Published in final edited form as:

J Mol Biol. 2020 January 17; 432(2): 427–447. doi:10.1016/j.jmb.2019.11.013.

X-ray crystallographic and molecular dynamic analyses of *Drosophila melanogaster* embryonic muscle myosin define domains responsible for isoform-specific properties

James T. Caldwell^{1,2}, Daniel J. Mermelstein^{3,4}, Ross C. Walker^{3,5}, Sanford I. Bernstein², Tom Huxford^{1,*}

¹Structural Biochemistry Laboratory, Department of Chemistry & Biochemistry, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-1030

²Department of Biology and Molecular Biology Institute, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-4614

³San Diego Supercomputer Center and Department of Chemistry & Biochemistry, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0505

⁴Present address: OpenEye Scientific Software, 9 Bisbee Court, Suite D, Santa Fe, NM 87508

⁵Present address: GlaxoSmithKline PLC, 1250 S. Collegeville Road, Collegeville, PA 19426

Abstract

Drosophila melanogaster is a powerful system for characterizing alternative myosin isoforms and modeling muscle diseases, but high-resolution structures of fruit fly contractile proteins have not been determined. Here we report the first x-ray crystal structure of an insect myosin: the *D. melanogaster* skeletal muscle myosin II embryonic isoform (EMB). Using our system for recombinant expression of myosin heavy chain (MHC) proteins in whole transgenic flies, we prepared and crystallized stable proteolytic S1-like fragments containing the entire EMB motor domain bound to an essential light chain. We solved the x-ray crystal structure by molecular replacement and refined the resulting model against diffraction data to 2.2 Å resolution. The protein is captured in two slightly different renditions of the rigor-like conformation with a citrate

*Corresponding author: Phone: (619) 594-1606, thuxford@sdsu.edu.

CRedit Author Statement

James T. Caldwell: Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Visualization, Writing—Original Draft, Writing—Review & Editing. **Daniel J. Mermelstein:** Methodology, Investigation, Writing—Review & Editing. **Ross C. Walker:** Methodology, Investigation, Resources, Writing—Review & Editing. **Sanford I. Bernstein:** Conceptualization, Resources, Supervision, Funding Acquisition, Writing—Review & Editing. **Tom Huxford:** Conceptualization, Methodology, Data Curation, Formal Analysis, Visualization, Writing—Original Draft, Writing—Review & Editing

¹**Abbreviations used:** MHC-Myosin heavy chain, EMB-Embryonic muscle myosin isoform in *Drosophila*; IFI-Adult indirect flight muscle isoform of myosin, IFM-Indirect flight muscle, ELC-Essential light chain subunit of myosin, S1-Myosin proteolytic subfragment-1, RLC-Regulatory light chain subunit of myosin, U50-Upper 50 kDa myosin subdomain, L50-Lower 50 kDa myosin subdomain

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

Accession numbers

Coordinates and structure factors have been deposited in the Protein Data Bank with accession number 5W1A.

of crystallization at the nucleotide binding site and exhibits structural features common to myosins of diverse classes from all kingdoms of life. All atom molecular dynamics simulations on EMB in its nucleotide-free state and a derivative homology model containing 61 amino acid substitutions unique to the indirect flight muscle isoform (IFI) suggest that differences in the identity of residues within the relay and the converter that are encoded for by MHC alternative exons 9 and 11, respectively, directly contribute to increased mobility of these regions in IFI relative to EMB. This suggests the possibility that alternative folding or conformational stability within these regions contribute to the observed functional differences in *Drosophila* EMB and IFI myosins.

Graphical Abstract

Keywords

Molecular dynamics; Motor proteins; Protein engineering; Skeletal muscle; X-ray crystallography

INTRODUCTION

Myosin is a superfamily of motor proteins that converts the chemical energy of ATP into mechanical movement along actin filaments [1]. At least eighteen classes of myosin proteins have been identified to control diverse motile processes in eukaryotes [2]. The class II or “conventional” myosins work in conjunction with actin filaments as the principal catalytic components of contractile muscle [3]. Class II myosins are also involved in axon guidance, endothelial cell migration, cell division, and vesicle trafficking [4]. Although all class II myosins follow the same catalytic steps of ATP hydrolysis-driven conformational change, actin binding, inorganic phosphate release, power stroke, and ADP release, the rates with which they proceed through this actin- and strain-dependent “cross-bridge cycle” and the amount of work accomplished with each passage vary significantly [5]. This principle is classically illustrated by the contractile muscle myosin II system in the fruit fly *Drosophila melanogaster*.

Only one gene encoding for a class II muscle myosin heavy chain (MHC)¹ is contained within the *Drosophila* genome [6,7]. This appears to be the case with all arthropods for which full genome sequences are currently available [8,9]. In *Drosophila*, muscle myosin II from the single gene is expressed in cardiac, visceral, and exoskeletal muscles, the latter of which are denoted as skeletal muscles throughout this report. During the lifespan of a fly, at least 13 and as many as 21 different MHC II isoforms are produced via alternative splicing of up to six distinct exon sets [10,11]. For example, the *Drosophila* skeletal muscle embryonic and adult indirect flight muscle (IFM) myosin II protein heavy chain isoforms (referred to throughout this report as EMB and IFI, respectively) differ in amino acid sequence as a result of alternative mRNA processing. EMB mRNA uses exons 3a, 7a, 9b, 11c, 15b and excludes exon 18 while mRNA encoding IFI employs exons 3b, 7d, 9a, 11e, 15a and includes exon 18. The four alternative exons that code for portions of the motor

domain (exons 3, 7, 9, and 11) encode differences, respectively, at 18 of 48 amino acid residues (exon 3), 12 of 35 residues (exon 7), 5 of 47 residues (exon 9), and 26 of 49 residues (exon 11) [12]. These differences contribute to the observed greater than 2-fold decrease of basal ATPase and actin-activated ATPase activities, and roughly 4-fold reduced maximum power produced by EMB myosin expressed transgenically in IFM relative to IFI myosin [13,14]. Furthermore, transgenic flies engineered to express EMB against a MHC null background (*Mhc¹⁰*), though viable, were incapable of flight, exhibited impaired jump ability, poor ambulation, and difficulty in mating, and displayed severe deterioration of muscle myofibrils resulting in a “wings up” phenotype [15]. The *Drosophila* IFI myosin, but not EMB, is capable of generating extremely fast motion in the asynchronous and stretch-activated IFM, sustaining wing beat frequencies of approximately 200 Hz [14,16]. The mechanisms by which these changes in amino acid identity and placement contribute to the unique biophysical and physiological properties of *Drosophila* embryonic and indirect flight muscles remain unclear.

In addition to providing a genetically tractable experimental system for observing how differences in selection of MHC alternative exons give rise to the unique physiologies of muscles required throughout the insect life cycle, the IFM system in *Drosophila* has more recently been developed as a surrogate for testing *in vivo* the effects of MHC mutations associated with skeletal and cardiac muscle diseases as well as in pathology of aging muscle [17–23].

A major barrier to studying *in vitro* the effects of alternative isoforms or disease-causing mutations on myosin II structure and function has been a lack of reliable methods for large scale recombinant expression and purification of myosin proteins. Researchers in the Winkelmann Lab have shown that proper folding of the myosin motor domain requires chaperone proteins that are specific to striated muscle cells [24]. This discovery prompted development of methods for generating recombinant MHC proteins from adenovirus-transfected murine C2C12 myocytes in culture and the structural study of an engineered human β -cardiac cell MHC motor domain-GFP fusion protein [25,26]. Recent studies reported the successful adaptation of HEK cells for low level recombinant expression of human embryonic myosin II and production of unconventional myosin 15 from baculovirus-infected Sf9 insect cells by co-expression with dedicated chaperone proteins [27–29]. In contrast to these cell culture-based systems, we previously reported a robust and inexpensive alternative method in which whole transgenic flies can be used to express specific MHC isoforms within intact IFM at sufficient levels and purity to support biophysical and structural studies [30].

In support of ongoing efforts aimed at exploiting the power of the *Drosophila* genetic model system to understand how changes in amino acid sequence give rise to functional myosin motor proteins with altered physiological properties, we have crystallized and determined the 2.2 Å resolution x-ray crystal structure of an S1-like proteolytic fragment containing the entire EMB motor domain and essential light chain (ELC). This represents the first x-ray crystal structure of an insect myosin. The crystallographic model contains two complete copies of EMB in its rigor-like conformation and reveals the structures of and interactions between its alternative exon-encoded regions. Atomic resolution molecular dynamics

simulations conducted on similarly prepared EMB and IFI models in their nucleotide-free states suggest that variation within the sequences of amino acids encoded by alternative exons 9 and 11 directly contribute to differences in the dynamic modes of the respective protein isoforms and that these regions may contribute to their isoform-specific properties.

RESULTS

EMB motor domain preparation and x-ray crystal structure determination

The *Drosophila melanogaster* embryonic isoform of skeletal muscle myosin II (EMB) was expressed in the IFM of transgenic flies using a system we developed previously for recombinant overexpression of histidine-tagged exogenous MHC proteins [30]. An engineered myosin II heavy chain transgene encoding a protease-cleavable N-terminal hexahistidine-tag, the coding region for the globular head domain including alternative exons specific for EMB (3a, 7a, 9b, and 11c), and genomic DNA encoding the rod and C-terminal domain of MHC was placed under the *Actin88F* promoter and inserted into the fly genome by the PhiC31 integrase-mediated system (Figure 1) [31]. The resulting transgenic flies were crossed into the *Mhc*¹⁰ background, which does not accumulate MHC in its IFM [32]. The homozygous flies expressing the His-tagged EMB MHC are viable and develop normally but are incapable of flying. By 24 hours after eclosure their wings move into a fixed position pointing upward (Figure 1A). This has been shown to correlate with deterioration of the flight muscle ultrastructure in adult flies expressing the EMB isoform [15].

EMB subfragment-1 (S1)-like fragments were prepared via extraction of MHC from whole flies using a scaled up modification of standard protocols coupled with affinity and size exclusion chromatography [33,34]. Limited proteolysis of purified MHC with chymotrypsin resulted in cleavage of the motor domain from the rod and also removal of the histidine tag. Rod fragments were precipitated in low salt buffer and the soluble motor domain-containing S1 fragments were purified by size exclusion chromatography, concentrated, and crystallized. Details on EMB expression, purification and crystallization are given in the Materials & Methods section.

Complete diffraction data were collected and the x-ray crystal structure was solved by molecular replacement and refined to 2.2 Å resolution. The resulting EMB crystallographic model (PDB ID: 5W1A), which contains two copies of the motor domain, each bound to separate ELC proteins, has an R_{cryst} of 18.4% and R_{free} of 22.1% (Table 1).

The EMB crystallographic model

The EMB crystallographic model contains two copies of the *Drosophila* embryonic isoform of MHC into which could be modeled amino acids 8-807 (chain A) and 9-810 (chain C). Electron density corresponding to the loop 2 region was not well defined in either myosin motor domain. Therefore, MHC amino acid residues 628-641 (chain A) and 628-644 (chain C) were not included in the final crystallographic model (Supplementary Figure 1). EMB heavy chains A and C are each non-covalently associated with a separate copy of the *Drosophila* IFM ELC, the models of which contain amino acids 1-154 (chain B) and 5-155

(chain D), respectively. Each EMB motor domain also contains one molecule of citrate at its nucleotide binding site. A total of four glycerol molecules, introduced during cryopreservation, are contained within the final refined model as well as 1,063 ordered water molecules.

