MAC155), raised to *Lethocerus* Z-discs,³⁹ the antibody reacts with the recombinant linker-Ig-linker fragment of *Drosophila* kettin that includes KIg16.¹⁵ A mouse monoclonal antibody raised to KIg34 (called Ket94; mouse ID 540715) was used in some immunolabelling of cryosections for electron microscopy.⁴⁰

SDS-gel electrophoresis and immunoblotting

SDS-PAGE with 12% acrylamide and Laemmli sample buffer was used for detecting proteins up to 200 kDa. High molecular weight proteins in *Drosophila* muscles were analysed by SDS-PAGE using 2.5% acrylamide gels strengthened with 1.5% agarose.⁴¹ Thoraces were dissected from about 100 flies and frozen immediately in liquid N₂; they were homogenised while frozen and suspended in 100 µl Laemmli sample buffer containing 20 µM leupeptin.¹⁷ Legs were removed from about 200 flies and processed similarly. IFMs were dissected from 50 thoraces and homogenised in 50 µl sample buffer without freezing. After heating at 95°C for 3 min, samples were centrifuged and the supernatant was run on mini-gels 8 cm long; gels were stained with Coomassie brilliant blue R. SIs isoforms in *Drosophila* thoraces and IFMs were identified in immunoblots of 2.5% gels. Proteins were transferred to nitrocellulose by electrophoresis at 700 mA for 6 h in a buffer containing 0.1% SDS. The nitrocellulose was incubated in antibodies to different regions of the SIs sequence

(Figure 2); then in goat anti-rabbit or goat anti-rat secondary antibody (Dianova). Blots were developed with a chemiluminescent substrate (ECL, Amersham).

Immunofluorescence microscopy

Drosophila thoraces were cut in half by a vertical cut through the mid-line and incubated in relaxing solution (20 mM phosphate buffer (pH 7.0), 5 mM MgCl₂, 5 mM EGTA, 5 mM ATP, 5 mM DTT, 1 mM PMSF, 0.2 mM leupeptin) with 50% glycerol and 0.5% Triton X-100 for 2 h on ice. Dorsal longitudinal muscles were dissected from ten thoraces and homogenised gently in an Eppendorf tube. Myofibrils were washed twice in relaxing solution without glycerol, then in buffer without glycerol and Triton. A drop of myofibrils on a microscope slide was incubated in blocking buffer (relaxing solution with 1% BSA) for 30 min. For double labelling, a mixture of two antibodies (diluted 1:50 in blocking buffer) was added to the myofibrils; after 1 h, myofibrils were washed and incubated in mixed secondary antibodies (diluted 1:50). Myofibrils were also prepared from flies expressing GFP-tagged SIs;⁴² these myofibrils were incubated in anti-obscurin and anti-rabbit secondary antibody. Secondary antibodies (Dianova) were anti-rabbit and anti-rat, conjugated with FITC or Texas Red. Slides were examined in a Zeiss Axioscope microscope with a Photonic Science Coolview colour CCD camera. Confocal microscopy was carried out in the EMBL Advanced Light Microscopy Facility. Image series were acquired on a Leica SP2 laser scanning confocal microscope using a 100 x oil-immersion objective. Images were analysed and displayed in 3D using Leica and EMBL software.

Immunoelectron microscopy

Drosophila thoraces were cut in half by a vertical cut, which exposed the dorsal longitudinal muscles, or by a horizontal cut, which exposed more of the non-flight muscles in the ventral half of the thorax. Half thoraces were fixed in 4% paraformaldehyde in relaxing solution and infused with 2.1 M sucrose; they were then placed on copper stubs with the cut surface uppermost and frozen in liquid N₂.³⁹ Groups of three legs were removed from a fly and

treated in the same way; they were aligned side-by-side on the stub before freezing.

Cryosections were blocked and labelled with IgG (diluted 1:10 to 1:100) in rigor buffer with 0.1% Triton X-100, followed by 10 nm Protein A gold. In the case of the rat monoclonal antibody, a rabbit anti-rat bridging antibody (Cappel Laboratories) was used before Protein A gold.^{15,39} Sections were picked up dry on carbon-coated copper grids, thawed and stained with methyl cellulose and uranyl acetate.⁴³ They were examined in a Fei Biotwin electron microscope operating at 80 or 100 kV and images taken with a Gatan 1Kx1K CCD camera.

Mass spectroscopy

Proteins in the major bands in 2.5% polyacrylamide gels were identified by mass spectroscopy. Samples were cut from gel bands and digested with trypsin.⁴⁴ The tandem mass spectrometer used at EMBL for MALDI-TOF (MS-MS) was a Q-ToF 2TM (Waters, Manchester, UK). Protein identification was by sequence tag analysis against SwissProt and TREMBL databases using 'Peptide Search'. The tandem mass spectrometer used at the University of York was an Applied Biosystems 4700 Proteomics Analyser. Proteins were identified by searching the NCBI database, using MASCOT. The *sIs* gene was searched for peptide sequences in order to identify the ORFs containing particular peptides.

Protein binding assay

Binding of zornin B1 fragment (three Ig domains) to thick filaments and actin was measured by co-sedimentation. Thick filaments were prepared from the *Act88F^{KM88}* mutant *Drosophila*, which lack thin filaments in the IFM.⁴⁵ Thick filaments from IFMs of 35 flies were washed and resuspended in 100 μ l of binding buffer (50 mM NaCl, 10 mM Tris-Cl, pH 7.1, 2 mM MgCl₂, 2 mM DTT). The suspension (10 μ l) was added to B1 (7 μ M) in a total volume of 80 μ l of binding buffer, and incubated at room temperature for 1 h. Actin was prepared from rabbit back muscle.⁴⁶ Actin (4 μ M) was added to B1 (7 μ M) in 100 μ l of binding buffer and incubated at room temperature for 1 h. Samples of thick filaments alone, actin alone and B1 alone were treated similarly. The mixtures were centrifuged in a Beckman Airfuge for 30 min

at 28 psi. SDS-PAGE sample buffer containing 6 M urea was added to supernatants and pellets.

