2021

Structural Characterization Of The Novel Flightin Domain Wyr And Its Defining Role In The Thick Filament Structure And Mechanics

Lynda Menard
University of Vermont

STRUCTURAL CHARACTERIZATION OF THE NOVEL FLIGHTIN DOMAIN WYR AND ITS DEFINING ROLE IN THE THICK FILAMENT STRUCTURE AND MECHANICS

A Dissertation Presented

by

Lynda Menard

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
Specializing in Biology

January, 2021

Defense Date: October 30, 2020
Dissertation Examination Committee:

Jim O. Vigoreaux, Ph.D., Advisor
Matthew D. Liptak, Ph.D., Chairperson
Bryan A. Ballif, Ph.D.
Ying Wai Lam, Ph.D.
Teresa Ruiz, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

ABSTRACT

The evolutionary success of Insecta has been attributed largely to the development of efficient means of motility: flight powered by muscle architecture harboring a largely conserved yet tunable system of power relay. The indirect flight muscle (IFM) of *Drosophila melanogaster* is a well-studied model for dissection of the structural and mechanical means by which muscle operates and evolves. Striated muscle, conserved throughout Animalia, is demarcated by an ordered array of thick- and thin-filaments prominently composed of the proteins myosin and actin. Flightin (fln) is a myosin binding thick filament protein essential for IFM stability, structure and function. The manner by which fln contacts myosin and relevance of its highly conserved domain (WYR) has not been fully elucidated. This dissertation presents the culmination of an effort to elucidate fln's role in the thick filament and the nature and involvement of the novel WYR domain. Cardiac myosin binding protein-C (cMyBP-C), exclusive to vertebrates, and fln, exclusive to Pancrustacea bind a common site in the light meromyosin (LMM) region of myosin and have been hypothesized to have partially overlapping functions within the thick filament. To evaluate this, IFM sarcomeres and thick filaments from *D. melanogaster* mutant and transgenic strains with and without additional cMyBP-C expression were examined by transmission electron microscopy (TEM) and atomic force microscopy (AFM), respectively. cMyBP-C, like fln, is found to influence sarcomere length and contribute to thick filament flexural rigidity. This suggests a shared influence on thick filament properties though cMyBP-C did not fully rescue the fln^0 phenotype. Adding depth to the fln-LMM relationship, we examined the structure and function of WYR. The structure of WYR, determined by circular dichroism (CD), is mostly aperiodic, with 30% antiparallel β content. A putative model of WYR secondary structure is presented, derived from CD findings and interpreted on the basis of WYR's primary sequence and the potential contributions of its aromatic and polar residue electronic state transitions. Employing both cosedimentation and CD, we find that WYR binds the LMM and induces structural change. The WYR-LMM structure depict the LMM as decreasing in α-helical nature and increasing in coiled-coil character and sedimentation assays demonstrate increased prevalence of macroscopic assemblies upon the association. Data from a structural study of the waterbug IFM thick filament was processed to reveal fln association to regions depicting coiled-coil unwinding. The portions of the LMM interfacing with fln were associated to the myosin sequence, revealing specific amino acids over which fln is in close proximity. We identify five interfaces, one of which is heptad mapped and reveals an LMM binding region shared between fln and cMyBP-C. Given the importance of fln to IFM function and the conservation of the WYR domain through Pancrustacea, the convergent effects of fln and cMyBP-C along with LMM structural change induced by WYR presents a positional and structural basis over which the thick filament experiences context-dependent tuning. Our findings depict fln as a cinch connecting multiple myosin dimers via the LMM, and support its intimate involvement in thick filament assembly. This work describes WYR on a multiscale, considering the nanoscopic mechanisms that underpin macroscopic biological phenomena. WYR is an important agent by which structural and mechanical adaptations are incorporated into the IFM hierarchy, relevant to the rise of flight within Insecta. Further dissection of WYR's function and relationship to the LMM should provide insight pertinent to the scaling of mechanical processes by structural design and have bearing in studies beyond the IFM and insect adaptation.

CITATIONS

Material from this dissertation has been published in the following form:

Menard L.M., Nyland L.R., Vigoreaux J.O.. The Structural and Biomechanical Properties of Insect Thick Filaments Expressing Flightin and Cardiac Myosin Binding Protein-C. Microscopy and Microanalysis. 2013;19:80-1

ACKNOWLEDGEMENTS

I am greatly appreciative to my advisor Dr. Jim O. Vigoreaux for his contributions and pivotal support throughout my growth as a scientist and towards the completion of my PhD work. I am also grateful to my committee members, Dr. Bryan A. Ballif, Dr. Ying Wai Lam, Dr. Matthew D. Liptak, and Dr. Teresa Ruiz for providing helpful critique and sharing their pertinent scientific knowledge and expertise. Additional thanks to Dr. David Maughan, his optimism and encouragement was much appreciated. Past and present members of the Vigoreaux lab, including Dr. Andrew Mead, Ravi Nagori, and Emily Price, provided both helpful critique and useful insight and ideas. I would also like to acknowledge the substantial time and effort of contributing undergraduate students including, but not limited to, Julia Boettger, Campbell Dickson, Nathan Gasek, and Neil Wood. Neil Wood was especially instrumental to the work within this Dissertation.

Thank you to my friends and family for their patience, consideration, encouragement and understanding. I have immense appreciation for my friends whom did not allow my social sparsity to deteriorate our connection, offered their helpful insight, included me when able and held an open mind. Eternal thanks to Teagan and Esper, without whom I could not have accomplished this. Organizations such as the Vermont Biosciences Alliance and ASCB are recognized for adding diversification to my PhD experience and encouraging inclusivity.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

The aim of this review is to describe how muscle function, in the context of striated muscle hierarchy, is sensitively tuned within Arthropoda by means of the small protein flightin, where flightin fits in to the innovation of stretch activation, and what next steps are needed to bring this picture a step further. Understanding of the thick filament near-crystalline ultrastructure and non-linear viscoelastic properties has reached a new level in the past few years [1-5]. This raises new opportunities to elucidate the complex coordinated structural and mechanical triggers, stops, and shifts necessary for unabated success of stretch-activated muscle function of which flightin plays a major role.

This review will also relate flightin to the initial development of the sarcomere, its capacity in regards to understanding mammalian muscle pathologies and its changing/adaptable behavior from an ecological perspective. There has been extensive research on the cohesive components of the sarcomere and reviews exist for the evaluation of sarcomere function as a whole [6-9], the role of the Z-disc [10] and M-band [11], along with some reviews much more specifically targeting the thick [12, 13] and thin filament [14, 15], or major non-myosin constituents such as titin/sallimus [16-18] and obscurin [19]. There are a number of other good reviews on muscle development, regeneration and maintenance, including the mechanical properties involved [20-25], and a few recent publications on the use of invertebrates, such as Drosophila, as models for muscle-based diseases [26-29].

Universally, muscle is defined, in part, by the presence of myosin-containing thick filaments that are necessary for the functionality of the tissue as a whole. The thick filaments and their interacting components are varied between muscle types and between organisms as a way of tuning the biomechanical properties to suit the organisms' functional needs. The sarcomere is the basic conserved structural unit of contraction whereupon interdigitating myosin thick filaments and actin thin filaments engage with accessory proteins to fulfill the dynamic needs of the muscle as a whole.

During muscle maturation, the myofibrils develop characteristic sarcomeric units whose structure is dictated by actin and myosin content as well as early-expressing accessory proteins. The character and quantity of the accessory proteins change along with muscle development and continue to play a role in modification, and maintenance, of contractile properties. While there is a wide familiarity with the basic components of muscle tissue, details regarding integral accessory proteins are still coming to light.

Invertebrates make up more than 95% of animal species with insects representing 80% of all animal species and 60% of all known species of living organisms [30, 31]. As more than half of all characterized species are within Insecta, there are many hypotheses as to why such richness in diversity developed, many of which implicate the development of muscle-driven flight as a primary driver [32-34]. The innovation of flight enhanced the

motile properties of these organisms. This permitted increased accessibility to food sources, dispersal, predator evasion, and reproductive specialization.

Insect flight muscles are highly variable, multi-functional and possess the highest power outputs among all kinds of muscle. They are capable of functioning at very high contraction rates (>1000Hz) and represent a substantial metabolic investment by the organism. This is due to the energetic demands of flight itself but the flight muscle also serves various additional functions, including playing a major role in courtship [35, 36]. The highly-ordered indirect flight muscle (IFM) of *Drosophila melanogaster* has been instrumental in elucidating the molecular basis of these functions and has proven an excellent venue for examining flight muscle ultrastructure and molecular composition [37].

THE SARCOMERE – AN EVOLUTIONARILY CONSERVED DESIGN

Structural Organization

The basic sarcomere structure is conserved in striated muscle across both vertebrates and invertebrates and all contain strategically arranged interdigitating myosin thick filaments and actin thin filaments (**Fig 1-1**). The full sarcomere is viewed from Z-disc to Z-disc in which the Z-disc is an area of high density containing alpha-actinin and other proteins that anchor the thin filaments. Immediately adjacent to the Z-discs are regions known as “I bands” in which no myosin thick filaments are present. Towards the center of the

sarcomere, the thick and thin filaments overlap in a region denoted as the “A band”. In the middle of the “A band” is the M-line which is an anchoring point for the thick filaments. Flanking the M-line, is the H-zone where thick filaments exist without overlap with the thin filaments.

The repeating sarcomeric units, all in parallel, along a myofibril permits the amplification of force generation along the filament axis that ultimately correlates to the whole function of the muscle fiber. The sliding filament theory, in which the simultaneous hydrolysis of ATP on the myosin motor along with regulated exposure of the actin binding site, represents the conserved contraction mechanism among all striated muscle. Sarcomere structure is tightly linked to muscle function; defects in proteins dictating sarcomere formation and stability are associated with many disease states that influence full body function [38-40].

Sarcomere structure is dictated largely by myosin and actin associating proteins that give rise to the classical striated pattern. Spanning the sarcomere is titin or other members of the titin family: D-titin/sallimus, kettin, and projectin. In the M line, myomesin, M-protein, obscurin and other proteins crosslink antiparallel myosin rods, interact with each other and are involved in enzyme recruitment [41, 42]. The Z-disc is formed by interactions between alpha-actinin, Sallimus, and Zasp52 [43-47]. While not all of these proteins retain the exact same function across all muscle, modulators of the sarcomere structure such as LASP and LASP-like proteins [48, 49], actin capping proteins [50, 51],

and giant titin-like proteins [52, 53] are functionally conserved in all vertebrate and invertebrate striated muscle.

The sarcomere through evolution

Evidence of sarcomeres have been found beyond Metazoa. Cnidaria is believed to be the earliest phylum exhibiting muscle cells with identifiable sarcomeres. The existence of muscle cells is one of the morphological traits that unite Ctenophores with Tripoblasts and their presence in Cnidaria has supported Cnidaria as a sister phylum to Tripoblasts [54]. The earliest fossil believed to contain muscle is the Cnidarian *Haootia quadriformis* [55]. While most muscle within Cnidaria is smooth, striated muscle has been identified in hydrozoan medusa [54, 56]. Early Ctenophore muscles lacked an H band, suggesting that the cells underwent only one contraction due to an inability to then relax [57]. There is still dispute regarding muscle origins for Metazoa and there is growing evidence of independent origins for the muscles existent among Cnidaria, Cteniphora, and Tripoblastica [54, 58].

Sarcomere components have been conserved throughout life history. The characteristic myosin heavy chain motor proteins existed in unicellular organisms, pre-muscle. In these organisms, a duplication of the myosin heavy chain gene resulted in formation of the striated isoform (ST-MHC) and smooth/non-muscle isoform (SM-MHC). The Bilaterian striated isoform was ultimately maintained in the muscles of protostomes [59]. Striated

muscles in Bilaterians are distinctly characterized by the presence of a troponin complex (I,C & T) [59, 60].

Molecular data from evolutionary studies serve to emphasize the importance of specific protein components within the sarcomere. Orthologs to actomyosin machinery - including actin, myosin II heavy chain, myosin light chain, tropomyosin and calmodulin - all predating muscle - have been found to be conserved among metazoans, likely having been developed in Holozoa [61, 62]. Within Bilateria, conserved Z-disc components include alpha-actinin, Lim and Zasp proteins and giant Titin-related proteins- characterized by Immunoglobulin/Fibronectin type III super repeats [62].

Primitive muscle cells contained an epithelial component but as evolution gave rise to more complex organisms, ‘true muscle cells’ appeared and then became further tuned to more elaborate functions through compartmentalization. Regulation of electro-chemical signals through compartmentalization and tuned protein specificities permits greater precision in communication, giving rise to a larger functional repertoire. Compartmentalization in the muscle ranges from traditional membrane-bound organelles, such as the SR and mitochondria, to dynamic protein-level complexes that can transiently utilize other structural components, such as the thick and thin filaments [63-66].

The context of muscle varies between vertebrates and invertebrates giving rise to differences in the proteins that are required for sarcomere structure and function. Proteins

present in both, but only necessary for sarcomere structure in invertebrates or possessing altered function in invertebrates include such proteins as obscurin [67, 68] and Fhos [69]. Novel proteins present only in invertebrates that have been found to be integral include paramyosin, miniparamyosin, projectin, kettin [52], and flightin [70].

THICK FILAMENT STRUCTURE – VARIATIONS ON A THEME

Thick filaments are organized with tightly packed myosin dimers oriented such that the globular heads are extended outwards. Thick filaments of vertebrate striated muscles are organized such that myosin heads extend in sets of three of tripartite sections and are spaced apart by 143 Å; these regions are known as crowns [71]. This axial spacing lengthens to ~145 Å upon muscle activation, before tension development. The rod diameter is 2 nm and the length of the myosin rod is ~1600 Å, extending through 11 crowns. Invertebrates possess tightly packed myosin dimers that have a constant crown spacing of ~145 Å but great variability in the number of myosin heads associated per crown [21, 72]. Vertebrates tend to have shorter thick filaments of ~1.6 μM in length [73-75] compared to Invertebrates which trend towards longer thick filament lengths along with increasing paramyosin content and thick filament diameter [21].

The thick filament, resultant from the highly conserved MHC within Metazoa, coordinates structure and function along with complexes – both dynamic and persistent- that tune the sarcomere. Null mutations in the muscle myosin and actin genes in

Drosophila have demonstrated that myosin and actin stoichiometry is important for formation of sarcomere width and lengths and double heterozygotes still do not exhibit the full WT phenotype [76]. While these mutations causing myosin or actin depletion all affect flight muscle function, increased production of myosin and actin do not [77, 78]; the sarcomere remains unaffected. This strongly supports stability and assembly of both actin and myosin filaments being dependent on the available accessory proteins produced only at levels sufficient for the assembly of 'normal' sarcomeres. Indeed, importance of the stoichiometry of affiliated proteins has been demonstrated [79-81].

Thick filaments are modulated firstly by expression of various myosin stage- and tissue-specific isoforms dependent on the muscle type and organism. While many organisms have a MHC multigene family from which isoforms are derived [82], Drosophila has one MHC gene, with isoforms coming about by alternative RNA splicing [83-85]. Replacement of the IFM isoform with the embryonic isoform does not change the ultrastructure of the sarcomere in the IFM suggesting that those differences are not involved in characteristics that differentiate sarcomere assembly between the two tissues (e.g., sarcomere length, myofilament packing) [86]. The embryonic isoform in the context of the IFM, however, does show lessened stability that is suspected to be due to improper interactions with the thin filament. From this, it appears that the developmental differences are brought on by myofilament binding partners rather than intrinsic distinction between MHC isoforms.

The myosin thick filament plays a major role in defining the properties of the sarcomere and a number of MHC point mutations have been found to result in major myopathies [87] which have been modelled in Drosophila [28, 88-90]. *Mhc*^{9}, in which a lysine is substituted for a glutamic acid at aa482, was the first missense mutation identified in Drosophila MHC gene, which resulted in failure of MHC to accumulate in the IFM specifically [91]. *Mhc*^{10} similarly prevents myosin accumulation, specifically in the jump and indirect flight muscles by preventing use of exon 15a [92]. Use of headless myosin mutants [93-96] and myosin rod specific mutants [97, 98] has been instrumental in differentiation of phenotypes generated by actomyosin engagement versus effects stemming from the light meromyosin (LMM) region that myosin head motors rely upon as a highly conserved, indispensable support. Such mutations have allowed study of myosin rod binding proteins in both the presence and absence of contractile forces.

The myosin rod is a highly conserved archetype of coiled-coiled structure that directly connects to thick filament function [99]. The myosin coiled-coil, produced by supercoiling of adjacent α helical myosin rods, exhibits a classic heptad repeat structure over which positions 1, “a”, and 4, “d”, form a hydrophobic interface along which the helical axis runs. The myosin coiled-coil dimer is further separated into 28-mer zones of alternating positive and negative charge [100]. This tertiary structure is known for its mechanical rigidity and capability for force sensing and transmission [101]. Deviations from the classical heptad repeat, known as ‘staggers’, ‘stutters’, and ‘skips’, result in alterations to coiled-coil pitch and radius. There are four skip residues conserved across

the striated muscle myosin rod of both vertebrates and invertebrates [100, 102]. These local distortions, while not negating the overall coiled-coil structure, introduce areas of flexibility that are involved with thick filament formation [100] and binding of accessory proteins to the myosin rod [103].

Changes in the engagement of the myosin rod by thick filament accessory proteins near skip residues play a major role in disease states impacting the sarcomere and scaling to whole muscle dysfunction. Residues identified as important for mammalian cardiac myosin binding protein-C (cMyBP-C) binding to the rod include aa1554 and aa1581 [104], which are very close to the third skip residue (E1582). The mutation E1554K in Drosophila also prevents accumulation of flightin, resulting in aberrant structure within the sarcomere and a flightless phenotype [97, 105]. While very different proteins structurally, flightin and cMyBP-C share some functional similarities [106] and the altered coiled-coil properties around skip 3 may be connected to their binding capacity. Cryo-EM studies have identified areas of non-myosin density within the myosin rod in *Lethocerus* in connection with unwound portions of the myosin rod [2], supporting the possibility of such associations.

Flightin is a particularly interesting thick filament protein as it is found in the asynchronous indirect flight muscle which is depended on by 75% of known insect species [107], recognized to possess unusually fast actomyosin kinetics [108] and an enhanced stretch activation response. Drosophila has a highly crystalline myofilament

lattice throughout the IFM myofibril to accommodate enhanced oscillatory work and the high power output per gram body weight needed for flight [109]. It is within the myosin coiled-coils of the thick filament that flightin binds and modulates muscle function.

MOLECULAR STRUCTURE OF FLIGHTIN

Flightin is a thick filament accessory protein that has been identified in members of Pancrustacea and demonstrated to be required for flight in *D. melanogaster*. It is a ~20 kDa myosin rod binding protein with a pI of 5.2-5.3, first found in the IFM of *D. melanogaster* [110]. It has been found to localize to the A band of the sarcomere [70, 110] and associate with the LMM of the thick filament [70, 105, 111]. Since its discovery, flightin has undergone characterization of its binding profile and function within the thick filament. As additional information regarding flightin's structure and influence on muscle become available, it is becoming possible to further hypothesize on its influence beyond the LMM, with other components of the thick filament, and its evolutionary profile.

While flightin's 182-aa sequence does not reveal any previously known, characterized, protein domains, it can be divided into three regions of differing evolutionary conservation. Alignment of sequences from 12 Drosophila species [112], show the 65aa NH_2-terminal region is poorly conserved with less than 15% identity and the COOH-terminal region from aa137 to aa182 (*D. melanogaster* numbering) possesses

intermediate conservation with 60% identity. The highest conservation is within the midsection of the protein from aa66 to aa136 with 93% identity. An examination of flightin among members of Pancrustacea show this region to possess an invariant tryptophan at aa85 with two tyrosines at aa93 and aa104 and arginine at aa131. Hence, the highly conserved 52 aa middle region of flightin is referred to as "WYR" [113]. R87 and P123 are also invariant, while Y103 and E130 are invariant with the exception of one species. Overall, 23% of WYR is conserved to some degree throughout Pancrustacea, 48% in insects. Given that these three regions are undergoing separate levels of selection, it has been hypothesized that they also represent different functional regions of flightin.

Flightin's secondary structure has not been elucidated. It has been predicted to be predominantly in a random coil conformation with two regions of α-helical character between residues 88-114 and 149-180 [114]. It also has low predicted hydrophobicity that is typical of natively unfolded proteins. Secondary structure prediction programs predict >50% α helical content within WYR but do not take into account aromatic character which represent the residues of highest conservation. There is consensus between prediction programs (Jpred, RaptorX2, Phyre2, and i-tasser) that the region of Y104-K114 in the WYR sequence is most likely to be helical when the influence of strictly nonpolar character at the aromatic positions is removed. This would represent ~19% of the WYR sequence. The portion of the sequence that precedes this contains alternating tyrosines which would be unfavorable for helical structure.

Eleven isoelectric variants have been identified in the native flightin protein in the adult fruit fly. The number of detectable phospho-variants increase during development with the final compliment being identified 2-3 hours post eclosion [111]. Initial MALDI-TOF analysis identified Ser 139, Ser 141, Ser 145, Thr 158 and Ser 162 as potential phosphorylation sites [115] though more sites in the N-terminus have been since identified by LC-MS-MS [116]. It has been estimated that mature adults possess a 1:1 ratio of phosphorylated to nonphosphorylated flightin. There is a shift towards the more acidic variants during flight, however, suggesting that this ratio shifts in the favor of more phosphorylated variants.

Characterization

Flies heterozygous for the flightin gene (*Df(3L)fln*^{1}) have ~20% reduced flightin accumulation associated with a 26% decrease in myofibril diameter with flightin most deplete from the myofibril periphery. While *Df(3L)fln*^{1} was not found to interfere with myofibril assembly in the pupal stages, as myofibril transverse sections and sarcomere structure were comparable to that of WT, peripheral myofilaments are not well integrated into the lattice in adult IFM and are removed during typical skinning protocols. Frequency required to achieve maximal power was increased for *Df(3L)fln*^{1} compared to WT although maximum power output for the fibers was unaltered. Accelerated actomyosin kinetics are observed and the flies are flight capable. Although myofibril

packing and sarcomere length is unaffected, some sarcomeres experience hypercontraction or extension in the adult.

Transgenic flies completely lacking flightin, fln^0, exhibit a much more dramatic impact on the IFM. The flightin null strain of *D. melanogaster* holds their wings ventrolaterally as opposed to dorsally and shows complete loss of flight ability accompanied by major structural defects down to the sarcomere level [70]. The dorsal longitudinal muscles (DLMs) go from long and wavy in pupa to shortened and torn in the adult. There are fewer thick filaments across the myofibril diameter indicating decreased myofilament packing. Sarcomeres are ~25% longer and more variable in pupa as are the thick filaments. Sarcomeres feature "triple" M-lines, as two transverse electron-dense stripes appear flanking the M-line. The sarcomere quickly degenerates in the adult with fragmentation of the Z-discs, a broad obscuring of the M-line and myofilaments ultimately falling apart with the entire IFM exhibiting a hypercontractile phenotype. Site-specific proteolytic cleavage at the hinge of the LMM is observed in the fln^0 and akin to that observed in the Mhc^{13} mutant that does not accumulate flightin.

In addition to the structural detriment evident in the flightin null, a number of mechanical deficiencies have also been identified. The structural degradation associated with hypercontractile phenotype of fln^0 is abolished in myosin motor domain mutants that cannot engage actomyosin force production [96]. This indicates that IFM breakdown in the fln^0 is partially resultant from the inability to maintain fiber structure and organization

during such activity. On the fiber level, stretch to 4-5% original length in WT flies is sufficient to produce a resting stress of ~1 kN/m^2 whereas in *fln*^{0}, the required stretch is approximately twice this [117]. Thick filaments from *fln*^{0} have decreased stiffness [106]. Axial stiffness is expected to decrease with increased thick filament length, and this is evident in the *fln*^{0} as stiffness is also decreased along with the increase in sarcomere length.

Role of the C-terminal region

To investigate the role of the C-terminal region of flightin, a transgenic line was generated lacking the C-terminal 44 amino acids (*fln*$^{\Delta C44}$) and structural components of the IFM and flight capacity were evaluated [112, 118]. The produced protein had an estimated pI of 4.9 and size of 16.23 kDa. The truncated flightin continued to incorporate within the thick filament, indicating that the C-terminus was not required for the protein's binding capacity. A reduced phosphorylation profile was noted, with 8 isoelectric variants identified rather than the 11 found in native flightin of adult flies. The flight ability of the *fln*$^{\Delta C44}$ line was abolished although *fln*$^{\Delta C44}$ expressed with endogenous full length flightin was not found to interfere with flight ability.

Absence of the flightin C-terminal results in distinct abnormalities in sarcomere structure. As seen by TEM, the M-line is much weaker in intensity or entirely absent while the Z-disc remains but with occasional irregularities along its length in which the density

associated with the Z-disc extends variably into the I band. Sarcomere length was ~9% shorter compared to the flightin rescue line (*fln*$^{+}$). Thick filaments are also significantly shorter and more compliant although flexural rigidity affiliated with bending propensity is inconsistent along the length of the thick filament in the *fln*$^{\Delta C44}$ [118]. Myofibril cross-sections reveal a disordered filament lattice associated with *fln*$^{\Delta C44}$ in which there was a ~1% decrease in thick filament spacing. X-ray diffraction shows a repositioning of the myosin heads away from the backbone and an overall decreased thick filament: thin filament mass in the myofibril.

Cross-bridge cycling kinetics are also influenced by the absence of the flightin C-terminal. Relaxed steady-state isometric tension of muscle fibers increase and the stretch required to achieve maximal oscillatory work is doubled. Viscoelastic mechanical properties for IFM fibers under relaxed and rigor conditions and duration of crossbridge attachment were not found to be different from *fln*$^{+}$. At maximal calcium activation (pCa 5.0), fibers have an increased elastic moduli and decreased viscous moduli along with decreased frequencies associated with maximum power and work output indicative of decreased cross-bridge recruitment during active contraction. Both rate of cross bridge attachment and number of strongly bound cross bridges are compromised in *fln*$^{\Delta C44}$. Although the IFM is non-functional in the *fln*$^{\Delta C44}$, the integrity of the IFM is improved compared to the flightin null (*fln*^{0}) with fewer disrupted sarcomeres. The C-terminal is considered to be predominantly responsible for promoting myofilament lattice order and sarcomeric structure, including thick filament length, with the compromising of these

structures concordant with altered cross-bridge kinetics, decreased frequency of IFM operations and reduced maximal oscillatory work and power. While flightin successfully incorporates in the myofibril, the C-terminal is a necessary component to maintain the structural and mechanical properties necessary for IFM function.

Role of the N-terminal region

The N-terminal region of flightin is described as not being essential for IFM function but important for the tuning of male courtship song and under different selection pressure than the other two flightin segments. A transgenic model lacking the N-terminal 62 aa of flightin (*fln*$^{\Delta N62}$) has a decreased capacity for flight by ~1/3 and oscillatory power output is decreased by 57% [119]. Courtship song exhibits higher sine song frequency and longer pulse song with longer inter-pulse intervals. This impacts female mating choice greatly (92% preference for WT males) indicating the involvement of the N-terminus in species-specific tuning of courtship song.

Structural and mechanical properties of the IFM are still impacted in *fln*$^{\Delta N62}$ although flight capability is preserved. Both the elastic and viscous moduli are compromised in relaxed and rigor conditions and a greater number of fibers from *fln*$^{\Delta N62}$ are unable to withstand tension in rigor. The myofilament lattice is less ordered and more compact, containing 11% more thick filaments [119]. Sarcomeres are 13% shorter, lack an evident H-zone and exhibit a narrower M-line. The thick filaments retain the same length as in

the *fln*$^{+}$ but have a markedly decreased stiffness, beyond that observed for *fln*$^{\Delta C44}$ [118]. Unlike *fln*$^{\Delta C44}$ flies, the N-terminal mutant does not exhibit any change to thick filament length nor to wingbeat frequency at maximal power output indicating that, while force transmission is impacted, actomyosin kinetics are unchanged.

Although flight is preserved in *fln*$^{\Delta N62}$ flies, it is not optimal and the altered male song may be due to the decreased dampening of a less resilient, less stiff, myofilament network. The acidic N-terminal region of flightin may function to maintain normal myofilament spacing of the lattice by extending from the thick filament backbone and engaging with the negatively charged actin thin filaments by electrostatic repulsion. The responsibilities of the N-terminus underscore flightin as a good example of the co-evolution of immediate individual tissue function and long-term species-specific survival within a single protein.

POSSIBLE DYNAMICS OF FLIGHTIN INTERACTIONS

Implications of flightin phospho-variants

Eleven isoelectric variants of flightin are present in the adult that include at least nine phosphovariants and are designated N1, N2, and P1-9. When myofibrillogenesis is nearly complete in the late pupal stage of Drosophila, flightin is present in predominantly unphosphorylated variants. The full complement of flightin isovariants is present in young adult flies 5-6 hours post eclosion and are retained throughout life. The most

prominent isovariant accounting for 20-45% of the total in adults, is primarily unmodified flightin, with a pI of 5.2 while most others are increasingly acidic, except for one minor variant with a more basic pI [111]. Phosphorylation is expected to be sequential as an increase in the most acidic variants are associated with a decrease in less acidic variants.

The shift in pI between P1/2 and P3 is much greater than expected for addition of a single phosphate and it is possible that this is due to phosphorylation events taking place, for P3, close to another phosphorylated residue or in a part of the protein that is already acidic. Flightin contains five consensus sites for casein kinase II, cAMP, and cGMP-dependent protein kinase but these would not account for all 9 phosphovariants observed; the kinase responsible is still unknown.

The least conserved of the flightin segments, the N-terminal, is predicted to be the most disordered and contain a cluster of phosphorylation sites which represent the most conserved portion of the N-terminal across Drosophila species [116]. In association with changes in phosphovariants during flight, flightin N-terminal phosphorylation sites S21, T22, T26, S31, S35, S38, S44, T49 were identified by MS-MS separated by 2-DE. C-terminal residues S139 and S141 were identified in the acidic variants that did not show significant change during flight. T25 and T49 are proposed to be phosphorylated first, followed by S44 and then by either S35 or S38 with T22 occurring late. Most variants were found to possess multiple of these phosphorylated sites including N1.

Flightin appears to become unphosphorylated during flight after 90 seconds as there is a decrease in the less acidic (P1-P4) forms with a correlating increase in the unphosphorylated form, N1 [116]. Residues S31 and S44 have been proposed to be responsible for the dephosphorylation observed during flight. This could be a side effect of continued IFM activity or serve a functional purpose, compensating for needs that arise for endurance flight.

Phosphoryation does not appear to be involved in the initial structural organization of the IFM but could be important for a number of other reasons. Phosphorylation may influence modulation of flightin's positioning in the thick filament, interaction with other sarcomeric components or structural influence of flightin in the LMM itself to accommodate varying viscoelastic needs in the functional actomyosin environment. It is possible that while initial binding is accomplished without phosphorylation, phosphorylation is necessary to cement flightin's positioning in the LMM sufficiently enough to endure the mechanical challenge experienced by its substrate during muscle activity.

Supported & possible protein-protein interactions

Myosin binding proteins frequently have been found to bind one another and, while there is no direct evidence of such behavior for flightin, protein-protein interactions beyond the myosin LMM may be partially responsible for the phenotypes in flightin mutants. The

definition of the Z-disc and M-line necessarily describe multiple contacts for all proteins. For example, abba interacts with alpha-actinin, kettin, and mlp84B [120] though an abba mutation doesn't produce the identical mutant phenotype as an alpha-actinin, kettin or mlp84B mutation [52, 121-124] as all these proteins require various additional contacts for complete function. Multiple contacts for the protein components of the thin and thick filaments are the norm in the involved network that forms the sarcomere, making additional binding partners to flightin worth consideration.

The phenotype of flightin mutants provide direction when evaluating other binding partners. Flightin mutants exhibit sarcomere and thick filament length change, aberrant and weak or absent M-lines, disruption of the myofilament lattice, increased thick filament packing within the lattice and decreased stability of the myosin dimers. Z-discs break apart post-contraction, along with the entirety of the myofibril. Oscillatory work is also compromised alongside decreased passive stiffness in all flightin mutants. The pre-contractile phenotype is somewhat different in lines lacking N- and C-terminal domains of flightin though all cases feature shorter sarcomeres in the adult IFM, shorter thick filaments and changes to the M-line and lattice. Changes in the myofilament lattice, and decreased thick filament length have been associated with the absence of proteins serving as adaptors between the thin and thick filaments and connecting filaments, including Lasp [48], and formins DAAM [125] and Fhos [69]. Matters of the M-line are almost exclusively focused on obscurin, required for the presence of the M-line, H-zone, and

thick/thin filament symmetry within the Drosophila sarcomere [126], though a smaller isoform of sls, zormin, can also be found in the M-line.

This subsection is arranged to address potential protein-protein interactions by relationship type rather than by type of partner as many proteins have multiple, or unknown, functions or are present in multiple areas of the sarcomere (e.g. zormin in both the Z disc and M line). It is also important to recognize that not all protein-protein interactions involve direct binding and some connections may be more/less involved than others. Hence, this section is separated into three categories: i) binding site exclusion, ii) repulsive effects and iii) binding interactions. The last subsection (iv) addresses flightin as a component of a multi-protein process.

Binding Site Exclusion

Flightin's binding site in the myosin LMM has been increasingly clarified in recent years, enabling comparisons of the relevant region between organisms. Flightin is known to be a myosin-binding A-band resident with the region around E1554 being necessary for this incorporation. The E1554 (E1563K in Lethocerus) residue is found to be part of a site that extends from (Lethocerus) E1547 to R1582 (**Chapter 4**), corresponding to E1538-R1575 in Drosophila. Additional binding sites are expected to be around (Lethocerus) S972-L996, E1254-A1284, S1759-T1786 and S1851-Q1873 (**Chapter 4**). Gold labelling indicated the densities associated with flightin are separated by 800 nm, centered at the Z-disc. Using calculated values of the sarcomere component widths from Szikora et al

2019 [127], in which the Z-disc and I band together are about 304 nm, this indicates that flightin shows up about 250 nm in to the A band. Further analysis of thick filament cryo-EMs featuring flightin (**Chapter 4**) indicate that the binding site involving E1554 simultaneously involves interfaces S1759-T1786 and S1851-Q1873 from two other myosin dimers and that this "multiface" first becomes possible ~250 nm into the A-band, supporting these regions as being of primary importance, and likely to be the most secure of the flightin contact points.

M-band associated proteins in vertebrates have been known to contact the LMM region occupied by flightin and similar proteins in Drosophila may be excluded from the A band as a result. The region surrounding E1554 is a hot spot for binding in vertebrate LMM as M-protein, myomesin and MyBP-C all bind in this area [104, 128, 129]. M-protein and MyBP-C bind the region with Ig domains while myomesin makes contact via a disordered region. All three of these proteins are known to also bind obscurin, making this region of further relevance to the M band. Obscurin-like-1 (Obsl1) also may bind the LMM, but further research is required to verify [130]. Except for obscurin, none of these proteins have homologs in Drosophila, though other proteins in Drosophila have Ig domains that may find this area attractive, including sallimus (sls) isoforms, projectin, and stretchin. Of these, the major sls isoform kettin is solidly established in the Z-disc [45, 127, 131] and does not require exclusion from myosin sites for its localization, while other sls isoforms and projectin have a more complicated relationship.

Distribution and timing of expression may provide a hint as to the likelihood of binding site exclusion by flightin occupancy. Sallimus isoform zormin is considered to have a more unusual distribution in the sarcomere compared to kettin as zormin is present in both the Z-disc and M-line. Though zormin could be binding a site unoccupied by flightin in the M-line, it is more likely to be associated with obscurin as it becomes diffuse when obscurin is depleted by RNAi [126]. Zormin is also expressed earlier than flightin, permitting it the opportunity of flightin site exposure before flightin placement [132]. Projectin is a much more probable as it crosslinks the thin and thick filaments in the I band [127, 133-135] and is found in the A-band in synchronous muscles [59], suggesting that further binding of myosin is possible in the absence of IFM-specific proteins, such as flightin. Projectin also is expressed maximally after flightin expression, making it less likely to be exposed to unoccupied flightin binding sites on the LMM [132]. Stretchin, a titin-related Ig-heavy A-band resident, is also a candidate for exclusion from the flightin's binding regions. They are both under the jurisdiction of transcription factor spalt major (Salm) and expressed at pupal stage 8 [132]. Notably, this might place them in competition for binding sites if they shared affinity, but a mixture of binding regions is not observed between the two in recent studies [136]. While the co-expression may allude to collaboration between the two proteins, it is unlikely that they compete for binding sites.

M-line associated proteins obscurin and paramyosin may also be considered. Drosophila obscurin binds myosin [126] and it is currently unknown why it is restricted to the M-

line. Obscurin/UNC-89 of *C.elegans* is not strictly restricted to the M-line and is found within the A-band [137]. Miniparaymosin also binds myosin and is detectable in both the M-line and A-band in muscles lacking flightin, but only in the M-line of the IFM [138, 139], and is expressed later in muscle maturation [140]. Obscurin further shares a connection with flightin via transcription regulator spalt major [132]; it becomes detectable earlier than flightin but is expressed at higher levels later on. This would suggest that, if a binding site were shared, it would be only secondary to obscurin's primary binding site still most accessible within the M-line. While it is likely that obscurin, and/or (mini)paramyosin, is being excluded from a secondary myosin binding site in the A-band, it may be that this is a factor of dense association of myosin dimers, later secured by flightin- or more dominant binding sites available elsewhere, such as with paramyosin.

Taken together, if flightin were to be involved in exclusion of other myosin-binding proteins from the thick filament, projectin is the most likely candidate. Though stretchin is also capable of engaging in such a function as the LMM region known to be bound by flightin is the area of affinity for Ig-domain containing proteins in other organisms. It must be noted that flightin may be masking binding opportunities within myosin by more indirect means. Other ways in which it could exclude myosin binding in areas outside of its binding site include i) blocking regions by repulsive interactions with proximal LMM-binding or non-binding regions of flightin or ii) impacting the structure of the myosin

dimer sufficiently to alter availability of an alternate interface, such as by securing the coiled-coil in areas that may otherwise be more flexible.

Repulsive effects

Repulsive forces between the protein components are known to be a major player in muscle structure and function and flightin has been hypothesized to function, in part, based on electrostatic repulsive effects. For instance, phosphorylation of the myosin regulatory light chain (MLC2) enhances the number of actomyosin cross-bridges through electrostatic repulsion between the MLC2 and myosin heavy chains in mammalian muscle [141, 142] and it is generally accepted that electrostatic forces are the driving factor for the association of myosin dimers.

It has been previously suggested that the highly acidic N-terminus (pI: 3.78) of flightin may be involved in electrostatic repulsive force with the actin thin filament leading to maintenance of proper interfilament spacing [143]. The N-terminus is predicted to be the most disordered region of flightin and is subject to phosphorylation, possibly responsible for the multiple phosphor-variants associated with flightin [116] in Drosophila and further exacerbating its repulsive effects. It has been proposed that, if extended, the N-terminus may be able to reach up to 27 nm, while only 18 nm must be traversed from the surface of the thick filament to reach the thin filament. As the lattice of $fln^{\Delta N62}$ is significantly more compact, this is a feasible function of the N-terminal region. Small changes in the myofilament lattice may explain the mechanical differences between the

fln$^{\Delta N62}$ and WT and changes of even 1 nm have been known to be a determining factor in functional capacity between otherwise identical muscles [144]. Such nuance may be the driving factor in the differences in song production seen between Drosophila species in which flightin's N-terminus exhibits the most variation and is under greater selective pressure than the rest of the protein.

This kind of interaction may include involvement of an intermediary protein that engages with the thin filament in the A band, such as Lasp, the only nebulin repeat protein in Drosophila and is found throughout the A band and I band. The myofilament lattice is impacted in Lasp null lines but conversely to that observed in the *fln*$^{\Delta N62}$ with fewer filaments per um^2 as opposed to the more compact lattice [48]. It is hypothesized that Lasp presence in the A-band involves interaction with both the thick and thin filaments to maintain interfilament spacing; it is possible that the flightin N-terminus participates in this effort.

Binding Interactions

When it comes to consideration of candidates for direct binding to flightin, we can conveniently limit ourselves to either known or suspected A band residents. Known thick filament proteins of Drosophila IFM A-band include stretchin-klp, myofilin and paramyosin with suspected residents, such as Lasp. However, myofilin can be removed from the list of flightin-binding suspects as it has been found to not bind flightin *in vitro* and is not expected to bind *in vivo* [145]. As some myosin antibodies can not detect

myosin in the A band due to the packing of the dimers, it is expected that some myosin binding proteins within the A band likewise have their epitope masked [48].

Stretchin-klp which shows the same expression pattern as flightin, and may connect via collaborative functions, is a possible binding partner. Stretchin-klp has two isoforms of 225 kDa and 231 kDa prevalent at 72 hr APF, grouping with cluster 22 [146] and is found to be dispersed within the A-band of the sarcomere by gold labelling in a strikingly similar pattern as flightin [147]. Both are down-regulated during aging and after paraquat [148]. The larger isoform only accumulates in the presence of the adult myosin rod, though the smaller isoform requires neither myosin nor actin [147]. Stretchin-klp has been hypothesized to bind aa1120-1200 of the LMM based off of recent cryo-EM studies of Drosophila thick filament [136] in which non-myosin proteins are proximal to the area of difference in hinge switch mutants (1216-1241). In hinge-switch mutants, the larger stretchin isoform fails to accumulate; altered actomyosin kinetics, decreased flight ability, increased sarcomere lengths, and disruption of the peripheral myofilaments of the lattice over time [149, 150] are incurred in the mutant. The smaller stretchin isoform may be more reliant on paramyosin, which still accumulates in *Mhc*^{7} and *Act88F* mutants, but still lack the additional contacts necessary for full function. Contact between stretchin and flightin, both aligned temporally and spatially within the A band, may be involved in mutual functions involving thick-thin filament alignment.

Paramyosin (PM) should be considered based off comparable functional roles along with its similarity to flightin's main binding partner, myosin. It has been proposed that PM interacts with myosin and acts as a scaffold, or chaperone, during thick filament development [21]. It's possible that the LMM binding region is retained sufficiently on PM to promote flightin binding to PM dimers in the core, though high sequence homology may not be necessary for direct contact. In Lethocerus thick filament cryo-EM studies [2], PM was identified in contact with non-myosin densities, including one designated as flightin, strongly supporting an interaction between the two. As paramyosin connects to the M-line and has both mechanical and structural impacts on the thick filament and sarcomere, flightin's engagement with it may explain some of the M-line associated defective characteristics of the flightin transgenic mutants.

Direct contact of flightin with the thin filament can also be considered. Just as the N-terminus has been hypothesized to engage in a repulsive interaction with Lasp, so too could it be involved in a more dynamic relationship, inclusive of direct contact, possibly dependent on phosphorylation state. Flightin has a region of sequence similarity to the F-actin binding motif found in F-actin bundling proteins [151, 152] and may be involved in direct binding to actin. This region could also be conditionally involved in a collaborative actin binding effort with another binding protein in which an initial binding interaction with flightin enhances the affinity of an actin binding site on that protein.

The top candidates for flightin engagement include stretchin, lasp, actin, and paramyosin but it is also important to keep in mind that more than one of these proteins, or less characterized members of the A-band, may be binding flightin. For instance, binding to stretchin may enable an interface to become available that is now conducive to actin binding. Limpet (Lmpt) found in the I-band and Z-disc in Drosophila muscle [153], is homologous to mammalian four-and-a-half LIM (Fhl1) [154]. Lmpt is the only known homolog to Fhl1, which is known to bind MyBP-C, and possibly myosin, localizing to the M-line, I band, and partially into the A band [155]. Flightin may be either involved in a direct binding interaction that masks an epitope needed for fluorescent detection in the A band or involved in an exclusion activity to such a less-characterized protein as Limpet.

As part of a system/process

Flightin may also be part of a modular system of proteins that serve a purpose in thick-thin filament scaffolding during sarcomere maturation and ensuring proper mechanical relay and stability during muscle operation through adulthood. In such a system, a multi-protein process uses different protein connectors for linking the thin filament to the outside of the thick filament with distinct connections into the thick filament core (**Fig. 1-2**). Internal thick filament protein(s) would be responsible for connecting the external protein cord/process to the internal cord/process. Connections in the I-band and in the H-zone/M-line serve a prominent purpose in initial alignment of thick-thin filaments during

muscle development but, alone, provide insufficient stability in the face of the continuous, regular, contractile forces present in the operation of the adult IFM. This proposed modular system may be functionally likened, and considered an alternative, to the large titin and nebulin proteins in vertebrates [156, 157]. Titin is considered a primary component in communicating stretch through the A band. Titin operates as an interface that contacts both proteins of the thick filament and proteins of the thin filament, creating an additional elastic process by which strain is relayed. Force can be conditionally dissipated or relayed depending on titin as mediator. By incorporating larger components (such as titin), with their own viscoelastic landscapes, the mode (e.g. rate of relay, extent of dissipation) by which mechanical communication occurs is determined first by the large protein itself, and next by altering effects upon its state via its contact to other proteins. Having regions of one large protein structurally, and functionally, separated from each other by outside-contacts, may allow each contact to be a liability in terms of its potential to interrupt a neighboring segment. In a more modular invertebrate system in which this structural separation is now enabled by distinct, smaller and more local, protein modules the locality of potential disruption at contact sites may be minimized. This may also permit adaption/evolution to occur on the level of each module, without having a dramatic effect on the other modules. In this way, actomyosin interactions, short range thick filament stiffness/mechanical properties of the A band, and long range thick filament stiffness/mechanical properties extending to the H- and M-lines can be tuned with greater isolation than if dependent on another overarching large

filament in which changes in stoichiometry, sequence, and temporal expression would have more impactful, multi-level and far-reaching effects.

The top candidates for flightin interaction are also the most likely to be involved in a thin-to-thick filament core process. On the thin filament end, Lasp and Projectin, sometimes called mini-Titin [135, 158, 159], are prominent. Projectin is found to link the thin and thick filaments, spanning from the Z-disc to the A band [160] and is likely to be contacting the thin filament within the I band as well as in the Z-disc [59, 161]. The only nebulin protein in Drosophila, Lasp, is hypothesized to be involved with myosin contacts coordinated with the periodicity of the myosin heads [48], though the impact on contractile processes still remains to be examined. Stretchin-klp, a multi-Ig domain titin-like protein, is found to bind near the first skip domain on the myosin rod and modulate interaction between the myosin head and rod [136]. Stretchin-klp is expected to be extensible and would represent an elastic relay between the thin-and-thick filaments. Stretchin-klp, or another isoform of stretchin, could be contacting actin directly or be in contact with Lasp. Stretchin does not appear to reach the thick filament core, but both flightin and myofilin do, and appear to extend into the thick filament core to contact paramyosin. Stretchin is large enough to come in contact with either, or both, flightin and myofilin. In such a case, flightin would be involved in the bridge between thin filament contacting proteins and the paramyosin scaffold of the center of the thick filament.

Modules of this proposed system could be separated into three functional categories: thin-filament connection, the inter-thick filament myosin and the myosin-core connections (**Fig. 1-2**). The primary purpose of the figure is to show how invertebrates may be incorporating a connective process, involving flightin, that features two main elements, i) increased modularity of thick-thin filament connecting process and ii) progression of this process through the myosin layers into the thick filament core. The increased modularity may be a requirement for a filament system that transverses the thick filament while still accommodating the different, and dynamic, requirements of the thick and thin filament operating within the temporal constraints of sarcomere development. A minimalist approach is taken in **Fig. 1-2** in order to emphasize the comparison of the invertebrate modular system to titin of the vertebrate system and its integration through the thick filament core. The generalized, approximate positions of the protein constituents shown are backed empirically with the exception of stretchin's connection to the thin filament, which is hypothetical and demarcated with an asterisk. From top to bottom in **Fig. 1-2** (A), projectin connecting the Z-ward edge of the thick filament to the thin filament or Z-disc, stretchin connecting the thin filament to the thick filament in the A-band, stretchin contacting flightin and the thick filament to create the linkage into the core, myofilin and flightin connecting myosin to paramyosin, and paramyosin extending to connect to the M-band. From top to bottom in **Fig. 1-2** (B), titin connects myosin both to the Z-disc and to the thin filament in the I band, titin connects to myosin and other myosin affiliated proteins and titin extends to connect myosin to the M-line and M-line associated proteins. The entire process in vertebrates takes place external to the myosin assemblies of the

thick filament while this process permeates into the core of the thick filament in invertebrates. In such a system it is very likely that other protein components are at play, including- for instance- Lasp, which is not included due to both its localization on the thin filament within the A band and its proposed interaction with stretchin being hypothetical. The figure does not portray the Z-disc or M-line or involve the complex network of proteins within each.

Flightin's tripartate organization based off of different evolutionary pressures of the N-terminal, WYR, and C-terminal may support its segmented role in the modular invertebrate system at its intersection. The flightin transgenic lines support involvement in all three categories, thin filament association, inter-thick, and core to M-line, without being the primary regulator in all of them. Though flies are still flight-capable in the $fln^{\Delta N62}$ transgenic lines, the N-terminus is suggested for some function on the surface of the thick filament with the thin filament or thin-filament associated proteins, where phosphovariants may arise. The highly conserved WYR region is indicated as important for the thick filament myosin contacts among the, likewise, highly conserved regions of myosin, and the C-terminus may be more closely related to function within the thick filament core and to contacts that ultimately have bearing in the M-line.

A similar arrangement of proteins was proposed in the schematic of a portion of the Drosophila myofilament lattice at the A-I junction in Maughan & Vigoreaux, 2005 [109]. The authors propose that both projectin and kettin are involved in linking the thick

filament to the Z-disc and are likely to also connect to the thin filament of the I band in the process. Projectin is found in the A band and both the Z-disc and I band in asynchronous IFM [133-135]. While it is possible that projectin and kettin are connected to each other in some way, it is known that kettin does not require myosin for its localization to the Z-disc and does not bind myosin *in vitro* [162]. Kettin has been clearly established as a Z-disc specific resident by antibodies to its N-terminus [45], central region [163], and C-terminus [45, 131]. The only support for kettin binding close to the A band exists in cases of non-native conditions, such as the absence of its native binding partner in an Act88F mutant [164]. It has been repeatedly established that Kettin is not found in direct contact with the thick filament of the native sarcomere [45, 127, 131, 162]. A direct interaction between kettin and myosin has been described as highly unlikely [127] though an indirect one-mediated by another protein-is possible, with projectin proposed as the prime candidate [127]. Kettin may still be involved in such a linkage between the thick and thin filaments via projectin [163] and this would be well aligned with our proposed modular model. However, as kettin-projectin affiliation is lacking in empirical support, we focus on projectin as the primary linker of the thick filament to the thin filament and Z-disc though kettin may be playing a supporting role more indirectly, closer to the Z-disc.

Modularity has been proposed to be both the product, and facilitator, of evolution [165]. Proteins contain sets of residues that form a domain, capable of a specific function or binding interaction each of which is, in itself, a module with the set of domains- the full

protein- being a higher-level module. In a protein-module, the component domains are subject to circumscription as a whole. When sets of domains are organized into separately expressed proteins, rather than belonging to one large protein, the modularity of the system is increased and the sets of domains gain a level of flexibility in the way their functions can be realized. Discretion of functional components increases the degree by which these components may be recombined, titrated or otherwise adapted. For instance, changing the stoichiometry of a component, such as stretchin or paramyosin, may further involve an adaption to a connecting protein- such as flightin- to undergo changes in affinity, but not function, to limit changes within other connecting modules. Increased modularity permits variation of independent/interdependent behavior among the delineated constituents and may factor in to both the high diversity and evolutionary success of Invertebrates. The proposed modular system being employed within the protein filaments of the sarcomere is representative of such an evolutionary technique. As party to this arrangement, flightin is proposed to contribute as the mediator between the protein modules outside the thick filament, ‘Z-ward’ or in extension to the thin filament, and inside the thick filament, towards the securing elements of the M-line. Further examination of any of the components for such a process would benefit from a keen eye towards co-characterization of potential interacting partners.

Forward to chapters 2-4

The next few chapters address the structure and interactions of flightin's novel WYR domain and expands on characteristics of flightin function. Chapter 2 compares the properties of cardiac Myosin Binding Protein C (cMyBP-C) and flightin in the the context of the Drosophila flight muscle to evaluate whether their operation within muscle may have similarities. Chapter 3 has two major sections, 'WYR SEQUENCE AND STRUCTURE' and 'LMM AND C600 SEQUENCE AND STRUCTURE'. In the first of these two sections, the CD profile of WYR is presented and accompanied by an analysis of the WYR sequence as it pertains to the development of secondary structure. The structural content estimated by BeStSel is put into more particular context and a hypothetical secondary structure of WYR is proposed. In the second of these two sections, the alpha helical coiled-coil of the LMM is described and compared to our findings by circular dichroism and placed adjacent to the examination of WYR binding to the LMM and associated structural changes. In chapter 4, additional analysis of Cryo-EM studies done by Hu et al. (2016) are incorporated together with our findings to examine flightin behavior within the context of the thick filament.

FIGURES

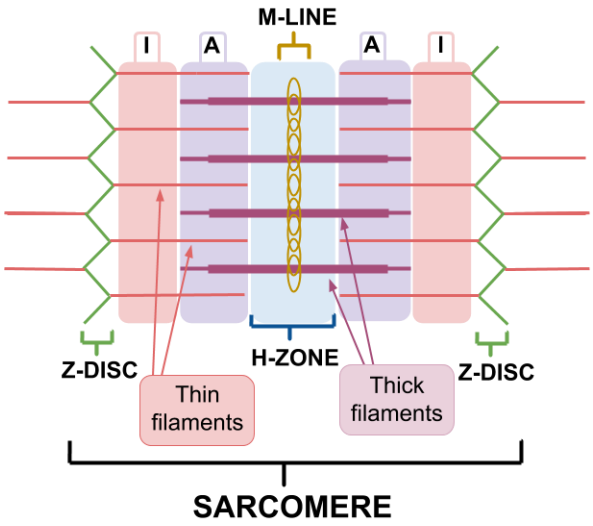


Figure 1-1: Muscle Sarcomere and Terminology. The Sarcomere is made up of interdigitating thick and thin filaments. Different regions are designated in affiliation with the relationship to these filaments. The "A" and "I" bands are designated towards the top the figure over corresponding shaded regions. The center of the sarcomere contains the M-line, flanked by the H-zone. The whole sarcomere unit is flanked by thin filament anchoring zones known as Z-discs. For simplicity, not all filament systems are shown.

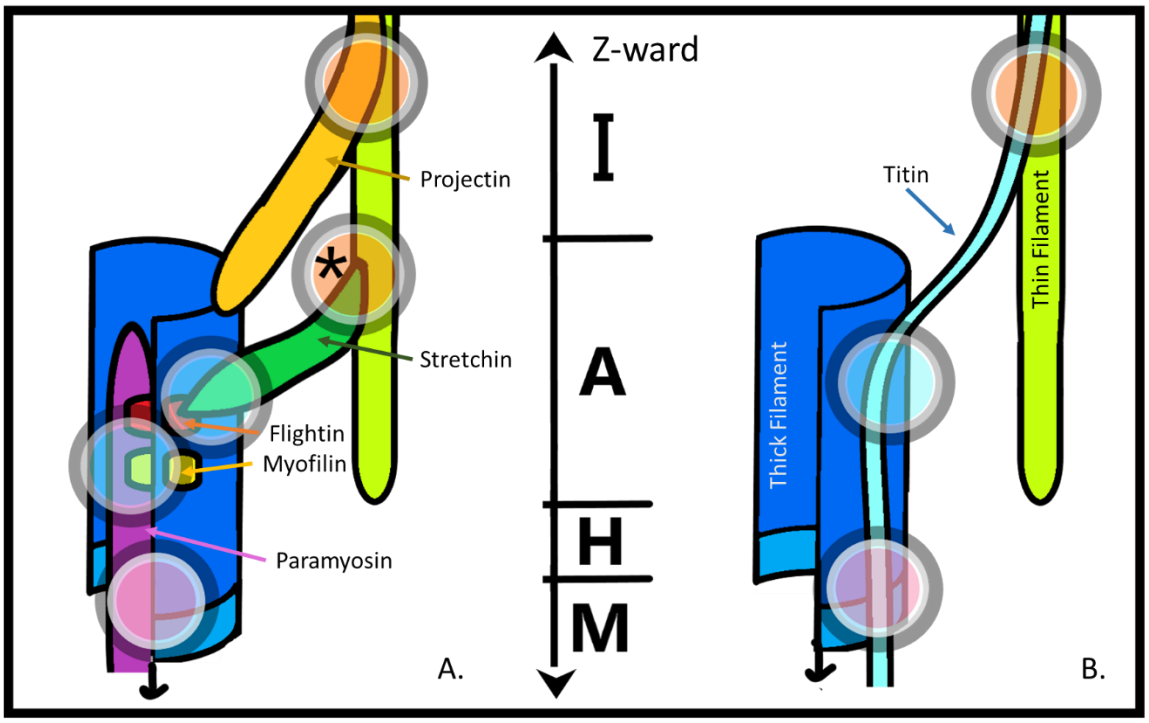


Figure 1-2: Increased modularity in the thick-thin filament connecting process of Invertebrate systems vs Vertebrate systems – concept diagram. (A) Invertebrates may employ a system in which a connecting process is constituted by distinct protein components that enable stabilization through connections from the thin to the thick filament and thick filament core, extending into the M-line from the thick filament core. (B) The vertebrate system relies heavily on domain-based connection modules within a single large extensive protein, titin, which connects the thick filament to both the thin filament and the M-line by forming a process that remains external to the thick filament itself. Other proteins exist within both systemss to support these distinct methods of integration are not shown (e.g. MyBP-C). Circle denote connective modules with an internal coloring of orange for thick-to-thin, blue for within thick-filament or between myosin dimers, and pink for thick to M-line or core.

BIBLIOGRAPHY

[1] Hu Z, Taylor DW, Edwards RJ, Taylor KA. Coupling between myosin head conformation and the thick filament backbone structure. Journal of structural biology. 2017;200:334-42.

[2] Hu Z, Taylor DW, Reedy MK, Edwards RJ, Taylor KA. Structure of myosin filaments from relaxed Lethocerus flight muscle by cryo-EM at 6 A resolution. Science advances. 2016;2:e1600058.

[3] Alamo L, Qi D, Wriggers W, Pinto A, Zhu J, Bilbao A, et al. Conserved Intramolecular Interactions Maintain Myosin Interacting-Heads Motifs Explaining Tarantula Muscle Super-Relaxed State Structural Basis. Journal of molecular biology. 2016;428:1142-64.

[4] Lee KH, Sulbaran G, Yang S, Mun JY, Alamo L, Pinto A, et al. Interacting-heads motif has been conserved as a mechanism of myosin II inhibition since before the origin of animals. Proceedings of the National Academy of Sciences of the United States of America. 2018;115:E1991-e2000.

[5] Mansson A, Persson M, Shalabi N, Rassier DE. Nonlinear Actomyosin Elasticity in Muscle? Biophysical journal. 2019;116:330-46.

[6] Rassier DE. Sarcomere mechanics in striated muscles: from molecules to sarcomeres to cells. American journal of physiology Cell physiology. 2017;313:C134-c45.

[7] Nishikawa K. Eccentric contraction: unraveling mechanisms of force enhancement and energy conservation. The Journal of experimental biology. 2016;219:189-96.

[8] Nishikawa KC, Monroy JA, Tahir U. Muscle Function from Organisms to Molecules. Integrative and comparative biology. 2018;58:194-206.

[9] Lin BL, Song T, Sadayappan S. Myofilaments: Movers and Rulers of the Sarcomere. Comprehensive Physiology. 2017;7:675-92.

[10] Gautel M, Djinovic-Carugo K. The sarcomeric cytoskeleton: from molecules to motion. The Journal of experimental biology. 2016;219:135-45.

[11] Lange S, Pinotsis N, Agarkova I, Ehler E. The M-band: The underestimated part of the sarcomere. Biochimica et biophysica acta Molecular cell research. 2019.

[12] Irving M. Regulation of Contraction by the Thick Filaments in Skeletal Muscle. Biophysical journal. 2017;113:2579-94.

[13] Alamo L, Koubassova N, Pinto A, Gillilan R, Tsaturyan A, Padron R. Lessons from a tarantula: new insights into muscle thick filament and myosin interacting-heads motif structure and function. Biophysical reviews. 2017;9:461-80.

[14] Moore JR, Campbell SG, Lehman W. Structural determinants of muscle thin filament cooperativity. Archives of biochemistry and biophysics. 2016;594:8-17.

[15] Lehman W. Thin Filament Structure and the Steric Blocking Model. Comprehensive Physiology. 2016;6:1043-69.

[16] Lindstedt S, Nishikawa K. Huxleys' Missing Filament: Form and Function of Titin in Vertebrate Striated Muscle. Annual review of physiology. 2017;79:145-66.

[17] Myhre JL, Pilgrim D. A Titan but not necessarily a ruler: assessing the role of titin during thick filament patterning and assembly. Anatomical record (Hoboken, NJ : 2007). 2014;297:1604-14.

[18] Tskhovrebova L, Trinick J. Titin and Nebulin in Thick and Thin Filament Length Regulation. Sub-cellular biochemistry. 2017;82:285-318.

[19] Manring HR, Carter OA, Ackermann MA. Obscure functions: the location-function relationship of obscurins. Biophysical reviews. 2017;9:245–58.

[20] Gunage RD, Dhanyasi N, Reichert H, VijayRaghavan K. Drosophila adult muscle development and regeneration. Seminars in cell & developmental biology. 2017;72:56-66.

[21] Hooper SL, Hobbs KH, Thuma JB. Invertebrate muscles: thin and thick filament structure; molecular basis of contraction and its regulation, catch and asynchronous muscle. Progress in neurobiology. 2008;86:72-127.