The x-ray crystal structure reveals the now familiar myosin motor domain fold (Figure 1E) [35]. This consists of five separate major structural elements: an N-terminal domain (defined for this study as EMB residues 21-200), the upper (residues 214-460 and 603-622) and lower (residues 468-597 and 648-708) 50 kDa subdomains, the converter (residues 709-767), and the extended α -helical lever arm (residues 768-807) (Figure 1B–E). The crystallized EMB lever arm contains only one IQ motif and interacts with one calmodulin-like ELC [36]. The absence of a second IQ motif and bound regulatory light chain (RLC), which in *Drosophila* carries a unique 46 residue N-terminal extension, is a result of treatment with chymotrypsin during motor domain fragment production and suggests exposure of a protease-sensitive site between the ELC and RLC. Similar sized S1-like proteolytic fragments have been prepared and studied previously by x-ray crystallography, including from chicken smooth muscle myosin II and myosin V proteins [37,38].

Conformation of the *Drosophila* myosin embryonic isoform

Proper functioning of myosin requires changes in the state of bound nucleotide and phosphate at the intersection of the N-terminal, upper, and lower 50 kDa subdomains (referred to throughout the remainder of this paper as “U50” and “L50,” respectively), which are relayed to the actin binding site at one extreme and to the converter of the lever arm at the opposite end of the motor domain (Figure 1E, F). The actin binding site is created upon juxtaposition of the 50 kDa subdomains and disrupted by their separation, which results in the formation of a large cleft that splits the actin binding site and divides the globular motor domain roughly in half [38,39]. On the other end of the motor domain, the converter connects the lever arm to the rest of the motor domain through the relay helix (EMB residues 472-503) and SH1 helix (residues 687-707). These two helices undergo dramatic conformational changes during the cross-bridge cycle that reorient the lever arm by $\sim 60^\circ$, moving the actin binding site by a distance of roughly 10 nm and accounting for the motile properties of myosin [40].

Despite having been built and refined independently of one another, the two MHC motor domain:ELC complexes contained within the EMB crystallographic asymmetric unit are nearly identical. The root-mean-squared deviation (rmsd) for C α positions after superposition of the two heavy chains is 0.588 Å² with the slight difference resulting primarily from movement of the entire U50 relative to the remainder of the motor domain (Figure 2). This view is supported by the fact that superposition of U50 subdomains only from heavy chains A and C results in an rmsd of 0.216 Å² while exclusion of the U50 during superposition results in an rmsd of 0.289 Å² for the roughly 520 amino acids that remain. Translation of the entire heavy chain A U50 by 1.1 Å and rotation about its center of mass by 10.6° places it almost perfectly on the similar region of heavy chain C. The variation in precise placement of U50 within the two EMB complexes arises as a consequence of the sum of relatively small differences in the polypeptide backbone phi/psi angles in the amino

acids that border the 50 kDa subdomains rather than any major observable change in backbone conformation at any single residue. Since the polypeptide crosses over between the U50 and NTD or L50 at four different places, the greatest differences in polypeptide backbone geometry occur at amino acids 212 and 455, again at residue 600, and within the disordered loop 2 (Supplementary Figure 1). Only one amino acid residue within proximity of these four apparent pivot points, Glu216, adopts a different side chain rotamer in heavy chain C relative to heavy chain A. Therefore, it appears that modest hinge movement within the polypeptide backbone at or near the U50 borders is responsible for the slightly different conformations of the two myosin motor domains observed in the refined EMB myosin crystallographic model.

During contraction of skeletal muscle, the MHC motor domain adopts different conformations as it passes through structurally and kinetically defined steps of the cross-bridge cycle. We identified the functional conformation of the EMB crystallographic model by the method described by Yang, *et al.* [41]. In this approach, distances and angles between common reference points within the motor domain are measured and used to classify conformation. Among these measures are the angles of either the U50 or L50 relative to a straight line connecting the Ca atoms of Arg148 at the base of the 50 kDa cleft and Asn188 near the nucleotide binding site as well as the distances between Ca atoms of conserved amino acid positions throughout the cleft that separates the 50 kDa subdomains (Table 2). In addition to comparing our structure to models exhibiting the well-established pre-power stroke, rigor-like, and post-rigor conformations, we analyzed 5JLH and 5H53, near atomic resolution cryo-electron microscopy structures of human and rabbit skeletal myosin, respectively, in complex with actin filaments [42,43]. These actomyosin models were included to establish cleft orientation and closure parameters for the unique nucleotide-free actin-bound conformation that we refer to in Table 2 as “rigor”.

Analysis of the EMB structure suggests that both copies of the myosin motor domain are most consistent with the “rigor-like” conformation. This stable conformation, which represents the high affinity actin-bound post-power stroke form of myosin, has been observed previously in other x-ray crystal structures of myosin motor domain proteins including squid myosin II, human nonmuscle myosin IIb, myosin II from the slime mold *Dictyostelium discoideum*, striated muscle myosin II from sea scallop, chicken myosin V, pig myosin VI, and myosin A from *Plasmodium falciparum* [38,41,44–47]. The signature features include a smaller angle between the U50, L50, and the N-terminal domain, greater closure of the cleft that separates the U50 and L50, and a “fully twisted” core seven-stranded β -sheet, known the transducer [48], that supports counterclockwise rotation of the U50 relative to the rest of the entire motor domain. In comparison to other rigor-like myosin crystallographic models, the slightly different conformations exhibited by the two MHC motor domains in the EMB x-ray structure represent extremes of the rigor-like conformation available to EMB (Table 2). Therefore, the small differences we observe in the two independent EMB models might capture the complete range of relative motion available to myosin motor domains within the rigor-like conformation. We note that the rigor-like *Dictyostelium* myosin II structure (PDB ID: 2AKA) differs from rigor to an even greater degree than the two EMB MHC heavy chains vary from one another [45]. It is not clear whether this suggests that an even greater range of motion is available to the *Dictyostelium*

protein in its rigor-like conformation or if the actual rigor state of this protein differs from that observed in actomyosin structures containing other myosins.

Interestingly, although the degree of cleft closure and relative orientations of its U50 and L50 observed in EMB clearly align it with other rigor-like models, the EMB actin binding site more closely resembles that of the human non-muscle myosin 2C in its rigor conformation taken from the 3.9 Å cryo-electron microscopy structure of its complex with polymerized actin (PDB ID: 5JLH) [43]. This is particularly noticeable by the placement and orientation of the cardiomyopathy loop (EMB heavy chain residues 401-414) after L50 superposition of MHC motor domain models in different conformations (Supplementary Figure 2). Previously, the only MHC crystallographic model reported to exhibit a rigor-like conformation so similar to the actual rigor conformation exhibited by actin-bound myosin was chicken myosin V [38]. Thus, we conclude that the conformation exhibited by the MHC motor domains in the EMB x-ray crystal structure is strongly rigor-like with its actin binding site in a conformation that is nearly identical to that of actin-bound myosin in its true rigor conformation.

Citrate at the nucleotide binding site

The EMB crystallographic model was determined from crystals grown in the absence of any known ligands and does not contain nucleotide at the nucleotide binding sites in either of the two MHC motor domains. However, from our earliest difference electron density maps we noted a strong positive peak over the P loop in both EMB nucleotide binding sites that persisted throughout the course of model building and refinement. The position of the peak corresponds well with the location of a bound sulfate ion in some previously determined myosin x-ray crystal structures [49,50]. However, sulfate was absent from EMB protein preparation, crystallization, or crystal stabilization buffers and sulfate did not refine well when placed over the experimental density. On the basis of its multiple negative charges, size, and the necessary inclusion of citrate at concentrations in the 10 mM range throughout the process of crystal growth and harvesting, we modeled one citrate molecule into the density at each of the EMB nucleotide binding sites. The citrates refined well with good fit to the $2F_o - F_c$ density, similar orientations with respect to both heavy chain A and heavy chain C motor domains, and *B* factors that fit the profile for the model overall and that are in good agreement with amino acids that reside within range of contact. Omit and polder maps support our assignment of citrate to this nucleotide binding site density (Supplementary Figure 3) [51].

The bound citrate makes contact with EMB amino acid residues within the glycine-rich P loop (residues 180-184). The combination of rotational versatility available to glycines 182 and 184 and initiation of an α -helix by residues Lys185, Thr186, and Glu187 affords the large, polyanionic citrate ion residence within proximity (~ 4 Å) of six consecutive amide nitrogen atoms from Gly182 through Glu187. Comparison with the crystallographic model for scallop myosin II (PDB ID: 1S5G), also in a rigor-like conformation but with ADP and sulfate ion bound at the nucleotide binding site, reveals that the larger citrate polyanion fills space occupied by both the sulfate and the α -phosphate of ADP in the scallop myosin (Figure 3 and Supplementary Figure 4) [50]. The nearby switch I (EMB amino acids

233-243) resides in close proximity over switch II (amino acids 461-471) exhibiting an interaction that has been shown to correlate with the high degree of transducer twist and cleft closure that characterize the rigor-like conformation [41]. Interestingly, the previously studied nucleotide-free myosins that adopt rigor-like structures were purified in the presence of Mg-ADP that was subsequently removed via dialysis prior to crystallization. The EMB x-ray crystal structure therefore suggests that the larger citrate polyanion might serve to stabilize myosin for crystallization in its rigor-like conformation independent of bound nucleotide.

Essential light chain structure

Electron density calculated from $2F_o - F_c$ difference maps with the final refined model coordinates is excellent for both copies of ELC protein in the asymmetric unit (Supplementary Figure 5). This allows for a detailed analysis of *Drosophila* ELC structure and its interactions with the MHC lever arm. The ELC adopts its familiar calmodulin-like fold. The entirely α -helical protein wraps around one IQ motif within the C-terminal stalk-like lever arm of MHC between amino acids Arg776-Leu803. It also makes contacts with several amino acids in the range of Lys720-Pro728 within the converter (Supplementary Figures 1 and 6).

Structural comparison of EMB with scallop striated muscle myosin in its pre-power stroke conformation (PDB ID: 1QVI) and the rigor-like chicken smooth muscle myosin V, which was crystallized in complex with a human ELC, (PDB ID: 1OE9) reveals that their respective ELC proteins are highly similar with rmsd for C α positions after superposition of 2.12 Å (EMB vs. 1QVI) and 1.77 Å (EMB vs. 1OE9) [38,52]. The first of two significant structural differences is that *Drosophila* ELC diverges significantly in its sequence at the end of its most N-terminal α -helix (Supplementary Figure 6). The amino acid sequence between residues Gly20-Gly25 of *Drosophila* ELC is three residues shorter than the scallop ELC and one residue shorter than human ELC. The result is that its first helix is shortened by one entire turn and lacks the conserved Asp residue that is necessary for calcium binding in scallop ELC [53]. The loop that follows helix 1 connects to helix 2 at precisely the same point as in the other ELC proteins and contacts to MHC lever arm residues throughout this region are not altered as a consequence of this missing turn of helix (Figure 4A). Following the third helix, *Drosophila* ELC closely resembles the scallop protein, while human ELC adopts a short alpha helix between its amino acids 56-62.

A second aspect unique to the structure of *Drosophila* ELC involves its extreme C-terminal region. As is also the case with muscle myosin II, the *Drosophila melanogaster* genome contains only one ELC gene and relies upon alternative splicing to generate mRNA isoforms. *Drosophila* have only two ELC protein isoforms. One is reserved exclusively for IFM myosin while the other is employed in all other muscles. The two ELC protein isoforms differ in the amino acid sequence of twelve amino acids at their respective C-termini (Supplementary Figure 6) [54]. Due to the fact that our expression system relies upon targeting recombinant MHC expression to adult IFM, the EMB x-ray crystal structure contains ELC with the 12 C-terminal amino acid residues unique to the *Drosophila* IFM isoform.