Acknowledgements

We are grateful to Dr Thomas Franz, Proteomic Core Facility, EMBL and Adam Dowle, Technology Facility, Department of Biology, University of York for mass spectroscopy measurements; to Natascha Bushati, EMBL, for RNA from different stages of *Drosophila* development; to Pieranna Chiarella and Alan Sawyer, Monoclonal Antibody Core Facility, EMBL for monoclonal antibody; and to Dr Timo Zimmermann, Advanced Light Microscopy Facility, EMBL for assistance with confocal microscopy. This work was partly funded by an EU 6th Framework NOE grant (Myores) and by the DFG (LA668/7-2).

References

1. Dickinson, M. H. & Tu, M. S. (1997). The Function of Dipteran Flight Muscle. *Comp. Biochem. Physiol.*, 116A, 223-238.
2. Lehmann, F. O. (2006). Muscle Systems Design and Integration. In *Nature's Versatile Engine: Insect Flight Muscle Inside and Out* (Vigoreaux, J., ed.), pp. 230-239. Landes Bioscience/Eurekah.com, Georgetown.
3. Granzier, H. L. & Labeit, S. (2005). Titin and its associated proteins: the third myofilament system of the sarcomere. *Adv Protein Chem* 71, 89-119.
4. Tskhovrebova, L. & Trinick, J. (2003). Titin: properties and family relationships. *Nat Rev Mol Cell Biol* 4, 679-89.
5. Bullard, B., Linke, W. A. & Leonard, K. (2002). Varieties of elastic protein in invertebrate muscles. *J Muscle Res Cell Motil* 23, 435-47.
6. Ayme-Southgate, A., Saide, J., Southgate, R., Bounaix, C., Cammerato, A., Patel, S. R. & Wussler, C. (2005). In indirect flight muscles *Drosophila* projectin has a short PEVK domain and its NH₂-terminus is embedded at the Z-band. *J Muscle Res Cell Motil* 26, 467-477.
7. Southgate, R. & Ayme-Southgate, A. (2001). Alternative splicing of an amino-terminal PEVK-like region generates multiple isoforms of *Drosophila* projectin. *J Mol Biol* 313, 1035-43.
8. Bullard, B., Burkart, C., Labeit, S. & Leonard, K. (2005). The function of elastic proteins in the oscillatory contraction of insect flight muscle. *J Muscle Res Cell Motil* 26, 479-85.
9. Labeit, S. & Kolmerer, B. (1995). Titins: giant proteins in charge of muscle ultrastructure and elasticity. *Science* 270, 293-296.
10. Machado, C. & Andrew, D. J. (2000). D-Titin: a giant protein with dual roles in chromosomes and muscles. *J Cell Biol* 151, 639-652.
11. Zhang, Y., Featherstone, D., Davis, W., Rushton, E. & Broadie, K. (2000). *Drosophila* D-titin is required for myoblast fusion and skeletal muscle striation. *J Cell Sci* 113 (Pt 17), 3103-3115.
12. Machado, C., Sunkel, C. E. & Andrew, D. J. (1998). Human autoantibodies reveal titin as a chromosomal protein. *J Cell Biol* 141, 321-333.
13. Kolmerer, B., Clayton, J., Benes, V., Allen, T., Ferguson, C., Leonard, K., Weber, U., Knekt, M., Ansorge, W., Labeit, S. & Bullard, B. (2000). Sequence and expression of the kettin gene in *Drosophila melanogaster* and *Caenorhabditis elegans*. *J Mol Biol* 296, 435-48.
14. Hakeda, S., Endo, S. & 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.
15. Lakey, A., Labeit, S., Gautel, M., Ferguson, C., Barlow, D. P., Leonard, K. & Bullard, B. (1993). Kettin, a large modular protein in the Z-disc of insect muscles. *Embo J* 12, 2863-71.
16. van Straaten, M., Goulding, D., Kolmerer, B., Labeit, S., Clayton, J., Leonard, K. & Bullard, B. (1999). Association of kettin with actin in the Z-disc of insect flight muscle. *J Mol Biol* 285, 1549-62.
17. Kulke, M., Neagoe, C., Kolmerer, B., Minajeva, A., Hinssen, H., Bullard, B. & Linke, W. A. (2001). Kettin, a major source of myofibrillar stiffness in *Drosophila* indirect flight muscle. *J Cell Biol* 154, 1045-57.

