Molecular Analysis of the Muscle Protein Projectin in Lepidoptera

Authors: Ayme-Southgate, A. J., Turner, L., and Southgate, R. J.

Source: Journal of Insect Science, 13(88) : 1-14

Published By: Entomological Society of America

URL: https://doi.org/10.1673/031.013.8801

Journal of Insect Science
www.insectscience.org

Molecular analysis of the muscle protein projectin in Lepidoptera

A. J. Ayme-Southgate[1a]*, L. Turner[1,2b], R. J. Southgate[1c]

[1]Department of Biology, College of Charleston, 66 George Street, Charleston, SC 29401
[2]Current address: Central Carolina Technical College, Sumter, SC

Abstract

Striated muscles of both vertebrates and insects contain a third filament composed of the giant proteins, namely kettin and projectin (insects) and titin (vertebrates). All three proteins have been shown to contain several domains implicated in conferring elasticity, in particular a PEVK segment. In this study, the characterization of the projectin protein in the silkmoth, *Bombyx mori* L. (Lepidoptera: Bombycidae), and the monarch butterfly, *Danaus plexippus* L. (Lepidoptera: Nymphalidae), as well as a partial characterization in the Carolina sphinx, *Manduca sexta* L. (Lepidoptera: Sphingidae), are presented. This study showed that, similar to other insects, projectin's overall modular organization was conserved, but in contrast, the PEVK region had a highly divergent sequence. The analysis of alternative splicing in the PEVK region revealed a small number of possible isoforms and the lack of a flight-muscle specific variant, both characteristics being in sharp contrast with findings from other insects. The possible correlation with difference in flight muscle stiffness and physiology between Lepidoptera and other insect orders is discussed.

Keywords: alternative splicing, elastic filaments, flight muscle, insect, titin
Abbreviations: Fn, fibronectin type III; **Ig**, immunoglobulin; **NTCS-1**, N-terminal conserved sequence 1
Correspondence: [a] southgatea@cofc.edu, [b] turnerl@cctech.edu, [c] southgater@cofc.edu,
*Corresponding author
Received: 27 January 2012 **Accepted:** 8 May 2012

ISSN: 1536-2442 | Vol. 13, Number 88

Cite this paper as:
Ayme-Southgate AJ, Turner L, Southgate RJ. 2013. Molecular analysis of the muscle protein projectin in Lepidoptera. *Journal of Insect Science* 13:88. Available online: http://www.insectscience.org/13.88

Introduction

How insect flight originated and how structures associated with flight evolved are still mostly unknown, and as groups of insects acquire additional features, they are considered "derived" compared to earlier basal groups. The flight musculature of derived insects relies on indirect flight muscles, which are known to contain connecting C-filaments providing a direct mechanical link between the muscle's sarcomeric Z-discs and the ends of the thick filaments (Trombitas 2000). In these muscles, the passive myofibrillar elasticity attributed to the C-filaments resides with several proteins, projectin, and several isoforms from the sallimus gene, including the most abundant form, which is known as kettin. Projectin, kettin, and other isoforms of sallimus are therefore proposed to be responsible for the high resting stiffness of indirect flight muscles (Granzier and Wang 1993; Moore et al. 1999; Bullard et al. 2000, 2002, 2005; Trombitas 2000; Vigoreaux et al. 2000; Kulke et al. 2001; Burkart et al. 2007).

The complete amino acid sequence of projectin is currently available in several insect species from five different orders and reveals that projectin's modular organization is highly conserved with its specific pattern of repeated motifs and unique sequences (Ayme-Southgate et al. 2008). This modular structure, as well as the arrangement of motifs, are actually common to all invertebrate projectins characterized so far, including twitchin in *Caenorhabditis elegans* and projectin in crayfish, *Procambarus clarkii*, even though the number of Ig domains at the NH_2-terminus is lower (for example 7 rather than 8 in the crayfish; Benian et al. 1989; 1993, Oshino et al 2003).

The NH_2-terminus of projectin contains a unique region that in previous studies was delineated into two segments: a PEVK region followed by the so-called NTCS1 segment (N-terminal conserved sequence 1; Ayme-Southgate et al. 2011). Sequence comparison of the PEVK segments across several insect species revealed a series of unique features: an enrichment in 4 specific amino acids (Proline (P), Glutamic acid (E), Valine (V), and Lysine (K)), a highly divergent primary sequence, and a complex pattern of alternative splicing (Southgate and Ayme-Southgate 2001; Ayme-Southgate et al. 2008, 2011). In all the species investigated so far, alternative splicing of the projectin PEVK region has been shown to generate isoforms ranging in lengths from 34 to 624 amino acids. and a P, E, V, and K composition from 42% to 100% (Southgate and Ayme-Southgate 2001; Ayme-Southgate et al 2008, 2011).

The vertebrate protein, titin, makes up the third, elastic filament of striated muscles. Although titin is larger in size, it contains the same domains as projectin, in particular a PEVK region, which is longer and more complex (Labeit et al. 1992; Labeit et al. 1997). The titin PEVK region undergoes extensive alternative splicing events, and variable lengths of the PEVK region found in different muscle types are associated with significant divergence in passive tension (Cazorla et al. 2000; Freiburg et al. 2000; Granzier and Labeit 2002, 2005).

