The structural role of high molecular weight tropomyosins in Dipteran indirect flight muscles and the effect of phosphorylation

Jesús MATEOS[1], Raúl HERRANZ[1], Alberto DOMINGO[1,*], John SPARROW[2] and Roberto MARCO[1,#]

[1]Departamento de Bioquímica, Facultad de Medicina, Universidad Autónoma de Madrid, Madrid, Spain.

[2]Department of Biology, University of York, York, YO10 5DD, United Kingdom.

[*] Present Address: Departamento de Bioquímica y Biología Molecular, Universidad de Alcalá de Henares, Spain

[#] Corresponding author: Prof. Dr. Roberto Marco, Departamento de Bioquímica, Facultad de Medicina (UAM), C/ Arzobispo Morcillo, 4 CP 28029 Madrid (SPAIN).
Email: roberto.marco@uam.es, Phone: 34 91 497 54 09 / 54 02, Fax: 34 91 585 44 01

Abbreviations: **IFM:** Indirect Flight Muscles, **ORFs:** Open Reading Frames, **Tn:** Troponin, **Tm:** Tropomyosin **MLC-2:** Myosin Light Chain-2, **GST-2:** Glutathione-S-transferase-2, **MW**: Molecular Weight.

Heavy Tropomyosin phosphorylation state on insects IFM

ABSTRACT

In *Drosophila melanogaster* two indirect flight muscle (IFM)-specific heavy troponin isoforms, proteins 33 and 34, are encoded by the *Tm1* gene. They are fusions of a conventional tropomyosin at their N-termini and proline-alanine-rich domain at their C-termini, so should be known as tropomyosin-H (TmH; heavy tropomyosin). We report that the C-terminal domains are post-translationally phosphorylated domains in mature flies, but not in recently emerged flies that are incapable of flight. From stoichiometric measurements of thin filament proteins, and cysteine bridges formed by TmH isoforms and a standard tropomyosin muscle isoform in *Drosophila*, we conclude that the N-terminal part is integral to the thin filament structural unit as tropomyosin dimers while the C-terminal, phosphorylated domain could be involved in interfilamental cross-links. Comparisons of the *Tm1* and *Tm2* gene organization show complete conservation across other Drosophilidae, including *Drosophila pseudoobscura,* while in *Anopheles gambiae* a single TmH-like exon exists though the overall gene organization is maintained. Thus, in *Diptera* isoforms such as *Drosophila* 33 and 34 are not troponins but the heaviest tropomyosins. Interestingly, in hymenopterans (*Apis*), while most *Tm1* and *Tm2* gene features are maintained, *Tm1* lacks the tropomyosin carboxyterminal extension sequences that are instead found encoded at the 3' end of the troponin I gene. In this order, as in *Lethocerus* (hemipterans), the original name of troponin H (TnH) should be retained. So, in parallel with the idea that the stretch activation mechanism, responsible of the asynchronicity of the IFM contractions and the flight in many high insects, has independently evolved several times, at least two different but related molecular configurations in the different insect orders can be identified, that likely share similar functional properties.

INTRODUCTION

In vertebrate striated muscle, thick and thin filament interactions are regulated by a thin filament protein complex **(Gordon *et al.*, 2000; Geeves and Holmes, 1999).** This troponin-tropomyosin (Tn-Tm) complex acts as a sensor of Ca^{2+} concentration in the fibers. In skeletal muscle, Ca^{2+} binding to Tn-Tm promotes conformational changes which permit the movement of the continuous tropomyosin filament across the F-actin surface, exposing the myosin-binding site of actin and allowing crossbridge cycling to begin **(Craig and Lehman, 2001)**. The troponin complex consists of a tropomyosin binding subunit, troponin T (TnT), an inhibitory subunit, troponin I (TnI), and a Ca^{2+}-binding subunit, troponin C (TnC). The Tn-Tm complex components have been well characterised in the indirect flight muscles (IFM) of *Drosophila* and the giant water-bug, *Lethocerus* **(Wendt and Leonard, 1999)**. They show sequence homology to their vertebrate counterparts, but IFM-specific isoforms of each are common. This variability is produced either by IFM-expression of one or more alternate gene copies, e.g, five in the case of TnC **(Herranz *et al.*, 2004; Herranz *et al.*, 2005; Qiu *et al.*, 2003),** or by IFM-specific splicing of gene transcripts from single genes, e.g. TnI and TnT **(Herranz et al, in review; Barbas *et al.*, 1991; Fyrberg *et al.*, 1990)**. Tropomyosin (Tm) IFM isoforms are generated from two contiguous genes, *Tm1* and *Tm2* genes that are both alternatively spliced **(Karlik and Fyrberg, 1986; Bautch *et al.*, 1982).**

Muscle contraction in insect IFM is oscillatory but asynchronous, i. e., not in phase with nervous stimulation **(Thorson and White, 1983; Pringle, 1978)**. Although Ca^{2+} is necessary for IFM contractions to be initiated and maintained, it is the periodic

alternative stretch of the two sets of opposing fibres on the thorax cuticle movement which transmits a strain from one set of fibres to the other. The fibres respond to this strain by developing tension following a delay that is species-specific. So in the presence of sufficient calcium these alternate out-of-phase contractions produce an oscillatory resonant system that can maintain wing beat frequencies in excess of 200 Hz. This mechanism is an evolutionary adaptation in higher insects **(Vigoreaux, 2001; Sparrow, 1995)** that allows wing beat frequencies well above the nervous impulse frequency limit, but probably as a consequence of this uncoupling, reduces the volume of muscle fibres that would be otherwise be occupied by the sarcoplasmic reticulum and mitochondria. Stretch-activation is an inherent property of the contractile apparatus **(Peckham *et al.*, 1990)**. The evolution of new IFM specific proteins as well as new exons and N- and C-terminal extensions in a variety of IFM-specific proteins compared to their vertebrate homologues may have been important in this regard. This point was made by Reedy et al. **(1994)** for the TmH extensions. More recently, the evidence of the importance of phosphorylation of the N-terminal extension of *Drosophila* regulatory light chain, DmMLC-2 for flight led to proposals that this protein is directly involved in establishing of actomyosin cross-bridges **(Tohtong *et al.*, 1995)**. The existence and role of other potential cross-links between thick and thin filaments involving IFM-specific variants of TnT and TnI and tissue-specific protein isoforms such as flightin **(Vigoreaux and Perry, 1994)** and troponin H **(Bullard *et al.*, 1988)** remain to be examined critically.

Historically, *Drosophila* muscle protein isoforms 33 and 34 were included as members of the insect troponin complex because they reacted to monoclonal antibodies raised against

Lethocerus TnH. Qiu et al **(2003)** have proposed that *Lethocerus* TnH is a modification of the TnI gene, to which a proline-alanine rich extension had been added, justifying in this case the name of troponin H. However, molecular analysis of the *Drosophila Tm1* gene showed that isoforms 33 and 34 are fusion proteins of a tropomyosin-homologous N-terminal part with proline-alanine rich C-terminal extensions **(Hanke and Storti, 1988; Karlik and Fyrberg, 1986)** which share epitopes within the *Lethocerus* proline-alanine rich sequence. Apart from this, the *Drosophila* and *Lethocerus* TnH are not homologous proteins. The *Drosophila melanogaster* IFM-specific isoforms 33 and 34 **(Mogami *et al.*, 1982)**, show exclusive C-terminal extensions encoded by exons 16 and 15 of the *Tm1* gene, respectively. The *Tm2* gene **(Karlik and Fyrberg, 1986)**, originally named *TMI* **(Hanke and Storti, 1988)** encodes standard muscle tropomyosins, isoforms 127 and 129. By alternative splicing the *Tm1* gene also produces a standard muscle tropomyosin isoform, mTm1 (tropomyosin 128), found in several muscles, and a cytoskeletal tropomyosin (cTm1) expressed in non-muscle cells. The two genes located in tandem in chromosome region 88F, show a high degree of homology in those exons encoding the common protein sequence, the N-terminal part, and in similar regulatory elements within intron 1 **(Gremke *et al.*, 1993; Meredith and Storti, 1993)**.