18. Shapiro, M. B. & Senapathy, P. (1987). RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res* 15, 7155-74.
19. Bang, M. L., Centner, T., Fornoff, F., Geach, A. J., Gotthardt, M., McNabb, M., Witt, C. C., Labeit, D., Gregorio, C. C., Granzier, H. & 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.
20. Vigoreaux, J. O., Saide, J. D. & Pardue, M. L. (1991). Structurally different *Drosophila* striated muscles utilize distinct variants of Z-band-associated proteins. *J Muscle Res Cell Motil* 12, 340-354.
21. Patel, S. R. & Saide, J. D. (2005). Stretchin-klp, a novel *Drosophila* indirect flight muscle protein, has both myosin dependent and independent isoforms. *J Muscle Res Cell Motil* 26, 213-24.
22. Saide, J. D., Chin-Bow, S., Hogan-Sheldon, J., Busquets-Turner, L., Vigoreaux, J. O., Valgeirsdottir, K. & Pardue, M. L. (1989). Characterization of components of Z-bands in the fibrillar flight muscle of *Drosophila melanogaster*. *J Cell Biol* 109, 2157-2167.
23. Champagne, M. B., Edwards, K. A., Erickson, H. P. & Kiehart, D. P. (2000). *Drosophila* stretchin-MLCK is a novel member of the Titin/Myosin light chain kinase family. *J Mol Biol* 300, 759-777.
24. Small, T. M., Gernert, K. M., Flaherty, D. B., Mercer, K. B., Borodovsky, M. & Benian, G. M. (2004). Three new isoforms of *Caenorhabditis elegans* UNC-89 containing MLCK-like protein kinase domains. *J Mol Biol* 342, 91-108.
25. Sutter, S. B., Raeker, M. O., Borisov, A. B. & Russell, M. W. (2004). Orthologous relationship of obscurin and Unc-89: phylogeny of a novel family of tandem myosin light chain kinases. *Dev Genes Evol* 214, 352-9.
26. Djinic-Carugo, K., Gautel, M., Ylanne, J. & Young, P. (2002). The spectrin repeat: a structural platform for cytoskeletal protein assemblies. *FEBS Lett* 513, 119-23.
27. Bullard, B., Leake, M. C. & Leonard, K. (2006). Some Functions of Proteins from the *Drosophila* *sallimus* (*sls*) Gene. In *Nature's Versatile Engine: Insect Flight Muscle Inside and Out* (Vigoreaux, J., ed.), pp. 177-184. Landes Bioscience/Eurekah.com, Georgetown.
28. Trinick, J., Knight, P. & Whiting, A. (1984). Purification and properties of native titin. *J Mol Biol* 180, 331-356.
29. Young, P., Ehler, E. & Gautel, M. (2001). Obscurin, a giant sarcomeric Rho guanine nucleotide exchange factor protein involved in sarcomere assembly. *J Cell Biol* 154, 123-36.
30. Kontogianni-Konstantopoulos, A., Jones, E. M., van Rossum, D. B. & Bloch, R. J. (2003). Obscurin Is a Ligand for Small Ankyrin 1 in Skeletal Muscle. *Molecular Biology of the Cell* 14, 1138-1148.
31. Kreisköther, N., Reichert, N., Buttgerit, D., Hertenstein, A., Fischbach, K. F. & Renkawitz-Pohl, R. (2006). *Drosophila* Rolling pebbles colocalises and putatively interacts with alpha-Actinin and the Sls isoform Zormin in the Z-discs of the sarcomere and with Dumbfounded/Kirre, alpha-Actinin and Zormin in the terminal Z-discs. *J Muscle Res Cell Motil.* 27, 93-106.
32. Peckham, M., Molloy, J. E., Sparrow, J. C. & White, D. C. (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.

33. Henkin, J. A., Maughan, D. W. & Vigoreaux, J. O. (2004). Mutations that affect flightin expression in *Drosophila* alter the viscoelastic properties of flight muscle fibers. *Am J Physiol Cell Physiol* 286, C65-72.
34. Liu, H., Miller, M. S., Swank, D. M., Kronert, W. A., Maughan, D. W. & Bernstein, S. I. (2005). Paramyosin phosphorylation site disruption affects indirect flight muscle stiffness and power generation in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 102, 10522-7.
35. Hao, Y., Miller, M. S., Swank, D. M., Liu, H., Bernstein, S. I., Maughan, D. W. & Pollack, G. H. (2006). Passive Stiffness in *Drosophila* Indirect Flight Muscle Reduced by Disrupting Paramyosin Phosphorylation, but not by Embryonic Myosin S2 Hinge Substitution. *Biophys J*.
36. Leake, M. C., Wilson, D., Bullard, B. & Simmons, R. M. (2003). The elasticity of single kettin molecules using a two-bead laser-tweezers assay. *FEBS Lett* 535, 55-60.
37. Bullard, B., Garcia, T., Benes, V., Leake, M. C., Linke, W. A. & Oberhauser, A. F. (2006). The molecular elasticity of the insect flight muscle proteins projectin and kettin. *Proc Natl Acad Sci U S A* 103, 4451-6.
38. Harlow, E. & Lane, D. (1988). *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
39. Lakey, A., Ferguson, C., Labeit, S., Reedy, M., Larkins, A., Butcher, G., Leonard, K. & Bullard, B. (1990). Identification and localization of high molecular weight proteins in insect flight and leg muscle. *Embo J* 9, 3459-67.
40. De Masi, F., Chiarella, P., Wilhelm, H., Massimi, M., Bullard, B., Ansorge, W. & Sawyer, A. (2005). High throughput production of mouse monoclonal antibodies using antigen microarrays. *Proteomics* 5, 4070-4081.
41. Tatsumi, R. & Hattori, A. (1995). Detection of giant myofibrillar proteins connectin and nebulin by electrophoresis in 2% polyacrylamide slab gels strengthened with agarose. *Anal Biochem* 224, 28-31.
42. Morin, X., Daneman, R., Zavortink, M. & Chia, W. (2001). A protein trap strategy to detect GFP-tagged proteins expressed from their endogenous loci in *Drosophila*. *PNAS* 98, 15050-15055.
43. Tokayasu, K. T. (1980). Immunocytochemistry on ultrathin frozen sections. *Histochem. J.* 12, 381-390.
44. Rosenfeld, J., Capdevielle, J., Guillemot, J. C. & Ferrara, P. (1992). In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal Biochem* 203, 173-9.
45. Clayton, J. D., Cripps, R. M., Sparrow, J. C. & Bullard, B. (1998). Interaction of troponin-H and glutathione S-transferase-2 in the indirect flight muscles of *Drosophila melanogaster*. *J Muscle Res Cell Motil* 19, 117-27.
46. Bullard, B., Bell, J., Craig, R. & Leonard, K. (1985). Arthrin: a new actin-like protein in insect flight muscle. *J Mol Biol* 182, 443-54.

