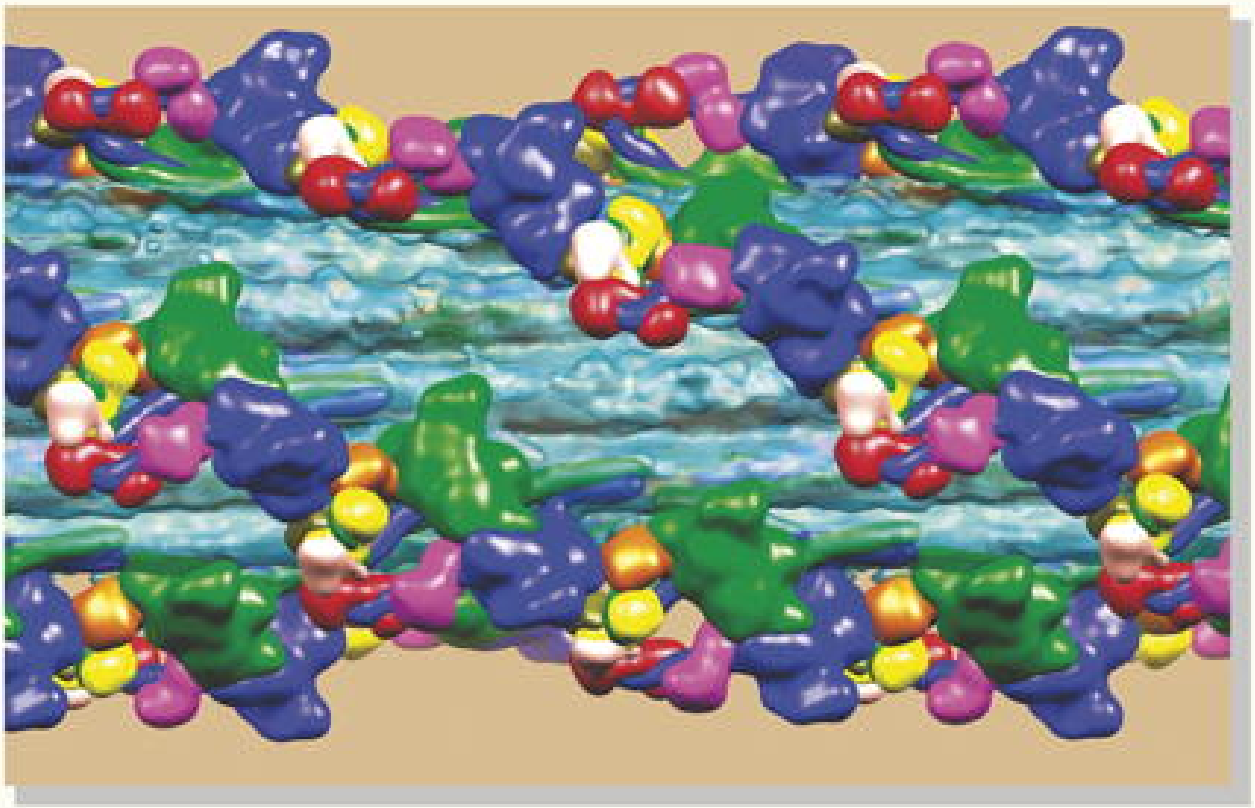
Volume 384, Number 4, 26 December 2008

ISSN 0022-2836

ELSEVIER

JMB

Journal of Molecular Biology



0022-2836(20081226)384:4;1-5

doi:10.1016/j.jmb.2008.10.013 J. Mol. Biol. (2008) 384, 780–797

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Three-Dimensional Reconstruction of Tarantula Myosin Filaments Suggests How Phosphorylation May Regulate Myosin Activity

Lorenzo Alamo[1], Willy Wriggers[2], Antonio Pinto[1], Fulvia Bártoli[1], Leiria Salazar[1], Fa-Qing Zhao[3], Roger Craig[3] and Raúl Padrón[1]*

[1]*Departamento de Biología Estructural, Instituto Venezolano de Investigaciones Científicas, Apdo. 20632, Caracas 1020A, Venezuela*

[2]*University of Texas Health Science Center at Houston, Houston, TX 77030, USA*

[3]*Department of Cell Biology, University of Massachusetts Medical School, Worcester, MA 01655, USA*

Received 26 June 2008; received in revised form 27 September 2008; accepted 2 October 2008
Available online
14 October 2008