The C-terminal 12 amino acids adopt a configuration that appears unique among ELC protein structures. The alternative C-terminus associated with flight muscle ELC contains the motif Arg150-Pro151-Asp152-Gln153, which in the ELC adopts a hook-like protrusion that is stabilized by two nearly ideal hydrogen bonds between its side chain carboxylate oxygens and atoms Ne and Nη of Arg150 and a third hydrogen bond between Asp152 and the Ne atom of Gln 153 (Figure 4B). Pro151 appears to play a critical role in orienting the polypeptide to allow for this conformation. Both copies of the ELC present in the EMB crystallographic model adopt identical conformations at their respective C-termini, although thermal *B* factors for this region in light chain D are considerably lower than for the same region in light chain B. This is almost certainly a consequence of the involvement of the C-terminus of light chain D in mediating close packing with a crystallographic neighboring complex, while in the crystal light chain B contacts only solvent. It bears mentioning that *Drosophila* RLC, which is not present in the EMB crystallographic model, contains a 46-amino acid N-terminal extension, not observed in other RLC proteins, and which has been shown to be required for optimal power output and contraction frequency in chemically skinned IFM fibers [55]. It is possible that the unique structure of the ELC C-terminal tail mediates interactions with RLC that support IFM mechanics. It is also worth noting that the *Drosophila* ELC isoform expressed in all other muscles exhibits an altogether different and significantly less polar amino acid sequence at its C-terminus (Supplementary Figure 6).

Alternative exons unique to EMB

The EMB x-ray crystal structure confirms that the MHC and ELC from insect muscle, as represented by an isoform present in *Drosophila* embryonic body wall muscle, bear all the hallmarks of a typical conventional myosin motor domain. As such, the *Drosophila* genetic system appears well suited for measuring structural, biochemical, physiological, and phenotypic consequences of muscle disease-causing mutations from diverse species, including humans. Interestingly, the *Drosophila* system presents an ideal platform for studying myosin structure/function as alternative selection of four variable exons can produce the EMB motor domain or that of IFI, which is the extremely rapid acting myosin II isoform associated with adult IFM that powers flight. To begin identifying how amino acid changes at particular locations might contribute to the drastically altered physiological properties of diverse insect myosin isoforms, we analyzed the structure of portions of EMB encoded for uniquely by exons 3a, 7a, 9b, and 11c by comparing them with similar portions of bay scallop striated muscle myosin in both its detached and pre-power stroke conformations (PDB ID: 1KK8, 1QVI), *Dictyostelium* myosin II in its rigor-like conformation (PDB ID: 2AKA), and post-rigor chicken skeletal muscle myosin II (PDB ID: 2MYS) [45,49,52,56].

Exon 3a—Exon 3a encodes EMB amino acids 69-116. In total, there are eighteen amino acid differences within this span compared to exon 3b, which is employed in IFI (Supplementary Figure 1). By swapping the regions encoded for by exons 3a and 3b and expressing the resulting chimeric proteins in flies against the *Mhc*¹⁰ background, previous studies revealed that introduction of exon 3b significantly increased the *in vitro* actin sliding rate in the resultant chimeric EMB protein. However, this replacement had little effect on EMB solution ATPase rates. Flies expressing the EMB-exon 3b chimera exhibited improved

flight muscle ultrastructure and increased fiber power output relative to EMB flies harboring the native exon 3a, though that was not enough to rescue their flightless phenotype. Interestingly, incorporation of the exon 3a encoded region into IFI resulted in generally decreased rates of ATPase activity, power output, and flight ability while not affecting the *in vitro* actin sliding rate of the IFI-exon 3a chimera relative to IFI [13,57].

The EMB x-ray crystal structure reveals that the portion of the protein encoded for by exon 3a begins within the final β -strand of the SH3-like β -barrel fold within the N-terminal domain and snakes its way through a long loop and α -helix before ending within the first β -strand of the transducer (Figure 5A). Comparison of our two copies of the EMB motor domain reveals that they are virtually identical in structure throughout this region, with the only discernable differences being altered side chain conformers exhibited by Ile72 and Arg107. Furthermore, the backbone structure exhibited by EMB residues 69-116 is the same as that in homologous regions of scallop striated muscle myosin in both its detached and pre-power stroke conformations, rigor-like *Dictyostelium* myosin II, and post-rigor chicken skeletal muscle myosin II (Figure 5A).

Analysis of the amino acid substitutions that distinguish EMB from IFI in the exon 3 encoded region yields two clear observations. First, most of the amino acid differences occupy surface exposed positions. Second, almost all substitutions are conservative in nature. Three exceptions are Lys73, Val96, and Ala112 encoded by exon 3a, which in exon 3b are Leu, Tyr, and Asn, respectively (Supplementary Figure 1). Therefore, it is not obvious from the structure of *Drosophila* EMB how substitutions at any of the eighteen variable positions in the region encoded by exon 3 might influence the biochemical and/or physiological properties of embryonic body wall versus indirect flight muscle.

Exon 7a—Exon 7a encodes EMB amino acids 298-332. There are twelve amino acids within this stretch that differ from the exon 7d encoded region employed by IFI. As a consequence of its location within the outer surface of the U50 subdomain directly above switch I, the exon 7 encoded region has been referred to as an “upper lip” over the nucleotide binding site [58]. Swapping exons 7 between EMB and IFI and expressing the resulting protein chimeras in *Mhc¹⁰* flies had the intriguing consequence of increasing the basal and actin-dependent ATPase rates of each chimera relative to either EMB or IFI of native sequence. However, neither change had any significant effect on actin filament sliding rates [59]. Incorporation of exon 7d significantly increased maximum power generation and active stiffness in EMB muscle fibers, as well as improving muscle ultrastructure [60]. Taken together, and in conjunction with transient kinetic measurements including ADP release rates, these studies suggested that the exon 7 encoded portions modulate the kinetic rates of transition between different actin-bound states [58].

Structurally, the amino acids within the region encoded by exon 7a constitute an extended, ordered surface loop that begins and ends within α -helical segments. Other than an alternative side chain rotamer conformation adopted by Glu326, the structures of both myosin chains in the EMB model are identical throughout this region. The polypeptide backbone fold is shared among the other myosins used for comparison in this study. A slight difference in the placement of amino acids 298-309 is observed relative to *Dictyostelium*

myosin II. As is the case for the exon 3 encoded region, most of the amino acid differences between EMB and IFI in the region encoded by exon 7 are conservative in nature. One exception is Ile311, which is one of several residues whose side chain is buried in the hydrophobic core that appears to hold this loop in place. This position is conserved as a Phe, Tyr, or Leu in other myosins, however, it is replaced by the small polar Asn residue in IFI (Supplementary Figure 1). Another curious substitution is the IFI amino acid Gly in place of Ala325 in EMB. None of the other myosins studied contain a Gly at this position. Were the loop of the exon 7 encoded region to adopt an alternative fold in IFI, then it is possible that this unique Gly position could function to relieve backbone strain without affecting the structure of subsequent portions of the molecule.

Exon 9b—Exon 9b, which encodes EMB amino acids 469-525, exhibits the lowest degree of primary sequence variability in comparison with IFI-specific exons, as only five residues differ from exon 9a. Exon 9 encodes for the entire relay (Supplementary Figure 1). This well-studied component of myosin motor domains is one of the few structural elements that undergoes significant conformational change during the cross-bridge cycle [61]. Replacement of exon 9b with 9a in transgenic flies expressing EMB within the *Mhc¹⁰* background resulted in severe muscle assembly defects and degeneration. The resulting chimeric protein exhibited a significant decrease in actin-activated Mg-ATPase activity relative to EMB, showed increased actin binding affinity, and failed to move actin filaments by *in vitro* sliding assay. In contrast, fibers expressing this chimera produced increased power compared to EMB, possibly due to increased strain sensitivity [62]. Interestingly, there was almost no difference in flight ability or muscle structure observed in transgenic flies expressing IFI and a chimeric version of IFI containing exon 9b, although fiber studies showed a decrease in maximum power production [62,63]. Steady-state and transient kinetics measurements on S1 fragments revealed that reciprocal substitution of exons 9 between EMB and IFI resulted in decreased rates of ATP-induced actomyosin dissociation and decreased affinity for ADP in both chimeras [64].

The exon 9 encoded relay is composed of a long α -helical relay helix (residues 472-503) followed by an ordered loop (residues 504-515) and a small α -helix (residues 516-523). It protrudes from the base of the L50 and contacts the converter (Figure 5C). Structures of the myosin head in its pre-power stroke conformation revealed uncoiling and bending of the relay helix at a point that corresponds approximately to Met492 of EMB due to interaction with the SH1 helix (residues 699-708 in EMB) [52]. This movement, which occurs as a result of changes in nucleotide binding conveyed through the transducer, is transmitted through the converter to the lever arm in preparation for high affinity actin binding and the power stroke [35].

The two independent molecules in the EMB x-ray crystal structure display significant differences in their respective peptide backbone geometries within the range of Asp506 to Ile510 as well as alternative side chain rotamers at Met514, both within the relay loop. These differences arise almost certainly as a consequence of crystal contacts between neighboring myosin heads in the crystallographic asymmetric unit (discussed below). However, it is worth noting that the differences include residues Asp506 and Ala508, which are two positions that differ in exon 9a of IFI and that bracket Trp507. Despite extremely

high sequence conservation throughout the relay segments across diverse classes of MHC, these two positions display a high degree of variability and both differ between EMB and IFI, with IFI containing Asn506 and Asp508 (Supplementary Figure 1). The other three differences between EMB and IFI in the region encoded for by exon 9 reside within the relay helix. His491, which in the rigor-like EMB model is slightly buried against β -strand 4 of the transducer in the N-terminal domain, is an Ile in IFI, whereas solvent exposed EMB residues Leu495 and Arg502 undergo conservative changes in IFI to Met and Lys, respectively.

Exon 11c—The fourth and final alternative exon in the EMB motor domain coding region is exon 11c, which encodes amino acids 723-761 comprising the majority of the converter (Supplementary Figure 1). This is by far the most chemically distinct portion of EMB relative to the IFI motor, as there are 26 amino acid differences, including a stretch between Lys729-Pro752 in which 21 of 24 residues differ. The converter is part of the lever arm, which moves between two extreme conformations during the cross-bridge cycle. In light of this fundamental role, it is not surprising that amino acid differences within the region of the converter have a significant effect on MHC function. Swapping IFI exon 11e for 11c and expression of the resulting chimeric EMB protein in the IFM-null *Mhc*¹⁰ background partially rescued the myofibril deterioration phenotype of EMB. Muscle fibers isolated from these flies generated more than double the power at double the maximal power frequency of native EMB. In contrast, introduction of EMB exon 11c into the IFI coding region reduced the power output by one-third, with a maximum at half the frequency of IFI. These data suggest that the respective contribution of the converter region encoded by exon 11 to motor domain kinetics is both sequence and context dependent [14,65]. Recent studies on *Drosophila* myosins expressing five alternative exon 11 versions suggest that this region directly influences kinetics of actin detachment and the power stroke [66].

Structurally, the exon 11c encoded region begins near the end of the first of two α -helices in the converter. An extended ordered loop that includes one turn of helix connects to a second α -helix (residues 737-747) and then continues through the edge strand of a three-stranded anti-parallel β -sheet. Most of the amino acid differences between exons 11c and 11e map to this converter loop-second helix-loop motif, which is surface exposed and distant from the rest of the molecule. Comparison of the two independent motor domains in the EMB crystallographic model reveals alteration of the backbone structures between amino acids 749-753. This difference arises as a result of crystal contacts between Asn151 and Asp153 in EMB heavy chain C and the ELC from a crystallographic symmetry-related complex, which causes a 180° flip of the psi angle in Thr748. The polypeptide backbone of EMB differs significantly between amino acids 727-736 from the *Dictyostelium* myosin II model (PDB ID: 2AKA) and from chicken skeletal muscle myosin II (PDB ID: 2MYS) and scallop striated muscle myosin (PDB ID: 1QVI and 1KK8), which each contain breaks in the polypeptide chain between equivalent amino acids 730-735 (Figure 5D). With respect to differences in the amino acid sequence within the regions encoded by exon 11c in EMB and exon 11e in IFI, near the beginning of this stretch there are two glycine residues in EMB (Gly730 and Gly733) that are Ile and Ala in IFI. IFI contains only one glycine residue at position Gly742. It is possible that the distinctive placement of glycine residues in this

region supports unique structural variations between regions encoded by exons 11c and 11e and influence the structure, flexibility, and/or dynamics of the respective myosin isoforms that contain them.