Figure legends

Figure 1 . Diagram showing RACE sequencing of zornin cDNA. The numbers below open reading frames are the 17 exons. Numbers above the larger open reading frames are the *sls* ORF numbers (see Figure 2). RACE extensions are shown as blue arrows; sequences obtained from EST **GH18167.5** and **GH18167.3** are shown by green arrows. This is the zornin-PA isoform, which lacks exon 6. The red line is the region used to express the peptide for producing B1 antibody.

Figure 2 . Schematic structure of the *Drosophila sls* gene. Zornin and kettin are isoforms derived from the *sls* gene. Zornin RNA can be spliced to kettin RNA, which can be spliced to sequence downstream to produce larger isoforms. The exon links were identified by RT-PCR, and subsequently confirmed by sequencing. Splicing pathways are indicated by green arrows. Sequence domains are shown, and the location of peptides to which antibodies were raised (black arrows). Internal promoters and terminations and poly-A sites could not be detected by the RT-PCR method used. The poly-A site near the end of zornin RNA was found in an EST sequence (Figure 1); the poly A sites at the end of kettin and Sls RNA were identified by RACE (see Methods).

Figure 3 . Expression of Sls RNA at different developmental stages. Oligonucleotides derived from the coding region of the *sls* gene were used in RT-PCR reactions with cDNA from developmental stages. The following regions were amplified: zornin (*sls* 3 to 4); zornin to kettin (*sls* 8 to 9); kettin (*sls* 9 to 10); ORFs in the PEVK region (*sls* 13 to 14 and 18 to 19); ORFs at the end of Sls (*sls* 22 to 23); and regions in which ORFs are skipped (*sls* 12 to 16 and 15 to 21). PCR reactions were run for 30 or 34 cycles; PCR products run in an agarose gel are

shown. *RP49* was used to check that approximately equal amounts of cDNA were used for amplification. Embryonic stages are: 1 h, 12 h, 24 h (lanes 1 to 3). Larval stages are: first instar, second instar, and early third instar (lanes 4 to 6). Pupal stage is in lane 7 and the adult in lane 8.

Figure 4. High molecular weight proteins in *Drosophila* muscles. Proteins in IFM, leg and thorax are compared by SDS-PAGE (2.5 % polyacrylamide gel). Lanes 1 and 4, IFM; lane 2, whole thorax (including IFM and legs); lane 3, leg; lane 5, thorax from which IFMs have been removed. IFM has a lower molecular weight isoform of projectin than the other muscles. Obscurin is in IFM and whole thorax, but was not detected in leg and non-IFM thoracic muscles. There are two isoforms of obscurin (1 and 2); the amount of obscurin 1 varies in different IFM samples (lanes 1 and 4). Asterisks show high molecular weight proteins, which are visible in the thorax without IFMs. Brackets show bands analysed by mass spectroscopy (Table 1).

Figure 5. Isoforms of SIs in *Drosophila* IFM. Immunoblots of IFM separated by 2.5 % SDS-PAGE were incubated in antibodies to SIs; the epitopes of the antibodies are shown in Figure 2 and above each lane. The splice isoforms (based on RT-PCR) corresponding to gel bands, are shown on the left; arrows indicate antibody-binding sites. IFM contains two isoforms of zornin (antibody to B1). Kettin is the major isoform of SIs in IFM (antibodies to KIg3, KIg16 and KIg34) and there are lesser amounts of SIs(700). SIs(700) has kettin sequence and the Ig domains in *sls* 22, but not those in *sls* 13 (antibody to B2 reacts but antibody to KIg34 does not).

Figure 6. Isoforms of SIs in *Drosophila* thoracic muscles. Immunoblots of whole thoraces separated by 2.5% SDS-PAGE were incubated in antibodies to SIs; epitopes of antibodies are

shown in Figure 2 and above each lane. The splice isoforms (based on RT-PCR) corresponding to gel bands, are shown on the left; arrows indicate antibody-binding sites. (a) Gels were loaded to show the major bands on blots. Thorax has two isoforms of zornin (antibodies to Spec and B1). Kettin is the major isoform of SIs in the thorax (antibodies to KIg3, KIg16 and KIg34). SIs(700) and SIs(1000) have kettin sequence and the Ig domains derived from *s/s* 22, but not those from *s/s* 13 (antibody to B2 reacts and antibody to KIg34 does not). There are higher molecular weight SIs isoforms (see below) that include the Ig domains from *s/s* 15 (antibody to B3) and the Ig domains from *s/s* 22 (antibody to B2). (b) Gels were loaded with three times the amount of sample used in (a), to show minor high molecular weight bands; only bands above 1000 kDa are shown. The three high molecular weight isoforms contain kettin sequence (antibodies to KIg3 and KIg34), and sequence from *s/s* 14,15,17, and 22 (antibodies to PIg, B3, B4, and B2). The smallest of the three isoforms lacks sequence from *s/s* 19 (antibody to Cb1) and the largest isoform is the only one that includes sequence from *s/s* 18 (antibody to B5). The approximate molecular mass of SIs isoforms is shown.