[22] Sanger JW, Wang J, Fan Y, White J, Mi-Mi L, Dube DK, et al. Assembly and Maintenance of Myofibrils in Striated Muscle. Handbook of experimental pharmacology. 2017;235:39-75.

[23] Baghdadi MB, Tajbakhsh S. Regulation and phylogeny of skeletal muscle regeneration. Developmental biology. 2018;433:200-9.

[24] Deora T, Gundiah N, Sane SP. Mechanics of the thorax in flies. The Journal of experimental biology. 2017;220:1382-95.

[25] Lemke SB, Schnorrer F. Mechanical forces during muscle development. Mechanisms of development. 2017;144:92-101.

[26] Kreipke RE, Kwon YV, Shcherbata HR, Ruohola-Baker H. Drosophila melanogaster as a Model of Muscle Degeneration Disorders. Current topics in developmental biology. 2017;121:83-109.

[27] Rai M, Nongthomba U, Grounds MD. Skeletal muscle degeneration and regeneration in mice and flies. Current topics in developmental biology. 2014;108:247-81.

[28] Kronert WA, Bell KM, Viswanathan MC, Melkani GC, Trujillo AS, Huang A, et al. Prolonged cross-bridge binding triggers muscle dysfunction in a Drosophila model of myosin-based hypertrophic cardiomyopathy. eLife. 2018;7.

[29] Potikanond S, Nimlamool W, Noordermeer J, Fradkin LG. Muscular Dystrophy Model. Advances in experimental medicine and biology. 2018;1076:147-72.

[30] Engel DGMS. Evolution of the Insects. 1 ed: Cambridge University Press; 2005.

[31] Wilson E. The Diversity of Life. Cambridge, MA and Norton, NY: Belknap Press; 1992.

[32] Mayhew PJ. Why are there so many insect species? Perspectives from fossils and phylogenies. Biol Rev Camb Philos Soc. 2007;82:425-54.

[33] Deora T, Singh AK, Sane SP. Biomechanical basis of wing and haltere coordination in flies. Proc Natl Acad Sci U S A. 2015;112:1481-6.

[34] Nicholson DB, Ross AJ, Mayhew PJ. Fossil evidence for key innovations in the evolution of insect diversity. Proc Biol Sci. 2014;281.

[35] Chakravorty S, Vu H, Foelber V, Vigoreaux JO. Mutations of the Drosophila myosin regulatory light chain affect courtship song and reduce reproductive success. PloS one. 2014;9:e90077.

[36] Chakravorty S, Wajda MP, Vigoreaux JO. Courtship song analysis of Drosophila muscle mutants. Methods (San Diego, Calif). 2012;56:87-94.

[37] Vigoreaux JO. Genetics of the Drosophila flight muscle myofibril: a window into the biology of complex systems. BioEssays : news and reviews in molecular, cellular and developmental biology. 2001;23:1047-63.

[38] Wallgren-Pettersson C, Sewry CA, Nowak KJ, Laing NG. Nemaline myopathies. Seminars in pediatric neurology. 2011;18:230-8.

[39] van Eldik W, Passier R. Signalling in sarcomeres in development and disease. Netherlands heart journal : monthly journal of the Netherlands Society of Cardiology and the Netherlands Heart Foundation. 2013;21:367-71.

[40] Ochala J, Sun YB. Novel myosin-based therapies for congenital cardiac and skeletal myopathies. Journal of medical genetics. 2016;53:651-4.

[41] Obermann WM, Gautel M, Steiner F, van der Ven PF, Weber K, Furst DO. The structure of the sarcomeric M band: localization of defined domains of myomesin, M-

protein, and the 250-kD carboxy-terminal region of titin by immunoelectron microscopy. The Journal of cell biology. 1996;134:1441-53.

[42] Hu LY, Ackermann MA, Kontrogianni-Konstantopoulos A. The sarcomeric M-region: a molecular command center for diverse cellular processes. BioMed research international. 2015;2015:714197.

[43] Luther PK. The vertebrate muscle Z-disc: sarcomere anchor for structure and signalling. Journal of muscle research and cell motility. 2009;30:171-85.

[44] Liao KA, Gonzalez-Morales N, Schock F. Zasp52, a Core Z-disc Protein in Drosophila Indirect Flight Muscles, Interacts with alpha-Actinin via an Extended PDZ Domain. PLoS genetics. 2016;12:e1006400.

[45] van Straaten M, Goulding D, Kolmerer B, Labeit S, Clayton J, Leonard K, et al. Association of kettin with actin in the Z-disc of insect flight muscle. Journal of molecular biology. 1999;285:1549-62.

[46] Chechenova MB, Bryantsev AL, Cripps RM. The Drosophila Z-disc protein Z(210) is an adult muscle isoform of Zasp52, which is required for normal myofibril organization in indirect flight muscles. The Journal of biological chemistry. 2013;288:3718-26.

[47] Katzemich A, Liao KA, Czerniecki S, Schock F. Alp/Enigma family proteins cooperate in Z-disc formation and myofibril assembly. PLoS genetics. 2013;9:e1003342.

[48] Fernandes I, Schock F. The nebulin repeat protein Lasp regulates I-band architecture and filament spacing in myofibrils. The Journal of cell biology. 2014;206:559-72.

[49] Witt CC, Burkart C, Labeit D, McNabb M, Wu Y, Granzier H, et al. Nebulin regulates thin filament length, contractility, and Z-disk structure in vivo. The EMBO journal. 2006;25:3843-55.

[50] Pappas CT, Bhattacharya N, Cooper JA, Gregorio CC. Nebulin interacts with CapZ and regulates thin filament architecture within the Z-disc. Molecular biology of the cell. 2008;19:1837-47.

[51] Mardahl-Dumesnil M, Fowler VM. Thin filaments elongate from their pointed ends during myofibril assembly in Drosophila indirect flight muscle. The Journal of cell biology. 2001;155:1043-53.

[52] Hakeda S, Endo S, Saigo K. Requirements of Kettin, a giant muscle protein highly conserved in overall structure in evolution, for normal muscle function, viability, and flight activity of Drosophila. The Journal of cell biology. 2000;148:101-14.

[53] Ayme-Southgate AJ, Turner L, Southgate RJ. Molecular analysis of the muscle protein projectin in Lepidoptera. Journal of insect science (Online). 2013;13:88.

[54] Burton PM. Insights from diploblasts; the evolution of mesoderm and muscle. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution. 2008;310B:5-14.

[55] Liu AG, Matthews JJ, Menon LR, McIlroy D, Brasier MD. Haootia quadriformis n. gen., n. sp., interpreted as a muscular cnidarian impression from the Late Ediacaran period (approx. 560 Ma). Proceedings Biological sciences. 2014;281.

[56] Seipel K, Schmid V. Evolution of striated muscle: jellyfish and the origin of triploblasty. Developmental biology. 2005;282:14-26.

[57] Mackie GO, Mills CE, Singla CL. Structure and function of the prehensile tentilla of Euplokamis (Ctenophora, Cydippida). Zoomorphology. 1988;107:319-37.

[58] Baguna J, Martinez P, Paps J, Riutort M. Back in time: a new systematic proposal for the Bilateria. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2008;363:1481-91.

[59] Hooper SL, Thuma JB. Invertebrate muscles: muscle specific genes and proteins. Physiological reviews. 2005;85:1001-60.

[60] Rui Y, Bai J, Perrimon N. Sarcomere formation occurs by the assembly of multiple latent protein complexes. PLoS genetics. 2010;6:e1001208.

[61] Sebe-Pedros A, Grau-Bove X, Richards TA, Ruiz-Trillo I. Evolution and classification of myosins, a paneukaryotic whole-genome approach. Genome biology and evolution. 2014;6:290-305.

[62] Steinmetz PR, Kraus JE, Larroux C, Hammel JU, Amon-Hassenzahl A, Houliston E, et al. Independent evolution of striated muscles in cnidarians and bilaterians. Nature. 2012;487:231-4.

[63] Menard L, Maughan D, Vigoreaux J. The structural and functional coordination of glycolytic enzymes in muscle: evidence of a metabolon? Biology. 2014;3:623-44.

[64] Lygren B, Tasken K. Compartmentalized cAMP signalling is important in the regulation of Ca(2+) cycling in the heart. Biochemical Society transactions. 2006;34:489-91.

[65] Mishra P, Varuzhanyan G, Pham AH, Chan DC. Mitochondrial Dynamics is a Distinguishing Feature of Skeletal Muscle Fiber Types and Regulates Organellar Compartmentalization. Cell metabolism. 2015;22:1033-44.

[66] Willis BC, Ponce-Balbuena D, Jalife J. Protein assemblies of sodium and inward rectifier potassium channels control cardiac excitability and arrhythmogenesis. American journal of physiology Heart and circulatory physiology. 2015;308:H1463-73.

[67] Katzemich A, West RJ, Fukuzawa A, Sweeney ST, Gautel M, Sparrow J, et al. Binding partners of the kinase domains in Drosophila obscurin and their effect on the structure of the flight muscle. Journal of cell science. 2015;128:3386-97.

[68] Borisov AB, Sutter SB, Kontrogianni-Konstantopoulos A, Bloch RJ, Westfall MV, Russell MW. Essential role of obscurin in cardiac myofibrillogenesis and hypertrophic response: evidence from small interfering RNA-mediated gene silencing. Histochemistry and cell biology. 2006;125:227-38.

[69] Shwartz A, Dhanyasi N, Schejter ED, Shilo BZ. The Drosophila formin Fhos is a primary mediator of sarcomeric thin-filament array assembly. eLife. 2016;5.

[70] Reedy MC, Bullard B, Vigoreaux JO. Flightin is essential for thick filament assembly and sarcomere stability in Drosophila flight muscles. The Journal of cell biology. 2000;151:1483-500.

[71] Skubiszak L, Kowalczyk L. Myosin molecule packing within the vertebrate skeletal muscle thick filaments. A complete bipolar model. Acta biochimica Polonica. 2002;49:829-40.

[72] Squire JM. Muscle myosin filaments: cores, crowns and couplings. Biophysical reviews. 2009;1:149.

[73] Al-Khayat HA. Three-dimensional structure of the human myosin thick filament: clinical implications. Global cardiology science & practice. 2013;2013:280-302.

[74] Otten E. Optimal design of vertebrate and insect sarcomeres. Journal of Morphology. 1987;191:49-62.

[75] Squire J. The Structural Basis of Muscular Contraction: Springer US; 1981.

[76] Beall CJ, Sepanski MA, Fyrberg EA. Genetic dissection of Drosophila myofibril formation: effects of actin and myosin heavy chain null alleles. Genes & development. 1989;3:131-40.

[77] Hiromi Y, Okamoto H, Gehring WJ, Hotta Y. Germline transformation with Drosophila mutant actin genes induces constitutive expression of heat shock genes. Cell.44:293-301.

[78] Homyk T, Jr., Emerson CP, Jr. Functional interactions between unlinked muscle genes within haploinsufficient regions of the Drosophila genome. Genetics. 1988;119:105-21.

[79] Firdaus H, Mohan J, Naz S, Arathi P, Ramesh SR, Nongthomba U. A cis-regulatory mutation in troponin-I of Drosophila reveals the importance of proper stoichiometry of structural proteins during muscle assembly. Genetics. 2015;200:149-65.

[80] Vigoreaux JO, Hernandez C, Moore J, Ayer G, Maughan D. A genetic deficiency that spans the flightin gene of Drosophila melanogaster affects the ultrastructure and function of the flight muscles. The Journal of experimental biology. 1998;201:2033-44.

[81] Barton B, Ayer G, Heymann N, Maughan DW, Lehmann FO, Vigoreaux JO. Flight muscle properties and aerodynamic performance of Drosophila expressing a flightin transgene. The Journal of experimental biology. 2005;208:549-60.

[82] Emerson CP, Jr., Bernstein SI. Molecular genetics of myosin. Annual review of biochemistry. 1987;56:695-726.

[83] O'Donnell PT, Collier VL, Mogami K, Bernstein SI. Ultrastructural and molecular analyses of homozygous-viable Drosophila melanogaster muscle mutants indicate there is a complex pattern of myosin heavy-chain isoform distribution. Genes & development. 1989;3:1233-46.

[84] George EL, Ober MB, Emerson CP, Jr. Functional domains of the Drosophila melanogaster muscle myosin heavy-chain gene are encoded by alternatively spliced exons. Molecular and cellular biology. 1989;9:2957-74.

[85] Orfanos Z, Sparrow JC. Myosin isoform switching during assembly of the Drosophila flight muscle thick filament lattice. Journal of cell science. 2013;126:139-48.

[86] Wells L, Edwards KA, Bernstein SI. Myosin heavy chain isoforms regulate muscle function but not myofibril assembly. The EMBO journal. 1996;15:4454-9.

[87] Tajsharghi H. Thick and Thin Filament Gene Mutations in Striated Muscle Diseases. International Journal of Molecular Sciences. 2008;9:1259-75.

[88] Suggs JA, Melkani GC, Glasheen BM, Detor MM, Melkani A, Marsan NP, et al. A Drosophila model of dominant inclusion body myopathy type 3 shows diminished myosin kinetics that reduce muscle power and yield myofibrillar defects. Disease models & mechanisms. 2017;10:761-71.

[89] Achal M, Trujillo AS, Melkani GC, Farman GP, Ocorr K, Viswanathan MC, et al. A Restrictive Cardiomyopathy Mutation in an Invariant Proline at the Myosin Head/Rod

Junction Enhances Head Flexibility and Function, Yielding Muscle Defects in Drosophila. Journal of molecular biology. 2016;428:2446-61.

[90] Achal M, Trujillo AS, Melkani GC, Farman GP, Ocorr K, Viswanathan MC, et al. A Restrictive Cardiomyopathy Mutation in an Invariant Proline at the Myosin Head/Rod Junction Enhances Head Flexibility and Function, Yielding Muscle Defects in Drosophila. Journal of molecular biology. 2016;428:2446-61.

[91] Kronert WA, O'Donnell PT, Bernstein SI. A Charge Change in an Evolutionarily-conserved Region of the Myosin Globular Head Prevents Myosin and Thick Filament Accumulation in Drosophila. Journal of molecular biology. 1994;236:697-702.

[92] Collier VL, Kronert WA, O'Donnell PT, Edwards KA, Bernstein SI. Alternative myosin hinge regions are utilized in a tissue-specific fashion that correlates with muscle contraction speed. Genes & development. 1990;4:885-95.

[93] Cripps RM, Suggs JA, Bernstein SI. Assembly of thick filaments and myofibrils occurs in the absence of the myosin head. The EMBO journal. 1999;18:1793-804.

[94] Iwamoto H, Trombitás K, Yagi N, Suggs JA, Bernstein SI. X-ray diffraction from flight muscle with a headless myosin mutation: implications for interpreting reflection patterns. Frontiers in physiology. 2014;5.

[95] Swank DM, Wells L, Kronert WA, Morrill GE, Bernstein SI. Determining structure/function relationships for sarcomeric myosin heavy chain by genetic and transgenic manipulation of Drosophila. Microscopy research and technique. 2000;50:430-42.

[96] Nongthomba U, Cummins M, Clark S, Vigoreaux JO, Sparrow JC. Suppression of muscle hypercontraction by mutations in the myosin heavy chain gene of Drosophila melanogaster. Genetics. 2003;164:209-22.

[97] Kronert WA, O'Donnell PT, Fieck A, Lawn A, Vigoreaux JO, Sparrow JC, et al. Defects in the Drosophila myosin rod permit sarcomere assembly but cause flight muscle degeneration. Journal of molecular biology. 1995;249:111-25.

[98] Salvi SS, Kumar RP, Ramachandra NB, Sparrow JC, Nongthomba U. Mutations in Drosophila myosin rod cause defects in myofibril assembly. Journal of molecular biology. 2012;419:22-40.

[99] White GE, Erickson HP. Sequence divergence of coiled coils--structural rods, myosin filament packing, and the extraordinary conservation of cohesins. Journal of structural biology. 2006;154:111-21.

[100] Taylor KC, Buvoli M, Korkmaz EN, Buvoli A, Zheng Y, Heinze NT, et al. Skip residues modulate the structural properties of the myosin rod and guide thick filament assembly. Proceedings of the National Academy of Sciences of the United States of America. 2015;112:E3806-15.

[101] Yogurtcu ON, Wolgemuth CW, Sun SX. Mechanical response and conformational amplification in α-helical coiled coils. Biophysical journal. 2010;99:3895-904.

[102] Offer G. Skip residues correlate with bends in the myosin tail. Journal of molecular biology. 1990;216:213-8.

[103] Houmeida A, Holt J, Tskhovrebova L, Trinick J. Studies of the interaction between titin and myosin. The Journal of cell biology. 1995;131:1471-81.

[104] Flashman E, Watkins H, Redwood C. Localization of the binding site of the C-terminal domain of cardiac myosin-binding protein-C on the myosin rod. The Biochemical journal. 2007;401:97-102.

[105] Ayer G, Vigoreaux JO. Flightin is a myosin rod binding protein. Cell biochemistry and biophysics. 2003;38:41-54.

[106] Menard L, Nyland L, Vigoreaux J. The Structural and Biomechanical Properties of Insect Thick Filaments Expressing Flightin and Cardiac Myosin Binding Protein-C. Microscopy and Microanalysis. 2013;19:80-1.

[107] Josephson RK, Malamud JG, Stokes DR. Asynchronous muscle: a primer. The Journal of experimental biology. 2000;203:2713-22.

[108] Swank DM, Vishnudas VK, Maughan DW. An exceptionally fast actomyosin reaction powers insect flight muscle. Proceedings of the National Academy of Sciences of the United States of America. 2006;103:17543-7.

[109] Maughan D, Vigoreaux J. Nature's Strategy for Optimizing Power Generation in Insect Flight Muscle. Boston, MA: Springer US; 2005. p. 157-67.

[110] Vigoreaux JO, Saide JD, Valgeirsdottir K, Pardue ML. Flightin, a novel myofibrillar protein of Drosophila stretch-activated muscles. The Journal of cell biology. 1993;121:587-98.

[111] Vigoreaux JO, Perry LM. Multiple isoelectric variants of flightin in Drosophila stretch-activated muscles are generated by temporally regulated phosphorylations. Journal of muscle research and cell motility. 1994;15:607-16.

[112] Tanner BC, Miller MS, Miller BM, Lekkas P, Irving TC, Maughan DW, et al. COOH-terminal truncation of flightin decreases myofilament lattice organization, cross-bridge binding, and power output in Drosophila indirect flight muscle. American journal of physiology Cell physiology. 2011;301:C383-91.

[113] Soto-Adames FN, Alvarez-Ortiz P, Vigoreaux JO. An evolutionary analysis of flightin reveals a conserved motif unique and widespread in Pancrustacea. Journal of molecular evolution. 2014;78:24-37.

[114] Barton B, Vigoreaux J. Novel Myosin Associated Proteins. Nature's Versatile Engine: Insect Flight Muscle Inside and Out. Boston, MA: Springer US; 2006. p. 86-96.

[115] Barton B, Ayer G, Maughan DW, Vigoreaux JO. Site directed mutagenesis of Drosophila flightin disrupts phosphorylation and impairs flight muscle structure and mechanics. Journal of muscle research and cell motility. 2007;28:219-30.

[116] Lemas D, Lekkas P, Ballif BA, Vigoreaux JO. Intrinsic disorder and multiple phosphorylations constrain the evolution of the flightin N-terminal region. Journal of proteomics. 2016;135:191-200.

[117] Henkin JA, Maughan DW, Vigoreaux JO. Mutations that affect flightin expression in Drosophila alter the viscoelastic properties of flight muscle fibers. American journal of physiology Cell physiology. 2004;286:C65-72.

[118] Gasek NS, Nyland LR, Vigoreaux JO. The Contributions of the Amino and Carboxy Terminal Domains of Flightin to the Biomechanical Properties of Drosophila Flight Muscle Thick Filaments. Biology. 2016;5.

[119] Chakravorty S, Tanner BCW, Foelber VL, Vu H, Rosenthal M, Ruiz T, et al. Flightin maintains myofilament lattice organization required for optimal flight power and courtship song quality in Drosophila. Proceedings Biological sciences. 2017;284.

[120] Domsch K, Ezzeddine N, Nguyen HT. Abba is an essential TRIM/RBCC protein to maintain the integrity of sarcomeric cytoarchitecture. Journal of cell science. 2013;126:3314-23.

[121] Clark KA, Bland JM, Beckerle MC. The Drosophila muscle LIM protein, Mlp84B, cooperates with D-titin to maintain muscle structural integrity. Journal of cell science. 2007;120:2066-77.

[122] Clark KA, Kadrmas JL. Drosophila melanogaster muscle LIM protein and alpha-actinin function together to stabilize muscle cytoarchitecture: a potential role for Mlp84B in actin-crosslinking. Cytoskeleton (Hoboken, NJ). 2013;70:304-16.

[123] Clark KA, Lesage-Horton H, Zhao C, Beckerle MC, Swank DM. Deletion of Drosophila muscle LIM protein decreases flight muscle stiffness and power generation. American journal of physiology Cell physiology. 2011;301:C373-82.

[124] Sjoblom B, Salmazo A, Djinovic-Carugo K. Alpha-actinin structure and regulation. Cellular and molecular life sciences : CMLS. 2008;65:2688-701.

[125] Molnár I, Migh E, Szikora S, Kalmár T, Végh AG, Deák F, et al. DAAM is required for thin filament formation and Sarcomerogenesis during muscle development in Drosophila. PLoS genetics. 2014;10:e1004166.

[126] Katzemich A, Kreiskother N, Alexandrovich A, Elliott C, Schock F, Leonard K, et al. The function of the M-line protein obscurin in controlling the symmetry of the sarcomere in the flight muscle of Drosophila. Journal of cell science. 2012;125:3367-79.

[127] Szikora S, Gajdos T, Novák T, Farkas D, Földi I, Lenart P, et al. Nanoscopy reveals the layered organization of the sarcomeric H-zone and I-band complexes. The Journal of cell biology. 2020;219.

[128] Obermann WM, van der Ven PF, Steiner F, Weber K, Fürst DO. Mapping of a myosin-binding domain and a regulatory phosphorylation site in M-protein, a structural protein of the sarcomeric M band. Molecular biology of the cell. 1998;9:829-40.

[129] Obermann WMJ, Gautel M, Weber K, Fürst DO. Molecular structure of the sarcomeric M band: mapping of titin and myosin binding domains in myomesin and the identification of a potential regulatory phosphorylation site in myomesin. The EMBO journal. 1997;16:211-20.

[130] Blondelle J, Marrocco V, Clark M, Desmond P, Myers S, Nguyen J, et al. Murine obscurin and Obsl1 have functionally redundant roles in sarcolemmal integrity, sarcoplasmic reticulum organization, and muscle metabolism. Communications Biology. 2019;2:178.

[131] Burkart C, Qiu F, Brendel S, Benes V, Haag P, Labeit S, et al. Modular proteins from the Drosophila sallimus (sls) gene and their expression in muscles with different extensibility. Journal of molecular biology. 2007;367:953-69.

[132] Spletter ML, Barz C, Yeroslaviz A, Zhang X, Lemke SB, Bonnard A, et al. A transcriptomics resource reveals a transcriptional transition during ordered sarcomere morphogenesis in flight muscle. eLife. 2018;7.

[133] Saide JD, Chin-Bow S, Hogan-Sheldon J, Busquets-Turner L, Vigoreaux JO, Valgeirsdottir K, et al. Characterization of components of Z-bands in the fibrillar flight muscle of Drosophila melanogaster. The Journal of cell biology. 1989;109:2157-67.

[134] Vigoreaux JO, Saide JD, Pardue ML. Structurally different Drosophila striated muscles utilize distinct variants of Z-band-associated proteins. Journal of muscle research and cell motility. 1991;12:340-54.

[135] Royuela M, Fraile B, De Miguel MP, Cervera M, Paniagua R. Immunohistochemical study and western blotting analysis of titin-like proteins in the striated muscle of Drosophila melanogaster and in the striated and smooth muscle of the oligochaete Eisenia foetida. Microscopy research and technique. 1996;35:349-56.

[136] Daneshparvar N, Taylor DW, O'Leary TS, Rahmani H, Yeganeh FA, Previs MJ, et al. CryoEM Structure of Drosophila Flight Muscle Thick Filaments at 7Å Resolution. bioRxiv. 2020:2020.06.05.136580.

[137] Qadota H, Mayans O, Matsunaga Y, McMurry JL, Wilson KJ, Kwon GE, et al. The SH3 domain of UNC-89 (obscurin) interacts with paramyosin, a coiled-coil protein, in Caenorhabditis elegans muscle. Molecular biology of the cell. 2016;27:1606-20.

[138] Becker KD, O'Donnell PT, Heitz JM, Vito M, Bernstein SI. Analysis of Drosophila paramyosin: identification of a novel isoform which is restricted to a subset of adult muscles. The Journal of cell biology. 1992;116:669-81.

[139] Maroto M, Arredondo J, Goulding D, Marco R, Bullard B, Cervera M. Drosophila paramyosin/miniparamyosin gene products show a large diversity in quantity, localization, and isoform pattern: a possible role in muscle maturation and function. The Journal of cell biology. 1996;134:81-92.

[140] Maroto M, Arredondo JJ, San Roman M, Marco R, Cervera M. Analysis of the paramyosin/miniparamyosin gene. Miniparamyosin is an independently transcribed, distinct paramyosin isoform, widely distributed in invertebrates. The Journal of biological chemistry. 1995;270:4375-82.

[141] Levine RJ, Kensler RW, Yang Z, Stull JT, Sweeney HL. Myosin light chain phosphorylation affects the structure of rabbit skeletal muscle thick filaments. Biophysical journal. 1996;71:898-907.

[142] Levine RJC, Yang Z, Epstein ND, Fananapazir L, Stull JT, Sweeney HL. Structural and Functional Responses of Mammalian Thick Filaments to Alterations in Myosin Regulatory Light Chains. Journal of structural biology. 1998;122:149-61.

[143] Chakravorty S. Role of the Drosophila Melanogaster Indirect Flight Muscles in Flight and Male Courtship Song: Studies on Flightin and Mydson Light Chain - 2. Graduate College Dissertations and Theses: University of Vermont; 2013.

[144] Tune TC, Ma W, Irving TC, Sponberg S. A nanometer difference in myofilament lattice spacing of two cockroach leg muscles explains their different functions. bioRxiv. 2019:656272.

[145] Qiu F, Brendel S, Cunha PM, Astola N, Song B, Furlong EE, et al. Myofilin, a protein in the thick filaments of insect muscle. Journal of cell science. 2005;118:1527-36.

[146] Schnorrer F, Schonbauer C, Langer CC, Dietzl G, Novatchkova M, Schernhuber K, et al. Systematic genetic analysis of muscle morphogenesis and function in Drosophila. Nature. 2010;464:287-91.

[147] Patel SR, Saide JD. Stretchin-klp, a novel Drosophila indirect flight muscle protein, has both myosin dependent and independent isoforms. Journal of muscle research and cell motility. 2005;26:213-24.

[148] Zou S, Meadows S, Sharp L, Jan LY, Jan YN. Genome-wide study of aging and oxidative stress response in Drosophila melanogaster. Proceedings of the National Academy of Sciences of the United States of America. 2000;97:13726-31.

[149] Suggs JA, Cammarato A, Kronert WA, Nikkhoy M, Dambacher CM, Megighian A, et al. Alternative S2 hinge regions of the myosin rod differentially affect muscle function, myofibril dimensions and myosin tail length. Journal of molecular biology. 2007;367:1312-29.

[150] Miller MS, Dambacher CM, Knowles AF, Braddock JM, Farman GP, Irving TC, et al. Alternative S2 hinge regions of the myosin rod affect myofibrillar structure and myosin kinetics. Biophysical journal. 2009;96:4132-43.

[151] Doering DS, Matsudaira P. Cysteine Scanning Mutagenesis at 40 of 76 Positions in Villin Headpiece Maps the F-Actin Binding Site and Structural Features of the Domain. Biochemistry. 1996;35:12677-85.

[152] Vardar D, Buckley DA, Frank BS, McKnight CJ. NMR structure of an F-actin-binding "headpiece" motif from villin. Journal of molecular biology. 1999;294:1299-310.

[153] Sarov M, Barz C, Jambor H, Hein MY, Schmied C, Suchold D, et al. A genome-wide resource for the analysis of protein localisation in Drosophila. eLife. 2016;5:e12068.

[154] Maier D. The evolution of transcriptional repressors in the Notch signaling pathway: a computational analysis. Hereditas. 2019;156:5.

[155] McGrath MJ, Cottle DL, Nguyen MA, Dyson JM, Coghill ID, Robinson PA, et al. Four and a half LIM protein 1 binds myosin-binding protein C and regulates myosin filament formation and sarcomere assembly. The Journal of biological chemistry. 2006;281:7666-83.

[156] Chu M, Gregorio CC, Pappas CT. Nebulin, a multi-functional giant. The Journal of experimental biology. 2016;219:146-52.

[157] Linke WA. Titin Gene and Protein Functions in Passive and Active Muscle. Annual review of physiology. 2018;80:389-411.

[158] Hu DH, Matsuno A, Terakado K, Matsuura T, Kimura S, Maruyama K. Projectin is an invertebrate connectin (titin): isolation from crayfish claw muscle and localization in crayfish claw muscle and insect flight muscle. Journal of muscle research and cell motility. 1990;11:497-511.

[159] Nave R, Weber K. A myofibrillar protein of insect muscle related to vertebrate titin connects Z band and A band: purification and molecular characterization of invertebrate mini-titin. Journal of cell science. 1990;95 (Pt 4):535-44.

[160] Ayme-Southgate A, Saide J, Southgate R, Bounaix C, Cammarato A, Patel S, et al. In indirect flight muscles Drosophila projectin has a short PEVK domain, and its NH2-terminus is embedded at the Z-band. Journal of muscle research and cell motility. 2005;26:467-77.

[161] Weitkamp B, Jurk K, Beinbrech G. Projectin-thin filament interactions and modulation of the sensitivity of the actomyosin ATPase to calcium by projectin kinase. The Journal of biological chemistry. 1998;273:19802-8.

[162] Lakey A, Labeit S, Gautel M, Ferguson C, Barlow DP, Leonard K, et al. Kettin, a large modular protein in the Z-disc of insect muscles. The EMBO journal. 1993;12:2863-71.

[163] Bullard B, Burkart C, Labeit S, Leonard K. The function of elastic proteins in the oscillatory contraction of insect flight muscle. Journal of muscle research and cell motility. 2005;26:479-85.

[164] Kulke M, Neagoe C, Kolmerer B, Minajeva A, Hinssen H, Bullard B, et al. Kettin, a major source of myofibrillar stiffness in Drosophila indirect flight muscle. The Journal of cell biology. 2001;154:1045-57.

[165] Wagner GP, Altenberg L. Perspective: Complex Adaptations and the Evolution of Evolvability. Evolution. 1996;50:967-76.

CHAPTER 2 JOURNAL ARTICLE

The Structural and Biomechanical Properties of Insect Thick Filaments Expressing Flightin and Cardiac Myosin Binding Protein-C

DOI:10.1017/S1431927613002390

Lynda Menard[1*], Lori Nyland[2], Jim Vigoreaux[1]

Department of Biology[1] and Molecular Physiology and Biophysics[2]

University of Vermont, Burlington, Vermont

* Corresponding author. Address:

Department of Biology,

University of Vermont,

120 Marsh Life Science Bldg,

109 Carrigan Drive, Burlington, VT 05405,

USA.

Fax: +1 802 656 2914.

E-mail address: lmmenard@uvm.edu (Lynda Menard).

Cardiac myosin binding protein-c (cMyBP-C) of mammalian cardiac muscle and flightin (FLN) of invertebrate indirect flight muscle (IFM) have been shown to contribute to thick filament stiffness, as determined by calculations of persistence length (PL), an index of flexural rigidity [1, 2] in their corresponding muscle systems. FLN and cMyBP-C *in vitro* bind to a common site in the coiled-coil region of myosin II, and both proteins are known to be regulated by phosphorylation [3, 4]. To test the hypothesis that FLN and cMyBP-C are functionally homologous, we have determined the extent to which cMyBP-C can rescue the phenotypes manifested in the *Drosophila* FLN knockout strain fln^0. Structural characteristics of flight muscle sarcomeres were analyzed by transmission electron microscopy (TEM) and the contour and end-to-end length of isolated, hydrated native thick filaments was measured by atomic force microscopy (AFM).

Experiments were carried out on four *D. melanogaster* mutant and transgenic strains: (i) FLN knockout strain (fln^0), (ii) a knockout rescued transgenic strain ($fln^0;fln^+$), (iii) a transgenic cMyBP-C strain without FLN expression ($fln^0;cMyBPC^+$), and (iv) a transgenic strain with FLN expression alongside cMyBP-C expression ($fln^+;cMyBPC^+$). In preparation for TEM, thoraces from newly eclosed (<1 hour) *D. melanogaster* were bisected, fixed, dehydrated, infiltrated, embedded, sectioned and imaged by TEM [5]. The length of sarcomeres from 4-5 flies for each *Drosophila* strain was measured using ImageJ. AFM data of isolated thick filaments were evaluated using the parameters and programs described by [6]. Statistical analysis was done using JMP 9 software.

The TEM results confirmed both the sarcomere length measurements and level of structural order previously seen for fln^0 and $fln^0;fln^+$, while revealing shorter sarcomeres in the transgenic lines involving cMyBP-C alone (**Fig. 2-1, Table 2-1**). When cMyBP-C is expressed alongside FLN, sarcomere length is slightly but significantly longer than sarcomere length in the control $fln^0;fln^+$. These results support the idea of cMyBP-C binding to myosin in thick filaments of *D. melanogaster* and influencing the length of the filaments. However, the length regulation exerted by cMyBP-C is surpassed by FLN when FLN is present, either by direct binding competition to a common myosin binding site or another regulatory mechanism.

The PL for $fln^+;cMyBPC^+$ obtained by AFM was significantly higher than PL for $fln^0;fln^+$ (2.56mm compared to 1.67mm; $p<0.05$), suggesting that cMyBP-C contributes to filament stiffness when expressed ectopically in IFM. However, the cMyBP-C effect is seen only in the presence of FLN as PL of $fln^0;cMyBPC^+$ was not different that PL of fln^0. Our observations suggest that the presence of FLN influences the effects that cMyBP-C has on the mechanical properties of the thick filaments. This may possibly be due to FLN stabilizing the thick filaments to permit a more ideal environment for cMyBP-C binding. From these studies we conclude that ectopically expressed cMyBP-C influence sarcomere length and thick filament stiffness in the flight muscle, consistent with the hypothesis that cMyBP-C and FLN have convergent functions.

The mechanical AFM data complemented the structural TEM data in that it promoted the possibility that competition between the two proteins was likely occurring in the *fln*$^{+}$*;cMyBPC*$^{+}$ line. While FLN and cMyBP-C have been shown to bind to a common myosin site *in vitro*, a perspective supported by these AFM/EM observations, cMyBP-C is a very different protein compared to FLN in sequence and size (130 kDa vs. 20kDa) [1, 7]. Our results demonstrate that cMyBP-C may be adjusting structural and mechanical characteristics of myosin thick filaments in the same way as FLN, but not to the same degree.

Our data supports the hypothesis of cMyBP-C being a vertebrate functional homolog to invertebrate FLN. This brings us closer to understanding the role of myosin binding proteins in dictating the structural and mechanical properties of thick filaments, an important determinant of muscle functional properties. Further insight can be gained by elucidation of the molecular interaction between these proteins and the myosin coiled coil.

FIGURES

Figure 2-1: TEM displaying sarcomere morphologies.

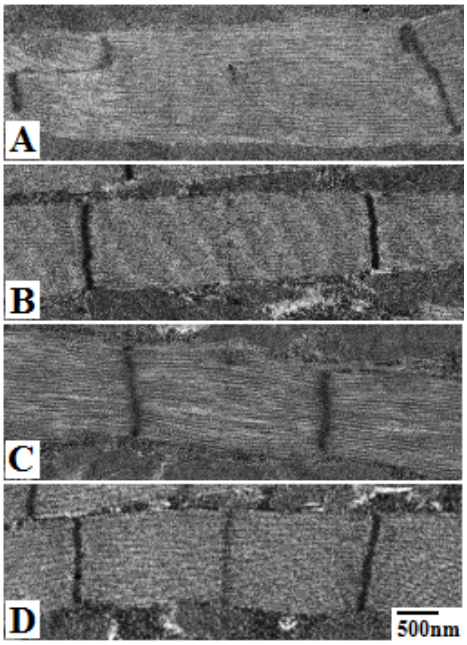


Figure 2-1. TEM displaying sarcomere morphologies of lines: fln^{0} (A), $fln^{0};fln^{+}$ (B), $fln^{0};MyBPC^{+}$ (C), and $fln^{+};MyBPC^{+}$ (D). 2500x magnification at 60kV. Scale bar is at 500 nm.

TABLES

Table 2-1: Sarcomere and Persistence Length Measurements

Table 2-1: Sarcomere and Persistence Length Measurements

Transgenic Strain	Mean Sarcomere Length (μM) ± SEM	Mean Persistence Length (mM) ± SEM
fln^0	3.65 ± 0.03 (n = 141)	0.98 ± 0.14^ (n = 105)
fln^0;fln$^+$	3.17 ± 0.01* (n = 436)	1.67 ± 0.27# (n = 26)
fln^0;cMyBPC$^+$	2.34 ± 0.02* (n = 545)	0.85 ± 0.21^ (n = 44)
fln$^+$;cMyBPC$^+$	3.30 ± 0.01* (n = 985)	2.56 ± 0.22* (n = 40)

*statistically distinct from *fln^0* and all other groups ($p<0.05$)
^statistically distinct from *fln^0;fln$^+$* ($p<0.05$)
#statistically distinct from *fln^0;cMyBPC+* ($p<0.05$)

BIBLIOGRAPHY

[1] Ayer G, Vigoreaux JO. Flightin is a myosin rod binding protein. Cell biochemistry and biophysics. 2003;38:41-54.

[2] Nyland LR, Palmer BM, Chen Z, Maughan DW, Seidman CE, Seidman JG, et al. Cardiac myosin binding protein-C is essential for thick-filament stability and flexural rigidity. Biophysical journal. 2009;96:3273-80.

[3] Vigoreaux JO. Alterations in flightin phosphorylation in Drosophila flight muscles are associated with myofibrillar defects engendered by actin and myosin heavy-chain mutant alleles. Biochem Genet. 1994;32:301-14.

[4] Coulton AT, Stelzer JE. Cardiac myosin binding protein C and its phosphorylation regulate multiple steps in the cross-bridge cycle of muscle contraction. Biochemistry. 2012;51:3292-301.

[5] Barton B, Ayer G, Heymann N, Maughan DW, Lehmann FO, Vigoreaux JO. Flight muscle properties and aerodynamic performance of Drosophila expressing a flightin transgene. The Journal of experimental biology. 2005;208:549-60.

[6] Kreplak L, Nyland LR, Contompasis JL, Vigoreaux JO. Nanomechanics of Native Thick Filaments from Indirect Flight Muscles. Journal of molecular biology. 2009;386:1403-10.

[7] Miyamoto CA, Fischman DA, Reinach FC. The interface between MyBP-C and myosin: site-directed mutagenesis of the CX myosin-binding domain of MyBP-C. Journal of muscle research and cell motility. 1999;20:703-15.

The authors acknowledge the funding of NSF grant MCB 1050834 to J.Vigoreaux.

CHAPTER 3 JOURNAL ARTICLE (Extended Version)

Secondary structure of the novel myosin binding domain WYR and implications within myosin structure

Lynda Menard*, Neil Wood and Jim O. Vigoreaux

Department of Biology

The University of Vermont

* Corresponding author. Address:

Department of Biology,

University of Vermont,

120 Marsh Life Science Bldg,

109 Carrigan Drive, Burlington, VT 05405,

USA.

Fax: +1 802 656 2914.

E-mail address: lmmenard@uvm.edu (Lynda Menard).

ABSTRACT

Structural changes in the myosin II light meromyosin (LMM) that influence thick filament mechanical properties and muscle function are modulated by LMM-binding proteins. Flightin (fln) is an LMM-binding protein indispensable for the function of Drosophila indirect flight muscle (IFM). Fln has a three domain structure that includes WYR, a novel 52 aa domain conserved throughout Pancrustacea. In this study we (1) test the hypothesis that WYR binds the LMM, (2) characterize the secondary structure of WYR, and (3) examine the structural impact WYR has on the LMM. Circular dichroism at 260-190 nm reveals a structural profile for WYR and supports an interaction between WYR and LMM. A WYR-LMM interaction is supported by cosedimentation with a stoichiometry of ~2.4:1. The WYR-LMM interaction results in an overall increased coiled-coil content while curtailing alpha helical content. WYR is found to be composed of 15% turns, 31% antiparallel beta, and 48% 'other' content. A hypothetical structure of WYR including an antiparallel beta hairpin between Q92-K114 centered on an ASX or beta turn around N102, with a G1 bulge at G117, is proposed. The Drosophila LMM segment used, V1346-I1941, encompassing conserved skip residues 2-4, is found to possess a traditional helical profile but is interpreted as having <30% helical content by multiple methods of deconvolution. This is the largest segment of the Drosophila LMM characterized by CD and this low helicity may be affiliated with dynamic behavior of the structure in solution or inclusion of a known non-helical region in the C-terminus. Our results support the hypothesis that WYR binds the LMM and that this interaction brings

about structural changes in the coiled-coil. These studies implicate fln, via the WYR domain, for distinct shifts in LMM secondary structure that could influence structural properties and stabilization of the thick filament, scaling to modulation of whole muscle function.

INTRODUCTION

The insect indirect flight muscle (IFM) is highly ordered, stretch activated and known to propel wing beats up to 1000 times per second [1, 2]. The IFM of *Drosophila melanogaster* has been a valuable informant to structure-function relationships of myofibrillar proteins, mechanical parameters modulated for stretch activation, and as a model for investigating the molecular underpins of muscle and cardiac diseases [3-7]. Drosophila IFM has also been utilized for the discovery and characterization of novel contractile proteins necessary to tune the structural and viscoelastic properties to optimize function [7]. A well-studied example is flightin (UniProtKB - P35554), a 20 kDa protein shown to be a myosin-binding component of the thick filament by genetic, biochemical, and structural studies [8-11]. While in *D. melanogaster* flightin is exclusive to the IFM, its wide-raging presence in hexapods and crustaceans (Pancrustacea, *sensu stricto*) and its deep evolutionary history suggest a broader role in muscles, and perhaps in other tissues [12] A comparative sequence analysis of flightin revealed a tripartite organization characterized by a 52 amino acid conserved domain in Drosophila (from H84 to T135,

denoted as WYR) that dates to the origin of Pancrustacea, flanked at the N-terminal and C-terminal sides by less conserved regions of variable length [12].

The analyses of flightin mutants in *D. melanogaster* have revealed its important role in flight and courtship, two behaviors that underscore the evolutionary success and prolific speciation of insects [7, 13, 14]. Specifically, mutants that express a truncated flightin missing the C-terminal region ($fln^{\Delta C44}$) are incapable of generating a courtship song or a wing beat to propel flight, similar to the flightin null mutant, fln^{0}. In contrast, mutants that express a truncated flightin missing the N-terminal region ($fln^{\Delta N62}$) have impaired flight mechanics and produce an abnormal courtship song that lessens the male's mating success [14]. Despite the truncations, the mutant flightin are integral components of the fiber indicating that neither the N-terminal or C-terminal region is necessary for flightin incorporation into the thick filament. In this study, we test the hypothesis that the conserved WYR domain harbors a myosin rod binding sequence.

Flightin has been shown to deeply impact the stability, structure and organization of IFM thick filaments, sarcomeres, myofibrils, and fibers. Fibers in mutants lacking flightin (fln^{0}) bunch-up upon eclosion [11] and fray and break when exposed to rigor regardless to prior exposure to mechanical activation [15]. Sarcomeres of the fln^{0} pupal IFM are 25-30% longer than normal and there are fewer thick filaments across the diameter of myofibrils (17-19 vs 25-26) [11]. The C-terminal mutant $fln^{\Delta C44}$ exhibits slightly shorter (~9%) sarcomeres with dispersed or absent M-lines and a myofilament lattice that is less

ordered and more compact (i.e., reduced inter-thick filament spacing) compared to a transgenic null rescued control line [16, 17]. The N-terminal mutant $fln^{\Delta N62}$ also exhibits slightly shorter sarcomeres that lack an evident H-zone and have a narrower M-line. Myofibrils contain more thick filaments and a more compact and less "crystalline" (regular) myofilament lattice [14, 18].

IFM fibers that either lack flightin (fln^{0}) or express a mutant form ($fln^{\Delta N62}$ and $fln^{\Delta C44}$) manifest a variety of mechanical defects that include alterations in cross-bridge cycling kinetics, in viscous and elastic moduli, and in power output [14-16]. Native IFM thick filaments isolated from these mutant strains more clearly delineate the complex role that flightin plays in dictating thick filament structure, integrity and mechanical properties [18, 19]. While the absence of flightin results in thick filaments that are substantially longer (and more fragile) than normal, truncation of the N-terminal region has no effect on thick filament length while truncation of the C-terminal region results in shorter thick filaments. The absence of flightin also results in a significant decrease in filament stiffness (persistence length) with the N-terminal region making a larger contribution to stiffness than the C-terminal region.

Decreased accumulation of flightin is found in mutants of the myosin rod Mhc^{6}, R1559H and Mhc^{13}, E1554K [20] suggestive that these residues, or influential interactions in the local light meromyosin (LMM) coiled-coil, may be part of an LMM flightin binding site, as supported by *in vitro* studies [8]. Here, we use circular dichroism to test the hypothesis

that WYR binds myosin LMM and to characterize the secondary structural changes associated with the C-terminal 600 amino acids of the LMM, which encompass the residues R1559 and E1554, engaging with WYR. This study also represents the first experimental structural characterization for the novel domain, WYR. The WYR:LMM interaction is put into context by further analysis of the Lethocerus IFM thick filament cryo-EM model structure that includes non-myosin densities along the length of the thick filament [21].

WYR SEQUENCE AND STRUCTURE

The fully conserved residues, throughout Pancrustacea, of WYR, H84-T135, include W85, R87, Y93, Y103, Y104, P123, E130, R131 (**Fig. 3-1**). Many of these completely conserved residues are aromatic and many other residues in the sequence that are highly conserved are also aromatic. WYR from *Drosophila melanogaster* flightin contains an additional 6 aromatic residues, most of which are clustered together and among positively charged residues. Several highly conserved prolines (P88, P123, P125), one of which is completely conserved (P123) and another completely conserved in hexapods and Branchiopods (P88), are also present in the sequence. While it is not part of the WYR sequence, it should be noted that a series of prolines precede the WYR sequence in Drosophila and in other organisms.

As aromatic character, particularly from the side groups of tryptophan and tyrosine, and conformationally restricted prolines are prevalent in WYR structure, it is likely that these residues are critical for its secondary structure. In addition to favorable hydrophobic effect, proline and aromatic residues are capable of interacting with each other by an interaction between the negative π face of the aromatic side chain and the positively polarized C-H bonds of proline. Such an interaction is called a CH/π interaction and is a type of nonpolar hydrogen bonding. Tryptophan and tyrosine are more likely to be involved in stabilization of cis-amide bonds than phenylalanine [22]. Pi-cation interactions should also be considered as tryptophan and tyrosine have the propensity to form these kinds of bonds with lysine or arginine [23] of which there are multiple in the WYR sequence, especially if tyrosine's OH is additionally involved in polar H-bonding.

CD profile of WYR

The CD profile of WYR is characterized by a near positive band at 190 nm, a negative band at 200 nm and a shoulder centered ~220 nm with the overall profile residing in the positive range >230 nm (**Fig. 3-2**). This profile is predicted to be predominantly composed of combination of 'other' (48%), antiparallel β (31%), and turn structures (15%) via BeStSel (**Fig. 3-3**) and is most reminiscent of right-twisted antiparallel β strands and unordered, irregular structures [24] that can be further dissected.

Increasing WYR concentration leads to changes in the spectra profile at nm <200, decreasing the magnitude of the negative band and shifting it to 205 nm, plateauing and apparent splitting of the 190 nm band to 200 nm while retaining the shoulder around 220 nm (**Fig. 3-4**). The turn content is not predicted to change at the highest concentration (40 μM) but antiparallel β content and 'other' content experience decreases and helical content is shown to increase (**Fig. 3-5**).

A notable characteristic of the WYR CD profile is the lack of distinct aromatic bands in the far UV (data not shown). Exciton coupling between the π-π* transitions in aromatics occur when they are in close proximity and this gives rise to distinct band patterns [25, 26]. The aromatic contribution has been considered to be "idiosyncratic" among proteins [27] and could be masked by periodicity in the peptide backbone exhibiting dominant bands in an overlapping region. Both ionization of the aromatic groups and hydrogen bonding, along with other interactions can change the absorption profile for tyrosine and tryptophan [28-30], the most prevalent aromatic residues of WYR. The absence of far UV band patterns is suggestive that the aromatic residues of WYR are engaging in contacts that either ablate or substantially alter the presence and orientation of aromatic π-π* transitions. This is valuable information when considering the driving factors and possible inter-residue contacts involved with WYR structure and its behavior in the presence of ionic solvents and binding partners.

The WYR structure, broken down into antiparallel β sheet, turn and 'other' content can be further elucidated by combining information from its CD profile with examination of the primary sequence to provide insight into specific regions likely to be responsible for these structural segments. Model β sheets are generally characterized by a negative band at 218 nm and positive band between 190–220 nm [31, 32], with more highly twisted β sheets exhibiting more intense bands.

Beta rich proteins have been further characterized into β_I- and β_{II}- types. β_{II} exhibit a poly(Pro)II-type (P_2)-like profile [32] in which the positive band generally associated with β sheets is countered by the negative band from P_2 content. P_2 and β_{II} profiles are frequently characterized as having high 'other' content as P_2 is associated with having the character of non-periodic or denatured proteins. β_{II}-types have a characteristic 200 nm negative band, some with a small positive band around 190 nm and negative shoulder around 220 nm. While the WYR sequence contains several prolines, P_2-type structure is characterized by (φ, ψ) of approximately -70^o, $+150^o$ and doesn't require prolines. β_{II} structure is characterized by >15% of P_2 structure alongside <40% β sheet content. The fraction of P_2 to β content must be >0.4. The β_{II}-type profile fits WYR well and 30% β content would imply at least 12% P_2 content which would be 6-7 residues, though there may be more than 12% of the WYR structure adopting such a format.

We will now examine the probable location and nature of these components in the context of both the CD profile of WYR and the character of constituent residues.

Antiparallel Beta Content

Antiparallel β sheets differ from parallel β sheets by the positive band at ~190 nm [33] while parallel is negative at this wavelength. A semblance of this is exhibited by WYR structure when the mixture of P_2 content is considered. The simplest antiparallel β sheet is frequently considered the β hairpin where two antiparallel strands are connected via a short turn, of which there are three dominant types, 2:2, 3:5, and 4:4 [34, 35] with 2:2 and 3:5 hairpins being strongly preferred for right-twisted antiparallel sheets [36, 37]. The first numerical designation reflect the number of residues involved in the typical binding pattern of β sheets in which either the amide or carboxyl is involved in H-bonding and the second reflects the number of residues over which both the amide and carboxyl are participating in H-bonding. As the WYR sequence is short and dominated by 'other' 2^o structure, it is most likely that the antiparallel component is in the form of a β hairpin.

The antiparallel β character of WYR is similar to some amyloid proteins and may inform consideration of WYR multimerization under certain conditions. The periodic elements of antiparallel β strands and sheets arise from 3 aa, 6.5-7 Å, long repeats and hydrogen bonding between adjacent strands in which the Cα-Cα distance between H-bonded pairs is 5.3 Å [38]. In fibrous structures, these are usually parallel to the fibril axis but are perpendicular in amyloid fibrils and denoted "cross-β" [39]. CD profile of WYR is

similar to CD profiles of amyloid structures, in which there is a plateau ending at ~210 nm and minima at ~200 nm, purported to be a combination of random coil/β structure [40, 41]; WYR β content may be capable of forming similar cross-β fibrils under some conditions.

Turns & Loops

Turns are composed of n residues over which the distance between the residue α carbons must be shorter than 7.5 Å. Types of turns include γ-turns (n=3), β-turns (n=4), α-turns (n=5) and π-turns (n=6) [42]. Most turns are associated with typical φ and ψ angles for each participating residue that are allowed to vary by up to 30^{o}, unlike the repeating peptide conformations of β sheets and α helices that are associated with more specific φ and ψ angles and H-bonding patterns. Although turns are described as showing preference for hydrophilic amino acids as they are frequently on the outside of proteins, aromatic residues such as tyrosine and tryptophan have been known to 'protect' the turns from solvation [37, 43].

Beta turns are the most common type of nonrepetitive secondary structure [44]. They are classically defined as possessing hydrogen bonding between the carbonyl of the residue at position i and NH of i+2 or i+3 with a distance of <7 Å between residues i and i+3 [45]. They are easy to form and break and are considered unstable unless further stabilized by side chain interactions [36]. There are 8 types of β turns with type VI being characterized by the presence of proline. Type III and III' are similar to the 3_{10} helix with

a repeating φ of -60° and ψ of -30°. Type I, I' and IV turns are most likely to be involved in double turns with double type I being the most common and similar to two turns of a 3_{10} helix [46]. Beta turns are often a precursor for β-hairpin structure with the turn type playing a role in the twist exhibited in the strands. Type I' and II' are most conducive to right-twisted β strands [36, 37], the conformation indicated by CD/BeStSel. Beta type II and II' structure specifically has preference for prolines and tyrosines as well as lysines within the i to i+3 positions. Beta turns generally contain hydrophilic amino acids as they are frequently located externally on a folded protein. As antiparallel β hairpin structures strongly prefer type I' or II' turns, with these turns being not overly reliant on the presence of proline and glycine, this kind of β turn content is more likely to be present within the antiparallel β sheet of WYR, as the conserved prolines are towards the N- and C-terminal edges of the peptide and not proximal to glycine.

NetTurnP is a program that evaluates the propensity of a sequence, based off of the individual amino acid propensities and would-be associations (e.g. i, i+3, etc) with nearby residues, to possess different types of β turns [47]. The prediction along the length of the WYR sequence can be seen in **Fig. 3-6**. Type II and Via1 and Vib are considered the most probable within the region of greatest β turn likelihood. Types Via1 and Vib are characterized by a proline being in the cis conformation in either position 2 or 3 of the turn [48] and this prediction is driven by the positioning of P123 and P125. Type IV shares a similar pattern with the general β turn propensity pattern but Type IV is unlikely, despite the presence of prolines, as an X-pro cis amide bond is required between the i+1

and i+2 residues with X likely being an aromatic residue [22]; the prolines of WYR are not immediately neighboring the conserved aromatic residues.

Gamma-turns are defined by H-bonding between the carboxyl of residue i and the amine of the i+2 residue and come in 'classical' and 'inverse' varieties. The 'inverse' variant makes up the bulk of γ turns found in proteins though only the 'classical' variant can give rise to the chain reversal necessary in an antiparallel β hairpin [49]. Gamma-turns can form very rare 2.2_7 ribbons when consecutive [50, 51] but are not very stable on their own unless enclosed in Ω loops ($n \geq 6$) in pseudo-cyclic structures. Like β turns, γ turns can be protected by the presence of aromatic residues [52]. A 2.2_7 ribbon is unlikely to form with the variable structure of WYR but it is possible that a γ-type turn is present. If a γ turn is present it would likely be the 'classic' type (for i+1, φ +75±40; ψ = -64±40) if participating in the turn of the β hairpin, or exist as part of loop structures.

It is worth noting that loops and turns can incorporate Poly(Pro)II-type angles (-70°, +150°) and the first residue of a β type-II turn is known to possess this conformation. As WYR has a β_{II}-like spectra and is still more likely to have β type II' or I' turns; this orientation, if present, is likely within longer loop/turn-like structures unaffiliated with antiparallel β sheet.

Helical Content

Alpha helices, also known as 3.6_{13}, in which the average number of residues per turn are given with the number of atoms in between hydrogen bonding denoted in subscript, are characterized by i, i+4 hydrogen bonding, and are the most well-defined of the periodic structures probed by circular dichroism. They are associated with a distinct positive band (~190-195 nm) two negative bands at 208 nm and 222 nm that rarely vary. Some differently ordered helical content is categorized as 'Other' in BeStSel interpretations (e.g. 3_{10} and π helices, or 4.4_{16}) [24]. Although single α helical (SAH) domains exist in proteins, they are predominantly found stabilized by tertiary structure [53], as the residue side chains are otherwise completely exposed to solvent.

There is consensus among secondary structure prediction programs (Jpred, RaptorX2, Phyre2, and i-tasser) that the WYR region Y104-K114 is helical. This would represent ~19% of the WYR sequence. The portion of the sequence that precedes it contains alternating tyrosines which would be unfavorable for helical structure, while the portion that follows Y104-K114 contains a number of helix breakers (glycine, prolines) and a bulky tryptophan that is, likewise, unfavorable for helical structure. Y104-K114, while not normally helical in WYR at the primary concentration we've used for experiments (10 μM), may be the region over which some helicity develops at higher concentrations. Y104-K114 may also shift to a helical format upon changes in the solvent or in the presence of a binding partner, conditions that may provide stabilization of an induced helix.

BeStSel predicts some distorted helix content, 'Helix2' that may be present in the region of consensus among prediction programs based strictly on primary sequence, perhaps with conformational changes that take place. However, since multiple residues are involved for a helix (>3) and 3 of the 52 WYR residues would be >5% of the total structure, the low content predicted in the baseline WYR structure is likely to be from non-helical contributions. This could be due to confounding spectra from β content but may also be due to helical-like turns contributing to Helix2 (disordered helix) content. While 3_{10} helical-like orientations, categorized as 'Other' by BeStSel, may be present in turn/loop content, it's unlikely that Ω- (4_{13}) or π- (4.4_{16}) helical content exists in WYR as this would be expected to be present over longer spans or among a more ordered α helix.

Considerations for 'Other' structural elements

Protein loops – of which more structured short variants are considered turns – have the planes of the peptide bonds frequently forming angles of 90^{o} [54]. While not ideal for classical backbone hydrogen bonding, it is the ideal angle for aromatic cation-π or aromatic-aromatic interaction. Notably, the peptide bond carboxyl and nitrogen can compose a π-plane that can accept the hydrogen of a separate amide for polar hydrogen bonding. The energy of the H-π interaction increases with the size of the conjugate producing the π cloud, hence interaction energies are higher when this same propensity is at play in aromatic residues as they have a much larger π system [54]. These types of structures would fall under the 'Others' category as they are aperiodic or have short

segments of unusual periodicity. Given that 'Others' accounts for nearly half of the WYR structure, their importance for WYR function merits further studies.

Unlike extended periodic structures and turns, loops are not defined by their hydrogen-bonding pattern and tend to have fewer hydrogen bonds and so have been categorized by their 'flatness' described by geometrical relationships in the x-y direction and z-y direction [55]. Strap loops are on the more linear end of the spectra while Ω loops are represented on the non-linear end. Loops tend to exist in protein regions that vary in sequence between isoforms and between orthologs, but can, nevertheless, be responsible for enabling important secondary, tertiary structure or binding specificity. For instance, Ω loops have been found to form an "lid" that operates about a hinge to modulate accessibility of a binding or catalytic domain in enzymes such as tryptophan synthase, phosphoenolpyruvate carboxykinase (PEPCK), and triose phosphate isomerase (TPI/TIM) [55-57].

Additional insights and directions

Beta content of WYR further examined

While programs that predict secondary structure from primary sequence have suggested WYR to be predominantly α helical or random coil, a deeper look into the character of the prevalent and conserved residues make the high β content shown by CD unsurprising. Highly conserved residues found in WYR, including the canonical W, Y, (**Fig. 3-1**) along

with F, I, V, and T [12], are known to prefer β strand conformations over α helical [58]. There are two alanines, known for being pro-helical [59], in WYR of Drosophila flightin. A129 is well conserved and located between highly conserved residues W128 and E130. When not conserved, A129 is often replaced by a glycine [12]. The other, A119, represents the least conserved of all residues of WYR, even among Drosophila species, and is frequently replaced by a serine, threonine, aromatic or proline. The next least-conserved residue is a threonine (T116) that neighbors a much more highly conserved glycine and is changed to a valine, another pro-β residue, in other Drosophila species, and usually a negatively charged residue in non-Drosophila species.

Beta sheets are made up of extended strands with antiparallel sheets having hydrogen bonds perpendicular to the strands with alternating narrowly and widely spaced bond pairs, while parallel sheets have evenly spaced hydrogen bond pairing. In nature, mixed antiparallel and parallel sheets are possible but there is a bias towards either pure parallel or pure antiparallel [60] so it is unsurprising that parallel β content was not predicted alongside the antiparallel β content. Antiparallel β sheets can occur with fewer strands than are typical of parallel β structure and, unlike parallel sheets, tend to have one side exposed to solvent and frequently exhibit alternating hydrophobicity in the sequence.

Although far-UV CD spectra can’t determine tertiary structure, it is worth considering if the WYR β content might be participating in a higher order structure as inclusion/exclusion of possibilities could shed light on both β and other elements of

secondary structure [61], protein buffer sensitivity or likelihood of transitions upon oligomerization. Higher order structures more specific to parallel β sheets can be ruled out, along with structures that require many β sheets which are unable to be accommodated by WYR's short sequence. Beta helices, or 'β prisms', and β spirals involve parallel β sheets while β propellers are known to contain multiple antiparallel β sheets. Beta barrels, able to be formed by both antiparallel and parallel β sheets, are made up of at least 5 strands for the smallest known and at least 60 residues, at a minimum [62]. Most characterized structural β sandwich domains involve 100+ residues. These include the Immunoglobulin domains consisting of 7-9 antiparallel β strands that form a β sandwich with a Greek-key topology and the C2 domain, an 8-stranded β sandwich known for a propensity to bind calcium, among others. These fall into the category of β-solenoids. Beta solenoid is an umbrella term for a β strand-only containing higher order structures with the associated turns described as β-arcs [63]. Immunoglobulin domains, such as that associated with C10, the LMM binding domain of MyBP-C, fall into the categorization of β-solenoid. As these structures are not able to be considered for WYR, simpler higher order structures, such as β-meander or antiparallel two-stranded β helix, sometimes described as β-hairpin or β-ladder, are worth serious consideration.