Analysis of the projectin PEVK region in insects such as dragonflies, *Pachydiplax longipennis* Burmeister (Odonata: Libellulidae) and *Libellula pulchella* Drury, *Apis mellifera* L. (Hymenoptera: Apidae), and *Drosophila melanogaster* Meigen (Diptera: Drosophilidae) has shown that their flight

muscles contain a short isoform that is absent from other muscles in the same insect (Ayme-Southgate et al. 2005, 2011). These same insect orders are also known to have flight muscles with relatively high stiffness (Thorson and White 1983; White 1983; Peckham et al. 1990, 1992). To evaluate a possible correlation between projectin PEVK variants and muscle stiffness, the features of the PEVK region and splicing pattern in insects with lower flight muscle stiffness need to be established. It is generally proposed that synchronous flight muscles would display lower passive tension (reviewed in Pringle 1977 and Dudley 2000). To this purpose, insects from the Lepidoptera order, which have synchronous muscles (Pringle 1981), were used. Furthermore, the Lepidoptera order was used in order to take advantage of the availability of the genomes of the silk moth, *Bombyx mori* L. (Bombycidae) (Mita et al. 2004; International Silkworm Genome Consortium 2008), and the monarch butterfly, *Danaus plexippus* L. (Nymphalidae) (Zhan et al. 2011), to establish the overall structure of their projectin genes. The sequence for NH_2-terminal region of projectin in the Carolina sphinx, *Manduca sexta* L. (Sphingidae), was determined, and the analysis of its PEVK region was completed.

The analysis showed that the overall domain pattern of the projectin protein was conserved, and the PEVK region followed the features previously identified. However, even though alternative splicing of the *M. sexta* PEVK segment occurred, the number of variants was low, and there was no evidence for the presence of a short, flight muscle-specific PEVK isoform.

Materials and Methods

Insects and RNA sample preparation

B. mori and *M. sexta* were purchased as larvae and/or pupae from Educational Science (www.educationalscience.com) and reared up to emergence of the imago, at which point they were dissected. Total RNA was purified from whole animals, isolated body parts (legs, heads, thoraces), and from flight muscles using Trizol (www.invitrogen.com) as previously described by Ayme-Southgate et al. (2008).

Degenerate primers and splicing analysis

Degenerate primers have been described elsewhere (Ayme-Southgate et al. 2011). RT-PCR reactions were performed as described before with different RNA preparations and primer sets (Southgate and Ayme-Southgate 2001). Annealing for both the RT and PCR reactions were tested using a range of temperatures to optimize each primer set. DNA fragments were isolated after agarose gel electrophoresis and subcloned into the pGEM-T easy shuttle vector (Promega, www.promega.com), which was then followed by sequencing (Genewiz Inc., www.genewiz.com). Sequences from overlapping clones were manually assembled into contigs.

Bioinformatics analysis

For the *B. mori* projectin, contigs were isolated following tblastn searches of GenBank WGS database using a series of gene fragments from *D. melanogaster* projectin. The tblastn algorithm compares a query protein sequence to the 6-frame-translation of a DNA sequence; in this case, the contigs available for the *B. mori* genome (Altschul et al. 1990). The EST database from the two silkworm transcriptome projects, available through SILKBASE and GenBank, was also searched, and resulting EST sequences were aligned on the genomic sequence. For the *D. plexippus* projectin, a BLAST search of the recently

published genome was performed using *B. mori* projectin cDNA as query. Sequence comparisons were carried out using the CLUSTALW algorithm, and the alignments were viewed in Jalview (Thompson et al. 1994, 1997).

Supplementary Table. Summary of contigs from the *B. mori* genome database assembled to generate projectin genomic sequence.

AB01141722	Ig1	
AADK01016653, AADK01028815	Ig1 2nd half + overlap with 1141722	
	Gap in intron between Ig1 and unique 1 , cDNA bridged	
BABH01025785	U1-Ig2 to mid PEVK (just before 2nd YERP)	
AADK01022116	Ig2, Ig3, Ig4, Ig5, Ig6, Ig7	
BAAB01027333	end Ig2 + Ig3, Ig4	
AADK01055162	Ig5, Ig6, Ig7	
BAAB01089380	Ig5, Ig6, Ig7, Ig8 + begin PEVK	
AADK01007952	Ig8 end	
AADK01036325	Ig8 + mid PEVK	
BAAB01155026	PEVK, overlap 36325	
AADK01028285	PEVK, overlap 25785 +extend slightly	
AADK01040746	PEVK	
BAAB01049506	PEVK	
BABH01025784	PEVK	
AADK01015981	PEVK, Ig9, FRAM	
	Gap, cDNA bridged	
BABH01025782	Ig10	
BAAB01127199	Ig10 etc	
AADK01011457	Ig 10, etc Fn1, Fn2, Ig 15	
BAAB01051365.	Fn1	
BAAB01212557	Ig15	
AADK01005986	Fn2	
BAAB01146667	Ig15	
AADK01057796	Fn3	
	Gap in intron between Fn3 and Fn4 , cDNA bridged	
BAAB01169175.	Fn4	
AADK01007952	Fn4, Ig16, Fn5-6, Ig17,Fn 7-8beg, Fn8end, Ig18,Fn 9-10	Fn4 to beg Fn8 on reverse but Fn8end-to Fn10 on forward
BAAB01107024	Ig16, Fn5	
BAAB01151021	Fn6	
BAAB01031504	Ig17	
BAAB01150558	Fn9	
	Gap in intron between Fn10 and Ig19 , cDNA bridged	
AADK01005878	Ig 19, Fn11 , Ig22 to mid Fn19	
BAAB01126214	Fn11	
BAAB01025862	Fn13 to Ig28	
BAAB01061332	Fn15	
BAAB01044120	Fn16, 17	
AADK01024425	mid Fn19	
AADK01024425	to Fn21	
BABH01047754	end Fn21-Fn23	
AADK01005986	Fn22-Ig25 onward to begin kinase	deletion seems to remove Ig 34 and Fn39.
BAAB01122158	Fn39-mid kinase	
BABH01025779	Fn24-mid kinase	
	gap in kinase coding, cDNA bridged	
BABH01025778	end kinase	
AADK01016853	end kinase// Ig35 through first half Ig37	mixed up Ig35,36 on one strand, end kinase on other strand
	incertainty in intron between end kinase and Ig35, cDNA bridged	
BAAB01015278	Ig35	
BABH01025776	2nd half Ig37-end	overlap with AADK01016853 in intron within Ig37
BAAB01065217	2nd half Ig37	
EST clones		
NRPG1349	Ig5-Ig7	
BmNP08_FL5_A11 and BmNP08_T7_G21	Fn20-Ig24-Fn21-Fn22	
NRPG1646	2nd half of kinase, Ig35	