Figure 7. The position of SIs isoforms and obscurin in the sarcomere. (a) The position of zornin and obscurin in IFM relative to kettin. Immuno-fluorescence micrographs in each panel show myofibrils labelled with rabbit antibody to peptides in zornin (B1), kettin (KIg3 and KIg34) or obscurin (Obs) (top images). Myofibrils incubated with antibodies to B1, KIg3 and KIg34 were double-labelled with rat monoclonal antibody to kettin KIg16 (middle image), to mark the Z-disc (arrow). Myofibrils incubated with antibody to obscurin expressed GFP-SIs (see Methods), which marked the Z-disc. The merged image is at the bottom of each panel. Antibody to zornin B1 labels the Z-disc and the M-line; antibodies to kettin KIg3 and KIg34 label the Z-disc exclusively, and antibody to obscurin labels the M-line. (b) The position of the C-terminus of SIs in IFM. Confocal fluorescence micrographs show

myofibrils double-labelled with rabbit antibody to B2, (top image), and rat monoclonal antibody to kettin K1g16 (middle image) to mark the Z-disc; the merged image is at the bottom. Confocal image layers were combined to give the cross-sectional view on the right. Antibody to B2 labels the core of the IFM myofibril, close to the Z-disc, and antibody to kettin labels across the diameter of the myofibril. (c) The position of the C-terminus of SIs in a bundle of non-IFM myofibrils in the thorax. Confocal fluorescence micrographs show myofibrils labelled as in (b). Both B2 and kettin antibodies label across the diameter of the myofibril. B2 has a wider distribution along the axis of the sarcomere than kettin, and the antibody labels the I-band either side of the Z-disc (arrow). Secondary antibodies were conjugated with fluorescein isothiocyanate or Texas red. Scale bars 5 μm .

Figure 8. The layout of SIs and obscurin in the IFM sarcomere. Electron micrographs of cryosections that were labelled with antibodies and Protein A-gold. The peptides to which antibodies were raised are shown in Figure 2. Antibody to zormin B1 is at the M-line and there is also some labelling in the Z-disc region (see Supplementary data Figure S2 for a histogram of gold particle distribution). Antibody to K1g3 is at the Z-disc; antibody to K1g16 is on the edges of the Z-disc; and antibody to K1g34 is close to the ends of the thick filaments. Antibody to B2 (near the end of SIs(700)) is in the core region of the myofibril, at the end of the thick filaments. Antibody to obscurin (Obs) is on the M-line. Scale bar 0.5 μm .

Figure 9. The layout of SIs and obscurin in leg and other non-IFM sarcomeres. Electron micrographs of cryosections that were labelled with antibodies and Protein A-gold. Peptides to which antibodies were raised are shown in Figure 2. Antibody to zormin B1 is at the Z-disc. Antibodies to kettin K1g3, K1g16 and K1g34 are at the Z-disc, or either side of the Z-disc. The distribution of gold particles is more spread out than the labelling in the IFM sarcomere. Antibodies to B4 and B5 extend further from the Z-disc than the kettin antibodies,

and there is label on the ends of the thick filaments (arrows). Antibody to B2 labels close to the ends of the thick filaments, across the entire myofibril diameter. Antibody to obscurin (Obs) is in the middle of the sarcomere; there is no clear M-line and the label has a wider spread than in IFM. Cryosections are all of legs, except the one labelled with antibody to obscurin, which is thoracic muscle. Scale bar 0.5 μm .

Figure 10. Schematic diagram summarizing the position of SIs isoforms and obscurin in the IFM and non-IFM sarcomere. The positions of antibody epitopes in different isoforms of SIs are shown, based on reaction of the proteins in immunoblots, and labelling observed by immuno-electron microscopy. The approximate molecular mass of SIs isoforms is shown. Isoforms in IFM are zornin, kettin, and the minor isoform SIs(700) in the myofibril core. Zornin has been detected in both the M-line and the Z-disc. SIs isoforms in non-IFMs, with longer sarcomeres, are zornin, kettin and larger isoforms with elastic PEVK sequence; two of the larger isoforms, SIs(1700) and SIs(1900) are shown.

Figure 11. Binding of zornin fragment B1 to thick filaments and actin. B1 has three Ig domains in the zornin sequence (Figure 2). Binding of B1 to thick filaments and actin are compared by SDS-PAGE (12 % polyacrylamide gel). Thick filaments or actin (4 μM) were incubated with and without B1 (7 μM) and then pelleted by centrifuging. Lanes 1 and 2 show pellets and supernatants of thick filaments alone and with B1. Lanes 3, show B1 alone. Lanes 4 and 5 show actin with and without B1. The amount of the pellets loaded on the gels is equivalent to 5 times the amount of the supernatants. The B1 fragment binds to both thick filaments and to actin. P, pellet; S, supernatant; M, molecular mass markers; Mhc, myosin heavy chain; Mlc1 and 2, myosin light chain 1 and 2.

Table 1. Proteins identified by mass spectroscopy of gel bands

p250 Stretchin-Mlck	p350 Obscurin 1
VYAAQADGDESEPIFALPLR	FLIDIEPNGLLR
FSVQQAQISDSGTYFVVAR	VLNTVAGGPTPTQLLR
GQTPAVQWFK	
LEVYENPGTGDVPPTFTR	
WFFGDQPIAFGPR	
CG18255	CG33519 or CG30171
p450 Obscurin 2	p700 Sl(700)
VDASQIASESELILHLPQR	FAQGGNALFEGR (9)
IENYYLTLNLR	VIEPEPIPGPEIILYLR (10)
SHQGELSLSGIAEYR	APVFTVPLENIENLR (12)
FLIDIEPNGLLR	DLATIGLLLR (16)
EGYPPFFR	LTVEEPLVDFVIR (24)
CG33519 or CG30171	CG1915
p1000 Projectin 1	p1000 Projectin 2
FLKPHIDR	ATIPDLVEGQTYK
IFADNVYGR	
IQGYQIEYR	
FTVPSPPGAPQVTR	
YVGDDYYFIINR	
CG1479 or CG10285	CG1479 or CG10285
p1500 Sl(1500)	
APVFTVPLENIENLR (12)	
LTVEEPLVDFVIR (24)	
CG1915	

The CG numbers for the genes identified are given below the peptides. For Sl(700) and Sl(1500), the ORF numbers that contain the peptides are given (in brackets).