The exon 11 encoded segment is also unique in contacting both the relay (encoded for by exon 9) and the ELC. Both of the portions encoded for in exon 11 that mediate these contacts are highly conserved between EMB and IFI (Supplementary Figure 1). However, altered properties of the converter that arise as a consequence of differences in EMB exon 11c and IFI exon 11e could influence myosin action at a distance through a network communicating the nucleotide binding state to the lever arm through exons 9, 11, and the ELC [67]. Mutation of Arg756 from the converter to Glu was previously shown to correlate with a 60% decrease in ATPase rates and 35% decrease in actin sliding motility as well as loss of flight ability without affecting actin binding affinity. Even more intriguingly, the Arg756Glu mutant failed to exhibit increased inherent fluorescence by Trp507 from the relay upon binding to ATP [68–70]. Subsequently, mutation of Asn506 to Lys was shown to suppress these defects [70,71]. The EMB x-ray crystal structure also reveals how Arg756 interacts with the side chain of Ile505 in the rigor-like conformation. It bears mentioning that the electron density around Trp507, particularly for heavy chain C, is one of the very few regions of ambiguity in the otherwise exceptional electron density maps generated during the process of EMB crystallographic refinement (Supplementary Figure 7). In light of these observations, it seems reasonable to conclude that the ATP-dependent changes in Trp507 fluorescence likely result from structural rearrangement in this region and that the residues that frame Trp507, including the variable positions 506 and 508 from the relay and Arg756 from the converter, serve to tune the kinetics of this structural transition.

Asymmetric dimer

The interaction between the two EMB molecules within the crystallographic model has both molecules stacked next to one another such that they are nearly parallel with their C-terminal lever arm and ELC both projected in one direction and their actin binding sites poised adjacent to one another at the other end (Figure 6). The two molecules are not identically disposed as one is rotated slightly about its long axis. Superposition of heavy chain A and light chain B onto chains C and D requires a translation of 56.3 Å, or roughly one half the unit cell along the *a* axis, and rotation of 15.6° about the center of mass of the complex.

The surfaces on the two myosin heads that mediate their interaction involves close packing contacts of amino acids from three of the regions encoded by EMB-specific exons. One molecule employs residues from the region encoded by exon 7a to contact the second EMB motor domain through amino acids encoded by exons 9b and 11c. There is not an extensive network of specific interactions nor is there a significant amount of surface area buried upon interaction. Of the residues that participate in this interaction, only Ala508, encoded for by exon 9b, is unique to EMB (Supplementary Figure 1).

The arrangement of myosin heads in the EMB asymmetric dimer is interesting in that the two MHCs have globular head domains oriented with their actin binding surfaces adjacent to one another, both nucleotide binding sites accessible to solvent, and a structurally dynamic segment from the relay resides at the interface. However, we know of no experimental

evidence to suggest that this orientation of myosin motor domains is physiologically relevant.

Molecular dynamics comparison of *Drosophila* EMB and IFI

In an effort to begin to identify the consequences of differences at amino acid positions encoded by the alternative exons in *Drosophila* EMB and IFI motor domains, we performed atomic resolution molecular dynamics simulations. For this purpose, a single, nucleotide-free, continuous polypeptide EMB heavy chain model (chain C) was prepared by removing the lever arm and ELC (chain D) and adding residues 628-644 to close the discontinuity in loop 2. At this point, a second model was prepared by mutating the 61 amino acid side chains that differ between EMB and IFI. During this process, the placement and orientation of the mutated amino acid was kept the same as in EMB unless this resulted in clashes, in which case the highest probability side chain rotamer was selected. Steepest descent followed by conjugate gradient energy minimization was then carried out on each structure yielding the starting models for the molecular dynamics. Excepting differences at the 61 amino acid positions encoded for by alternative exons 3, 7, 9, and 11, the resulting EMB and IFI models are practically identical with less than 0.010 Å average rmsd for C α positions upon superposition. The models were subjected independently to 510 ns of molecular dynamics simulation using explicit solvent and the AMBER 14SB force field [72]. All simulations were run with the GPU accelerated version [73–75] of the AMBER software package (v14) [76,77].

There were no gross structural changes over the course of the simulation for either EMB or IFI. Rather, small and reversible structural rearrangements were observed throughout the models. An analysis of the average root-mean-squared fluctuation (rmsf) for C α positions throughout the time course of the simulation highlights the positions of some of the more mobile regions in the respective models (Figure 7A). By far, the most conformationally variable portion in both isoforms resides between amino acids 626 and 645. This corresponds to the loop 2 region that was absent from the EMB crystallographic model and needed to be built by hand prior to running the simulations. This loop, which resides near the actin binding site and contains the low complexity sequence GQSGGGEQAKGGRGKKGGG, is likely highly mobile. It also corresponds to a well-known site of *in vitro* proteolytic sensitivity that separates the central 50 kDa fragment from the C-terminal 20 kDa fragment that harbors the converter and lever arm [78].

Another observation from the dynamics data is that the relative positional variation throughout the two models is roughly similar until approximately amino acid position 500, after which the IFI model displays consistently higher average fluctuation (Figure 7B). This C-terminal portion of the molecule contains the amino acid sequences encoded for by exons 9 and 11 including the relay portion of L50 and the converter. Interestingly, when the root-mean-squared deviation (rmsd) for all C α atoms within the portions of the models encoded by exon 11 (amino acids 723-761) are compared over the time of the simulation, we observe an immediate and prolonged increase in the IFI model relative to EMB (Figure 7C). Thus it appears that the region encoded by exon 11 in IFI, which corresponds to a major portion of

the converter, is particularly prone to movement and/or structural rearrangement within the context of the experimental variables tested in this simulation.

As judged by measurement of change in the angle and degree of opening of the 50 kDa domain cleft, neither the EMB or IFI models undergo significant change from the rigor-like conformation throughout the course of the simulation. However, both models oscillate between two degrees of cleft opening. These two states correspond well with the differences in cleft opening observed in the two independently refined motor domains of the EMB crystallographic model, suggesting that these two slightly different arrangements might represent the limits of cleft opening and closure within the stable rigor-like conformation exhibited by the protein (Figure 2). Interestingly, we observe in the dynamics simulations that IFI oscillates more frequently between these two extremes, as evidenced by the increased frequency in change of distance between points across the 50 kDa cleft in IFI relative to EMB (Figure 7D). When combined with the observed increased rmsd of the IFI converter, this difference in the rate of cleft opening and closing suggests that amino acid differences within the relay and converter resulting from selection of alternative exons 9 and 11 could affect *Drosophila* myosin heads in ways that alter their biochemical properties (Figure 8).

DISCUSSION

In this study, we report the x-ray crystal structure of the *Drosophila melanogaster* embryonic isoform of the skeletal muscle myosin II motor domain and ELC at 2.2 Å resolution. The electron density maps are excellent, yielding a high-resolution model of overall exceptional quality (Supplementary Figure 5). Our purpose in determining the three-dimensional structure of a skeletal muscle myosin II protein from *Drosophila* is three-fold. First, the study serves as a validation of our novel method for recombinant expression and purification of myosin motor proteins from whole transgenic fruit flies at levels that can support biophysical and high-resolution structural studies. Second, as we have employed *Drosophila* to model deleterious effects in muscle function and ultrastructure that result from myosin mutations that are linked to human muscle diseases [17,22], it is important that we confirm that the *Drosophila* myosins are a valid surrogate for modeling myosin-based human disease and hence studying structural changes that might result as a consequence of disease-causing mutations. Third, insects make up roughly 80% of the known animal kingdom and myosin II plays diverse, critical roles during each insect developmental stage [79]. Notwithstanding this fact and despite numerous published molecular studies on insect myosin structure and function, no high-resolution x-ray crystal structure of an insect myosin motor domain has been determined previously.

We employed recombinant DNA and *Drosophila* genetics technologies to engineer a transgenic fly line robustly expressing EMB exclusively in the IFM of adult flies that lack the endogenous protein [30]. Inclusion of an engineered, cleavable histidine-tag resulted in relatively straightforward purification of a functional MHC protein at the milligram scale from which diffraction quality crystals could be prepared. Therefore, one important conclusion we draw from this study is that the *Drosophila melanogaster* IFM expression system is well suited to support structural and biophysical studies. We propose it as a cost-

effective strategy for the recombinant expression and purification of diverse myosin protein isoforms and mutants.

The EMB x-ray crystal structure that we report in this study conforms well with the existing standard structural understanding of myosin motor proteins. This suggests that the *Drosophila* EMB system is viable for modeling mutations that have been discovered to contribute to musculoskeletal or cardiovascular diseases in other organisms, including humans. On this topic, it is important to acknowledge that modeling MHC mutations across organisms can yield contradictory results, as has been observed for the R403Q mutation identified in human β -cardiac myosin [80–82]. Moreover, myosin of the same type from different organisms can have different biochemical properties based upon their structural differences [83]. Consequently, modeling human disease-causing mutations in *Drosophila* MHC proteins requires multidimensional approaches that combine structural biochemistry with computational modeling and physiological measurements. Along these lines it is worth noting that, with respect to human β -cardiac myosin, X-ray crystallography combined with molecular dynamics has led to predictions that were confirmed by *in vitro* biochemical studies with the recombinant protein [84,85]. In the absence of crystallographic models for *Drosophila* IFI, we cannot speculate on whether it or EMB might be a preferred system for modeling human disease mutations in flies.

The myosin motor domain is an ancient fold and appears to have changed little throughout the course of evolution [35]. This is evidenced by the startlingly high degree of structural homology in myosin motor proteins of different functional classes and from species as diverse as humans, mollusks, and slime molds. In large part, this conservation of structure can be explained by the fact that the actin-activated ATP-dependent mechanism for myosin motility is a highly concerted process to which numerous individual structural entities from throughout the entire roughly 800 amino acids long motor domain contribute [35]. The series of mutational studies to which we refer throughout this paper in which individual exon encoded segments are swapped between *Drosophila* myosins with different kinetic and power producing properties support this thinking. Whereas, swapping of exons usually, but not always, affects phenotype, muscle ultrastructure, and enzymatic properties of the resulting chimera, never is there observed a complete gain or loss of any measurable function that is attributable entirely to one particular exon or structural element. Therefore, context and allosteric communication among the individual subdomains and structural elements of myosin appear to be paramount for proper functioning of the resulting motor proteins.

A possible exception to this strict rule of myosin motor domain structural conservation might be the MHC of insect IFM. This specialized tissue actuates the extremely fast contraction of skeletal muscle to allow for rapid wing beat and enable flight. In *Drosophila*, IFM is capable of supporting wing beat frequencies of up to 220 Hz and in other insects that number can reach as high as 1000 Hz [86]. Many investigations into insect IFM structure have focused upon three-dimensional image reconstructions from electron microscopy of muscle from giant water beetles of the genus *Lethocerus* [87]. This ongoing work recently yielded a 6 Å resolution cryo-electron microscopy model of the whole thick filament from *L. indicus* asynchronous flight muscle in its relaxed state [88]. A novel arrangement between

the two heads of the MHC dimer, a clear view of MHC rod packing, and additional non-myosin muscle proteins including the central paramyosin core are all observable at this resolution. The model is not, however, of sufficient detail to reveal whether or not flight muscle myosin relies upon a uniquely folded motor domain.