Muscle contraction involves the interaction of the myosin heads of the thick filaments with actin subunits of the thin filaments. Relaxation occurs when this interaction is blocked by molecular switches on these filaments. In many muscles, myosin-linked regulation involves phosphorylation of the myosin regulatory light chains (RLCs). Electron microscopy of vertebrate smooth muscle myosin molecules (regulated by phosphorylation) has provided insight into the relaxed structure, revealing that myosin is switched off by intramolecular interactions between its two heads, the free head and the blocked head. Three-dimensional reconstruction of frozen–hydrated specimens revealed that this asymmetric head interaction is also present in native thick filaments of tarantula striated muscle. Our goal in this study was to elucidate the structural features of the tarantula filament involved in phosphorylation-based regulation. A new reconstruction revealed intra- and intermolecular myosin interactions in addition to those seen previously. To help interpret the interactions, we sequenced the tarantula RLC and fitted an atomic model of the myosin head that included the predicted RLC atomic structure and an S2 (subfragment 2) crystal structure to the reconstruction. The fitting suggests one intramolecular interaction, between the cardiomyopathy loop of the free head and its own S2, and two intermolecular interactions, between the cardiac loop of the free head and the essential light chain of the blocked head and between the Leu305–Gln327 interaction loop of the free head and the N-terminal fragment of the RLC of the blocked head. These interactions, added to those previously described, would help switch off the thick filament. Molecular dynamics simulations suggest how phosphorylation could increase the helical content of the RLC N-terminus, weakening these interactions, thus releasing both heads and activating the thick filament.

Keywords: myosin regulation; thick filament; cryo-EM; myosin regulatory light chain; phosphorylation

Edited by J. Karn

Introduction

Striated muscle is formed by two sets of overlapping filaments, the thick myosin-containing filaments and the thin actin-containing filaments. During contraction, the two sets actively slide past each other, shortening the sarcomere. The myosin heads are helically ordered on the backbone of relaxed thick filaments. When filaments are activated, the heads detach from the backbone and become disordered. Sliding force is produced when these heads cyclically attach to and pull on the thin filaments.[1]

**Corresponding author.* E-mail address: padron@ivic.ve.

Present address: W. Wriggers, D. E. Shaw Research, 39th Floor, 120 West Forty-Fifth Street, New York, NY 10036, USA.

Abbreviations used: C loop, cardiac loop; CM loop, cardiomyopathy loop; ELC, essential light chain; EM, electron microscopy; FSC, Fourier shell correlation; HMM, heavy meromyosin; I loop, interaction loop; IHRSR, iterative helical real-space reconstruction; MD, molecular dynamics; MLCK, myosin light-chain kinase; NTF, N-terminal fragment; RLC, regulatory light chain; S1, subfragment 1; S2, subfragment 2.

Contraction of most muscles is regulated via Ca^{2+}-controlled molecular switches located on either or both sets of filaments.[2] While the mechanism of actin-linked regulation by troponin and tropomyosin is well understood,[3,4] our knowledge on the mechanism of myosin-linked regulation is less complete.[5,6] It can occur by direct Ca^{2+} binding to the essential light chains (ELCs; see below)[5] (scallop striated muscle) or by regulatory light chain (RLC) phosphorylation (invertebrate striated muscle[7,8] and vertebrate smooth muscle[9]). While RLC phosphorylation occurs in many muscles, its functional importance varies. In the case of vertebrate striated muscle, for example, phosphorylation appears to modulate contraction, but it is not essential for activity.[6]