Figure

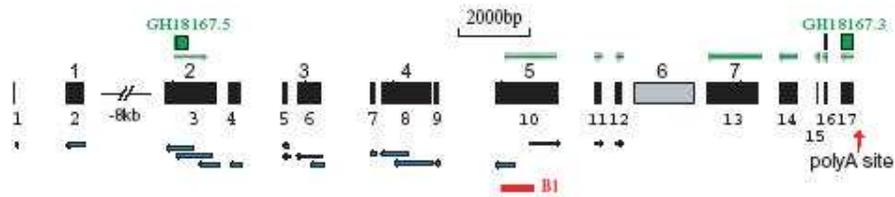


Figure 1

Figure

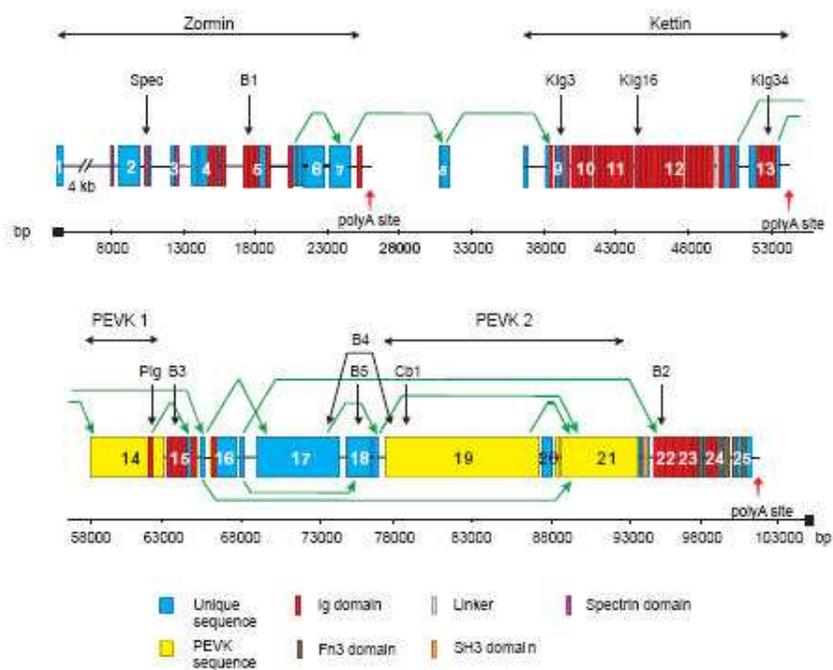


Figure 2

Figure

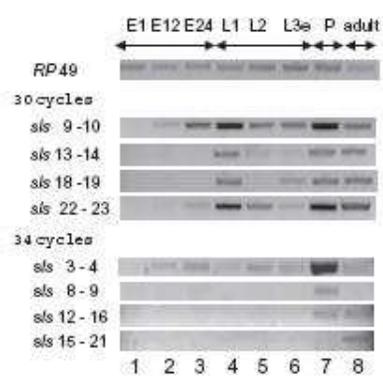


Figure 3

Figure

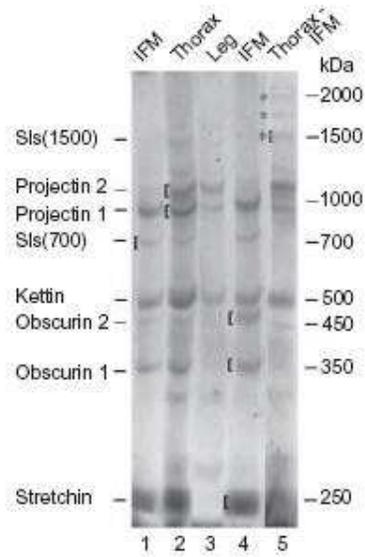