Results

B. mori and *D. plexippus* projectin sequences

The annotation in the GenBank database reported a *B. mori* projectin homolog based on genome analysis of the *B. mori* Z chromosome (Koike et al. 2003). This annotation predicted a series of short proteins with names such as D-titin, kettin, and projectin-like containing immunoglobulin (Ig) and/or fibronectin type III (Fn) domains (GenBank ID: NM_001114995 for projectin-like). None of these entries represented the correct gene for the projectin protein, as they were too short and did not contain the characteristic pattern of Fn and Ig domains expected for projectin. At the start of this project, the *B. mori* projectin gene was assembled *de novo* following tblastn searches of *B. mori* supercontigs (Mita et al. 2004) available in the GenBank database using a series of peptide fragments from the *D. melanogaster* projectin. Extensive overlap was established for most of the contigs representing the *B. mori* projectin gene (Supplementary Table 1). Some of the contigs available in GenBank were probably incorrectly assembled, however, as projectin domains known to be adjacent are coded on opposite strands or in reverse order. In other cases, some contigs contained a deletion covering several domains, which were present in other contigs. Many of these issues could have resulted from the presence of repetitive sequences. Despite our efforts, five gaps remain in the genomic sequence, as repeated searches of the GenBank database fail to return contigs overlapping these gaps, and we assume that these sequences were not obtained during the original genome sequencing. We have not attempted to "clean" the original assembly or to sequence the missing genomic DNA (see below).

The exon-intron pattern was predicted over most of the gene by performing translation in all three frames and visual alignment with the *D. melanogaster* projectin amino acid sequence. The PEVK region could not be entirely predicted by this approach (see below). Ambiguous splice sites were resolved by RT-PCR using *B. mori* total RNA followed by sequencing of cDNA products. In cases where the assembly of the genomic contigs was ambiguous, the prediction was verified by RT-PCR amplification across the gaps/misalignment and cDNA sequencing (see Materials and Methods for details; data not shown). EST sequences from two silkworm transcriptome projects available through SILKBASE and GenBank (Mita et al. 2003) were retrieved. The few EST sequences available were consistent with the exon-intron prediction. Of the five gaps remaining in the genomic sequence, all but one occurred within intron sequences and were bridged by our own cDNA sequencing or EST data, so it is indeed a continuous gene. The one gap falling within a coding region encompassed a segment of the kinase domain. The corresponding cDNA sequence was obtained following RT-PCR, but the exact exon-intron pattern for this part of the gene is uncertain because the genomic sequence was unavailable.

The projectin gene for *D. plexippus* was retrieved from the recently available genome data (Zhan et al. 2011). The combined sequences from two super-contigs (GenBank ID # AGBW01006173.1 and AGBW01009765.1) cover the entire projectin gene, except for half of the second Ig domain, which is located in the gap between the two super-contigs and for which no genomic data are available.

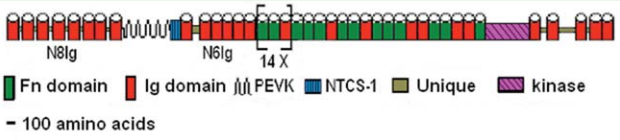

Figure 1. Diagram of the conserved projectin domain organization. This pattern is common to all projectin characterized so far, including the lepidopteran proteins from this study. The Ig and FnIII domains are represented as barrels to reflect their globular nature. The [Fn-Fn-Ig] module is repeated 14 times within the central core region. The NH_2-terminus is composed of two tracts of eight (N8Ig) and six (N6Ig) Ig domains separated by the PEVK-NTCS1 region. High quality figures are available online.

Table 1. Characteristics of the projectin gene in Lepidoptera. The lepidopteran genes are compared to the projectin genes in other insect species. *Dmel: Drosophila melanogaster; Amel: Apis mellifera; Tcas: Tribolium castaneum; Api: Acyrtosiphon pisum; Bmo: Bombyx mori,* and *Dple: Danaus plexippus.*

Characteristics of the projectin gene in Lepidoptera

	A. pisum	*A. mel*	*T. cas*	*D. mel*	*B. mori*	*D. ple*
gene size (bp)	>70,000	66,952	37,040	51,073	>116,150	>72,600
cDNA size (bp)	26,132	26,012	26,184	27,154	26,211	~26,800
# exons	144	96	64	46	124	122
largest exon (bp)	1,054	2,325	2,670	9,462	4,989	5,015
amino acids #	8,607	8,730	8,379	8,923	8,737	~ 8,700
calculated MW	960,331	995,346	933,354	998,292	973,530	~ 970,000

The gene characteristics for *B. mori* and *D. plexippus* projectins are summarized in Table 1 and compared with *D. melanogaster*, *Tribolium castaneum* Herbst (Coleoptera: Tenebrionidae), *A. mellifera*, and *Acyrthosiphum pisum* Harris (Hemiptera:Aphididae) (Ayme-Southgate et al. 2008). The overall domain organization of projectin in the two lepidopteran insects was identical to all the projectin proteins characterized so far (Figure 1; Ayme-Southgate et al. 2008, 2011). Even though the full length of the two lepidopteran projectin genes could not be ascertained completely, they were the largest of all the projectin genes characterized, with one of the highest number of exons (Table 1). Similar to the situation found in *D. melanogaster*, the largest exon in both genes contained approximately half of the domains for the core region (the section of the protein composed of the repeated Fn-Fn-Ig modules). For the remainder of the protein, individual Ig and Fn domains were often split between two exons, a situation more similar to the one found in insects from more basal orders such as *A, pisum* (Ayme-Southgate et al. 2008).