While the *Drosophila* IFI is specially adapted to support flight, its motor domain differs in amino acid composition from EMB only as a consequence of the selection of alternative exons 3, 7, 9, and 11. Our analysis of sequence differences between *Drosophila* EMB and IFI within these alternative exon-encoded regions suggested several points where amino acid substitutions could alter the structure of IFI relative to EMB and other myosins in ways that are as yet unprecedented. Interestingly, we observe that portions of the protein encoded for by three of these exons, 7, 9, and 11, participate in mediating interactions between the two MHC motor domain:ELC complexes in our EMB crystallographic model.

It is important to clarify that the dimer we observe crystallized in the EMB asymmetric unit is not the previously well-characterized interacting head motif (IHM) dimer. The IHM is adopted by neighboring motor domains of myosin in its relaxed state and involves ATP-bound myosin proteins in their pre-power stroke conformation [89]. Originally characterized in inhibited forms of smooth muscle myosin, the IHM has since been observed in skeletal muscle where, by burying the actin binding site of one motor domain and occluding the nucleotide binding site of the other, this asymmetric association of motor domains is thought to facilitate relaxation of muscle fibers [90–92]. The IHM (as well as the EMB crystallographic dimer) employs the continuous surface generated by the juxtaposition of relay and converter (encoded for in *Drosophila* MHC by exons 9 and 11, respectively) [89]. As mentioned previously, this portion of the relay, in the immediate vicinity of Trp507, is one of the only portions of the EMB model in which the two motor domains differ structurally from one another. It is also an area of significantly poorer electron density, which suggests high conformational variability (Supplementary Figure 7). Conformational plasticity and propensity to associate in protein-protein interactions are hallmarks of protein surfaces that engage in allosteric regulation of enzyme activity [93,94]. It seems likely that, in addition to occluding the actin and nucleotide binding sites, adoption of its IHM conformation disrupts relay dynamics and inhibits myosin ATPase activity. In contrast to the IHM noted in various relaxed muscles, cryo-electron microscopy of *Drosophila* IFM in its relaxed state reveals greater disorder and a lack of repeating structure of MHC head domains [95]. It is possible that organization into the stable IHM conformation is not a requirement of the extremely rapid *Drosophila* IFM and that unique structural differences in the IFI myosin head could preclude stable IHM formation. It is important to note, however, that both EMB and IFI are capable of adopting the IHM as isolated molecules, although the latter appears less stable [92].

To experimentally identify portions of IFI that might differ as a result of the changes in exons, we performed all atom molecular dynamics simulations on identically prepared nucleotide-free myosin motor domain models each with amino acids encoded for uniquely by the alternative exons specific for their respective isoform (Figure 7). This study was never intended to simulate function of the proteins in the context of skeletal muscle. Rather, it served to expose regions wherein the differences in amino acid sequence might cause altered

dynamics on account of chemical and/or structural differences. According to the data obtained through this approach, the regions encoded for by exons 9 and 11, effectively the relay and the converter, respectively, were the areas most affected. We observed evidence of increased local and global dynamic rates in the model bearing the IFI sequence relative to EMB, particularly for the converter region and also for the oscillation of actin binding cleft opening and closing (Figure 8). It remains to be seen whether the extreme functional differences between *Drosophila* IFI and EMB arise simply as a result of altered rates of individual steps within the cross-bridge cycle or if unique structural rearrangements specific to IFI also contribute to the phenomenon of extremely rapid skeletal muscle contraction associated with this MHC isoform. Currently, we are focused on experimentally determining high resolution crystallographic models of *Drosophila* IFI at different stages of its mechanochemical cycle. In the future, we plan to use our crystallographic models to support computational and physiological studies aimed at uncovering how this motor protein is uniquely adapted to enable insect flight.

MATERIALS AND METHODS

Expression and purification of *Drosophila* skeletal muscle myosin II embryonic isoform subfragment-1

The *Drosophila melanogaster* embryonic isoform of skeletal muscle myosin II heavy chain (EMB MHC) was expressed in transgenic flies using a system developed previously [30] and described briefly here. An artificial gene encoding a TEV protease-cleavable N-terminal His-tag and cDNA for myosin II heavy chain exons 1-12, including alternative exons specific for EMB (3a, 7a, 9b, 11c), was combined with the genomic DNA for the myosin rod and placed under the *Actin88F* promoter in a pattB transfer vector [96]. The plasmid was injected into fly embryos (BestGene Inc.) and the resulting fly line was crossed with *Mhc¹⁰*, a fly line null for endogenous myosin in the indirect flight muscle and jump muscle, yielding a homozygous strain of flies expressing the His-tagged embryonic MHC isoform [32]. The *Drosophila* EMB MHC was prepared via extraction from ~10 g whole flies in high salt (0.5 M NaCl) buffer. Two modifications to the published protocol improved yield when handling such a large amount of starting material: First, initial homogenization was done in 3 parts with ~3.3 g of flies in 30 mL homogenization buffer (12.5 % (w/v) sucrose, 40 mM NaCl 10 mM imidazole-HCl, 2 mM MgCl₂, 0.2 mM EGTA, 0.5 % (v/v) Triton X-100, 1 mM dithiothreitol, Roche protease inhibitor cocktail) using a Wheaton dounce homogenizer with the metal shaft of the Teflon pestle mounted in a DeWalt cordless drill. The pestle was rotated slowly for three minutes while the 30 mL glass mortar tube was raised to mechanically lyse the fly tissues and cells by shear stress. Second, after low salt precipitation of myosin, 10 mL of resuspension buffer (20 mM sodium phosphate, 1.5 M sodium chloride, 30 mM imidazole-HCl, pH 7.4) was added to each of six centrifuge tubes and the pellet was carefully scraped off the tube with a pipet tip. After the pellet became completely dislodged, the mixture was decanted into a chilled beaker and placed on a stir plate for five hours for complete dissolution of the pelleted material. His-tagged myosin was filtered and purified in high salt by Ni affinity chromatography on a 5 mL HisTrap column followed by treatment with alpha-chymotrypsin to generate S1 fragments that, after precipitation of the rods from the solution, were concentrated and purified by size exclusion

chromatography on a SuperDex 200 16/600 prep grade column (GE Healthcare) before concentration to between 3-10 mg/mL and flash freezing in liquid nitrogen for long term storage at -80°C .

Crystallization

Initial EMB microcrystals were grown by hanging drop vapor diffusion and optimized in sitting drop format using Slice pH and Silver Bullets additive screens (Hampton Research). A TTP Labtech Mosquito robot and MRC 2-well crystallization plates (Swissci) were employed to set up 200 nL drops composed of 100 nL protein and 100 nL reservoir solution (0.032 M citric acid, 0.049 M BIS-TRIS propane (pH 7.15), 17.8% (w/v) PEG 3,350, 0.09 M Tris-HCl (pH 8.1) 0.02% (w/v) Leu-Gly-Gly, 0.02% (w/v) Leu-Leu, 0.02% (w/v) Met-Ala-Ser, 0.02% (w/v) Ala-Ala-Ala, 0.02% (w/v) Gly-Gly-Gly, 0.02% (w/v) Trp-Gly-HCl, and 2 mM Na-HEPES (pH 6.8)) sealed with 80 μL of reservoir solution. A single $30 \times 30 \times 140 \mu\text{m}$ crystal grew during two weeks of incubation at 20°C . The crystal was harvested from its mother liquor with a nylon loop (Hampton Research) and soaked for 10 seconds in cryo-protectant stabilizer solution consisting of the reservoir solution and 20% ethylene glycol before flash cooling in liquid nitrogen.

X-ray diffraction data collection and processing

X-ray diffraction data were collected in continuous mode with a Pilatus-6MF detector at the Northeastern Collaborative Access Team (NE-CAT) beamline (24-ID-C) located at the Advanced Photon Source in Argonne National Laboratory. The unit cell parameters ($a = 108.55$, $b = 148.58$, $c = 148.73 \text{ \AA}$; primitive orthorhombic Bravais lattice) suggested two protein subunits in the asymmetric unit (53.0 % solvent content; Matthews coefficient $V_M = 2.62 \text{ \AA}^3$). Diffraction data were processed and individual reflection intensities integrated and scaled by XDS within the RAPD suite. Systematic absences along a , b , and c indicated space group $P2_12_12_1$. Data collection statistics are reported in Table 1.

Structure solution and refinement

The crystal structure was partially solved by molecular replacement using the online BALBES server [97], which returned a starting model containing, among its six polypeptide chains, one with more than 60% of one motor domain. After deletion of the five incorrectly placed amino acid chains, this initial model was used as a probe to identify the placement of a second molecule within the asymmetric unit in MolRep [98]. After rigid body and restrained refinement of the two motor domains in Refmac5 (50-3.1 $\text{\AA}$ data, group B) the lever arm and one ELC (chain D) was built into $2F_O - F_C$ electron density with Coot [99,100]. ELC chain B was built by overlaying heavy chain C and light chain D on heavy chain A. Refinement against all data was carried out with Refmac5 and Phenix [101,102] and additional model building was completed using Coot and model validation included monitoring stereochemistry with MolProbity [99,100]. The resulting model, which contains two complete copies of the EMB motor domain each bound to its own essential light chain (ELC), has an R_{cryst} of 18.4 % and R_{free} of 22.4% with exceptional stereochemistry for an asymmetric model of this size (Table 1). Figures were created in PyMol

Generation of EMB and IFI motor domain models for MD simulations

Coordinates for the EMB motor domain chain C were prepared for computational molecular dynamics simulations by truncating the myosin lever arm after residue Asp777. The missing loop 2 residues 628-644 were built *de novo* in Coot as poly-alanine, except where glycine is present in the protein primary sequence, and modeled as an extended polypeptide with correct stereochemistry. Coordinates for IFI were generated by replacement of side chains in the EMB coordinate file to their IFI sequence.

Molecular dynamics simulations of EMB and IFI motor domain models

The AMBER FF14SB force field was used for modeling the protein sequences for the all atom molecular dynamics simulations [72]. Each system was solvated in an orthorhombic box of TIP3P water molecules such that no solute atom was within 10 Å of any box edge [104]. Sodium or chloride ions were placed randomly within the solvent box to neutralize the system. Energy in each system was minimized using 2,000 steps of steepest descent followed by a total of 35,000 steps of conjugate gradient energy minimization in order to remove steric clashes caused by hydrogenation and solvation. Heating was then conducted over a 250 ps simulation [103], at constant volume using a Langevin thermostat with a collision frequency of 5.0 ps⁻¹. The target temperature was scaled linearly from 0°K to 300°K over the course of the simulation. After heating, the system was further equilibrated with two consecutive 500 ps anisotropic constant pressure simulations at 1 atm, adding a Monte Carlo barostat with a relaxation time of 1.0 ps. A Langevin thermostat was still used, but the collision frequency was increased from 5.0 ps⁻¹ in the first constant pressure simulation to 2.0 ps⁻¹ for the second 500 ps constant pressure simulation. Production simulations were then run for 510 ns in the NVT ensemble using a Langevin thermostat with a further increased collision frequency of 1.0 ps⁻¹. A time step of 2.0 fs was used for all simulations. Shake was used to constrain all bonds involving hydrogen [105]. A direct-space and vdW cut off of 8.0 Å was used for all simulations. Periodic boundaries coupled with the Particle Mesh Ewald (PME) were used to include long-range electrostatic interactions. Structural and energy data were recorded every 2 ps. All calculations were run with the SPFP precision model using the PMEMD.cuda MD engine (up to and including bugfix.19) from the AMBER 14 software suite using Quadro M6000 GPUs on in house resources [73–77]. Post simulation analysis was conducted using the cpptraj program from the AmberTools v14 software suite [106].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors thank G. C. Melkani, J. A. Suggs, A. Melkani, W. A. Kronert, and B. Cudney for technical assistance and encouragement, J. Jenkins for support in crystallization, J. Schuermann for assistance with synchrotron data collection, J. Headd and B. Stec for help with refinement, J. Vertrees for help with PyMol, and A. Wild for photography. We also thank the reviewers for many helpful suggestions throughout the peer review process. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Pilatus 6M detector on 24-ID-C beam line is funded by a NIH-ORIP HEI grant (S10 RR029205). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User

Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. This research is supported by NIH Grant R37 GM032443 to SIB. Biochemistry research at SDSU is supported in part by the California Metabolic Research Foundation.