Figure 4

RIPT

Figure

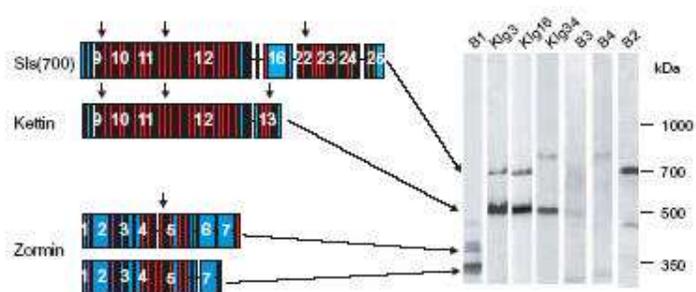


Figure 5

Figure

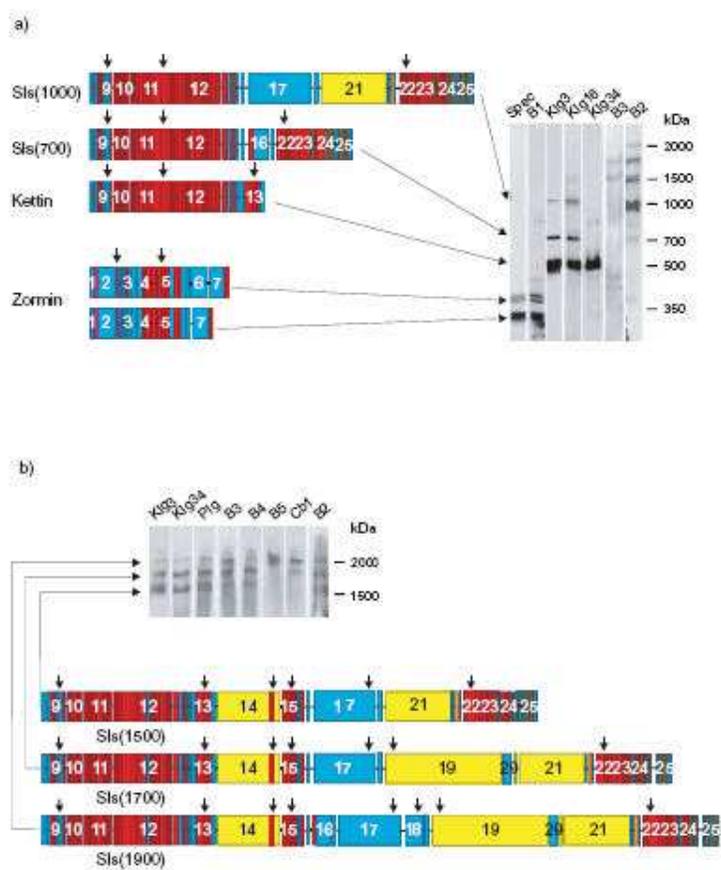


Figure 6

SCRIPT

Figure

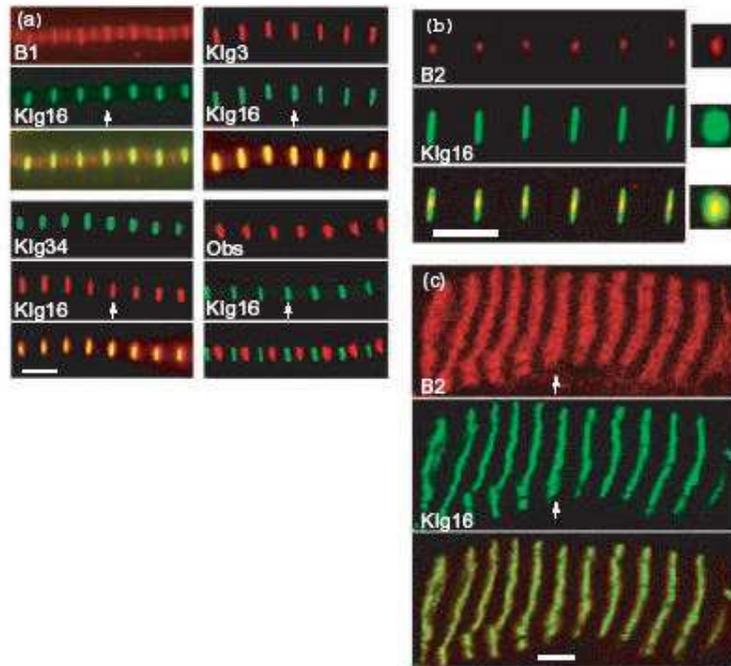


Figure 7

ISCRIP

Figure