Beta hairpins, the most likely β content present in WYR, require two β strands and are typically short in length. Beta strands are usually made up of at least 5 amino acids with short strands being more likely to exhibit twist. With the 52aa WYR segment, 31% of residues potentially engaging in antiparallel β form would equate to the involvement of

16 residues which could involve 3-3 β strands. Salt bridges often stabilize the surface of β strands and β sheets so the central charged residues (K94, R100, D105, D106, D109, D112, K113, K114) are conducive to this structure. Aromatic residues are favored in the middle of β sheets while prolines are favored towards the edges and have been frequently involved in the nucleation of β structure [64, 65].

The alternating pattern of tyrosine residues present within the first one third of the WYR sequence is strongly supportive of an area central to a β strand which would permit a solvent-exposed tyrosine ladder on one side. This alternating tyrosine pattern spans 7 residues (Y93-Y99) although it is likely that the strand would include residues at either end. The prolines in the WYR sequence (P88, P123, P125) are not permissive to β structure, unless they are participants of turns. The next likely strand is C-terminal to the first strand, closer to the central region of WYR, and requires separation from the first strand by a turn. The region of T101-Y104 is permissive of a β turn, possibly a type I' or II' β turn with T101 in the i position and N102 in the i+1 or N102 in the i position [49, 66]. A turn at T101-Y104 would align multiple hydrophobic and complimentary charged residues to occupy the center of the antiparallel β structure, whose formation is known to be turn-driven [67, 68]. T101-Y104 would be conducive to tyrosine-asparagine stacking [69, 70] while still permitting favorable interactions between the two tyrosines.

Expected twist and distance between residues strongly support favorable interactions along the tyrosine ladder, extending to T101. In a β strand, twist can vary from 0-30^{o} per

residue and the distance between residues is ~3.5 Å, so each residue is about 7 Å from the next residue with a same-side-facing R-group. For three residues, the total twist could be 90^{o} and is ideal for tyrosine-tyrosine interactions. While π-π stacking interactions are possible, T-shaped interactions in which the hydroxyl of one tyrosine is positioned over the pi face of another (OH-π) is the most stable conformation and is most probable between orientations of 80-100^{o} [71]. The threonine (T101) at the start of the twist is positioned excellently to 'cap' the last tyrosine in the sequence (Y98) in the same manner. In the format most ideal for T-shaped tyrosine contacts, the β hairpin would be in a coiled format [72].

Overall, the combination of CD and sequence analysis suggest the ~30% of residues engaged in right-twisted antiparallel β structure includes at least Y93-D112. The proposed structure includes a favorable line up of charged residues: R100 aligns with D105 and K94 aligns with D109, the latter of which may be reinforced by, or involved in potentiating, cross-β contact between Y95 and Y110. There are no unfavorable charged contacts observed within Y93-D112 in a β hairpin format. There is only one possible outward facing hydrophobic residue (I108), which may be involved with further contacts outside of the β hairpin. Hence, the predominant β content is expected to encompass Y93-D112; with the β turn excluded, this would be the equivalent of ~30% of WYR structure, matching BeStSel's suggested proportion.

The proposed structure also includes a tyrosine ladder extending from Y93 to Y99 in which a more central tyrosine, Y95 may be involved in cross-β contact with Y110. Cross-β Tyrosine ladders, related over multiple antiparallel β strands to form a row of tyrosines, have been found to be involved in amyloid proteins, docking surfaces for protein-protein interactions [73] and have been the basis for development of self-assembly mimics [74]. Cross-β Tyrosine ladders found in amyloid structures are not resultant from interactions along a single solvent-side of a β strand. Likewise, we surmise it is not likely that WYR is exhibiting extensive cross-β character from its own putative solvent-exposed tyrosine ladder, though cross-β is possible in the one instance in the area center of the proposed hairpin between Y95 and Y110. How this would impact the charge distribution along the tyrosine ladder between Y93-Y99 is unclear but a cross-β connection between Y95-Y110 would act as a stabilizer for the hairpin structure [75, 76], with Y95 rendering the stabilizer sensitive to any change relayed through the exposed tyrosine ladder, be it due to solvent or binding. The possibility that WYR Y93-Y99 is a protein-binding site deserves further study.

'Other' & 'Turn' content of WYR further examined

The predicted β hairpin encompasses at least one turn and the rest of the sequence is expected to be predominantly 'turn' or 'other' content and make up most of the WYR sequence. While turn and loop content has historically been considered to be mostly important for 'linking' other secondary structures in a protein, a greater understanding

and appreciation of their roles has led to further characterization and provides some ground for hypothesizing on the type(s) present in WYR structure.

It is likely that T101-Y104 encompasses a turn responsible for the majority of right-twisted β hairpin content in WYR. This could be in the form of a β-type turn but other turn types are open for consideration, including the Asx turn, and ST turn. Asx turns fall under similar categorization as β turns but the side chain of residue i (Asn or Asp) hydrogen bonds to the backbone NH group of i+2. Asparagine has been observed making such turn-based contacts with tyrosine at a position +2 away, and the N102 and Y104 would be amenable to this [77]. Similarly, in ST turns, the serine or threonine at i frequently will form a hydrogen bond with the main chain NH of i+2. Given the expected 30^{o}/aa rotation along the right twisted strand, T101 is in an ideal position to cap the end of the stacked OH/π bonds along the tyrosine ladder. If an ST turn were present, threonine's R-group would not be available for engagement of the tyrosine ladder. However, a β type I' turn would accommodate the expected twist well while leaving the threonine's R-group to engage with tyrosine's aryl group, as would an Asx turn. Beta turns can also incorporate Asx turns and asparagine is strongly over-represented in β type I' turns [78]; these are not mutually exclusive designations.

Tyrosine corners are common to β sandwich domains and are highly conserved in fibronectin type III and immunoglobulin superfamilies [79] which are present in cMyBP-C. Tyrosine corners contain one tyrosine at the beginning or end of an antiparallel β

strand in which the tyrosine hydroxyl H-bonds to a backbone NH or CO of a residue at Y-3, Y-4, or Y-5 with the Y-4 variant being most common. The consensus sequence is LxPGxY with a hydrophilic residue in the x position representing Y-4, although other residues, such as tryptophan or histidine can also form similar "corners" [80]. The WYR sequence doesn't harbor the consensus sequence or contain like-residues for previously found tryptophan and histidine corners so this type of turn/loop structure is not supported though there is nothing clearly prohibiting an aromatic outside of the tyrosine ladder from being involved in H-bonding to the peptide backbone.

Other possible turns can incorporate more or fewer residues compared to β turns such as α and γ turns. Gamma turns (n=3) have a hydrogen bond between i and i+2 and the φ, ψ angles must be within 40^{o} of either the inverse (-79^{o}, 69^{o}) or classic (75^{o},-64^{o}) variant, and may be more amenable to incorporation of diverse residues than β turns [81]. Alpha-turns (n=5) have been characterized by ϕ, ψ angle distribution as extensively as β turns, though are less common [82, 83]. A turn of n=3 (γ, Asx) or n=4 (β) are most favorable for the proposed central WYR hairpin but modelling would be necessary to fully articulate the possible conformations.

Matters become more complicated in that there can be overlapping turn segments in loops or other turns, similar to β turns encompassing Asx turns. π-turns (n= 6; i, i+5 H bonding) are most often present at the end of helical structures and often internalize a β-turn or are composed of multiple types of β-turns [84] as do over 30% of smaller α-turns

[85]. Pi turns are not likely for any part of WYR, partially due to the absence of helical content but also due to a general lack of glycines, strongly preferred in π-turns, within the sequence. Aromatics, of which there are many in WYR, are preferred at π-turn positions at i-1, i+6, but are also not appropriately positioned to other π-turn preferred residues that do exist within the WYR sequence. Loop structure on the outskirts of the central antiparallel β content (Q93-K113), however, could include γ-, β- or α- turn content.

Beta bulges may also be contributing to irregularity designated by 'Other' content. They are known to occur in β sheets composed of at least two strands. In antiparallel β structure, 'classic' bulges occur with an additional residue on only one strand positioned between two narrowly spaced pairs of hydrogen bonds while a 'wide' bulge is characterized by the additional residue lying between a pair of the more widely spaced hydrogen bond pairings. More rare variants are the 'bent', 'G1' and 'Special': 'bent' bulges involve an additional residue present on both strands, "G1" bulges occur at the end of an antiparallel strand and are characterized by a glycine, and 'Special' involve insertions of up to three residues in one strand. Of all the types of bulges, the G1 bulge is most often found on the outside ends of an antiparallel sheet, rather than being internal [86]. The most probable place for a 'bulge' to show up in WYR structure would be towards the end of the antiparallel β hairpin at G117, a residue only lacking in paraWYR of decapods and chelicerates. A type I turn followed by a G1 bulge is a trademark of 3:5 hairpins and could be taking part in further brief antiparallel β contact N-terminal to the conserved proline (P123).

Although we hypothesize the β hairpin of WYR to be centered on a type I'/II' β turn or Asx turn, other portions of WYR are expected to exhibit additional turn or loop behavior given the BeStSel projected turn and other content. Such series-turn or loop structures may dominate in the C-terminal of WYR and be involved with additional contacts with the central hairpin. Omega loops are the most well characterized of loop structures but remain somewhat loosely categorized. More rare are ψ-loops in which two antiparallel strands are connected by a "+2 connection", an additional strand in between that is hydrogen bonded to both.

Omega loops, due to their variability in H-bonding and hydrophobic contacts and undefined end points, have been characterized by purpose as "functional", "stability" or "folding" loops [55] and are frequently solvent-exposed. Omega loops are so named by their shape resulting from the reversal in direction of the polypeptide after 6 or more residues. "Functional" loop regions may be extremely important to substrate specificity and binding and may be highly conserved for this purpose. Hypervariable loop structures in immunoglobulins, responsible for antigen specificity and trypsin and chymotrypsin substrate specificity, can be interchanged by swapping of the loop segment. There are multiple instances of Ω loop lids that operate alongside a hinge, participating in transition states and moderating access to primary binding sites [55]. "Stability" loops are so named as conservation within loops are not found to be important for general protein structure but can be important for stability in different environmental contexts (e.g. temperature,

presence of ions). “Folding” loops are loop types that may be considered to be microdomains and maintain independent structure and, like “stability” loops, can be involved in transitions, specifically for protein folding. “Stability” and “Folding” loops are still in need of further characterization with least information being available for loops responsible for the folding of periodic structure in the remainder of the protein.

Given the high conservation of WYR sequence, loop content is likely to be functional Ω- or ψ-loops towards the C-terminal and involve the rigid segment around P123. Psi-loops can be involved in H-bonding a coil or helical structure to a β hairpin and are not limited to additional β strands. Cation-π interactions with tryptophan are very common and form the strongest cation-π between the aromatic residues [87], which can be augmented by proline. Prolines and basic residues (R,K) are located proximal to both tryptophans in the WYR sequence and are likely to be interacting and stabilizing for loop structure. The conserved tryptophans could also be participating in anion-quadrupole interactions with aspartate (D) or glutamate (E) residues [88] if, the participating groups were >5 Å apart with hydrogen bonding accommodated. Intra-sequence contacts may also be necessary to decrease the inter-protein reactivity of tryptophan in the N- and C-terminal regions of WYR to promote contacts closer to the central region. If the WYR structure is able to instigate β to helical transitions, these flanking loop structures may be involved in the regulation of this process.

Overall, the turn associated with the central β hairpin is expected to be a β type I'/ASX turn, possibly a β type II', that may be involved in transitions in secondary structure with 'other' loop/turn content predominantly located in the C-terminal of WYR. C-terminal loop structures, likely encompassing additional turns, following a G1 β-bulge towards the end of the central hairpin, is also proposed. In addition to loops, β bridges, defined as β sheet hydrogen bond formation between a single pair of residues, and 3_{10} helical content contribute to the 'Other' BeStSel categorization. Both β-bridges and Type I β-turns contribute to 'Other' content found in WYR as β type I turns are similar to the format of a 3_{10} helix and small contacts with the central β hairpin embedded in the C-terminal loops or G1 β-bulge are likely. Unsupported "Other" content for the entire WYR sequence includes π-turns, Tyr/His/Trp-corners, and ST turns/motifs.

The hypothesized WYR structure we propose contains T101-Y104 as encompassing the turn central to WYR's β hairpin whose strands flank this turn, with an additional G1 β bulge present on the C-terminal strand with 'other' structure flanking this β hairpin (**Fig. 3-7**). The β hairpin is deduced in part by the probable location of the defining turn, by favorable contacts between and within hypothetical strands and through exclusion of the more N- and C-terminal ends of WYR as being inauspicious to β structure. The 'other' structure is likely to be functional in nature and include rigid loops and turns that contribute to the P_2 component of WYR's overall β_{II}-like CD profile. The proposed structure and contacts can be tested and some suggestions for future studies are proposed in a later section.

Comparison to Secondary Structure Prediction Programs Based on AA Sequence

There are a plethora of secondary structure prediction programs that aim to accurately identify regions of periodic and aperiodic structure on the basis of amino acid sequence alone. Programs we've used to attain secondary structure estimates for WYR include CFSSP [61], YASPIN [89], PHDpsi [90], PSIPRED [91], Phyre2 [92], Jpred [93, 94], and I-TASSER [95-97]. We examined the output of these along with all the secondary structure prediction methods offered by Network Protein Sequence Analysis (NPS@) of Pôle Bioinformatique Lyonnais [98], PredictProtein [99] and the Proteus Structure Prediction Server [100] amounting to a total of 22 programs (**Fig. 3-8**).

Compared to the β sheet content estimated from the WYR CD profile, 21/22 programs predicted substantially less or none with a much higher projection of helical content. YASPIN, PsiPred, Reprofsec (PredictProtein), Porter 4.0 and Jpred (Jnet) all predicted no β strand content along the entirety of the WYR sequence (**Fig. 3-9A**) and ranged from a high of 71% (YASPIN) to a low of 40% (Jpred) helical content with the remainder being random coil (not shown).

To see if propensities changed within the programs that only predicted helical/coil content, tyrosines were changed to alanines, a change that in theory should greatly impact the predicted structure under the condition that the side chain of the tyrosines are

considered in the secondary structure predictions (**Fig. 3-9B**). Alanine was used as it is able to participate in H-bonding of both α and β content, and while it is frequently associated with some helical preference over β sheets, the relative orientation of the alanines as a result of replacement of all the tyrosines in the WYR Y-ladder would be more favorable for β content. The resulting alternating hydrophobic pattern would be more conducive to an inward facing portion of a β strand whereas the pattern would be rare in a helix unless that helix were very internal in a tertiary structure. Hence, β would be favored over α helical content. Change to alanine allows removal of tyrosine's R-group yet is not as invasive as replacement by phenylalanine which would be a much larger nonpolar group. The change to alanine would be expected to increase predicted β-content and allow us to observe how the programs treat the structure in the context of the absence of tyrosine's bulky side group, without the conformational freedom that glycine permits, or the stronger, more hydrophobic, pressures of β-proponent residues (Ile, Val, Phe).

Upon this change, Porter 4.0 designated one more residue as helical (P88) with all others unchanged. Reprofsec (PredictProtein) found two more residues to be helical as opposed to coil (Y93A, K94) and designated L133 as coil rather than helical. PsiPred newly designated Q92 as helical and N98 and Y99A as coil. Jpred (Jnet) found the most change with Q92 becoming the only strand-predicted residue among all five programs; K94, Y99A, Y104A, D105, and T127 were changed to a helical designation and R100 and T101 were changed to a coil designation. Three programs re-designated some tyrosine

positions within the ladder as either helical or coil. Overall, there was minimal change and most changes were at the beginning or end of a helical span with the exception of the additional 3-residue coil in PsiPred. Jpred (Jnet), which initially had the lowest prediction of helical content, exhibited the most change. This could be due to greater fragmentation of these segments in the original prediction compared to the other programs; i.e., there were more start/end points of helical segments than the other programs and these were hot spots of change. Types of change in the structure was very variable between the programs; only one pair of programs (JPred & ReProfSec) were in consensus about a designation change of one residue (K94 becoming helical). It would be expected that, if the aromatic character of the tyrosines were being considered, greater spans of residues would be impacted and there would be more instances of shift for the mutated residues and nearest neighbors.

The literature has recognized that caveats of secondary structure prediction based on primary sequence stem from limited basis datasets, and datasets sourced using methods not representative of realistic *in vivo* conditions [61]. Algorithms vary in regards to basis datasets as well as whether they take the conformational probability of each residue alone or in context of its neighbors, and how that context is gated. Algorithms commonly do

not take into consideration orientation driven by side-chain interactions, which, in a highly aromatic residue-containing sequence, is expected to be substantial.

Of the programs that detect turns, GOR I and SOPM from NPS@, and CFSSP, turn regions were proposed at positions K89, R100-N102, K113-T116, R124-P125, and R134 (**Fig. 3-10**). K114 attains the highest chance of turn content, in consensus between all three programs. N102 and R134 are unique in that these programs attributed three different structures to these positions: helix, strand and turn for N102 and helix, coil or turn for R134. There are only three other positions in which all three programs give different predictions (R87, V118, L133). Amongst 22 algorithms (**Fig. 3-8**), of the residues suggested to be in turn conformation by GOR I, SOPM, or CFSSP, R134 is equally divided between coil and helix prediction while R100-N102, Q115-T116 are the only segments that are designated at β strand by any of the programs. The segment of R100-N102 is the only segment to have a turn designation that is flanked by multi-residue spans predicted to be in β strand format by at least 2 algorithms. T101 and N102 have been hypothesized as possible positions of turn initiation, with particular interest for initiation at N102. R100 lies between the last Y of the tyrosine ladder and the, theoretical, capping T101. No programs have been found capable of taking into account a possible tyrosine ladder and OH-π interactions hypothesized, which would impact the probability of a turn within the area. Hence, with consideration of the tyrosine ladder, the most support for a turn residue within the WYR sequence is N102.

The algorithm that categorizes N102 as a turn is GOR I and is the only program that considers circular dichroism data for their basis structures in the absence of x-ray crystallography, which would accommodate the concern that crystal structures are not always representative of those in solution or *in vivo*. It's arguable that this may be more representative for structures not able to form crystal structures and could be the best choice for evaluating WYR structure via sequence. This program gives high estimates of β structure (**Fig. 3-11**), 67% of total, with the longest β segment flanking N102. The largest helical segment is I108-K110 which is in consensus between all turn-predicting programs and is towards the end of the longest helical segment predicted by most other algorithms, additional encouragement towards this region having potential to undergo helical transition.

Despite caveats for each of the predictive methods, some areas of interest can be identified when the output of all the programs are considered. Between the many algorithms, regions that stick out include N102, I108-Y110, and a segment starting at W128. There is also support for regularity in spans of 8+ residues flanking N102, even if there is no agreement regarding the type of regularity. N102 and the region of drastic transition that takes place just past I111-K113, at K114, may represent turns leading into changes in periodicity. I111-K113 and a segment at W128 display highest helical propensities, even for the β-heavy GOR 1 prediction, and may be most likely to exhibit helical periodicity if β-to-helical transitions factor into WYR structural change upon binding or context changes.

How might the structure of WYR be changing in context of buffer or binding?

The CD profile between our buffered WYR and preliminary findings with WYR in just dH2O are slightly different, with WYR in water exhibiting lower helical and higher antiparallel β content (**Fig. 3-3, Fig. 3-5**). Preliminary experiments also demonstrated WYR to be more soluble in pure dH2O (observations). Though further experimentation is needed to confirm these observations, the hydration properties of WYR may play a role in its solubility/aggregation propensity in ionic buffers and allude to a possible role for intrinsic charge interactions. Proteins in solution exhibit conformational flexibility that can encompass a range of hydration states not found in a crystal, a form avoided by proteins *in vivo*. Water activity and fugacity is decreased by the presence of salts. In pure water, some proteins that are normally insoluble under buffered conditions, become soluble due to permissible, unshielded, interactions between a protein's intrinsic charges [101]. This could factor in the case of WYR as there is evident segmentation of positive and negative charges, such that the complimentary charged residues may be proximal only when folded, but exposed to solvent. Tyrosine, tryptophan and proline are slightly kosmotropic and can stabilize the structure by stabilizing low-density water, thereby increasing the local density of water and providing some protection.

WYR has a high aromatic (19%) and hydrophobic (34%) content that could enable the formation of ordered water in clathrate-like structures or in linear structures. Structured water, in turn, can impact the function of WYR. Linear clusters have been used as proton

wires for charge transfer in several proteins [102]. In BPTI, a linear chain of water forms transiently but repeatedly; this ‘aqueduct mechanism’ has been found to pull in ligands [103]. The formation of clathrate-like hydration of antifreeze proteins (AFP), via both hydrophobic residues and compartmentation, have been known to be important for their function [104]. An AFP from Chironomidae (midge) has been found to involve a π-stacked series of highly conserved tyrosines between β strands and its function is completely ablated upon mutagenesis of the central tyrosine [105]. In the case of the midge AFP, threonine and asparagine were also important for anchoring the clathrate cages that formed.

The prevalent tyrosine content, especially in the suggested format of a tyrosine ladder, is expected to be very sensitive to the presence of cations, more so than other aromatics [106]. Aromatics are known to be over-represented in regions of binding, partially due to their ability to engage in cation−π interactions [107]. Multiple individual tyrosines have been shown to develop into higher order structures due to their high reactivity and aromatic-aromatic involvement [76, 108] and are focal points for development of synthetic binding proteins due to their high prevalence in protein-protein binding regions [109]. There are multiple examples of solvent-exposed tyrosine stacking in antiparallel β structure purposed for dimerization or receptor binding [110]. Taken together, Y93-Y99 of WYR is a strong candidate for being both sensitive to cations in solution and acting as binding participants, whether that be to a positively charged region of the LMM, self-interactions or to other muscle proteins. Cations in solutions, such as Ca++, may even

modulate affinity for protein binding with higher [cation] in solution releasing an otherwise secure interface.

A tyrosine ladder involved in H-bonding of the tyrosine OH group or involved in edge-to-face OH-π interactions can greatly potentiate cation-π contacts [111]. A tyrosine ladder formed by such contacts can be extra attractive to cations causing a transition to more dominant cation-π type interactions. This transition can be initiated by changes in the protein milieu such as in ionic conditions or presence of other proteins. Arginine is preferred in cation-π interactions with tyrosine, over lysine [87]. As lysines are the preferred positively charged residues among the tyrosine ladder and expected antiparallel β segment, it's likely that cation-π interactions are not prioritized between these residues in the native WYR structure. However, this could develop in the presence of exogenous cations that could destabilize the highly right twisted-form of the tyrosine ladder strand and stabilizing salt bridges. In such a situation, the central β structure may be less twisted and more reliant on H-bonding. A transition to strict reliance on antiparallel H-bonding could result in competition from the multiple aspartic acids and asparagines for such contacts and lead to a disconnect between the two original strands and ablation of antiparallel character.

Aromatic and hydrophobic regions are hot spots for protein-protein interactions with arginine being over-represented for contacts involving β strands [112, 113] suggesting that most of the WYR sequence is conducive to binding interactions. There may be a

specific region involved in first contact with the remainder of WYR undergoing sequential binding stimulated by the first. Although arginine is over-represented in the contacts for β strand binding, in the formation of helix-β strand interfaces, arginine is under-represented with a much greater representation for tyrosine and tryptophan [114], most frequently in contact with hydrophobic residues. Proline and lysine are also under-represented at such interfaces, lending greater consideration towards the central antiparallel β segment of WYR being involved in such an interface, even if smaller segments of β character are present outside of this region.

Oligomerization would represent a very different scenario compared to buffer or alternate protein binding. Assuming two or more equivalently stabilized structures, longer distance ionic interactions that play only a lesser role in the stabilization of the β component may play a much larger role in oligomerization, in which stagger of the β strands allow contacts not otherwise available in the native structure. Such behavior is utilized and well researched in the creation of hydrogels and can impart new mechanical character [115, 116]. If multiple domains were to stagger, it would result in a fibrillar tertiary structure and potentially stabilize larger, more-reactive, interfaces such as the tyrosine ladder. The WYR sequence is amenable to dimorphic β-to-β oligomerization but may not be limited to oligomerization via a matching of the central β strands. Beta-to-beta dimerization has been most frequently found to occur in homodimers, antiparallel at the interface, with aromatic residues and prolines favoring a dimorphic relationship [117], in which one strand of each monomer is interacting. This is as opposed to a standard format

in which at least two strands from each monomer are interacting with each other or hybrid format in which one strand from one monomer and two strands from another monomer are interacting. Arginine and lysine are also the most common charged residues for dimorphic interfaces and, if involved in additional short antiparallel content outside of the central hairpin, could be important for such an interface. Most contact areas for dimorphic antiparallel β-to-β dimers are also short, so β contacts outside of the central pair of β strands hypothesized for WYR, can be considered as possible interfaces for such contact.

Small structural changes resultant from binding, temperature change, or ionic domination in the solvent could result in more dramatic shifts or change the exposed component(s). This could lead to further accessibility for binding partners and trigger larger shifts. Tryptophan and proline-heavy loop structures with strong interactions are rigid and less prone to change, while more flexible regions that may act as a 'hinge', or lead-in to such regions, can be modulated to change the exposed landscape more dramatically. Re-orientation of the rigid component via changes in the 'hinge' without impacting the stability or structure of that component could result in exposure of additional segments of the WYR structure previously masked by positioning of the rigid component. This could increase availability of reactive residues. Alternately, repositioning of the rigid component could 'protect' normally exposed regions and put a halt to further conformational change.

In cases of binding contact with a conformationally different target, regions of instability or inconsistency within the target may be attractive as an interface and the binding interaction could stabilize the complex. As an example, a classical helical coiled-coil contains a complementary pattern of hydrophobic (H), polar (P) and charged (C) residue types: H P P H C P C, in which each position is designated a-g. Disruptions in this pattern can vary in extremity and result in exposure of reactive residues, sometimes completely shifting the target's structure and allowing greater accessibility for binding. Disruptions range from stutters (deletion of 3 residues) and stammers (deletion of 4 residues) to more disruptive insertions of 1 or 5 residues that can lead to formation of π-helix segments or break the helix, resulting in short β-strand content, called a β layer [118]. Binding of WYR to such a region could allow re-orientation of destabilizing contacts that could either allow local re-structuring or provide a secure to these regions of "weak links" that may permit even greater stability than the native heptad pattern.

What can be done in future studies to further examine WYR structure?

There are multiple ways in which to examine what residues are solvent-exposed in WYR at different [WYR] or in conjunction with binding studies. N-acetylimidizole (NAI) acetylates K,R,N,S,T and Y residues and can be used along with SILAC mass spectroscopy to determine if these residues are exposed in solution [119]. To examine whether the multiple tyrosines are present in a structure that permits same-side/face tyrosine-tyrosine interaction, Ru(bpy)32+-catalyzed photo-cross-linking can be used to form dityrosine adducts [120] which can then be identified and characterized by ESI

tandem mass spectrometry [121]. Dyes such as Congo Red and Methylene Yellow may permit similar evaluation of WYR structure in the absence or presence of binding partners, different buffer solutions (e.g. pH, ionic strength) or in conjunction with other context changes (e.g. increased molecular crowding, temperature). Tyrosine ladders composed of 4 or more tyrosines in a row also provide a binding site for methylene yellow (Thioflavin T) [73, 122] that emits a weak signal at ~530 nm when unbound, and a strong signal at 482 nm when bound and excited at 450 nm.

The role of salt bridges can be examined broadly or specifically. Changing pH (e.g. deprotonating positively charged residues lysine and arginine at higher pH) and by changing ionic strength to introduce charge screening [123] are traditional means that can be easily coupled with Circular Dichroism or other methods. Temperature and the dielectric constant (e.g. glycerol vs water) of the solution can also be used to evaluate the prevalence of charge interactions and allow greater selectivity of charged residues impacted [123-125]. Emerging methodologies, such as the use of Nanosecond pulsed electric fields [126] may be relevant in the future, as WYR structure becomes more established and structure-coupled charge interactions, responsible for more nuanced conformational shift or binding contact, are of deeper interest.

Colloidal stability can be evaluated using kosmotropes (e.g. glycerol, betaine, ectoine, trehalose) and chaotropes (e.g. urea, guanidine) can elucidate what types of interactions are important for structure and reveal new context-dependent transition states that may be

relevant for *in vivo* binding or other environmental conditions. Kosmotropes are polar and possess negligible net charge and so act to extensively H-bond with water molecules while chaotropes have weaker H-bonding and can simulate “more dense” water. This can probe WYR’s reliance on H-bonding, aggregation propensity and reversibility. This can provide information on the protein-water interface for the native structure and can be coupled with analytical ultracentrifugation, circular dichroism, static/dynamic light scattering, self-interaction chromatography, various other types of absorption spectroscopy (e.g. NMR, IR, Raman), ultrasonic shear rheometry and many other techniques [127, 128].

To examine if the structural changes induced by WYR can be related to mechanical capacity, single molecule force measurements can be performed with or without a binding partner, comparing bound and free constructs, in which the binding partner is covalently attached to an AFM tip by a PEG tether. As information about the structure grows, the change in structure of oligomers, or in conjunction with other binding partners, can be coupled to mechanical impact on the oligomer or solution behavior (e.g. viscosity) using rheometry. Possible fibrillar tertiary structure may be able to be detected by EM and allow further mechanical dissection.

Post-translational modifications (PTMs) may play a role for WYR function *in vivo* and the function of side-groups within the structure may be able to be probed by phosphorylation, glycosylation, acetylation, oxidation or other PTMs [129]. Oxidation is

most likely to occur on methionine, serine, or tryptophan. Oxidation of the methionine neighboring a tyrosine, can be involved in tyrosyl oxidation or nitrotyrosine formation [130, 131] with negatively charged residues facilitating this process and positively charged residues impeding it [132]. In the WYR sequence, M96 is sandwiched between Y95 and Y97 of the tyrosine ladder and has been found to be oxidized by MS (Emily Price, unpublished data). Even with a possible oxidized methionine nearby, the tyrosines in a ladder form for the opposite side of a β-hairpin would be more protected from methionine-mediated modification and Y95 is further neighboring K94, which would discourage this process. An additional consideration is that M96 is not very highly conserved and is frequently changed to a leucine or isoleucine outside of Drosophila, more suggestive towards a primary role as a supportive non-polar internal residue.

Asparagine (N102) would have its R-group H-bonding with the main chain in the proposed Asx turn/motif and this may be able to be evaluated by inhibiting this type of contact by modification of the side group. N-linked glycosylation can be performed on the amine group of asparagine located at N102, which may be enabled by the nearby threonine residue. Deamination can also disrupt the H-bonding pattern if the side-chain is involved and can further reveal information of the level of exposure for asparagine residues within the structure [133]. The propensity for Asp/Asn to form such bonding has been found to be a conformational and functional driver in other proteins [134], coupled with neighboring prolines.

Many of the methods described up to this point can provide information on structure with and without the context of ligand binding interactions, but additional techniques can be utilized with a more direct focus on the protein-ligand interaction. Near UV CD can determine packing of aromatic residues and Förster resonance energy transfer (FRET) using intrinsic fluorescence of aromatic residues [135] can determine distance correlations for interacting regions. Thermodynamic properties (e.g. enthalpy (ΔH), entropy (ΔS)) of binding and stoichiometry can be measured without the need for molecular labels by isothermal titration calorimetry and kinetics by using surface plasmon resonance or biolayer interferometry [136]. If tags or cross-linking is an option, MS methods become feasible [137, 138].

In muscle, structures on the minute level of protein folding may result in structural shifts for the entire fiber. Protein structure has bearing on function, dictating the roles of residues by their precise arrangements. Such arrangements dictate the level of exposure and reactivity that the residues engage in individually and as a collective. The study of flightin, known to impact structure and organization of the thick filament, necessitates an understanding of the structure, and structural changes, within both binding partners: the flightin WYR region and the myosin LMM. In order to properly articulate the impact of this relationship, we will now discuss the structure of the LMM before transitioning into a dissection of the structure formed by the LMM in the context of WYR.

LMM AND C600 SEQUENCE AND STRUCTURE

Myosin II is a member of a protein superfamily that shares three functional domains that include (i) an N-terminal globular domain that binds actin and hydrolyzes ATP to perform mechanical work, (ii) a helical coiled-coil tail domain and (iii) a partially helical hinge region that connects the two. Proteolysis of myosin II separates the head 'S1' and hinge 'S2' regions, together designated heavy meromyosin (HMM), from the coiled-coil tail, designated light meromyosin (LMM). Myosin II is highly conserved throughout Animalia and evolutionary analysis suggests that these three domains co-evolved and are functionally interdependent [139].

The LMM's coiled-coil structure is necessary for myosin to assemble into the thick filaments. Some of this assembly is inherent, driven by an assembly competence domain (ACD) and an alternating charge repeat pattern that spans 28 residues. In low ionic strength solutions (~150 mM), the LMM from rabbit striated muscle will form ordered aggregates (paracrystals) *in vitro* with a similar periodicity, as found *in vivo*, at 43 nm detected by EM [140] though the assemblies vary in length and width. This behavior has also been identified for the Drosophila LMM [7]. Regular assembly according to species-specific filament parameters, however, require orchestration on the part of particular LMM-associating proteins, or molecular 'rulers' such as titin, present *in vivo*. Both α helical coiled-coiled structure and LMM-associating proteins are necessary for thick filament species-specific attunement.

Heptad repeat

The structural character of the LMM is driven, in part, by a heptad repeat, a character of primary structure, in which amino acids of each participating α helix in a dimer are arranged in a 'HPPHCPC' pattern in which H represents hydrophobic residues, P for polar residues and C for charged residues. The heptad is described as being in positions 'abcdefg' (**Fig. 3-12**). Deviation from this pattern is described as a 'stutter' if there is a deletion of 3 residues from the repeat, a 'stammer' if there is a deletion of 4 residues or a 'skip' for a deletion of 6 residues (or an extra 1 residue). The charged e and g positions of each helix stabilize the seam of the coiled-coil/super helix but can be responsible for instability in a single α helix. Hydrophobic interactions at the 'a' and 'd' positions form an apolar core that may engage in other structures when not in a coiled-coil form. The heptad repeat of the myosin II coiled-coil further exhibits higher order alternating 28 aa repeats of 14 positively charged and 14 negatively charged segments. However, the myosin heptad repeat contains conserved discontinuities inclusive of four skip residues [141], equating to positions T1187, E1384, E1581, G1806 in Drosophila.

Extended heptad repeats result in left-handed supercoiling but tend to have less regularity. With the introduction of stammers, stutters and skips, structural character can be inconsistent along the length of the protein. Stutters can result in unwinding of the coiled-coil towards right handed-ness and stammers can result in accentuated left handed-ness [142]. Crystal structures on portions of the LMM driven to coiled-coil formation by non-LMM elements can be seen to possess imperfect heptads [143]. A good example is

found in the referenced study, Xrcc4-H1590-L1657 from human beta cardiac myosin. Xrcc4 is a ~140 aa mostly β-stranded globular segment of a DNA-binding protein attached to improve formation of a LMM helical coiled coil segment. This segment (H1590-L1657) is described as having 10 heptads. However, even if the looser description is used for a traditional heptad repeat (HXXHXXX) and spacing between them is ignored, the segment contains 6 at maximum (underlined with ***bold/italics***; hydrophobic residues are orange; every seven residues are separated by a period in accordance to the reported heptad positions), counting overlapping:

Homo sapiens (P12883) – H1590-L1657

HLRVVDSLQTSLDAETRSRNEALRVKKKMEGDLNEMEIQLSHANRMAAEAQKQVKSLQSLLKDTQIQL
XHXH.***HXXHXXX.***HXHXXXX.***XXXHHXH.XXX***HXXX. ***HXXHXHX***. ***HXXHXXH.HHXHXXX.HXXHXXH.***HXXXXHX.H
1 **2** **3** **4** **5** **6**
defg abcdefg abcdefg abcdefg abcdefg abcdefg abcdefg abcdefg abcdefg abcdefg a
Heptad: HXXHXXX

The description of this being 10 heptads is more representative of there being almost 70 amino acids in the stretch (7*10). When the positions are listed in heptad order as designated in the literature, five of the six are aligned (**1**, and **3-6** as listed above). As this is not commonly discussed in the literature, the implication is that the only impactful interruptions to a regular heptad repeat in the LMM are at, or around, the four conserved skip residues. In experimental application, this does not appear to be the case. Experiments using these LMM fragments, and others [144], required attachment to portions of other non-myosin coiled-coil forming segments such as Xrcc4.

This sequence from human beta cardiac myosin aligns to H1589-L1666 of *Drosophila melanogaster* myosin II with 56% identity (identical) and 72% positives (conserved character). Even with more hydrophobic (27) residues compared to the beta cardiac myosin sequence (25), there are fewer heptads following the HXXHXXX pattern (5) with four of those five aligned with the heptad order. These counts don't include aromatic amino acids as 'a' and 'd' position possibilities as they are non-ideal for the core of the coiled-coil.

Drosophila melanogaster (P05661) - H1589-L1666

HQRALDSMQASLEAEAKGKAEALRMKKKLEADINELEIALDHANKANAEAQKNIKRYQQQLKDIQTAL
XXXH.HXXHXHXH.HXHXHXX.XHXHHXH.XXXHXHX. HXXHXHH.HXXHXXH.XHXHXXX.HXXXXXX.HXXHXXH.H
1 2 3 4 5
defg abcdefg abcdefg abcdefg abcdefg abcdefg abcdefg abcdefg abcdefg abcdefg a
Heptad: HXXHXXX

As these examples from Human beta cardiac and Drosophila myosin demonstrate, the LMM does feature a traditional heptad repeat but it is not as regular as the literature may lead one to believe and the differences between species are likely to bring about diversity in helical content *in vivo*, possibly more so *in vitro*. It is important to remember that the near crystalline myofibril, with regular organized myosin coiled-coils making up the thick filament, becomes such only after extensive *in vivo* processing and with the assistance of a myriad of associated proteins. As such, in addition to different protein addendums, studies on coiled-coil-forming proteins sometimes have their heptad repeats re-enforced by non-native pro-helical aa substitutions; this has been done for tropomyosin [145].

The full sequence of the LMM region we examine by CD, referred to here as C600, spans the Drosophila muscle myosin II (P05661) from V1346-I1941 and contains an additional N-terminal 6xHis tag that results in a total length of 602 aa with a molecular weight of 68,903.48 daltons. The Drosophila LMM extends to F1962 but the C-terminal end (V1928-F1962) is not expected to form a helical coiled-coil, containing an additional five prolines. Notably, V1346-I1941 overlaps with the non-helical C-terminal region from V1928-I1941 and contains two prolines. The proximity of the N-terminal start point for this segment (V1346) to the second of the four myosin LMM skip residues (E1384) may also factor in to the propensity of the segment to form α helices and α helical coiled-coils.

LMM helicity by CD

Helical content can be determined either using programs that incorporate libraries of model proteins with known structure (CONTIN, CDSSTR, BeStSel, etc) or by a calculation based off of the magnitude of the MRE output at key positions known to be characteristic of α helices. The latter is most common in literature examining myosin and the LMM by CD and both evaluation at 208 nm and 222 nm has been used. In the more common case of MRE at 222 nm, the values associated with maximal helicity range from 36,000-49,000 and originate from equations developed for smaller synthetic peptides. The actual value attained at 222 nm for the peptide or protein of interest is then divided

by the maximal MRE predicted for a fully helical protein and that fraction becomes the predicted fraction of helical content.

Aside from guided helicity of the LMM by additional residues [143], some research has been done on purified LMM by circular dichroism (CD) and results in a range of output. Reports on the helical content of LMM segments range from <30% to 99% (**Table 3-1**). Early studies on LMM derived from cleaved rabbit myosin calculated >90% helicity for the LMM and 78% for whole myosin [146]. LMM cleaved from walleye pollack myosin was found to have 56% helicity while the recombinant LMM gave rise to 80% helicity in high ionic strength (600 mM) based off of the magnitude at 222 nm [147]. Other studies of the LMM, specifically beta cardiac MyHC found a range of <30% to up to 99% based on the magnitude of ellipticity at 222 nm [148-151]. The one study found that examined LMM segments of Drosophila myosin predicted 81% helicity by K2D [152]. Predicted helicities are dependent on buffer conditions in addition to the myosin sequence. Skeletal myosin extracted from rabbit to have ~56% helicity and 79% helicity from bovine with values decreasing by 11-33% in the presence of calcium [153] though such a study has not been done with the LMM alone.

Studies examining myosin II and LMM have all exhibited traditional helical character in which a positive band is present at ~195 nm and two negative bands exist at 208 nm and 222 nm [143, 146, 147, 149-153]. This has been our finding with the C600 as well (**Fig. 3-13**). Shown are C600 profiles for lower concentrations of C600 (2 μM) and for higher

concentrations of C600 (10 μM). However, we do not find the intensities at 208 nm or 222 nm to be intense enough to suggest bulk α helical content if we use the typical methods of calculation (see **Table 3-2**).

Multiple equations calculated fractional helicity (FH) based off of a comparison of the experimentally observed MRE at a given wavelength (Θ_λ^{exp}), and the predicted ellipticity for a protein with 0% (Θ_λ^{u}) to 100% helical content (Θ_λ^{h}) at that wavelength [154, 155]. The equation for FH is as follows:

$$FH = (\Theta_\lambda^{exp} - \Theta_\lambda^{u})/(\Theta_\lambda^{h} - \Theta_\lambda^{u}) \qquad \textbf{Eq.1}$$

The method for calculating Θ_λ^{h} used for short helices can be adapted for the LMM when considering that it is unlikely that the helix is continuous throughout the entire length (**Eq.2**). T is temperature in celsius, n is the number of helical units, and k is a finite length correction between 2.4-4.6 (see [156]). The shorter the helical segments or smaller the protein, the smaller the maximum magnitude for a completely helical protein. In many cases, variation on the following equation is used to determine a value to associate with 100% helicity:

$$\Theta_\lambda^{h} = (-44\,000 + 250T)(1 - k/n)] \qquad \textbf{Eq.2}$$

A value is also frequently assigned for a random coil to subtract from the intensity at 222 nm since random coils also have some magnitude in the negative at 222 nm. The value to subtract for random coil has been described as 2220-53T (Θ_λ^u). These values can be inserted into the FH equation as shown previously (**Eq.1**).

There isn't unanimous agreement regarding what values are best used for evaluating expected helicity and new methods are continuing to be developed. Some studies have focused on other points along the spectra in the 220-230 region [157], the slope between 230-240 nm [158] or the negative band at 208 nm [150]. To evaluate C600 helicity, we tried several methods including a 222 nm method described above using -36,000 to represent Θ_λ^h, a 208 nm method [150], the 230-240 nm slope method, and several CD processing programs based on various algorithms and gating parameters (K2D, CONTIN, CDSSTR, BeStSel) (**Table 3-2**).

The helical prediction for C600 at either the higher or lower concentration is similar with the higher concentration ranging from 8.3-28.7% and the lower concentration ranging from 8-28.5% α helical content. The greatest difference between the two is observed for CDSSTR and BeStSel. If results from programs that rely on data from protein libraries are removed, the range for high concentration (10 μM) is 15.3-28.7% and 14.15-28.5% helicity for low concentration (2 μM). The 222/208 ratio is helical in nature for both but with the ratio being more indicative of coiled-coil content for the higher [C600].

The values found for C600 indicate lower helicity than most LMM and LMM fragments published with the exception of a study by Wolny et al (2013) [149] suggestive of one possible explanation. In their study, beta cardiac myosin LMM peptides of 55-111 aa in length were examined for helicity, leading to the finding that only peptides that included amino acids 1301-1329 had >30% helicity even among wild type peptides with the appropriate hydrophobic residues at the 'a' and 'd' positions of the heptad. This was suggested as a possible trigger sequence for the LMM. A portion of this sequence between T1309-E1322 for beta cardiac myosin shares 64% identity and 71% positives with T1308-E1321 in Drosophila. The Drosophila LMM peptide used in our studies, V1346-I1941, excludes the trigger sequence.

Exclusion of a possible α helical trigger sequence (T1308-E1321) in the Drosophila sequence is only one of several considerations for the resulting low α helicity. As mentioned previously, C600 contains a C-terminal region that isn't predicted to be coiled-coil forming and contains two prolines (P1934, P1936); this region could contribute to the dilution of helical magnitudes or inhibit the stability of preceding helices. There is only one study that evaluates Drosophila LMM helicity by CD in the literature [152] but a segment of our size has not previously been evaluated by CD. Drosophila LMM is associated with thick filaments that are arranged differently from the striated muscle of vertebrates and require some different contacts; more processing may be required *in vivo*, not attained by *in vitro* processing, to achieve the α helical coiled-coil nature of the rods found in the thick filament.

Comparison to Drosophila LMM CD studies

Only one other study has examined a portion of the *Drosophila melanogaster* myosin LMM by CD [152]. Here, I will examine the parameters used in that study that result in a much higher observed helicity (81%). The most dramatic difference between the segment used by [152] and our own is the size and span. Our segment is ~600 residues in length (V1346-I1941) and includes both the 2nd skip residue and a portion of the non-helical tail region (V1928-I1941) which normally extends from V1928 to F1962 in Drosophila myosin. In [152], the non-helical tail is excluded along with more of the C-terminal, giving them a segment of 375 residues in length from E1495-L1870. The termination at L1870 is about eight heptads N-terminal from the non-helical region and nine heptads C-terminal from the last skip residue (4th - G1806), placing the terminal site as far as possible from the nearest disruptive regions and the local unwinding affiliated with them. Neither the first (T1187) or second (E1384) skip residues are included in their sequence and the 3rd skip residue, E1581 is still twelve heptads away from the first residue (E1495).

Some elements of purification are similar though the buffer used for CD is quite different. Salvi et al [152] used a T7 tag in addition to a His tag, though it appears that only the His tag is required for their purification procedure which was very similar to ours [152]. It is unclear how a T7 tag might influence the nucleation of a helix, but this remains attached to the LMM segment through purification. The final buffer used for CD

is 20 mM NaPO4, pH 7.4 w/ 500 mM NaCl. Sodium chloride is not recommended for use with CD as the chloride anions absorb strongly below 200 nm, demarcating 200 nm as their lower cut off. This higher ionic strength may be more limiting of aggregation than ours (20 mM Sodium Phosphate, pH 7 with 215 mM NaF, 1 mM TCEP) though no sedimentation experiments were done to examine this in either their study or ours. Salvi et al [152] studied aggregate size in "low salt" conditions, though the methods cite the same buffer as that used for CD (500 mM NaCl) so it is unclear if this is an error, or if filamentous aggregates were forming in this high salt buffer.

Among all the LMM studies, [152] is unique in their use of the Drosophila LMM, initiating the LMM segment well-in to the helical coiled-coil, and employment of K2D, the neural-network based program for analyzing the CD profile. Although K2D can sometimes give higher values for % helicity, as in our case (**Table 3-2**), it is second only to the 208 nm magnitude method in estimation of % helicity, the reason behind the higher helicity in [152] is more likely to be due to the strategy of placing the N- and C-terminal positions far away from areas of known destabilization and limiting the size of the fragment. It would be worth examining if the T7 tag may operate as an assist in nucleation of helical structure, especially as proline substitutions in regions expected to be sensitive (A1662P-g position; L1705P-a position; L1793P-d position) display no change in- or slightly increased- helicity in their CD profiles.

Other structural considerations

It is generally accepted that fragments of coiled-coils rarely fold without additional, often exogenous, trigger sequences [144, 159]. This may be an important characteristic for coiled-coils *in vivo*, particularly for the LMM. Folding only under guidance of additional scaffolding may be what enables the exact orientations and parameters necessary for the thick filament to operate. Stability of a coiled-coil is dependent on favorable ionic interactions either between i, i+3 and i, i+4, a tight hydrophobic core and shielding of that core from solvent. The number of helices in a structure also greatly increases its stability partially due to enhanced shielding of the hydrophobic core, with the most stable known coiled-coils being tetramers [159]. Close association of coiled-coil dimers is found within the thick filament and the LMM *in vitro* naturally trends to self-association by way of forming paracrystals at low ionic strengths. Without these conditions, the α helices and coiled-coil may be more likely to be in an unstable or transition state with variable structure.

Sensitivity, such as to ionic and mechanical forces, is also a known characteristic of the thick filament and it would follow that the components, LMM dimers, would likewise need to be sensitive in order to remain extensible. Myosin belongs to a group of fibrous proteins that exhibit reversible intramolecular transformation and may naturally exhibit fluctuating structure in solution which could dilute the spectral profile due to less absolute secondary structure being in existence. The helix-coil transition theory suggests that all residues can exist in either a helical or non-helical state and only once three

consecutive residues are in a helical state is a helix nucleated, though not necessarily in a traditional α helical form [155]. If a long sequence possesses multiple smaller segments of helices, frayed ends more similar to a coil or disordered helix may dominate. Circular dichroism doesn't give information on protein segments but provides a bulk character that constitutes total structural content averaged.

Although there are many reasons why helical structure might be diluted in the C600, it is important to recognize that there is α helical content identified. While the N-terminal trigger sequence is not present, the purported assembly competence domain (ACD) is. The ACD of human beta cardiac myosin is L1871-V1899 [160] and shares 71% identity, 78% positives with Drosophila muscle myosin II at L1870-K1897 included in C600. Once an α helix is initiated, continuation of the helix is cooperative. Thus, the purported ACD, or residues near it, may be where most of the helical content in the *in vitro* C600 lies. Once the triggered helix is broken, disordered "frayed" helix or coil states may dominate. Disordered or non-traditional α helical content (ex. 3_{10} helical) would fall into the 'Other' category so imperfect helical content may be existent within the LMM even with a lower <30% predicted specifically α helical component and this may also explain the dominance of an α helical profile despite the lower α helical predicted proportion.

For future studies, shortening the examined segment, adding a trigger sequence/helix-promoting peptide, or making adjustments to the buffer solutions to be more similar to the dense native milieu, such as increasing packing using beta-alanine [161, 162] or

trehalose [163], including calcium or excluding reducing agents [145] may allow increased stability of α helical content. This may allow α helical content, and the impacts of binding upon it, to be more resolved.

Behavior of C600 in the context of WYR

One way of examining protein-protein binding is to look at the Actual experimental output of C600+WYR and compare it to a Theoretical output obtained from the combination of the experimental values for C600 alone and WYR alone before conversion to MRE. The Actual and Theoretical profiles are expected to be the same under conditions of non-binding. As shown in **Fig. 3-14**, the magnitudes are distinctly different between Actual and Theoretical profiles while the pattern remains reminiscent of an α helix. Specifically, the 222/208 ratios are different (Actual: 1.13±0.06 ; Theoretical: 1.05±0.07; p=0.03). Additional qualitative information can be gleaned after MRE conversion that informs on the structure formed by the combination of LMM and WYR experimentally.

When looking at the combined experimental mdeg (Actual) and combined theoretical mdeg (Theo), the signals for WYR and C600 are not separated out and the proportion of the signal due to each component is not evident. As WYR is smaller than C600 by a factor of ~10, WYR alone spectra experiences a greater magnitude adjustment (~10) than C600 spectra when these are converted to MRE. At the mdeg level, C600 spectra would

be expected to naturally dominate and persistence of the helical-like profile seen in **Fig. 3-14** is expected for 'theoretical' and not surprising for 'actual'.

In addition to non-equal elliptical shifts due to a size difference, the WYR alone spectra is distinctly qualitatively different from the C600 spectra (see **Figs. 3-15, 3-16**). Because of this, the mdeg proportions between WYR and C600 over different wavelengths are distributed differently. Under the condition of non-binding, the proportionality is determined using the Theo spectra in which, WYR/(C600+WYR) and C600/(C600+WYR) is evaluated to attribute a non-binding proportion at each wavelength. When taking this into consideration, the mdeg values of Actual can be separated into theoretical non-binding proportions. At that point, each of those values could be converted to MRE and then added together, referred here as Act(sep), and this could be directly compared to the theoretical MRE (LMM-alone MRE value plus WYR-alone MRE value added together). The Act(sep) output could then also be evaluated for structure on its own using, for example, BeStSel. It is important to recognize that the calculation of Act(sep) as it is does not accommodate variation on the y-axis between C600 and WYR, leading to large standard deviations in regions of sign change (positive to negative or vice versa) (**Fig. 3-17**). An accommodation for this is shown later.

Aside from the Act(sep), the other option is to treat the WYR and C600 components as a unit which is a realistic consideration since, upon binding, WYR and C600 would behave as a unit and so retention of the theoretical proportionalities is unlikely. In this case MRE

conversion of the WYRC600 ellipticity uses “combined” parameters that include a weighted averaging of concentrations and MW. This version of Actual, Act(comb) (**Fig. 3-17**), can then be used to evaluate structure of its own using BeStSel.

Centrifugal sedimentation experiments showed that ~30% of LMM sediments in the presence of WYR in the same buffer and WYR:C600 ratio as used for CD (**Table 3-3**). Given that the Act(comb) profile is similar to C600 alone (**Fig. 3-18**), we examined whether the signal intensity decrease seen in Act(comb) relative to C600 alone matched what might be expected if sedimentation occurred. Act(comb) contains both WYR and C600 so this comparison contains the assumption that WYR is incorporated into LMM structure. We find that, between 200-240 nm, the average signal intensity is 55% for Act(comb) compared to C600 alone with a median relative decrease to 67%. This would suggest that 33-45% of the structural component would have to be removed (sedimented out) to account for the shift in magnitude.

This same procedure was employed between Theo(sep) and Act(sep) (**Fig. 3-19**) that exhibit similar profiles. Between 200-240 nm, the average signal intensity is 50% for Act(comb) compared to Theo(sep) with a median relative decrease to 59%. This would suggest that 41-50% of the structural component would have to be removed (sedimented out) to account for the shift in magnitude. Given that ~30% of LMM is found to sediment with WYR under centrifugal conditions, this suggests that the decrease in magnitude reflects a structural component shift and is not strictly due to sedimentation.

In order to achieve a version of Act(sep) without the high error brought on by the extreme differences in sign when C600 and WYR proportionality is calculated, we created a version (Act(sep*)) in which the baseline of the raw data (mdeg) was moved. By moving the baseline such that no value could be negative, proportions for C600 and WYR were re-calculated and applied to the Actual mdeg. The baseline was re-established, returning the sign character to the spectra. The resultant Act(sep) with accompanied Theo(sep) and Act(comb) with accompanied Theo(comb) is shown in **Fig 3-20**.

Using the same calculations that we used to evaluate helicity for C600, we obtained a range of helicities from 5-23.7% for Act(comb) and 14.1-45.6% for Act(sep*) which can be compared to 15.1-29.8% for Theo(comb) and 19.3-51.6% for Theo(sep) (**Table 3-4**).

If we view Act(sep*) through Act(comb) as a range of structural possibilities from conserved proportionality to single unit calculations, BeStSel reports 7.4-16.9% helicity; 17.9-31% antiparallel β, 0-1.2% parallel β, 15.3-17.6% turn, and 45.3-47.6% 'other' content. Greatest variation is seen for helical and antiparallel β content.

The 222/208 ratios, indicative of coiled-coil nature for higher values, are higher in the actual experimental combination of LMM and WYR compared to the theoretical non-binding scenarios in conserved proportionality 'Sep' and single unit 'Comb' methods. The 222/208 ratio is 0.7 for Theo(sep) and 1 for Act(sep). This ratio is also found to be

increased compared to the theoretical nonbinding scenario with a 222/208 ratio of 1.1 for Theo(comb) and 1.2 for Act(comb).

As magnitudes at ~195 nm, 208 nm and 222 nm largely dominate the predictions of helical structure, it is unsurprising that helical predictions for Theo(sep) are larger than those for C600 alone though 222/208 ratios may be considered to contrast the suggestion that helical content is decreased in the experimental combination of WYR and LMM. If we turn to 222/208 ratios observed in the raw data, the Theoretical exhibits a 222/208 ratio of 1.05±0.07 while the Actual experimental 222/208 ratio is 1.13±0.06; as a ratio of ~0.9 is indicative of α helical content and above 1.1 is indicative of coiled-coil content this would suggest more coiled-coil content in the binding scenario, a tertiary structure known to be important for stabilizing helices. C600 alone exhibits a 222/208 ratio of 1.08-1.17 which appears to be (at least) retained when it is considered that WYR and the LMM bind and can be considered a unit as Act(comb) has a 222/208 ratio of 1.16. More likely, as comparisons between corresponding Actuals & Theoreticals are more appropriate, there is increased presence of coiled-coils alongside a decrease in alpha helicity, when programs using established algorithms with basis datasets are given emphasis (CONTIN, CDSSTR, BeStSel).

Taken together, the significantly increased 222/208 ratio in the Actual experimental output and variable but generally decreased α helical content proposed by program methods suggest that i) α helices are either decreased over the entire length of the LMM

or further segmented, ii) helical content may be more disordered or associated with non-α type H-bonding, and iii) more of the remaining alpha helices are taking the form of a coiled-coil.

Possible role of Beta content

BestSel processing of Act(comb) also suggests a much higher antiparallel β content than Theo(sep) (31% vs. ~21.3%) and Act(sep*), a slight increase (17.9% vs 17%). Several possibilities come to mind that may account for this. This could be due to (i) a transition to higher β, rather than α, content due to disrupted polarity of the LMM's α helix as β-like transitions become more abundant due to an increase in side-side interactions with WYR or due to higher order oligomerization, (ii) stabilization and increase of WYR's own antiparallel β content as it establishes contacts with the LMM, or (iii) some combination.

The dilution of helical structure could be due to such increased β content. This could be due to the simple increase of antiparallel β content in WYR or the LMM, in addition to segmentation of α helical content, or overall loss of α helical structure affiliated with a change in solvation properties. It is known that helical coiled-coil proteins can form helices composed of β sheets which can either be responsible for mechanical functions or involved in a structural transition state [164-166] and the myosin LMM is capable of forming β structure at the expense of the α helix [167].

ADDITIONAL MATERIALS AND METHODS

Purification of MHC Fragment

A 68.9 kDa 602-amino-acid peptide encompassing V1346 through I1941 of *D. melanogaster* myosin heavy-chain (herein C600) with an N-terminal 6x His-tag cloned into pET-23a vector was transformed into *E. coli* BL-21(DE3) pLysS cells [8]. Cells were grown in Luria broth (LB) containing 100 µg/ml ampicillin and 50 µg/mL chloramphenicol. Upon reaching A_{600} of 0.8-1.0, culture was chilled on ice for 15 minutes, and expression was induced by adding IPTG to 0.75 mM and incubating for 16 hours at 25˚C with gentle rocking. Induced cells were collected by centrifugation at 10,000xG for 10min and stored as pellets at -40˚C. Pellets were resuspended in lysis buffer described by Korkmaz et al [143], and lysed by sonication. To purify the MHC peptide, the lysate was rocked at 4˚C with Ni-NTA-agarose resin for one hour in a 20mL column before being rinsed with 8 column volumes of 20 mM imidazole wash-buffer. The resin was then rinsed in two 15mL portions of each 40- and 80 mM imidazole wash buffer, with 10 minutes of rocking in each 15mL portion. The peptide was eluted in six 1ml fractions and one 12 mL fraction for a total volume of 18 mL 200 mM imidazole elution buffer. All wash and elution buffers are variants of 'Buffer A' described by Korkmaz et al. with the following modifications: imidazole concentration was adjusted to 20 mM, 40 mM or 80 mM for wash buffer and to 200 mM for elution buffer [143]. The relative purity of each fraction was assessed via SDS-PAGE stained with Krypton (Thermo Scientific). Selected fractions were further purified and concentrated with an Amicon Ultra-15 50k centrifugal filter, and buffer exchanged into 400 mM NaF, 20 mM

sodium phosphate pH 7.0, 1 mM TCEP for storage. Protein concentration was determined via absorbance at 280 nm using an extinction coefficient of 18450 cm^{-1} M^{-1}. Densitometry using ImageJ was done on Krypton stained gels to ascertain purity of the sample [168] and the peptide identity verified by mass spectrometry.

WYR Peptide

Synthetic, 6.6 kDa 52-amino-acid WYR peptide encompassing H84-T135 of *D. melanogaster* flightin was sourced from Genscript (Piscataway, NJ). Peptide was suspended in ddH_2O, filtered through a 0.22 µm Millipore filter, and concentration determined by micro BCA assay and absorbance at 280 nm using an extinction coefficient of 21430 cm^{-1} M^{-1}.

Circular Dichroism Sample Preparation

WYR and C600 samples ranging from 0.5 to 10 µM were prepared in 700 µl of 215 mM NaF, 20 mM Sodium Phosphate, pH 7, 1 mM TCEP. C600 preparations were centrifuged at 10,000 rcf for 10 minutes and filtered through a 0.22 µm Millipore filter. Blank samples used an equivalent volume of filtrate from the final flow through from C600 concentrators as used in the experimental samples to account for any possible negligible contaminant. These blanks were checked against equivalent freshly prepared buffer. A rectangular STARNA quartz cuvette with a pathlength of 0.2 cm was used throughout. In between experiment days the cuvette received a full wash with the provided STARNA

detergent in accordance to STARNA protocol [169]. In between samples, 1x 60% EtOH and 3x dH2O rinses were done and residual fluid was evaporated with pressurized nitrogen. All samples were preserved on ice after each experiment until further use, or discarded.