Myosin contains two heavy chains and two pairs of light chains (the ELC and the RLC).[10,11] One RLC and one ELC are noncovalently bound to the heavy-chain α-helix in the lever arm of each myosin head,[10] stabilizing it and forming the regulatory domain.[11,12] The RLC has two domains[9] connected by a "linker helix" and an N-terminal fragment (NTF) or extension; all are important for regulation. The NTF varies in length, depending on species, and includes phosphorylation sites. While the NTF is absent from S1 crystal structures,[10] its structure has emerged from EPR of smooth muscle myosin;[13] revealing that it acts as a distinct *phosphorylation domain*, changing from solvent inaccessible and weakly helical when unphosphorylated to solvent accessible with helical order and increased rotational mobility when Ser19 is phosphorylated. Electron microscopy (EM) of chicken smooth muscle myosin [14–17] revealed that, in the switched-off state (dephosphorylated), the two heads establish an asymmetric "interacting-head" structure in which actin binding activity of one head ("blocked") is sterically blocked by binding of its actin binding interface to the converter domain of the other head. Actin binding activity of this head is not blocked (it is therefore called the "free" head), while its ATPase activity is inhibited, by prevention of converter movements needed for phosphate release. In addition to interaction between the two motor domains, there is also an interaction between the blocked-head motor domain and the free-head ELC. This interacting-head structure, deduced from two-dimensional crystals[14–16] and single-molecule[17] studies, has also been shown to be present in native thick filaments of tarantula striated muscle.[18] It is therefore not an artifact of myosin isolation. In the filament, three additional interactions are seen:[18] one intramolecular interaction between the blocked head and its subfragment 2 (S2); subsequently also seen in single-molecule studies[17] and two intermolecular interactions between the blocked-head SH3 domain and the S2 from its axially adjacent neighbor as well as between the blocked-head ELC and the axially neighboring free-head motor domain. The interacting-head motif has now been observed in other isolated myosin molecules (scallop,[19] tarantula, *Limulus*, and mouse skeletal and cardiac muscles[20]) and in thick filaments from mouse cardiac muscle,[21] *Limulus* and scallop striated muscle,[22] and scorpion striated muscles (Sanchez *et al.*, unpublished results), supporting the concept[18,23,24] that this motif is highly conserved and underlies the relaxed state of thick filaments in both smooth and striated muscles over a wide range of species.

When regulated, thick filaments are activated by phosphorylation of their RLCs, the interaction between the heads is broken,[14,16] and the heads are released from the filament surface, becoming disordered.[8,25,26] Breakage of the interaction could occur through a disorder-to-order transition of the phosphorylation domain,[13] apparently by establishing an Arg16–Ser19 salt bridge,[27] a sequence known to be essential for regulation by Ser19 phosphorylation.[28]

Our goal in this study was to elucidate the mechanism by which phosphorylation-induced changes in the RLC activate the thick filament by releasing the heads from each other and from the filament surface so that they can interact with the thin filament. To achieve this, we calculated a new three-dimensional map of tarantula filaments that better defines the RLC region, revealing two new interactions with functional significance. We propose a mechanism to explain how RLC phosphorylation weakens these inhibitory head interactions, releasing them so they can interact with the thin filament.

Results

Three-dimensional reconstruction of frozen–hydrated tarantula thick filaments

We carried out a new three-dimensional reconstruction of tarantula filaments by the iterative helical real-space reconstruction (IHRSR) technique, using an initial reference viewed at 0° or tilted up to ±12°, to account for possible filament tilt in the images (see Materials and Methods).[29–31] The reconstruction was more detailed, clearly showing two new interactions in addition to those seen previously[18] (see Introduction)—an intramolecular density ("a" in Fig. 1) between the motor domain of the free head and its own S2 and an intermolecular density ("b" in Fig. 1) between the RLC of the blocked head and the motor domain of the free head of the axially adjacent molecule.

We carried out atomic fitting to aid in the interpretation of these interactions. With the use of the atomic model determined previously,[18] part of the RLC volume is not filled ("b" in Fig. 1). A simple explanation is that the atomic model used[18] had a smaller RLC than tarantula. The molecular mass of the tarantula RLC has been reported to be 26 kDa,[8] whereas the chicken smooth muscle RLC[32] used is only 20 kDa. To determine whether this difference could account for the extra volume, we determined the sequence of the tarantula striated muscle RLC and performed an improved atomic fitting of the three-dimensional map.