REFERENCES

- [1]. Vale RD. The molecular motor toolbox for intracellular transport. *Cell*. 112 (2003) 467–480. [PubMed: 12600311]
- [2]. Hartman MA, Spudich JA. The myosin superfamily at a glance. *J. Cell Sci* 125 (2012) 1627–1632. [PubMed: 22566666]
- [3]. Reggiani C, Bottinelli R. Myosin II: Sarcomeric myosins, the motors of contraction in cardiac and skeletal muscles In: Coluccio LM, editor. *Myosins*. Dordrecht: Springer; 2008 p. 125–169.
- [4]. Chantler PD, Wylie SR, Wheeler-Jones CP, McGonnell IM. Conventional myosins - unconventional functions. *Biophys. Rev* 2 (2010) 67–82. [PubMed: 28510009]
- [5]. De La Cruz EM, Ostap EM. Relating biochemistry and function in the myosin superfamily. *Curr. Opin. Cell Biol* 16 (2004) 61–67. [PubMed: 15037306]
- [6]. Rozek CE, Davidson N. *Drosophila* has one myosin heavy-chain gene with three developmentally regulated transcripts. *Cell*. 32 (1983) 23–34. [PubMed: 6402306]
- [7]. Bernstein SI, Mogami K, Donady JJ, Emerson CP Jr. *Drosophila* muscle myosin heavy chain encoded by a single gene in a cluster of muscle mutations. *Nature*. 302 (1983) 393–397. [PubMed: 6403869]
- [8]. Odronitz F, Kollmar M. Comparative genomic analysis of the arthropod muscle myosin heavy chain genes allows ancestral gene reconstruction and reveals a new type of ‘partially’ processed pseudogene. *BMC Mol. Biol* 9 (2008) 21. [PubMed: 18254963]
- [9]. Kollmar M, Hatje K. Shared gene structures and clusters of mutually exclusive spliced exons within the metazoan muscle myosin heavy chain genes. *PLoS One*. 9 (2014) e88111. [PubMed: 24498429]
- [10]. Zhang S, Bernstein SI. Spatially and temporally regulated expression of myosin heavy chain alternative exons during *Drosophila* embryogenesis. *Mech. Dev* 101 (2001) 35–45. [PubMed: 11231057]
- [11]. George EL, Ober MB, Emerson CP Jr. Functional domains of the *Drosophila melanogaster* muscle myosin heavy-chain gene are encoded by alternatively spliced exons. *Mol. Cell. Biol* 9 (1989) 2957–2974. [PubMed: 2506434]
- [12]. Bernstein SI, Milligan RA. Fine tuning a molecular motor: The location of alternative domains in the *Drosophila* myosin head. *J. Mol. Biol* 271 (1997) 1–6. [PubMed: 9300050]
- [13]. Swank DM, Knowles AF, Kronert WA, Suggs JA, Morrill GE, Nikkhoy M, et al. Variable N-terminal regions of muscle myosin heavy chain modulate ATPase rate and actin sliding velocity. *J. Biol. Chem* 278 (2003) 17475–17482. [PubMed: 12606545]
- [14]. Swank DM, Knowles AF, Suggs JA, Sarsoza F, Lee A, Maughan DW, et al. The myosin converter domain modulates muscle performance. *Nat. Cell Biol* 4 (2002) 312–316. [PubMed: 11901423]
- [15]. Wells L, Edwards KA, Bernstein SI. Myosin heavy chain isoforms regulate muscle function but not myofibrillar assembly. *EMBO J*. 15 (1996) 4454–4459. [PubMed: 8887536]
- [16]. Josephson RK, Malamud JG, Stokes DR. Asynchronous muscle: A primer. *J. Exp. Biol* 203 (2000) 2713–2722. [PubMed: 10952872]
- [17]. 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*. 7 (2018).
- [18]. Viswanathan MC, Tham RC, Kronert WA, Sarsoza F, Trujillo AS, Cammarato A, et al. Myosin storage myopathy mutations yield defective myosin filament assembly in vitro and disrupted myofibrillar structure and function in vivo. *Hum. Mol. Genet* 26 (2017) 4799–4813. [PubMed: 28973424]
- [19]. Cannon L, Zambon AC, Cammarato A, Zhang Z, Vogler G, Munoz M, et al. Expression patterns of cardiac aging in *Drosophila*. *Aging Cell*. 16 (2017) 82–92. [PubMed: 28090760]

- [20]. Bell KM, Kronert WA, Huang A, Bernstein SI, Swank DM. The R249Q hypertrophic cardiomyopathy myosin mutation decreases contractility in *Drosophila* by impeding force production. *J. Physiol* 597 (2019) 2403–2420. [PubMed: 30950055]
- [21]. Das S, Kumar P, Verma A, Maiti TK, Mathew SJ. Myosin heavy chain mutations that cause Freeman-Sheldon syndrome lead to muscle structural and functional defects in *drosophila*. *Dev. Biol* 449 (2019) 90–98. [PubMed: 30826400]
- [22]. Rao DS, Kronert WA, Guo Y, Hsu KH, Sarsoza F, Bernstein SI. Reductions in atpase activity, actin sliding velocity, and myofibril stability yield muscle dysfunction in *Drosophila* models of myosin-based Freeman-Sheldon syndrome. *Mol. Biol. Cell* 30 (2019) 30–41. [PubMed: 30379605]
- [23]. Dahl-Halvarsson M, Olive M, Pokrzywa M, Ejleskar K, Palmer RH, Uv AE, et al. *Drosophila* model of myosin myopathy rescued by overexpression of a trim-protein family member. *Proc. Natl. Acad. Sci. USA* 115 (2018) E6566–E6575. [PubMed: 29946036]
- [24]. Srikakulam R, Winkelmann DA. Chaperone-mediated folding and assembly of myosin in striated muscle. *J. Cell. Sci* 117 (2004) 641–652. [PubMed: 14709723]
- [25]. Deacon JC, Bloemink MJ, Rezavandi H, Geeves MA, Leinwand LA. Identification of functional differences between recombinant human alpha and beta cardiac myosin motors. *Cell. Mol. Life Sci* 69 (2012) 4239–4255. [PubMed: 23001010]
- [26]. Winkelmann DA, Forgacs E, Miller MT, Stock AM. Structural basis for drug-induced allosteric changes to human beta-cardiac myosin motor activity. *Nat. Commun* 6 (2015) 7974. [PubMed: 26246073]
- [27]. Bird JE, Takagi Y, Billington N, Strub MP, Sellers JR, Friedman TB. Chaperone-enhanced purification of unconventional myosin 15, a molecular motor specialized for stereocilia protein trafficking. *Proc. Natl. Acad. Sci. USA* 111 (2014) 12390–12395. [PubMed: 25114250]
- [28]. Zieba J, Zhang W, Chong JX, Forlenza KN, Martin JH, Heard K, et al. A postnatal role for embryonic myosin revealed by MYH3 mutations that alter TGF β signaling and cause autosomal dominant spondylotarsal synostosis. *Sci. Rep* 7 (2017) 41803. [PubMed: 28205584]
- [29]. Hellerschmied D, Lehner A, Franicevic N, Arnese R, Johnson C, Vogel A, et al. Molecular features of the unc-45 chaperone critical for binding and folding muscle myosin. *Nat. Commun* 10 (2019) 4781. [PubMed: 31636255]
- [30]. Caldwell JT, Melkani GC, Huxford T, Bernstein SI. Transgenic expression and purification of myosin isoforms using the *Drosophila melanogaster* indirect flight muscle system. *Methods*. 56 (2012) 25–32. [PubMed: 22178692]
- [31]. Groth AC, Fish M, Nusse R, Calos MP. Construction of transgenic *Drosophila* by using the site-specific integrase from phage Φ C31. *Genetics*. 166 (2004) 1775–1782. [PubMed: 15126397]
- [32]. 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 Dev*. 4 (1990) 885–895. [PubMed: 2116987]
- [33]. Gilmour D. Myosin and adenylpyrophosphatase in insect muscle. *J. Biol. Chem* 175 (1948) 477. [PubMed: 18873325]
- [34]. Margossian SS, Lowey S. Preparation of myosin and its subfragments from rabbit skeletal muscle. *Methods Enzymol*. 85 Pt B (1982) 55–71. [PubMed: 6214692]
- [35]. Sweeney HL, Houdusse A. Structural and functional insights into the myosin motor mechanism. *Annu. Rev. Biophys* 39 (2010) 539–557. [PubMed: 20192767]
- [36]. Cheney RE, Mooseker MS. Unconventional myosins. *Curr. Opin. Cell Biol* 4 (1992) 27–35. [PubMed: 1558751]
- [37]. Dominguez R, Freyzon Y, Trybus KM, Cohen C. Crystal structure of a vertebrate smooth muscle myosin motor domain and its complex with the essential light chain: Visualization of the pre-power stroke state. *Cell*. 94 (1998) 559–571. [PubMed: 9741621]
- [38]. Coureux PD, Wells AL, Menetrey J, Yengo CM, Morris CA, Sweeney HL, et al. A structural state of the myosin V motor without bound nucleotide. *Nature*. 425 (2003) 419–423. [PubMed: 14508494]