N-terminus sequence determination in *M. sexta*

To gain access to the NH_2-terminal sequence of projectin in *M. sexta*, a series of degenerate primers based on sequence alignment of several Ig domains from the N8Ig and N6Ig tracts was used (Figure 1; Ayme-Southgate et al. 2011), as well as primers based on *B. mori* sequences. RT-PCR amplifications using *M. sexta* RNA were performed with these different primer sets, and these products were cloned and sequenced. The NH_2-terminus region for *M. sexta* followed the standard pattern found in all projectin genes, which is two separate tracts of 8 and 6 Ig domains respectively with small interspersed linker sequences of 5 to 46 amino acids in length, separated by a unique sequence (see Figure 1).

PEVK structure in *B. mori, D. plexippus,* and *M. sexta* genes

Species-specific primers from *M. sexta* Ig8 and Ig9 domains were used to amplify the unique sequence between the N8Ig and N6Ig regions. Internal primers were used in a second stage to try and amplify larger cDNA products for the *M. sexta* PEVK-NTCS-1 segments. The cDNA sequence was aligned to the corresponding genomic sequences in *B. mori* and *D. plexippus.* Using this approach,

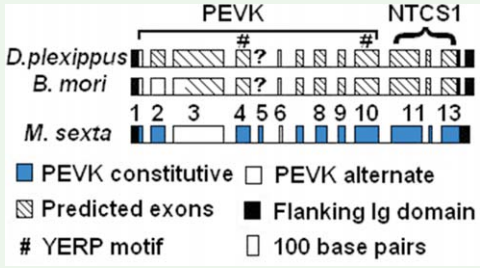

Figure 2. Exon-intron pattern for the PEVK-NTCS-1 segments in lepidopteran projectins. The striped boxes indicate exons that are only predicted by alignment to *Manduca sexta.* The ? indicates the exon that was not found in *Bombyx mori* and *Danaus plexippus.* PEVK exon numbers are indicated above the *M. sexta* map. Introns are not to scale. High quality figures are available online.

the PEVK and NTCS-1 exons for both insects, as well as the intron-exon boundaries in all three species, were predicted (Figure 2). Only one of the *M. sexta* exons could not be identified in either *B. mori* or *D. plexippus*, possibly because this exon was very short, with only 24 nucleotides (exon #5 in Figure 2). Several of the splice sites were confirmed in *M. sexta* through the sequencing of alternate splice products (see below).

The possibility of additional exons in the PEVK region does exist, but these exons would probably be rarely expressed, as this region was thoroughly amplified in *M. sexta*, and all resulting cDNAs were sequenced. Also, the available continuous *B. mori* and *D. plexippus* PEVK genomic sequences were translated in all three frames and visually scanned for open stretches of at least 10 amino acids with elevated PEVK content. No additional exons were predicted using this approach (data not shown).

Alignment of PEVK-NTCS-1 segments from *B. mori*, *D. plexippus,* and *M. sexta* with those available from some of the other species was performed using CLUSTALW and viewed with Jalview (see Materials and Methods for details). As shown by the alignment presented in Figure 3, the current subdivision of this unique sequence into two segments is supported; there is a highly divergent PEVK region and the conserved NTCS-1 segment, which is positioned just before the second stretch of six Ig domains (solid black line above alignment in Figure 3). Contrary to the PEVK segments found in other proteins (human titin, *C. elegans* TTN-1, and *Drosophila* sallimus), the projectin PEVK regions described here did not contain any repeating pattern. This was consistent with all other insect projectin PEVK segments, except for a short repeat found exclusively in *A. mellifera*