Tropomyosins normally form a continuous dimeric coiled-coiled coil filament that extends along the thin filament. The location and role of isoforms 33 and 34 in the IFM sarcomeres are unknown. Both may be able to substitute for or dimerize with the standard *Tm2* product (isoform 127) as has been suggested to explain the effects of some *Tm2* gene mutations**(Kreuz *et al.*, 1996)**. The major proposal that the additional carboxy-

terminal 200 amino acids may form, or be part of, an extended link between the thick and thin filaments (**Reedy *et al.*, 1994**), lacks experimental evidence**.**

In this paper we show that both TmH 33 and 34 isoforms are phosphorylated in their IFM-specific sequences in fully mature flies. Stoichiometric measurements of the different sarcomeric components support calculations which show that the number of different Tm isoforms, including both 33 and 34, are necessary to generate sufficient Tm dimers as filaments along F-actin core of the thin filaments. Morever, TmH 33 and 34 form disulfide complexes *in vivo* between themselves and with the standard Tm isoform, consistent with the presence of cysteine residues in the N-terminal part of all these isoforms. Therefore, we propose that isoforms 33 and 34 are integrated into the thin filament by their N-terminal parts, forming homo- and heterodimers between themselves and with the standard Tm 127. We postulate that in both cases, the C-terminal parts act as extensions projecting towards the thick filament. Such lateral extensions could act either to maintain the thick/thin filament geometry out of phase before stretch or act directly as stretch sensors feeding directly into thick or thin filament regulatory mechanisms. While this arrangement is found in dipterans, in other holometabolous insects such as hymenopterans (*Apis*) and hemipterans (*Lethocerus*) a similar result is attained by a combination of similar motifs in TnI in accordance with the idea that stretch-activation may have evolved independently in different high insect orders (**Dudley, 2000; Pringle, 1981)**

MATERIAL AND METHODS

Fly stocks

Oregon R strains were used as wild type *Drosophila* for all experiments.

2D electrophoresis

Single female *Drosophila* thoraces were dissected in cold acetone and homogenised in 8M urea, 0.5 % IGEPAL CA-630, 10 mM dithiothreitol, and 0.1% ampholytes (pH range 3-10). Solubilization was performed for 3h by agitation at room temperature. Samples were centrifuged at 13000 rpm in a tabletop centrifuge and the supernatants subjected to 2-D electrophoresis using an IPGphor system (Amersham Pharmacia Biotech). Second dimension separations were performed on SDS-PAGE gels with 10% acrylamide for experiments of *in vivo* phosphorylation, staining, and mass spectrometry.

***In vivo* phosphorylation of isoforms 33 and 34**

Flies were fed for 20h with 1mCi of ^{32}P added in the food. Thoraces were dissected and processed as described above. The 2-D gels were developed using X-ray films.

For phosphoaminoacid analysis spots 33 and 34 were cut from a dry 2-D gel and processed as previously described (**Bonato *et al.*, 1984**). The amino acid mixture was separated by thin layer chromatography and radioactivity was detected using X-ray film exposure.

Stains-all staining

After the second dimension, 2-D gels were fixed with 50% 2-propanol with microwave heating and staining in the dark with 30% 2-propanol (Merck), 0.3 M TrisHCl pH 8.8, 10 % formamide (Riedel) and 0.1% Stains–all (Serva).

Mass spectrometry experiments

Spots 33 and 34 were cut out from silver stained 2-D gels of single thoraces and trypsin digested *in gel*. The resulting peptide mixtures were extracted several times with sonication in 50% acetonitrile (Sigma) and 0.5% trifluoroacetic acid (Pierce). Samples were concentrated in a Speedvac centrifuge prior to MALDI-TOF and nanospray analyses.

MALDI-TOF of proteins extracted from 2-D gels was performed alternatively using either a Kompact Probe (Kratos-Shimadzu) or a BrukerReflexIII (Bruker-Franzen Analytic GmbH) mass spectrometers, as previously described **(Paradela *et al.*, 2000),** working in positive ion mode. Samples were subjected to microdesalting in a 300 μm of diameter nanocolumn (Vydac C-18) in a Smart-HPLC (Amersham Pharmacia Biotech). The peptides with theoretical phosphorylations were isolated and sequenced by quadrupole nanoelectrospray-ionic trap in a LCQ instrument (Finnigan ThermoQuest) as described previously **(Paradela *et al.*, 2000).** The peptide maps obtained were compared with the Swiss-Prot database.

Stoichiometry analyses

Thoraces and IFMs from recently emerged flies were dissected and prepared for SDS gel electrophoresis. For ^{35}S-labeling, newly eclosed flies (with still visible meconium) were etherised and injected with 1 μl of a high specific activity (1 Ci/ liter) ^{35}S-methionine aqueous solution (Amersham) to which 1/10 of 10 fold concentrated *Drosophila* Ringer solution had been added. The gels were autoradiographed and also stained with Coomassie blue. The autoradiographs and the blue-stained gel bands were quantitized by densitometry and peaks were analysed taking into account their surface/height/width and the quantitities of methionines and/or lysines/arginines in the different thin filament components. Information on the relative masses of IFMs and non-IFMs thorax muscles has been obtained from the quantitation of Figures 26 and 27 of the Miller´s Chapter assuming each muscle can be approximated by a cylinder or a truncated cone **(Demerec, 1950)**.

Protein Immunodetection in reducing/non-reducing monodimensional gels and in diagonal gels

In both cases immunodetection was performed using two rat monoclonal antibodies developed by Dr. Bullard group **(Clayton *et al.*, 1998)** and kindly provided by her. MAC144 recognizes one of the seven residue repeats in the IFM-specific part of isoforms 33 and 34 and, by sequence homology, a similar sequence in GST-2. MAC141 recognizes the common sequence shared with standard Tm.

Dissected thoraces were homogenised in York buffer **(Clayton *et al.*, 1998)** with or without dithiothreitol (Sigma). After centrifugation, pellets were solubilized in Laemmli sample buffer with or without β-mercaptoethanol and heated at 100ºC for 5 min. Samples were run on 10% SDS-PAGE gels and transferred onto nitrocellulose in a semi-dry blotting apparatus (Pharmacia). Membranes were incubated with MAC 141 or MAC 144 antibodies, then with goat-anti-rat serum and developed using a chemiluminescent substrate for the ECL system (Amersham).

Diagonal gels were performed by running a non-reducing dimension in a tube gel and a reducing second dimension in a slab gel as described in **(Creighton and Chothia, 1989)**. Detection of cross-linked complexes using MAC144 and MAC141 antibodies was carried out as indicated above.

Protein and genome sequence analysis and comparison

For the protein sequence analysis the ClustalW program from ExPASy was used **(Combet *et al.*, 2000)**. Comparisons of the sequence of all the Tm sequences were performed using either ClustalW or GeneJockey II programs. Secondary structure predictions were made with PSI-pred program from ExPASy **(McGuffin *et al.*, 2000)**. *Drosophila melanogaster* and *Anopheles gambiae* sequences were obtained from the NCBI BLAST server **(www.ncbi.nlm.nih.org)**. *Drosophila pseudoobscura* and *Apis* sequences were obtained from the Baylor College of Medicine BLAST server **(www.hgsc.bcm.tmc.edu)**.

RESULTS

Properties of muscle isoforms 33 and 34

Figure 1 shows the sequence comparison of both isoforms 33 and 34 and a prediction of the secondary structure of the consensus sequence. The Tm-homologous N-terminal part, encoded by the common exons, will have a predicted α-helical structure identical to the Tm one. In this part, Cys residues appear within the first 27 amino acids involved in Tm dimerization (indicated with an arrowhead). These occur in all Tm isoforms encoded by the *Tm1* and *Tm2* genes **(Holtzer, 1986).** By comparison, the C-terminal part is an Ala/Pro rich domain with an unstructured random coiled extension containing several Thr residues (shown with asterisks).

The 2-D gels of *Drosophila* thoraces show that both isoforms 33 and 34 are differentially stained in blue with *stains-all* (Figure 2A), to be compared with other proteins also staining blue such as TnT or MLC. Calcium overlay experiments do not show any signal from isoforms 33 and 34 (data not shown). Phosphorylated proteins or Ca^{2+}-binding proteins are known to stain blue with this cationic dye **(Domingo *et al.*, 1998)**.

Both isoforms 33 and 34 are *in vivo* phosphorylated

The autoradiograph of a 2-D gel of thoraces from flies fed with ^{32}P shows a relatively strong signal from both isoforms TmH 33 and 34 (Figure 2B). This feature is shared with other proteins classically involved in the mechanism of contraction such as the myosin light chains **(Domingo *et al.*, 1998)**. The phospho-aminoacid analysis of the protein spots by thin layer chromatography of the residues and comparison with non-radioactive markers showed that the phosphorylation occurs on threonine residues (data not shown).

Isoforms 33 and 34 are phosphorylated in fully mature flies

Mass spectrometry techniques were used to identify the exact amino acid that was phosphorylated. Trypsin digestions of spots 33 or 34 extracted from silver stained 2D gels (Figure 2C) were used for MALDI-TOF analysis. The peptides maps obtained (Figure 3) were compared with the theoretical map derived from databases, confirming that the proteins were isoforms 33 and 34. In addition, to establish the physiological significance of the post-translational modification, we have compared two different functional stages of IFM, working with two types of fibers, two day old, fully mature flies, and recently emerged flies, minutes after eclosion from the pupal case and incapable of flying. Mass spectrometric data of TmH 33 from fully mature flies showed two new peaks not present in the reference set of peptides, with molecular weights consistent with two contiguous peptides, Asn293-Lys300 and Leu301-Arg316, if a single phosphate moiety was included in each. These two peptides were detected in recently-emerged flies as non-phosphorylated peptides. Moreover, TmH 34 from fully mature flies showed a discordant mass peak with a molecular weight consistent with a single peptide, Tyr292-Lys305, also including a single phosphorylation. The last peptide is similar and occupies an analogous place in the sequence to the Asn293-Lys300 found modified in TmH 33 (See alignment in Figure 1). So the two analogous peptides, Tyr292-Lys305 from TmH 34 and Asn293-Lys300 from TmH 33, are both phosphorylated after emergence of flies. Note that all three phosphorylated peptides are located in the IFM-specific part of each protein. Several Thr residues exist in this IFM-specific region of both isoforms (Figure 1), but only specific residues seem to be capable of being phosphorylated.