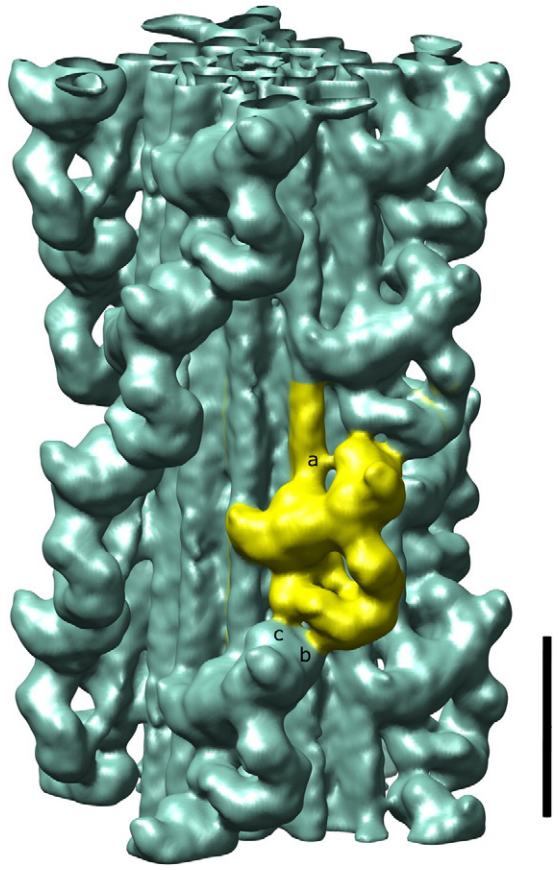


Fig. 1. Three-dimensional reconstruction of the frozen–hydrated tarantula thick filament, filtered to 2-nm resolution (see Supplementary Fig. 3 and Methods), showing four helices of interacting-head motifs (one in yellow). The three-dimensional map segment shows four 14.5 crowns of interacting heads. Bar represents 14.5 nm.

Sequence of tarantula RLC

We cloned the RLC gene of the tarantula *Avicularia avicularia* from RNA isolated from the flexor metatarsus longus striated muscle of the legs. The sequence consists of 196 aa and has a molecular mass of 21,679.2 Da, including a 52-aa NTF (Fig. 2). This is about 4 kDa smaller than the 26 kDa estimated by SDS-PAGE[8] and about 2 kDa bigger than the 20 kDa of the smooth muscle RLC.[32] The NTF molecular mass of 5671.4 Da, about one-quarter of the RLC molecule, is absent from the RLC atomic structure used for the previous fitting.[18] A histogram of the amino acid length of the NTFs from the reported 110 RLC amino acid sequences (Supplementary Fig. 1) shows that these fragments can be either short (8–27 aa, 100 species) or long (43–61 aa, 10 species) (Fig. 3a), as in tarantula. In almost all reported RLC sequences with short NTFs, the putative phosphorylatable serine is located in the same homologous position with a closely similar myosin light-chain kinase (MLCK) consensus sequence,[44] KKRXXSXBB (where "X" is any amino acid and "B" is any hydrophobic amino acid A, V, F, I, or L). Tarantula myosin has two phosphorylation sites,[46] and mass spectrometry studies show that the first site phosphorylated is Ser35 and the second is Ser45 and that some basal (mono-)phosphorylation occurs (at Ser35) even in relaxed filaments (Brito *et al.*, unpublished results). The agreement with the MLCK consensus region (Fig. 2) is better for Ser45 than for Ser35. In common with tarantula, all the RLC sequences with long NTFs (Fig. 3a) have one serine homologous to Ser45, which is located inside an MLCK consensus region,[44] and all but one (*Riftia*) have an additional putative phosphorylatable serine located within ±1 aa of the homologous Ser35 tarantula position.

Predicted atomic structure of the tarantula RLC

As the crystallographic structure of the tarantula RLC has not been solved, we attempted to predict its tertiary structure using homology modeling methods on the basis of known template structures[33] with the sequence we have determined (see Materials and Methods). The predicted structure consists of three domains (Fig. 4). Domains 1 and 2 each comprises two EF-hand helix–loop–helix motifs, so that there are four helices A–D in domain 1 and four helices E–H in domain 2 connected by a linker helix (cf., Ref. 9). The known divalent cation (Ca^{2+} or Mg^{2+}) binding site in the chicken skeletal RLC is located in the first EF motif (helices A and B)[10] (although sequence analysis suggests that this is nonfunctional as a cation binding site in tarantula; Zhu *et al.*, unpublished results). Domain 3, or the "phosphorylation domain,"[13] comprises the NTF, whose structure cannot be predicted using such homology methods because the NTF has not been reported in any myosin crystallographic studies. Therefore, we used programs to predict its secondary structure (Fig. 2). These algorithms predicted 10 helical regions for the complete RLC sequence, with 8 of them confirming helices A–H as revealed in the X-ray structure[10] and 2 (helices L and P; Figs. 2 and 4) located in the NTF domain, in agreement with previous predictions.[27] Helix L is rich in positively charged Lys residues and is connected through a Pro–Ala repeat *linker* to the *phosphorylation helix* P, which contains the two phosphorylatable serines, Ser35 and Ser45. This NTF secondary structure (Fig. 2) was used to predict the structure of domain 3 using the *ab initio* tertiary structure predictor PredictProtein server.[33] As shown in Fig. 4, helix L, the linker, and helix P of the final predicted domain 3 comprised Asp3–Gly21, Gly22–Pro33, and Pro34–Gln41, respectively.