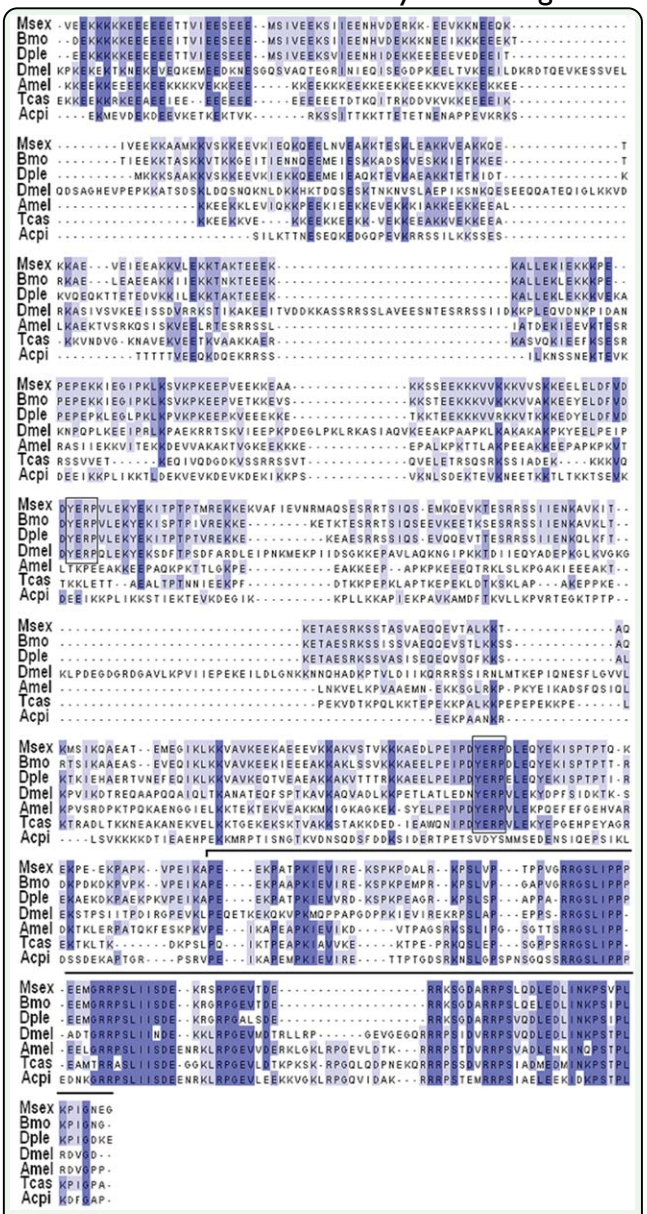

Figure 3. CLUSTAL-W-generated amino acid sequence alignment between PEVK-NTCS-1 regions in different insects. The intensity of the shading indicates the level of amino acid identity. Black boxes indicate the position of the YERP motifs. The solid black line corresponds to the NTCS-1 segment. *Dmel: Drosophila melanogaster; Amel: Apis mellifera; Tcas: Tribolium castaneum; Api: Acyrtosiphon pisum; Bmo: Bombyx mori,* and *Dple: Danaus plexippus*. High quality figures are available online.

projectin (Ayme-Southgate et al. 2011). The NTCS-1 region was described as the largest conserved region when comparing projectins from basal (dragonfly) and more derived insects (Ayme-Southgate et al. 2011). In derived insects, including the three lepidopteran sequences described here, this conserved segment was slightly longer (by 28–35 amino

acids) and began with a "YERP" motif (boxed residues in the alignment; Figure 3). This conserved block was not present in the sequences of *Pediculus humanus* L. (Phthiraptera: Pediculidae), *A. pisum*, or the two dragonfly species, *P. longipennis* and *L. pulchella* (Ayme-Southgate et al. 2011). Two YERP motifs were present in the lepidopteran and *D. melanogaster* PEVK regions (black boxes in Figure 3 and # in Figure 4).

Alternative isoforms of the PEVK-NTCS-1 segments in different muscle types

The alternative splicing pattern for the PEVK-NTCS-1 region was also ascertained, and the analysis indicated that the PEVK region was the site of several alternative splicing combinations. As shown in Figure 4, there were only two exons that could be alternatively spliced to generate the shortest form (exons # 3 and 6), compared to 11 exons in *D. melanogaster* and *T. castaneum* and up to 21 in *A. mellifera*. Only 4 PEVK variants were detected in *M. sexta* PEVK, compared to at least 10 in *D. melanogaster* (Southgate and Ayme-Southgate 2001). The longest splice variant identified in *M. sexta* would encode a 377 amino acid-long PEVK region, and the shortest form would be 205 amino acids. The length difference between the longest and shortest variants was therefore not as "striking" in *M. sexta* as it is in other insects, for example 75 and 530 amino acids for the shortest and longest variants respectively in *D. melanogaster* (Table 2). Also, the YERP motifs were included in all of the *M. sexta* PEVK variants, whereas they were excluded from the short variant in other insects (Figure 4). The characteristics of the PEVK and NTCS-1 segments for all three lepidopteran genes are summarized in Table 2 together with the corresponding regions in other insect projectin proteins.

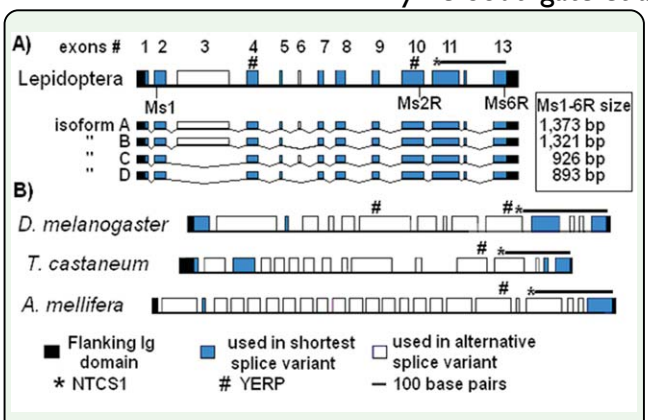

Figure 4. Diagram representation of the PEVK exons in different insect species. A) Exon pattern in Lepidoptera, together with the isoforms detected in *Manduca sexta*. Exon composition for the alternative variants is presented below the map together with their expected sizes following amplification with primers Ms1-6R. B) Pattern in other insects (Ayme-Southgate et al 2011). The exons used in the shortest variants are shaded in gray. Intron sizes are not to scale. High quality figures are available online.

The long form of the PEVK segment in all lepidopteran genes was most similar in length and PEVK content to the *T. castaneum* sequence. In contrast, the PEVK content of the short form, but not its length, was closer to the short variant of *D. melanogaster*. The short isoform in *M. sexta* also had a lower P, E, V, and K content than the longest isoform; this was in sharp contrast to the situation in all other studied insects, where the P, E, V, and K content of the short variant can be as high as 100 percent (see Table 2).