Circular Dichroism Measurements and Analysis

Samples were measured at 25˚C using a Jasco J-1700 spectropolarimeter at a scanning-speed of 20 nm/min with a digital integration time (DIT) of 8 s and bandwidth of 0.5-1 nm over six accumulations, minimally from 260-190 nm on continuous scan mode. A minimum n=6 was used for experimental combination of C600 and WYR and an n=10 for WYR alone measurements in 215 mM NaF, 20 mM Sodium Phosphate, pH 7, 1 mM TCEP. Wavelengths at maxima and minima, as well as $\theta_{222/208}$ and $\theta_{192/208}$ ellipticity ratios were recorded for all spectra. Structural interpretations were generated using BeStSel, after conversion to Mean Residue Ellipticity (MRE), for the range of 190-250 nm [170, 171].

Independent spectra for each peptide were evaluated after subtracting baseline spectra. To account for concentration and peptide molecular weight, the resultant ellipticity in millidegrees (mdeg) was converted to MRE using the following equation:

$$\text{Ellipticity} = mo*MRW/(10*L*C) \qquad \textbf{Eq. 3}$$

Where mo is the millidegrees ellipticity at a given wavelength, MRW is the mean residue weight which is molecular weight of the peptide divided by number of amino acids in length minus one (M/(N-1)) in daltons, L is the cell pathlength in cm and C is the peptide concentration in g/L.

For experiments in which the two peptides were combined (WYRC600), the experimental output ("Actual") was compared to a value representing nonbinding conditions ("Theoretical") attained by adding the independent ellipticities of C600 and WYR with buffer subtracted. Differences in magnitudes between "Actual" and the additive "Theoretical" CD profiles was indicative of binding.

To evaluate the structural profile of C600 and WYR together, the combined (WYRC600) spectra with buffer subtracted was converted to molar ellipticity using two distinct methods: A Separated Parameters Method and a Combined Parameters Method.

Separated Parameters Method

For this method, we calculated the Theoretical (nonbinding) proportions of the spectra that WYR and C600 contribute at each wavelength ($\{\lambda EZ | 190 \leq \lambda \leq 260\}$) by dividing the total Theoretical mdeg output by WYR alone or C600 alone at each wavelength observed. This fraction was then used to separate the Actual mdeg output into two data sets, one representing the C600 proportion and the other representing the WYR

proportion. The resultant set of values for the WYR proportion was then subjected to MRE conversion to WYR specifications (molecular weight, concentration) and the resultant set of values for the C600 proportion was subjected to MRE conversion to C600 specifications. These two sets of values were then added to attain Actual(separate) or "Act(sep)". To avoid high error arising from large differences in sign (+/-) between the C600 only and WYR only profiles used to estimate proportions, the baseline was shifted to evade sign change at the mdeg level and then re-established after MRE conversion.

Combined Parameters Method

This method treats C600 and WYR of the WYRC600 experimental output as one unit. In this case, the mdeg of WYRC600 is converted to MRE (**Eq. 3**) using combined parameters: the average molecular weight of C600 and WYR and the sum of their concentrations.

Further accommodation was made for differences in C600 concentration between experiments by equalizing the [C600] for all experiments, also involved in the calculation of theoretical and actual outputs. To do this, the median mdeg value at the wavelength that exhibited the least variation between all C600 only outputs (215 nm) was used as a marker by which all C600 outputs were compared. The relative value attained, representing deviation in magnitude either above or below the median, was used to scale mdeg values for each experiment.

Additional measurements for helicity used equations based on the 222 nm magnitude [149], 208 nm magnitude [150] or 230-240 nm slope [158] and several programs, in addition to BeStSel, from Dichroweb [172, 173] were used: K2D, CONTIN and CDSSTR. CONTIN and CDSSTR output were based on the SMP180 reference dataset.

Cosedimentation Assays

Cosedimentation assays were performed with C600 titrated by WYR in quintuplicate. Components were added to a final buffer of 215 mM NaF, 20 mM Na-P, 0.5 mM TCEP in a total volume of 60 µl. Twenty microliters were removed and labelled Pre-Spin (PS). Solutions were incubated overnight at 4˚C and then centrifuged at 15,000 rcf at 4˚C for 10 minutes. Twenty microliters of supernatant (S) were separated without disturbing the pellet. The remaining 20 µl of solution were included in the pellet (P) fraction and accounted for in the calculations, as described in the next section. Sedimentation of C600 with Insulin (5.8 kDa, 51 aa) was used as a control, done in triplicate, to determine non-specific binding/sedimentation.

Samples were combined with 5 µl of 5x Sample Buffer (50% glycerol; 300 mM Tris HCl, pH 6.8; 10% SDS; 0.05% Bromo Blue, 125 mM DTT) and boiled for 10 minutes. The samples were loaded into 15 well 15% SDS PAGE minigels and run at 170 V. Gels were fixed in 40% Ethanol, 10% Acetic acid with one exchange at 20 min. They were rocked in this solution overnight. Krypton (Thermo Fisher) was used to stain the gels for

2.5 hrs and destained in accordance to Krypton protocol prior to viewing. Gels were viewed by Gel Doc (Bio-Rad) and the .tiff files were exported for analysis.

Densitometry

Image Studio Lite Ver 5.2 (LICOR) was used to assess band density of C600 and WYR. Intensity values from 'S' were subtracted from 'P' to attain adjusted P (P-S). To correct for non-binding precipitation, the ratio (P-S)/P was calculated for all trials and the (P-S)/P of C600 alone was subtracted for C600 intensity values in the presence of WYR; (P-S)/P of WYR alone was subtracted for WYR intensity values in the presence of C600. Insulin (20 μM) was used as a negative control and any C600 that sedimented with Insulin was subtracted from the C600 sedimenting with WYR as non-specific binding. The fraction that pelleted beyond baseline for both WYR and C600 was converted to μM quantities based on loading amount.

Graphpad PRISM was used to fit the specific binding and calculate parameters (Kd and Bmax).

FIGURES

Figure 3-1: The D. melanogaster WYR sequence. Conserved residues throughout Pancrustacea underlined, aromatic residues with a circle above, a + mark above positively charged residues, a – mark above negatively charged residues. Characteristics of the *D. melanogaster* WYR sequence include high aromatic content (19%), with just over 30% charged residues, non-alternating hydrophobic residues (25%), 31% polar residues (including tyrosines) and 3 prolines (6%).

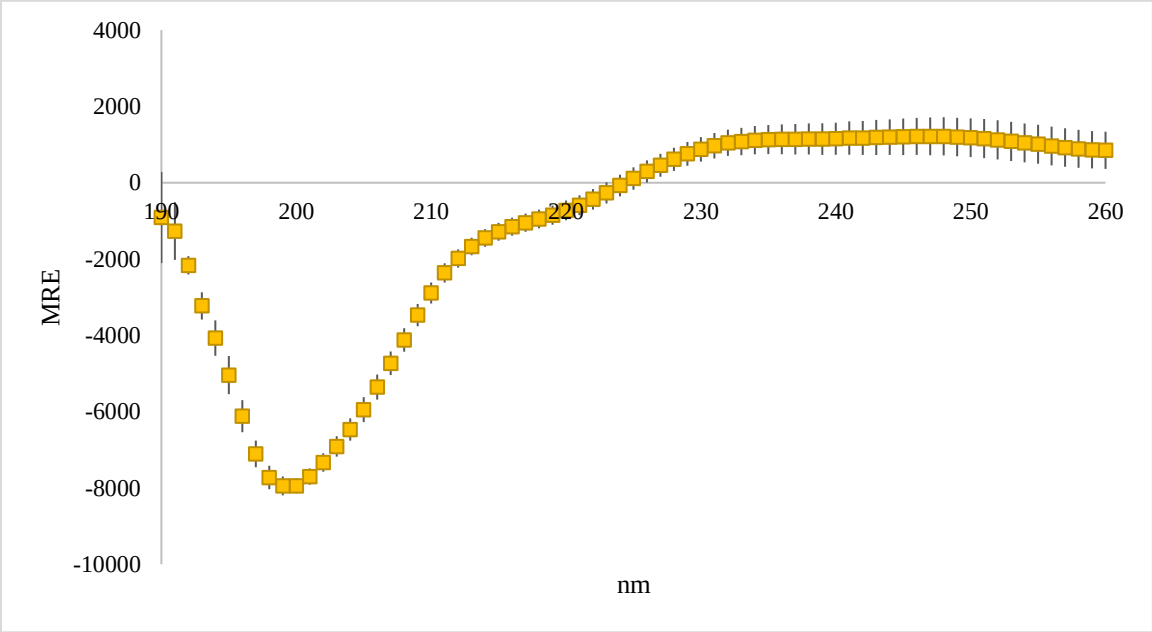


Figure 3-2: Spectra of [WYR] at 10 µM in 215 mM Sodium Fluoride, 20 mM Sodium Phosphate buffer. This displays a near-zero rising band at 190 nm, negative band at 200 nm and a shoulder centered ~220 nm. Each point represents mean ± SEM, n=10.

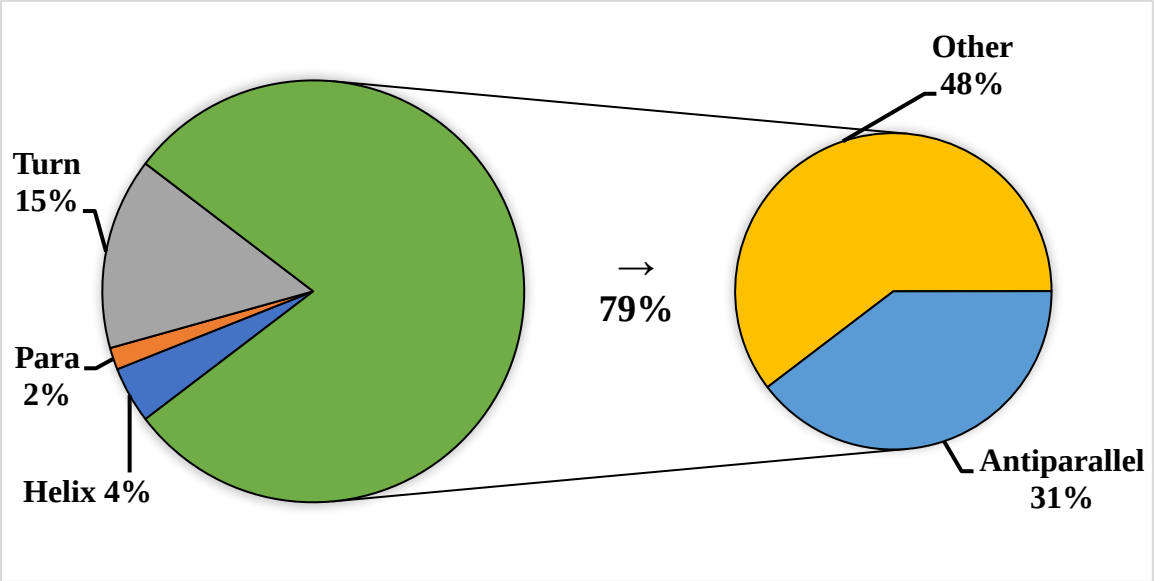


Figure 3-3: BeStSel interpretation for WYR (10 µm) spectra. This shows predominantly "Antiparallel β" and "Other" content, followed by "Turn" content. Parallel β and α Helix content is not accurate due to the size of WYR (52 aa).

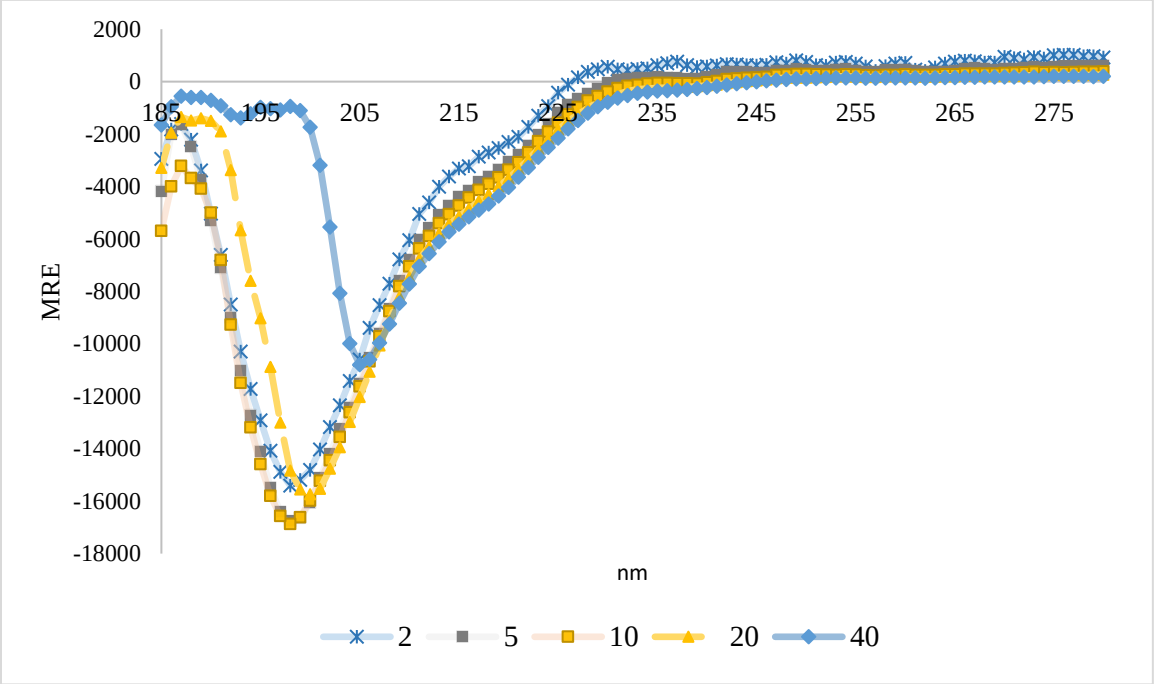


Figure 3-4: Increasing [WYR] from 2 µM to 40 µM in dH2O. This results in a shift of the 190 nm band to ~200 nm and ~200 nm negative band to ~205 nm while the shoulder centered at ~220 nm remains stable.

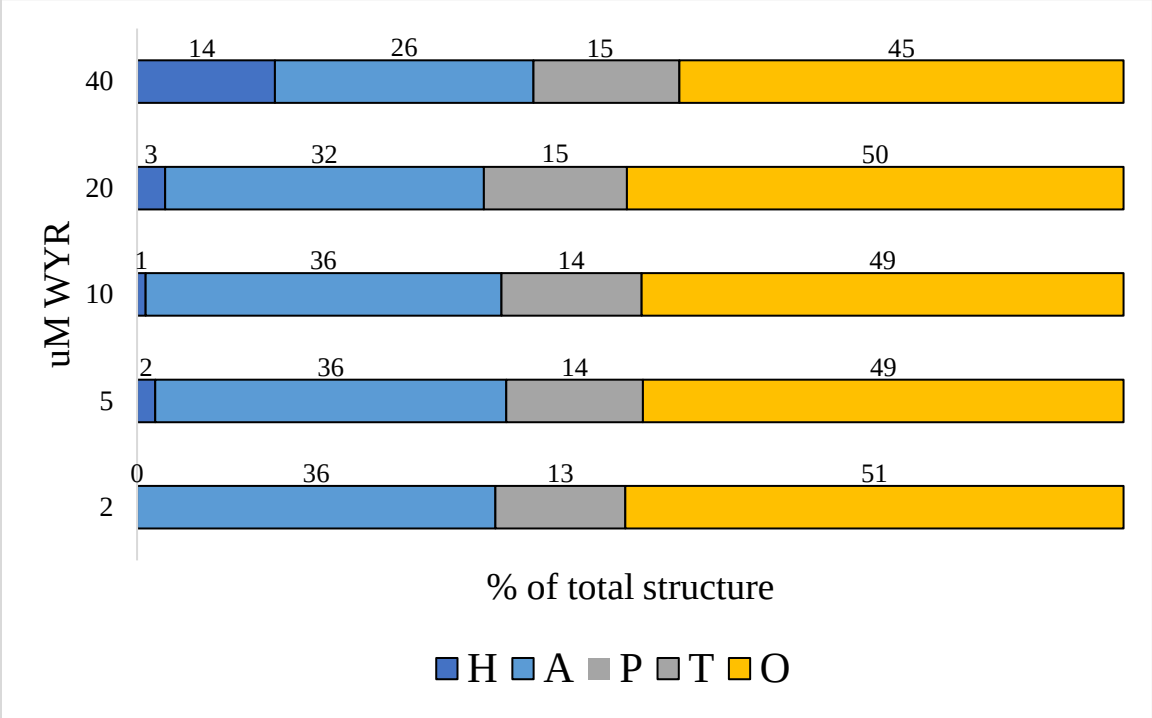


Figure 3-5: BeStSel interpretation for increasing [WYR] in dH2O. Structural change in WYR over 3-40 µM determined by BeStSel (190-250 nm) indicates increasing α helical content (H) at the expense of Antiparallel β (A) content and some Other (O) content while the proportion of turns (T) remains largely unchanged. No parallel (P) content was observed. Values are rounded.

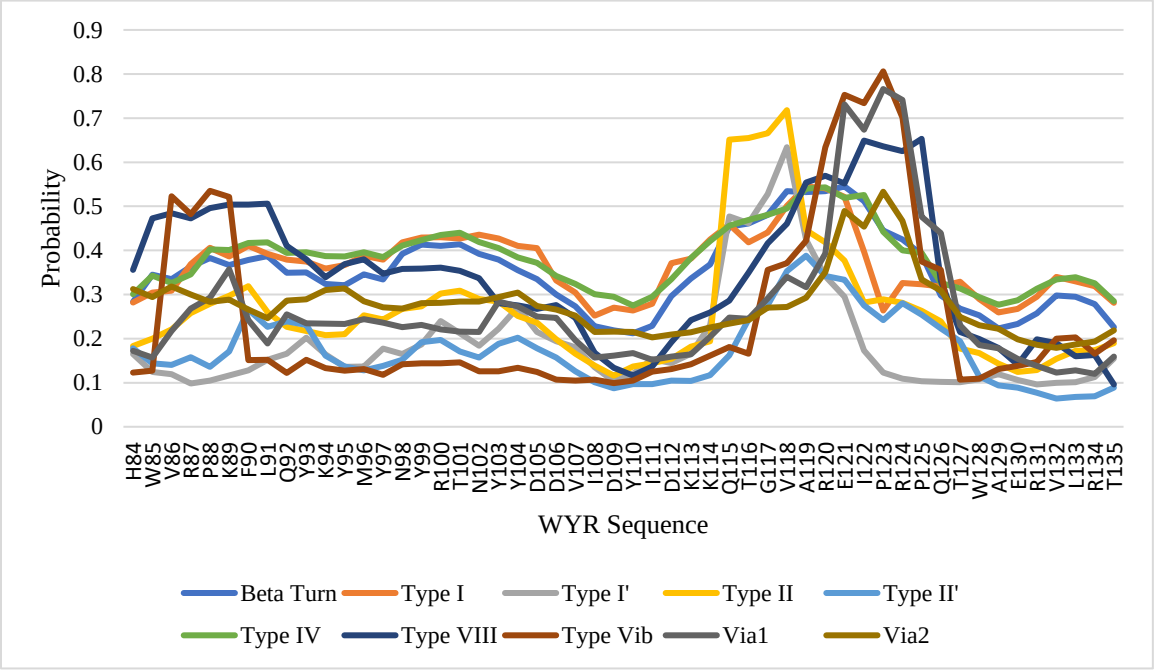


Figure 3-6: NetTurnP prediction of β turn propensity along the *D. melanogaster* flightin WYR sequence. The greatest propensity is predicted between K114-Q126 of the flightin sequence.

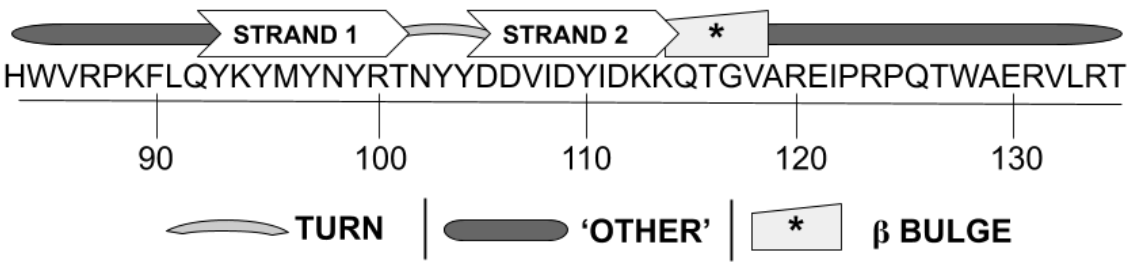


Figure 3-7: Pictograph of hypothesized location of secondary structure of WYR. The numerical axis represents the #aa of the *D. melanogaster* flightin sequence. Strand 1 is the first strand of a 3-strand antiparallel β hairpin that encompasses the tyrosine ladder between Q93-T101 and is separated from Strand 2 by a turn segment expected to encompass N102-Y104. A β turn/G1 β bulge reliant on G117 is hypothesized immediately C-terminal to Strand 2. Flanking the β and turn components is additional 'Other' content which may include loop and further turn structures.

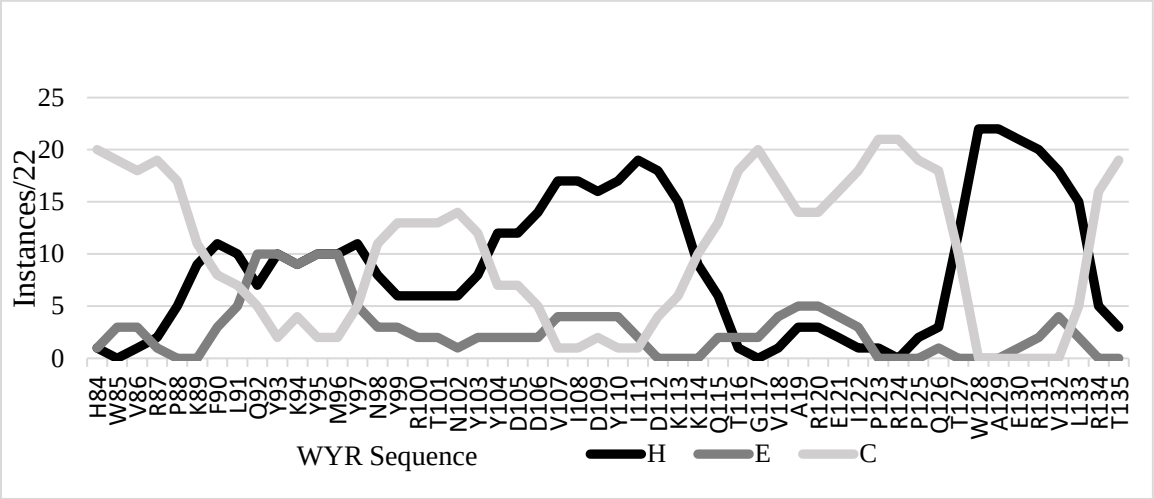


Figure 3-8: Secondary designation of the WYR sequence for 22 secondary structure prediction programs. Turns not shown. H = helix, E = extended (β) , C = coil.

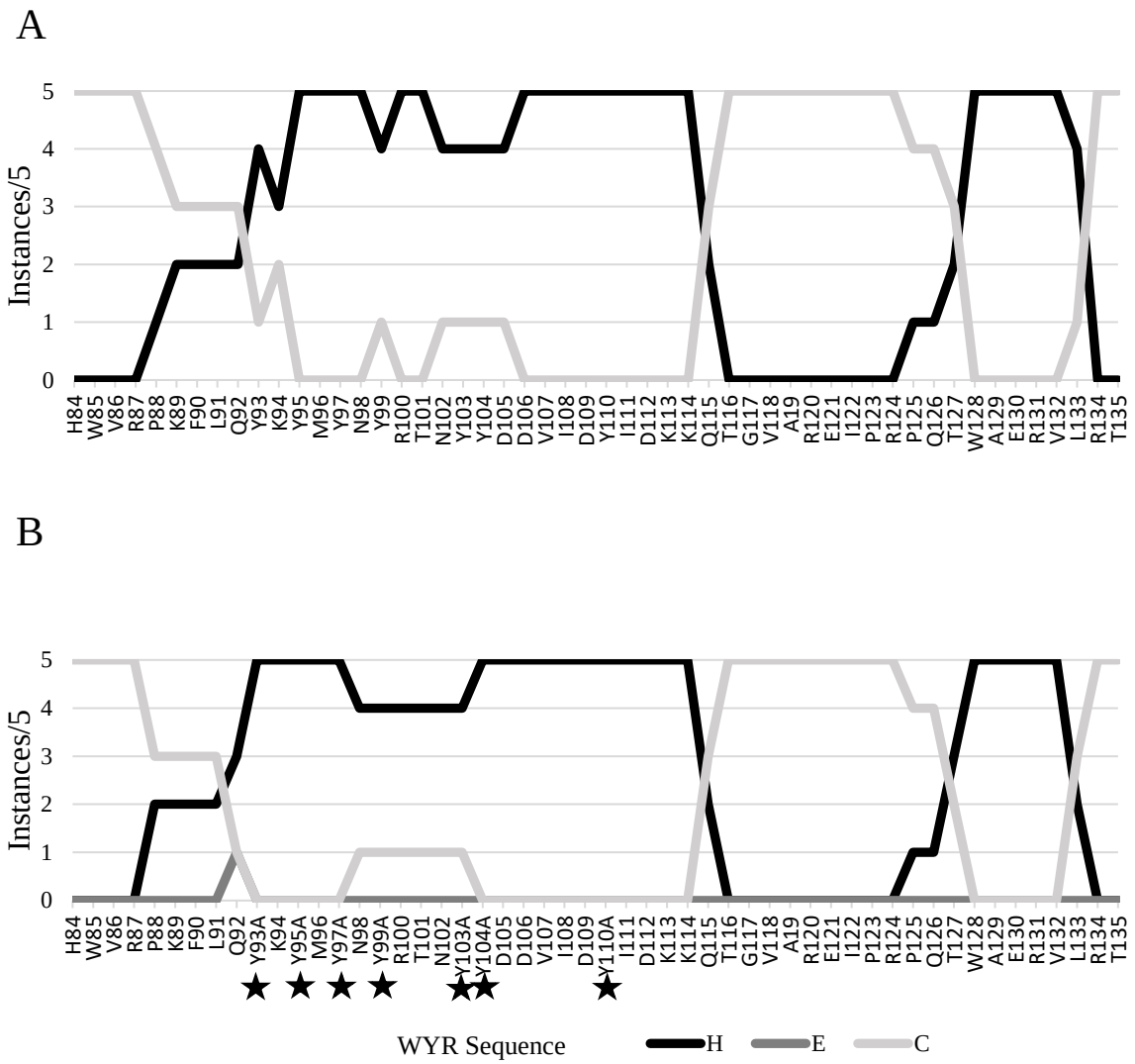


Figure 3-9: Secondary structure designation for WYR structure of 5 programs capable of finding Extended β strand content but only detecting Helical or Random Coil content. (A) Original designations. (B) Secondary structure designations for the 5 programs originally predicting only H/C content in WYR when all tyrosines have been changed to alanines, shown on the x-axis (stars). H = helix, E = extended (β), C = coil.

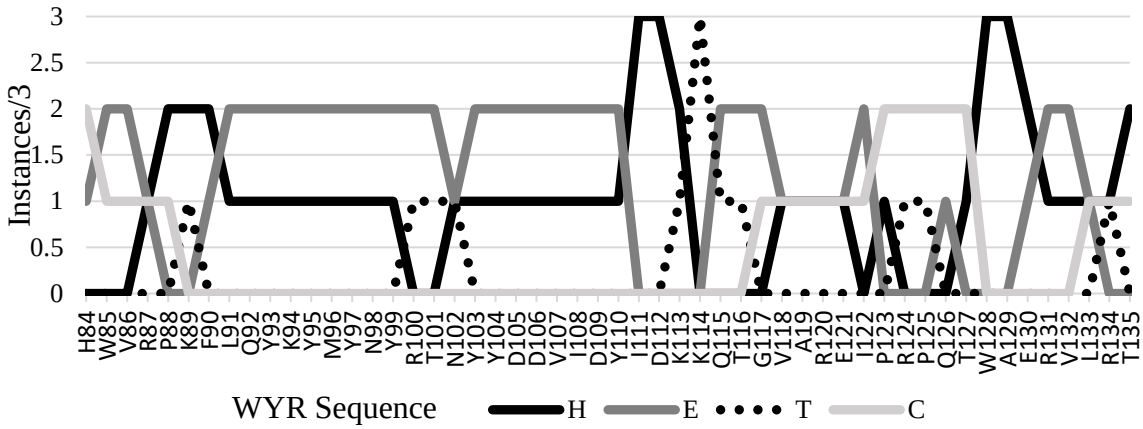


Figure 3-10: Secondary designation from turn-inclusive programs. H = helix, E = extended (β), T = turn, C = coil

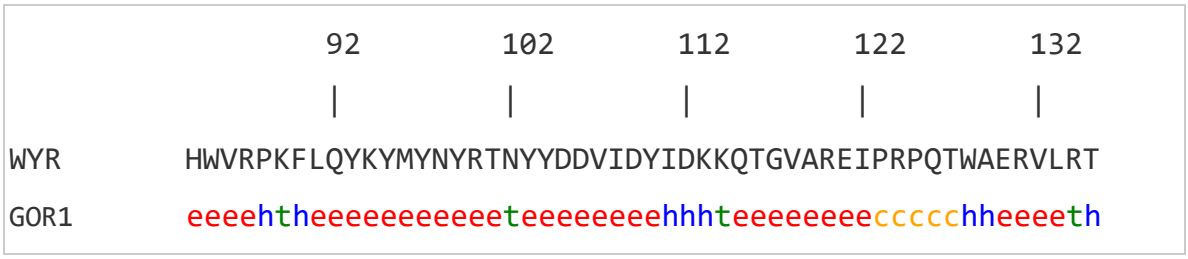


Figure 3-11: Secondary structure designation of WYR by GOR1. h = helix, e = extended (β), t = turn, c = coil. Numbers at the top indicate *D. melanogaster* amino acid positions.

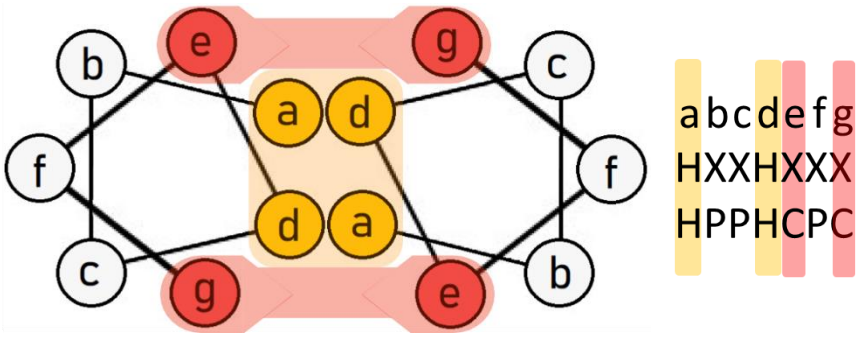


Figure 3-12: Two α helices form a coiled-coil via their 'heptad' repeats. Positions 'a' and 'd' form a hydrophobic core and charged residues at positions 'e' and 'g' add further stabilization. Although HPPHCPC is the ideal heptad pattern, HXXHXXX is also an accepted heptad pattern. X can be any type of residue.

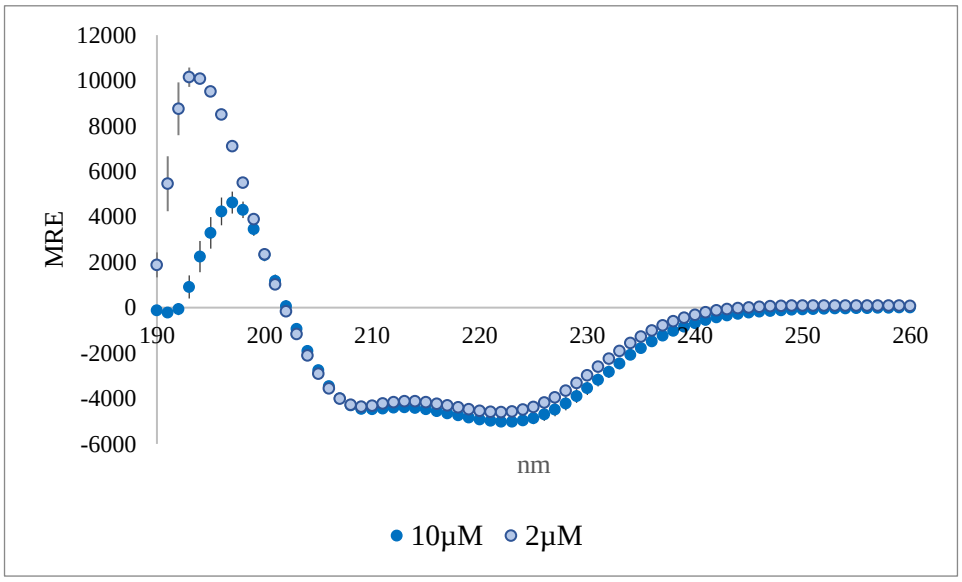


Figure 3-13: MRE CD profile of C600 at 10 µm and 2 µM from 190-260 nm. Both display a characteristic α helical profile. The positive band for lower concentrations of C600 is at 193-194 nm while higher concentrations show a bathochromic shift to 197-198 nm. Each symbol represents mean ± SEM N=3 for 2 µM; N=6 for 10 µM.

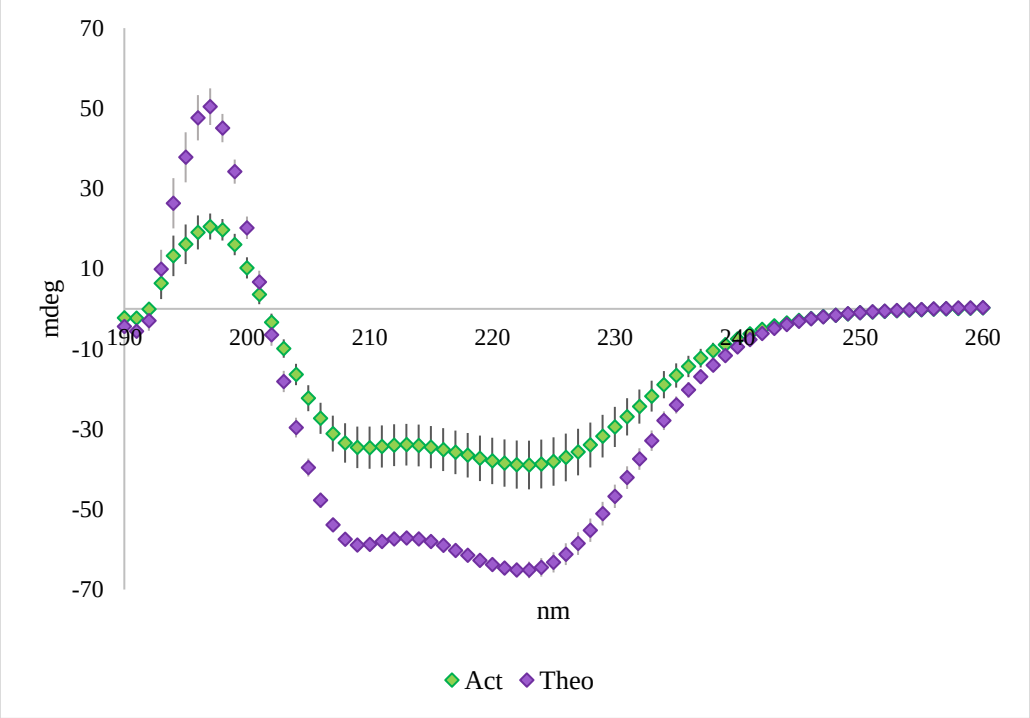


Figure 3-14: Experimental (Act) and Theoretical (Theo) C600+WYR CD profiles. Averaged Actual experimental output of C600+WYR compared to the Theoretical expected for a nonbinding C600+WYR combination suggests that binding is taking place. Each symbol represents mean ± SEM N=6.

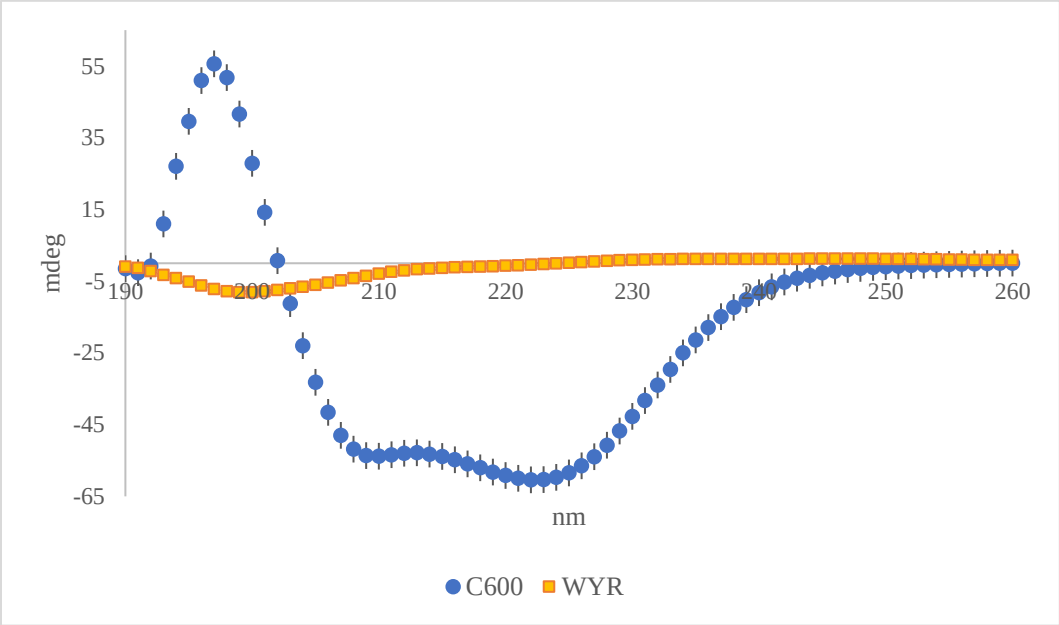


Figure 3-15: Comparison of CD mdeg profile of C600 (10 µM) to WYR (10 µM). Each symbol represents mean ± SEM. N=6 for C600; N=10 for WYR.

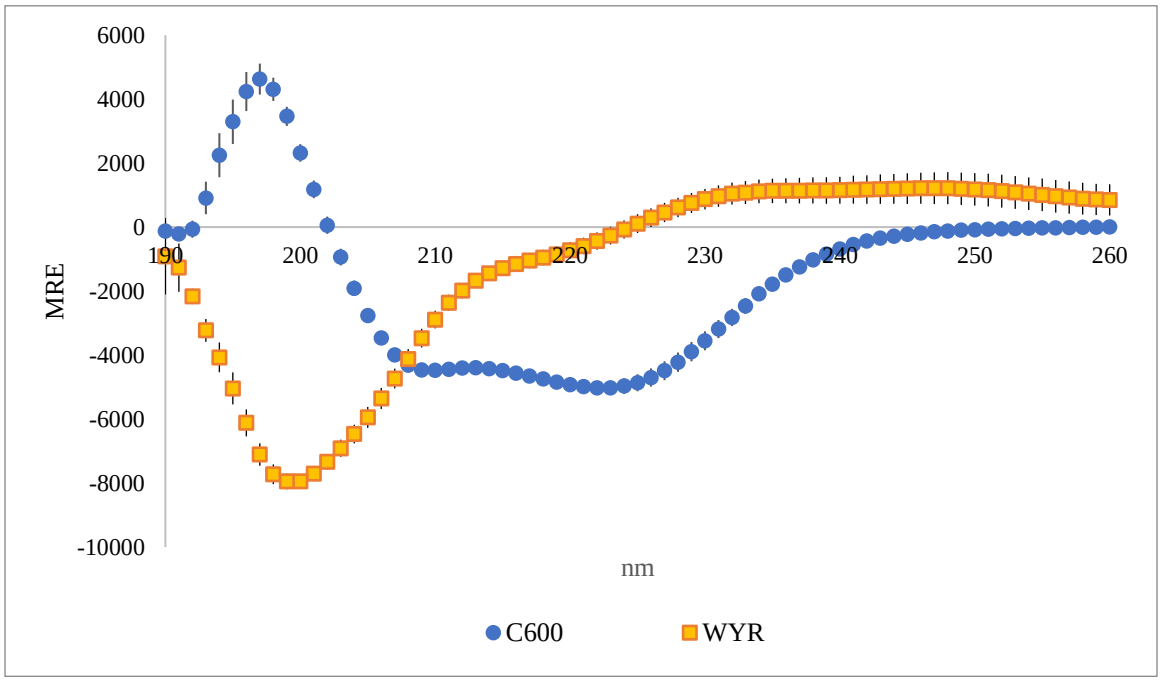


Figure 3-16: Comparison of the CD MRE profile of C600 (10 µM) to WYR (10 µM). Each symbol represents mean ± SEM N=6 for C600; N=10 for WYR.

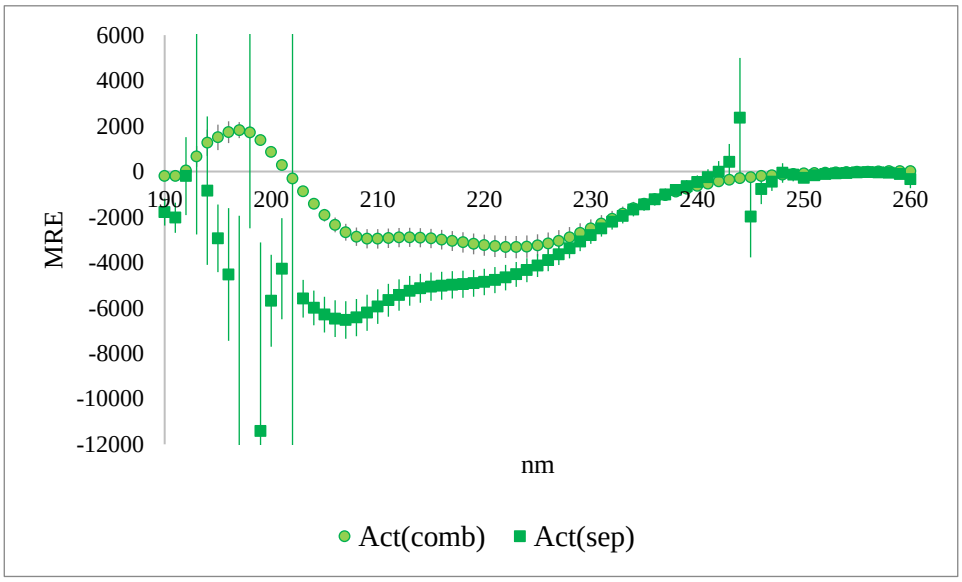


Figure 3-17: Act(comb) and Act(sep) CD profiles for C600+WYR. Both profiles can be used to evaluate structural possibilities present in the experimental combination of C600WYR. Large variation in Act(sep) is present in regions where either the WYR and/or C600 spectra normally undergo a sign change. Each symbol represents mean ± SEM. N=6.

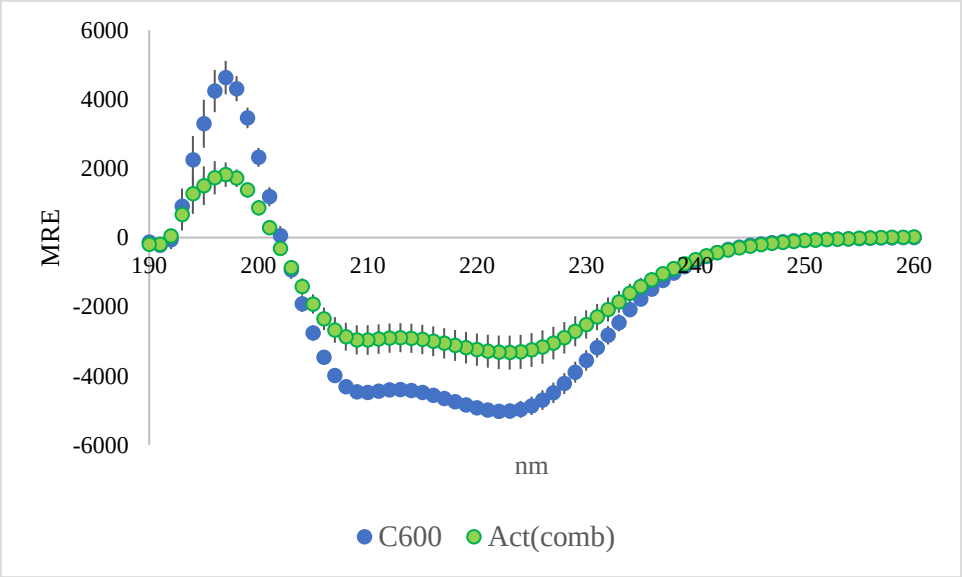


Figure 3-18: C600 vs. Act(comb) CD profiles. These profiles can be quantitatively compared due to their similar helical-type structures. Each symbol represents mean ± SEM. N=6.

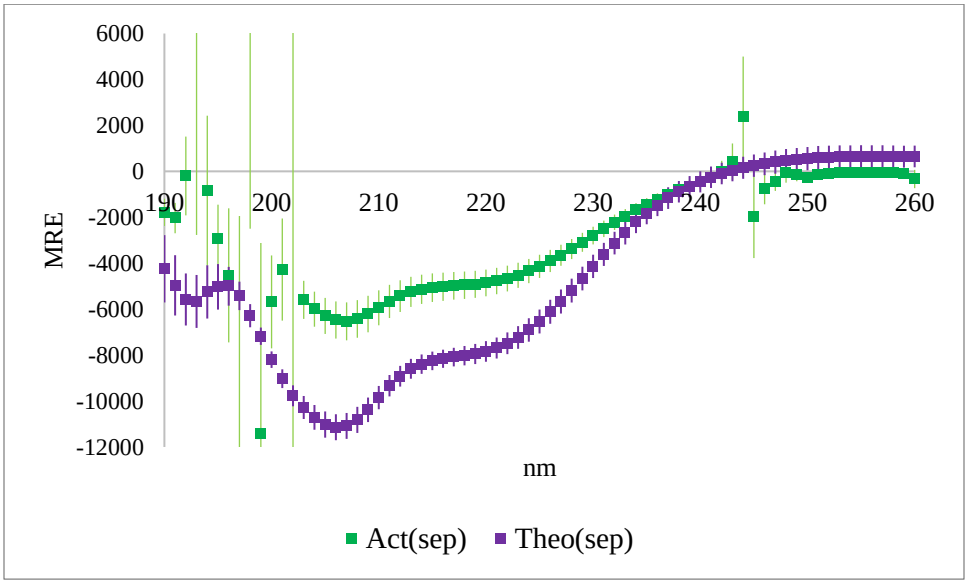


Figure 3-19: Act(sep) vs Theo(sep) CD profiles. These can be both qualitatively and quantitatively compared as they both contain information for C600 and WYR together and possess similar profiles.

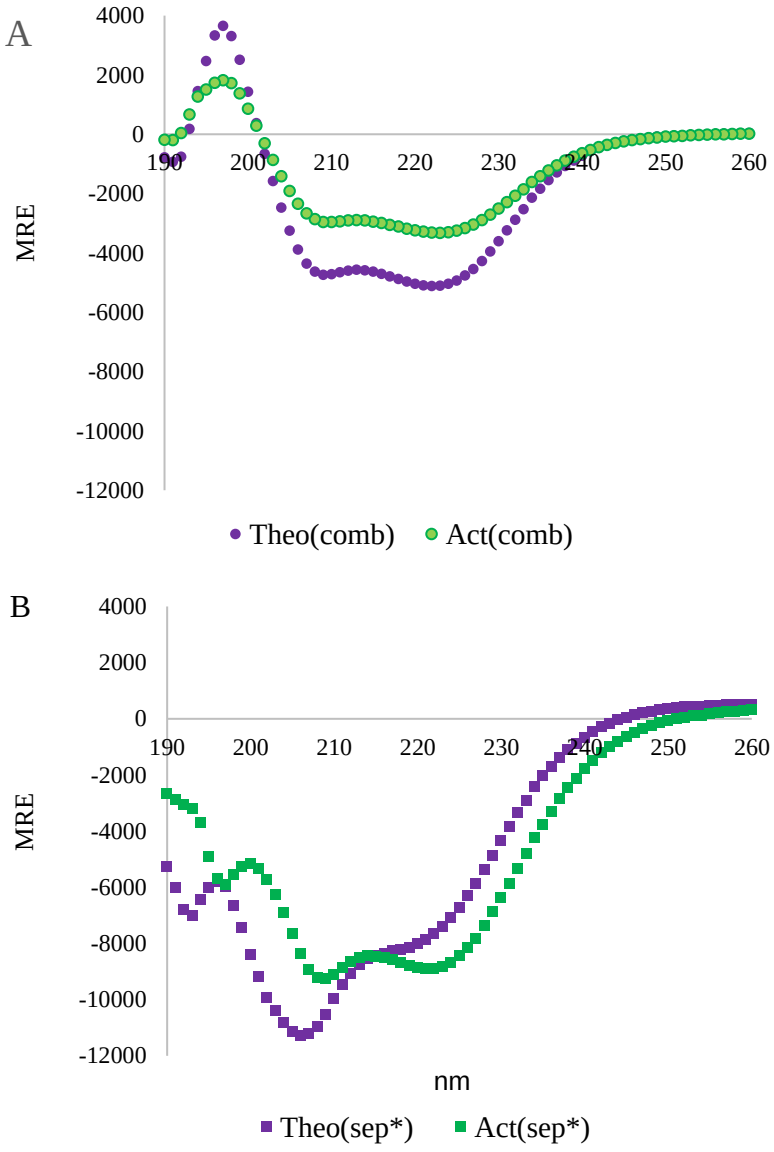


Figure 3-20: Nonbinding theoretical profiles compared to the corresponding actual C600+WYR CD profiles for the final iteration of "Combined" (A) and "Separate" (B) methods .

TABLES

Table 3-1: Helicity of myosin LMM found by Circular Dichroism (various sources).

Helicity	>90%	~80%	56%	88%	91.50%	81%
Method	all helical magnitudes	222 magnitude	222 magnitude	208 magnitude	not specified	K2D
Source	Extraction: back muscle of rabbit	Recombinant	Extraction: frozen surimi	Recombinant	Recombinant	Recombinant
LMM range	not specified	not specified	not specified	6xHIS-1231–1938	6xHIS-1231–1938	6xHIS-E1495-L1870
Organism	rabbit	walleye pollack	walleye pollack	human	human	drosophila melanogaster
Uniprot ID				P12883	P12883	P05661
Citation	Wu & Yang, 1976 [146]	Togashi et al., 2002 [147]	Togashi et al., 2002 [147]	Armel & Leinward, 2009 [150]	Armel & Leinward, 2010 [148]	Salvi et al., 2012 [152]

Table 3-1: Helicity of myosin LMM found by Circular Dichroism (various sources). CONTINUED (1)

Helicity	57%	<30%	<30%	<30%	<30%	<30%
Method	222 magnitude	222 magnitude	222 magnitude	222 magnitude	222 magnitude	222 magnitude
Source	Recombinant (tag removed)	Recombinant (tag removed)	Recombinant (tag removed)	Recombinant (tag removed)	Recombinant (tag removed)	Recombinant (tag removed)
LMM range	7H1 (L1301-E1350)	7H2 (L1329-D1378)	7H3 (L1477-L1526)	7H4 (L1533-E1582)	7H5 (A1744-L1793)	7H6 (L1357-V1407)
Organism	human	human	human	human	human	human
Uniprot ID	P12883	P12883	P12883	P12883	P12883	P12883
Citation	Wolny et al., 2013 [149]	Wolny et al., 2013 [149]	Wolny et al., 2013 [149]	Wolny et al., 2013 [149]	Wolny et al., 2013 [149]	Wolny et al., 2013 [149]

Table 3-1: Helicity of myosin LMM found by Circular Dichroism (various sources). CONTINUED (2)

Helicity	79%	<30%	91%	<30%	99%	79%
Method	222 magnitude	222 magnitude	222 magnitude	222 magnitude	222 magnitude	222 magnitude
Source	Recombinant (tag removed)	Recombinant (tag removed)	Recombinant (tag removed)	Recombinant (tag removed)	Recombinant	Recombinant (tag removed)
LMM range	11H1 (L1301-D1378)	11H2 (L1329-V1407)	15H1 (L1301-V1407)	15H2 (L1477-E1582)	GST-1280–1936	11H1 (L1301-D1378)
Organism	human	human	human	human	human	human
Uniprot ID	P12883	P12883	P12883	P12883	P12883	P12883
Citation	Wolny et al., 2013 [149]	Wolny et al., 2013 [149]	Wolny et al., 2013 [149]	Wolny et al., 2013 [149]	Parker et al., 2018 [151]	Wolny et al., 2013 [149]

Table 3-2: Percent α helical composition for 10 µM and 2 µM C600 spectral profiles according to various methods of calculation.

Method	[C600] 10 µM	[C600] 2 µM
208 magnitude	28.7%	28.5%
222 magnitude	19.1%	18.1%
230-240 slope	15.3%	14.2%
K2D	22%	19%
CONTIN (SMP180)	8.2%	N/A
CDSSTR (SMP180)	17.9%	8%
BeStSel (190-260)	15.6%	11.3%
222/208 ratio*	1.17	1.08

The SPM180 dataset is used for CONTIN & CDSSTR.
* a ratio of 1-1.1 is indicative of alpha helical content and >1.1 is indicative of coiled-coil content.

Table 3-3: Increasing WYR to C600 ratio increases cosedimentation.

	[μM] pelleting*			[μM] ratio sedimented
WYR:C600	WYR(μM)	C600(μM)		
0.5:1	0.21	0.34		0.62
1:1	0.96	0.62		1.55
2.5:1	2.67	1.38		1.93
5:1	4.12	1.67		2.47
7.5:1	5.89	1.75		3.36
10:1	6.61	1.72		3.84

*Amount [μM**]** found in the pellet after baseline pelleting is subtracted. The [μM] ratio is shown on the right.

Table 3-4: Calculated helicity for Act(sep), Act(comb) and Theo(sep), Theo(comb) using various methods.

Method	Act(sep*)	Theo(sep*)	Act(comb)	Theo(comb)
208 magnitude	45.6%	51.6%	23.7%	29.8%
222 magnitude	24.5%	21.5%	16.2%	15.1%
230-240 slope	24.3%	19.3%	10.1%	15.4%
K2D	30%	21%	10%	23%
Contin	14.1%	19.7%	9.9%	17%
CDSSTR	22%	24%	5%	16.6%
BeStSel (190-250)	16.9%	21.4%	7.4%	16.5%

Table 3-5: BeStSel structural predictions for the separate (sep*) and combined (comb) methods for theoretical (Theo) and actual (Act) CD profiles for the range 190-250 nm.

	Act(sep*)	Theo(sep*)	Act(comb)	Theo(comb)
Helix	16.9%	21.4%	7.4%	16.5%
Antiparallel	17.9%	17%	31%	21.3%
Parallel	0%	2.1%	1.2%	0.4%
Turn	17.6%	13.3%	15.2%	15%
Others	47.6%	46.1%	45.2%	46.8%

BIBLIOGRAPHY

[1] Sotavalta O. Recordings of High Wing-Stroke and Thoracic Vibration Frequency in Some Midges. Biological Bulletin. 1953;104:439-44.

[2] Iwamoto H. Structure, function and evolution of insect flight muscle. Biophysics. 2011;7:21-8.

[3] Kreipke RE, Kwon YV, Shcherbata HR, Ruohola-Baker H. Drosophila melanogaster as a Model of Muscle Degeneration Disorders. Current topics in developmental biology. 2017;121:83-109.

[4] Kronert WA, Bell KM, Viswanathan MC, Melkani GC, Trujillo AS, Huang A, et al. Prolonged cross-bridge binding triggers muscle dysfunction in a Drosophila model of myosin-based hypertrophic cardiomyopathy. eLife. 2018;7.

[5] Mohr SE. First in Fly: Drosophila Research and Biological Discovery: Harvard University Press; 2018.

[6] Sink H. Muscle Development in Drosophila. New York, NY: Springer; 2006.

[7] Vigoreaux J. Nature's Versatile Engine: Insect Flight Muscle Inside and Out. New York, New York: Springer Science & Business Media, Inc; 2007.

[8] Ayer G, Vigoreaux JO. Flightin is a myosin rod binding protein. Cell biochemistry and biophysics. 2003;38:41-54.

[9] Vigoreaux JO, Hernandez C, Moore J, Ayer G, Maughan D. A genetic deficiency that spans the flightin gene of Drosophila melanogaster affects the ultrastructure and function of the flight muscles. The Journal of experimental biology. 1998;201:2033-44.

[10] Vigoreaux JO, Saide JD, Valgeirsdottir K, Pardue ML. Flightin, a novel myofibrillar protein of Drosophila stretch-activated muscles. The Journal of cell biology. 1993;121:587-98.

[11] Reedy MC, Bullard B, Vigoreaux JO. Flightin is essential for thick filament assembly and sarcomere stability in Drosophila flight muscles. The Journal of cell biology. 2000;151:1483-500.

[12] Soto-Adames FN, Alvarez-Ortiz P, Vigoreaux JO. An evolutionary analysis of flightin reveals a conserved motif unique and widespread in Pancrustacea. Journal of molecular evolution. 2014;78:24-37.

[13] Engel DGMS. Evolution of the Insects. 1 ed: Cambridge University Press; 2005.

[14] Chakravorty S, Tanner BCW, Foelber VL, Vu H, Rosenthal M, Ruiz T, et al. Flightin maintains myofilament lattice organization required for optimal flight power and courtship song quality in Drosophila. Proceedings Biological sciences. 2017;284.

[15] Henkin JA, Maughan DW, Vigoreaux JO. Mutations that affect flightin expression in Drosophila alter the viscoelastic properties of flight muscle fibers. American journal of physiology Cell physiology. 2004;286:C65-72.

[16] Tanner BC, Miller MS, Miller BM, Lekkas P, Irving TC, Maughan DW, et al. COOH-terminal truncation of flightin decreases myofilament lattice organization, cross-bridge binding, and power output in Drosophila indirect flight muscle. American journal of physiology Cell physiology. 2011;301:C383-91.

[17] Barton B, Ayer G, Heymann N, Maughan DW, Lehmann FO, Vigoreaux JO. Flight muscle properties and aerodynamic performance of Drosophila expressing a flightin transgene. The Journal of experimental biology. 2005;208:549-60.

[18] Gasek NS, Nyland LR, Vigoreaux JO. The Contributions of the Amino and Carboxy Terminal Domains of Flightin to the Biomechanical Properties of Drosophila Flight Muscle Thick Filaments. Biology. 2016;5.

[19] Contompasis JL, Nyland LR, Maughan DW, Vigoreaux JO. Flightin is necessary for length determination, structural integrity, and large bending stiffness of insect flight muscle thick filaments. Journal of molecular biology. 2010;395:340-8.

[20] Kronert WA, O'Donnell PT, Fieck A, Lawn A, Vigoreaux JO, Sparrow JC, et al. Defects in the Drosophila myosin rod permit sarcomere assembly but cause flight muscle degeneration. Journal of molecular biology. 1995;249:111-25.

[21] Hu Z, Taylor DW, Reedy MK, Edwards RJ, Taylor KA. Structure of myosin filaments from relaxed Lethocerus flight muscle by cryo-EM at 6 A resolution. Science advances. 2016;2:e1600058.

[22] Zondlo NJ. Aromatic-proline interactions: electronically tunable CH/π interactions. Accounts of chemical research. 2013;46:1039-49.

[23] Dougherty DA. Cation-π Interactions Involving Aromatic Amino Acids. The Journal of Nutrition. 2007;137:1504S-8S.

[24] Micsonai A, Wien F, Kernya L, Lee Y-H, Goto Y, Réfrégiers M, et al. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. Proceedings of the National Academy of Sciences. 2015;112:E3095-E103.

[25] Mukherjee D, Gai F. Exciton circular dichroism couplet arising from nitrile-derivatized aromatic residues as a structural probe of proteins. Anal Biochem. 2016;507:74-8.

[26] Khan MA, Neale C, Michaux C, Pomès R, Privé GG, Woody RW, et al. Gauging a hydrocarbon ruler by an intrinsic exciton probe. Biochemistry. 2007;46:4565-79.

[27] Woody RW. [4] Circular dichroism. Methods in enzymology: Academic Press; 1995. p. 34-71.

[28] Antosiewicz JM, Shugar D. UV-Vis spectroscopy of tyrosine side-groups in studies of protein structure. Part 1: basic principles and properties of tyrosine chromophore. Biophysical reviews. 2016;8:151-61.

[29] Noronha M, Lima JC, Lamosa P, Santos H, Maycock C, Ventura R, et al. Intramolecular Fluorescence Quenching of Tyrosine by the Peptide α-Carbonyl Group Revisited. The Journal of Physical Chemistry A. 2004;108:2155-66.

[30] Fornander LH, Feng B, Beke-Somfai T, Nordén B. UV Transition Moments of Tyrosine. The Journal of Physical Chemistry B. 2014;118:9247-57.

[31] Greenfield NJ. Using circular dichroism spectra to estimate protein secondary structure. Nature protocols. 2006;1:2876-90.

[32] Sreerama N, Woody RW. Structural composition of betaI- and betaII-proteins. Protein Science : A Publication of the Protein Society. 2003;12:384-8.

[33] Bush CA, Sarkar SK, Kopple KD. Circular dichroism of β turns in peptides and proteins. Biochemistry. 1978;17:4951-4.

[34] J. Cooper JW, A. Mills. Chapter 2: Super-Secondary Structure. Principles of Protein Structure, Comparative Protein Modelling, and Visualisation. SWISS-MODEL: Biozentrum.

[35] Sibanda BL, Blundell TL, Thornton JM. Conformation of beta-hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering. Journal of molecular biology. 1989;206:759-77.

[36] Blanco F, Ramirez-Alvarado M, Serrano L. Formation and stability of beta-hairpin structures in polypeptides. Current opinion in structural biology. 1998;8:107-11.

[37] de Alba E, Rico M, Jiménez MA. Cross-strand side-chain interactions versus turn conformation in beta-hairpins. Protein Science : A Publication of the Protein Society. 1997;6:2548-60.