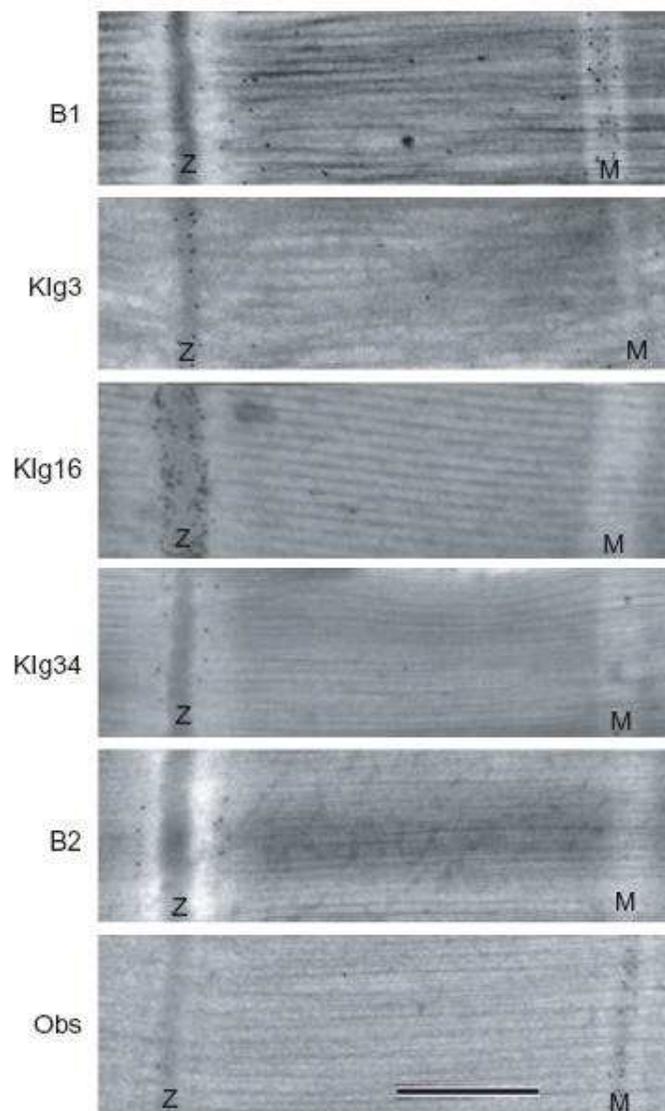


Figure 8

Figure

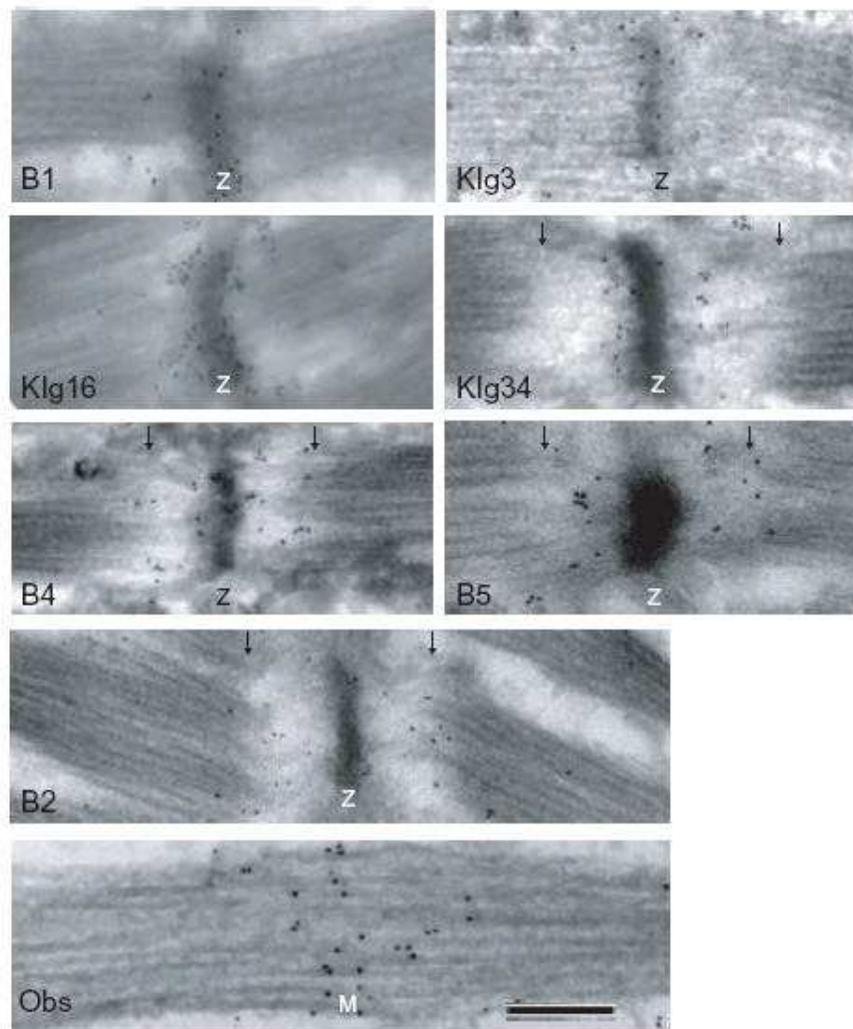


Figure 9

Figure

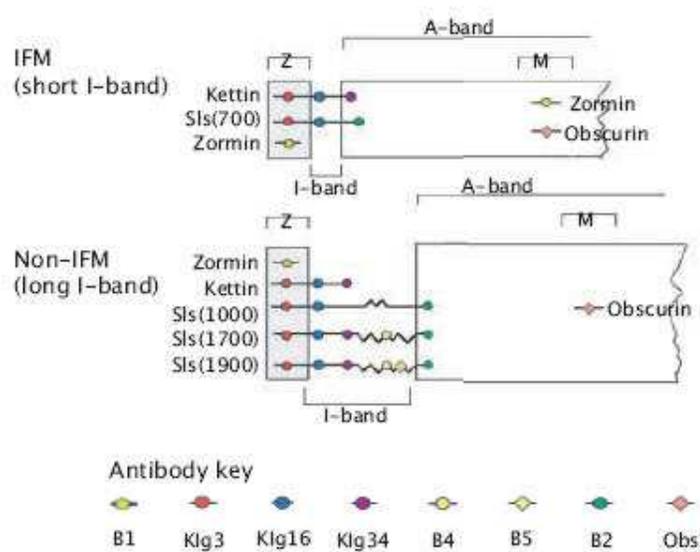


Figure 10

Figure

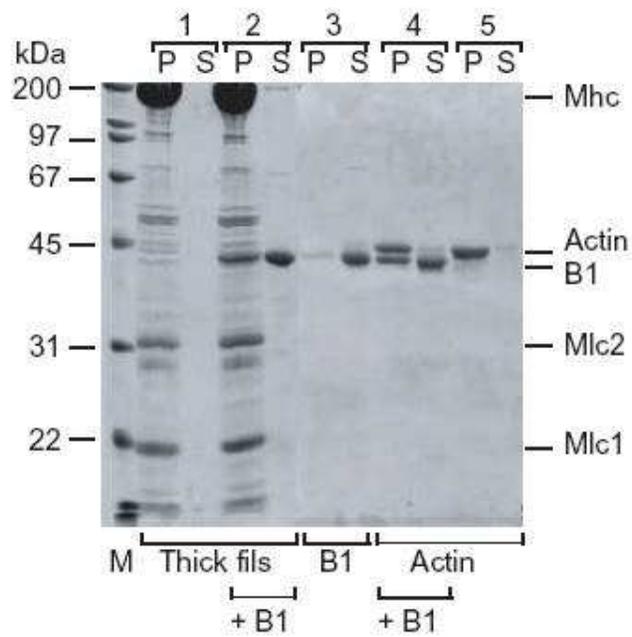


Figure 11