Phosphorylations detected by the molecular weight change, were corroborated empirically. Prior to analysis, samples of the fractions corresponding to the discordant peptides of fully mature flies were concentrated and desalted by HPLC and then subjected to nanoelectrospray ionic trapping. Sequencing of the isolated peptides confirmed that the phosphorylations were located at specific Thr residues. The fragmentation spectra of each peptide are shown in Figure 3.

Stoichiometric proportion of the sarcomere components in *Drosophila* IFM.

By volume the indirect flight musculature of dipterans is the major component of the thorax. It fills almost completely its interior. It exceeds 75% of the total muscle volume, even without taking into account the solidity of the fibrillar muscles vs the more sparse organization of the non-fibrillar ones. The hypodermal musculature is insignificant by comparison and in sections the major tubular muscle in the thorax, the tergal depressor of the trochanter (TDT) covers only at most 15% of the volume. As a first step in analyzing the stoichiometry of the sarcomere components in *Drosophila* flight muscles, we compared the protein bands of dissected IFMs and the thoraxes and found them indistinguishable. Therefore, it was decided to continue the stoichiometry analysis with dissected thoraxes. Using different approaches, ^{35}S-methionine labelling of the sarcomeric components normalized to their actual content in this amino acid or the extent of Coomassie blue staining related to the actual proportion of basic amino acids (lysine and arginine) in these components, the stoichiometry of the main different structural components of *Drosophila* IFM thin filaments have been established. The data (Table 1) show some variation depending on the experimental approach and method of analysis

used, but overall indicate common stoichiometric ratios of the components. If the numbers are normalized to fourteen actin molecules, the number of actin monomers in the repeating structure of the thin filament **(Filatov *et al.*, 1999)**, with some variation, the mean value gives two molecules of arthrin (ubiquitinated actin) per twelve molecules of actin. Note that 14 actins, as the basic repeating unit of the thin filament should accommodate 4 molecules of Tm, one dimer in each groove, each associated with two troponin complexes. This fits with a TnT number close to this value, but Tm does not reach a value of 4, unless the TmH 33 and 34 molecules are included. There are variations in the values for the components, the more important ones relating to the levels of arthrin that can range from less than two to three per 14 actins and the number of TmH 33 and 34 molecules. Although they approach the stoichiometric levels given above, the values vary slightly from experiment to experiment. Usually, but not always, TmH 33 is more prominent than protein 34. The model suggested by these measurements is also supported by the following analysis.

Isoforms 33 and 34 form protein complexes with standard Tropomyosins *in vivo*

An unresolved question is how TmH 33 and 34 are integrated into the thin filament structure. Sequence homology of the N-terminal part of all the *Tm1* and *Tm2* gene products argues that all the IFM isoforms will form homodimers or heterodimers. The stoichiometry results described in the previous section support both these possibilities, but make more likely the suggestion that TmH 33 and 34 could form heterodimers with protein 127, the standard Tm isoform expressed in IFM. Cysteine residues appear in the conserved N-terminal part of these *Drosophila* Tms. It is possible that, as has been

described in skeletal muscle (**Holtzer, 1986**), Tm coiled-coils are stabilized by intermolecular disulfide bridges, as has been established for paramyosin (**Vinos *et al.*, 1991**). To determine whether this is true and also whether TmH 33 and 34 dimerize with the standard Tm (protein 127) two monoclonal antibodies were employed to detect the proteins in gels run in reducing and non-reducing conditions. The monoclonal antibodies MAC141 and MAC144 recognize the common Tm part and the IFM-specific part, respectively. In both cases a higher apparent MW complex could be detected in non-reducing gels (Figure 4A) but not seen in the reducing gels, indicating the presence of a disulfide bond held protein complex.

Diagonal gel electrophoresis is a very useful technique to study this type of interaction. A non-reducing "first dimension" is performed in which disulfide bridge containing complexes appear delayed. The entire gel is then subjected to a reducing slab "second dimension". Proteins previously delayed, finally appear off the theoretical diagonal, given by a straight line from the origin of the gel and the signal from the monomeric low MW isoform (GST-2 and tropomyosin respectively). Immunodetection of two parallel diagonal gels using both monoclonal antibodies are represented in Figure 4B. MAC144 reveals a large spot of a small molecular weight protein on the diagonal which corresponds to GST-2 and a minor GST-2 fraction corresponding to the previously described interaction between TmH 33 and/or 34 and GST-2 (**Clayton *et al.*, 1998**) that is apparently only released under reducing conditions. The MAC144 antibody further identifies dimers of 33 and/or 34. The MAC141 antibody clearly shows that a large fraction of tropomyosin 127 lies on the diagonal and must therefore derive from dimers in

which cysteine bridges did not exist *in vivo* or were dissociated in non-reducing conditions. This makes it impossible to completely track the whole dimeric Tm complement. The MAC141 antibody clearly detected heterodimers of TmH 33 or 34 with Tm 127. The intensity of the spots detected off the diagonal with the antibodies indicates that, ignoring the monomeric Tm 127, at least 25% of Tm (low and high molecular weight) is homodimers of Tm 127 forms, 50% is present in heterodimers of Tm127 with the TmH 33 and 34 IFM-specific isoforms, and 25% is present as a significative fraction of Tm forms complexes *in vivo* as either homodimers of Tm 127 or heterodimers of Tm 127 and TmH 33 and 34 or homodimers of these large isoforms. The interactions determined in these experiments confirm that a significant fraction of TmH 33 and 34 is integrated into the thin filament as high molecular weight tropomyosins, as was already implied from the high homology of their N-terminal parts. Secondary structure predictions of a high scoring α-helical coiled coil domain for the N-terminal portion of TmH 33 and 34 (Figure 1) similar to that of the standard Tm (protein 127) agree with this model. The data from the cysteine residues confirm that these proteins dimerise through these domains. A model in which each Tm-Tn complex contains a heterodimeric tropomyosin (Tm 127 + TmH isoform 33 or 34) is made untenable by our detection of Tm 127 homodimers. Together with the existence of Tm 127 + TmH 33/34 heterodimers and heavy Tm dimers, this observation means that Tm-Tn complexes along the thin filament will be non-identical. TmH 33 contains a cysteine within the C-terminal part of the Tm domain which could conceivably generate cysteine bridges in the Tm-overlap region and thereby some of the heterodimers, but in the absence of data on the structure of the overlap region, we consider this an unlikely explanation.

The *Drosophila melanogaster* Tm genes organization is conserved in *Drosophila pseudooboscura* and *Anopheles gambiae* but not *in Apis mellifera*

If TmH 33 and 34 play essential roles in the activation of IFM contraction and flight ability of *Drosophila*, then the *Tm1* gene and its IFM-specific exons are likely to be conserved and functional in dipterans. This idea is sustained by previous data indicating that high molecular weight proteins similar to TmH 33 and 34 are present in the flight muscles of many insects **(Reedy *et al.*, 1994; Peckham *et al.*, 1990)**, but only in the Diptera are they also detected by an antibody recognizing standard Tm sequences. To confirm that isoforms of TmH 33 and 34 are widely present in other Diptera, we screened the recently sequenced genomes of *Anopheles gambiae* and *D. pseudoobscura*. A BLAST search was carried out using the *Drosophila Tm1* and *Tm2* gene exons and the entire *Tm1/Tm2* genomic muscle complement was identified in these two dipterans (Figure 5A). The homology of the exons encoding the tropomyosin sequences is very high extending into the IFM-specific TmH exons 15 and 16. The overall *Tm1-Tm2* gene structure and sequences are similar but slight differences are detected: a) the sequence containing the two genes is larger in *D. pseudoobscura* (37 kB) and in *A. gambiae* (49 kB) than in *D. melanogaster* (33.6 kB*)*, especially in *A. gambiae* where the *Tm1* gene homologue covers more than 38kB. The larger size of the genes is due largely to the longer introns, as well as to the intergenic space between the two *Anopheles* Tm genes, b) in *A. gambiae* there appears to be only one IFM-specific exon, more closely related to *D. melanogaster* exon 16 than to exon 15, c) in *Anopheles* the *Tm1* exons 2 (including the initiating codon) and 3 appear fused. The cytoplasmic *D. melanogaster* isoform (cTm1) and the muscle

specific isoform (mTm1) organization is conserved in Diptera. The *Tm2* organization is also conserved closely linked following the *Tm1* gene in dipteran chromosomes. In *Anopheles*, the first large *Tm2* translated exon in the *Drosophilidae* is divided into two exons.