The presence and specificity of alternative isoforms in the PEVK-NTCS-1 region were ascertained by performing RT-PCR amplification using RNAs extracted from several body parts, as well as from isolated flight muscles from *M. sexta* (see Materials and Methods for details). Primers were designed from exons

Table 2. Characteristics of the PEVK-NTCS-1 regions in Lepidoptera as compared to other insects. *L.pul: Libellula pulchella, A.mel: Apis mellifera; T.cas: Tribolium castaneum; D.mel: Drosophila melanogaster; M.sex: Manduca sexta; B.mor: Bombyx mori;* and *D.ple: Danaus plexippus.*

Characteristics of the PEVK-NTCS1 domains in Lepidoptera

	L. pul		*A. mel*		*T. cas*		*D. mel*		*M. sex*		*B. mor*	*D. ple*
Length	long	short	long	short	long	short	long	short	long	short		
entire region	332	130	542	39	456	93	655	131	489	317	494	483
PEVK	194	31	422	6	339	44	530	75	377	205	381	374
NTCS1	138	99	120	33	117	49	125	56	112	112	113	109
% P.E.V.K												
in PEVK	68.0	100	62.8	na	60.2	100	41.9	64.0	60.0	53.0	58.0	58.0
in NTCS1	42.7	39.4	38.3	33.3	38.5	34.7	42.0	28.1	46.0	46.0	45.0	43.1
# PEVK exons	na		23	2	16	4	15	4	13	11	12	12

flanking and internal to the PEVK-NTCS-1 segments.

The variant composition was qualitatively identical across muscle types irrespective of the primer pair used for the amplification. As presented in Figure 5A in the Ms1-6R reaction, the isoforms A and B (1,373 and 1,321 bp respectively; see Figure 4) were present in all muscles, and the C/D isoforms (926 and 893 bp respectively) were present in thorax and flight samples, as well as faintly in leg and head samples. The Ms1-6R primer set also faintly amplified two products around 500 bp, which were present in all muscle types. These products were sequenced and correspond to *M. sexta* hemocytin gene as identified by a BLAST homology search (Kotani et al. 1995; Tanaka et al. 2008). The reason for this reproducible amplification is unknown.

The Ms1-6R primer set amplified the entire PEVK-NTCS-1 region. Because short PCR products are favored in PCR amplification reactions, the absence of projectin products shorter than 800 bp in the Ms1-6R reaction was unlikely due to the low abundance of any short isoform. For the same reason, relative contribution of short and longer PEVK variants could not be ascertained completely from these data, even though isoforms A/B seemed to be the most abundant variants in both head and leg RNA samples.

Isoforms A, B, and C were also detected in all muscle types using primers Ms1-2R (Figure 5B). The shortest 480 bp (corresponding to isoform D) product in the Ms1-2R reaction was not detectable in the leg RNA sample. This product only differed from isoform C (corresponding to the 500 bp product) by the exclusion of exon 6 (see Figure 4), which was only 33 bp in length. So, isoform D could be considered a flight muscle-specific isoform.

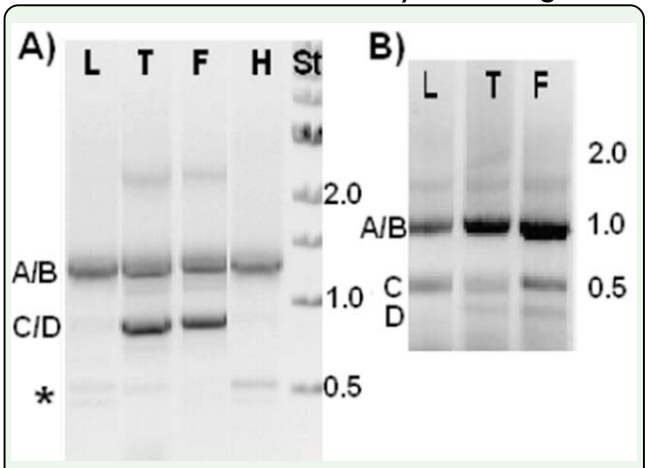

Figure 5. The PEVK variants are identical in all muscle types in *Manduca sexta*. Products of RT-PCR reactions were analyzed by gel electrophoresis. A): primers Ms1-6R and B): primers Ms1-2R (see Figure 4). The muscle types are as follows: L: leg, T: whole thorax, F: flight muscle, and H: head. St: 1 kb ladder from New England Biolabs (www.neb.com). Sizes are indicated on the right. A/B and C/D refer to the isoforms depicted in Figure 4. * is a contamination product corresponding to *M. sexta* hemocytin gene as identified by sequencing and BLAST homology search. High quality figures are available online.

Discussion

In this study, further evidence for the organization of the NH_2-terminal region of projectin into two tracts of 8 and 6 Ig domains separated by a unique sequence is provided. This unique region of the protein can be further divided into two segments. One segment showed little to no sequence conservation *per se*, but displayed biased amino acid content with predominantly P, E, V, and K residues and is considered the "true" PEVK region. In contrast, the second segment, found just before the second stretch of six Ig domains, was a highly conserved sequence of 112–138 residues in length and has been named NTCS-1. The importance of this region for positioning over the length of the sarcomere and/or post-translational modifications is unknown at this time. BLASTp search of the reference protein database yielded no significant homology other than projectin proteins.

The potential for alternative splicing in *M. sexta* muscles is limited, with only 2 alternatively spliced exons identified and 4 possible variant combinations. This is in contrast to 11 exons in *D. melanogaster* and *T. castaneum* and 21 in *A. mellifera*. In *M. sexta*, the length of the shortest PEVK variant (isoform D) represented 54% of the longest variant. In contrast, the same ratio varied from approximately 15% in dragonflies, *T. castaneum*, and *D. melanogaster* to less than 2% in *A. mellifera*. Therefore, the shortest isoform was not considerably shorter than the long isoform in the flight muscle of *M. sexta*. The shortest isoform (D) in *M. sexta* also had a lower P, E, V, and K content than the longest isoform (A); this was in sharp contrast to the situation in all other studied insects, where the short variant can be as high as 100% P, E, V, and K (see Table 2).