- [39]. Holmes KC, Angert I, Kull FJ, Jahn W, Schroder RR. Electron cryo-microscopy shows how strong binding of myosin to actin releases nucleotide. *Nature*. 425 (2003) 423–427. [PubMed: 14508495]
- [40]. Geeves MA, Holmes KC. The molecular mechanism of muscle contraction. *Adv. Protein Chem* 71 (2005) 161–193. [PubMed: 16230112]
- [41]. Yang Y, Gourinath S, Kovacs M, Nyitray L, Reutzel R, Himmel DM, et al. Rigor-like structures from muscle myosins reveal key mechanical elements in the transduction pathways of this allosteric motor. *Structure*. 15 (2007) 553–564. [PubMed: 17502101]
- [42]. Fujii T, Namba K. Structure of actomyosin rigour complex at 5.2 Å resolution and insights into the ATPase cycle mechanism. *Nat. Commun* 8 (2017) 13969. [PubMed: 28067235]
- [43]. von der Ecken J, Heissler SM, Pathan-Chhatbar S, Manstein DJ, Raunser S. Cryo-EM structure of a human cytoplasmic actomyosin complex at near-atomic resolution. *Nature*. 534 (2016) 724–728. [PubMed: 27324845]
- [44]. Menetrey J, Bahloul A, Wells AL, Yengo CM, Morris CA, Sweeney HL, et al. The structure of the myosin VI motor reveals the mechanism of directionality reversal. *Nature*. 435 (2005) 779–785. [PubMed: 15944696]
- [45]. Reubold TF, Eschenburg S, Becker A, Kull FJ, Manstein DJ. A structural model for actin-induced nucleotide release in myosin. *Nat. Struct. Biol* 10 (2003) 826–830. [PubMed: 14502270]
- [46]. Münnich S, Pathan-Chhatbar S, Manstein DJ. Crystal structure of the rigor-like human non-muscle myosin-2 motor domain. *FEBS Lett*. 588 (2014) 4754–4760. [PubMed: 25451231]
- [47]. Robert-Paganin J, Robblee JP, Auguin D, Blake TCA, Bookwalter CS, Kremmentsova EB, et al. Plasmodium myosin a drives parasite invasion by an atypical force generating mechanism. *Nat. Commun* 10 (2019) 3286. [PubMed: 31337750]
- [48]. Coureux PD, Sweeney HL, Houdusse A. Three myosin V structures delineate essential features of chemo-mechanical transduction. *EMBO J*. 23 (2004) 4527–4537. [PubMed: 15510214]
- [49]. Rayment I, Rypniewski WR, Schmidt-Base K, Smith R, Tomchick DR, Benning MM, et al. Three-dimensional structure of myosin subfragment-1: A molecular motor. *Science*. 261 (1993) 50–58. [PubMed: 8316857]
- [50]. Risal D, Gourinath S, Himmel DM, Szent-Gyorgyi AG, Cohen C. Myosin subfragment 1 structures reveal a partially bound nucleotide and a complex salt bridge that helps couple nucleotide and actin binding. *Proc. Natl. Acad. Sci. USA* 101 (2004) 8930–8935. [PubMed: 15184651]
- [51]. Liebschner D, Afonine PV, Moriarty NW, Poon BK, Sobolev OV, Terwilliger TC, et al. Polder maps: Improving omit maps by excluding bulk solvent. *Acta Crystallogr. D Struct. Biol* 73 (2017) 148–157. [PubMed: 28177311]
- [52]. Gourinath S, Himmel DM, Brown JH, Reshetnikova L, Szent-Gyorgyi AG, Cohen C. Crystal structure of scallop myosin S1 in the pre-power stroke state to 2.6 Å resolution: Flexibility and function in the head. *Structure*. 11 (2003) 1621–1627. [PubMed: 14656445]
- [53]. Fromherz S, Szent-Gyorgyi AG. Role of essential light chain of hand domains in calcium binding and regulation of scallop myosin. *Proc. Natl. Acad. Sci. USA* 92 (1995) 7652–7656. [PubMed: 7644472]
- [54]. Falkenthal S, Graham M, Wilkinson J. The indirect flight muscle of *Drosophila* accumulates a unique myosin alkali light chain isoform. *Dev. Biol* 121 (1987) 263–272. [PubMed: 3106119]
- [55]. Miller MS, Farman GP, Braddock JM, Soto-Adames FN, Irving TC, Vigoreaux JO, et al. Regulatory light chain phosphorylation and N-terminal extension increase cross-bridge binding and power output in *Drosophila* at in vivo myofilament lattice spacing. *Biophys. J* 100 (2011) 1737–1746. [PubMed: 21463587]
- [56]. Himmel DM, Gourinath S, Reshetnikova L, Shen Y, Szent-Gyorgyi AG, Cohen C. Crystallographic findings on the internally uncoupled and near-rigor states of myosin: Further insights into the mechanics of the motor. *Proc. Natl. Acad. Sci. USA* 99 (2002) 12645–12650. [PubMed: 12297624]
- [57]. Swank DM, Kronert WA, Bernstein SI, Maughan DW. Alternative N-terminal regions of *Drosophila* myosin heavy chain tune muscle kinetics for optimal power output. *Biophys. J* 87 (2004) 1805–1814. [PubMed: 15345559]

- [58]. Miller BM, Bloemink MJ, Nyitrai M, Bernstein SI, Geeves MA. A variable domain near the atp-binding site in *Drosophila* muscle myosin is part of the communication pathway between the nucleotide and actin-binding sites. *J. Mol. Biol* 368 (2007) 1051–1066. [PubMed: 17379245]
- [59]. Miller BM, Zhang S, Suggs JA, Swank DM, Littlefield KP, Knowles AF, et al. An alternative domain near the nucleotide-binding site of *Drosophila* muscle myosin affects atpase kinetics. *J. Mol. Biol* 353 (2005) 14–25. [PubMed: 16154586]
- [60]. Swank DM, Braddock J, Brown W, Lesage H, Bernstein SI, Maughan DW. An alternative domain near the ATP binding pocket of *Drosophila* myosin affects muscle fiber kinetics. *Biophys. J* 90 (2006) 2427–2435. [PubMed: 16399836]
- [61]. Kintses B, Yang Z, Malnasi-Csizmadia A. Experimental investigation of the seesaw mechanism of the relay region that moves the myosin lever arm. *J. Biol. Chem* 283 (2008) 34121–34128. [PubMed: 18854311]
- [62]. Yang C, Ramanath S, Kronert WA, Bernstein SI, Maughan DW, Swank DM. Alternative versions of the myosin relay domain differentially respond to load to influence *Drosophila* muscle kinetics. *Biophys. J* 95 (2008) 5228–5237. [PubMed: 18805920]
- [63]. Kronert WA, Dambacher CM, Knowles AF, Swank DM, Bernstein SI. Alternative relay domains of *Drosophila melanogaster* myosin differentially affect ATPase activity, in vitro motility, myofibril structure and muscle function. *J. Mol. Biol* 379 (2008) 443–456. [PubMed: 18462751]
- [64]. Bloemink MJ, Dambacher CM, Knowles AF, Melkani GC, Geeves MA, Bernstein SI. Alternative exon 9-encoded relay domains affect more than one communication pathway in the *Drosophila* myosin head. *J. Mol. Biol* 389 (2009) 707–721. [PubMed: 19393244]
- [65]. Littlefield KP, Swank DM, Sanchez BM, Knowles AF, Warshaw DM, Bernstein SI. The converter domain modulates kinetic properties of *Drosophila* myosin. *Am. J. Physiol. Cell Physiol* 284 (2003) C1031–1038. [PubMed: 12477668]
- [66]. Glasheen BM, Ramanath S, Patel M, Sheppard D, Puthawala JT, Riley LA, et al. Five alternative myosin converter domains influence muscle power, stretch activation, and kinetics. *Biophys. J* 114 (2018) 1142–1152. [PubMed: 29539400]
- [67]. Kronert WA, Melkani GC, Melkani A, Bernstein SI. Alternative relay and converter domains tune native muscle myosin isoform function in *Drosophila*. *J. Mol. Biol* 416 (2012) 543–557. [PubMed: 22226837]
- [68]. Yengo CM, Chrin LR, Rovner AS, Berger CL. Tryptophan 512 is sensitive to conformational changes in the rigid relay loop of smooth muscle myosin during the MgATPase cycle. *J. Biol. Chem* 275 (2000) 25481–25487. [PubMed: 10827189]
- [69]. Kronert WA, Melkani GC, Melkani A, Bernstein SI. Mutating the converter-relay interface of *Drosophila* myosin perturbs ATPase activity, actin motility, myofibril stability and flight ability. *J. Mol. Biol* 398 (2010) 625–632. [PubMed: 20362584]
- [70]. Bloemink MJ, Melkani GC, Bernstein SI, Geeves MA. The relay/converter interface influences hydrolysis of ATP by skeletal muscle myosin II. *J. Biol. Chem* 291 (2016) 1763–1773. [PubMed: 26586917]
- [71]. Kronert WA, Melkani GC, Melkani A, Bernstein SI. Mapping interactions between myosin relay and converter domains that power muscle function. *J. Biol. Chem* 289 (2014) 12779–12790. [PubMed: 24627474]
- [72]. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput* 11 (2015) 3696–3713. [PubMed: 26574453]
- [73]. Götz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with amber on GPUs. 1. Generalized born. *J. Chem. Theory Comput* 8 (2012) 1542–1555. [PubMed: 22582031]
- [74]. Le Grand S, Götz AW, Walker RC. SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput. Phys. Commun* 184 (2013) 374–380.
- [75]. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *J. Chem. Theory Comput* 9 (2013) 3878–3888. [PubMed: 26592383]

- [76]. Case DA, Babin V, Berryman JT, Betz RM, Cai Q, Cerutti DS, et al. AMBER 14: University of California, San Francisco; 2014.
- [77]. Salomon-Ferrer R, Case DA, Walker RC. An overview of the AMBER biomolecular simulation package. *Wiley Interdiscip. Rev. Comput. Mol. Sci* 3 (2013) 198–210.
- [78]. Balint M, Wolf I, Tarcsfalvi A, Gergely J, Sreter FA. Location of SH-1 and SH-2 in the heavy chain segment of heavy meromyosin. *Arch. Biochem. Biophys* 190 (1978) 793–799. [PubMed: 152606]
- [79]. Stork NE. How many species of insects and other terrestrial arthropods are there on earth? *Annu. Rev. Entomol* 63 (2018) 31–45. [PubMed: 28938083]
- [80]. Witjas-Paalberends ER, Ferrara C, Scellini B, Piroddi N, Montag J, Tesi C, et al. Faster cross-bridge detachment and increased tension cost in human hypertrophic cardiomyopathy with the R403Q MYH7 mutation. *J. Physiol* 592 (2014) 3257–3272. [PubMed: 24928957]
- [81]. Nag S, Sommese RF, Ujfalusi Z, Combs A, Langer S, Sutton S, et al. Contractility parameters of human β -cardiac myosin with the hypertrophic cardiomyopathy mutation R403Q show loss of motor function. *Sci. Adv* 1 (2015) e1500511. [PubMed: 26601291]
- [82]. Spudich JA, Aksel T, Bartholomew SR, Nag S, Kawana M, Yu EC, et al. Effects of hypertrophic and dilated cardiomyopathy mutations on power output by human β -cardiac myosin. *J. Exp. Biol* 219 (2016) 161–167. [PubMed: 26792326]
- [83]. Johnson C, McGreig J, Vera C, Mulvihill D, Ridout M, Leinwand L, et al. Cardiac contraction velocity has evolved to match heart rate with body size through variation in β -cardiac myosin sequence. (2019) doi: 10.1011/680413.
- [84]. Vera CD, Johnson CA, Walklate J, Adhikari A, Svicevic M, Mijailovich SM, et al. Myosin motor domains carrying mutations implicated in early or late onset hypertrophic cardiomyopathy have similar properties. *J. Biol. Chem* (2019) doi:10.1074/jbc.RA119.010563
- [85]. Robert-Paganin J, Auguin D, Houdusse A. Hypertrophic cardiomyopathy disease results from disparate impairments of cardiac myosin function and auto-inhibition. *Nat. Commun* 9 (2018) 4019. [PubMed: 30275503]
- [86]. Sane SP. The aerodynamics of insect flight. *J. Exp. Biol* 206 (2003) 4191–4208. [PubMed: 14581590]
- [87]. Taylor KA, Rahmani H, Edwards RJ, Reedy MK. Insights into actin-myosin interactions within muscle from 3D electron microscopy. *Int. J. Mol. Sci* 20 (2019).
- [88]. Hu Z, Taylor DW, Reedy MK, Edwards RJ, Taylor KA. Structure of myosin filaments from relaxed *Lethocerus* flight muscle by cryo-EM at 6 Å resolution. *Sci. Adv* 2 (2016) e1600058. [PubMed: 27704041]
- [89]. 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. *J. Mol. Biol* 428 (2016) 1142–1164. [PubMed: 26851071]
- [90]. Pinto A, Sanchez F, Alamo L, Padron R. The myosin interacting-heads motif is present in the relaxed thick filament of the striated muscle of scorpion. *J. Struct. Biol* 180 (2012) 469–478. [PubMed: 22982253]
- [91]. Wendt T, Taylor D, Messier T, Trybus KM, Taylor KA. Visualization of head-head interactions in the inhibited state of smooth muscle myosin. *J Cell Biol* 147 (1999) 1385–1390. [PubMed: 10613897]
- [92]. 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. *Proc. Natl. Acad. Sci. USA* 115 (2018) E1991–E2000. [PubMed: 29444861]
- [93]. Polley S, Huang DB, Hauenstein AV, Fusco AJ, Zhong X, Vu D, et al. A structural basis for I κ B kinase 2 activation via oligomerization-dependent trans auto-phosphorylation. *PLoS Biol.* 11 (2013) e1001581. [PubMed: 23776406]
- [94]. Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell.* 125 (2006) 1137–1149. [PubMed: 16777603]
- [95]. Menetret JF, Schroder RR, Hofmann W. Cryo-electron microscopic studies of relaxed striated muscle thick filaments. *J. Muscle Res. Cell Motil* 11 (1990) 1–11. [PubMed: 2351744]