[38] Merkel JS, Sturtevant JM, Regan L. Sidechain interactions in parallel beta sheets: the energetics of cross-strand pairings. Structure. 1999;7:1333-43.

[39] Sunde M, Blake CCF. From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. Quarterly Reviews of Biophysics. 1998;31:1-39.

[40] Terzi E, Hoelzemann G, Seelig J. Reversible Random Coil-β-Sheet Transition of the Alzheimer β-Amyloid Fragment (25-35). Biochemistry. 1994;33:1345-50.

[41] Juszczyk P, Kolodziejczyk AS, Grzonka Z. Circular dichroism and aggregation studies of amyloid beta (11-8) fragment and its variants. Acta biochimica Polonica. 2005;52:425-31.

[42] de Brevern AG. Extension of the classical classification of β-turns. Scientific reports. 2016;6:33191.

[43] Jackups R, Jr., Liang J. Interstrand pairing patterns in beta-barrel membrane proteins: the positive-outside rule, aromatic rescue, and strand registration prediction. Journal of molecular biology. 2005;354:979-93.

[44] Eckhardt B, Grosse W, Essen LO, Geyer A. Structural characterization of a beta-turn mimic within a protein-protein interface. Proceedings of the National Academy of Sciences of the United States of America. 2010;107:18336-41.

[45] Ball JB, Hughes RA, Alewood PF, Andrews PR. β-turn topography. Tetrahedron. 1993;49:3467-78.

[46] Hutchinson EG, Thornton JM. A revised set of potentials for beta-turn formation in proteins. Protein Science : A Publication of the Protein Society. 1994;3:2207-16.

[47] Petersen B, Lundegaard C, Petersen TN. NetTurnP--neural network prediction of beta-turns by use of evolutionary information and predicted protein sequence features. PloS one. 2010;5:e15079.

[48] Shapovalov M, Vucetic S, Dunbrack RL, Jr. A new clustering and nomenclature for beta turns derived from high-resolution protein structures. PLoS computational biology. 2019;15:e1006844.

[49] Guruprasad K, Rajkumar S. Beta-and gamma-turns in proteins revisited: a new set of amino acid turn-type dependent positional preferences and potentials. Journal of biosciences. 2000;25:143-56.

[50] Milner-White EJ. Situations of gamma-turns in proteins. Their relation to alpha-helices, beta-sheets and ligand binding sites. Journal of molecular biology. 1990;216:386-97.

[51] Mazzier D, Grassi L, Moretto A, Aleman C, Formaggio F, Toniolo C, et al. En route towards the peptide gamma-helix: X-ray diffraction analyses and conformational energy calculations of Adm-rich short peptides. Journal of peptide science : an official publication of the European Peptide Society. 2017;23:346-62.

[52] Crisma M, De Zotti M, Moretto A, Peggion C, Drouillat B, Wright K, et al. Single and multiple peptide γ-turns: literature survey and recent progress. New Journal of Chemistry. 2015;39:3208-16.

[53] Swanson CJ, Sivaramakrishnan S. Harnessing the unique structural properties of isolated α-helices. The Journal of biological chemistry. 2014;289:25460-7.

[54] Du QS, Chen D, Xie NZ, Huang RB, Chou KC. Insight into a molecular interaction force supporting peptide backbones and its implication to protein loops and folding. Journal of biomolecular structure & dynamics. 2015;33:1957-72.

[55] Fetrow JS. Omega loops: nonregular secondary structures significant in protein function and stability. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 1995;9:708-17.

[56] Johnson TA, Holyoak T. The Ω-loop lid domain of phosphoenolpyruvate carboxykinase is essential for catalytic function. Biochemistry. 2012;51:9547-59.

[57] Schneider TR, Gerhardt E, Lee M, Liang P-H, Anderson KS, Schlichting I. Loop Closure and Intersubunit Communication in Tryptophan Synthase. Biochemistry. 1998;37:5394-406.

[58] Griffiths-Jones SR, Sharman GJ, Maynard AJ, Searle MS. Modulation of intrinsic phi,psi propensities of amino acids by neighbouring residues in the coil regions of protein structures: NMR analysis and dissection of a beta-hairpin peptide. Journal of molecular biology. 1998;284:1597-609.

[59] Rohl CA, Fiori W, Baldwin RL. Alanine Is Helix-Stabilizing in Both Template-Nucleated and Standard Peptide Helices. Proceedings of the National Academy of Sciences of the United States of America. 1999;96:3682-7.

[60] Richardson JS. The Anatomy and Taxonomy of Protein Structure. In: Anfinsen CB, Edsall JT, Richards FM, editors. Advances in protein chemistry: Academic Press; 1981. p. 167-339.

[61] Ashok Kumar T. CFSSP: Chou and Fasman Secondary Structure Prediction server. 2013.

[62] Youkharibache P, Veretnik S, Li Q, Stanek KA, Mura C, Bourne PE. The Small β-barrel Domain: A Survey-based Structural Analysis. bioRxiv. 2018:140376.

[63] Kajava AV, Steven AC. β-Rolls, β-Helices, and Other β-Solenoid Proteins. Advances in protein chemistry: Academic Press; 2006. p. 55-96.

[64] Bhattacharjee N, Biswas P. Position-specific propensities of amino acids in the β-strand. BMC Struct Biol. 2010;10:29.

[65] Smith CK, Withka JM, Regan L. A Thermodynamic Scale for the β-Sheet Forming Tendencies of the Amino Acids. Biochemistry. 1994;33:5510-7.

[66] Eswar N, Ramakrishnan C. Secondary structures without backbone: an analysis of backbone mimicry by polar side chains in protein structures. Protein Engineering, Design and Selection. 1999;12:447-55.

[67] Du D, Zhu Y, Huang C-Y, Gai F. Understanding the key factors that control the rate of β-hairpin folding. Proceedings of the National Academy of Sciences of the United States of America. 2004;101:15915-20.

[68] Soranno A, Cabassi F, Orselli ME, Cellmer T, Gori A, Longhi R, et al. Dynamics of Structural Elements of GB1 β-Hairpin Revealed by Tryptophan–Cysteine Contact Formation Experiments. The Journal of Physical Chemistry B. 2018;122:11468-77.

[69] Nishio M, Umezawa Y, Fantini J, Weiss MS, Chakrabarti P. CH–π hydrogen bonds in biological macromolecules. Physical Chemistry Chemical Physics. 2014;16:12648-83.

[70] Flocco MM, Mowbray SL. Planar Stacking Interactions of Arginine and Aromatic Side-Chains in Proteins. Journal of molecular biology. 1994;235:709-17.

[71] Anjana R, Vaishnavi MK, Sherlin D, Kumar SP, Naveen K, Kanth PS, et al. Aromatic-aromatic interactions in structures of proteins and protein-DNA complexes: a study based on orientation and distance. Bioinformation. 2012;8:1220-4.

[72] Efimov AV. Structure of coiled β-β-hairpins and β-β-corners. FEBS Letters. 1991;284:288-92.

[73] Biancalana M, Makabe K, Koide A, Koide S. Molecular mechanism of thioflavin-T binding to the surface of beta-rich peptide self-assemblies. Journal of molecular biology. 2009;385:1052-63.

[74] Biancalana M, Makabe K, Koide A, Koide S. Aromatic cross-strand ladders control the structure and stability of beta-rich peptide self-assembly mimics. Journal of molecular biology. 2008;383:205-13.

[75] Scheerer D, Chi H, McElheny D, Samer A, Keiderling TA, Hauser K. Role of Aromatic Cross-Links in Structure and Dynamics of Model Three-Stranded beta-Sheet Peptides. The journal of physical chemistry A. 2018;122:543-53.

[76] Lee J, Ju M, Cho OH, Kim Y, Nam KT. Tyrosine-Rich Peptides as a Platform for Assembly and Material Synthesis. Adv Sci (Weinh). 2018;6:1801255-.

[77] Franklin MW, Slusky JSG. Tight Turns of Outer Membrane Proteins: An Analysis of Sequence, Structure, and Hydrogen Bonding. Journal of molecular biology. 2018;430:3251-65.

[78] Ramírez-Alvarado M, Blanco FJ, Niemann H, Serrano L. Role of β-turn residues in β-hairpin formation and stability in designed peptides. Journal of molecular biology. 1997;273:898-912.

[79] Hamill SJ, Cota E, Chothia C, Clarke J. Conservation of folding and stability within a protein family: the tyrosine corner as an evolutionary cul-de-sac. Journal of molecular biology. 2000;295:641-9.

[80] Hemmingsen JM, Gernert KM, Richardson JS, Richardson DC. The tyrosine corner: A feature of most greek key β-barrel proteins. Protein Science. 1994;3:1927-37.

[81] Némethy G, Printz MP. The γ Turn, a Possible Folded Conformation of the Polypeptide Chain. Comparison with the β Turn. Macromolecules. 1972;5:755-8.

[82] Chou K-C. Prediction and classification of α-turn types. Biopolymers. 1997;42:837-53.

[83] Chou K-C, Liu W-M, Maggiora GM, Zhang C-T. Prediction and classification of domain structural classes. Proteins: Structure, Function, and Bioinformatics. 1998;31:97-103.

[84] Dasgupta B, Chakrabarti P. pi-Turns: types, systematics and the context of their occurrence in protein structures. BMC Struct Biol. 2008;8:39.

[85] Dasgupta B, Pal L, Basu G, Chakrabarti P. Expanded turn conformations: Characterization and sequence-structure correspondence in α-turns with implications in helix folding. Proteins: Structure, Function, and Bioinformatics. 2004;55:305-15.

[86] Craveur P, Joseph AP, Rebehmed J, de Brevern AG. β-Bulges: Extensive structural analyses of β-sheets irregularities. Protein Science. 2013;22:1366-78.

[87] Gallivan JP, Dougherty DA. Cation-π interactions in structural biology. Proceedings of the National Academy of Sciences. 1999;96:9459-64.

[88] Chakravarty S, Ung AR, Moore B, Shore J, Alshamrani M. A Comprehensive Analysis of Anion-Quadrupole Interactions in Protein Structures. Biochemistry. 2018;57:1852-67.

[89] Lin K, Simossis VA, Taylor WR, Heringa J. A simple and fast secondary structure prediction method using hidden neural networks. Bioinformatics. 2005;21:152-9.

[90] Rost B, Sander C. Improved prediction of protein secondary structure by use of sequence profiles and neural networks. Proceedings of the National Academy of Sciences of the United States of America. 1993;90:7558-62.

[91] Buchan DWA, Jones DT. The PSIPRED Protein Analysis Workbench: 20 years on. Nucleic Acids Research. 2019;47:W402-W7.

[92] Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nature protocols. 2015;10:845.

[93] Cuff JA, Barton GJ. Application of multiple sequence alignment profiles to improve protein secondary structure prediction. Proteins. 2000;40:502-11.

[94] Cuff JA, Clamp ME, Siddiqui AS, Finlay M, Barton GJ. JPred: a consensus secondary structure prediction server. Bioinformatics. 1998;14:892-3.

[95] Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nature protocols. 2010;5:725-38.

[96] Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nature methods. 2015;12:7-8.

[97] Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008;9:40.

[98] Combet C, Blanchet C, Geourjon C, Deleage G. NPS@: network protein sequence analysis. Trends in biochemical sciences. 2000;25:147-50.

[99] Yachdav G, Kloppmann E, Kajan L, Hecht M, Goldberg T, Hamp T, et al. PredictProtein--an open resource for online prediction of protein structural and functional features. Nucleic Acids Res. 2014;42:W337-43.

[100] Montgomerie S, Sundararaj S, Gallin WJ, Wishart DS. Improving the accuracy of protein secondary structure prediction using structural alignment. BMC Bioinformatics. 2006;7:301.

[101] Li M, Liu J, Ran X, Fang M, Shi J, Qin H, et al. Resurrecting abandoned proteins with pure water: CD and NMR studies of protein fragments solubilized in salt-free water. Biophysical journal. 2006;91:4201-9.

[102] Maurer M, Oostenbrink C. Water in protein hydration and ligand recognition. J Mol Recognit. 2019;32:e2810.

[103] Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, et al. Atomic-level characterization of the structural dynamics of proteins. Science. 2010;330:341-6.

[104] Parui S, Jana B. Factors Promoting the Formation of Clathrate-Like Ordering of Water in Biomolecular Structure at Ambient Temperature and Pressure. The Journal of Physical Chemistry B. 2019;123:811-24.

[105] Basu K, Wasserman SS, Jeronimo PS, Graham LA, Davies PL. Intermediate activity of midge antifreeze protein is due to a tyrosine-rich ice-binding site and atypical ice plane affinity. The FEBS Journal. 2016;283:1504-15.

[106] Remko M, Soralova S. Effect of water coordination on competition between pi and non-pi cation binding sites in aromatic amino acids: L-phenylalanine, L-tyrosine, and L-tryptophan Li+, Na +, and K+ complexes. Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry. 2012;17:621-30.

[107] Dougherty DA. Cation-pi interactions involving aromatic amino acids. J Nutr. 2007;137:1504S-8S; discussion 16S-17S.

[108] Anand BG, Prajapati KP, Shekhawat DS, Kar K. Tyrosine-Generated Nanostructures Initiate Amyloid Cross-Seeding in Proteins Leading to a Lethal Aggregation Trap. Biochemistry. 2018;57:5202-9.

[109] Koide S, Sidhu SS. The importance of being tyrosine: lessons in molecular recognition from minimalist synthetic binding proteins. ACS chemical biology. 2009;4:325-34.

[110] Garnham CP, Campbell RL, Walker VK, Davies PL. Novel dimeric β-helical model of an ice nucleation protein with bridged active sites. BMC Structural Biology. 2011;11:36.

[111] Mecozzi S, West AP, Dougherty DA. Cation-pi interactions in aromatics of biological and medicinal interest: electrostatic potential surfaces as a useful qualitative guide. Proceedings of the National Academy of Sciences. 1996;93:10566-71.

[112] Watkins AM, Arora PS. Anatomy of β-strands at protein-protein interfaces. ACS chemical biology. 2014;9:1747-54.

[113] Guharoy M, Chakrabarti P. Secondary structure based analysis and classification of biological interfaces: identification of binding motifs in protein–protein interactions. Bioinformatics. 2007;23:1909-18.

[114] Hu C, Koehl P. Helix-sheet packing in proteins. Proteins: Structure, Function, and Bioinformatics. 2010;78:1736-47.

[115] Schneider JP, Pochan DJ, Ozbas B, Rajagopal K, Pakstis L, Kretsinger J. Responsive Hydrogels from the Intramolecular Folding and Self-Assembly of a Designed Peptide. Journal of the American Chemical Society. 2002;124:15030-7.

[116] Ozbas B, Rajagopal K, Schneider JP, Pochan DJ. Semiflexible Chain Networks Formed via Self-Assembly of β-Hairpin Molecules. Physical Review Letters. 2004;93:268106.

[117] Laibe J, Caffrey A, Broutin M, Guiglion S, Pierscionek B, Nebel J-C. Coil conversion to β-strand induced by dimerization. Proteins: Structure, Function, and Bioinformatics. 2018;86:1221-30.

[118] Hartmann MD, Mendler CT, Bassler J, Karamichali I, Ridderbusch O, Lupas AN, et al. α/β coiled coils. eLife. 2016;5.

[119] Kato M, McKnight SL. Cross-beta Polymerization of Low Complexity Sequence Domains. Cold Spring Harbor perspectives in biology. 2017;9.

[120] Ding Y, Li Y, Qin M, Cao Y, Wang W. Photo-Cross-Linking Approach to Engineering Small Tyrosine-Containing Peptide Hydrogels with Enhanced Mechanical Stability. Langmuir. 2013;29:13299-306.

[121] Mukherjee S, Kapp EA, Lothian A, Roberts AM, Vasil'ev YV, Boughton BA, et al. Characterization and Identification of Dityrosine Cross-Linked Peptides Using Tandem Mass Spectrometry. Analytical Chemistry. 2017;89:6136-45.

[122] Xue C, Lin TY, Chang D, Guo Z. Thioflavin T as an amyloid dye: fibril quantification, optimal concentration and effect on aggregation. Royal Society open science. 2017;4:160696.

[123] Zhou HX, Pang X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. Chem Rev. 2018;118:1691-741.

[124] Qi HW, Nakka P, Chen C, Radhakrishnan ML. The effect of macromolecular crowding on the electrostatic component of barnase-barstar binding: a computational, implicit solvent-based study. PloS one. 2014;9:e98618.

[125] Bibi F, Villain M, Guillaume C, Sorli B, Gontard N. A Review: Origins of the Dielectric Properties of Proteins and Potential Development as Bio-Sensors. Sensors (Basel, Switzerland). 2016;16.

[126] Della Valle E, Marracino P, Pakhomova O, Liberti M, Apollonio F. Nanosecond pulsed electric signals can affect electrostatic environment of proteins below the threshold of conformational effects: The case study of SOD1 with a molecular simulation study. PloS one. 2019;14:e0221685.

[127] Yamniuk AP, Ditto N, Patel M, Dai J, Sejwal P, Stetsko P, et al. Application of a kosmotrope-based solubility assay to multiple protein therapeutic classes indicates broad use as a high-throughput screen for protein therapeutic aggregation propensity. Journal of Pharmaceutical Sciences. 2013;102:2424-39.

[128] Quigley A, Williams DR. The second virial coefficient as a predictor of protein aggregation propensity: A self-interaction chromatography study. European journal of pharmaceutics and biopharmaceutics : official journal of Arbeitsgemeinschaft fur Pharmazeutische Verfahrenstechnik eV. 2015;96:282-90.

[129] Baslé E, Joubert N, Pucheault M. Protein Chemical Modification on Endogenous Amino Acids. Chemistry & Biology. 2010;17:213-27.

[130] JBergès PT, and C Houée-Levin. Oxidation of protein tyrosine or methionine residues: From the amino acid to the peptide. Journal of Physics: Conference Series. 2011;261.

[131] Zhang H, Zielonka J, Sikora A, Joseph J, Xu Y, Kalyanaraman B. The effect of neighboring methionine residue on tyrosine nitration and oxidation in peptides treated with MPO, H2O2, and NO2(-) or peroxynitrite and bicarbonate: role of intramolecular electron transfer mechanism? Archives of biochemistry and biophysics. 2009;484:134-45.

[132] Souza JM, Daikhin E, Yudkoff M, Raman CS, Ischiropoulos H. Factors determining the selectivity of protein tyrosine nitration. Archives of biochemistry and biophysics. 1999;371:169-78.

[133] Yang H, Zubarev RA. Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. Electrophoresis. 2010;31:1764-72.

[134] Windsor IW, Gold B, Raines RT. An n-->pi* Interaction in the Bound Substrate of Aspartic Proteases Replicates the Oxyanion Hole. ACS catalysis. 2019;9:1464-71.

[135] Ghisaidoobe AB, Chung SJ. Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Forster resonance energy transfer techniques. International Journal of Molecular Sciences. 2014;15:22518-38.

[136] Bartlett AI, Radford SE. An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. Nature structural & molecular biology. 2009;16:582-8.

[137] Chavez JD, Liu NL, Bruce JE. Quantification of protein-protein interactions with chemical cross-linking and mass spectrometry. Journal of proteome research. 2011;10:1528-37.

[138] Sinz A. Divide and conquer: cleavable cross-linkers to study protein conformation and protein-protein interactions. Analytical and bioanalytical chemistry. 2017;409:33-44.

[139] Korn ED. Coevolution of head, neck, and tail domains of myosin heavy chains. Proceedings of the National Academy of Sciences of the United States of America. 2000;97:12559-64.

[140] Bennett PM. The structure of spindle-shaped paracrystals of light meromyosin. Journal of molecular biology. 1981;146:201-21.

[141] Offer G. Skip residues correlate with bends in the myosin tail. Journal of molecular biology. 1990;216:213-8.

[142] Guzenko D, Strelkov SV. CCFold: rapid and accurate prediction of coiled-coil structures and application to modelling intermediate filaments. Bioinformatics. 2017;34:215-22.

[143] Korkmaz EN, Taylor KC, Andreas MP, Ajay G, Heinze NT, Cui Q, et al. A composite approach towards a complete model of the myosin rod. Proteins. 2016;84:172-89.

[144] Trybus KM, Freyzon Y, Faust LZ, Sweeney HL. Spare the rod, spoil the regulation: necessity for a myosin rod. Proceedings of the National Academy of Sciences of the United States of America. 1997;94:48-52.

[145] Greenfield NJ, Hitchcock-DeGregori SE. Conformational intermediates in the folding of a coiled-coil model peptide of the N-terminus of tropomyosin and alpha alpha-tropomyosin. Protein Science : A Publication of the Protein Society. 1993;2:1263-73.

[146] Wu C-SC, Yang JT. Reexamination of the conformation of muscle proteins by optical activity. Biochemistry. 1976;15:3007-14.

[147] Togashi M, Kakinuma M, Nakaya M, Ooi T, Watabe S. Differential scanning calorimetry and circular dichroism spectrometry of walleye pollack myosin and light meromyosin. Journal of agricultural and food chemistry. 2002;50:4803-11.

[148] Armel TZ, Leinwand LA. A mutation in the beta-myosin rod associated with hypertrophic cardiomyopathy has an unexpected molecular phenotype. Biochemical and biophysical research communications. 2010;391:352-6.

[149] Wolny M, Colegrave M, Colman L, White E, Knight PJ, Peckham M. Cardiomyopathy mutations in the tail of beta-cardiac myosin modify the coiled-coil structure and affect integration into thick filaments in muscle sarcomeres in adult cardiomyocytes. The Journal of biological chemistry. 2013;288:31952-62.

[150] Armel TZ, Leinwand LA. Mutations in the beta-myosin rod cause myosin storage myopathy via multiple mechanisms. Proceedings of the National Academy of Sciences of the United States of America. 2009;106:6291-6.

[151] Parker F, Batchelor M, Wolny M, Hughes R, Knight PJ, Peckham M. A1603P and K1617del, Mutations in β-Cardiac Myosin Heavy Chain that Cause Laing Early-Onset Distal Myopathy, Affect Secondary Structure and Filament Formation In Vitro and In Vivo. Journal of molecular biology. 2018;430:1459-78.

[152] Salvi SS, Kumar RP, Ramachandra NB, Sparrow JC, Nongthomba U. Mutations in Drosophila myosin rod cause defects in myofibril assembly. Journal of molecular biology. 2012;419:22-40.

[153] Kuhn ER, Naik AR, Lewis BE, Kokotovich KM, Li M, Stemmler TL, et al. Nanothermometry Reveals Calcium-Induced Remodeling of Myosin. Nano Letters. 2018;18:7021-9.

[154] Gans PJ, Lyu PC, Manning MC, Woody RW, Kallenbach NR. The helix-coil transition in heterogeneous peptides with specific side-chain interactions: theory and comparison with CD spectral data. Biopolymers. 1991;31:1605-14.

[155] Rohl CA, Baldwin RL. Deciphering rules of helix stability in peptides. Methods in enzymology. 1998;295:1-26.

[156] Wallimann P, Kennedy RJ, Kemp DS. Large Circular Dichroism Ellipticities for N-Templated Helical Polypeptides Are Inconsistent with Currently Accepted Helicity Algorithms. Angewandte Chemie (International ed in English). 1999;38:1290-2.

[157] Walters J, Milam SL, Clark AC. Practical approaches to protein folding and assembly: spectroscopic strategies in thermodynamics and kinetics. Methods in enzymology. 2009;455:1-39.

[158] Wei Y, Thyparambil AA, Latour RA. Protein helical structure determination using CD spectroscopy for solutions with strong background absorbance from 190 to 230nm. Biochimica et biophysica acta. 2014;1844:2331-7.

[159] Lupas AN, Gruber M. The structure of alpha-helical coiled coils. Advances in protein chemistry. 2005;70:37-78.

[160] Sohn RL, Vikstrom KL, Strauss M, Cohen C, Szent-Gyorgyi AG, Leinwand LA. A 29 residue region of the sarcomeric myosin rod is necessary for filament formation. Journal of molecular biology. 1997;266:317-30.

[161] Ghassemizadeh R, Moore B, Momose T, Walter M. Stability and IR Spectroscopy of Zwitterionic Form of β-Alanine in Water Clusters. The Journal of Physical Chemistry B. 2019;123:4392-9.

[162] Yancey PH, Somero GN. Counteraction of urea destabilization of protein structure by methylamine osmoregulatory compounds of elasmobranch fishes. The Biochemical journal. 1979;183:317-23.

[163] Kaushik JK, Bhat R. Why is trehalose an exceptional protein stabilizer? An analysis of the thermal stability of proteins in the presence of the compatible osmolyte trehalose. The Journal of biological chemistry. 2003;278:26458-65.

[164] Zhmurov A, Kononova O, Litvinov RI, Dima RI, Barsegov V, Weisel JW. Mechanical transition from α-helical coiled coils to β-sheets in fibrin(ogen). Journal of the American Chemical Society. 2012;134:20396-402.

[165] Qin Z, Kreplak L, Buehler MJ. Hierarchical structure controls nanomechanical properties of vimentin intermediate filaments. PloS one. 2009;4:e7294.

[166] Apostolovic B, Danial M, Klok HA. Coiled coils: attractive protein folding motifs for the fabrication of self-assembled, responsive and bioactive materials. Chemical Society reviews. 2010;39:3541-75.

[167] Minin KA, Zhmurov A, Marx KA, Purohit PK, Barsegov V. Dynamic Transition from alpha-Helices to beta-Sheets in Polypeptide Coiled-Coil Motifs. Journal of the American Chemical Society. 2017;139:16168-77.

[168] Briggs M. Densitometry using ImageJ. 2013.

[169] Starna Cells I. Suggestions for Cleaning Cells.

[170] Micsonai A, Wien F, Kernya L, Lee YH, Goto Y, Refregiers M, et al. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. Proc Natl Acad Sci U S A. 2015;112:E3095-103.

[171] Micsonai A, Wien F, Bulyaki E, Kun J, Moussong E, Lee YH, et al. BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. Nucleic Acids Res. 2018;46:W315-W22.

[172] Whitmore L, Wallace BA. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers. 2008;89:392-400.

[173] Whitmore L, Wallace BA. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res. 2004;32:W668-73.

CHAPTER 4 JOURNAL ARTICLE (Extended Version)

Contiguity & structural impacts of a non-myosin protein within the thick filament myosin layers

Lynda Menard*, Neil Wood and Jim O. Vigoreaux

Department of Biology

The University of Vermont

* Corresponding author. Address:

Department of Biology,

University of Vermont,

120 Marsh Life Science Bldg,

109 Carrigan Drive, Burlington, VT 05405,

USA.

Fax: +1 802 656 2914.

E-mail address: lmmenard@uvm.edu (Lynda Menard).

LETTER ABSTRACT

Myosin dimers arranged in layers and interspersed with non-myosin densities were first described by cryo-EM 3D reconstruction of the thick filament in Lethocerus at ~5.5 Å resolution by Hu et al (2016). One of the non-myosin densities, denoted the 'red density', is hypothesized to be flightin, an LMM-binding protein essential to the structure and function of Drosophila indirect flight muscle (IFM). Here, we build upon the 3D reconstruction results specific to the red density and its engagement with the myosin coiled-coil rods that form the backbone of the thick filament. Each independent red density winds its way through the myosin dimers such that it links four dimers in a layer and one dimer in a neighboring layer. This area in which three distinct interfaces within the myosin rod are contacted at once and the red density extends to the thick filament core is designated the "multiface". Present within the multiface is a contact area inclusive of E1563 and R1568. Mutations in the corresponding Drosophila residues (E1554K and R1559H) are known to interfere with flightin accumulation and phosphorylation in Drosophila. We further examine the LMM area in direct apposition to the red density and identified potential binding residues spanning up to ten helical turns. We find that the red density is associated within an expanse of the myosin coiled-coil that is unwound by the third skip residue and the coiled-coil is re-oriented while in contact with the red density. These findings suggest a mechanism by which flightin induces ordered assembly of myosin dimers through its contacts with multiple myosin dimers and reinforcement on the level of a single myosin dimer by stabilization of the myosin coiled-coil.

LETTER MAIN TEXT

Molecular-level muscle structure amongst both vertebrates and invertebrates employs many of the same building blocks and strategies for structural and mechanical attunement per organism. Striated muscle is known for its organized subcellular arrangement of protein filaments into regularly repeating structures known as sarcomeres. Attunement of largely conserved thick filaments, prominently composed of myosin dimers, is accommodated by changes within the myosin sequence and assembly and protein addendums. The packing of myosin within the thick filament backbone is known to vary between vertebrates and invertebrates, and among invertebrates [1, 2]. The significance of these differences and their implications in thick filament function and mechanobiology are not fully understood but are likely to underpin muscle-type functional differences and locomotory modalities. Such understanding can be realized in model systems for which information from molecular structures can be interpreted in light of mechanical, physiological, and organismal functional properties.

Cryo-EM studies by Hu et al (2016) [3] have revealed the thick filaments of Lethocerus (Hemiptera) to be arranged in layers of associating myosin dimers through engagement of their light meromyosin (LMM) regions - long C-terminal coiled-coiled rods. These layers of myosin dimers include additional proteins winding their way through the dimers of each layer and between layers. The pitch of the coiled-coil was found to be variable (60-126Å) with areas of unwinding. Four unconnected non-myosin densities were found and

assigned different colors (red, yellow, blue, and green) (figure 5 and movie S3 in [3]) as their identities have not been confirmed. The ratio of each of the densities to myosin was found to be 1:1 with the combined volume of a full set amounting to ~20 kDa of a polypeptide, resulting in an expectation that each density represents a more ordered segment of a protein whose less ordered regions are not visible. The red and yellow densities both connect to adjacent rods and contact the paramyosin core. The red density was found to pass through from the surface to the center of the thick filament, and the yellow density appeared to 'stitch' together multiple layers. The former was hypothesized as flightin [4] and the latter as myofilin [5].

The position of the red density is the primary reason for flightin's attribution. Flightin is known to be a component of the Drosophila thick filament [6] and to bind a 600 amino acid segment of the LMM *in vitro* [7]. The mutation E1554K in Drosophila myosin prevents flightin accumulation *in vivo* [8] and binding *in vitro* [7]. The red density is found on the outside of the filament, consistent with flightin antibody labelling in Lethocerus [9], and in close proximity with the rod at the corresponding E1554 residue in Lethocerus (E1563). The flightin to myosin stoichiometry was calculated to be approximately 1:1 to 1:2 [7], in alignment with the 1:1.4 to 1:2 ratio suggested by Hu et al [3].

We set out to determine the specific amino acid ranges and the pattern of red density contacts with the LMM as it winds its way through the myosin dimers of the thick filament. The 3D model of the thick filament, inclusive of non-myosin densities, is provided in movie S3: a video fly-through that follows the complete path of a myosin dimer in the M-ward direction as viewed from the globular head to the end of the coiled-coil rod [3]. The key to the video provided in fig. S8 [3] was used along with a manual matching procedure, using ApowerEdit [10], to properly orient the LMM sequence encompassed in each frame. The video was sorted into three 435Å segments and one 292Å segment and the primary region of interest (G1528-A1628) was determined within the boundaries of its 435Å segment. The frames for this section had a representative dimer isolated from the rest of the image using GIMP [11]. The resultant images were evaluated in ImageJ [12] (see letter methods).

The winding path that the red density takes along the length of the myosin rod brings it in close contact with five different sections of the LMM. Among these are four sections within the same layer (S972-L996; E1254-A1284; E1547-R1582, and S1851-Q1873) and one section in a neighboring layer (S1759-T1786) (**Fig. 4-1**). A single red density contacts each of these regions once along five different myosin dimers, alluding to a possible role in tying or clasping them together. The last three contact areas (E1547-R1582; S1759-T1786; S1851-Q1873) are termed the “multiface” because a single red density is simultaneously contacting three distinct myosin dimers (**Fig. 4-2**). The

multiface is of further interest as this is where the red density links layers and reaches the thick filament core to contact paramyosin.

The linking of four dimers within a layer to each other and to one dimer in a neighboring layer may represent a mechanism whereupon flightin directs and secures ordered assembly of myosin into the thick filament. Skinned IFM fibers from the mutant *Df(3L)fln*^{1}, which result in ~20% less flightin, exhibit a loss of thick filaments from the myofibril periphery [13], as if these myosin molecules were not firmly secured in the outwardly developing myofibril [14]. In the flightin null mutant *fln*^{0}, there are decreased thick filaments across the myofibril diameter and sarcomeres and thick filaments are ~25% longer and more variable in pupa with breakdown occurring shortly after eclosion [6]. Such change in the arrangement of thick filaments within the myofibril coupled with instability throughout the *fln*^{0} muscle system speaks to flightin's role as conducive and secure to higher order structure myosin assemblies.

Given the prevalence of coiled-coils in proteins with mechanical roles [15, 16] and the importance of the myosin coiled-coil in assembly of the thick filament [17], we asked whether the red density was associated with any changes in the winding (pitch) of coiled-coil structure. The structure of the coiled-coil is guided by a heptad repeat, residues in the pattern of HPPHCPC in which hydrophobic (H), polar (P) and charged (C) residues dictate the left-handed supercoiling of the right handed helices [18]; however, deviations

from this pattern are common and contribute to imperfections in the coiled-coil pitch. Rotational angle change relative to the major axis of the dimer was measured M-ward along the length of the LMM, from G1528 to A1628, for several matching layers and averaged. To do this, an isolated dimer was fit to an ellipse in ImageJ and the angle change between frames was recorded and graphed against the associated amino acid range (see Letter Supplementary Methods).

We find negative slope and stasis of rotation between S1574 and R1582, representing a change in apparent direction of the dimer turn to be slightly right-handed (**Fig. 4-3**). Generally, there is an M-ward left-handed turning of the dimer. This indicates change in the winding of the coiled coil: a local relaxation of pitch. This area is especially interesting as the red density making contact with the dimer also contacts paramyosin over a short span from E1572 to Q1575. Once the red density disappears, around R1582, the slope recovers its typical rotation. This precedes the third skip residue (E1590). When rotation is mapped along the entire myosin rod, similar shifts are evident only in locations associated with skip 1 (T1196) and, possibly, skip 4 (G1815), centering around T1196 and Q1802 (not shown).

The association of the red density with a change in pitch proximal to E1590 could indicate the involvement of the red density in stabilization of an otherwise unstable area for coiled-coil formation. Few crystal structures of portions of the LMM exist though it

has been shown that the skip residues are responsible for disruption in the coiled-coil that extends beyond a heptad both N- and C-terminally [17, 19] in the absence of non-myosin thick filament proteins. If the red density stabilizes the area N-terminal to the 3rd skip residue, it can reinforce the coiled-coil by preventing the disturbance from radiating further. Such stabilization of the coiled-coil may be taking place in other areas of the LMM in connection to the red density, or other non-myosin densities, as the LMM is known to harbor additional deviations (stammers, stutters) from the heptad ideal for coiled-coil formation [18]. Securing areas of heptad disruption would increase the overall coiled-coil integrity of the LMM and resilience in the context of contractile forces.

The LMM region in the vicinity of E1563 (E1554 in Drosophila), spanning from residues E1547 to R1582, was further examined to identify potential residues in direct juxtaposition to the red density. The specific interface of the LMM to the red density was defined based on the angular relationship of the red density to each monomer in the LMM dimer. Notation is not taken beyond a 5 pixel distance (~3.2 Å); this relationship is estimated using fig. 4A in [3].

The residues identified are shown in **Figure 4-4**. Other residues in this area may be important to the orientation of the LMM relative to the red density as this region is part of the multiface and is stabilized by other dimers. The contact region borders E1563 (E1554 in Drosophila) and contains R1568 (R1559 in Drosophila). The inclusion of these two

residues within the interface, in contact with the bulk of the red density, provides an explanation for the depletion of flightin in *Mhc*^{13} and *Mhc*^{6}, two Drosophila strains that carry point mutations E1154K and R1559H, respectively [8, 20]. These mutations significantly diminish power output while differing in their effects on fibrillar passive and dynamic viscoelastic properties [12]. The revelation of this interface allows further exploration on the nature of the flightin-myosin interaction and consequences on thick filament structure, fibers mechanics, and muscle function.

The estimated mass of the red density is less than the mass of flightin [4]. Mutant Drosophila flightin lacking the N-terminal 62 amino acid region [21, 22] or the C-terminal 44 amino acid region [23] are incorporated into the fiber, indicating that the region between amino acids 63 and 138, encompassing the conserved WYR domain [24], harbors an essential myosin binding site. We hypothesize that the red density is mostly or exclusively WYR, a hypothesis further supported by the predicted unstructured nature of the larger N-terminal region [25]. The myosin sequence encompassing the red density interface is well conserved between vertebrates and invertebrates. Comparison of the Drosophila MHC rod sequence to its human cardiac counterpart reveals 56% identity, 74% positives while the interface area of flightin binding between I1534-E1586 shares 68% identity and 85% positives. Vertebrate proteins that influence thick filament stability and alignment, including M-protein, myomesin and Myosin Binding Protein-C (MyBP-C), have been shown to bind to this region [26-28]. A shared myosin coiled-coil binding

region raises the prospect of a conserved binding mechanism for these divergent proteins. Furthermore, studies with exogenously expressed cardiac MyBP-C in wild-type and fln^0 flies suggest that flightin and cMyBP-C have partially convergent functions, both in contributing to the mechanical properties of the thick filament (flexural rigidity) and in assembly (thick filament and sarcomere length) [29].

The findings from this study provide insight into the mechanism by which flightin provides stability and rigidity to the thick filament as arising from securing of the LMM coiled-coil and enforcement of dimer-dimer contacts within the thick filament. We propose that flightin behaves as a 'cinch' to stabilize the LMM structure and partition the coiled-coil thereby influencing the thick filament's capacity for mechanical relay and stretch activation. The highly conserved region of binding on the LMM may further allude to a shared strategy between invertebrate and vertebrate striated muscle for tuning thick filament properties. The information presented can inform molecular dynamics and structural studies to shed light on a possible conserved mode of molecular interaction between the myosin coiled-coil and its binding partners.

LETTER SUPPLEMENTARY METHODS

Overview

This supplement contains further processing information of Supplementary movie 3 [3], "Cross-sectional thick filament fly-through", an M-line directed fly-through of the

myosin dimers participating in the 12 curved ribbon-like layers of the thick filament. The video follows a single dimer as it winds its way through the thick filament to its core. Along the way, it makes contact with multiple other myosin dimers and non-myosin densities. This single dimer is termed ‘cyan’ for its coloration in the video and is referred to as it was used for orientation to the LMM sequence. It is a member of the ‘blue’ layer. Initial measurements were fit to the Drosophila myosin sequence, as was used in Hu et al., 2016 [3] and were later accommodated to fit the Lethocerus myosin sequence. The heptad designations for the LMM residues were limited to G1528-A1628, based on Taylor et al 2015 [17].

We address three main processes, organized as follows:

1. Identifying start/end points for the selected dimer
    a. Characterizing the structure of the video
    b. Selecting the region of interest
2. Mapping the rotation of the selected dimer
    a. Isolating the dimer fit to an ellipse
    b. Attaining coiled-coil rotation over set aa range
3. Attaining angle of interface of dimer to the red density

Identifying start/end points for the selected dimer

Characterizing the structure of the video

Values used for calculations over the duration of the movie are shown in **Table 4-1**. The movie S3 is described as going through 1700 Å which encompasses an entire myosin dimer. The entire video is just over 28 seconds long with 15 fps (~420 frames).

While the video itself is ~420 frames, 1.3 to 27.9 sec (26.6s) represents the extent of the cyan dimer rod region. The full length of the rod, including tether is reported as 1598 Å. At 15 fps, 26.6 sec calculates to 399 frames (26.6 sec*15 fps). With each frame being 4 Å apart, the expected number of frames is 399.5 (1598Å /4Å) which supports the previous calculation.

There are crowns every 145 Å with a 435 Å axial repeat for crowns along the length of the thick filament for which dimers in any individual layer laterally associate. When following the cyan dimer, each blue crown is 435 Å from the previous and the following blue crown. After the cyan crown, the next blue crown is 435 Å from that start point and so forth. As there is a 1.485 Å rise per residue, the number of amino acids in 435 Å is approximately 293 (435Å /1.485Å = 292.93). This corresponds to the number of amino acids along a dimer between crowns originating from the same layer.

Figure S8 [3] is a ‘key’ to the S3 movie and designates 1.3, 8.6, 15.7, and 22.9 seconds as crowns spaced 435 Å apart along the cyan myosin dimer tracked in Movie S3. The time between these sections are not consistent but averages to 7.2 seconds indicating that 108 (7.2*15) frames as representative of a 435 Å region. As there are 7.2 seconds between crowns and 15 fps, there are 108 frames per 435 Å. This is in close alignment with each frame being 4 Å along (435Å /4Å = 108.75).

Table 4-2 shows the span time and duration key provided in Figure S8 [3]. Using 1.3 sec as the starting point, we used the video editing software, ApowersoftEdit, to more precisely assign the span and duration through closest matching of the myosin head densities at crown positions, every 435 Å using 1.3 sec as the starting point (**Table 4-2**, right-most columns). These were considered our checkpoints. The 7.25 second durations fits better with the calculated number of frames per 435 Å (108.75 frames).

Note, in ApowersoftEdit, seconds are separated into 1/20s with a notation that completes at “0.2”. For example, 4:30 seconds, which equates to 4.5 seconds in decimal notation, is expressed as 4.1 in ApowersoftEdit.

The detectable rod length and tether is 1077 residues (R843-F1919). This is 1598 Å with ~293 aa per 435 Å segment. 1598 Å is 3.67 times 435 Å so the extra portion corresponds to (0.67*435) 292 Å, or ~196 aa.

Calculations are as follows:

Distance after last same-layer crown until dimer disappears: 0.67* 435Å = 291.45Å

Frames covering last ~292Å: 291.45Å/4Å = 72.8625 frames

Expected duration from last crown to last frame: 72.86 frames/15fps = 4.86s

Expected end point after last participating crown following Figure S8 key [3]: 22.9s+4.86s =27.76s

Expected end point for adjusted crown positions by timepoint matching: 23.05s+4.86s = 27.91s

The expected end point with adjusted crown positions corresponds exactly with the last frame in which the cyan dimer is present. This supports the accuracy of the modified time points/frames associated with the residue ranges.

Selecting the region of interest

The combined rod and tether region is 1598 Å in contour length and identified, after fitting with the Drosophila sequence, to be 1077 aa in length from R843-F1919 with a rise-per-residue of 1.485 Å. At the time of the Hu et al. 2016 publication [3], the Drosophila sequence was used to fit the model as the Lethocerus sequence had not yet been published. The Drosophila region of interest selected for our analysis (R1520-D1620) falls within the third 435 Å region, with the rod being segmented into three 435 Å regions plus an additional 292 Å (**Table 4-2**).

Within the selected region exists two specific amino acids that are highly conserved and suspected to be involved in interaction between flightin and the LMM: Drosophila myosin residues E1554 and E1580 [8]. This area also contains the third skip residue (E1581). In Lethocerus, E1563 and E1589 correspond to Drosophila myosin residues E1554 and E1580. The skip residue is E1590 for Lethocerus. The total range from Drosophila R843-F1919 corresponds to Lethocerus N852-F1928. The R1520-D1620 Drosophila region corresponds to Lethocerus R1529-D1629.

Frames were separated using VLC. The region of interest aimed for, Lethocerus amino acids N1530-I1630, was manually separated out and designated frame #1-37. This ended up specifically being positions 1528.6 to 1628.5 and annotated as G1528-A1628.

Each frame was considered to be representative of a 4 Å region, which has 2.7 amino acid increments. The positioning of G1528-A1628 was calculated based on distance from the closest crown checkpoint.

Mapping the rotation of the selected dimer

The cyan dimer isolation is confounded by the black line overlay designed to show the path of movement through its layer over time. The diagonally opposite dimer to the cyan dimer was selected as representative. This horizonatally and vertically mirrored dimer moves in the exact same pattern as the cyan dimer.

After isolation of the 37 frames, the representative dimer was separated from the rest of the image using GIMP and saved as .png and .jpg files. The following progression was done to convert the selected dimer to neonpink to distinguish the selected dimer from the other dimers in the ribbon in each frame:

GIMP Protocol

1. (Colors -> Adjust Brightness and Contrast) Added 10% contrast
2. Selected dimer of interest
3. (Colors -> Adjust Brightness and Contrast) Added 10% contrast
4. (Colors -> Adjust Color Balance) Set to Grey (Midtones, maxed)

5. (Colors -> Adjust Color Balance) Highlights, maxed to Red
6. (Colors -> Colorize) Maxed Hue and Saturation

In the above progression, the contrast and brightness of the entire image was enhanced. A small area, inclusive only of the dimer of interest and background, was selected and its brightness and contrast were further increased. Color balance was then adjusted to starken the color contrast of the dimer from the rest of the image. Frames were saved as PNG files and evaluated in ImageJ.

Isolating the dimer to fit an ellipse

ImageJ Protocol

In order to attain an axis for each frame over which position and rotational change can be tracked, we fitted each dimer to an ellipse that encompassed both myosin monomers.

1. (Image>Adjust->Color Threshold) Threshold Color set to point just before selection circle is not evident around dimer of interest: Hue 0,36; Sat 5,255; Brightness 130,255
2. (Analyze->Set Measurements) ‘Area’, ‘centroid’ and ‘fit ellipse’ was checked.

3. (Analyze->Analyze Particles) Set min pixel size to 35 to reduce off-target measures
4. Set measurements to fit ellipse and centroid. This automatically selected the two globular units as a single ellipse. Measurements of the major axis angle and the length of the major (long) and minor (short) axes of the ellipse were provided by the program under these settings.
    - The “Draw Ellipse” macro was used to visualize the measured ellipse. (Edit->Selection->Fit Ellipse)

Attaining coiled-coil rotation over amino acid range

We used the measurements of the ellipse to track movement of the dimer over a series of frames. ImageJ fitting provided X,Y coordinates of center of ellipse, pixel length of major and minor axis line, and angle of major axis.

The angle of the major axis represents the major axis line angle to the horizontal of the image out of 180 degrees. To check, the angle function was used while going over the line used to form Ellipse using the Macro, "Draw Ellipse" function. These values were collected from each frame and the angle change between frames was recorded and graphed against the associated amino acid range with each frame progressing 2.7 amino acids. For simplicity, the starting angle in frame 1 was set to zero degrees. This was then graphed in sequence to visualize the rotational shift. We did this for the 100 amino acid

range for multiple dimers within the blue layer and dimers within the yellow and pink layers (n=6) to confirm conservation among various dimers within the same region (not shown).

When mapping the entire dimer over the full range for the visible dimer there will be some discrepancy due to the increased span over which drift will occur. Drift is resultant from the actual frame span being very slightly off from 4 Å and 2.7 aa per frame being a rounded value. Small imperfections are magnified over long spans that are not evident within a 100 aa range. The discrepancy is less than 1 frame. Frames 60-411 have been evaluated both with and without a forced fit to the check points. There is no substantial loss of information and the regions over which there are rotational shifts are consistent. We chose a forced fit to the check points as this equalizes the drift between segments and decreases the overall drift.

Linear regression analysis was done using Graphpad PRISM. R-square values were very similar regardless of whether residue numbers are fixed to checkpoint-values (0.9977 vs 0.9976) suggesting that the methods are comparable for making observations over larger amino acid regions. Checkpoints are useful for aligning heptad positions (next section) as they provide better accuracy by ablating drift.

Attaining angle of interface of dimer to the red density

ImageJ Protocol

1. Calculated the X,Y coordinates for the start and end points of the major axis (**Table 4-3**), manually inserted those values into a new Draw Line macro, and ran the macro on the image.
2. Used the reflex angle feature to trace over the drawn line representing the major axis such that the angle parallel to the line is 180° to 360°.
3. Measured the angle range over which the dimer is facing the density of interest.
    - Subtracted 360° from any value that passes the 360° point parallel to the major axis to get the 'actual' value.
4. Compared the angle ranges to the heptad positions
    - The coiled coil is left-handed and composed of two right-handed helices. A predominant left handed rotation is seen when viewing the frames in video S3 [3]. The arrangement of the two helices in accordance to heptad positions is show in **Figure 4-12**.

Within the 37 frames of the selected region, the red density appears from frame 7 to 19. The dimers turn counter-clockwise in this M-ward view, therefore the alpha helices are turning clockwise [30].

To determine the amino acids that are in close contact with the red density, the dimer was identified as having alpha helices “1” and “2” since at different points only one of the two might be close to the red density or the two helices may be close to the red density at different angles. Helix “1” was designated to be the left helix (**Fig. 4-12**). For the duration of the 37 frames, the helices do not switch relative position (ex. helix 1 is always left-most).

Although the rods rotate over time, the major line of the ellipse formed between the two densities was considered to be zero to 180° as the horizontal major axis in **Figure 4-12**. If, for instance, the red density is closest to Helix 1, between the angles of 230°-300° then positions ‘c’ and ‘g’ are noted as being in closest proximity and it depended on the range encompassed by the frame whether ‘c’ or ‘g’ is determined to be the closest contact.

Notation is not taken past a 5 pixel distance (~3.2 Å); this relationship is estimated using Fig. 4A in [3].

Mapping of the myosin dimer to the red density across frames 7-19 allowed positions to be identified for amino acids involved in the interfacing between the two. This combined with dimer rotation measurements allowed correlation of pitch change along the region of interest.

IMAGE PROCESSING - EXTENDED RESULTS & DISCUSSION

Areas of Interest

G1528-A1628 was selected to be analyzed for several reasons. Firstly, the red density, proposed to be flightin, appears to interact within this region predominantly. This is a region where specific mutations have been shown to result in decreased accumulation of flightin such as R1559H (R1568H in Lethocerus) and Mhc^{13}, corresponding to a charge change, E1554K (E1563K in Lethocerus), in the heptad 'e' position [7, 8]. *Ifm(2)RU1*, E1570K (E1579K in Lethocerus) also exists within this region with mutants exhibiting a less extreme phenotype than Mhc^{13} mutants but still alongside some reduction in flightin accumulation [31]. There are also a series of myopathy associated mutations within this range in conserved residues in mammalian muscle at sites: Q1541, A1549, E1564, E1573, N1589, L1597 and R1606 [32-38]. These correspond to Q1549, A1557, E1572, D1581, N1597, M1605, K1614 in Lethocerus.

This region is also home to the third skip residue and surrounding point mutations in mammalian heart that result in cardiomyopathies (E1573K, R1588P, L1591P) [17]. In Beta human cardiac myosin, the region surrounding the skip residue has been shown to exhibit 'local unwinding' and, unlike with skip 1 and 2, deletion of skip 3 causes myosin aggregation in the cytoplasm and this has been connected with specific structural changes. The unwound, but stable, structure surrounding skip 3 was determined to be of functional importance [17] and exists at E1590 in Lethocerus.

A further advantage to looking at this region over other regions is that this is the only area of the myosin light meromyosin (LMM) that has been mapped so as to form a complete model through expression of segments expressed *in vitro* [19]. While *in vitro* information lacks the additional binding proteins existent in structure of *in vivo* models, this study provides a good starting point for mapping the heptad along this range of residues and provides data regarding exposed or internal residues in the absence of other binding proteins. Corroborating studies on flightin, paramyosin, the modelling studies of Korkmaz et al (2016) [19] and Hu et al (2016) [3], provide grounds for hypotheses regarding structural/mechanical modulation that occurs within this region.

Considering the Interface

Each layer is considered to be composed of myosin dimers participating in crowns, extensions of the myosin heads, +/- 3 crowns away and when observed by transverse section, 3 or 4 myosin dimers in a crown can be viewed in association with each other (**Fig. 4-5**). The dimer whose observed position is closest to its myosin head (more N-terminal on the LMM) is on the outer edge of the thick filament and is referred to as the 1st dimer here. The '1st dimer' constitutes a region within N852-A1145, the first 435 Å segment of the LMM. The '2nd dimer' constitutes a region within A1145-Q1437, the second 435 Å segment. The '3rd dimer' encompasses the area of interest and includes Q1437-S1730, the third 435 Å segment. The '4th dimer' constitutes a region within

S1730-F1928, the C-terminal remainder of the LMM. Over the span in which the red density is present, all four dimers are identifiable. These ultimately correspond to 35aa regions within each of the (1^{st}, 2^{nd}, 3^{rd}, 4^{th}) segments for an individual red density within the same layer: K960-E995 for the 1^{st} dimer, T1253-L1288 for the 2^{nd} dimer, K1545-I1580 for the 3^{rd} dimer (**Fig. 4-4**), and Q1838-Q1873 for the 4^{th} dimer.

The individual red densities are predominantly associated with a single layer but do make a contact outside the originating layer. For simplicity, we focus on a red density associated with the 'blue' layer, the color designation as it relates to Hu et al (2016) movie S3. As there are interactions outside of the blue layer, the dimers are designated "B" for originating in the blue layer, "Y" for belonging to the yellow layer or "P" for belonging to the pink layer. This leads to notation such as "2ndB" which would mean 2^{nd} dimer from the outside of the filament, belonging to the blue layer (**Fig. 4-5**). There are no direct interactions of the pink layer within our area of interest in context of the blue layer and red density. However, the pink layer would interact with a red density as blue is observed to in the Z-ward direction from the red density's appearance in the blue layer. This is because the interfaces involved with the red density are staggered as the layers are staggered. In such a case, 4thB takes on the role of the observed 4thY. In either M-ward or Z-ward view, the yellow layer is to the blue layer as the blue layer is to the pink layer and the pink layer is to the yellow layer for red density interactions between them.

Observations of the red density are split into four sections: (i) pre-split in which there is a single red density furthest from the M-line, (ii) path of the smaller red density that appears after the split, (iii) path of the larger (bulk) red density, which incorporates a local smaller red density due to proximity, and (iv) the point at which the larger red density makes contact with paramyosin.

Red Density Pre-Split

Within the range K1545-I1580, the red density shows up first at E1547, in contact with only one of the helices of the 3rd dimer in the blue layer (3rdB) and one helix of 2ndB from E1254-V1264 of that dimer (**Fig. 4-6**). The position of E1547 is focused on as it is in the proper heptad position (c) for the amino acid range of the frame (See Methods).

Within the area in which the 2ndB and 3rdB are both contacting the red density, there are predominant positive charges within 2ndB and predominant negative charges within 3rdB. Hydrophobic residues of 3rdB are also within this region. At the most clear split point of the red density (frames 9 and 10), negatively charged residues are in close proximity to the red density from 3rdB and 4thY and the smaller separated red density is furthest from all surrounding dimers. It is possible that the regions of negative charge between 3rdB and 4thY promote the bifurcation of the red density.

Path of the Red Density 'hook'

Around A1551-E1553 (frame 9), the red density starts to split into two densities. Once the red density splits into two, the smaller of the two densities begins to make contact with one helix of 1stB starting around S972 (**Fig. 4-7**). This smaller density remains in close proximity with 1stB until the disappearance of both portions of the red density. The 1stB is in close proximity from S972-L996. One helix of 2ndB is in close proximity with one helix of 1stB which may be stabilizing the position of 1stB. This smaller portion of the red density also makes contact with 2ndB around R1276-A1284. Between A991-L996 the remainder of the small red density only makes contacts with 1stB, close to the center of the two helices, and does not appear to be making any other contacts. During the period in which 2ndB is making contact with the red density, it is exposed to multiple close contacts, including dimers of the yellow layer and the non-myosin blue density on its other helix, which may be involved in stabilizing this interaction.

This smaller red density 'hook' is its only connection to 1stB, the dimer involved in the closest crown. Notably, in the area in which the red density interacts with 1stB, there are no additional stabilizing contacts for either the 1stB dimer or the red density. This would support mechanical information relay being most purely translated between 1stB through that portion of the red density.

Path of the Larger Red Density

After the split, the larger red density makes contact with one helix of 3rdB (E1556-E1572) along with one of the helices of the 4thB from S1851-K1867 (**Fig. 4-8**) and 4thY (E1761-L1777). Right before paramyosin contact is made, potentially stabilizing interactions include both helices of 4thY which is also interacting on its opposite side with the non-myosin blue density. This area is unique in that both of the helices of the 3rdB and 4thB dimers are in close contact with the red density starting at L1571 for 3rdB, and A1862 for 4thB. The red density is at its largest relative volume compared to other slices from this point onward until it disappears.

When the red density first makes contact with 4thB, it initially contacts a region of predominantly positive charge which then shifts to negative and then positive again. When the red density starts to be engaged in multiple close contacts, the area of proximity of 4thB contains a large amount of hydrophobic residues that exist between the regions of positive charge, among the interim of negative charge. This may reflect a hydrophobic span of residues on the part of the red density as well.

Contact with Paramyosin

At the point of contact with paramyosin (L1573-E1579 for 3rdB; K1867-E1870 for 4thB) the red density has resumed contact with only one of the two helices of each of the three

contacting dimers (**Fig. 4-9**). The area over which there is paramyosin interaction is short, only existing over two frames suggesting a region of maximally 8 Å in length.

Within the region in which paramyosin is contacted, there appears to be an extension from one of the helices of 3rdB that exists between a helix of 2ndB and 4thY while the same helix of 3rdB and the same helix of 4thY make contact with the red density. The heptad position corresponding to the extension of the 3rdB helix is closest to what would be the ‘b’ position which is outside the range for this frame. The closest ‘b’ position is S1574. It could also be from the ‘e’ position which would be R1577. However, change in the pitch here may place a different heptad position at that location. Notably, the other myosin helix of the 3rdB dimer has an extension at approximately the same area but more clearly in the ‘e’ position that appears to make contact with the ‘g’ position of the other dimer in the pair and this positioning is appropriate for an ionic interaction as the ‘e’ position is R1577 and the ‘g’ position is E1579. The extended residue of 3rdB also appears to ‘shove’ 2ndB away from close proximity to 4thY and into closer proximity to the nonmyosin blue density and 3rdY. It is at this specific point that 3rdB is most closely connected with 4thY. It is just after S1574 that the coiled-coil is found to be unwound, with a greater pitch.

A connection of a myosin accessory protein to paramyosin may have distinct structural and mechanical consequences. Paramyosin is involved with passive stiffness and

maximum power output in Drosophila with significant loss being found in mutants with putative phosphorylation sites mutated [39, 40]. The ratio of myosin/paramyosin is associated with thick filament length with those with higher paramyosin content being longer [41-43] and capable of higher isometric tension. Isometric tension is especially important in catch muscles of molluscs, having long sarcomeres and the greatest paramyosin content, less-so with insects that require a 'warm-up' period pre-flight such as Lethocerus and even less in insects that do not require such a warm up period, as in Drosophila.

Flightin has been known to play a role in regulation of thick filament length and, in the fln^0, has resulted in longer thick filaments formed during development [6, 44]. Flightin has been found to restrict myosin incorporation/dissociation [45] which may be done by blocking a potential site of interaction or securing the connections between rods to exclude other conformations. As observations of the red density implicate flightin involved in connection to the paramyosin core in addition to multiple interaction sites between myosin rods, flightin and paramyosin may be acting in concert to regulate myosin incorporation and thick filament length.

In addition to establishing the thick filament length, these connections have a strong implication for impacting passive stiffness, compromised in flightin mutants [20]. The positioning of flightin at five points within a rod (S972-L996; E1254-A1284; E1547-

R1582; S1759-T1786; S1851-Q1873) in which three of these points (E1547-R1582; S1759-Q1785; R1849-Q1873) may be further secured or enabled by the red density's connection to paramyosin. Higher stability of the flightin-LMM contact in this area of multiple-interface ('multiface') involvement may contribute to myosin's series elastic behavior, with the red density effectively acting as a 'clasp' or 'cinch' between multiple dimers and paramyosin. Just as flightin may be acting in concert with paramyosin to elicit developmental structural parameters of the thick filament, both may be working together to dictate passive stiffness.

Rotational Observations

Rotational angle change relative to the major axis of the dimer was measured along the length of the LMM from G1528-A1628 for several myosin dimers and averaged (**Fig. 4-3**). There is an M-ward left-handed turning of the dimer. It was found that there was a negative slope and stasis of rotation between S1574-R1582 which represents a change in apparent direction of the dimer turn to be slightly right-handed. The difference looks dramatic around S1574-R1577 though this is likely to be partially due to the previously described extension of one of the helix densities. Even without the change at this point, however, there is a slight reversal and pause in dimer rotation. This indicates change in the winding of the coiled coil: a local relaxation of pitch. This area is especially interesting as the red density making contacts with the LMM also makes contact with paramyosin at the point of the LMM at E1572-Q1575. Once the red density disappears,

around R1582, the slope has recovered its typical rotation. This area of relaxed pitch immediately precedes the third skip residue (E1590).

The angle change between each frame represents a regular orientation shift that can be associated with the helices ceasing to wind around each other or becoming bent, strained, and/or extended. In Hu et al (2016) [3], studies were done on super-relaxed muscle; a change in helical pitch could represent an area that may be of pivotal importance during transition to an active state. Between V1576-E1579, an extension appears to be protruding from expected heptad position ‘b’ or ‘e’ with R1577 in the ‘e’ position. This does not appear to be solely responsible for the rotational changes seen in this region though this was not compensated for in the rotational measurements. Such a shift is evident only in two other locations, centering at Q1192 and E1800, when rotation is mapped along the entire myosin rod (**Fig. 4-10**).

E1590 is the third skip residue in Lethocerus and lies within a region where there is a relative absence of helical winding between the two helices of the dimer and is very close to the point at which the rotation recovers from the shift seen at Q1575. Taylor et al. (2015) [17] has found an increase in helical pitch surrounding skip 3 which could also be observed as a lack of rotation in this region. An increase in super helical pitch surrounding the third skip residue has been described as starting at F1565 in beta cardiac myosin [19] which corresponds to L1573 in Lethocerus. This is close to the point at

which the shift in rotation is observed, as well as red density and paramyosin contact, in Lethocerus. Although the residue is not conserved, distortion appears to be.