In other insects orders **(Peckham *et al.*, 1990)**, high molecular weight protein containing the anti-TnH epitope is not linked to the standard Tm sequence. When similar BLAST searches were carried out on the sequenced genome of the hymenopteran *Apis mellifera*, two genes were also found, one homologous to *Tm1* and a second homologous to *Tm2*. One of these shows specific exons homologous to the dipteran cytoplasmic tropomyosin as well as muscle-specific *Tm1* isoforms and common exons as can be seen in Figure 5A. The *Tm2* orthologous gene shows a further division of the first translated exon 2 in *Apis* into three exons (2a, b and c). The availability of the *Apis mellifera* genome has allowed us to confirm the absence of the TmH specific exons in the *AmTm1* and the *AmTm2* genes. We have reported the presence of IFM-specific additional exons in the *AmTnI* gene **(Herranz et al, in review)**, that encode a carboxy-terminal region containing a repetitive APPAEGA motif identical to the one found in the dipteran TmH. The presence of a similar Ala-Pro-rich extension fused to TnI has been also cited in Qiu et al **(2003)** for the hemipteran giant water bug, *Lethocerus*. Interestingly, the *Apis Tm1* orthologue shows two final muscle-specific exons, the first one (15) highly homologous to the initial amino acids of the dipteran proline-alanine rich exon 16, pointing to this sequence as the target for the evolutionary insertion of the extra proline-rich specific sequences in the dipteran flight muscles. The final exons of the bee TnI encode sequences homologous to

the C-terminal extensions of both TmH 33 and 34 (Figure 5B). The phosphorylated residues in *D. melanogaster* are conserved in the other two species (indicated by asterisks), as well as the patterns of predicted protein secondary structure (data not shown). As already mentioned, in *Anopheles* only one exon (14 as a consequence of the fusion of exons 3 and 4), more similar in sequence to the drosophilid exon 16, is found in between the homologous final cytoplasmic exon and the final standard muscle exon. Part of the *Apis* troponin I gene sequence is also included in Figure 5B. Although somewhat different from the dipteran tropomyosin H extensions, its similarity to them can be easily seen, including the conservation of the threonine-rich region that we have shown here to be phosphorylated.

DISCUSSION

Drosophila Troponin H or Tropomyosin H?

A monoclonal antibody raised against a component, named heavy troponin, of the flight muscle troponin complex from the giant water-bug, *Lethocerus indicus* **(Reedy *et al.*, 1994)** cross-reacted with two proteins of similar size from *Drosophila* IFM myofibrils, designated as IFM proteins 33 and 34. Although, as described above, these are *Tm1* gene products, from the cross-reaction they were referred to as TnH 33 and TnH 34. Now that the *Apis* TnH and *Lethocerus* TnH have been identified as TnI isoforms **(Herranz et al, in review; Qiu *et al.*, 2003)** while *Drosophila* TnH 33 and TnH 34 show clear sequence identity as tropomyosins, we propose that these ‘heavy tropomyosin’ *Tm1* gene products

in Dipterans be referred to as TmH 33 and 34 in *D. melanogaster*, while in other orders it can retain the original designation of TnH.

Physiological roles of phosphorylation of TmH and other IFM proteins

We have described the *in vivo* phosphorylation of the two IFM-specific products of *Tm1* gene in *Drosophila*, muscle isoforms 33 and 34. The modified Thr residues are located, in both cases, in a random-coil part of the IFM-specific sequence. Phosphorylation is readily detectable in two day old mature flies but not in recently emerged ones incapable of flight. Phosphorylation of other *Drosophila* muscle proteins such as the myosin regulatory light chain (**Tohtong *et al.*, 1995**), flightin **(Vigoreaux, 2001; Vigoreaux and Perry, 1994),** troponin T **(Nongthomba *et al.*, In review; Domingo *et al.*, 1998)**, miniparamyosin **(Maroto *et al.*, 1996)** and projectin **(Ayme-Southgate *et al.*, 1995; Maroto *et al.*, 1992)** has previously been described, and may act in a coordinated way to cause conformational changes of the IFM proteins. Mutants in which the *Drosophila* regulatory light chain (DmMLC-2) is truncated or in which the phosphorylatable serines are mutationally substituted with alanines **(Tohtong *et al.*, 1995)** reduced the power output and flight ability, although myofibrillogenesis and the calcium-activated force developed appeared normal. The effects are probably due to an interaction between regulatory light chain and actin, similar to the essential light chain–actin interaction in vertebrates **(Henkin *et al.*, 2004; Trayer *et al.*, 1987)**. It is interesting to point out that as just described, a group of these changes involve phosphorylation and/or are specific of the IFM. Together with the expression of an almost IFM specific actin isoform and arthrin, of TnI and TnC IFM specific isoforms as well as a specific myosin heavy chain

isoform **(Hastings and Emerson, 1991)**, all these changes may contribute to the especial physiology of these muscles. White **(1983)** proposed that stretch-activation might be a property of strain sensed by attached myosin cross-bridges. With the discovery that in *Lethocerus* and *Drosophila* the IFM homologues of vertebrate striated muscle proteins had unusual additional domains, novel proteins were present and that many of these components were post-translationally modified, molecular explanations for stretch-activation have focused on these proteins and the potential crosslinks between thick and thin filaments.

IFM thin filament stoichiometry and structure

Protein stoichiometry is important in defining the structure of a macromolecular complex. Our data confirm that within the acknowledged 14-monomer actin repeat of the muscle thin filament there are two molecules of each troponin component (TnI, TnC and TnT). What is more revealing is that by homology to vertebrate thin filament structures there should be two Tm dimers associated with this length of thin filament. The stoichiometry shows only two molecules of standard length Tm per repeat. However, there are two molecules of the TmH 33 and 34, each with a standard Tm sequence as their N-terminal halves. This suggests that the IFM thin filament repeat structure could contain – 14 actins (including two arthrins), with two heterodimeric (Tm + TmH) Tms and two troponin complexes. The analysis of cysteine crosslinks between Tm and TmH isoforms suggest that the repeats more likely differ as homodimers of Tm and TmH isoforms occur as well as Tm/TmH heterodimers as was argued by Kreuz et al **(1996)**. So non-identical thin filament repeats in terms of the Tm strand content are probable.

Current evidence in support of proposals that TmH and other IFM-specific proteins are involved in inter-filament cross-links, independent of the cross-bridges, is scanty. No *Drosophila Tm1* gene mutant affecting flight muscle function has been recovered so far. In genetic studies using a deletion which removes both *Tm1* and *Tm2* genes and nonsense mutations within the *Tm2* gene, Kreuz *et al* **(1996)** concluded that while flies heterozygous for the deletion affected important IFM functions (from wing beat frequency, myofibrillar organization, and skinned fibre mechanics), by comparison to the *Tm1* mutant heterozygotes most of these effects were explicable by a 50% gene dose reduction of *Tm2* rather than a similar *Tm1* reduction. This suggests that a single *Tm1* gene copy is sufficient to cover the role(s) of TmH 33 and TmH 34 in asynchronous muscle contraction. In addition to the TmH-Tm interactions demonstrated in this report, it has been earlier reported that TmH also binds glutathione-S-transferase 2, GST-2, through the C-terminal domain **(Clayton *et al.*, 1998)** as also detected in our diagonal gels (Figure 4B left). This interaction may be structural, but equally the enzymatic activity of GST-2 towards lipid substrates **(Singh *et al.*, 2001)** may provide protection against the deleterious effects of oxidative stress in IFM.

From this information we propose a model in which the N-terminal parts of TmH 33 and TmH 34 are integrated into the thin filament as true tropomyosins, while the phosphorylated C-terminal regions extend towards the thick filaments. While there is no direct structural evidence for this, immuno-gold labelling data **(Reedy *et al.*, 1994)**

showed a close association of the *Lethocerus* TnH epitope, shared with TmH 33 and TmH 34, and the myosin crossbridges.

Comparative evolution of the Tm1/Tm2 gene complements in Holometabolous Insects

Comparison of the Tm sequences within the *Tm1* and *Tm2* genes shows that they likely arose by duplication from a single gene, although the situation with the *Tm1* gene is more complex as different gene products, a non-muscle cytoplasmic and several muscle specific isoforms, are coded by both common and alternatively spliced exons. The *Tm1* and *Tm2* genes represent the complete Tm gene content of the *Drosophila* genome.

By sequence homology, we have identified the entire Tm gene complement in two additional dipterans, *D. pseudoobscura* and *A. gambiae* as well as in the hymenopteran *Apis*, whose genomes have been completely or almost completely sequenced. The high similarity of the homologous exons and the analysis of the splicing signals suggest that the *Tm1* gene is conserved and functional in these organisms. In dipterans, they encode isoforms similar in sequence and structure to TmH 33 and 34 (Figure 5). In *D. pseudoobscura* homology is very high, with the only differences being the slightly larger introns. The increased size of the *Tm1* gene in *Anopheles,* with longer introns and intergenic separation, may also reflect the larger genome of this species compared to *D. melanogaster*. The presence of a single IFM-specific exon, more closely related to the *D. melanogaster Tm1* exon 16 rather than to exon 15 in *Anopheles* and the homology in the sequences flanking exons 15 and 16 in *Drosophila*, support the proposal that these exons arose by duplication of an ancestral exon equivalent to the one currently found in

Anopheles. The functional implications of this putative duplication and the insertion of an intron between exons 2 and 3 in the drosophilid evolutionary line are unknown.