Isoform D was the only splice variant specific for flight muscle, and it differed from isoform C by the exclusion of only 11 amino acids encoded by exon #6. So, even though isoform D could be considered a flight muscle-specific isoform, it was not very different from the other isoforms, which were present in all muscle types. This is in contrast with previous observations carried out in other insects, namely that the flight-muscle-specific isoform is strikingly shorter than any other PEVK variants (Ayme-Southgate et al. 2004, 2011).

In vertebrates, the sarcomeric passive tension of striated muscles have been correlated with the size and composition of titin's extensible regions, in particular its PEVK segments (for example Cazorla et al. 2000; Freiburg et al. 2000; Trombitas et al. 2000; Granzier and Labeit 2005; Granzier et al. 2007). In the titin model, for a given sarcomere length, shorter PEVK segments lead to a high resting tension, whereas a longer extensible region results in a lower force (reviewed in Granzier and Labeit 2005).

In derived insects with asynchronous flight muscle, the C filaments composed of kettin/sallimus and projectin have been shown to be a source of the myofibrillar stiffness in flight muscles (Moore et al. 1999; Bullard et al. 2000; Hakeda et al. 2000; Kulke et al. 2001; Bullard et al. 2006). In these indirect flight muscles, according to the titin model, projectin molecules with a short PEVK region would contribute to their elevated stiffness. On the other hand, it is generally proposed that synchronous flight muscles, such as those of lepidopteran, have higher muscle strain and lower passive tension (reviewed in Dudley 2000). Additional studies of other insects with synchronous flight muscles will be required before a precise correlation can be established, but the current study leads us to propose that, by analogy to the spring model described for titin, long PEVK variants will be associated with muscles with low passive stiffness, higher strain, and synchronous physiology, whereas a short PEVK sequence would contribute to high myofibrillar passive stiffness.

Acknowledgements

This research was supported by the College of Charleston faculty and undergraduate research grants, as well as by a grant to the College of Charleston from the HHMI Undergraduate Science Education Program for L. Turner.

Bibliography

Altschul SF, Gish W, Miller W, Meyers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.

Ayme-Southgate A, Saide J, Southgate R, Bounaix C, Camarato A, Patel S, Wussler C. 2005. In indirect flight muscles *Drosophila* projectin has a short PEVK domain, and its NH_2-terminus is embedded at the Z-band. *Journal of Muscle Research and Cell Motility* 26: 467–477.

Ayme-Southgate A, Southgate RJ, Philipp RA, Sotka AA, Kramp C. 2008. The myofibrillar protein, projectin, is highly conserved across insect evolution except for its PEVK domain. *Journal of Molecular Evolution* 67(6): 653–669. DOI: 10.1007/s00239-008-9177-2.

Ayme-Southgate A, Philipp RA, Southgate RJ. 2011. The projectin PEVK domain, splicing variants and domain structure in basal and derived insects. *Journal of Insect Molecular Biology* 20(3): 347–356. DOI: 10.1111/j.1365-2583.2011.01069.x. PubMed PMID: 21349121.

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

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

Bullard B, Goulding D, Ferguson C, Leonard K. 2000. Links in the chain: the contribution of kettin to the elasticity of insect muscles. In: Pollack GH, Granzier H, Editors. *Proceedings: Elastic Filaments of the Cell*. pp. 207–220. Kluwer Academic/Plenum Publishers.

Bullard B, Linke W, Leonard K. 2002. Varieties of elastic protein in invertebrate muscles. *Journal of Muscle Research and Cell Motility* 23: 435–447.

Bullard B, Burkart C, Labeit S, Leonard K. 2005. The function of elastic proteins in the oscillatory contractions of insect flight muscle. *Journal of Muscle Research and Cell Motility* 26: 479–485.

Bullard B, Garcia T, Benes V, Leake M, Linke W, Oberhauser A. 2006 The molecular elasticity of the insect flight muscle proteins projectin and kettin. *Proceeding of the National Academy of Science USA* 103: 4451–4456.

Burkart C, Qiu F, Brendel S, Benes V, Hååg P, Labeit S, Leonard K, 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): 953–969.

Cazorla O, Freiburg A, Helmes M, Centner T, McNabb M, Wu Y, Trombitás K, Labeit S, Granzier H. 2000. Differential expression of cardiac titin isoforms and modulation of cellular stiffness. *Circulation Research* 86: 59–67.

Dudley R. 2000. *The Biomechanics of Insect Flight: Form, Function, Evolution*. Princeton University Press.

Freiburg A, Trombitás K, Hell W, Cazorla O, Fougerousse F, Centner T, Kolmerer B, Witt C, Beckmann JS, Gregorio CC, Granzier H, Labeit S. 2000. Series of exon-skipping events in the elastic spring region of titin as the structural basis for myofibrillar elastic diversity. *Circulation Research* 86: 1114–1121.

Granzier HL, Wang K. 1993. Passive tension and stiffness of vertebrate skeletal and insect flight muscles: the contribution of weak cross-bridges and elastic filaments. *Biophysical Journal* 65: 2141–2159.

Granzier HL, Labeit S. 2002. Cardiac Titin: An Adjustable Multi-Functional Spring. *Journal of Physiology* 541: 335–342.

Granzier HL, Labeit S. 2005. Titin and its associated proteins: the third myofilament system of the sarcomere. *Advances in Protein Chemistry* 71: 89–119.

Granzier H, Radke M, Royal J, Wu Y, Irving TC, Gotthardt M, Labeit S. 2007. Functional genomics of chicken, mouse and human titin supports splice diversity as an important mechanism for regulating biomechanics of striated muscle. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology* 293: 557–567.