- [96]. Bischof J, Maeda RK, Hediger M, Karch F, Basler K. An optimized transgenesis system for *Drosophila* using germ-line-specific Phic31 integrases. *Proc. Natl. Acad. Sci USA* 104 (2007) 3312–3317. [PubMed: 17360644]
- [97]. Long F, Vagin AA, Young P, Murshudov GN. BALBES: A molecular-replacement pipeline. *Acta Crystallogr. D* 64 (2008) 125–132. [PubMed: 18094476]
- [98]. Vagin A, Teplyakov A. Molecular replacement with MOLREP. *Acta Crystallogr. D* 66 (2010) 22–25. [PubMed: 20057045]
- [99]. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MOLPROBITY: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* 66 (2010) 12–21. [PubMed: 20057044]
- [100]. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of COOT. *Acta Crystallogr D*. 66 (2010) 486–501. [PubMed: 20383002]
- [101]. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: A comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr D*. 66 (2010) 213–221. [PubMed: 20124702]
- [102]. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D*. 67 (2011) 355–367. [PubMed: 21460454]
- [103]. DeLano WL. The PyMol molecular graphics system. Palo Alto, CA, USA: DeLano Scientific; 2002.
- [104]. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys* 79 (1983) 926–935.
- [105]. Ryckaert J-P, Ciccotti G, Berendsen HJC. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J. Comput. Phys* 23 (1977) 327–341.
- [106]. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput* 9 (2013) 3084–3095. [PubMed: 26583988]

Research Highlights

- A 2.2 Å x-ray crystal structure of the embryonic isoform of *Drosophila* myosin II
- This is the first x-ray crystal structure of an insect myosin motor domain
- The model closely resembles rigor-like myosins from other species
- 61 amino acids differ between embryonic and indirect flight muscle myosins
- Molecular dynamics reveal that relay and converter differ most between isoforms

Figure 1.

Expression and structure of the *Drosophila melanogaster* embryonic muscle myosin II heavy chain isoform subfragment-1 (EMB). A) Photographs of transgenic flies expressing EMB at 6 (left) and 24 (right) hours post-eclosure, with the latter displaying the raised wing phenotype associated with IFM degeneration. B) Cartoon representation of the dimeric myosin II structure with one subunit (above) colored by functional subdomains and the other (below) colored to show the locations of the exon encoded regions unique to EMB (see panel C for color coding). The portion that was isolated by proteolysis for crystallography is

labeled as “S1”. C) Two schematic representations of recombinant EMB MHC. In one representation (top) the encoded MHC subdomain borders are indicated (NTD, N-terminal domain; U50, upper 50 kDa subdomain; L50, lower 50 kDa subdomain; CVTR, converter; LVR, lever arm) as they are encoded by the plasmid used in generation of transgenic fly lines. The other (bottom) shows the plasmid with the specific alternative exons present in EMB transcripts listed. D) Schematic of the crystallized EMB MHC proteolytic fragment depicted as in panel C. E) Ribbon diagram model of the EMB x-ray crystal structure shown from two views. MHC subdomains and features are labeled and colors are consistent with the top portions of panels B-D. The essential light chain (ELC) is depicted in cyan. F) Ribbon diagram model of the EMB x-ray crystal structure oriented as in panel E. The regions encoded by alternative exons unique to EMB are labeled and colored as in the bottom portions of panels B-D and the ELC is cyan.

Figure 2.

Overlay of two MHC S1-like fragments from the asymmetric unit of the EMB x-ray crystal structure. Chain A (cyan) and chain C (black) are depicted as Ca traces in line form and rendered in cross-eyed stereo. ELC has been removed for clarity. Amino acid residues 214 and 451, which border the U50 subdomain, are labeled, which reveals *en bloc* movement of the entire U50 subdomain relative to the rest of the protein.

Figure 3.

Citrate in the nucleotide binding site of *Drosophila* EMB. A) Close up view of the EMB heavy chain A nucleotide binding site with bound citrate depicted as sticks in a $2F_o - F_c$ electron density map (blue mesh) contoured at 1.1σ . The EMB myosin II motor domain is rendered as a ribbon diagram with its respective subdomains labeled and colored as in Figure 1. The P loop (yellow) as well as switch I and switch II are also labeled. B) EMB heavy chain C oriented and represented in a similar manner to reveal its bound citrate ion. C) Similarly oriented and labeled crystallographic model of the myosin II protein from scallop (PDB ID: 1S5G) in a rigor-like conformation with sulfate ion (SO_4^{2-}) and ADP (stick models) in the nucleotide binding pocket.

Figure 4.

Essential light chain structure in EMB. A) The *Drosophila* ELC protein is depicted as a cyan ribbon diagram bound to the EMB lever arm in yellow. ELC is superposed upon ELC proteins from scallop myosin II (light grey) and human ELC from the chicken nonmuscle myosin V (dark grey) structures and shown from three orthogonal views. B) Close up view of the side chain orientations and hydrogen bonding at the C-terminus of *Drosophila* ELC with individual residues labeled.

Figure 5.

Comparison of alternative exon-encoded region structures in the *Drosophila* EMB crystallographic model and other MHC motor domains. Portions of EMB encoded by alternative exons are colored as in Figure 1F. A) EMB is depicted as a ribbon diagram representation with semi-transparent surface and the region encoded for by exon 3a highlighted in red. Small spheres represent the C α positions of amino acids that differ between EMB and IFI and homologous regions of bay scallop rigor-like (orange in A and C; red in B and D), and pre-power stroke (yellow), chicken post-rigor (green in A,B, and D; red in C), and *Dictyostelium* rigor-like (purple) myosin motor domains are shown superimposed as C α ribbons. B-D) Similar representations of EMB with regions encoded by exons 7a, 9b, and 11c, respectively, colored as in Figure 1F and with amino acid borders and other structural features labeled.

Figure 6. Interactions between neighboring EMB proteins in the crystallographic asymmetric unit. A) Cross-eyed stereo view of the arrangement of two independent EMB myosin heavy chain S1-like fragment models within the crystallographic asymmetric unit. Portions encoded by alternative exons and the ELC are colored as in Figure 1F and labeled. B) Close up view of the interaction interface with models depicted as cartoon ribbons. Portions of exons 7a and 9b that mediate interaction are colored consistent with Figure 1F and individual amino acid side chains are labeled.

Figure 7.

Molecular dynamics simulations on EMB and IFI models based on EMB x-ray crystal structure. A) Root-mean-squared fluctuation of $\text{C}\alpha$ positions during molecular dynamics. The graph monitors amino acid positional changes over the time of the simulation. Data for the EMB model are shown in blue and IFI data are plotted in orange. B) Difference root-mean-squared fluctuations (IFI minus EMB) expressed as a function of amino acid position. C) Time-resolved changes in rmsd for $\text{C}\alpha$ positions of amino acids 723-761, encoded for by exon 11 and comprising a major portion of the converter. D) Time-resolved changes in cleft

opening between U50 and L50 subdomains measured as the distance between residues 276 and 471 at the “inner” cleft.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 8. Schematic of *Drosophila* myosin II motor domain dynamics. A) Cartoon representation of EMB in rigor-like conformation. Structural features are labeled. The central seven-stranded β -sheet “transducer” is represented as arrows. B) Dynamics simulations on EMB and an IFI homology model reveal an increased rate of transition between the two rigor-like sub-states observed in the EMB x-ray crystal structure (represented by double-headed arrow at the opening of the cleft between L50 and U50 subdomains) as well as increased mobility of the converter (symbolized by multiple placements of converter) for the IFI model relative to EMB. Each dot corresponds to one amino acid change between EMB and IFI within portions of the respective proteins encoded for by exons 3, 7, 9 and 11.

Table 1.

Data collection and refinement statistics

EMB	
<i>Data collection</i>	
X-ray source	APS 24ID-C
Wavelength (Å)	0.9795
Space group	P2 ₁ 2 ₁ 2 ₁
Unit cell (Å)	
a	108.554
b	148.582
c	148.734
Molecules/asymm. unit	2
Resolution range (Å) ¹	105.1-2.23 (2.31-2.23)
R _{sym} (%)	9.0 (45.4)
Observations	523,304
Unique reflections	116,704
Completeness (%)	99.0 (97.3)
<I/σ>	11.0 (2.4)
<i>Refinement</i>	
Number of reflections	116,624
R _{cryst} (%)	18.4 (25.2)
R _{free} (%) ²	22.1 (29.3)
Protein atoms (No H)	14,995
Ligand atoms (No H)	50
Solvent atoms	1,063
Hydrogen atoms	14,946
R.m.s.d.	
Bond lengths (Å)	0.002
Bond angles (°)	0.464
Ramachandran plot	
Favored	98.1%
Allowed	1.7%
Disallowed	0.2%
MolProbity score ³	1.24
PDB accession code	5W1A

¹Data in parentheses are for highest resolution shell

²Calculated against a cross-validation set of 5.0% of data selected at random prior to refinement

³Combines clashscore, rotamer, and Ramachandran evaluations in to a single score, normalized to the same scale as x-ray resolution.

Table 2.

Conformational analysis of myosin II motor domain models

Conformation	Cleft orientation (°) ¹		Cleft closure (Å) ²			
	Upper 50	Lower 50	Far outer	Outer	Inner	Far inner
Pre-power stroke ³	148.3-154.6	127.4-133.1	22.2-25.4	13.3-13.9	11.8-12.5	8.8-8.9
Rigor-like ⁴	126.4-129.2	114.9-119.8	13.1-21.6	7.2-10.4	9.6-12.9	7.5-10.4
Rigor ⁵	120.8-125.5	112.4-113.8	14.1-16.5	6.5-7.7	7.9-12.7	7.9-9.5
Post-rigor ⁶	146.7-154.5	123.5-134.2	20.1-23.5	13.1-15.8	14.3-16.8	12.4-15.0
EMB chain A	126.5	116.1	15.6	8.0	11.7	10.6
EMB chain C	124.7	115.6	13.4	7.0	8.9	7.5

¹ Angle between the Cα of EMB N-terminal domain residues 148, 188 and either the upper 50 kDa domain residue 421 or lower 50 kDa domain residue 600.

² Distance between the Cα of EMB residues 370 and 540 (far outer/actin binding surfaces), 421 and 598 (outer/strut), 276 and 471 (inner), and 236 and 467 (far inner/switches).

³ Measured from models 1QVI (bay scallop with ADP and vanadate), 1VOM (*Dictyostelium* with ADP and vanadate), 1BR1 (chicken smooth muscle myosin with ADP and AF₄⁻).

⁴ Measured from models 3I5G (squid), 1OE9 (human myosin V), 2BKH (pig myosin VI), 2OS8 (sea scallop), 2AKA (*Dictyostelium*).

⁵ Measured from cryo-EM myosin:actin complex structures 5JLH (human non-muscle myosin 2C) and 5H53 (rabbit skeletal muscle).

⁶ Measured from models 3I5F (squid), 1SR6 (bay scallop), 1W7J (human myosin V with ADP-BeFx), 2OTG (sea scallop), 1MMD (*Dictyostelium* with ADP and BeF₃), 2MYS (chicken striated), 2Y8I (*Dictyostelium* with ADP).