The situation is different in *Apis*. The *Tm1* and *Tm2* genes are further apart (30 kB is the minimum value of separation based on their distance to the end of the unlinked sequence contigs where they are found). The *Tm2* gene appears subdivided into more exons in *Apis* while the *Tm1* gene also includes exons encoding the cytoplasmic isoform (not shown). Interestingly, the gene lacks the extra polyproline alanine sequences typical of TmH 33 and 34 in dipterans, but includes two final small exons, one of them highly homologous to the initial residues in the *Drosophilidae* exon 16 and *Anopheles* exon 14. The extra polyproline alanine residues in *Apis* are found in three exons (H1, H2 and H3) at the end of the TnI gene (Figure 5B and **(Herranz et al, in review)**) in agreement with earlier data **(Peckham *et al.*, 1990)** and the information obtained in the hemipteran *Lethocerus* (Genbank entry AJ621044). The phosphorylated Thr residues detected in *Drosophila* are conserved in the *Anopheles* TmH protein and possibly even in *Apis* TnI sequence (Figure 5B).

The data suggest an evolutionary pathway for the acquisition of these polyproline sequences in holometabolous insects. The primitive organization of the tropomyosin genes may have been similar to that found in *Apis*. In dipterans the extra sequence was initially inserted coupled to the penultimate muscle exon producing a configuration similar to that now found in *Anopheles*. Finally in the *Drosophilidae*, this modified exon was duplicated producing the two TmH 33 and TmH 34 isoforms.

Finally, it has to be pointed out that the distribution of asynchronous flight muscles in different insect taxa suggests that the stretch-activation mechanism may have evolved independently several times **(Pringle, 1981)**. The appearance of multiple IFM-specific components and isoforms **(Herranz *et al.*, 2005; Vigoreaux, 2001)** is an obvious consequence of this, though the structure-function relationship between these proteins and stretch-activation in asynchronous muscles remains elusive. Different evolutionary pathways may have been utilised to achieve similar physiological effects. An example of this is the conservation of the APPAEGA epitope in the Ala-Pro rich domain present in *Apis* and *Lethocerus* TnH and dipteran heavy tropomyosins, as well as in many different additional insects. This conservation argues that these domains are important for the IFM physiology but that there is flexibility in whether this domain is associated with TnI (*Apis*, *Lethocerus*) or Tm (Diptera). The different solutions may be related to the power requirements for flight in each species. If this power is directly related to the number of inter-filamental cross-bridges **(Wray *et al.*, 1988)**, and if the IFM-specific proline-rich extension is implied in the establishment of these cross-bridges, then a higher number of proline rich extensions could drive to a higher power output. This hypothesis agree with the fact that in dipterans, (*Drosophila* and *Anopheles*), the extension is part of a Tm (up to four molecules per 14-monomer actin) while in Hymenoptera (*Apis*) and Hemiptera (*Lethocerus*) the extension appears to be part of a TnI (up to two molecules per 14-monomer actins). The multiplicity and the co-evolution of the changes in the additional specific proteins in insect flight muscles suggest that as a consequence, the effects of the evolutionary or experimental elimination of any one of them may be of little functional

significance and only become evident under extreme performance conditions or when additional specific components of the thin/thick filament are also removed. The importance of completing the description of the molecular properties of the IFM sarcomere components and proposing realistic structural models is clear if wish to understand the subtleties involved in the structure-function relationship in these highly performance muscles.

Acknowledgements. We are grateful to Belinda Bullard for supplying the two monoclonal antibodies, to Dr. Jesús Vázquez and Anabel Marina for help with the analysis of the MALDI-TOF results and to Aida Villa for technical help. This research was supported by grants from the Spanish Government PB96-0069, ESP1999-0379-C02 and PNE-008/2001-C. JCS was funded by the BBSRC (United Kingdom).

REFERENCES

Ayme-Southgate, A., Southgate, R., Saide, J., Benian, G.M. & Pardue, M.L. (1995) Both synchronous and asynchronous muscle isoforms of projectin (the Drosophila bent locus product) contain functional kinase domains. *J Cell Biol*. **128**: 393-403.

Barbas, J.A., Galceran, J., Krah-Jentgens, I., de la Pompa, J.L., Canal, I., Pongs, O. & Ferrus, A. (1991) Troponin I is encoded in the haplolethal region of the Shaker gene complex of Drosophila. *Genes Dev*. **5**: 132-140.

Bautch, V.L., Storti, R.V., Mischke, D. & Pardue, M.L. (1982) Organization and expression of Drosophila tropomyosin genes. *J Mol Biol*. **162**: 231-250.

Bonato, M.C., da Silva, A.M., Maia, J.C. & Juliani, M.H. (1984) Phosphorylation of ribosomal protein S6 in the aquatic fungus Blastocladiella emersonii. *Eur J Biochem*. **144**: 597-606.

Bullard, B., Leonard, K., Larkins, A., Butcher, G., Karlik, C. & Fyrberg, E. (1988) Troponin of asynchronous flight muscle. *J Mol Biol*. **204**: 621-637.

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**: 117-127.

Combet, C., Blanchet, C., Geourjon, C. & Deleage, G. (2000) NPS@: network protein sequence analysis. *Trends Biochem Sci*. **25**: 147-150.

Craig, R. & Lehman, W. (2001) Crossbridge and tropomyosin positions observed in native, interacting thick and thin filaments. *J Mol Biol*. **311**: 1027-1036.

Creighton, T.E. & Chothia, C. (1989) Protein structure. Selecting buried residues. *Nature*. **339**: 14-15.

Demerec, M. (1950) *Biology of Drosophila*. New York: John Wiley & Sons, Inc.

Domingo, A., Gonzalez-Jurado, J., Maroto, M., Diaz, C., Vinos, J., Carrasco, C.*, et al* (1998) Troponin-T is a calcium-binding protein in insect muscle: in vivo phosphorylation, muscle-specific isoforms and developmental profile in Drosophila melanogaster. *J Muscle Res Cell Motil*. **19**: 393-403.

Dudley, R. (2000) The evolutionary physiology of animal flight: paleobiological and present perspectives. *Annu Rev Physiol*. **62**: 135-155.

Filatov, V.L., Katrukha, A.G., Bulargina, T.V. & Gusev, N.B. (1999) Troponin: structure, properties, and mechanism of functioning. *Biochemistry (Mosc)*. **64**: 969-985.

Fyrberg, E., Fyrberg, C.C., Beall, C. & Saville, D.L. (1990) Drosophila melanogaster troponin-T mutations engender three distinct syndromes of myofibrillar abnormalities. *J Mol Biol*. **216**: 657-675.

Geeves, M.A. & Holmes, K.C. (1999) Structural mechanism of muscle contraction. *Annu Rev Biochem*. **68**: 687-728.

Gordon, A.M., Homsher, E. & Regnier, M. (2000) Regulation of contraction in striated muscle. *Physiol Rev*. **80**: 853-924.

Gremke, L., Lord, P.C., Sabacan, L., Lin, S.C., Wohlwill, A. & Storti, R.V. (1993) Coordinate regulation of Drosophila tropomyosin gene expression is controlled by multiple muscle-type-specific positive and negative enhancer elements. *Dev Biol*. **159**: 513-527.

Hanke, P.D. & Storti, R.V. (1988) The Drosophila melanogaster tropomyosin II gene produces multiple proteins by use of alternative tissue-specific promoters and alternative splicing. *Mol Cell Biol*. **8**: 3591-3602.

Hastings, G.A. & Emerson, C.P., Jr. (1991) Myosin functional domains encoded by alternative exons are expressed in specific thoracic muscles of Drosophila. *J Cell Biol*. **114**: 263-276.

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.

Herranz, R., Díaz-Castillo, C., Nguyen, T.G., Lovato, T.L., Cripps, R.M. & Marco, R. (2004) Expression pattern characterization of the whole Troponin C gene repertoire during Drosophila development. *Gene Expression Pattern*. **4**: 183-190.

Herranz, R., Mateos, J. & Marco, R. (2005) Diversification and independent evolution of troponin C genes in insects. *J Mol Evol*. **60**: 31-44.

Herranz, R., Mateos, J., Mas, J.A., García-Zaragoza, E., Cervera, M. & Marco, R. (In review) The co-evolution of insect muscle TpnT and TpnI genes. *Mol Biol Evol*.

Holtzer, A. (1986) Phenomenological analysis of the kinetics of the production of interchain disulfide cross-links in two-chain, coiled-coil proteins by reaction with 5,5'-dithiobis(2-nitrobenzoate). *Biochemistry*. **25**: 3008-3012.