In Hu et al (2016) [3], the crystal structure of myosin was found from isolated super-relaxed Lethocerus myosin, while in Korkmaz et al (2016) [19], the crystal structure was determined from a series of *E. coli* expressed fragments of human beta cardiac myosin which would not be expected to possess the typical affiliated binding proteins. This conserved distortion may be important for non-myosin and paramyosin contacts in Lethocerus and could be a key position for myosin binding proteins in beta cardiac as there are several known proteins that bind proximally to this region in mammalian systems [26, 46]. The distortion in beta cardiac myosin is described as extending 11 aa C-terminal of the skip residue which equates to resolution at R1600 in Lethocerus. The rotational distortion observed in Lethocerus, however, appears to resolve itself before then, closer to the skip residue itself. This could be due to the non-myosin red entity, and paramyosin, securing the LMM at that point and permitting the following residues to more readily resume helical winding.

A primary finding of Hu et al (2016) [3] was clear evidence of an interacting head motif (IHM) that would need to be disrupted for muscle activation; this may be accomplished by stretch, with the rod involved in the translation. Torsion by formation of the IHM in normal mode analysis has found the rod distorted throughout its length [47]. Formation of

the IHM stores elastic energy throughout the rod making the IHM/rod spring loaded. Stretch could release stored elastic energy leading to an opposing torsional shift that could release the IHM to free up myosin heads for subsequent activation. The region where abnormal twist exists along with non-myosin and paramyosin interactions could represent an area where the binding of these entities cinches the coiled-coil in a formation that strains the within-dimer and dimer-dimer contacts that could then retract through elastic recoil, and more fully yield only once the other contacts are disrupted by stretch.

Overview of the Red Density Behavior & Implications for Mechanical Transduction

The red density winds between the rods of an individual layer and connects to a portion of the neighboring layer participating in the crown just M-ward, specifically to the 4th rod in the neighboring layer, and paramyosin. Each rod is connected to three other rods in the same layer via different red densities, one at each of the regions S972-L996; E1254-A1284; E1547-R1582; S1851-Q1873. Hence, all rods within a layer are connected by the red density in (i) four places to each other, (ii) to rods in the neighboring layer once between S1759-T1786 and (iii) paramyosin within the multiface. Each myosin dimer can be considered to connect to paramyosin most directly three times via three different red densities along its length, once at Q1571-Q1576, once at S1759-T1786, and once at S1851-Q1873. Notably, there would be five different red densities along the length of a given rod, all ultimately in contact with paramyosin, but the red density interfaces at S972-L996 and E1254-A1284 are more removed from those contact points.

Ordered Transduction to the IHM

The red density winding within layers may reinforce the arrangement of rods within a layer and allow each rod to sense and respond to stress or stretch in a coordinated manner and, in doing so, sequentially release or secure the associated IHMs. The IHM's head-head and backbone interactions have been found to be conserved from the emergence of the first animals [48]. Studies on active vertebrate striated muscle have shown that during a phase of contraction, myosin heads from crowns 43 nm away - therefore involving rods of the same layer - have identically 'unwrapped' myosin heads from the thick filament backbones [49]. This 'unwrapping' may be representative of IHM release states triggered by stretch or tension through the associated rods [50] and demonstrate the relevance of within-layer connections.

Beyond within-layer coordination, connection to each rod in the neighboring layer within that layer's rods' more C-terminal regions may distribute tension between layers and orient the layers to each other. In this way, each rod in a layer is connected to a set of rods laterally that participate in the crown +1 to itself thereby introducing a directionality of layer connections. In the context of Hu et al (2016) [3], this is such that the rods of the blue layer each connect to all rods in the pink layer within each blue rod's S1759-T1786 segment, rods in the pink layer each connect to all rods in the yellow layer within each pink rod's S1759-T1786 segment and rods in the yellow layer each connect to all the rods

in the blue layer within each yellow rod's S1759-T1786 segment. This directionality may reveal a sequence of crown participation in active muscle and is supported by the gradient of 'unwrapping' previously observed [49].

It is known that not all myosin heads are simultaneously active during contractile activity but if one considers that layers are being connected to each other in a +1 format, mechanical transduction from crowns with active myosin heads would move such that contractile activity of crown 0 would be most immediately relayed to crown 1 whose activity would then be relayed to crown 2. As this connection appears to exist towards the end of the rods in the connecting +1 layer (S1759-T1786), this might be necessary for strategically delaying the effect that stretch has on the IHM of that following layer. This position is straddled by two red densities at E1547-R1582 and S1851-Q1873 that are secured by +1 crown connections from there (**Fig. 4-11**). Stabilization and segmentation within the area of the LMM between E1547-R1582 and S1851-Q1873 via the connection point at S1759-T1786 may enable a mechanism by which energy storage and release is titrated. A more stabilized segment could store energy until a threshold was reached resulting in uncoupling of one of the cinch points (E1547-R1582, S1759-T1786 or S1851-Q1873), upon which that energy would be transferred. Strain could be produced in the segment between E1547-R1582 and S1851-Q1873 first by the red density at S1759-T1786 connected to the -1 crown, pre-straining the connection to the next series connected to the S1759-T1786 rod portion from that layer.

While Hu et al (2016) [3] has shown a first reveal of non-myosin entities entwined within the thick filament rods, other hints towards an organization within the invertebrate rods have supported the structural implications suggested here. Headless myosin in Drosophila IFM has been found to possess 14.5 nm reflections that are even stronger than those observed in the WT [51] for which involvement of flightin and paramyosin have been speculated. The strength of these reflections compared to WT may represent an order permitted to exist due to a lack of a myosin head induced force gradient along the rod-rod and layer-layer connections. The super-lattice found in Drosophila associated with the A-band has also been considered a possible adaptation for stretch activation [52].

Relationship to stretch activation

The high wing beat frequencies and efficiency of insect flight muscle is partially attributed to enhanced stretch activation (SA). In a calcium-activated state, a delayed increase in tension instigated by stretch is produced with the shortening of an antagonist muscle. A complimentary phenomenon of shortening deactivation (SD) exists in which there is a delayed drop in tension following shortening that serves to lower resistance during muscle lengthening [53]. These processes are responsible for increased power generation and permit the asynchronous muscles to forgo calcium cycling.

Just as IHM regulation through the rod appears to be conserved throughout animals, invertebrates have been expected to share a similar mechanism with vertebrates when it comes to stretch activation [54]. Stretch activation may be linked to the means by which the IHM is modulated. Insect flight muscles, which display enhanced stretch activation, are known to be highly ordered. It is suspected that the enhanced order of insect IFM is required for coordination of IHM states that may be necessary for such enhanced stretch activation. MHC isoforms alone only have a limited capacity to enhance stretch activation [55] and it has been proposed that non-myosin thick filament proteins provide some of the force enhancement [56-58]. In vertebrates, titin has been proposed as a parallel elastic component within muscle [57, 59-61]; paramyosin, along with the titin-like sallimus isoforms of invertebrates, and other myosin binding proteins, such as flightin, are candidates for engagement in a similar role.

Effective stretch activation/deactivation is necessarily related to the resonant frequency of the system which is impacted by the elasticity and damping within the system [62, 63] as well as temperature, as it relates to stiffness [64]. Wingbeat frequency (WBF) matches the resonant frequency of the system and increase in stiffness of the system corresponds to changes in WBF. Natural modulators of stiffness include the pleurosternal ‘control’ muscles, endogenous or exogenous temperature change within the system, variation of thick filament proteins and the myosin duty ratio. Per organism, combined factors of load, necessary power output and corresponding WBF require different viscoelastic

parameters for optimal operation. While asynchronous stretch activated IFM is a highly efficient system, it requires very particular operation settings so that mechanical transduction optimally distributes stress, permitting some separation of components for sequential activity and sustained force while preventing fracture. Means by which organisms attune the resonant system to permit stretch activation (e.g. isometric activation) may be further enhanced, or rendered obsolete, by titrating the involvement of thick filament proteins, such as flightin, in myosin and non-myosin crosslinking in the thick filament. Comparative biomechanics throughout Arthropoda, especially between species that utilize stretch activation with varying degrees of isometric involvement, may shed light on flightin's involvement within the resonant system.

Modulation within Invertebrates

Examination of flightin across species with different contractile needs is important for determining the means by which flightin itself is adapted to accommodate. It is known that both the flightin sequence and the extent of its post-translational modification vary throughout Arthropoda. Increasing accessibility to structural information, such as the cryo-EM studies by Hu et al (2016) [3], can lead study on binding partners and the relevance of expression levels/stoichiometry as additional factors of flightin adaptation.

Observations made by Hu et al (2016) [3] , expanded upon here, were in Lethocerus flight muscle but, recently, the same lab has published data on cryo-EM studies in

Drosophila flight muscle [65]. Lethocerus depends on isometric contractions to 'warm up', has greater passive tension and operates at much lower frequencies than Drosophila, while Drosophila IFM operates at higher frequencies with an augmented stretch activation response. The red density between Lethocerus and Drosophila was found to be very similar though, among other structural differences within the thick filament, other non-myosin densities became prevalent and the density associated with myofilin was lessened. The red density appeared to have conserved myosin interfaces between both Drosophila and Lethocerus; this is expected under the condition that the red density contacting LMM represents the highly conserved WYR region of flightin. *Lethocerus indicus* flightin (164 aa) shares only 40% identity with that of *Drosophila melanogaster* flightin (182 aa), though 65% identity (34/52) and 76% positives (40/52) are within the WYR region. The WYR region of Lethocerus flightin is also the same size as that in Drosophila, lending further support to the red density's identity as flightin WYR. Other, more variable, portions of flightin or additional non-myosin contacts are likely to be involved in flightin adaptation between these two organisms, though anchoring within the thick filament is attributed to WYR.

One difference between Drosophila and Lethocerus flight muscle is the preponderance of isometric tension generation in Lethocerus over Drosophila. Evaluation of viscoelastic properties in fln^0 led to the finding that the fiber generates lower passive stiffness, dynamic stiffness, elastic modulus, and a decreased frequency at which maximum work

was achieved [20]. A greater viscous modulus, compliance, resulted in more absorption of work. Interestingly, while oscillatory work output was greatly decreased, isometric tension was not greatly hindered in the fln^0. It has been proposed that this is due to viscous properties being important for stretch activation but not as involved for generation of isometric tension. Stretch activation occurs with the muscular system initially behaving as a compressed spring; the pull of the filaments then release the compression and decrease elastic energy sufficient to allow cross-bridge attachment [66]. Filaments are known to stretch slightly during isometric tension. Thermal energy, that may be attained during isometric 'warm-up' contractions, can additionally create strained pre-stroke states that allow detached myosin heads to access longitudinally displaced actin sites [67]. Isometric tension generation, though distinct from that of SA, is expected to be involved in setting the stage for SA. The action of the Lethocerus isometric 'warm up' may partially encompass a role filled by flightin in species for which isometric contractions are not prevalent.

Thick filaments in Lethocerus revealed the red density as having come in contact with paramyosin; though paramyosin was unable to be observed in Drosophila, if paramyosin binding is involved in flightin function, control of paramyosin stoichiometry may be another means by which flightin function is adapted between the two organisms. Higher paramyosin content has been associated with higher isometric tension; the relative stoichiometry of paramyosin to flightin may determine the locality of paramyosin's

effects on stiffness, or accessibility to other potential binding partners and kinases. Correlative studies between paramyosin and flightin may provide an additional lead towards dissection of their relationship with thick filament structure and viscoelastic attunement.

Order and Mechanical Transduction During Development

While mechanical relay is a requirement for active muscle, there is evidence that it is also required during development [68]. Flightin binding is integral to IFM development in the late pupal stages. This raises further questions regarding the nature of red density's binding profile. Is there a temporal order to red density interaction with the LMM? If so, how is that dictated? Are the interactions necessary during development maintained through adulthood? As expression profiles of different myosin-binding proteins and isoforms change throughout development, flightin may be in contact with more, or fewer, interfaces within the myosin LMM during its initial expression compared to adulthood.

In fln^{0}, IFM exhibits fiber hypercontraction after eclosion. Sarcomeres become disordered, accompanied by loss of the M-line and fragmentation of the Z-disc [6]. This phenotype requires actomyosin generated-tension [69]. Actomyosin kinetics are partially modulated by elastic strains. Hypercontraction could be due to force generation beyond the point of fracture from depletion of flightin as an important player in the series elastic component. Too great of a power output in a system designed to favor faster muscle

kinetics would also generate this phenotype and could be a factor if flightin is found to impact IHM formation-and-release timing.

Flightin is found as different phospho-variants throughout development before the full set is established in the adult [70] and this may be responsible for differences between binding partners or interfaces between pupal and adult stages. While there are complex contacts at the point of paramyosin interaction at the multiface, not all these contacts may be necessary for this interaction. Modulation by post translational modification of flightin or paramyosin could weaken or abolish a contact point with one of the dimers while enhancing (or opposing) one or both of the other points of contact. Paramyosin's function is drastically impacted by its own phosphorylation profile [40] though it is currently unknown how this varies from the time of its expression (mid-late pupa stages) to adulthood. Phosphorylation of either flightin or its binding partners may be another mode by which the number of segments in the thick filament series could change along with the segments' stiffness, viscoelasticity and, hence, capacity for energy storage/release.

Relevance of a disordered structural component

Another intrinsically disordered protein that binds in the LMM near E1554 in mammalian systems that does not exist in invertebrates, myomesin, encourages attribution of flightin to the red density and may hint at the system within which flightin operates. Myomesin contains an intrinsically disordered domain of 138aa (My1) that has been found to bind

myosin residues 1506-1674 [71] in mammalian striated muscle. This region alone has not been stably expressed but it has been identified that My1 is required for binding in this LMM region. Myomesin is known to connect titin with the thick filament and be involved in determination of lattice spacing. In the heart, an isoform of myomesin known as “EH-Myomesin” expresses an unstructured “EH segment” that is present early in development provides large extensibility to avoid damage from stretch [72]. WYR has now been indicated to have an uncommon structure that is disordered-like, though many other ‘disordered’ regions may be found to simply have a less characterized, or more dynamic, structure in the future. Though there is no sequence homology between the disordered LMM binding regions of myomesin and the WYR region of flightin, this area of the LMM is highly conserved and the similar ‘type’ of structure predicted for both binding domains in myomesin and flightin may be related to the means by which contact with this region is made.

Flightin is likely to be part of a system within which cinching together the myosin dimers is key. In such a system, flightin may be likened to a gear among several that serve the dual purpose of mechanical adjustment and structural shift, cable-like in nature when considered together. Flightin isn’t the only protein found within the A band that is involved in determination of myofilament structure and whether flightin is, indeed, interacting with paramyosin or other non-myosin thick filament proteins remains to be elucidated. The many LMM contacts that the red density partakes in, the extremity of

mutations that occur within the dimer within the area of the multiface, and the apparent connection between structural and mechanical modulation promotes flightin as a key member in orchestration of muscle synchrony scaling up from the myosin rod to the muscle whole. These findings further compel study of flightin's role in muscle attunement and can be utilized as a valuable tool for understanding muscle evolution among members of Pancrustacea and structure-function conservation throughout straited muscle.

EXTENDED SUPPLEMENTARY METHODS

This section is an extended version of the LETTER SUPPLEMENTARY METHODS section in which we build upon the processing methods used with more in-depth detail. For clarity, some description is repeated and new information is provided both in text and with additional supporting images.

Identifying start/end points for the selected dimer

In the Hu et al. Supplementary movie 3, “Cross-sectional thick filament fly-through”, an M-line directed series of frames are put in sequence to display a fly-through of the myosin dimers participating in the 12 curved ribbon-like layers of the thick filament. There are 3 distinct layers per quadrant and each of these layers are the same as those in the mirror-opposite quadrant to it. The layers that are in the same, mirrored, structure are colored light blue, light yellow and pink.

One particular dimer within one of the twelve layers (one of the 4 light blue layers) is separated out from the rest by being colored a darker blue (‘cyan’) than the light blue designating the other dimers in its layer.

A frame is representative of a 1 Å thick section and each frame is 4 Å apart. Values important to calculations over the duration of the movie are shown in **Table 4-1**.

The movie S3 is described as going through 1700 Å which encompasses an entire myosin dimer. The entire video is just over 28 seconds long with 15 fps (~420 frames).

While the video itself is ~420 frames, the period spanning 1.3 to 27.9 sec (the point at which the last cyan color is seen) is 26.6 seconds in duration. This region represents the center point of the cyan interacting head motif (IHM) to the end of its rod. The full length of the rod, including tether is reported as 1598 Å. At 15 fps, 26.6 sec calculates to 399 frames (26.6 sec*15 fps). With each frame being 4 Å apart, the expected number of frames is 399.5 (1598Å/4Å) which supports the previous calculation.

There are crowns every 145 Å with a 435 Å axial repeat for crowns along the length of the thick filament for which dimers in any individual layer laterally associate. When following the cyan dimer, each blue crown is 435 Å from the previous and the following blue crown. After the cyan crown, the next blue crown is 435 Å from that start point and so forth. As there is a 1.485 Å rise per residue, the number of amino acids in 435 Å is approximately 293 (435Å/1.485Å = 292.93). This corresponds to the number of amino acids along a dimer between crowns originating from the same layer.

Hu et al. figure S8 is a ‘key’ to the S3 movie and designates 1.3, 8.6, 15.7, and 22.9 seconds as crowns spaced 435 Å apart along the cyan myosin dimer tracked on in Movie

S3. The time between these sections is not consistent but averages to be 7.2 seconds suggesting that 108 (7.2*15) frames are representative of a 435 Å region. As there are 7.2 seconds between crowns and 15 fps, there are 108 frames per 435 Å. This is in close alignment with each frame being 4 Å along (435Å/4Å = 108.75).

The Hu et al. fig S8 'key' is approximate and using the video editing software, ApowersoftEdit, more exact designations are visually identified by closest matching of the IHM densities at every 435 Å using 1.3 sec as the starting point, as shown in the two right-most columns in **Table 4-2**. These are considered our checkpoints. Using the S8 key without matching, the segments are as shown in the second and third columns in **Table 4-2**. **Image 1** shows the screenshots for the best matches. The 7.25 second durations fit even better with the calculated number of frames per 435 Å (108.75 frames).

Note, ApowersoftEdit is used to make initial frame-by-frame visualizations. In ApowersoftEdit, seconds are separated into 1/20s with a notation that completes at ".2". For example, 4:30 seconds, which equates to 4.5 seconds in decimal notation, is expressed as 4.1 in ApowersoftEdit. ApowersoftEdit notation is show in the bottom right of the images in **Image 1**.

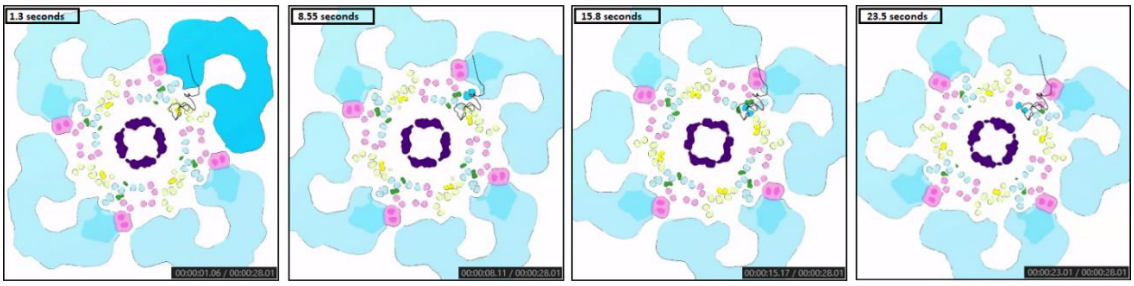


Image 1: Best matches by ApowersoftEdit. 1.3 is the start point suggested by fig.S8. The calculated increments for 435 Å from this point are best represented in 7.25 sec increments. Associated matched images are shown.

Rod length is supposed to be 1077 residues (R843-F1919). This is 1598 Å with ~293aa per 435 Å segment. 435 Å goes into 1598 Å 3.67 times so the extra portion corresponds to (0.67*435) 292 Å, or ~196 aa. Calculations are as follows:

Distance after last same-layer crown until dimer disappears: 0.67* 435Å = 291.45 Å

Frames covering last ~292 Å: 291.45Å/4Å = 72.8625 frames

Expected duration from last crown to last frame: 72.86 frames/15 fps = 4.86 seconds

Expected end point after last participating crown following S8 key: 22.9s+4.86s = 27.76s

Expected end point for adjusted crown positions by timepoint matching:

23.05s+4.86s = 27.91s

The expected end point with adjusted crown positions corresponds exactly with the last frame in which the cyan dimer is present. This supports the accuracy of the modified time points/frames associated with the residue ranges.

Frames are separated using VLC. The region of interest, Lethocerus amino acids N1530-I1630, is manually separated out and designated frame #1-37. This ends up actually being aa1528.6-aa1628.5 and is now to be described as G1528-A1628.

Each frame is considered to be representative of a 4 Å region, which has 2.7 (4/1.485) amino acid increments. The positioning of G1528-A1628 is calculated based on distance from the closest checkpoint (**Img. 1**; 15.8 sec).

Selected Region of Interest

The rod and tether is 1598 Å in length and identified, after fitting with the Drosophila sequence, to be 1077 aa in length from R843-F1919 with a rise-per-residue of 1.485 Å. At the time of the Hu et al. (2016) [3] publication, the Drosophila sequence was used to fit the model as the Lethocerus sequence had not yet been published. The Drosophila region of interest selected (R1520-D1620) corresponds to a portion of the last 435 Å region, with the rod being segmented into three 435 Å regions plus an additional 292 Å (**Table 4-2**).

Within the selected region exists two specific amino acids that are highly conserved and suspected to be involved in interaction between Flightin and the LMM: E1554 and E1580. This area also contains the third skip residue (E1581). In Lethocerus, E1563 and E1589 would correspond to Drosophila position E1554 and E1580. The skip residue is E1590 for Lethocerus.

The R1520-D1620 Drosophila region aligns with Lethocerus R1529-D1629 (**Img. 2**). To ensure homologous sequence-ranges are being aligned, we consider that the Drosophila rod and tether starts 6 amino acids C-terminal to the last invariant proline and we start at the same relative location for Lethocerus. That would change the total range from Drosophila R843-F1919 to Lethocerus N852-F1928 (**Table 4-4**).

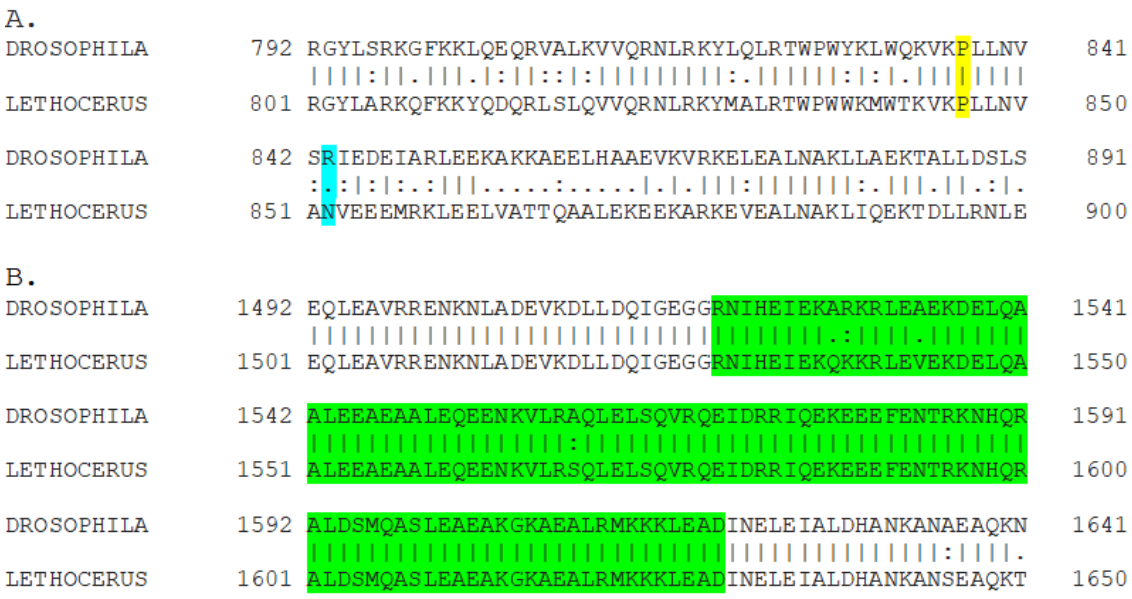


Image2: Waterbug and fruit fly myosin sequence comparison. Lethocerus (A0A222DXQ8) and Drosophila (MYSA_DROME) myosin II pairwise sequence alignments to show shifted numerical 'start' point (A) and region of interest (B) done using EMBOSS Stretcher. In A, the last invariant proline is highlighted in yellow and the corresponding 'start' residue is highlighted in blue. In B, the R1520-D1620 Drosophila region aligns with Lethocerus R1528-D1628 region. The region in which rotational observations are made is highlighted in green.

The region of interest, G1528-A1628 is within the 435 Å segment located between 15.8-23.05 seconds in the video. G1528-A1628 corresponds to 148.5 Å, or ~37 frames. The start of the region of interest is 91 aa (~135 Å) away from the start of its 435 Å segment at 18.1 sec and amounts to ~34 frames over a duration of 2.25 seconds. This places the start of this section at (15.8sec+2.25sec) where the closest frame is at 18.06 sec. The point that is 100 aa from 18.06 sec is (37/15fps) ~2.47 seconds away which leads to the time point of (18.06sec+2.47sec) 20.53 seconds representing the last frame in this 100 aa segment. The closest frame to this point is at 20.5 seconds.

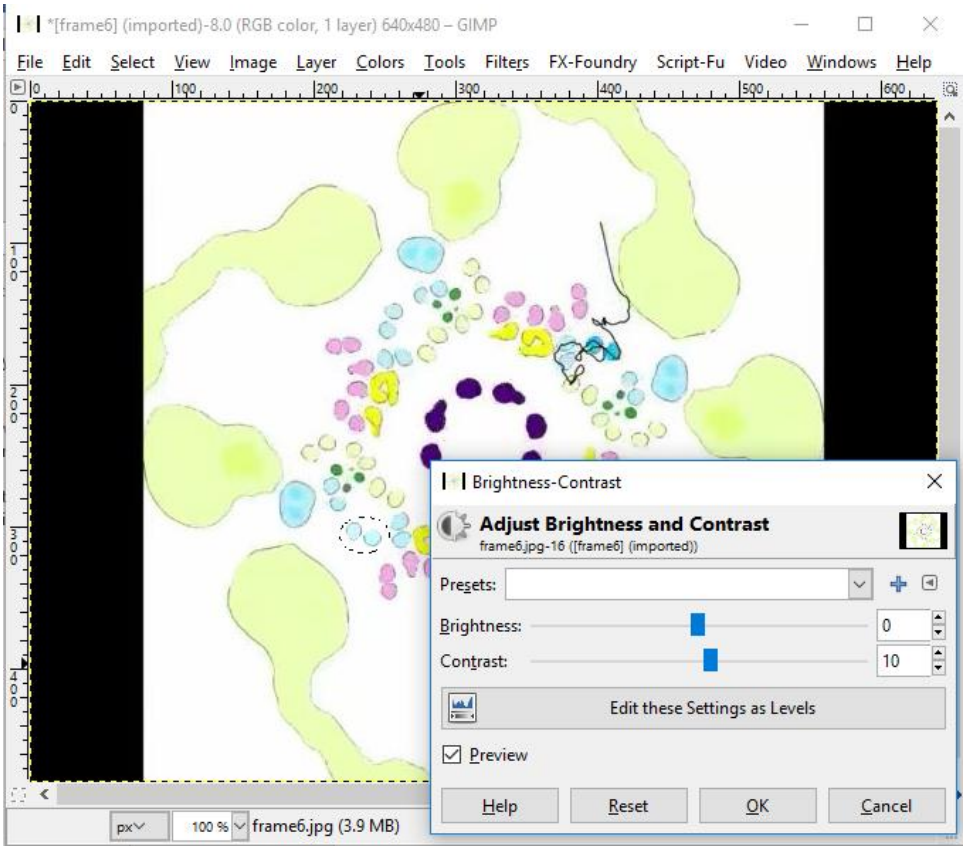
Mapping the rotation of the selected dimer

The cyan dimer isolation is confounded by the black line overlay designed to show the path of movement through its layer over time. The diagonally opposite dimer to the cyan dimer is selected as representative. This exactly opposite dimer moves in the exact same pattern as the cyan dimer.

After isolation of the 37 frames, the representative dimer is separated from the rest of the image using GIMP and saved as .png and .jpg files. The following progression is done to convert the selected dimer to neonpink to distinguish the selected dimer from the other dimers in the ribbon:

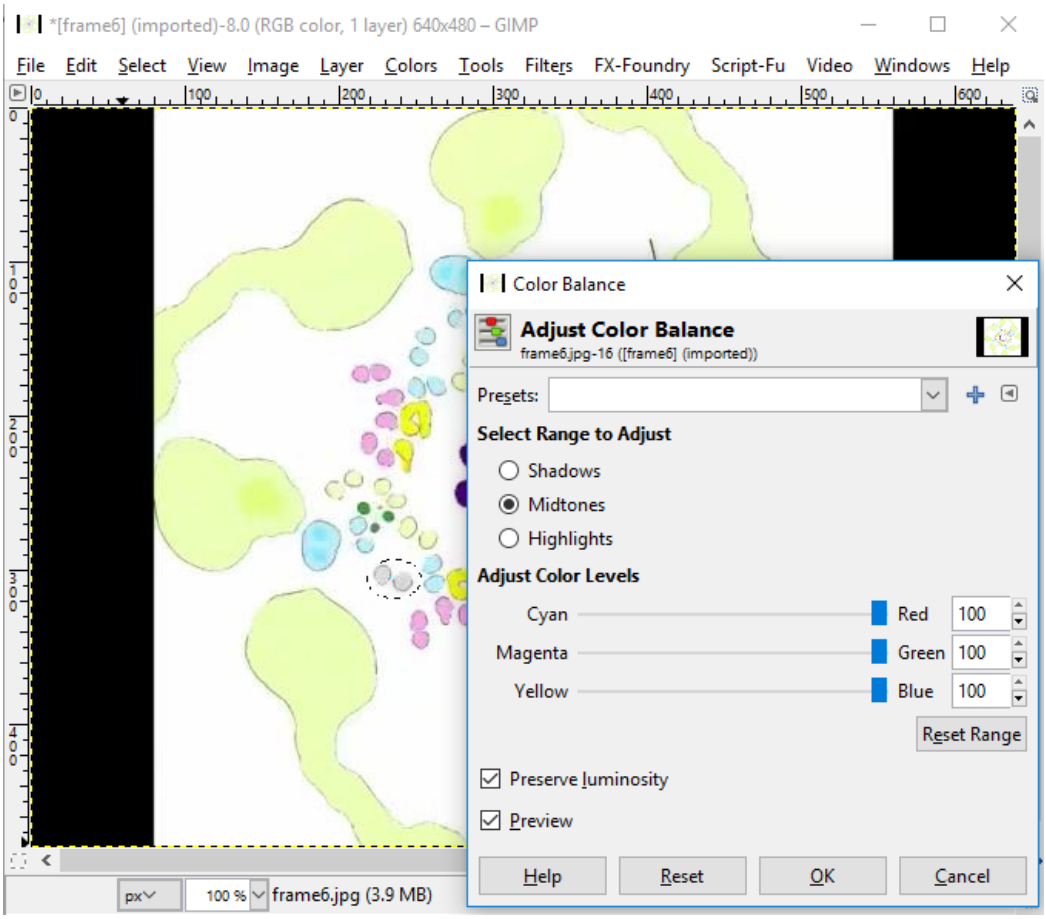
GIMP Protocol

1. (Colors -> Adjust Brightness and Contrast) Add 10% contrast
2. Select dimer of interest
3. (Colors -> Adjust Brightness and Contrast) Add 10% contrast (**Scn. 1**)



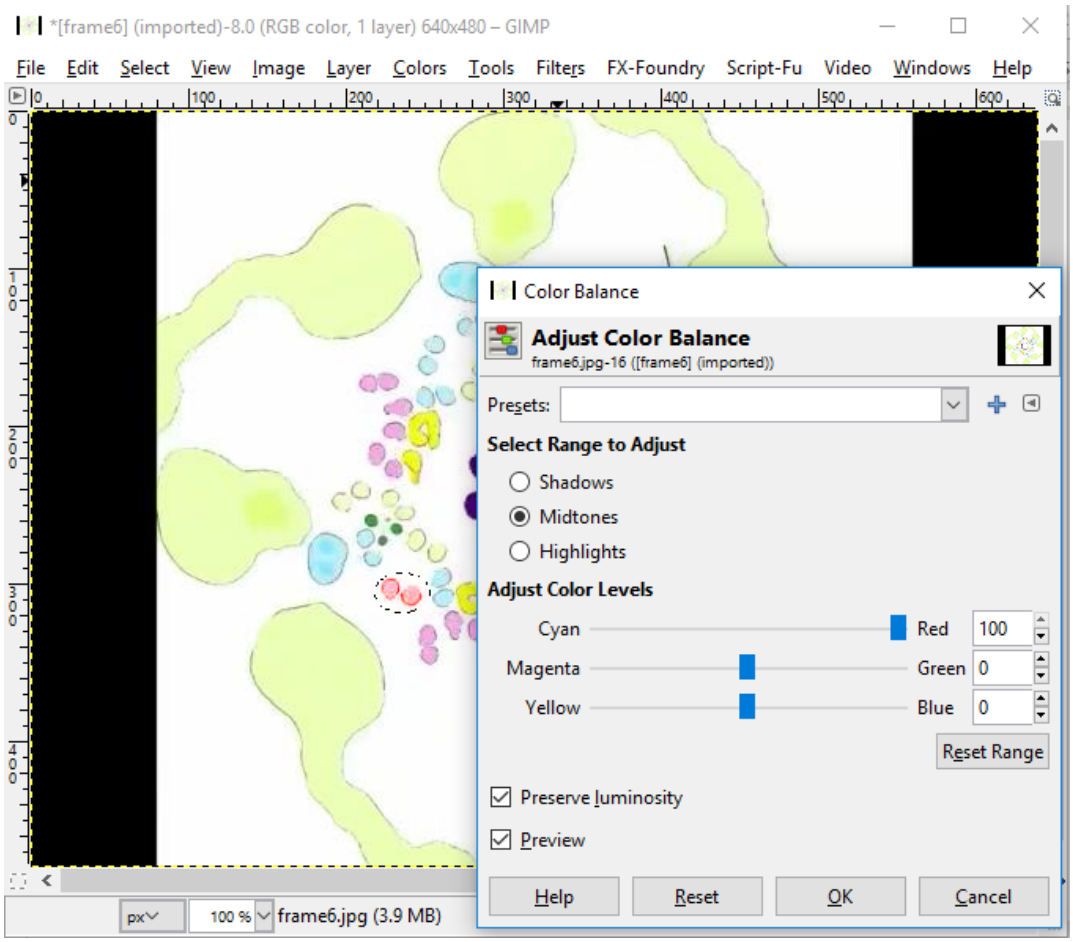
Screenshot 1: Contrast is increased for the background and then specifically for the dimer.

4. (Colors -> Adjust Color Balance) Set to Grey (Midtones, maxed) (**Scn. 2**)



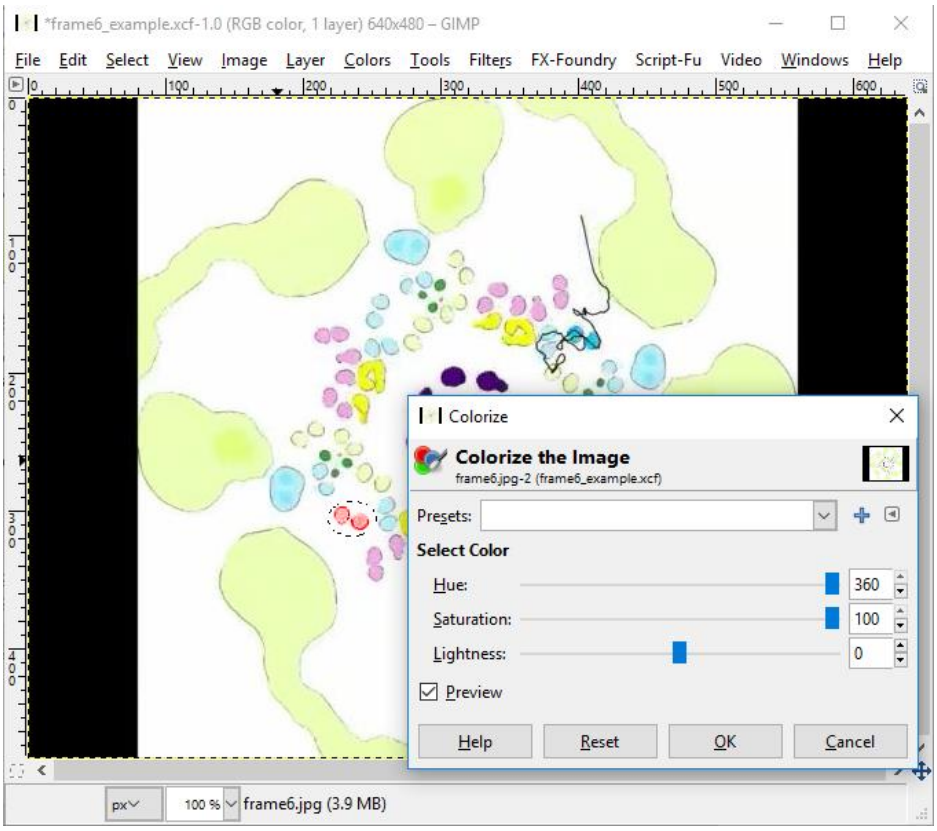
Screenshot 2: With the dimer still selected, midtones are set to maximum.

5. (Colors -> Adjust Color Balance) Highlights, max to Red (**Scn. 3**)



Screenshot 3: With the dimer still selected, red levels are maximized.

6. (Colors -> Colorize) Max Hue and Saturation (**Scn. 4**)



Screenshot 4: In the final step, the dimer is further emphasized by setting Hue and Saturation to max with the dimer still selected.

In the above progression, the contrast and brightness of the entire image is enhanced. A small area, inclusive only of the dimer of interest and background, was selected, and its brightness and contrast were further increased. Color balance is then adjusted to starken

the color contrast of the dimer from the rest of the image. This is done for each frame and then saved as a PNG file.

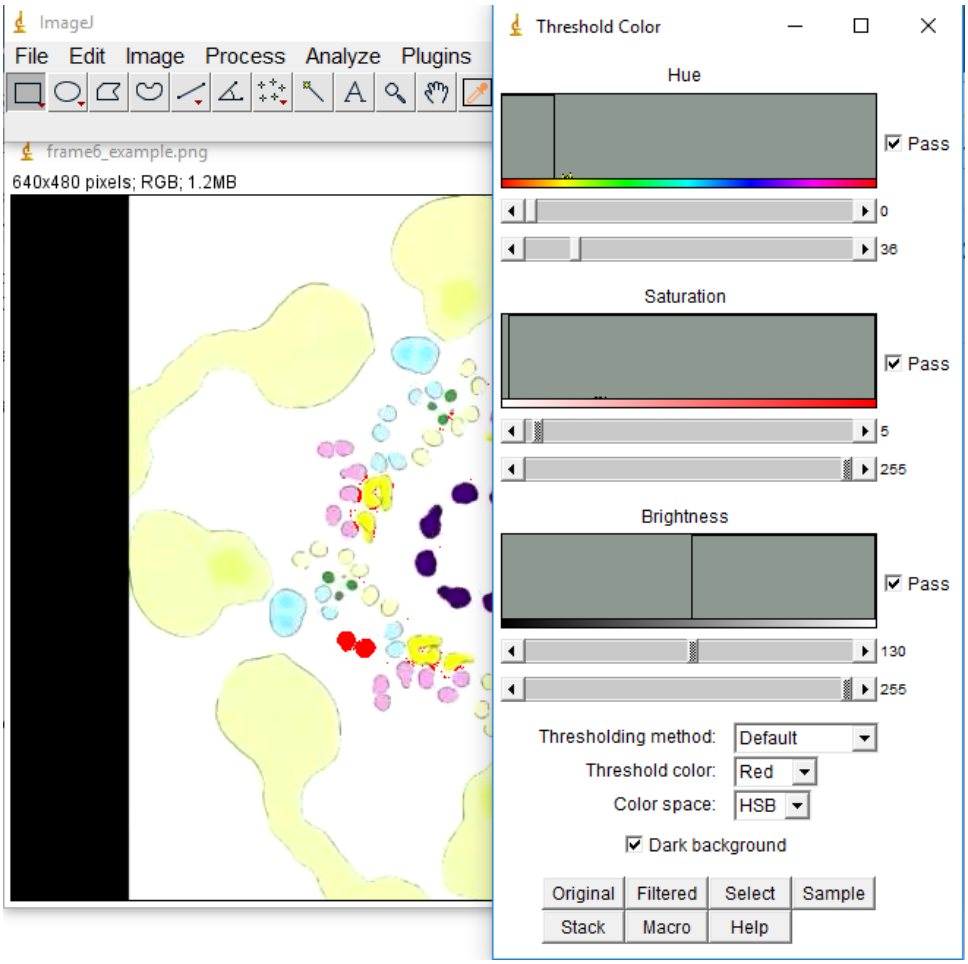
These frames are evaluated in ImageJ:

ImageJ Protocol

Part 1: Isolating the dimer fit to an ellipse

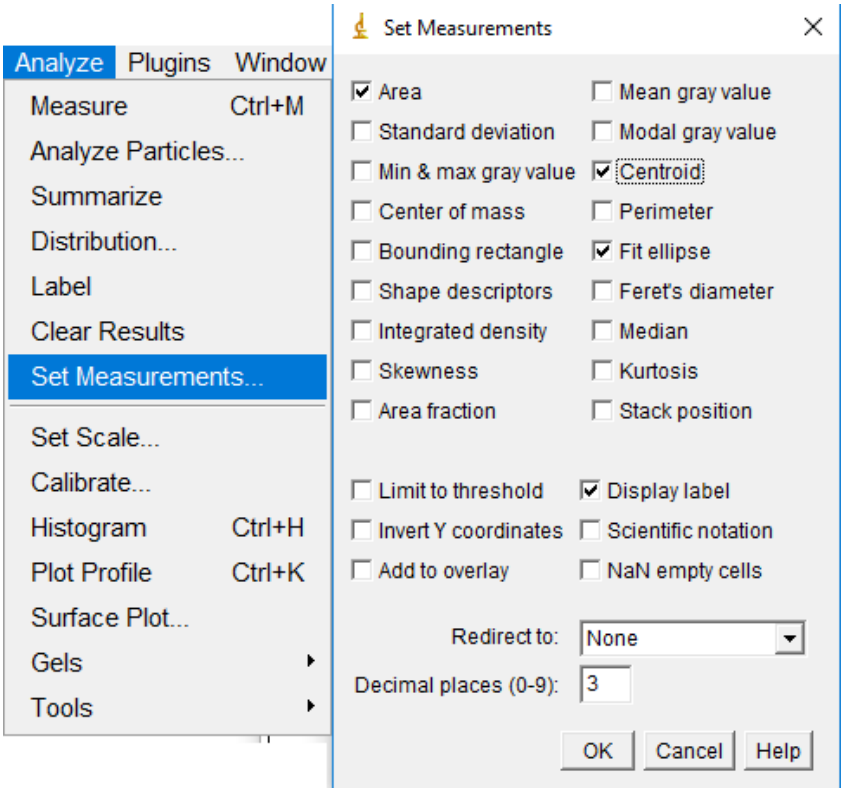
In order to attain an axis for each frame over which position and rotational change can be tracked, the dimer must first be fit to an ellipse that encompasses both myosin helices.

5. (Image>Adjust->Color Threshold) Threshold Color to point just before selection circle is not evident around dimer of interest: Hue 0, 36; Sat 5, 255; Brightness 130, 255 (**Scn. 5**).
    - Be sure ‘pass’ is selected for all options.



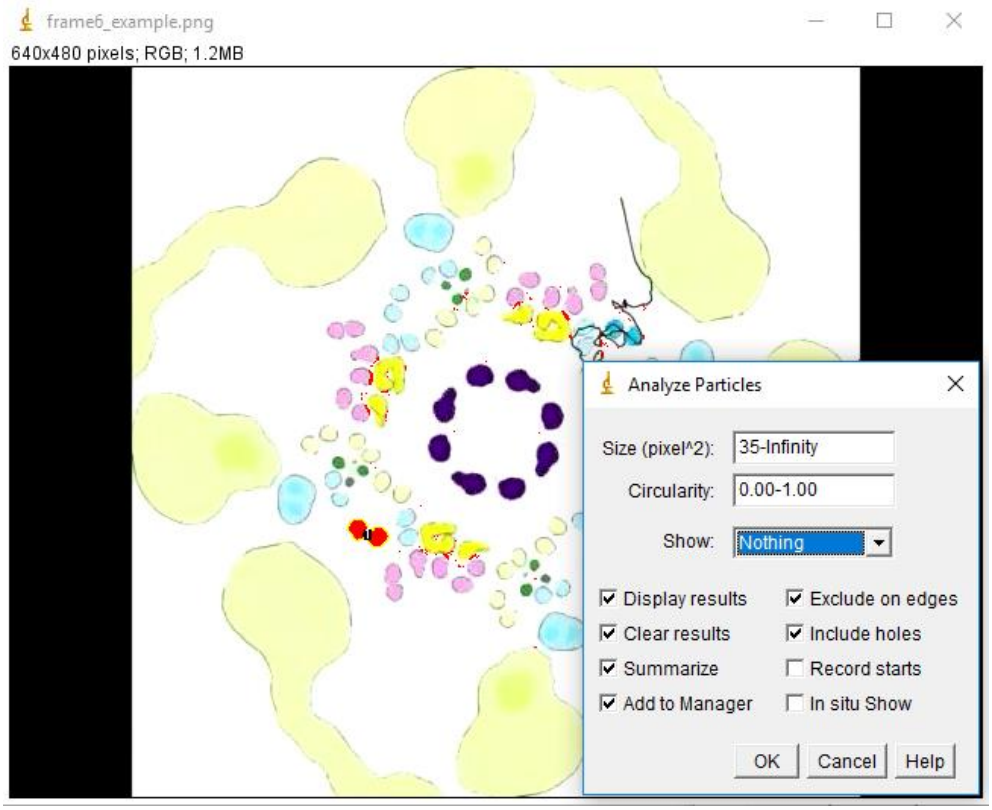
Screenshot 5: ImageJ threshold adjustment. Threshold is adjusted to eliminate surrounding selection circle and specifically select dimer of interest.

6. (Analyze->Set Measurements) Ensure that ‘area’, ‘centroid’ and ‘fit ellipse’ are checked. Check other parameters as desired (**Scn. 6**).



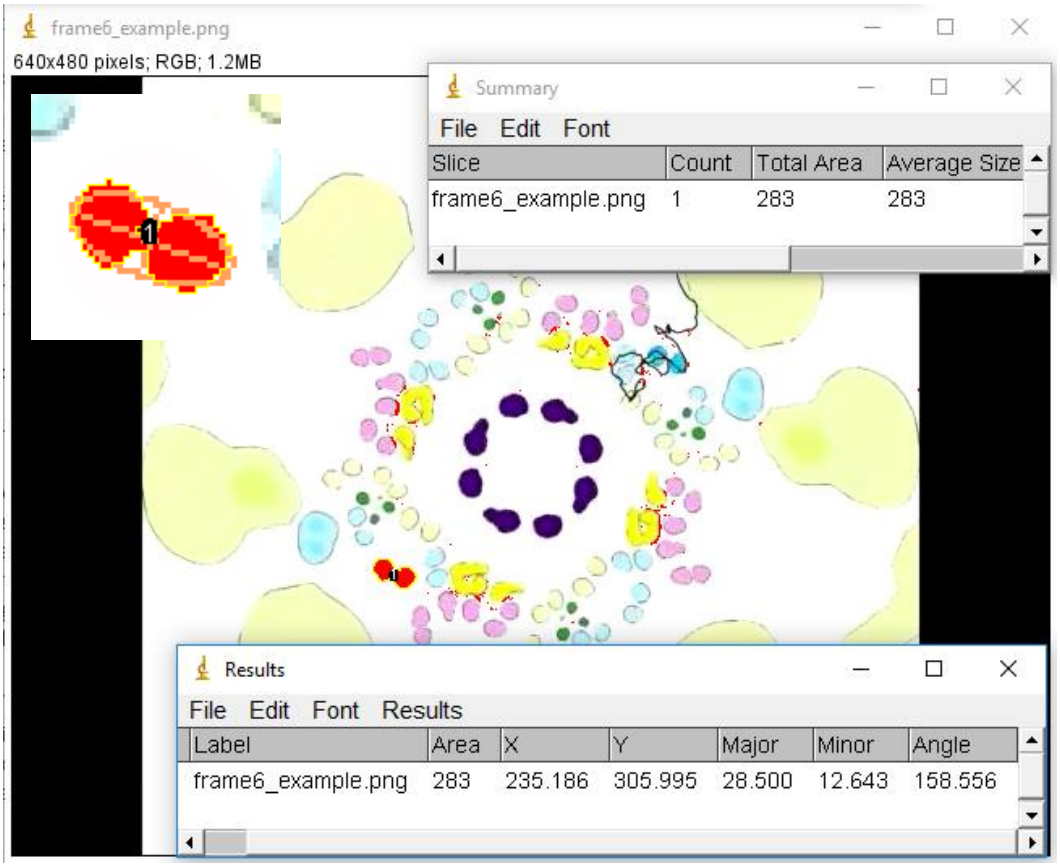
Screenshot 6: Set measurements to obtain values for an ellipse. Other selections can be useful for checking your calculations as desired.

7. (Analyze->Analyze Particles) Set min pixel size to 35* to reduce off-target measures (**Scn. 7**)
    - *This may need to be adjusted depending on dimensions of globular units and any external confounding portions of the image. Alternatively, selecting only the dimer of interest while using this function (Analyze Particles) will frequently suffice without pixel size adjustment.



Screenshot 7: Using the imageJ Analyze Particles function. All particles detected that are within threshold and the specified size range will be fit to ellipses. Adjusting pixel size will reduce off-target measures.

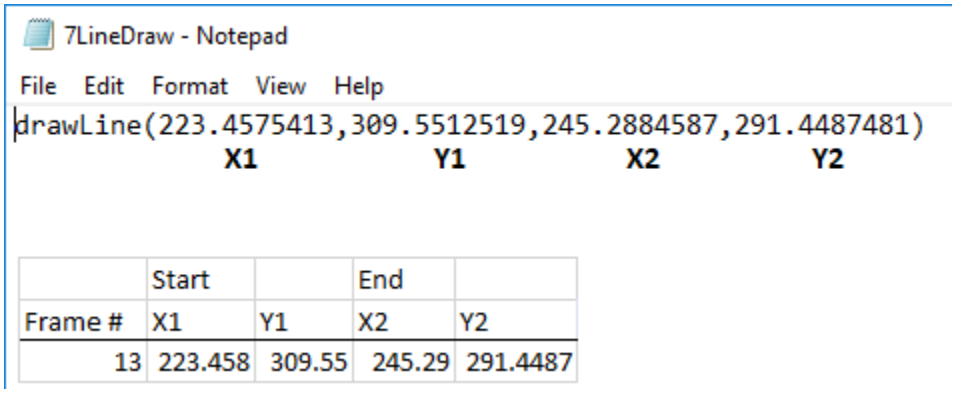
8. Set measurements to fit ellipse and centroid. This will automatically select the two globular units as a single ellipse. Measurements of the major axis angle and the length of the major (long) and minor (short) axes of the ellipse are provided by the program under these settings.
    - If an ellipse is not automatically mapped, the two globular units can be selected using the wand tool. Double click on the wand tool to set a threshold. Hold down shift-click to select both. Then “Measure”. The new output will have the appropriate measurements for a fitted ellipse.
    - The “Draw Ellipse” macro can be used to visualize your measured ellipse. (Edit->Selection->Fit Ellipse) An example of this is the upper left inset for **screenshot 8**.



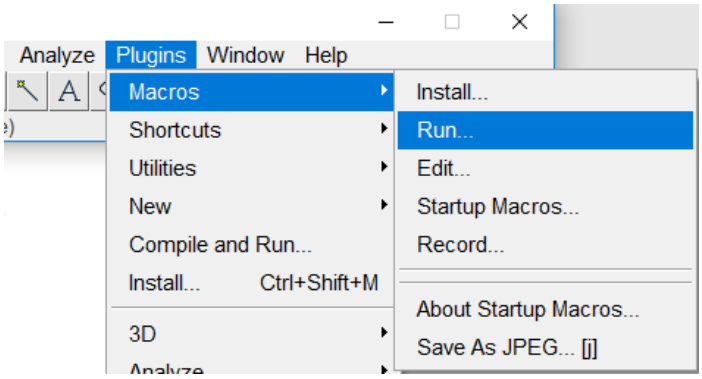
Screenshot 8: Output from fitting ellipse. Lengths of the Major and Minor axis, the X,Y coordinates of the mid-point, and Major axis angle of the fitted ellipse are provided upon measurement.

Part 2: Attaining angle of interface on dimer to non-myosin densities

9. Calculate the X,Y coordinates for the start and end points of the major axis (**Table 4-4**) and manually insert those values into a new Draw Line macro. Run that macro on your image (**Scn. 9-11**).



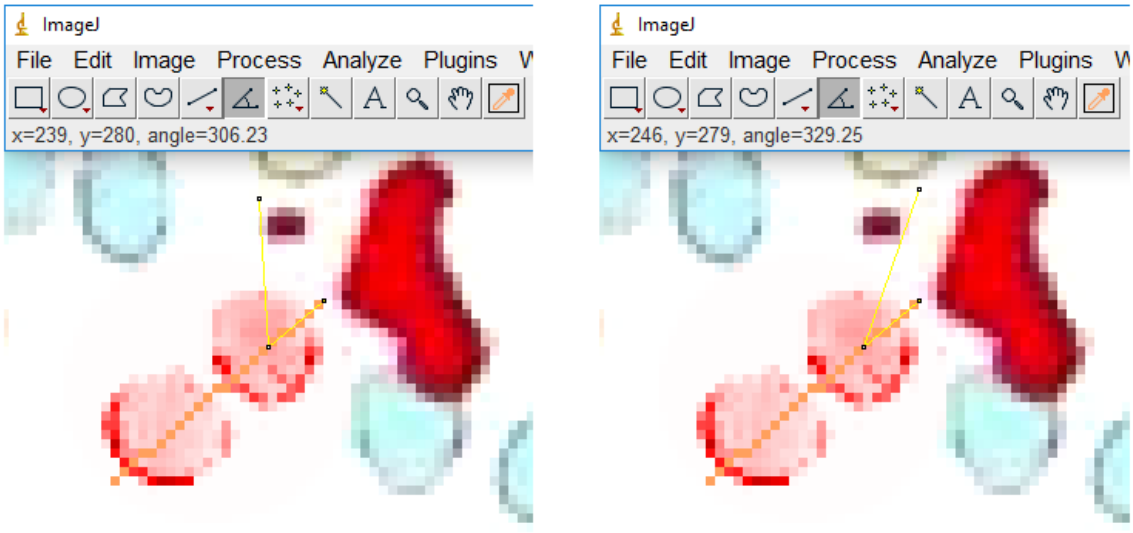
Screenshot 9: Draw a line through the Major Axis of the dimer. Macros are in text file format and must be adjusted for each frame. Inset (lower left) are example calculated end points for a Major line of a sample ellipse.



Screenshot 10: Make sure your version of ImageJ is configured to use Plugins to incorporate Macros.

Screenshot 11: The Draw Line macro will draw a line along the ellipse major axis.

10. Double click on the angle function and check the box “Measure reflex angle”. Use the angle feature to trace over the drawn line representing the major axis such that the angle parallel to the line is 180 to 360 (**Scn. 12**).



Screenshot 12: Trace and Angle function. Using the Angle function, trace the Major axis from the end closest to the helix you are measuring for to the center point of that helix. Here the trace is shown as a yellow line. When tracing, the start and end points can be approximate but the directionality must be maintained such that you draw from the edge of whichever globular unit you are looking for angles from, to the center. Here, the angles given (306.23 and 329.25) is from the view that the major axis line runs from 180 to 360 from left to right, diagonally in this case. This angle is specifically for the globular unit it is centered on. You can move the second leg of the trace after centering it to show the change in angle; this is shown in the two frames.

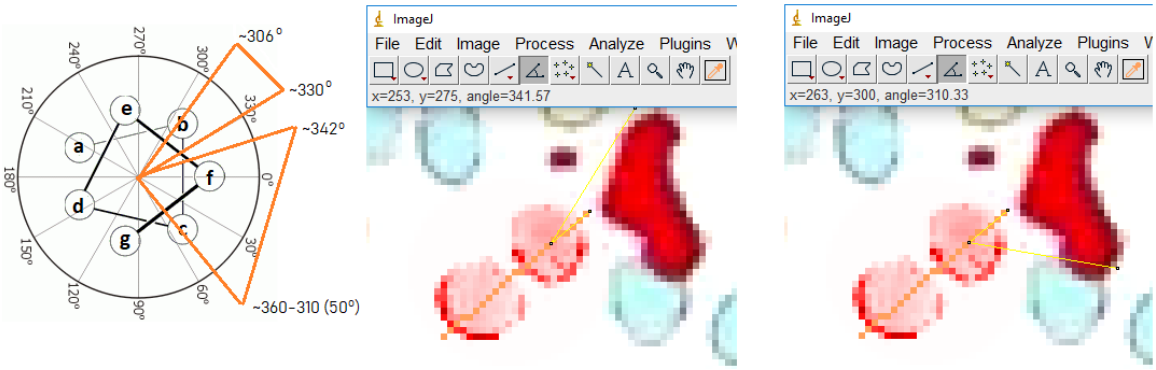
11. Measure the angle range over which the interacting helix is facing the density of interest (**Scn. 13**).

    - Subtract 360 from any value that passes the 360 point parallel to the major axis to get the ‘actual’ value.

12. Compare the angle ranges to the heptad positions (**Scn. 13)**

The coiled coil is left-handed and composed of two right-handed helices. If you view the frames as they are in video S3, M-ward, you see a predominant left handed rotation. The arrangement of the two helices in accordance to heptad positions is show in **Figure 4-12**.

- If selection numbers are getting in the way of drawing your line, delete from ROI manager



Screenshot 13: Heptad related orienting of the dimer. The heptad wheel on the left represents the position of the globular unit whose angles are being measured in the two right screenshots. It can be considered helix "2" referencing Figure 4-12. The angle interpretations for the middle and left screenshots and for Screenshot 12 are shown on the helical wheel. When reflex angle is selected, the values provided do not go below 180 and the zero to 180 axis is 180 to 360. This is accounted for by subtraction of 360 from all values below the major axis line.

To summarize, GIMP contrasting allows the neonpink dimer to be thresholded out of the rest of the image. It is then analyzed as particles. The particles are the globular units of the dimer and both particles are then selected and fit to an ellipse. At that point, it is possible to use the measurements of the ellipse to track movement of the dimer over a series of frames.

Part 3: Attaining coiled-coil rotation over amino acid range

ImageJ fitting provides X,Y coordinates of center of ellipse, pixel length of major and minor axis line, and angle of major axis as described in Part 2. The angle of the major

axis represents the major axis line angle to the horizontal of the image out of 180 degrees. The major axis angle is provided once the dimer is fit to an ellipse and can be seen in **Screenshot 8**. To check, the angle function can be used while going over the line used to form Ellipse using the Macro, "Draw Ellipse" function. These values are collected from each frame and the angle change between frames is recorded and graphed against the associated amino acid range with each frame progressing 2.7 amino acids. For simplicity, the starting angle in frame 1 is set to zero degrees. This can then be graphed in sequence to visualize the rotational shift. **Table 4-5** shows the first six frames for the cyan dimer. By using the method of calculating amino acid segments in Part 1, we did this for the 100 aa range of multiple dimers within the blue layer and single dimers within the yellow and pink layers (n=6) to confirm conservation among various dimers within the same region (not shown).

When mapping the entire dimer over the full range of the visible dimer there will be some discrepancy due to the increased span over which drift will occur. Drift is resultant from the actual frame span being very slightly off from 4 angstroms and 2.7 aa per frame also being a rounded value. Small imperfections are magnified over long spans that are not evident within a 100 aa range. The discrepancy is still small and ends up being less than 1 frame. Frames 60-411 have been evaluated both with and without a forced fit to the check points. There is no substantial loss of information and the regions over which there are rotational shifts are consistent (**Img. 3**). We chose a forced fit to the check points.

Using Graphpad PRISM, linear regression analysis was done (**Img. 4**). R-square values are very similar regardless of whether residue numbers are fixed to checkpoint-values (0.9977 vs 0.9976) suggesting that the methods are comparable for making observations over larger amino acid regions; though it should be stressed that goodness of fit, in this case, says nothing of either method being more or less representative of the coiled coil. Both methods would not be as comparable for more specific observations, as is done using the previously described methods for our smaller region of interest.