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*. *Journal of Cell Biology* 148: 101–114.

International Silkworm Genome Consortium. 2008. The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. *Insect Biochemistry and Molecular Biology* 38(12): 1036–1045.

Koike Y, Mita K, Suzuki MG, Maeda S, Abe H, Osoegawa K, deJong PJ, Shimada T. 2003. Genomic sequence of a 320-kb segment of the Z chromosome of *Bombyx mori* containing a kettin ortholog. *Molecular Genetics and Genomics* 269(1): 137–149. PubMed PMID: 12715162.

Kotani E, Yamakawa M, Iwamoto S, Tashiro M, Mori H, Sumida M, Matsubara F, Taniai K, Kadono-Okuda K, Kato Y. 1995. Cloning and expression of the gene of hemocytin, an insect humoral lectin which is homologous with the mammalian von Willebrand factor. *Biochimical and Biophysical Acta* 1260(3): 245–258. PubMed PMID: 7873598.

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

Labeit S, Gautel M, Lakey A, Trinick J. 1992. Towards a molecular understanding of titin. *EMBO Journal* 11: 1711–1716.

Labeit S, Kolmerer B, Linke WA. 1997. The giant protein titin. Emerging roles in physiology and pathophysiology. *Circulation Research* 80: 290–294.

Mita K, Morimyo M, Okano K, Koike Y, Nohata J, Kawasaki H, Kadono-Okuda K, Yamamoto K, Suzuki MG, Shimada T, Goldsmith MR, Maeda S. 2003. The construction of an EST database for *Bombyx mori* and its application. *Proceeding of the National Academy of Science USA* 100: 14121–14126.

Mita K, Kasahara M, Sasaki S, Nagayasu Y, Yamada T, Kanamori H, Namiki N, Kitagawa M, Yamashita H, Yasukochi Y, Kadono-Okuda K, Yamamoto K, Ajimura M, Ravikumar G, Shimomura M, Nagamura Y, Shin-I T, Abe H, Shimada T, Morishita S, Sasaki T. 2004. The genome sequence of silkworm, *Bombyx mori*. *DNA Research* 11: 27–35.

Moore JR, Vigoreaux JO, Maughan DW. 1999. The *Drosophila* projectin mutant, bent^{D}, has reduced stretch activation and altered flight muscle kinetics. *Journal of Muscle Research and Cell Motility* 20: 797–806.

Oshino T, Shimamura J, Fukuzawa A, Maruyama K, Kimura S. 2003. The entire cDNA sequences of projectin isoforms of crayfish claw closer and flexor muscles and their localization. *Journal of Muscle Research and Cell Motility* 24(7): 431–438.

Peckham M, Molloy JE, Sparrow JC, White DCS. 1990. Physiological properties of the dorsal longitudinal flight muscle and the tergal depressor of the trochanter muscle of *Drosophila melanogaster*. *Journal of Muscle Research and Cell Motility* 11: 125–136.

Peckham M, Cripps R, White D, Bullard B. 1992. Mechanics and protein content of insect flight muscles. *Journal of Experimental Biology* 168: 57–76.

Pringle JWS. 1977. The mechanical characteristics of insect fibrillar muscle. In: Tregear RT, Editor: *Insect flight muscles*. pp. 177-196. Elsevier/North Holland Biomedical Press.

Pringle JWS. 1981. The Bidder lecture, 1980: The evolution of fibrillar muscle in insects. *Journal of Experimental Biology* 94: 1–14.

Southgate R, Ayme-Southgate A. 2001. *Drosophila* projectin contains a spring-like PEVK region which is alternatively spliced. *Journal of Molecular Biology* 313: 1037–1045.

Tanaka H, Ishibashi J, Fujita K, Nakajima Y, Sagisaka A, Tomimoto K, Suzuki N, Yoshiyama M, Kaneko Y, Sunagawa T, Yamaji K, Asaoka A, Mita K, Yamakawa M. 2008. A genome-wide analysis of genes and gene families involved in innate immunity of *Bombyx mori*. *Insect Biochemistry and Molecular Biology* 38(12): 1087–1110. PMID: 18835443.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673–4680.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25: 4876–4882.

Thorson J, White DC. 1983. Role of cross-bridge distortion in the small-signal mechanical dynamics of insect and rabbit striated muscle. *Journal of Physiology* 343: 59–84. PubMed PMID: 6685767.

Trombitas K. 2000. Connecting filaments: a historical prospective. In: Pollack GH, Granzier H, Editors. *Proceedings: Elastic Filaments of the Cell*. pp. 1–23. Kluwer Academic/Plenum Publishers.

Trombitas K, Freiburg A, Greaser M, Labeit S, Granzier H. 2000. From connecting filaments to co-expression of titin isoforms. In: Pollack GH, Granzier H, Editors. *Proceedings: Elastic Filaments of the Cell.* pp. 405–418. Kluwer Academic/Plenum Publishers.

Vigoreaux JO, Moore JR, Maughan DW. 2000. Role of the elastic protein projectin in stretch activation and work output of Drosophila flight muscles. In: Pollack GH, Granzier H, Editors. *Proceedings: Elastic Filaments of the Cell.* pp. 237–247. Kluwer Academic/Plenum Publishers.

White DC. 1983. The elasticity of relaxed insect fibrillar flight muscle. *Journal of Physiology* 343: 31–57. PubMed PMID: 6557139; PubMed Central PMCID: PMC1193907.

Zhan S, Merlin C, Boore JL, Reppert SM. 2011. The monarch butterfly genome yields insights into long-distance migration. *Cell* 147(5):1171-85.