Karlik, C.C. & Fyrberg, E.A. (1986) Two Drosophila melanogaster tropomyosin genes: structural and functional aspects. *Mol Cell Biol*. **6**: 1965-1973.

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

Maroto, M., Vinos, J., Marco, R. & Cervera, M. (1992) Autophosphorylating protein kinase activity in titin-like arthropod projectin. *J Mol Biol*. **224**: 287-291.

Maroto, M., Arredondo, J., Goulding, D., Marco, R., Bullard, B. & Cervera, M. (1996) Drosophila paramyosin/miniparamyosin gene products show a large diversity in quantity, localization, and isoform pattern: a possible role in muscle maturation and function. *J Cell Biol*. **134**: 81-92.

McGuffin, L.J., Bryson, K. & Jones, D.T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics*. **16**: 404-405.

Meredith, J. & Storti, R.V. (1993) Developmental regulation of the Drosophila tropomyosin II gene in different muscles is controlled by muscle-type-specific intron enhancer elements and distal and proximal promoter control elements. *Dev Biol*. **159**: 500-512.

Mogami, K., Fujita, S.C. & Hotta, Y. (1982) Identification of Drosophila indirect flight muscle myofibrillar proteins by means of two-dimensional electrophoresis. *J Biochem (Tokyo)*. **91**: 643-650.

Nongthomba, U., Ansari, M., Stark, M. & Sparrow, J.C. (In review) The Drosophila melanogaster upheld1 mutation causes aberrant splicing of a newly identified jump and flight muscle-specific exon in the troponin-T gene.

Paradela, A., Alvarez, I., Garcia-Peydro, M., Sesma, L., Ramos, M., Vazquez, J. & Lopez De Castro, J.A. (2000) Limited diversity of peptides related to an alloreactive T cell epitope in the HLA-B27-bound peptide repertoire results from restrictions at multiple steps along the processing-loading pathway. *J Immunol*. **164**: 329-337.

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.

Pringle, J. (1981) The evolution of fibrillar muscle in insects. *J. Exp. Biol.* **94**: 1-14.

Pringle, J.W. (1978) The Croonian Lecture, 1977. Stretch activation of muscle: function and mechanism. *Proc R Soc Lond B Biol Sci*. **201**: 107-130.

Qiu, F., Lakey, A., Agianian, B., Hutchings, A., Butcher, G.W., Labeit, S., *et al* (2003) Troponin C in different insect muscle types: identification of two isoforms in Lethocerus, Drosophila and Anopheles that are specific to asynchronous flight muscle in the adult insect. *Biochem J*. **371**: 811-821.

Reedy, M.C., Reedy, M.K., Leonard, K.R. & Bullard, B. (1994) Gold/Fab immuno electron microscopy localization of troponin H and troponin T in Lethocerus flight muscle. *J Mol Biol*. **239**: 52-67.

Singh, S.P., Coronella, J.A., Benes, H., Cochrane, B.J. & Zimniak, P. (2001) Catalytic function of Drosophila melanogaster glutathione S-transferase DmGSTS1-1 (GST-2) in conjugation of lipid peroxidation end products. *Eur J Biochem*. **268**: 2912-2923.

Sparrow, J.C. (1995) Muscle. Flight and phosphorylation. *Nature*. **374**: 592-593.

Thorson, J. & White, D.C. (1983) Role of cross-bridge distortion in the small-signal mechanical dynamics of insect and rabbit striated muscle. *J Physiol*. **343**: 59-84.

Tohtong, R., Yamashita, H., Graham, M., Haeberle, J., Simcox, A. & Maughan, D. (1995) Impairment of muscle function caused by mutations of phosphorylation sites in myosin regulatory light chain. *Nature*. **374**: 650-653.

Trayer, I.P., Trayer, H.R. & Levine, B.A. (1987) Evidence that the N-terminal region of A1-light chain of myosin interacts directly with the C-terminal region of actin. A proton magnetic resonance study. *Eur J Biochem*. **164**: 259-266.

Vigoreaux, J.O. (2001) Genetics of the Drosophila flight muscle myofibril: a window into the biology of complex systems. *Bioessays*. **23**: 1047-1063.

Vigoreaux, J.O. & Perry, L.M. (1994) Multiple isoelectric variants of flightin in Drosophila stretch-activated muscles are generated by temporally regulated phosphorylations. *J Muscle Res Cell Motil*. **15**: 607-616.

Vinos, J., Domingo, A., Marco, R. & Cervera, M. (1991) Identification and characterization of Drosophila melanogaster paramyosin. *J Mol Biol*. **220**: 687-700.

Wendt, T. & Leonard, K. (1999) Structure of the insect troponin complex. *J Mol Biol*. **285**: 1845-1856.

Wray, J.S., Goody, R.S. & Holmes, K.C. (1988) Towards a molecular mechanism for the crossbridge cycle. *Adv Exp Med Biol*. **226**: 49-59.

FIGURE 1: Sequence comparison of both isoforms 33 and 34 and prediction of secondary structure for the consensus sequence.

ClustalW and PSI-pred programs were used to align the sequence of both isoforms and to predict their secondary structure, respectively. In both cases, the N-terminal half has a predicted α-helical structure similar to the sequence of the entire standard Tms. On the other hand, for the C-terminal extensions the predictions are for unstructured conformations with random coiled sequences. An arrowhead indicate the position of a Cys residue in the N-terminal part, while asterisks show the presence of potential Thr phosphorylation sites in the IFM-specific C-terminal extensions. The predicted α-helical structures are indicated with cylinders, the β-sheets with arrows and the unstructured parts with plain lines. High homology sequences are shown shadowed.

FIGURE 2: Isoforms 33 and 34 phosphorylation analysis.

A) A single female *Drosophila* thorax was subjected to 2-D electrophoresis (pH range 3-10) and stained with silver nitrate using the Mann method **(Trayer *et al.*, 1987)**, compatible with mass spectrometry analyses. Main muscle isoforms could be detected, as **myosin heavy chains** (MHC), **paramyosin** (P), **arthrin** (Art), **actin** (A), **troponin T** (TnT), **standard Tms** –(127, 129), **muscle Tm 2** –(128)-, **myosin light chains** (MLC-2, AMLC), and **TmH isoforms 33 and 34.**

B) Autoradiography of a 2-D gel of thoraces of flies fed with ^{32}P showing the *in vivo* phosphorylation of isoforms 33 and 34 (detail). Protein spots were excised from the gel, separated by thin layer chromatography and the radioactivity detected confirmed as phosphorylation of threonine residues (data not shown).

C) 2-D gel stained with the cationic dye *Stains-all*, showing a differential blue spot in both TmH isoforms 33 and 34. This differential staining is a property of phosphorylated and Ca^{2+}-binding proteins. Other phosphorylated isoforms such as TnT and myosin light chains appear differentially stained too.

FIGURE 3. Isoforms 33 and 34 Mass Spectrometry analysis.

Spots 33 and 34 were excised from silver-stained 2-D gels of single female thoraces from recently eclosed (R.E.) and mature (2D) flies and MALDI-TOF analyses were performed. Peptide maps from trypsin digestions of protein 33 (1, 3) and protein 34 (2, 4) of recently emerged flies (1, 2) and fully mature flies (3, 4) are shown. (1) M/Z = 878.58 fragment corresponding to the Asn293-Lys300 non-phosphorylated peptide. (2) M/Z= 1568.14 fragment, corresponding to the Tyr292-Lys305 non-phosphorylated peptide. (3) M/Z= 958.0 and M/Z= 1834.82 fragments corresponding to the Asn293-Lys300 and Leu301-Arg316 phosphorylated peptides, respectively. (4) M/Z= 1738.10 fragment, corresponding to the Tyr292-Lys305 phosphorylated peptide. The fragmentation spectra of the three isolated, phosphorylated peptides are shown at the bottom of the figure. Phosphorylated Thr residues from fully mature fly peptides are shown in grey.

FIGURE 4. **Analysis of the tropomyosin complexes in non-reducing and diagonal gels.**

A) Protein extracts from fully mature flies were run in polyacrylamide gels in reducing (R, 1 and 3) and non reducing (NR, 2 and 4) conditions. Proteins were transferred to membranes and immunodetection using Mac144 **(1, 2)** or Mac141 **(3, 4)** was carried out. Mac144, which recognizes the IFM-specific part, detects GST-2 and isoforms 33 and 34 (1), but in non-reducing conditions can detect a delayed complex with an apparent MW ≈ 120 kDa. Using Mac141, which recognizes the common sequence, isoforms 33 and 34 and standard Tm (protein 129) were detected (3), but in non-reducing conditions a delayed complex with an apparent MW ≈ 120 kDa was also detected. Both results strongly suggest that isoforms 33 and 34 are forming protein complexes *in vivo*.