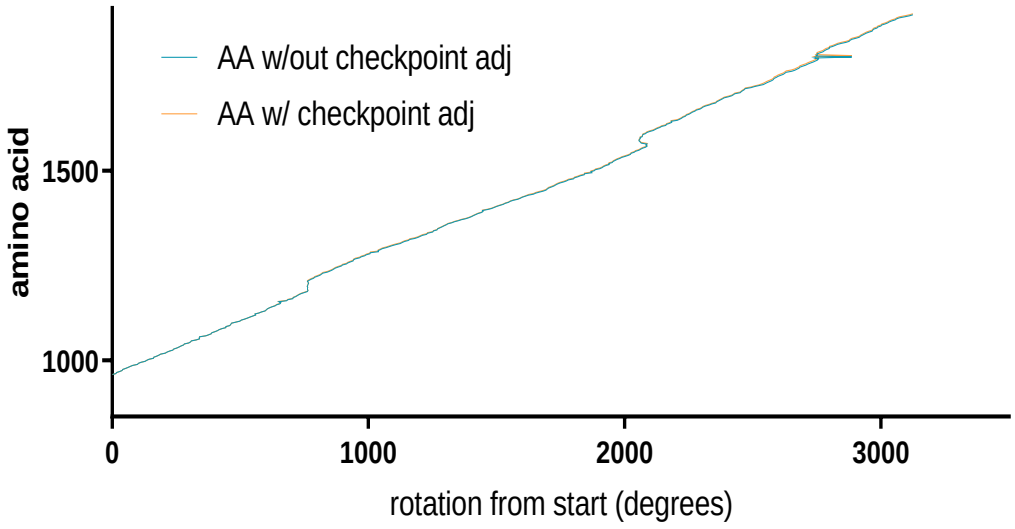


Image 3: Comparison of whole dimer rotation method choices. There is negligible deviation between forced-fit of the frames 60-411 aa range to the checkpoints (orange) and same frames aa range resultant from calculation strictly from the frame which the tether and rod measures from (frame 19, N852).

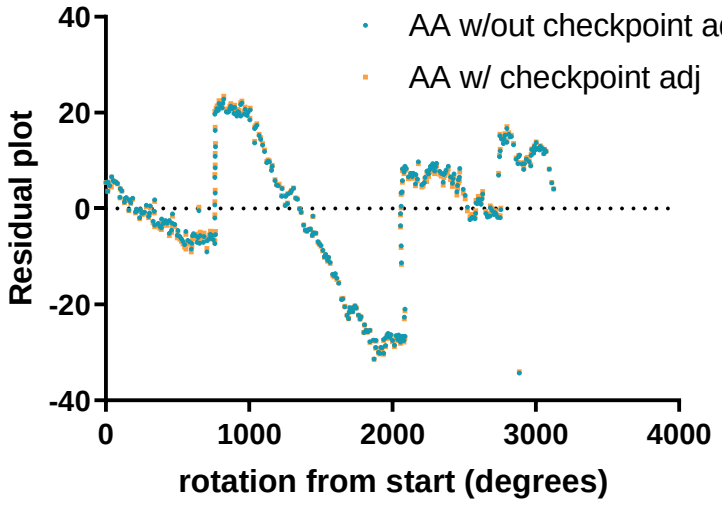


Image 4: Residuals of the linear regression of rotation methods. The differences here represent discrepancy between the two methods. R-squared values to the line of best fit are 0.9977 (without checkpoint) and 0.9976 (w/checkpoint).

Identifying specific residues of greatest proximity of the LMM dimer to Red density

Within the 37 frames of the selected region, the red density appears from frame 7 to 19. The dimers turn counter-clockwise in this M-ward view, therefore the alpha helices are turning clockwise [30]. The heptad positions are viewed to be in the format shown in **Figure 4-12**.

Although the rods rotate over time, the major line of the ellipse formed between the two densities is considered to be zero to 180 degrees as the horizontal major axis in **Figure 4-12**. If, for instance, the red density is closest to Helix 1 (the left helix in **Fig. 4-12**), between the angles of 230-300 then positions ‘c’ and ‘g’ are noted as being in closest

proximity. If the range for the frame (~3 aa) only includes the ‘g’ position, this position is noted in regular font and the closest ‘c’ position is noted in italics. If there are two ‘g’ positions equi-distant and outside the frame, they are both noted in italics. An example of what this notation looks like using an actual case is shown in **Image 5** in the AA Pos column.

Notation is not taken past a 5 pixel distance (~3.2 Å); this relationship is estimated using Hu et al. [3] Fig. 4A.

Image 6 shows a sample of notation for one frame.

Frame #	X1	Y1	X2	Y2	Helix #	Heptad Pc	AA Pos		Angle Range		AA Region	
10	222.5244	305.4268	248.1956	297.5072	1	c	E1554		212	240	1552.9	1555.6
					1	g	*A1558*	*A1551*	385	320		
						d	A1555					
					2	e	E1556		315	230		
						b	E1553					

Image 5: Notation for one frame. Example of recording with coordinates for the Major axis line, helix designation, associated heptad positions and amino acid positions followed by the angle range obtained and the entire region the frame encompasses (in this case, 1552-1556).

Mapping of the myosin dimer to the red density across frames 7-19 allows positions to be identified for probable amino acids involved in the interfacing between the two. This combined with dimer rotation measurements described in Part 3 allows correlation of pitch change along this region.

FIGURES

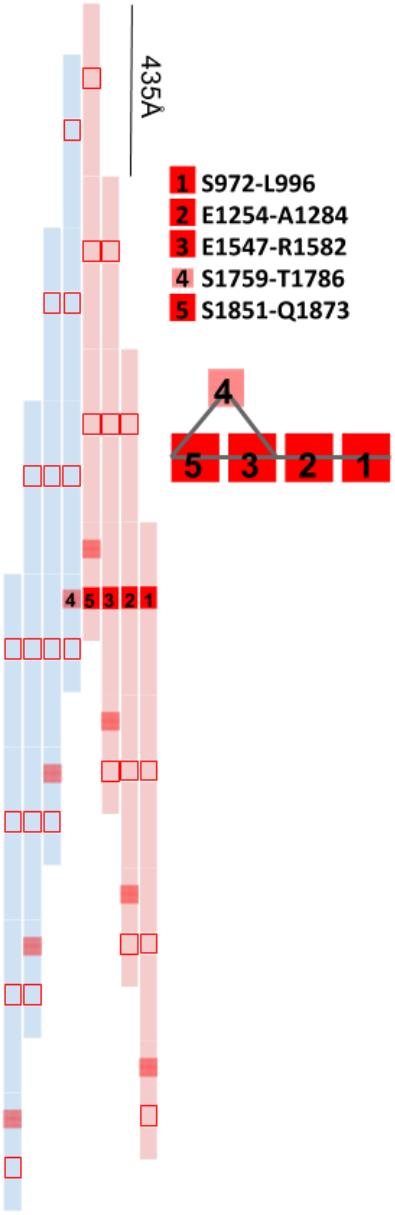


Figure 4-1. Flightin contacts between two layers. Each myosin contains a 1600 Å rod region that overlaps with each other such that 3-4 dimers (rods) are aligned at any one time at a stagger of 435 Å. The red rectangles represent the red density contact sites along the length of four dimers. The numbers 1-5 show the LMM interfaces for a single red density, linking each rod to the other rods in its layer and once to a neighboring layer (4, red stripes). The area of position 3-4-5 is the multiface.

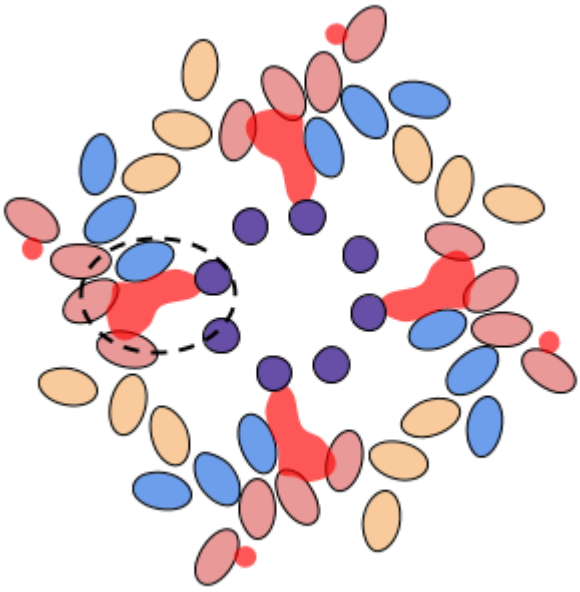


Figure 4-2. Graphical view down the filament axis (Redrawn from a frame of Hu et al, 2016 [3] movie S3). Ovals represent myosin dimers with the color (pink, yellow, blue) representing individual layers. The red density is shown in a translucent red and is circled at the multiface (dashed oval). Here, the red density contacts three LMM interfaces simultaneously with two dimers belonging to the 'pink' layer and one dimer belonging to the 'blue' layer. Contact with paramyosin (purple circles) is made at the center of the thick filament.

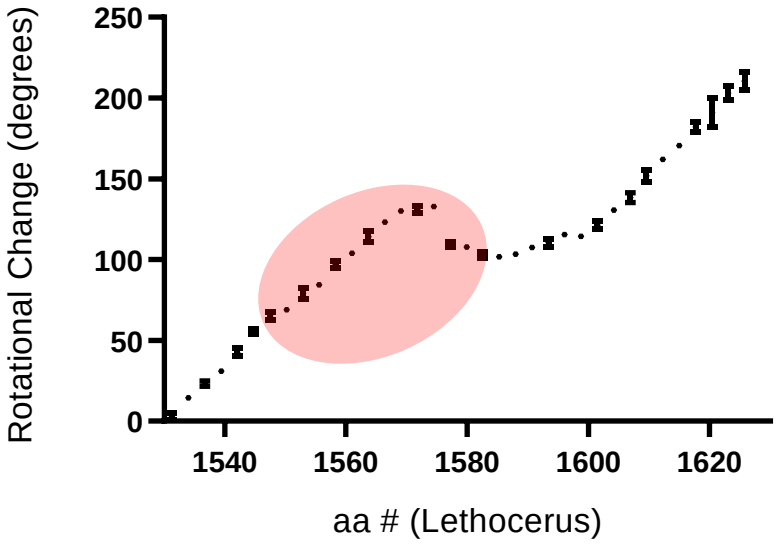
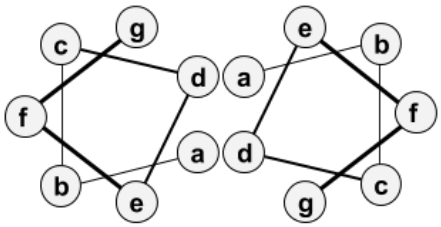


Figure 4-3. M-ward helical rotation of the coiled-coil from G1528-A1628. The change in angle represents rotational change between frames. Each point is averaged across three dimers within the same layer±SD. The points within red shading constitutes the area over which the red density is in contact with the myosin dimer.



d	a	e	b	f	c	g	d	a
I 1534		E 1535		K 1536		Q 1537	I 1534	
L 1541	K 1538	E 1542	K 1539	V 1543	R 1540	E 1544	L 1541	K 1538
L 1548	K 1545	Q 1549	D 1546	A 1550	E 1547	A 1551	L 1548	K 1545
A 1555	L 1552	E 1556	E 1553	A 1557	E 1554	A 1558	A 1555	L 1552
E 1562	L 1559	E 1563	E 1560	N 1564	Q 1561	K 1565	E 1562	L 1559
S 1569	V 1566	Q 1570	L 1567	L 1571	R 1568	E 1572	S 1569	V 1566
V 1576	L 1573	R 1577	S 1574	Q 1578	Q 1575	E 1579	V 1576	L 1573
R 1581	I 1580	I 1584	D 1581	Q 1585	R 1582	E 1586	R 1583	I 1580

Figure 4-4. Heptad-mapped positions for red density contact between I1534-E1586. (TOP) The heptad positioning over which the myosin monomers associate to form a coiled-coil dimer. A-g refer to heptad positions in which the ideal heptad harbors a pattern of HPPCPC where H=hydrophobic, P=polar, C=charged. (BOTTOM) Interface over which red density proximity was identified from perspective of heptad positions (a-g, top row) in the side-order the amino acids would be found in a helix. 'd' and 'a' positions are repeated on either side to communicate that ‘e’ positions are proximal to the ‘a’ positions Thick grey lines separate the amino acids that are in contact with the red density.

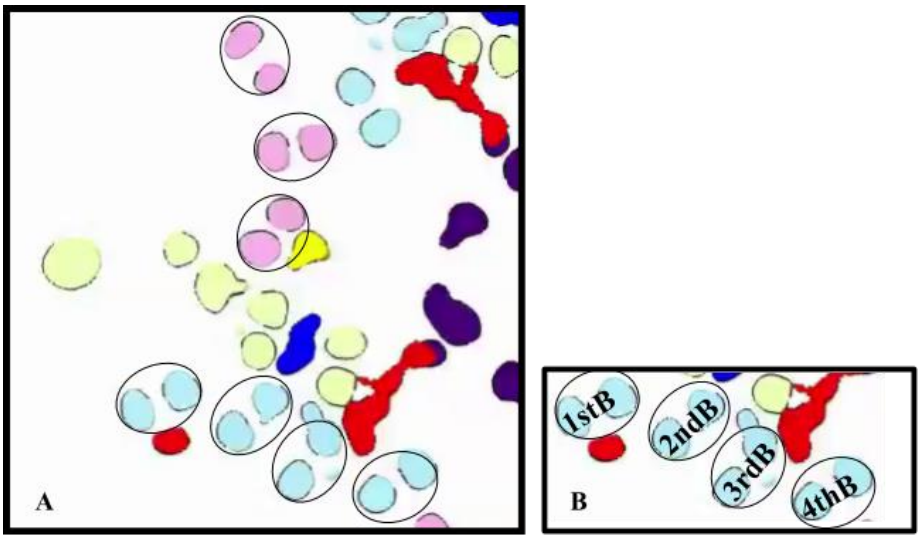
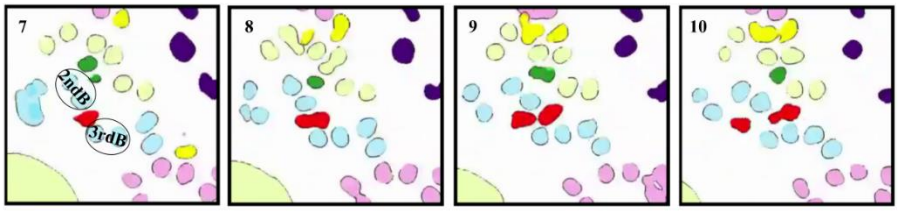
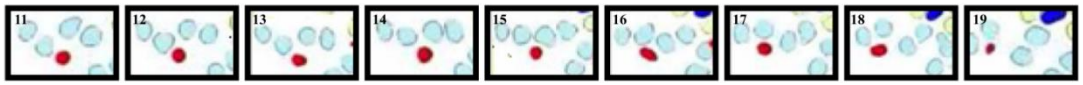


Figure 4-5: Three or four myosin dimers in a crown can be viewed in association with each other. In this frame, (A) the pink, yellow and blue layers are shown with circles around the dimers of the pink and blue layers. Note that there are three dimers associated with the pink layer, and four for the blue layer. In (B) the notation used for referring to dimers within the blue layer is provided.



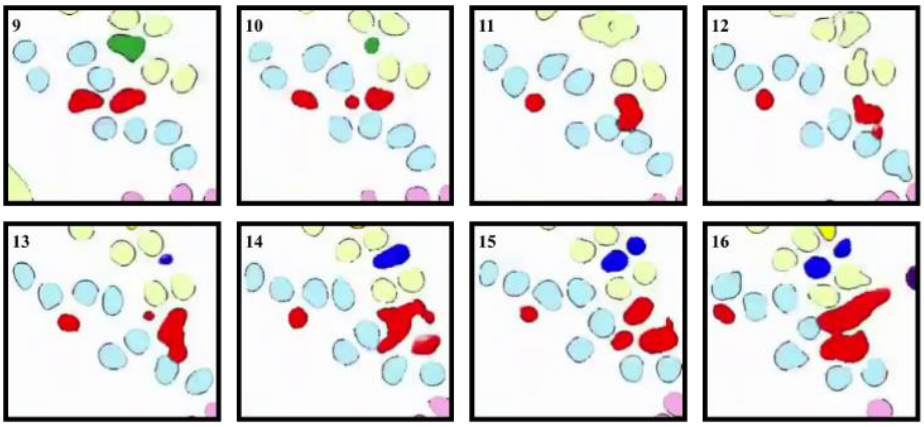
2ndB		3rdB		Frame	
E	1254	D	1546		
K	1255	E	1547	7	
I	1256	L	1548		
A	1257	Q	1549		8
K	1258	A	1550		
Q	1259	A	1551		
L	1260	L	1552	9	
Q	1261	E	1553		
H	1262	E	1554		
Q	1263	A	1555		10
V	1264	E	1556		
N	1265	A	1557		

Figure 4-6: Smaller red density association within LMM aa ranges E1254-N1265/D1546-A1557. (Left) The splitting of the red density takes place over 4 frames, #7-10 from the set analyzed). Frame 7 has the 2ndB and 3rdB dimers designated. The four squares represent portions of frames in sequence M-ward. (Right) Associated table shows the aa regions for 2ndB and 3rdB over this range. Residue colors correspond to hydrophilic (blue), hydrophobic (green), negative (yellow) and positive (red).



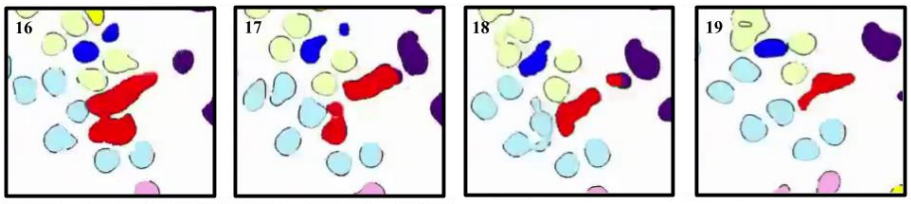
Frame				12					14					16					18					
	11					13					15					17					19				
1stB	S	D	Q	D	K	A	S	K	D	H	Q	I	R	N	L	N	D	E	I	A	H	Q	D	E	L
	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996
2ndB	N	E	I	Q	G	K	L	D	E	A	N	R	T	L	N	D	F	D	S	A	K	K	K	L	S
	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289

Figure 4-7: Smaller red density association within LMM aa ranges S972-L996/N1265-S1289. (TOP) Portions of frames #11-19 that show the 1st and 2nd (1stB, 2ndB) blue dimers proximal to the red density. (BOTTOM) The corresponding amino acid ranges associated with the frames #11-19.



Frame			10						12				14						16		
	9				11					13					15						
3rdB	A	L	E	E	A	E	A	A	L	E	Q	E	E	N	K	V	L	R	A	Q	L	E
	1551	1552	1553	1554	1555	1556	1557	1558	1559	1560	1561	1562	1563	1564	1565	1566	1567	1568	1569	1570	1571	1572
4thB	K	N	L	R	K	S	E	R	R	I	K	E	L	S	F	Q	A	D	E	D	R	K
	1846	1847	1848	1849	1850	1851	1852	1853	1854	1855	1856	1857	1858	1859	1860	1861	1862	1863	1864	1865	1866	1867

Figure 4-8: Larger red density association within LMM aa ranges A1551-E1572/K1846-K1867. (TOP, MIDDLE) The red densities to the right of the frame segments are considered the 'larger' red density. After the split (frame 9) interaction of the larger red density predominantly contacts 3rdB, 4thB and 4thY. (BOTTOM) Lower panel shows corresponding 3rdB and 4thB sequences associated with frames #9-16.



Frame			17						19		
	16					18					
3rdB	Q	L	E	L	S	Q	V	R	Q	E	I
	1571	1572	1573	1574	1575	1576	1577	1578	1579	1580	1581
4thB	D	R	K	N	H	E	R	M	Q	D	L
	1865	1866	1867	1868	1869	1870	1871	1872	1873	1874	1875
4thY	D	E	L	R	A	E	Q	D	H	A	Q
	1775	1776	1777	1778	1779	1780	1781	1782	1783	1784	1785

Figure 4-9: Larger red density association within LMM aa ranges Q1571-I1581/D1865-L1875/D1775-Q1785. (TOP) Portions from frames #16-19 focusing on the larger red density at the 'multiface'. (BOTTOM) The aligned amino acid ranges for 3rdB, 4thB and 4thY for the associated frames.

.

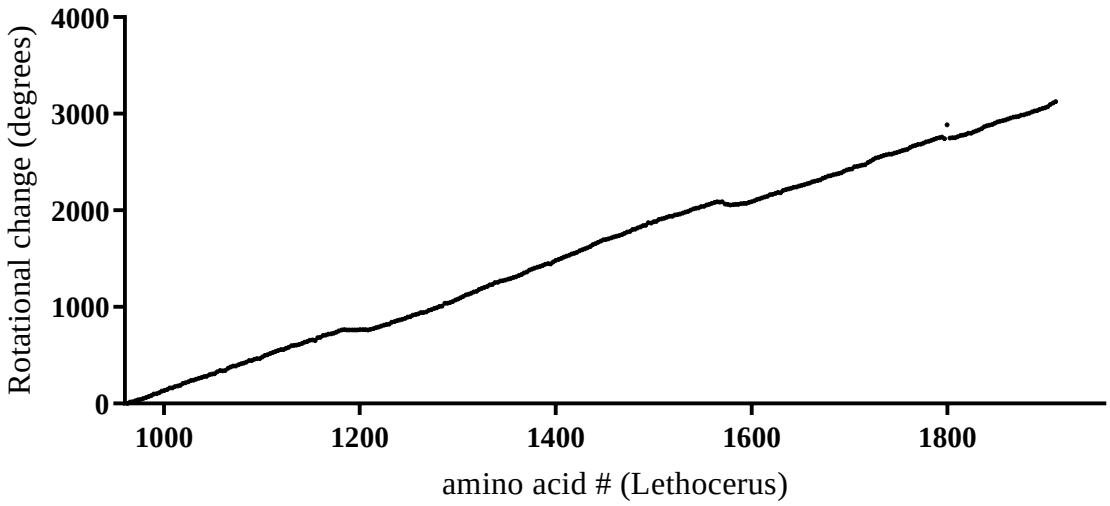


Figure 4-10: Dimer Rotation from I962-E1916. Large scale observations of dimer rotation show only two other disrupted regions that center around Q1192 and E1800

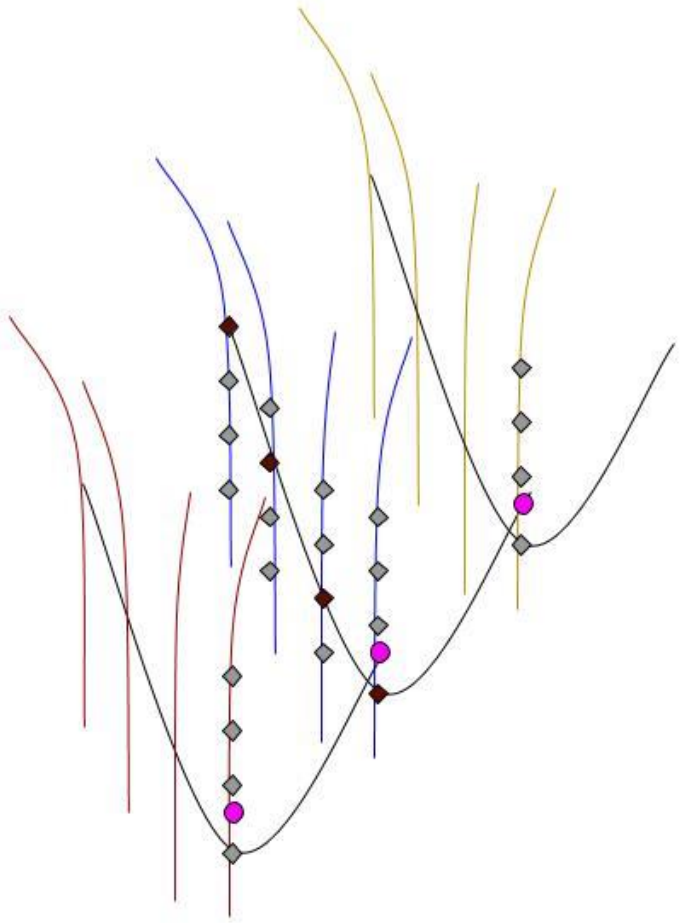


Figure 4-11: Red density contact among myosin dimer layers concept image. In this concept image, all rods in a given layer (ex. blue) are connected to each other at four points (black diamonds) by a red density (black line) and to a rod in the neighboring layer at 1759-1786 (pink circle) which would be located between two regions that have contacts with two different red densities that wind within that neighboring layer (not shown).

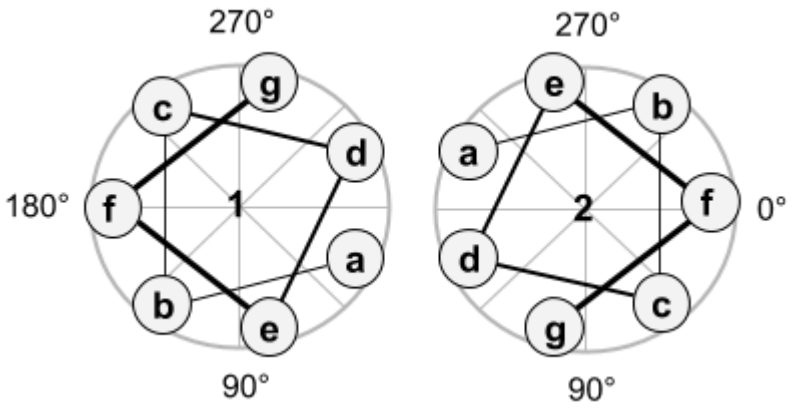


Figure 4-12: Heptad mapping angles. The heptad is mapped as two right handed α helices participating in a left-handed coiled coil. Because angle measurements are not the same for each helix along the major axis, the helices are given arbitrary "1" and "2" designations.

TABLES

Table 4-1: Å values used for calculations in the analysis of Hu et al. Movie S3.

1700 Å	Myosin distance encompassed by Movie S3
1598 Å	Contour length of myosin rod and tether
435 Å	Stagger between rods in a ribbon layer
4 Å	Distance between 2 frames
1.485 Å	Rise per residue relative to coiled
1 Å	Averaged slice shown per frame

Table 4-2: Drosophila aa associations to Hu et al. Movie S3*.

	S8 key without matching		Adjusted key with matching	
aa seg Dros	**Span (s)**	**Duration (s)**	**Span (s)**	**Duration (s)**
843-1136	1.3-8.6	7.3	1.3-8.55	7.25
1136-1428	8.6-15.7	7.1	8.55-15.8	7.25
1428-1721	15.7-22.9	7.2	15.8-23.05	7.25
1721-1919	22.9-27.76	4.86	23.05-27.91	4.86

*** The time spans and durations are shown for those provided by the fig.S8 legend for Sup.Vid.3 (gray columns) next to the span & duration identified by time-point matching from the start (2 right-most columns). Left-most column shows corresponding Drosophila amino acid segments.**

Table 4-3: How to calculate start/end coordinates for major axis using Excel.

Calculating Values for Major Axis Line	
Start X	Xg-(COS(Ag*PI()/180)*(Lma/2))
Start Y	Yg+(SIN(Ag*PI()/180)*(Lma/2))
End X	Xg+(COS(Ag*PI()/180)*(Lma/2))
End Y	Yg-(SIN(Ag*PI()/180)*(Lma/2))
Center point (Xg,Yg), Angle (Ag), and Length (Lma)are given:	
Xg	Given X
Yg	Given Y
Ag	Given Angle (Major)
Length for Major	Lma

Table 4-4: Durations corresponding to portions of the Drosophila and Lethocerus sequence.

Span (s)	Duration (s)	aa segment (Dros)	aa segment (Leth)
1.3-8.55	7.25	843-1136	852-1145
8.55-15.8	7.25	1136-1428	1145-1437
15.8-23.05	7.25	1428-1721	1437-1730
23.05-27.91	4.86	1722-1919	1731-1928

Table 4-5: Frame sampling*.

Frame	Angle	Deg change	aa range	
1	114.121	0	1528.6	1531.3
2	118.977	4.856	1531.3	1534
3	128.007	13.886	1534	1536.7
4	138.971	24.85	1536.7	1539.4
5	145.972	31.851	1539.4	1542.1
6	159.419	45.298	1542.1	1544.8

***Example of 6 frames associated with the corresponding angle for the major axis, degree change between frames and the amino acid range that the frame encompasses.**

BIBLIOGRAPHY

[1] Craig R, Woodhead JL. Structure and function of myosin filaments. Current opinion in structural biology. 2006;16:204-12.

[2] Squire JM. Muscle myosin filaments: cores, crowns and couplings. Biophysical reviews. 2009;1:149.

[3] Hu Z, Taylor DW, Reedy MK, Edwards RJ, Taylor KA. Structure of myosin filaments from relaxed Lethocerus flight muscle by cryo-EM at 6 A resolution. Science advances. 2016;2:e1600058.

[4] Vigoreaux JO, Saide JD, Valgeirsdottir K, Pardue ML. Flightin, a novel myofibrillar protein of Drosophila stretch-activated muscles. The Journal of cell biology. 1993;121:587-98.

[5] Qiu F, Brendel S, Cunha PM, Astola N, Song B, Furlong EE, et al. Myofilin, a protein in the thick filaments of insect muscle. Journal of cell science. 2005;118:1527-36.

[6] Reedy MC, Bullard B, Vigoreaux JO. Flightin is essential for thick filament assembly and sarcomere stability in Drosophila flight muscles. The Journal of cell biology. 2000;151:1483-500.

[7] Ayer G, Vigoreaux JO. Flightin is a myosin rod binding protein. Cell biochemistry and biophysics. 2003;38:41-54.

[8] Kronert WA, O'Donnell PT, Fieck A, Lawn A, Vigoreaux JO, Sparrow JC, et al. Defects in the Drosophila myosin rod permit sarcomere assembly but cause flight muscle degeneration. Journal of molecular biology. 1995;249:111-25.

[9] Ferguson C, Lakey A, Hutchings A, Butcher GW, Leonard KR, Bullard B. Cytoskeletal proteins of insect muscle: location of zeelins in Lethocerus flight and leg muscle. Journal of cell science. 1994;107:1115-29.

[10] Ltd A. ApowersoftEdit. https://www.apowersoft.com; 2017.

[11] Team TGD. GIMP. https://www.gimp.org; 2017.

[12] Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nature methods. 2012;9:671-5.

[13] Vigoreaux JO, Hernandez C, Moore J, Ayer G, Maughan D. A genetic deficiency that spans the flightin gene of Drosophila melanogaster affects the ultrastructure and function of the flight muscles. The Journal of experimental biology. 1998;201:2033-44.

[14] Lemke SB, Schnorrer F. Mechanical forces during muscle development. Mechanisms of development. 2017;144:92-101.

[15] Rose A, Meier I. Scaffolds, levers, rods and springs: diverse cellular functions of long coiled-coil proteins. Cellular and molecular life sciences : CMLS. 2004;61:1996-2009.

[16] Kreuzer SM, Elber R. Coiled-coil response to mechanical force: global stability and local cracking. Biophysical journal. 2013;105:951-61.

[17] Taylor KC, Buvoli M, Korkmaz EN, Buvoli A, Zheng Y, Heinze NT, et al. Skip residues modulate the structural properties of the myosin rod and guide thick filament assembly. Proceedings of the National Academy of Sciences of the United States of America. 2015;112:E3806-15.

[18] Brown JH, Cohen C, Parry DA. Heptad breaks in alpha-helical coiled coils: stutters and stammers. Proteins. 1996;26:134-45.

[19] Korkmaz EN, Taylor KC, Andreas MP, Ajay G, Heinze NT, Cui Q, et al. A composite approach towards a complete model of the myosin rod. Proteins. 2016;84:172-89.

[20] Henkin JA, Maughan DW, Vigoreaux JO. Mutations that affect flightin expression in Drosophila alter the viscoelastic properties of flight muscle fibers. American journal of physiology Cell physiology. 2004;286:C65-72.

[21] Chakravorty S, Tanner BCW, Foelber VL, Vu H, Rosenthal M, Ruiz T, et al. Flightin maintains myofilament lattice organization required for optimal flight power and courtship song quality in Drosophila. Proceedings Biological sciences. 2017;284.

[22] Gasek NS, Nyland LR, Vigoreaux JO. The Contributions of the Amino and Carboxy Terminal Domains of Flightin to the Biomechanical Properties of Drosophila Flight Muscle Thick Filaments. Biology. 2016;5.

[23] Tanner BC, Miller MS, Miller BM, Lekkas P, Irving TC, Maughan DW, et al. COOH-terminal truncation of flightin decreases myofilament lattice organization, cross-bridge binding, and power output in Drosophila indirect flight muscle. American journal of physiology Cell physiology. 2011;301:C383-91.

[24] Soto-Adames FN, Alvarez-Ortiz P, Vigoreaux JO. An evolutionary analysis of flightin reveals a conserved motif unique and widespread in Pancrustacea. Journal of molecular evolution. 2014;78:24-37.

[25] Lemas D, Lekkas P, Ballif BA, Vigoreaux JO. Intrinsic disorder and multiple phosphorylations constrain the evolution of the flightin N-terminal region. Journal of proteomics. 2016;135:191-200.

[26] Flashman E, Watkins H, Redwood C. Localization of the binding site of the C-terminal domain of cardiac myosin-binding protein-C on the myosin rod. The Biochemical journal. 2007;401:97-102.

[27] Obermann WM, van der Ven PF, Steiner F, Weber K, Fürst DO. Mapping of a myosin-binding domain and a regulatory phosphorylation site in M-protein, a structural protein of the sarcomeric M band. Molecular biology of the cell. 1998;9:829-40.

[28] Obermann WMJ, Gautel M, Weber K, Fürst DO. Molecular structure of the sarcomeric M band: mapping of titin and myosin binding domains in myomesin and the identification of a potential regulatory phosphorylation site in myomesin. The EMBO journal. 1997;16:211-20.

[29] Menard L, Nyland L, Vigoreaux J. The Structural and Biomechanical Properties of Insect Thick Filaments Expressing Flightin and Cardiac Myosin Binding Protein-C. Microscopy and Microanalysis. 2013;19:80-1.

[30] Peckham M, Knight PJ. When a predicted coiled coil is really a single α-helix, in myosins and other proteins. Soft Matter. 2009;5:2493-503.

[31] Salvi SS, Kumar RP, Ramachandra NB, Sparrow JC, Nongthomba U. Mutations in Drosophila myosin rod cause defects in myofibril assembly. Journal of molecular biology. 2012;419:22-40.

[32] Lamont PJ, Wallefeld W, Hilton-Jones D, Udd B, Argov Z, Barboi AC, et al. Novel mutations widen the phenotypic spectrum of slow skeletal/beta-cardiac myosin (MYH7) distal myopathy. Human mutation. 2014;35:868-79.

[33] Ferbert A, Zibat A, Rautenstrauss B, Kress W, Hugens-Penzel M, Weis J, et al. Laing distal myopathy with a novel mutation in exon 34 of the MYH7 gene. Neuromuscular disorders : NMD. 2016;26:598-603.

[34] Cecconi M, Parodi MI, Formisano F, Spirito P, Autore C, Musumeci MB, et al. Targeted next-generation sequencing helps to decipher the genetic and phenotypic heterogeneity of hypertrophic cardiomyopathy. International journal of molecular medicine. 2016;38:1111-24.

[35] Postma AV, van Engelen K, van de Meerakker J, Rahman T, Probst S, Baars MJ, et al. Mutations in the sarcomere gene MYH7 in Ebstein anomaly. Circulation Cardiovascular genetics. 2011;4:43-50.

[36] Wang J, Wang Y, Zou Y, Sun K, Wang Z, Ding H, et al. Malignant effects of multiple rare variants in sarcomere genes on the prognosis of patients with hypertrophic cardiomyopathy. European journal of heart failure. 2014;16:950-7.

[37] Clarke NF, Amburgey K, Teener J, Camelo-Piragua S, Kesari A, Punetha J, et al. A novel mutation expands the genetic and clinical spectrum of MYH7-related myopathies. Neuromuscular disorders : NMD. 2013;23:432-6.

[38] Helms AS, Davis FM, Coleman D, Bartolone SN, Glazier AA, Pagani F, et al. Sarcomere mutation-specific expression patterns in human hypertrophic cardiomyopathy. Circulation Cardiovascular genetics. 2014;7:434-43.

[39] Hao Y, Miller MS, Swank DM, Liu H, Bernstein SI, Maughan DW, et al. Passive stiffness in Drosophila indirect flight muscle reduced by disrupting paramyosin phosphorylation, but not by embryonic myosin S2 hinge substitution. Biophysical journal. 2006;91:4500-6.

[40] Liu H, Miller MS, Swank DM, Kronert WA, Maughan DW, Bernstein SI. Paramyosin phosphorylation site disruption affects indirect flight muscle stiffness and power generation in Drosophila melanogaster. Proceedings of the National Academy of Sciences of the United States of America. 2005;102:10522-7.

[41] Epstein HF, Miller DM, 3rd, Ortiz I, Berliner GC. Myosin and paramyosin are organized about a newly identified core structure. The Journal of cell biology. 1985;100:904-15.

[42] Hooper SL, Hobbs KH, Thuma JB. Invertebrate muscles: thin and thick filament structure; molecular basis of contraction and its regulation, catch and asynchronous muscle. Progress in neurobiology. 2008;86:72-127.

[43] Levine RJ, Elfvin M, Dewey MM, Walcott B. Paramyosin in invertebrate muscles. II. Content in relation to structure and function. The Journal of cell biology. 1976;71:273-9.

[44] Contompasis JL, Nyland LR, Maughan DW, Vigoreaux JO. Flightin is necessary for length determination, structural integrity, and large bending stiffness of insect flight muscle thick filaments. Journal of molecular biology. 2010;395:340-8.

[45] Orfanos Z, Sparrow JC. Myosin isoform switching during assembly of the Drosophila flight muscle thick filament lattice. Journal of cell science. 2013;126:139-48.

[46] Miyamoto CA, Fischman DA, Reinach FC. The interface between MyBP-C and myosin: site-directed mutagenesis of the CX myosin-binding domain of MyBP-C. Journal of muscle research and cell motility. 1999;20:703-15.

[47] Tama F, Feig M, Liu J, Brooks CL, 3rd, Taylor KA. The requirement for mechanical coupling between head and S2 domains in smooth muscle myosin ATPase regulation and its implications for dimeric motor function. Journal of molecular biology. 2005;345:837-54.

[48] Lee KH, Sulbaran G, Yang S, Mun JY, Alamo L, Pinto A, et al. Interacting-heads motif has been conserved as a mechanism of myosin II inhibition since before the origin of animals. Proceedings of the National Academy of Sciences of the United States of America. 2018;115:E1991-e2000.

[49] Skubiszak L. Geometrical conditions indispensable for muscle contraction. International Journal of Molecular Sciences. 2011;12:2138-57.

[50] Poglazov BF, Samokhin GP, Klibanov AM, Levitsky DI, Martinek K, Berezin IV. The effect of mechanical stretching of the myosin rod component (fragment LMMMM S-2) on the ATPase activity of myosin. Biochimica et biophysica acta. 1978;524:245-53.

[51] Iwamoto H, Trombitás K, Yagi N, Suggs JA, Bernstein SI. X-ray diffraction from flight muscle with a headless myosin mutation: implications for interpreting reflection patterns. Frontiers in physiology. 2014;5.

[52] Squire JM, Bekyarova T, Farman G, Gore D, Rajkumar G, Knupp C, et al. The myosin filament superlattice in the flight muscles of flies: A-band lattice optimisation for stretch-activation? Journal of molecular biology. 2006;361:823-38.

[53] Josephson RK, Malamud JG, Stokes DR. Asynchronous muscle: a primer. The Journal of experimental biology. 2000;203:2713-22.

[54] Iwamoto H, Yagi N. The Molecular Trigger for High-Speed Wing Beats in a Bee. Science. 2013;341:1243-6.

[55] Zhao C, Swank DM. The Drosophila indirect flight muscle myosin heavy chain isoform is insufficient to transform the jump muscle into a highly stretch-activated muscle type. American journal of physiology Cell physiology. 2017;312:C111-c8.

[56] Nishikawa K. Eccentric contraction: unraveling mechanisms of force enhancement and energy conservation. The Journal of experimental biology. 2016;219:189-96.

[57] Eckels EC, Tapia-Rojo R, Rivas-Pardo JA, Fernandez JM. The Work of Titin Protein Folding as a Major Driver in Muscle Contraction. Annual review of physiology. 2018;80:327-51.

[58] Colombini B, Nocella M, Bagni MA. Non-crossbridge stiffness in active muscle fibres. The Journal of experimental biology. 2016;219:153-60.

[59] Li Y, Lang P, Linke WA. Titin stiffness modifies the force-generating region of muscle sarcomeres. Scientific reports. 2016;6:24492.

[60] Pinniger GJ, Ranatunga KW, Offer GW. Crossbridge and non-crossbridge contributions to tension in lengthening rat muscle: force-induced reversal of the power stroke. The Journal of physiology. 2006;573:627-43.

[61] Nishikawa KC, Monroy JA, Uyeno TE, Yeo SH, Pai DK, Lindstedt SL. Is titin a 'winding filament'? A new twist on muscle contraction. Proceedings Biological sciences. 2012;279:981-90.

[62] Vemuri R, Lankford EB, Poetter K, Hassanzadeh S, Takeda K, Yu ZX, et al. The stretch-activation response may be critical to the proper functioning of the mammalian heart. Proceedings of the National Academy of Sciences of the United States of America. 1999;96:1048-53.

[63] Barry DT, Cole NM. Muscle sounds are emitted at the resonant frequencies of skeletal muscle. IEEE Trans Biomed Eng. 1990;37:525-31.

[64] Upadhye V, Agashe S. Effect of Temperature and Pressure Variations on the Resonant Frequency of Piezoelectric Material. Measurement and Control. 2016;49:286-92.

[65] Daneshparvar N, Taylor DW, O'Leary TS, Rahmani H, Yeganeh FA, Previs MJ, et al. CryoEM Structure of Drosophila Flight Muscle Thick Filaments at 7Å Resolution. bioRxiv. 2020:2020.06.05.136580.

[66] Thomas N, Thornhill RA. Stretch activation and nonlinear elasticity of muscle cross-bridges. Biophysical journal. 1996;70:2807-18.

[67] Smith DA, Geeves MA, Sleep J, Mijailovich SM. Towards a unified theory of muscle contraction. I: foundations. Annals of biomedical engineering. 2008;36:1624-40.

[68] Weitkunat M, Brasse M, Bausch AR, Schnorrer F. Mechanical tension and spontaneous muscle twitching precede the formation of cross-striated muscle in vivo. Development (Cambridge, England). 2017;144:1261-72.

[69] Nongthomba U, Cummins M, Clark S, Vigoreaux JO, Sparrow JC. Suppression of muscle hypercontraction by mutations in the myosin heavy chain gene of Drosophila melanogaster. Genetics. 2003;164:209-22.

[70] Vigoreaux JO, Perry LM. Multiple isoelectric variants of flightin in Drosophila stretch-activated muscles are generated by temporally regulated phosphorylations. Journal of muscle research and cell motility. 1994;15:607-16.

[71] Obermann WM, Gautel M, Steiner F, van der Ven PF, Weber K, Furst DO. The structure of the sarcomeric M band: localization of defined domains of myomesin, M-protein, and the 250-kD carboxy-terminal region of titin by immunoelectron microscopy. The Journal of cell biology. 1996;134:1441-53.

[72] Agarkova I, Auerbach D, Ehler E, Perriard JC. A novel marker for vertebrate embryonic heart, the EH-myomesin isoform. The Journal of biological chemistry. 2000;275:10256-64.

CONCLUDING REMARKS

This work introduces fresh insight into how flightin is intimately involved in architectural design of the thick filament through its novel WYR domain and how this may enable flightin's role in muscle structure, mechanics, and function. Invertebrates and vertebrates approach adaptation of the structural-mechanical framework within muscle differently and comparative studies on the level of the myosin LMM and myosin binding proteins suggest different methods of modulation while converging on a common purpose: to produce a resilient and effective work-producing system that scales biochemical to mechanical transitions. The structure of WYR identified in this study evokes inquiry into the role of aromatic residues within the thick filament. WYR may be an optimal model for how resonant behavior and capacity on the level of electron transfer can be employed in muscle operation. As the between-layer contacts place separate flightin densities proximal to each other and angled such that the contact points wrap around the hollow thick filament interior tangential to the thick filament axis, while propagating along it, WYR is positioned such that it potentially forms a coil within the thick filament core; could the WYR aromatic residues be aligned and have an energetic role? A growing understanding of quantum mechanics in biological systems means that information regarding domains reliant on aromatic structures, such as WYR, may increase in value substantially as research in this area matures.

A number of next steps come to mind when considering the potential analogous roles of cMyBP-C and flightin in impacting thick filament dimensions. MyBP-C is known to impact paracrystal formation of myosin in low ionic strength solutions though its destabilization of paracrystal periodicity alludes to its *in vivo* periodicity within the thick filament being dictated later on in the myosin assembly [1]. Paracrystal formation could be examined in the presence of flightin or WYR to evaluate the impact upon paracrystal dimensions and even mechanical properties, such as flexural rigidity of the assembly. The potential analogous roles between the two proteins also makes flightin's participation in a segmented connective process that runs from the thin filament to the core of the thick filament, and then to the M-line an enticing prospect. The participation of cMyBP-C and flightin in corresponding connective systems may explain how the functions of either protein become incorporated and realized throughout the sarcomere and relate to their potential analogous roles. Binding studies of flightin to stretchin and paramyosin or examining the myosin:paramyosin relationship both with and without flightin or with residues at the WYR-LMM multiface would more concretely expose the relationship between these elements and examine whether or not such a connective process involving these components is likely to exist.

Both flightin's role within muscle architecture and the specific nature of its conserved WYR domain may have far-reaching significance. If a process, involving flightin and paramyosin, that tunes the structural and mechanical properties of the thick filament/LMM associations is identified, this reveals a novel ingenuity of IFM systems

and lead to a greater understanding of invertebrate muscle function on a broader scale. Beyond its relevance to evolutionary study, such a strategy may be able to be employed in treatments for structural and mechanical deficits in muscle pathologies. Dissection of the role of the highly aromatic WYR domain is likely to expand our understanding of the larger role of aromatic residues and how they engage with each other or with other binding partners. This would serve to build our understanding on how ordering and alignment of resonant structures initiate or enable energetic conversion and relay. As a high proportion of pharmaceuticals are designed to mimic tyrosine or tryptophan-derived natural molecules (e.g. dopamine, norepinephrine, serotonin) and many more that are created *de novo* rely on aromatic structures [2], that the study of WYR may lead to new drug therapies or a greater understanding of the mechanism behind current effective treatment strategies is within reason.

In the near term, more CD and NMR studies could add greatly to our understanding of WYR structure and its effects on the LMM. The proposed 'cinch' function of WYR as pertains to stability can be evaluated in the melting of the LMM with and without the presence of WYR, observed by CD. The relevance of the high aromatic content in WYR necessitates the use of NMR to elucidate their orientation. Examination of its behavior by NMR in the context of various Ca++ and phosphate levels may lead to a realization of its *in vivo* operation. This could further be coupled with mechanical testing by AFM of myosin paracrystals that incorporate WYR. WYR interaction with myosin dimers in the context of strain may also be able to be performed by using centrifugal forces for control

of force applied with dimers or formed paracrystals bound to a grid or membrane layer. To this end, thick filaments from transgenic lines expressing WYR in the absence of flightin, could also be examined by AFM with the advantage of being able to couple such an examination with structural characterization by TEM.

Considerations for the development of transgenic lines

Examination of flightin function in the context of WYR structure and its connection to the LMM *in vivo* is a necessary effort that can greatly benefit from *in vitro* dissection of the WYR structure both alone and with its binding partner. Peptides of mutated WYR can be trialed by the same CD and cosedimention assays described in Chapter 3 and coupled with biolayer interferometry, mass photometry, or many other assays. Analysis of mutated WYR could first determine whether a change impacts the WYR structure, or alters the behavior in solution (e.g. solubility). In the context of the LMM, the mutated WYR can be examined to see whether such changes impact parameters of LMM engagement (e.g. stoichiometry, Km, coiled-coil propensity) or abolish binding altogether. Depending on the question being asked, different WYR mutants would be desired.

If the role of WYR in its specific structural state is of primary interest, the ideal residues to mutate would be ones considered to be most imperative to the antiparallel β character. As β hairpins are heavily reliant on their turns, it is expected that the residues involved

would be highly conserved- as we see for N102-Y104- and these residues come to mind above all others. Avoiding the absolutely conserved residues has value in that their role may have both binding and structural consequences. The pair of tyrosines may be involved in a turn but also integral for solubility or a conformational change upon contact to a binding partner. To target the relevance of structure, it would be prudent to mutate, firstly, a residue expected to be the structure's linchpin. As an ASX turn would be reliant on the N102 sidegroup H-bonding to the mainchain specifically, decreasing its likelihood to external engagement, N102 makes a very attractive position for the purpose of examining WYR structure and how the resultant structure impacts its relationship to the LMM. Glutamine is capable of H-bonding to the main-chain but at angles that are not conducive to ASX turn formation in a β hairpin, making N102Q an informative choice. Additionally, if Y104 is involved in backbone H-bonding to N102 and is otherwise important for binding or a secondary conformational shift, it would be expected that a Y104T mutation would not impact the ASX turn formation. This could be a possible lead-in to separating residues imperative for initial secondary structure, for binding to the LMM itself and for conformational shift important after binding.

A change that is not likely to impact the β hairpin formation or WYR solubility but may be very impactful for WYR-WYR and WYR-LMM interactions is altering one, or more, of the tyrosines present in the first β strand of the hypothetical WYR structure. Y95 may be acting in cross-tyrosine contact but all the tyrosines from Y93-Y99 may be engaging

with each other in a cumulative manner. Mutating the absolutely conserved Y93 residue, Y95 or all of the tyrosines in the strand (Y93, Y95, Y97, Y99) to threonine would be permissive to an aqueous-facing strand incapable of ring-associated interactions. If selecting only one, destabilizing a central tyrosine may be most informative (e.g. Y95T).

Focusing on the β hairpin allows us a more defined area for examination than dissection of potential Trp-Pro interactions that likely drive P_2 or other aperiodic structure but this can still be examined by proline to histidine substitutions. The absolutely conserved P123 may be engaged with W128 or W85, or both, depending on conditions (initial fold vs. fold upon binding). The conservative conversion to histidine would be expected to change this relationship and the P_2 content in WYR without impacting the β hairpin and behavior of the tyrosines. As this pertains to the aperiodic content, this would be best employed when use of NMR with WYR has been optimized.

Once point mutations in WYR have been characterized along with their coordinate LMM-binding behavior, residues that appear to impact binding stoichiometry or Kd without changing the structure would be of interest for examining LMM-LMM associations in the context of WYR. A different or weaker WYR-LMM relationship may translate into a shift in the relationship of myosin rods to each other in the thick filament and alter thick filament dimensions, stability, or stiffness. Both WT and mutant WYR peptides would benefit from experiments including the flightin N- or C-terminus as well,

since an association with either component may be important for proper incorporation into the thick filament *in vivo* or self-associations before, or after, incorporation.

Though pursuit of more nuanced dissection of WYR is desirable, exclusive expression of WYR as-is or flightin without WYR should not be discounted. A transgenic line exclusively expressing WYR may be able to determine whether WYR is sufficient *in vivo* for thick filament incorporation and a line expressing both the N- and C-terminal portions of flightin, with the exception of WYR, may be able to reveal if the WYR region is necessary for the association of flightin to the thick filament.

Future research of flightin's role in muscle and WYR's performance on the molecular level is sure to expand our understanding of the mechanistic ingenuity employed within the insect IFM and possibly related to the conservation of WYR throughout Pancrustacea. The perseverance of WYR throughout evolutionary history solidifies its contemporary relevance and permits us an opportunity to understand, build, incorporate, and advance our utilization of such natural invention that spans, and unites the fields of medicine, biology, chemistry, and physics.

1. Moos, C., et al., *Interaction of C-protein with myosin, myosin rod and light meromyosin.* Journal of Molecular Biology, 1975. **97**(1): p. 1-9.
2. Polêto, M.D., et al., *Aromatic Rings Commonly Used in Medicinal Chemistry: Force Fields Comparison and Interactions With Water Toward the Design of New Chemical Entities.* Frontiers in Pharmacology, 2018. **9**(395).

COMPREHENSIVE BIBLIOGRAPHY

[1] Achal M, Trujillo AS, Melkani GC, Farman GP, Ocorr K, Viswanathan MC, et al. A Restrictive Cardiomyopathy Mutation in an Invariant Proline at the Myosin Head/Rod Junction Enhances Head Flexibility and Function, Yielding Muscle Defects in Drosophila. Journal of molecular biology. 2016;428:2446-61.

[2] Agarkova I, Auerbach D, Ehler E, Perriard JC. A novel marker for vertebrate embryonic heart, the EH-myomesin isoform. The Journal of biological chemistry. 2000;275:10256-64.

[3] Alamo L, Koubassova N, Pinto A, Gillilan R, Tsaturyan A, Padron R. Lessons from a tarantula: new insights into muscle thick filament and myosin interacting-heads motif structure and function. Biophysical reviews. 2017;9:461-80.

[4] Alamo L, Qi D, Wriggers W, Pinto A, Zhu J, Bilbao A, et al. Conserved Intramolecular Interactions Maintain Myosin Interacting-Heads Motifs Explaining Tarantula Muscle Super-Relaxed State Structural Basis. Journal of molecular biology. 2016;428:1142-64.

[5] Al-Khayat HA. Three-dimensional structure of the human myosin thick filament: clinical implications. Global cardiology science & practice. 2013;2013:280-302.

[6] Anand BG, Prajapati KP, Shekhawat DS, Kar K. Tyrosine-Generated Nanostructures Initiate Amyloid Cross-Seeding in Proteins Leading to a Lethal Aggregation Trap. Biochemistry. 2018;57:5202-9.

[7] Anjana R, Vaishnavi MK, Sherlin D, Kumar SP, Naveen K, Kanth PS, et al. Aromatic-aromatic interactions in structures of proteins and protein-DNA complexes: a study based on orientation and distance. Bioinformation. 2012;8:1220-4.

[8] Antosiewicz JM, Shugar D. UV-Vis spectroscopy of tyrosine side-groups in studies of protein structure. Part 1: basic principles and properties of tyrosine chromophore. Biophysical reviews. 2016;8:151-61.

[9] Apostolovic B, Danial M, Klok HA. Coiled coils: attractive protein folding motifs for the fabrication of self-assembled, responsive and bioactive materials. Chemical Society reviews. 2010;39:3541-75.

[10] Armel TZ, Leinwand LA. Mutations in the beta-myosin rod cause myosin storage myopathy via multiple mechanisms. Proceedings of the National Academy of Sciences of the United States of America. 2009;106:6291-6.

[11] Armel TZ, Leinwand LA. A mutation in the beta-myosin rod associated with hypertrophic cardiomyopathy has an unexpected molecular phenotype. Biochemical and biophysical research communications. 2010;391:352-6.

[12] Ashok Kumar T. CFSSP: Chou and Fasman Secondary Structure Prediction server. 2013.

[13] Ayer G, Vigoreaux JO. Flightin is a myosin rod binding protein. Cell biochemistry and biophysics. 2003;38:41-54.

[14] Ayme-Southgate A, Saide J, Southgate R, Bounaix C, Cammarato A, Patel S, et al. In indirect flight muscles Drosophila projectin has a short PEVK domain, and its NH2-terminus is embedded at the Z-band. Journal of muscle research and cell motility. 2005;26:467-77.

[15] Ayme-Southgate AJ, Turner L, Southgate RJ. Molecular analysis of the muscle protein projectin in Lepidoptera. Journal of insect science (Online). 2013;13:88.

[16] Baghdadi MB, Tajbakhsh S. Regulation and phylogeny of skeletal muscle regeneration. Developmental biology. 2018;433:200-9.

[17] Baguna J, Martinez P, Paps J, Riutort M. Back in time: a new systematic proposal for the Bilateria. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2008;363:1481-91.

[18] Ball JB, Hughes RA, Alewood PF, Andrews PR. β-turn topography. Tetrahedron. 1993;49:3467-78.

[19] Barry DT, Cole NM. Muscle sounds are emitted at the resonant frequencies of skeletal muscle. IEEE Trans Biomed Eng. 1990;37:525-31.

[20] Bartlett AI, Radford SE. An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. Nature structural & molecular biology. 2009;16:582-8.

[21] Barton B, Ayer G, Heymann N, Maughan DW, Lehmann FO, Vigoreaux JO. Flight muscle properties and aerodynamic performance of Drosophila expressing a flightin transgene. The Journal of experimental biology. 2005;208:549-60.

[22] Barton B, Ayer G, Maughan DW, Vigoreaux JO. Site directed mutagenesis of Drosophila flightin disrupts phosphorylation and impairs flight muscle structure and mechanics. Journal of muscle research and cell motility. 2007;28:219-30.

[23] Barton B, Vigoreaux J. Novel Myosin Associated Proteins. Nature's Versatile Engine: Insect Flight Muscle Inside and Out. Boston, MA: Springer US; 2006. p. 86-96.

[24] Baslé E, Joubert N, Pucheault M. Protein Chemical Modification on Endogenous Amino Acids. Chemistry & Biology. 2010;17:213-27.

[25] Basu K, Wasserman SS, Jeronimo PS, Graham LA, Davies PL. Intermediate activity of midge antifreeze protein is due to a tyrosine-rich ice-binding site and atypical ice plane affinity. The FEBS Journal. 2016;283:1504-15.

[26] Beall CJ, Sepanski MA, Fyrberg EA. Genetic dissection of Drosophila myofibril formation: effects of actin and myosin heavy chain null alleles. Genes & development. 1989;3:131-40.

[27] Becker KD, O'Donnell PT, Heitz JM, Vito M, Bernstein SI. Analysis of Drosophila paramyosin: identification of a novel isoform which is restricted to a subset of adult muscles. The Journal of cell biology. 1992;116:669-81.

[28] Bennett PM. The structure of spindle-shaped paracrystals of light meromyosin. Journal of molecular biology. 1981;146:201-21.

[29] Bhattacharjee N, Biswas P. Position-specific propensities of amino acids in the β-strand. BMC Struct Biol. 2010;10:29.

[30] Biancalana M, Makabe K, Koide A, Koide S. Aromatic cross-strand ladders control the structure and stability of beta-rich peptide self-assembly mimics. Journal of molecular biology. 2008;383:205-13.

[31] Biancalana M, Makabe K, Koide A, Koide S. Molecular mechanism of thioflavin-T binding to the surface of beta-rich peptide self-assemblies. Journal of molecular biology. 2009;385:1052-63.

[32] Bibi F, Villain M, Guillaume C, Sorli B, Gontard N. A Review: Origins of the Dielectric Properties of Proteins and Potential Development as Bio-Sensors. Sensors (Basel, Switzerland). 2016;16.

[33] Blanco F, Ramirez-Alvarado M, Serrano L. Formation and stability of beta-hairpin structures in polypeptides. Current opinion in structural biology. 1998;8:107-11.

[34] Blondelle J, Marrocco V, Clark M, Desmond P, Myers S, Nguyen J, et al. Murine obscurin and Obsl1 have functionally redundant roles in sarcolemmal integrity, sarcoplasmic reticulum organization, and muscle metabolism. Communications Biology. 2019;2:178.

[35] Borisov AB, Sutter SB, Kontrogianni-Konstantopoulos A, Bloch RJ, Westfall MV, Russell MW. Essential role of obscurin in cardiac myofibrillogenesis and hypertrophic response: evidence from small interfering RNA-mediated gene silencing. Histochemistry and cell biology. 2006;125:227-38.

[36] Briggs M. Densitometry using ImageJ. 2013.

[37] Brown JH, Cohen C, Parry DA. Heptad breaks in alpha-helical coiled coils: stutters and stammers. Proteins. 1996;26:134-45.

[38] Buchan DWA, Jones DT. The PSIPRED Protein Analysis Workbench: 20 years on. Nucleic Acids Research. 2019;47:W402-W7.

[39] Bullard B, Burkart C, Labeit S, Leonard K. The function of elastic proteins in the oscillatory contraction of insect flight muscle. Journal of muscle research and cell motility. 2005;26:479-85.

[40] Burkart C, Qiu F, Brendel S, Benes V, Haag P, Labeit S, et al. Modular proteins from the Drosophila sallimus (sls) gene and their expression in muscles with different extensibility. Journal of molecular biology. 2007;367:953-69.

[41] Burton PM. Insights from diploblasts; the evolution of mesoderm and muscle. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution. 2008;310B:5-14.

[42] Bush CA, Sarkar SK, Kopple KD. Circular dichroism of β turns in peptides and proteins. Biochemistry. 1978;17:4951-4.

[43] Cecconi M, Parodi MI, Formisano F, Spirito P, Autore C, Musumeci MB, et al. Targeted next-generation sequencing helps to decipher the genetic and phenotypic heterogeneity of hypertrophic cardiomyopathy. International journal of molecular medicine. 2016;38:1111-24.

[44] Chakravarty S, Ung AR, Moore B, Shore J, Alshamrani M. A Comprehensive Analysis of Anion-Quadrupole Interactions in Protein Structures. Biochemistry. 2018;57:1852-67.

[45] Chakravorty S. Role of the Drosophila Melanogaster Indirect Flight Muscles in Flight and Male Courtship Song: Studies on Flightin and Mydson Light Chain - 2. Graduate College Dissertations and Theses: University of Vermont; 2013.

[46] Chakravorty S, Tanner BCW, Foelber VL, Vu H, Rosenthal M, Ruiz T, et al. Flightin maintains myofilament lattice organization required for optimal flight power and courtship song quality in Drosophila. Proceedings Biological sciences. 2017;284.

[47] Chakravorty S, Vu H, Foelber V, Vigoreaux JO. Mutations of the Drosophila myosin regulatory light chain affect courtship song and reduce reproductive success. PloS one. 2014;9:e90077.

[48] Chakravorty S, Wajda MP, Vigoreaux JO. Courtship song analysis of Drosophila muscle mutants. Methods (San Diego, Calif). 2012;56:87-94.

[49] Chavez JD, Liu NL, Bruce JE. Quantification of protein-protein interactions with chemical cross-linking and mass spectrometry. Journal of proteome research. 2011;10:1528-37.

[50] Chechenova MB, Bryantsev AL, Cripps RM. The Drosophila Z-disc protein Z(210) is an adult muscle isoform of Zasp52, which is required for normal myofibril organization in indirect flight muscles. The Journal of biological chemistry. 2013;288:3718-26.