B) For a more detailed analysis of these interactions, diagonal gel electrophoresis of protein extracts from fully mature flies were run. The gels were transferred and immunodetection was performed using Mac144 (1) and Mac141 (2). Several signals below the theoretic diagonal (given by a straight line from the origin of the gel to the signal of the monomeric low MW isoform) were detected. Using Mac144, homodimers of standard Tm were detected (short arrow), as well as heterodimers between standard Tm and TmH 33 and/or 34 (long arrow). Dimers between both TmH 33 and 34 are also present (asterisk). In addition, the previously described interaction between GST-2 and isoforms 33/34 is detected (short arrow) using Mac141. No signal from monomeric isoforms 33 or 34 is detected, while a significant fraction of Tm 127 appears to be on monomeric form.

FIGURE 5. The Tm gene complement is conserved in *Drosophila pseudoobscura* and *Anopheles gambiae* but not in *Apis mellifera.*

A) Comparison of the structure of the tropomyosin gene complement in Diptera and Hemiptera. Only the muscle exons are shown for clarity. All genes contain a similar set of additional exons encoding the cytoplasmic isoform (Cytoplasmic tropomyosin specific exons (4, 6, 10 and 14 of Drosophilidae are not shown). Exons common to both cytoplasmic and muscles isoforms are shown here (5, 10 and 13 of Drosophilidae). Homology is very high, and lower, but still significant, in the IFM-specific exons -outlined boxes-. Even though the structure is similar, slight differences are detected. In *Anopheles*, exons 2 and 3 are fused and there is only one IFM-specific exon, more closely related to exon 16, suggesting an exon duplication in the *Drosophila* lineage. The genomic region is also bigger in this organism, due to the larger size of its introns. The Tm gene complement in *Apis* lacks IFM specific-type sequences, which have been described in the TnI gene (**Herranz *et al.*, in review**). The last *Tm1* exons (2 in *Anopheles* and *Apis* and 3 in Drosophilidae) are alternatively spliced. The initial part of exon 16/14 of Diptera is highly analogous to the last but one exon of *Apis*.

B) Sequence comparison of the *Anopheles* differential exon, exons 15 and 16 from both *Drosophila* species and the sequence corresponding to exons H1, H2 and H3 of the *Apis* TnI, responsible of the TnH in this species. High homology regions are shown in boxes. The phosphorylated Thr residues detected in *Drosophila melanogaster* are conserved in the other two species (asterisks), as well as the peptide sequences.

Components	Methionines in the molecule	Exp 1	Exp 2	Exp 3	Exp 4	Lysines & arginines in the molecule	Exp 5	Exp6	Mean ± standard deviation	Model
Actin*	15	11.92	12.01	12.02	12.14	45	10.06	9.56	11.29±1.56	12
Arthrin*	16	2.08	1.9	1.98	1.86	57	2.94	4.44	2.53±1.02	2
Tn-T	6	2.47	2.2	2.52	1.05	85	1.9	1.74	1.98±0.55	2
Tm	7	1.26	1.54	1.42	2.66	49	3.23	2.84	2.16±0.85	2
TnI**	4	2.47**	2.2**	2.52	1.05	54	1.9	1.74	1.98±0.55**	2
Isoform 33	7	1.18	1.2	1.34	0.8	58	1.6	1.12	1.21±0.26	1
Isoform 34	9	0.96	0.92	0.86	0.62	56	0.96	1.68	1±0.36	1
Tm+ 33 + 34		3.40	3.64	3.64	4.08		5.85	5.64	4.38±1.09	4

TABLE I. Stoichiometry of thin filament components in *Drosophila* IFM

*The quantitative information obtained from the densitometry of the autoradiographies (in the case of the ^{35}S-labelling or the Coomassie gels) was normalized so that the addition of the arthrin and actin numbers sum to 14). The actual value of methionines, lysines and arginines of arthrin were obtained by adding to actin the value found in *Drosophila* ubiquitin.

**The band that allows the quantification of the low molecular weight Tm also contains TnI. To quantify the actual level of Tm in the different experiments, we have assumed that the actual number of TnI molecules match the number of TnT molecules and taken into account the number of amino acids, we have obtained the numbers of Tm molecules listed in the table by substracting the contribution of the TnI from the total of the band.

FIGURE 1

Dm TmH33 MDAIKKKMQAMKVDKDGALERALVCEQEARDANTRAEKAEEEARQLQKKIQTVENELDQTQEALTLVTGKLEEKNKALQN 80
Dm TmH34 MDAIKKKMQAMKVDKDGALERALVCEQEARDANTRAEKAEEEARQLQKKIQTVENELDQTQEALTLVTGKLEEKNKALQN 80

Dm TmH33 AESEVAALNRRIQLLEEDLERSEERLGSATAKLSEASQAADESERARKILENRALADEERMDALENQLKEARFLAEEADK 160
Dm TmH34 AESEVAALNRRIQLLEEDLERSEERLGSATAKLSEASQAADESERARKILENRALADEERMDALENQLKEARFLAEEADK 160

Dm TmH33 KYDEVARKLAMVEADLERAEERAEQGENKIVELEEELRVVGNNLKSLEVSEEKANQREEEYKNQIKTLNTRLKEAEARAE 240
Dm TmH34 KYDEVARKLAMVEADLERAEERAEQGENKIVELEEELRVVGNNLKSLEVSEEKANQREEEYKNQIKTLNTRLKEAEARAE 240

Dm TmH33 FAERSVQKLQKEVDRLEDDLIVEKERYCMIGDSLDEAFVDLIKGLEPFWNPRNPKPPTPKLPTPTPEELAAMEEAR-AAA 319
Dm TmH34 FAERSVQKLQKEVDRLEDEMIKEIEHYALVGDQLDWTFVEMM-GMPPFYNERYPKPPTPEL---TEEEKAALEAAAIAEA 316

Dm TmH33 EAAAAAEAEAAEAAAAAGEAGADGAP---AAPGEEKAPAKEP---TPPKEP--TPPPPPPPPFEYSIDLPPEGAEVPYVK 391
Dm TmH34 EAKARAEELAALGEEAGAEAGEGGAPAEGAAPGEPGA-ATEPGVEAPPAEPERIPTPPPPPPFEYSIDLPPEGAEVPYVR 395

Dm TmH33 NYEPPPPGSEPEPVPAAEGEAAPAAEGAAPPAEGAAPPAEGAVPPADGAAPPAEGAAPAAEGAAPPADGAAPPAEAAAAP 471
Dm TmH34 NAEP---GDFAPPAEAAPAEGAPPAEG-APAAEGAAP-AEG-APAAEG-APPAEGAP----APAPAEGEAAPPAPAAEGD 454

Dm TmH33 ADAAAPAAEAAPAEAPAAEATAA-EAPPAEAAPAEAAPAAAEGEAPPA 518
Dm TmH34 AAAAPPP---PPAEGEAAPAPAEGEAPPAP--PAEAAPAA---EAPPA 510

Legend:

= helix

= strand

= coil

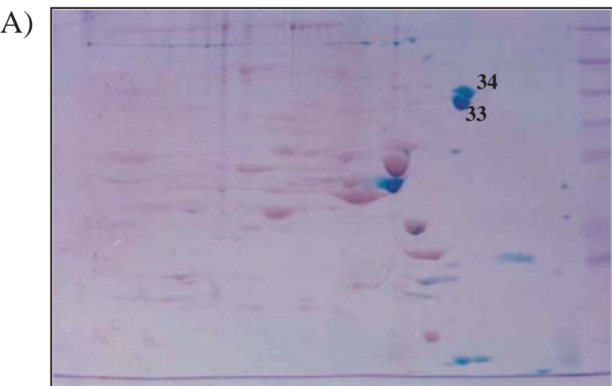
Conf: = confidence of prediction

\- +

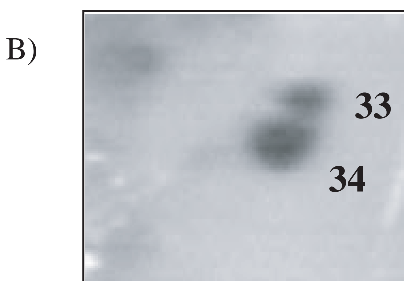
Pred: predicted secondary structure

FIGURE 2

A)



B)



C)

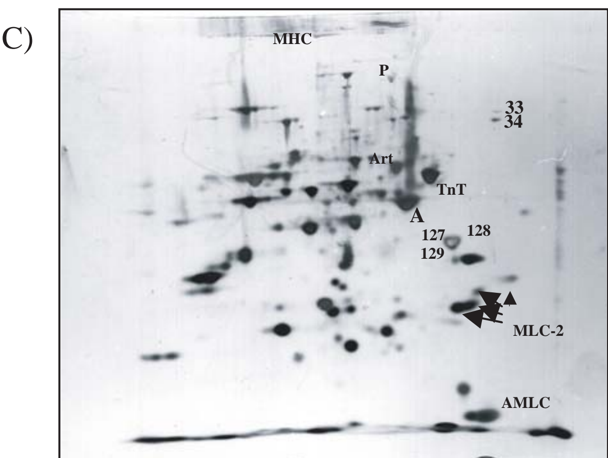


FIGURE 3

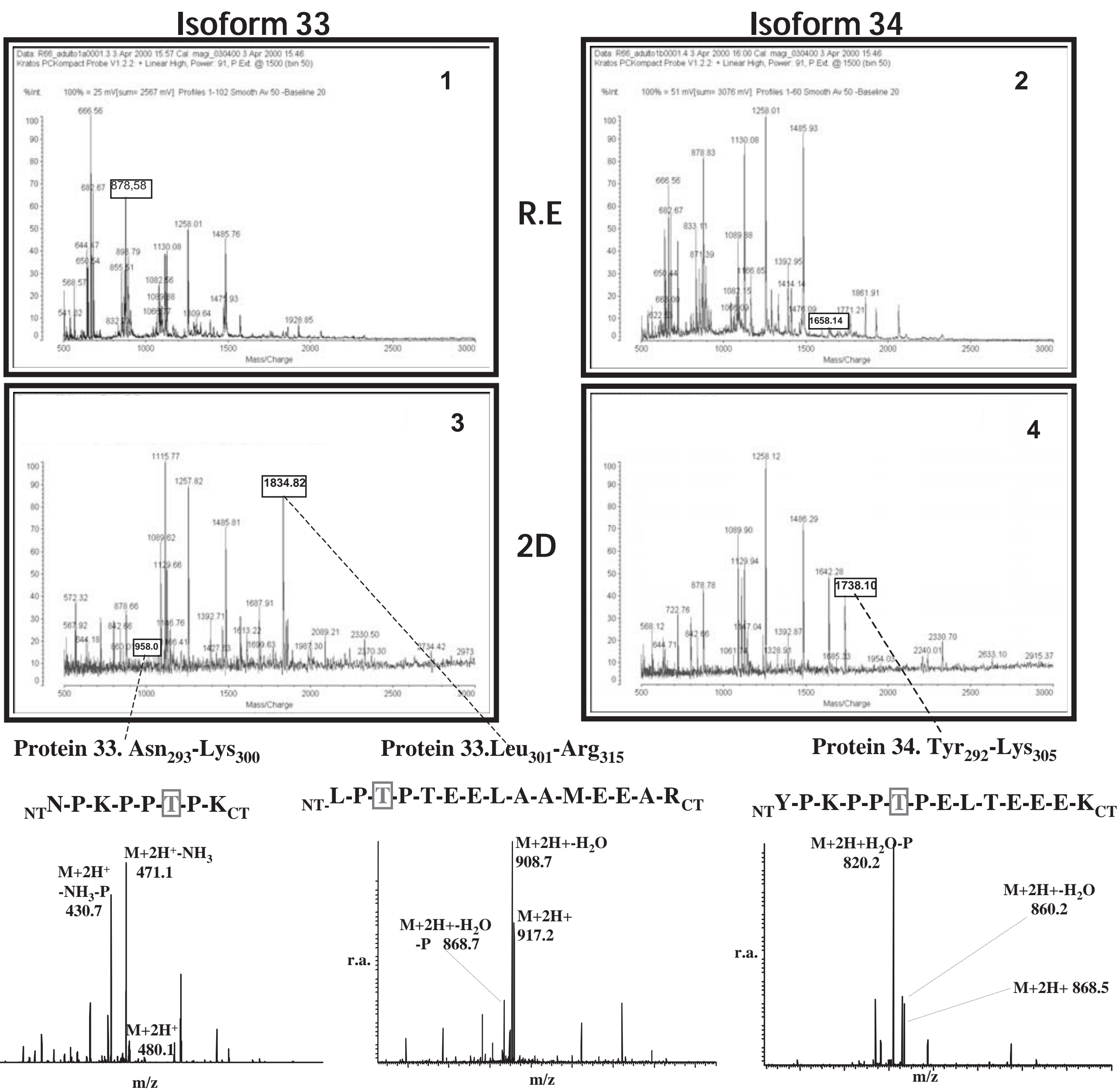


FIGURE 4

A) 1 2 3 4

184.5
121.3
85.9
68.8
52.5
40.0
28.4

R NR
Mac144

R NR
Mac141

B)

- BetaMEtOH

+ BetaMEtOH

184.5
121.3
85.9
68.8
52.5
40.0

Mac144

Mac141

FIGURE 5

A)

TM1
TM2

Drosophila melanogaster
15
16

1 Kb

Drosophila pseudoobscura
15
16

Anopheles gambiae
14

?

Apis mellifera

?

B)

```
                                                                   *           *
DmTm1-E15       1   ---EMIKEIE HYALVGDQLD WTFVEMM-GM PPFYNERYPK PPTPEL---- -TEEEK---A ALEAAAIAEA EAKARAEELA ALGEEAGAEA GEGGAPAEGA  88
DpTm1-E15       1   ---ETIVEKE HYALIGDSLD WTFVELM-GM PPFYNDRYPK PPTPEL---- -TEEDLARIA EAEAQAQAEA EAKAQAAEAK AQAAAALAEA LAAGVDPATL  91
DmTm1-E16       1   ---DLIVEKE RYCMIGDSLD EAFVDLIKGL EPFWNPRNPK PPTPKL---- PTPTPEELAA MEEARAAAEA AAAAEA---- -EAAEAAAAA GEAGADGA--  86
DpTm1-E16       1   ---DMIVEKE RYCLINDSLD EAFVDVIVGL EPFWNPRNPK PPTPEL---- -TEEDLARIA --EAKAAAEA KAEAEAAEQA ERVAAAVAAA LAAGIDPATV  90
AgTm1-E14       1   ---ELIVEKL RYAEIGDDLD FAFVDLIPGI DPIWTERRPK PKTPPPPPKL PTP-P----P P----PP--- APKEEA---- -------AA- -AAGGEGE--  70
AmTpnI-EH1,2,3  1   PDWSKKGEEK KEEPAPPPEP PQPTPSPTPA PEAAPTPSPE PPTPSP---- -TPSP---VP PAEPAPPSEP APPAEG---- ----AAPPPP PAEGAPAEG-  83

DmTm1-E15       89  APGEPGAATE PG----VEAP PAEP-ERIPT PPPPPPFEYS IDLPPEGAEV PYVRNAEP-- -GDFAPPAEA APAEGAPPAE GAPAAEGAAP AEGAP-AAEG  179
DpTm1-E15       92  DPALVEAATA GEEPVPIKIP TPPPKEPTPP PPPPPPFEYS IDLPPEGAEV PYVKNAEP-- -GDMLPPVEG AP-------E GEAAAVPPVE GEAAP-AAAP  180
DmTm1-E16       86  ----PAAPGE EK--APAKEP TPP-KEPTPP PPPPPPFEYS IDLPPEGAEV PYVKNYEPPP PGSEPEPVPA AEGEAAP--A -AEGAAPPAE GAAP--PAEG  174
DpTm1-E16       91  V--VEGAPGE PP--VPIKIP TPPPKEPTPP PPPPPPFEYS IDLPPEGAEV PYVKNYEPPP PGSEPEPVPA AEGEAA---- ---AAAPPAE GEAPAAPAEG  179
AgTm1-E14       70  ----AAAAGG EG------AP AGPTKARTPS P-----FELR KHFPPEAAEV PYVRSSTG-T TPKSSPAKEV APAEAAP--- -APEAAAPAE APAEG-GADA  149
AmTpnI-EH1,2,3  83  -----AAPPA EG------AP PAAPAEGAAP PAEGAPAAPP AEGAPPAESA PPAEGAPP-P EGAPPPAEGA PPAEGAP--- -PAEGAPPAE GAPAPPPTEA  167

DmTm1-E15       180 APPAEGAPAP APAEGEAA-- -PPAPAAEGD AAAAPPPPP- -----AEGEA APAPA---EG EAPPAPPAEA APAAEAPPA- ------- 246
DpTm1-E15       181 AEGEAAAPAA APAEGEAP-- -SVAAPAEGE AAAAPP---- ------AGEA PP-------- --PPPAAAEA APAAEAPPA- ------- 236
DmTm1-E16       175 AVPP-ADGAA PPAEGAAPAA EGAAPPADGA APPAEAAAAP ADAAAPAAEA APAEAPAAEA TAAEAPPAEA APAEAAPAAA EGEAPPA 260
DpTm1-E16       180 AAPPPADGAA PPAEGAAP-A EGAAPAAEGA APAAEAAPA- AEAAPAAAEA APAAEAAPAA EAAPA--AEA APAEAPPA-- ------- 253
AgTm1-E14       150 AAPP-AEAAA PAAEGEAP-- --AAPAAEGE AAAPAP---- ------EAAA AP-------- -AEEAAPPAA EPAPEAPAS- ------- 204
AmTpnI-EH1,2,3  168 APPAEGAPAP APPAEGTP-- --APPPAEGA PPPAEGAPP- ----AAPAEA APAAP---PA EAAPAPPPAE GNPLLAPFLT DYD---- 238
```