[51] Chou K-C. Prediction and classification of α-turn types. Biopolymers. 1997;42:837-53.

[52] Chou K-C, Liu W-M, Maggiora GM, Zhang C-T. Prediction and classification of domain structural classes. Proteins: Structure, Function, and Bioinformatics. 1998;31:97-103.

[53] Chu M, Gregorio CC, Pappas CT. Nebulin, a multi-functional giant. The Journal of experimental biology. 2016;219:146-52.

[54] Clark KA, Bland JM, Beckerle MC. The Drosophila muscle LIM protein, Mlp84B, cooperates with D-titin to maintain muscle structural integrity. Journal of cell science. 2007;120:2066-77.

[55] Clark KA, Kadrmas JL. Drosophila melanogaster muscle LIM protein and alpha-actinin function together to stabilize muscle cytoarchitecture: a potential role for Mlp84B in actin-crosslinking. Cytoskeleton (Hoboken, NJ). 2013;70:304-16.

[56] Clark KA, Lesage-Horton H, Zhao C, Beckerle MC, Swank DM. Deletion of Drosophila muscle LIM protein decreases flight muscle stiffness and power generation. American journal of physiology Cell physiology. 2011;301:C373-82.

[57] Clarke NF, Amburgey K, Teener J, Camelo-Piragua S, Kesari A, Punetha J, et al. A novel mutation expands the genetic and clinical spectrum of MYH7-related myopathies. Neuromuscular disorders : NMD. 2013;23:432-6.

[58] Collier VL, Kronert WA, O'Donnell PT, Edwards KA, Bernstein SI. Alternative myosin hinge regions are utilized in a tissue-specific fashion that correlates with muscle contraction speed. Genes & development. 1990;4:885-95.

[59] Colombini B, Nocella M, Bagni MA. Non-crossbridge stiffness in active muscle fibres. The Journal of experimental biology. 2016;219:153-60.

[60] Combet C, Blanchet C, Geourjon C, Deleage G. NPS@: network protein sequence analysis. Trends in biochemical sciences. 2000;25:147-50.

[61] Contompasis JL, Nyland LR, Maughan DW, Vigoreaux JO. Flightin is necessary for length determination, structural integrity, and large bending stiffness of insect flight muscle thick filaments. Journal of molecular biology. 2010;395:340-8.

[62] Coulton AT, Stelzer JE. Cardiac myosin binding protein C and its phosphorylation regulate multiple steps in the cross-bridge cycle of muscle contraction. Biochemistry. 2012;51:3292-301.

[63] Craig R, Woodhead JL. Structure and function of myosin filaments. Current opinion in structural biology. 2006;16:204-12.

[64] Craveur P, Joseph AP, Rebehmed J, de Brevern AG. β-Bulges: Extensive structural analyses of β-sheets irregularities. Protein Science. 2013;22:1366-78.

[65] Cripps RM, Suggs JA, Bernstein SI. Assembly of thick filaments and myofibrils occurs in the absence of the myosin head. The EMBO journal. 1999;18:1793-804.

[66] Crisma M, De Zotti M, Moretto A, Peggion C, Drouillat B, Wright K, et al. Single and multiple peptide γ-turns: literature survey and recent progress. New Journal of Chemistry. 2015;39:3208-16.

[67] Cuff JA, Barton GJ. Application of multiple sequence alignment profiles to improve protein secondary structure prediction. Proteins. 2000;40:502-11.

[68] Cuff JA, Clamp ME, Siddiqui AS, Finlay M, Barton GJ. JPred: a consensus secondary structure prediction server. Bioinformatics. 1998;14:892-3.

[69] Daneshparvar N, Taylor DW, O'Leary TS, Rahmani H, Yeganeh FA, Previs MJ, et al. CryoEM Structure of Drosophila Flight Muscle Thick Filaments at 7Å Resolution. bioRxiv. 2020:2020.06.05.136580.

[70] Dasgupta B, Chakrabarti P. pi-Turns: types, systematics and the context of their occurrence in protein structures. BMC Struct Biol. 2008;8:39.

[71] Dasgupta B, Pal L, Basu G, Chakrabarti P. Expanded turn conformations: Characterization and sequence-structure correspondence in α-turns with implications in helix folding. Proteins: Structure, Function, and Bioinformatics. 2004;55:305-15.

[72] de Alba E, Rico M, Jiménez MA. Cross-strand side-chain interactions versus turn conformation in beta-hairpins. Protein Science : A Publication of the Protein Society. 1997;6:2548-60.

[73] de Brevern AG. Extension of the classical classification of β-turns. Scientific reports. 2016;6:33191.

[74] Della Valle E, Marracino P, Pakhomova O, Liberti M, Apollonio F. Nanosecond pulsed electric signals can affect electrostatic environment of proteins below the threshold of conformational effects: The case study of SOD1 with a molecular simulation study. PloS one. 2019;14:e0221685.

[75] Deora T, Gundiah N, Sane SP. Mechanics of the thorax in flies. The Journal of experimental biology. 2017;220:1382-95.

[76] Ding Y, Li Y, Qin M, Cao Y, Wang W. Photo-Cross-Linking Approach to Engineering Small Tyrosine-Containing Peptide Hydrogels with Enhanced Mechanical Stability. Langmuir. 2013;29:13299-306.

[77] Doering DS, Matsudaira P. Cysteine Scanning Mutagenesis at 40 of 76 Positions in Villin Headpiece Maps the F-Actin Binding Site and Structural Features of the Domain. Biochemistry. 1996;35:12677-85.

[78] Domsch K, Ezzeddine N, Nguyen HT. Abba is an essential TRIM/RBCC protein to maintain the integrity of sarcomeric cytoarchitecture. Journal of cell science. 2013;126:3314-23.

[79] Dougherty DA. Cation-π Interactions Involving Aromatic Amino Acids. The Journal of Nutrition. 2007;137:1504S-8S.

[80] Du D, Zhu Y, Huang C-Y, Gai F. Understanding the key factors that control the rate of β-hairpin folding. Proceedings of the National Academy of Sciences of the United States of America. 2004;101:15915-20.

[81] Du QS, Chen D, Xie NZ, Huang RB, Chou KC. Insight into a molecular interaction force supporting peptide backbones and its implication to protein loops and folding. Journal of biomolecular structure & dynamics. 2015;33:1957-72.

[82] Eckels EC, Tapia-Rojo R, Rivas-Pardo JA, Fernandez JM. The Work of Titin Protein Folding as a Major Driver in Muscle Contraction. Annual review of physiology. 2018;80:327-51.

[83] Eckhardt B, Grosse W, Essen LO, Geyer A. Structural characterization of a beta-turn mimic within a protein-protein interface. Proceedings of the National Academy of Sciences of the United States of America. 2010;107:18336-41.

[84] Efimov AV. Structure of coiled β-β-hairpins and β-β-corners. FEBS Letters. 1991;284:288-92.

[85] Emerson CP, Jr., Bernstein SI. Molecular genetics of myosin. Annual review of biochemistry. 1987;56:695-726.

[86] Engel DGMS. Evolution of the Insects. 1 ed: Cambridge University Press; 2005.

[87] Epstein HF, Miller DM, 3rd, Ortiz I, Berliner GC. Myosin and paramyosin are organized about a newly identified core structure. The Journal of cell biology. 1985;100:904-15.

[88] Eswar N, Ramakrishnan C. Secondary structures without backbone: an analysis of backbone mimicry by polar side chains in protein structures. Protein Engineering, Design and Selection. 1999;12:447-55.

[89] Ferbert A, Zibat A, Rautenstrauss B, Kress W, Hugens-Penzel M, Weis J, et al. Laing distal myopathy with a novel mutation in exon 34 of the MYH7 gene. Neuromuscular disorders : NMD. 2016;26:598-603.

[90] Ferguson C, Lakey A, Hutchings A, Butcher GW, Leonard KR, Bullard B. Cytoskeletal proteins of insect muscle: location of zeelins in Lethocerus flight and leg muscle. Journal of cell science. 1994;107:1115-29.

[91] Fernandes I, Schock F. The nebulin repeat protein Lasp regulates I-band architecture and filament spacing in myofibrils. The Journal of cell biology. 2014;206:559-72.

[92] Fetrow JS. Omega loops: nonregular secondary structures significant in protein function and stability. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 1995;9:708-17.

[93] Firdaus H, Mohan J, Naz S, Arathi P, Ramesh SR, Nongthomba U. A cis-regulatory mutation in troponin-I of Drosophila reveals the importance of proper stoichiometry of structural proteins during muscle assembly. Genetics. 2015;200:149-65.

[94] Flashman E, Watkins H, Redwood C. Localization of the binding site of the C-terminal domain of cardiac myosin-binding protein-C on the myosin rod. The Biochemical journal. 2007;401:97-102.

[95] Flocco MM, Mowbray SL. Planar Stacking Interactions of Arginine and Aromatic Side-Chains in Proteins. Journal of molecular biology. 1994;235:709-17.

[96] Fornander LH, Feng B, Beke-Somfai T, Nordén B. UV Transition Moments of Tyrosine. The Journal of Physical Chemistry B. 2014;118:9247-57.

[97] Franklin MW, Slusky JSG. Tight Turns of Outer Membrane Proteins: An Analysis of Sequence, Structure, and Hydrogen Bonding. Journal of molecular biology. 2018;430:3251-65.

[98] Gallivan JP, Dougherty DA. Cation-π interactions in structural biology. Proceedings of the National Academy of Sciences. 1999;96:9459-64.

[99] Gans PJ, Lyu PC, Manning MC, Woody RW, Kallenbach NR. The helix-coil transition in heterogeneous peptides with specific side-chain interactions: theory and comparison with CD spectral data. Biopolymers. 1991;31:1605-14.

[100] Garnham CP, Campbell RL, Walker VK, Davies PL. Novel dimeric β-helical model of an ice nucleation protein with bridged active sites. BMC Structural Biology. 2011;11:36.

[101] Gasek NS, Nyland LR, Vigoreaux JO. The Contributions of the Amino and Carboxy Terminal Domains of Flightin to the Biomechanical Properties of Drosophila Flight Muscle Thick Filaments. Biology. 2016;5.

[102] Gautel M, Djinovic-Carugo K. The sarcomeric cytoskeleton: from molecules to motion. The Journal of experimental biology. 2016;219:135-45.

[103] George EL, Ober MB, Emerson CP, Jr. Functional domains of the Drosophila melanogaster muscle myosin heavy-chain gene are encoded by alternatively spliced exons. Molecular and cellular biology. 1989;9:2957-74.

[104] Ghassemizadeh R, Moore B, Momose T, Walter M. Stability and IR Spectroscopy of Zwitterionic Form of β-Alanine in Water Clusters. The Journal of Physical Chemistry B. 2019;123:4392-9.

[105] Ghisaidoobe AB, Chung SJ. Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Forster resonance energy transfer techniques. International Journal of Molecular Sciences. 2014;15:22518-38.

[106] Greenfield NJ. Using circular dichroism spectra to estimate protein secondary structure. Nature protocols. 2006;1:2876-90.

[107] Greenfield NJ, Hitchcock-DeGregori SE. Conformational intermediates in the folding of a coiled-coil model peptide of the N-terminus of tropomyosin and alpha alpha-tropomyosin. Protein Science : A Publication of the Protein Society. 1993;2:1263-73.

[108] Griffiths-Jones SR, Sharman GJ, Maynard AJ, Searle MS. Modulation of intrinsic phi,psi propensities of amino acids by neighbouring residues in the coil regions of protein structures: NMR analysis and dissection of a beta-hairpin peptide. Journal of molecular biology. 1998;284:1597-609.

[109] Guharoy M, Chakrabarti P. Secondary structure based analysis and classification of biological interfaces: identification of binding motifs in protein–protein interactions. Bioinformatics. 2007;23:1909-18.

[110] Gunage RD, Dhanyasi N, Reichert H, VijayRaghavan K. Drosophila adult muscle development and regeneration. Seminars in cell & developmental biology. 2017;72:56-66.

[111] Guruprasad K, Rajkumar S. Beta-and gamma-turns in proteins revisited: a new set of amino acid turn-type dependent positional preferences and potentials. Journal of biosciences. 2000;25:143-56.

[112] Guzenko D, Strelkov SV. CCFold: rapid and accurate prediction of coiled-coil structures and application to modelling intermediate filaments. Bioinformatics. 2017;34:215-22.

[113] Hakeda S, Endo S, Saigo K. Requirements of Kettin, a giant muscle protein highly conserved in overall structure in evolution, for normal muscle function, viability, and flight activity of Drosophila. The Journal of cell biology. 2000;148:101-14.

[114] Hamill SJ, Cota E, Chothia C, Clarke J. Conservation of folding and stability within a protein family: the tyrosine corner as an evolutionary cul-de-sac. Journal of molecular biology. 2000;295:641-9.

[115] Hao Y, Miller MS, Swank DM, Liu H, Bernstein SI, Maughan DW, et al. Passive stiffness in Drosophila indirect flight muscle reduced by disrupting paramyosin phosphorylation, but not by embryonic myosin S2 hinge substitution. Biophysical journal. 2006;91:4500-6.

[116] Hartmann MD, Mendler CT, Bassler J, Karamichali I, Ridderbusch O, Lupas AN, et al. α/β coiled coils. eLife. 2016;5.

[117] Helms AS, Davis FM, Coleman D, Bartolone SN, Glazier AA, Pagani F, et al. Sarcomere mutation-specific expression patterns in human hypertrophic cardiomyopathy. Circulation Cardiovascular genetics. 2014;7:434-43.

[118] Hemmingsen JM, Gernert KM, Richardson JS, Richardson DC. The tyrosine corner: A feature of most greek key β-barrel proteins. Protein Science. 1994;3:1927-37.

[119] Henkin JA, Maughan DW, Vigoreaux JO. Mutations that affect flightin expression in Drosophila alter the viscoelastic properties of flight muscle fibers. American journal of physiology Cell physiology. 2004;286:C65-72.

[120] Hiromi Y, Okamoto H, Gehring WJ, Hotta Y. Germline transformation with Drosophila mutant actin genes induces constitutive expression of heat shock genes. Cell. 1986;44:293-301.

[121] Homyk T, Jr., Emerson CP, Jr. Functional interactions between unlinked muscle genes within haploinsufficient regions of the Drosophila genome. Genetics. 1988;119:105-21.

[122] Hooper SL, Hobbs KH, Thuma JB. Invertebrate muscles: thin and thick filament structure; molecular basis of contraction and its regulation, catch and asynchronous muscle. Progress in neurobiology. 2008;86:72-127.

[123] Hooper SL, Thuma JB. Invertebrate muscles: muscle specific genes and proteins. Physiological reviews. 2005;85:1001-60.

[124] Houmeida A, Holt J, Tskhovrebova L, Trinick J. Studies of the interaction between titin and myosin. The Journal of cell biology. 1995;131:1471-81.

[125] Hu C, Koehl P. Helix-sheet packing in proteins. Proteins: Structure, Function, and Bioinformatics. 2010;78:1736-47.

[126] Hu DH, Matsuno A, Terakado K, Matsuura T, Kimura S, Maruyama K. Projectin is an invertebrate connectin (titin): isolation from crayfish claw muscle and localization in

crayfish claw muscle and insect flight muscle. Journal of muscle research and cell motility. 1990;11:497-511.

[127] Hu LY, Ackermann MA, Kontrogianni-Konstantopoulos A. The sarcomeric M-region: a molecular command center for diverse cellular processes. BioMed research international. 2015;2015:714197.

[128] Hu Z, Taylor DW, Edwards RJ, Taylor KA. Coupling between myosin head conformation and the thick filament backbone structure. Journal of structural biology. 2017;200:334-42.

[129] Hu Z, Taylor DW, Reedy MK, Edwards RJ, Taylor KA. Structure of myosin filaments from relaxed Lethocerus flight muscle by cryo-EM at 6 A resolution. Science advances. 2016;2:e1600058.

[130] Hutchinson EG, Thornton JM. A revised set of potentials for beta-turn formation in proteins. Protein Science : A Publication of the Protein Society. 1994;3:2207-16.

[131] Irving M. Regulation of Contraction by the Thick Filaments in Skeletal Muscle. Biophysical journal. 2017;113:2579-94.

[132] Iwamoto H. Structure, function and evolution of insect flight muscle. Biophysics. 2011;7:21-8.

[133] Iwamoto H, Trombitás K, Yagi N, Suggs JA, Bernstein SI. X-ray diffraction from flight muscle with a headless myosin mutation: implications for interpreting reflection patterns. Frontiers in physiology. 2014;5.

[134] Iwamoto H, Yagi N. The Molecular Trigger for High-Speed Wing Beats in a Bee. Science. 2013;341:1243-6.

[135] J. Cooper JW, A. Mills. Chapter 2: Super-Secondary Structure. Principles of Protein Structure, Comparative Protein Modelling, and Visualisation. SWISS-MODEL: Biozentrum.

[136] Jackups R, Jr., Liang J. Interstrand pairing patterns in beta-barrel membrane proteins: the positive-outside rule, aromatic rescue, and strand registration prediction. Journal of molecular biology. 2005;354:979-93.

[137] JBergès PT, and C Houée-Levin. Oxidation of protein tyrosine or methionine residues: From the amino acid to the peptide. Journal of Physics: Conference Series. 2011;261.

[138] Johnson TA, Holyoak T. The Ω-loop lid domain of phosphoenolpyruvate carboxykinase is essential for catalytic function. Biochemistry. 2012;51:9547-59.

[139] Josephson RK, Malamud JG, Stokes DR. Asynchronous muscle: a primer. The Journal of experimental biology. 2000;203:2713-22.

[140] Juszczyk P, Kolodziejczyk AS, Grzonka Z. Circular dichroism and aggregation studies of amyloid beta (11-8) fragment and its variants. Acta biochimica Polonica. 2005;52:425-31.

[141] Kajava AV, Steven AC. β-Rolls, β-Helices, and Other β-Solenoid Proteins. Advances in protein chemistry: Academic Press; 2006. p. 55-96.

[142] Kato M, McKnight SL. Cross-beta Polymerization of Low Complexity Sequence Domains. Cold Spring Harbor perspectives in biology. 2017;9.

[143] Katzemich A, Kreiskother N, Alexandrovich A, Elliott C, Schock F, Leonard K, et al. The function of the M-line protein obscurin in controlling the symmetry of the sarcomere in the flight muscle of Drosophila. Journal of cell science. 2012;125:3367-79.

[144] Katzemich A, Liao KA, Czerniecki S, Schock F. Alp/Enigma family proteins cooperate in Z-disc formation and myofibril assembly. PLoS genetics. 2013;9:e1003342.

[145] Katzemich A, West RJ, Fukuzawa A, Sweeney ST, Gautel M, Sparrow J, et al. Binding partners of the kinase domains in Drosophila obscurin and their effect on the structure of the flight muscle. Journal of cell science. 2015;128:3386-97.

[146] Kaushik JK, Bhat R. Why is trehalose an exceptional protein stabilizer? An analysis of the thermal stability of proteins in the presence of the compatible osmolyte trehalose. The Journal of biological chemistry. 2003;278:26458-65.

[147] Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nature protocols. 2015;10:845.

[148] Khan MA, Neale C, Michaux C, Pomès R, Privé GG, Woody RW, et al. Gauging a hydrocarbon ruler by an intrinsic exciton probe. Biochemistry. 2007;46:4565-79.

[149] Koide S, Sidhu SS. The importance of being tyrosine: lessons in molecular recognition from minimalist synthetic binding proteins. ACS chemical biology. 2009;4:325-34.

[150] Korkmaz EN, Taylor KC, Andreas MP, Ajay G, Heinze NT, Cui Q, et al. A composite approach towards a complete model of the myosin rod. Proteins. 2016;84:172-89.

[151] Korn ED. Coevolution of head, neck, and tail domains of myosin heavy chains. Proceedings of the National Academy of Sciences of the United States of America. 2000;97:12559-64.

[152] Kreipke RE, Kwon YV, Shcherbata HR, Ruohola-Baker H. Drosophila melanogaster as a Model of Muscle Degeneration Disorders. Current topics in developmental biology. 2017;121:83-109.

[153] Kreplak L, Nyland LR, Contompasis JL, Vigoreaux JO. Nanomechanics of Native Thick Filaments from Indirect Flight Muscles. Journal of molecular biology. 2009;386:1403-10.

[154] Kreuzer SM, Elber R. Coiled-coil response to mechanical force: global stability and local cracking. Biophysical journal. 2013;105:951-61.

[155] Kronert WA, Bell KM, Viswanathan MC, Melkani GC, Trujillo AS, Huang A, et al. Prolonged cross-bridge binding triggers muscle dysfunction in a Drosophila model of myosin-based hypertrophic cardiomyopathy. eLife. 2018;7.

[156] Kronert WA, O'Donnell PT, Bernstein SI. A Charge Change in an Evolutionarily-conserved Region of the Myosin Globular Head Prevents Myosin and Thick Filament Accumulation in Drosophila. Journal of molecular biology. 1994;236:697-702.

[157] Kronert WA, O'Donnell PT, Fieck A, Lawn A, Vigoreaux JO, Sparrow JC, et al. Defects in the Drosophila myosin rod permit sarcomere assembly but cause flight muscle degeneration. Journal of molecular biology. 1995;249:111-25.

[158] Kuhn ER, Naik AR, Lewis BE, Kokotovich KM, Li M, Stemmler TL, et al. Nanothermometry Reveals Calcium-Induced Remodeling of Myosin. Nano Letters. 2018;18:7021-9.

[159] Kulke M, Neagoe C, Kolmerer B, Minajeva A, Hinssen H, Bullard B, et al. Kettin, a major source of myofibrillar stiffness in Drosophila indirect flight muscle. The Journal of cell biology. 2001;154:1045-57.

[160] Laibe J, Caffrey A, Broutin M, Guiglion S, Pierscionek B, Nebel J-C. Coil conversion to β-strand induced by dimerization. Proteins: Structure, Function, and Bioinformatics. 2018;86:1221-30.

[161] Lakey A, Labeit S, Gautel M, Ferguson C, Barlow DP, Leonard K, et al. Kettin, a large modular protein in the Z-disc of insect muscles. The EMBO journal. 1993;12:2863-71.

[162] Lamont PJ, Wallefeld W, Hilton-Jones D, Udd B, Argov Z, Barboi AC, et al. Novel mutations widen the phenotypic spectrum of slow skeletal/beta-cardiac myosin (MYH7) distal myopathy. Human mutation. 2014;35:868-79.

[163] Lange S, Pinotsis N, Agarkova I, Ehler E. The M-band: The underestimated part of the sarcomere. Biochimica et biophysica acta Molecular cell research. 2019.

[164] Lee J, Ju M, Cho OH, Kim Y, Nam KT. Tyrosine-Rich Peptides as a Platform for Assembly and Material Synthesis. Adv Sci (Weinh). 2018;6:1801255-.

[165] Lee KH, Sulbaran G, Yang S, Mun JY, Alamo L, Pinto A, et al. Interacting-heads motif has been conserved as a mechanism of myosin II inhibition since before the origin of animals. Proceedings of the National Academy of Sciences of the United States of America. 2018;115:E1991-e2000.

[166] Lehman W. Thin Filament Structure and the Steric Blocking Model. Comprehensive Physiology. 2016;6:1043-69.

[167] Lemas D, Lekkas P, Ballif BA, Vigoreaux JO. Intrinsic disorder and multiple phosphorylations constrain the evolution of the flightin N-terminal region. Journal of proteomics. 2016;135:191-200.

[168] Lemke SB, Schnorrer F. Mechanical forces during muscle development. Mechanisms of development. 2017;144:92-101.

[169] Levine RJ, Elfvin M, Dewey MM, Walcott B. Paramyosin in invertebrate muscles. II. Content in relation to structure and function. The Journal of cell biology. 1976;71:273-9.

[170] Levine RJ, Kensler RW, Yang Z, Stull JT, Sweeney HL. Myosin light chain phosphorylation affects the structure of rabbit skeletal muscle thick filaments. Biophysical journal. 1996;71:898-907.

[171] Levine RJC, Yang Z, Epstein ND, Fananapazir L, Stull JT, Sweeney HL. Structural and Functional Responses of Mammalian Thick Filaments to Alterations in Myosin Regulatory Light Chains. Journal of structural biology. 1998;122:149-61.

[172] Li M, Liu J, Ran X, Fang M, Shi J, Qin H, et al. Resurrecting abandoned proteins with pure water: CD and NMR studies of protein fragments solubilized in salt-free water. Biophysical journal. 2006;91:4201-9.

[173] Li Y, Lang P, Linke WA. Titin stiffness modifies the force-generating region of muscle sarcomeres. Scientific reports. 2016;6:24492.

[174] Liao KA, Gonzalez-Morales N, Schock F. Zasp52, a Core Z-disc Protein in Drosophila Indirect Flight Muscles, Interacts with alpha-Actinin via an Extended PDZ Domain. PLoS genetics. 2016;12:e1006400.

[175] Lin BL, Song T, Sadayappan S. Myofilaments: Movers and Rulers of the Sarcomere. Comprehensive Physiology. 2017;7:675-92.

[176] Lin K, Simossis VA, Taylor WR, Heringa J. A simple and fast secondary structure prediction method using hidden neural networks. Bioinformatics. 2005;21:152-9.

[177] Lindstedt S, Nishikawa K. Huxleys' Missing Filament: Form and Function of Titin in Vertebrate Striated Muscle. Annual review of physiology. 2017;79:145-66.

[178] Linke WA. Titin Gene and Protein Functions in Passive and Active Muscle. Annual review of physiology. 2018;80:389-411.

[179] Liu AG, Matthews JJ, Menon LR, McIlroy D, Brasier MD. Haootia quadriformis n. gen., n. sp., interpreted as a muscular cnidarian impression from the Late Ediacaran period (approx. 560 Ma). Proceedings Biological sciences. 2014;281.

[180] Liu H, Miller MS, Swank DM, Kronert WA, Maughan DW, Bernstein SI. Paramyosin phosphorylation site disruption affects indirect flight muscle stiffness and power generation in Drosophila melanogaster. Proceedings of the National Academy of Sciences of the United States of America. 2005;102:10522-7.

[181] Ltd A. ApowersoftEdit. https://www.apowersoft.com; 2017.

[182] Lupas AN, Gruber M. The structure of alpha-helical coiled coils. Advances in protein chemistry. 2005;70:37-78.

[183] Luther PK. The vertebrate muscle Z-disc: sarcomere anchor for structure and signalling. Journal of muscle research and cell motility. 2009;30:171-85.

[184] Lygren B, Tasken K. Compartmentalized cAMP signalling is important in the regulation of Ca(2+) cycling in the heart. Biochemical Society transactions. 2006;34:489-91.

[185] Mackie GO, Mills CE, Singla CL. Structure and function of the prehensile tentilla of Euplokamis (Ctenophora, Cydippida). Zoomorphology. 1988;107:319-37.

[186] Maier D. The evolution of transcriptional repressors in the Notch signaling pathway: a computational analysis. Hereditas. 2019;156:5.

[187] Manring HR, Carter OA, Ackermann MA. Obscure functions: the location-function relationship of obscurins. Biophysical reviews. 2017;9:245–58.

[188] Mansson A, Persson M, Shalabi N, Rassier DE. Nonlinear Actomyosin Elasticity in Muscle? Biophysical journal. 2019;116:330-46.

[189] Mardahl-Dumesnil M, Fowler VM. Thin filaments elongate from their pointed ends during myofibril assembly in Drosophila indirect flight muscle. The Journal of cell biology. 2001;155:1043-53.

[190] Maroto M, Arredondo J, Goulding D, Marco R, Bullard B, Cervera M. Drosophila paramyosin/miniparamyosin gene products show a large diversity in quantity, localization, and isoform pattern: a possible role in muscle maturation and function. The Journal of cell biology. 1996;134:81-92.

[191] Maroto M, Arredondo JJ, San Roman M, Marco R, Cervera M. Analysis of the paramyosin/miniparamyosin gene. Miniparamyosin is an independently transcribed, distinct paramyosin isoform, widely distributed in invertebrates. The Journal of biological chemistry. 1995;270:4375-82.

[192] Maughan D, Vigoreaux J. Nature's Strategy for Optimizing Power Generation in Insect Flight Muscle. Boston, MA: Springer US; 2005. p. 157-67.

[193] Maurer M, Oostenbrink C. Water in protein hydration and ligand recognition. J Mol Recognit. 2019;32:e2810.

[194] Mazzier D, Grassi L, Moretto A, Aleman C, Formaggio F, Toniolo C, et al. En route towards the peptide gamma-helix: X-ray diffraction analyses and conformational energy calculations of Adm-rich short peptides. Journal of peptide science : an official publication of the European Peptide Society. 2017;23:346-62.

[195] McGrath MJ, Cottle DL, Nguyen MA, Dyson JM, Coghill ID, Robinson PA, et al. Four and a half LIM protein 1 binds myosin-binding protein C and regulates myosin filament formation and sarcomere assembly. The Journal of biological chemistry. 2006;281:7666-83.

[196] Mecozzi S, West AP, Dougherty DA. Cation-pi interactions in aromatics of biological and medicinal interest: electrostatic potential surfaces as a useful qualitative guide. Proceedings of the National Academy of Sciences. 1996;93:10566-71.

[197] Menard L, Maughan D, Vigoreaux J. The structural and functional coordination of glycolytic enzymes in muscle: evidence of a metabolon? Biology. 2014;3:623-44.

[198] Menard L, Nyland L, Vigoreaux J. The Structural and Biomechanical Properties of Insect Thick Filaments Expressing Flightin and Cardiac Myosin Binding Protein-C. Microscopy and Microanalysis. 2013;19:80-1.

[199] Merkel JS, Sturtevant JM, Regan L. Sidechain interactions in parallel beta sheets: the energetics of cross-strand pairings. Structure. 1999;7:1333-43.

[200] Micsonai A, Wien F, Kernya L, Lee Y-H, Goto Y, Réfrégiers M, et al. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. Proceedings of the National Academy of Sciences. 2015;112:E3095-E103.

[201] Miller MS, Dambacher CM, Knowles AF, Braddock JM, Farman GP, Irving TC, et al. Alternative S2 hinge regions of the myosin rod affect myofibrillar structure and myosin kinetics. Biophysical journal. 2009;96:4132-43.

[202] Milner-White EJ. Situations of gamma-turns in proteins. Their relation to alpha-helices, beta-sheets and ligand binding sites. Journal of molecular biology. 1990;216:386-97.

[203] Minin KA, Zhmurov A, Marx KA, Purohit PK, Barsegov V. Dynamic Transition from alpha-Helices to beta-Sheets in Polypeptide Coiled-Coil Motifs. Journal of the American Chemical Society. 2017;139:16168-77.

[204] Mishra P, Varuzhanyan G, Pham AH, Chan DC. Mitochondrial Dynamics is a Distinguishing Feature of Skeletal Muscle Fiber Types and Regulates Organellar Compartmentalization. Cell metabolism. 2015;22:1033-44.

[205] Miyamoto CA, Fischman DA, Reinach FC. The interface between MyBP-C and myosin: site-directed mutagenesis of the CX myosin-binding domain of MyBP-C. Journal of muscle research and cell motility. 1999;20:703-15.

[206] Mohr SE. First in Fly: Drosophila Research and Biological Discovery: Harvard University Press; 2018.

[207] Molnár I, Migh E, Szikora S, Kalmár T, Végh AG, Deák F, et al. DAAM is required for thin filament formation and Sarcomerogenesis during muscle development in Drosophila. PLoS genetics. 2014;10:e1004166.

[208] Montgomerie S, Sundararaj S, Gallin WJ, Wishart DS. Improving the accuracy of protein secondary structure prediction using structural alignment. BMC Bioinformatics. 2006;7:301.

[209] Moore JR, Campbell SG, Lehman W. Structural determinants of muscle thin filament cooperativity. Archives of biochemistry and biophysics. 2016;594:8-17.

[210] Moos C, Offer G, Starr R, Bennett P. Interaction of C-protein with myosin, myosin rod and light meromyosin. Journal of molecular biology. 1975;97:1-9.

[211] Mukherjee D, Gai F. Exciton circular dichroism couplet arising from nitrile-derivatized aromatic residues as a structural probe of proteins. Anal Biochem. 2016;507:74-8.

[212] Mukherjee S, Kapp EA, Lothian A, Roberts AM, Vasil'ev YV, Boughton BA, et al. Characterization and Identification of Dityrosine Cross-Linked Peptides Using Tandem Mass Spectrometry. Analytical Chemistry. 2017;89:6136-45.

[213] Myhre JL, Pilgrim D. A Titan but not necessarily a ruler: assessing the role of titin during thick filament patterning and assembly. Anatomical record (Hoboken, NJ : 2007). 2014;297:1604-14.

[214] Nave R, Weber K. A myofibrillar protein of insect muscle related to vertebrate titin connects Z band and A band: purification and molecular characterization of invertebrate mini-titin. Journal of cell science. 1990;95 (Pt 4):535-44.

[215] Némethy G, Printz MP. The γ Turn, a Possible Folded Conformation of the Polypeptide Chain. Comparison with the β Turn. Macromolecules. 1972;5:755-8.

[216] Nishikawa K. Eccentric contraction: unraveling mechanisms of force enhancement and energy conservation. The Journal of experimental biology. 2016;219:189-96.

[217] Nishikawa K. Eccentric contraction: unraveling mechanisms of force enhancement and energy conservation. The Journal of experimental biology. 2016;219:189-96.

[218] Nishikawa KC, Monroy JA, Tahir U. Muscle Function from Organisms to Molecules. Integrative and comparative biology. 2018;58:194-206.

[219] Nishikawa KC, Monroy JA, Uyeno TE, Yeo SH, Pai DK, Lindstedt SL. Is titin a 'winding filament'? A new twist on muscle contraction. Proceedings Biological sciences. 2012;279:981-90.

[220] Nishio M, Umezawa Y, Fantini J, Weiss MS, Chakrabarti P. CH–π hydrogen bonds in biological macromolecules. Physical Chemistry Chemical Physics. 2014;16:12648-83.

[221] Nongthomba U, Cummins M, Clark S, Vigoreaux JO, Sparrow JC. Suppression of muscle hypercontraction by mutations in the myosin heavy chain gene of Drosophila melanogaster. Genetics. 2003;164:209-22.

[222] Noronha M, Lima JC, Lamosa P, Santos H, Maycock C, Ventura R, et al. Intramolecular Fluorescence Quenching of Tyrosine by the Peptide α-Carbonyl Group Revisited. The Journal of Physical Chemistry A. 2004;108:2155-66.

[223] Nyland LR, Palmer BM, Chen Z, Maughan DW, Seidman CE, Seidman JG, et al. Cardiac myosin binding protein-C is essential for thick-filament stability and flexural rigidity. Biophysical journal. 2009;96:3273-80.

[224] Obermann WM, Gautel M, Steiner F, van der Ven PF, Weber K, Furst DO. The structure of the sarcomeric M band: localization of defined domains of myomesin, M-protein, and the 250-kD carboxy-terminal region of titin by immunoelectron microscopy. The Journal of cell biology. 1996;134:1441-53.

[225] Obermann WM, van der Ven PF, Steiner F, Weber K, Fürst DO. Mapping of a myosin-binding domain and a regulatory phosphorylation site in M-protein, a structural protein of the sarcomeric M band. Molecular biology of the cell. 1998;9:829-40.

[226] Obermann WMJ, Gautel M, Weber K, Fürst DO. Molecular structure of the sarcomeric M band: mapping of titin and myosin binding domains in myomesin and the identification of a potential regulatory phosphorylation site in myomesin. The EMBO journal. 1997;16:211-20.

[227] Ochala J, Sun YB. Novel myosin-based therapies for congenital cardiac and skeletal myopathies. Journal of medical genetics. 2016;53:651-4.

[228] O'Donnell PT, Collier VL, Mogami K, Bernstein SI. Ultrastructural and molecular analyses of homozygous-viable Drosophila melanogaster muscle mutants indicate there is a complex pattern of myosin heavy-chain isoform distribution. Genes & development. 1989;3:1233-46.

[229] Offer G. Skip residues correlate with bends in the myosin tail. Journal of molecular biology. 1990;216:213-8.

[230] Orfanos Z, Sparrow JC. Myosin isoform switching during assembly of the Drosophila flight muscle thick filament lattice. Journal of cell science. 2013;126:139-48.

[231] Otten E. Optimal design of vertebrate and insect sarcomeres. Journal of Morphology. 1987;191:49-62.

[232] Ozbas B, Rajagopal K, Schneider JP, Pochan DJ. Semiflexible Chain Networks Formed via Self-Assembly of β-Hairpin Molecules. Physical Review Letters. 2004;93:268106.

[233] Pappas CT, Bhattacharya N, Cooper JA, Gregorio CC. Nebulin interacts with CapZ and regulates thin filament architecture within the Z-disc. Molecular biology of the cell. 2008;19:1837-47.

[234] Parker F, Batchelor M, Wolny M, Hughes R, Knight PJ, Peckham M. A1603P and K1617del, Mutations in β-Cardiac Myosin Heavy Chain that Cause Laing Early-Onset Distal Myopathy, Affect Secondary Structure and Filament Formation In Vitro and In Vivo. Journal of molecular biology. 2018;430:1459-78.

[235] Parui S, Jana B. Factors Promoting the Formation of Clathrate-Like Ordering of Water in Biomolecular Structure at Ambient Temperature and Pressure. The Journal of Physical Chemistry B. 2019;123:811-24.

[236] Patel SR, Saide JD. Stretchin-klp, a novel Drosophila indirect flight muscle protein, has both myosin dependent and independent isoforms. Journal of muscle research and cell motility. 2005;26:213-24.

[237] Peckham M, Knight PJ. When a predicted coiled coil is really a single α-helix, in myosins and other proteins. Soft Matter. 2009;5:2493-503.

[238] Petersen B, Lundegaard C, Petersen TN. NetTurnP--neural network prediction of beta-turns by use of evolutionary information and predicted protein sequence features. PloS one. 2010;5:e15079.

[239] Pinniger GJ, Ranatunga KW, Offer GW. Crossbridge and non-crossbridge contributions to tension in lengthening rat muscle: force-induced reversal of the power stroke. The Journal of physiology. 2006;573:627-43.

[240] Poglazov BF, Samokhin GP, Klibanov AM, Levitsky DI, Martinek K, Berezin IV. The effect of mechanical stretching of the myosin rod component (fragment LMMMM S-2) on the ATPase activity of myosin. Biochimica et biophysica acta. 1978;524:245-53.

[241] Polêto MD, Rusu VH, Grisci BI, Dorn M, Lins RD, Verli H. Aromatic Rings Commonly Used in Medicinal Chemistry: Force Fields Comparison and Interactions With Water Toward the Design of New Chemical Entities. Frontiers in Pharmacology. 2018;9.

[242] Postma AV, van Engelen K, van de Meerakker J, Rahman T, Probst S, Baars MJ, et al. Mutations in the sarcomere gene MYH7 in Ebstein anomaly. Circulation Cardiovascular genetics. 2011;4:43-50.

[243] Potikanond S, Nimlamool W, Noordermeer J, Fradkin LG. Muscular Dystrophy Model. Advances in experimental medicine and biology. 2018;1076:147-72.

[244] Qadota H, Mayans O, Matsunaga Y, McMurry JL, Wilson KJ, Kwon GE, et al. The SH3 domain of UNC-89 (obscurin) interacts with paramyosin, a coiled-coil protein, in Caenorhabditis elegans muscle. Molecular biology of the cell. 2016;27:1606-20.

[245] Qi HW, Nakka P, Chen C, Radhakrishnan ML. The effect of macromolecular crowding on the electrostatic component of barnase-barstar binding: a computational, implicit solvent-based study. PloS one. 2014;9:e98618.

[246] Qin Z, Kreplak L, Buehler MJ. Hierarchical structure controls nanomechanical properties of vimentin intermediate filaments. PloS one. 2009;4:e7294.

[247] Qiu F, Brendel S, Cunha PM, Astola N, Song B, Furlong EE, et al. Myofilin, a protein in the thick filaments of insect muscle. Journal of cell science. 2005;118:1527-36.

[248] Quigley A, Williams DR. The second virial coefficient as a predictor of protein aggregation propensity: A self-interaction chromatography study. European journal of pharmaceutics and biopharmaceutics : official journal of Arbeitsgemeinschaft fur Pharmazeutische Verfahrenstechnik eV. 2015;96:282-90.

[249] Rai M, Nongthomba U, Grounds MD. Skeletal muscle degeneration and regeneration in mice and flies. Current topics in developmental biology. 2014;108:247-81.

[250] Ramírez-Alvarado M, Blanco FJ, Niemann H, Serrano L. Role of β-turn residues in β-hairpin formation and stability in designed peptides. Journal of molecular biology. 1997;273:898-912.

[251] Rassier DE. Sarcomere mechanics in striated muscles: from molecules to sarcomeres to cells. American journal of physiology Cell physiology. 2017;313:C134-c45.

[252] Reedy MC, Bullard B, Vigoreaux JO. Flightin is essential for thick filament assembly and sarcomere stability in Drosophila flight muscles. The Journal of cell biology. 2000;151:1483-500.

[253] Remko M, Soralova S. Effect of water coordination on competition between pi and non-pi cation binding sites in aromatic amino acids: L-phenylalanine, L-tyrosine, and L-tryptophan Li+, Na +, and K+ complexes. Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry. 2012;17:621-30.

[254] Richardson JS. The Anatomy and Taxonomy of Protein Structure. In: Anfinsen CB, Edsall JT, Richards FM, editors. Advances in protein chemistry: Academic Press; 1981. p. 167-339.

[255] Rohl CA, Baldwin RL. Deciphering rules of helix stability in peptides. Methods in enzymology. 1998;295:1-26.

[256] Rohl CA, Fiori W, Baldwin RL. Alanine Is Helix-Stabilizing in Both Template-Nucleated and Standard Peptide Helices. Proceedings of the National Academy of Sciences of the United States of America. 1999;96:3682-7.

[257] Rose A, Meier I. Scaffolds, levers, rods and springs: diverse cellular functions of long coiled-coil proteins. Cellular and molecular life sciences : CMLS. 2004;61:1996-2009.

[258] Rost B, Sander C. Improved prediction of protein secondary structure by use of sequence profiles and neural networks. Proceedings of the National Academy of Sciences of the United States of America. 1993;90:7558-62.

[259] Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nature protocols. 2010;5:725-38.

[260] Royuela M, Fraile B, De Miguel MP, Cervera M, Paniagua R. Immunohistochemical study and western blotting analysis of titin-like proteins in the striated muscle of Drosophila melanogaster and in the striated and smooth muscle of the oligochaete Eisenia foetida. Microscopy research and technique. 1996;35:349-56.

[261] Rui Y, Bai J, Perrimon N. Sarcomere formation occurs by the assembly of multiple latent protein complexes. PLoS genetics. 2010;6:e1001208.

[262] Saide JD, Chin-Bow S, Hogan-Sheldon J, Busquets-Turner L, Vigoreaux JO, Valgeirsdottir K, et al. Characterization of components of Z-bands in the fibrillar flight muscle of Drosophila melanogaster. The Journal of cell biology. 1989;109:2157-67.

[263] Salvi SS, Kumar RP, Ramachandra NB, Sparrow JC, Nongthomba U. Mutations in Drosophila myosin rod cause defects in myofibril assembly. Journal of molecular biology. 2012;419:22-40.

[264] Sanger JW, Wang J, Fan Y, White J, Mi-Mi L, Dube DK, et al. Assembly and Maintenance of Myofibrils in Striated Muscle. Handbook of experimental pharmacology. 2017;235:39-75.

[265] Sarov M, Barz C, Jambor H, Hein MY, Schmied C, Suchold D, et al. A genome-wide resource for the analysis of protein localisation in Drosophila. eLife. 2016;5:e12068.

[266] Scheerer D, Chi H, McElheny D, Samer A, Keiderling TA, Hauser K. Role of Aromatic Cross-Links in Structure and Dynamics of Model Three-Stranded beta-Sheet Peptides. The journal of physical chemistry A. 2018;122:543-53.

[267] Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nature methods. 2012;9:671-5.

[268] Schneider JP, Pochan DJ, Ozbas B, Rajagopal K, Pakstis L, Kretsinger J. Responsive Hydrogels from the Intramolecular Folding and Self-Assembly of a Designed Peptide. Journal of the American Chemical Society. 2002;124:15030-7.

[269] Schneider TR, Gerhardt E, Lee M, Liang P-H, Anderson KS, Schlichting I. Loop Closure and Intersubunit Communication in Tryptophan Synthase. Biochemistry. 1998;37:5394-406.

[270] Schnorrer F, Schonbauer C, Langer CC, Dietzl G, Novatchkova M, Schernhuber K, et al. Systematic genetic analysis of muscle morphogenesis and function in Drosophila. Nature. 2010;464:287-91.

[271] Sebe-Pedros A, Grau-Bove X, Richards TA, Ruiz-Trillo I. Evolution and classification of myosins, a paneukaryotic whole-genome approach. Genome biology and evolution. 2014;6:290-305.

[272] Seipel K, Schmid V. Evolution of striated muscle: jellyfish and the origin of triploblasty. Developmental biology. 2005;282:14-26.

[273] Shapovalov M, Vucetic S, Dunbrack RL, Jr. A new clustering and nomenclature for beta turns derived from high-resolution protein structures. PLoS computational biology. 2019;15:e1006844.

[274] Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, et al. Atomic-level characterization of the structural dynamics of proteins. Science. 2010;330:341-6.

[275] Shwartz A, Dhanyasi N, Schejter ED, Shilo BZ. The Drosophila formin Fhos is a primary mediator of sarcomeric thin-filament array assembly. eLife. 2016;5.

[276] Sibanda BL, Blundell TL, Thornton JM. Conformation of beta-hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering. Journal of molecular biology. 1989;206:759-77.

[277] Sink H. Muscle Development in Drosophila. New York, NY: Springer; 2006.

[278] Sinz A. Divide and conquer: cleavable cross-linkers to study protein conformation and protein-protein interactions. Analytical and bioanalytical chemistry. 2017;409:33-44.

[279] Sjoblom B, Salmazo A, Djinovic-Carugo K. Alpha-actinin structure and regulation. Cellular and molecular life sciences : CMLS. 2008;65:2688-701.

[280] Skubiszak L. Geometrical conditions indispensable for muscle contraction. International Journal of Molecular Sciences. 2011;12:2138-57.

[281] Skubiszak L, Kowalczyk L. Myosin molecule packing within the vertebrate skeletal muscle thick filaments. A complete bipolar model. Acta biochimica Polonica. 2002;49:829-40.

[282] Smith CK, Withka JM, Regan L. A Thermodynamic Scale for the β-Sheet Forming Tendencies of the Amino Acids. Biochemistry. 1994;33:5510-7.

[283] Smith DA, Geeves MA, Sleep J, Mijailovich SM. Towards a unified theory of muscle contraction. I: foundations. Annals of biomedical engineering. 2008;36:1624-40.

[284] Sohn RL, Vikstrom KL, Strauss M, Cohen C, Szent-Gyorgyi AG, Leinwand LA. A 29 residue region of the sarcomeric myosin rod is necessary for filament formation. Journal of molecular biology. 1997;266:317-30.

[285] Soranno A, Cabassi F, Orselli ME, Cellmer T, Gori A, Longhi R, et al. Dynamics of Structural Elements of GB1 β-Hairpin Revealed by Tryptophan–Cysteine Contact Formation Experiments. The Journal of Physical Chemistry B. 2018;122:11468-77.

[286] Sotavalta O. Recordings of High Wing-Stroke and Thoracic Vibration Frequency in Some Midges. Biological Bulletin. 1953;104:439-44.

[287] Soto-Adames FN, Alvarez-Ortiz P, Vigoreaux JO. An evolutionary analysis of flightin reveals a conserved motif unique and widespread in Pancrustacea. Journal of molecular evolution. 2014;78:24-37.

[288] Souza JM, Daikhin E, Yudkoff M, Raman CS, Ischiropoulos H. Factors determining the selectivity of protein tyrosine nitration. Archives of biochemistry and biophysics. 1999;371:169-78.

[289] Spletter ML, Barz C, Yeroslaviz A, Zhang X, Lemke SB, Bonnard A, et al. A transcriptomics resource reveals a transcriptional transition during ordered sarcomere morphogenesis in flight muscle. eLife. 2018;7.

[290] Squire J. The Structural Basis of Muscular Contraction: Springer US; 1981.

[291] Squire JM. Muscle myosin filaments: cores, crowns and couplings. Biophysical reviews. 2009;1:149.

[292] Squire JM, Bekyarova T, Farman G, Gore D, Rajkumar G, Knupp C, et al. The myosin filament superlattice in the flight muscles of flies: A-band lattice optimisation for stretch-activation? Journal of molecular biology. 2006;361:823-38.

[293] Sreerama N, Woody RW. Structural composition of betaI- and betaII-proteins. Protein Science : A Publication of the Protein Society. 2003;12:384-8.

[294] Starna Cells I. Suggestions for Cleaning Cells.

[295] Steinmetz PR, Kraus JE, Larroux C, Hammel JU, Amon-Hassenzahl A, Houliston E, et al. Independent evolution of striated muscles in cnidarians and bilaterians. Nature. 2012;487:231-4.

[296] Suggs JA, Cammarato A, Kronert WA, Nikkhoy M, Dambacher CM, Megighian A, et al. Alternative S2 hinge regions of the myosin rod differentially affect muscle function, myofibril dimensions and myosin tail length. Journal of molecular biology. 2007;367:1312-29.

[297] Suggs JA, Melkani GC, Glasheen BM, Detor MM, Melkani A, Marsan NP, et al. A Drosophila model of dominant inclusion body myopathy type 3 shows diminished myosin kinetics that reduce muscle power and yield myofibrillar defects. Disease models & mechanisms. 2017;10:761-71.

[298] Sunde M, Blake CCF. From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. Quarterly Reviews of Biophysics. 1998;31:1-39.

[299] Swank DM, Vishnudas VK, Maughan DW. An exceptionally fast actomyosin reaction powers insect flight muscle. Proceedings of the National Academy of Sciences of the United States of America. 2006;103:17543-7.

[300] Swank DM, Wells L, Kronert WA, Morrill GE, Bernstein SI. Determining structure/function relationships for sarcomeric myosin heavy chain by genetic and transgenic manipulation of Drosophila. Microscopy research and technique. 2000;50:430-42.

[301] Swanson CJ, Sivaramakrishnan S. Harnessing the unique structural properties of isolated α-helices. The Journal of biological chemistry. 2014;289:25460-7.

[302] Szikora S, Gajdos T, Novák T, Farkas D, Földi I, Lenart P, et al. Nanoscopy reveals the layered organization of the sarcomeric H-zone and I-band complexes. The Journal of cell biology. 2020;219.

[303] Tajsharghi H. Thick and Thin Filament Gene Mutations in Striated Muscle Diseases. International Journal of Molecular Sciences. 2008;9:1259-75.

[304] Tama F, Feig M, Liu J, Brooks CL, 3rd, Taylor KA. The requirement for mechanical coupling between head and S2 domains in smooth muscle myosin ATPase regulation and its implications for dimeric motor function. Journal of molecular biology. 2005;345:837-54.

[305] Tanner BC, Miller MS, Miller BM, Lekkas P, Irving TC, Maughan DW, et al. COOH-terminal truncation of flightin decreases myofilament lattice organization, cross-bridge binding, and power output in Drosophila indirect flight muscle. American journal of physiology Cell physiology. 2011;301:C383-91.

[306] Taylor KC, Buvoli M, Korkmaz EN, Buvoli A, Zheng Y, Heinze NT, et al. Skip residues modulate the structural properties of the myosin rod and guide thick filament assembly. Proceedings of the National Academy of Sciences of the United States of America. 2015;112:E3806-15.

[307] Team TGD. GIMP. https://www.gimp.org; 2017.

[308] Terzi E, Hoelzemann G, Seelig J. Reversible Random Coil-β-Sheet Transition of the Alzheimer β-Amyloid Fragment (25-35). Biochemistry. 1994;33:1345-50.

[309] Thomas N, Thornhill RA. Stretch activation and nonlinear elasticity of muscle cross-bridges. Biophysical journal. 1996;70:2807-18.

[310] Togashi M, Kakinuma M, Nakaya M, Ooi T, Watabe S. Differential scanning calorimetry and circular dichroism spectrometry of walleye pollack myosin and light meromyosin. Journal of agricultural and food chemistry. 2002;50:4803-11.

[311] Trybus KM, Freyzon Y, Faust LZ, Sweeney HL. Spare the rod, spoil the regulation: necessity for a myosin rod. Proceedings of the National Academy of Sciences of the United States of America. 1997;94:48-52.

[312] Tskhovrebova L, Trinick J. Titin and Nebulin in Thick and Thin Filament Length Regulation. Sub-cellular biochemistry. 2017;82:285-318.

[313] Tune TC, Ma W, Irving TC, Sponberg S. A nanometer difference in myofilament lattice spacing of two cockroach leg muscles explains their different functions. bioRxiv. 2019:656272.

[314] Upadhye V, Agashe S. Effect of Temperature and Pressure Variations on the Resonant Frequency of Piezoelectric Material. Measurement and Control. 2016;49:286-92.

[315] van Eldik W, Passier R. Signalling in sarcomeres in development and disease. Netherlands heart journal : monthly journal of the Netherlands Society of Cardiology and the Netherlands Heart Foundation. 2013;21:367-71.

[316] van Straaten M, Goulding D, Kolmerer B, Labeit S, Clayton J, Leonard K, et al. Association of kettin with actin in the Z-disc of insect flight muscle. Journal of molecular biology. 1999;285:1549-62.

[317] Vardar D, Buckley DA, Frank BS, McKnight CJ. NMR structure of an F-actin-binding "headpiece" motif from villin. Journal of molecular biology. 1999;294:1299-310.

[318] Vemuri R, Lankford EB, Poetter K, Hassanzadeh S, Takeda K, Yu ZX, et al. The stretch-activation response may be critical to the proper functioning of the mammalian heart. Proceedings of the National Academy of Sciences of the United States of America. 1999;96:1048-53.

[319] Vigoreaux J. Nature’s Versatile Engine: Insect Flight Muscle Inside and Out. New York, New York: Springer Science & Business Media, Inc; 2007.

[320] Vigoreaux JO. Alterations in flightin phosphorylation in Drosophila flight muscles are associated with myofibrillar defects engendered by actin and myosin heavy-chain mutant alleles. Biochem Genet. 1994;32:301-14.

[321] Vigoreaux JO. Genetics of the Drosophila flight muscle myofibril: a window into the biology of complex systems. BioEssays : news and reviews in molecular, cellular and developmental biology. 2001;23:1047-63.

[322] Vigoreaux JO, Hernandez C, Moore J, Ayer G, Maughan D. A genetic deficiency that spans the flightin gene of Drosophila melanogaster affects the ultrastructure and function of the flight muscles. The Journal of experimental biology. 1998;201:2033-44.

[323] Vigoreaux JO, Perry LM. Multiple isoelectric variants of flightin in Drosophila stretch-activated muscles are generated by temporally regulated phosphorylations. Journal of muscle research and cell motility. 1994;15:607-16.

[324] Vigoreaux JO, Saide JD, Pardue ML. Structurally different Drosophila striated muscles utilize distinct variants of Z-band-associated proteins. Journal of muscle research and cell motility. 1991;12:340-54.

[325] Vigoreaux JO, Saide JD, Valgeirsdottir K, Pardue ML. Flightin, a novel myofibrillar protein of Drosophila stretch-activated muscles. The Journal of cell biology. 1993;121:587-98.

[326] Wagner GP, Altenberg L. Perspective: Complex Adaptations and the Evolution of Evolvability. Evolution. 1996;50:967-76.

[327] Wallgren-Pettersson C, Sewry CA, Nowak KJ, Laing NG. Nemaline myopathies. Seminars in pediatric neurology. 2011;18:230-8.

[328] Wallimann P, Kennedy RJ, Kemp DS. Large Circular Dichroism Ellipticities for N-Templated Helical Polypeptides Are Inconsistent with Currently Accepted Helicity Algorithms. Angewandte Chemie (International ed in English). 1999;38:1290-2.

[329] Walters J, Milam SL, Clark AC. Practical approaches to protein folding and assembly: spectroscopic strategies in thermodynamics and kinetics. Methods in enzymology. 2009;455:1-39.

[330] Wang J, Wang Y, Zou Y, Sun K, Wang Z, Ding H, et al. Malignant effects of multiple rare variants in sarcomere genes on the prognosis of patients with hypertrophic cardiomyopathy. European journal of heart failure. 2014;16:950-7.

[331] Watkins AM, Arora PS. Anatomy of β-strands at protein-protein interfaces. ACS chemical biology. 2014;9:1747-54.

[332] Wei Y, Thyparambil AA, Latour RA. Protein helical structure determination using CD spectroscopy for solutions with strong background absorbance from 190 to 230nm. Biochimica et biophysica acta. 2014;1844:2331-7.

[333] Weitkamp B, Jurk K, Beinbrech G. Projectin-thin filament interactions and modulation of the sensitivity of the actomyosin ATPase to calcium by projectin kinase. The Journal of biological chemistry. 1998;273:19802-8.

[334] Weitkunat M, Brasse M, Bausch AR, Schnorrer F. Mechanical tension and spontaneous muscle twitching precede the formation of cross-striated muscle in vivo. Development (Cambridge, England). 2017;144:1261-72.

[335] Wells L, Edwards KA, Bernstein SI. Myosin heavy chain isoforms regulate muscle function but not myofibril assembly. The EMBO journal. 1996;15:4454-9.

[336] White GE, Erickson HP. Sequence divergence of coiled coils--structural rods, myosin filament packing, and the extraordinary conservation of cohesins. Journal of structural biology. 2006;154:111-21.

[337] Whitmore L, Wallace BA. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res. 2004;32:W668-73.

[338] Whitmore L, Wallace BA. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers. 2008;89:392-400.

[339] Willis BC, Ponce-Balbuena D, Jalife J. Protein assemblies of sodium and inward rectifier potassium channels control cardiac excitability and arrhythmogenesis. American journal of physiology Heart and circulatory physiology. 2015;308:H1463-73.

[340] Wilson E. The Diversity of Life. Cambridge, MA and Norton, NY: Belknap Press; 1992.

[341] Windsor IW, Gold B, Raines RT. An n-->pi* Interaction in the Bound Substrate of Aspartic Proteases Replicates the Oxyanion Hole. ACS catalysis. 2019;9:1464-71.

[342] Witt CC, Burkart C, Labeit D, McNabb M, Wu Y, Granzier H, et al. Nebulin regulates thin filament length, contractility, and Z-disk structure in vivo. The EMBO journal. 2006;25:3843-55.

[343] Wolny M, Colegrave M, Colman L, White E, Knight PJ, Peckham M. Cardiomyopathy mutations in the tail of beta-cardiac myosin modify the coiled-coil

structure and affect integration into thick filaments in muscle sarcomeres in adult cardiomyocytes. The Journal of biological chemistry. 2013;288:31952-62.

[344] Woody RW. [4] Circular dichroism. Methods in enzymology: Academic Press; 1995. p. 34-71.

[345] Wu C-SC, Yang JT. Reexamination of the conformation of muscle proteins by optical activity. Biochemistry. 1976;15:3007-14.

[346] Xue C, Lin TY, Chang D, Guo Z. Thioflavin T as an amyloid dye: fibril quantification, optimal concentration and effect on aggregation. Royal Society open science. 2017;4:160696.

[347] Yachdav G, Kloppmann E, Kajan L, Hecht M, Goldberg T, Hamp T, et al. PredictProtein--an open resource for online prediction of protein structural and functional features. Nucleic Acids Res. 2014;42:W337-43.

[348] Yamniuk AP, Ditto N, Patel M, Dai J, Sejwal P, Stetsko P, et al. Application of a kosmotrope-based solubility assay to multiple protein therapeutic classes indicates broad use as a high-throughput screen for protein therapeutic aggregation propensity. Journal of Pharmaceutical Sciences. 2013;102:2424-39.

[349] Yancey PH, Somero GN. Counteraction of urea destabilization of protein structure by methylamine osmoregulatory compounds of elasmobranch fishes. The Biochemical journal. 1979;183:317-23.

[350] Yang H, Zubarev RA. Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. Electrophoresis. 2010;31:1764-72.

[351] Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nature methods. 2015;12:7-8.

[352] Yogurtcu ON, Wolgemuth CW, Sun SX. Mechanical response and conformational amplification in α-helical coiled coils. Biophysical journal. 2010;99:3895-904.

[353] Youkharibache P, Veretnik S, Li Q, Stanek KA, Mura C, Bourne PE. The Small β-barrel Domain: A Survey-based Structural Analysis. bioRxiv. 2018:140376.

[354] Zhang H, Zielonka J, Sikora A, Joseph J, Xu Y, Kalyanaraman B. The effect of neighboring methionine residue on tyrosine nitration and oxidation in peptides treated with MPO, H2O2, and NO2(-) or peroxynitrite and bicarbonate: role of intramolecular electron transfer mechanism? Archives of biochemistry and biophysics. 2009;484:134-45.

[355] Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008;9:40.

[356] Zhao C, Swank DM. The Drosophila indirect flight muscle myosin heavy chain isoform is insufficient to transform the jump muscle into a highly stretch-activated muscle type. American journal of physiology Cell physiology. 2017;312:C111-c8.

[357] Zhmurov A, Kononova O, Litvinov RI, Dima RI, Barsegov V, Weisel JW. Mechanical transition from α-helical coiled coils to β-sheets in fibrin(ogen). Journal of the American Chemical Society. 2012;134:20396-402.

[358] Zhou HX, Pang X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. Chem Rev. 2018;118:1691-741.

[359] Zondlo NJ. Aromatic-proline interactions: electronically tunable CH/π interactions. Accounts of chemical research. 2013;46:1039-49.

[360] Zou S, Meadows S, Sharp L, Jan LY, Jan YN. Genome-wide study of aging and oxidative stress response in Drosophila melanogaster. Proceedings of the National Academy of Sciences of the United States of America. 2000;97:13726-31.