Interacting-head atomic model

The Wendt *et al.*[15] heavy meromyosin (HMM) atomic model included an S2 domain, but this was arbitrarily positioned as its density was not directly visible in the reconstruction. In contrast, in our previous three-dimensional reconstruction,[18] a clear rod-like volume of density, running from the junction of the light-chain domains toward the filament backbone, was identified as the first portion of the

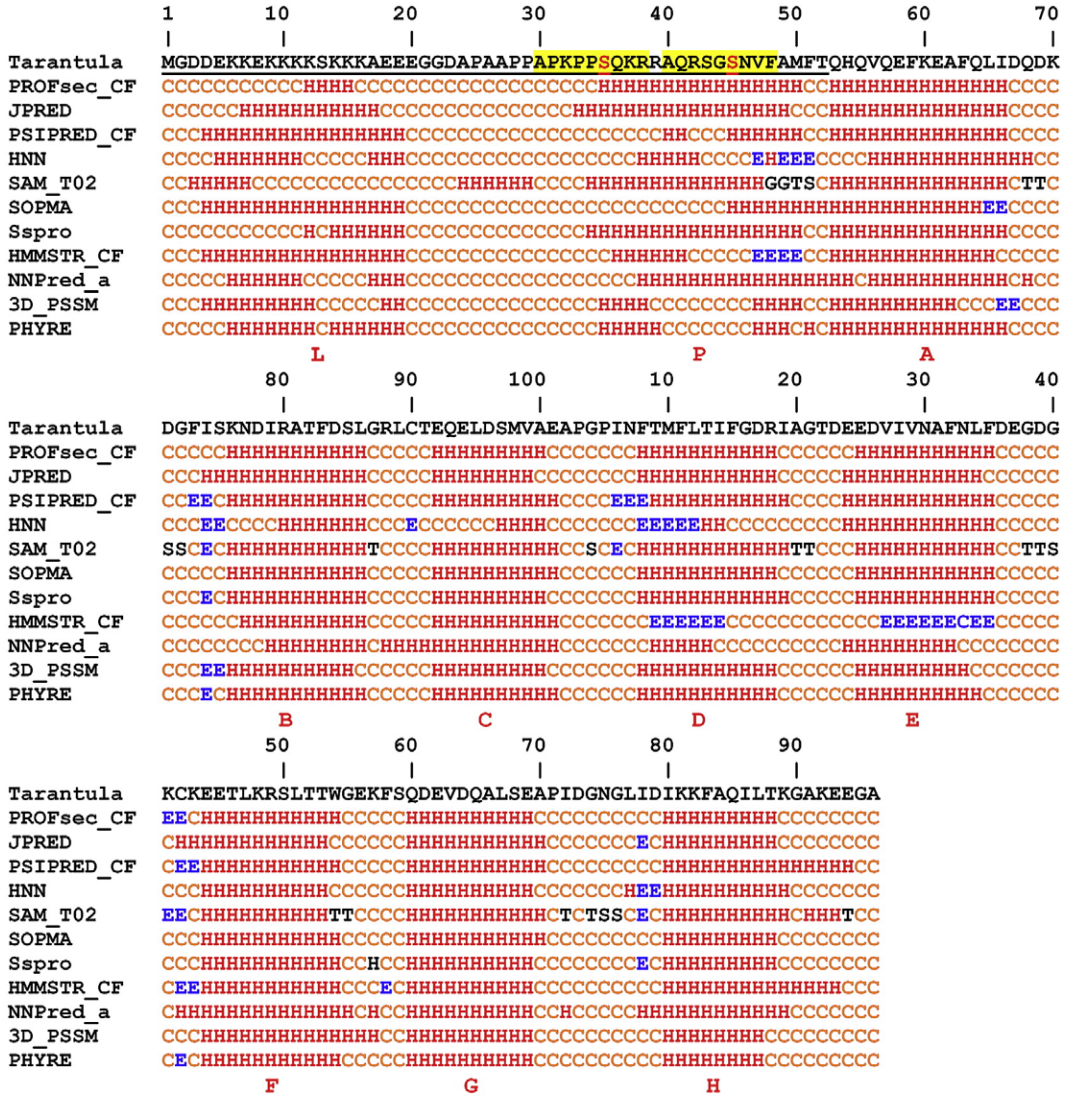


Fig. 2. The first line shows the amino acid sequence of tarantula RLC. The 52-aa NTF is underlined. Lines 2–12 show the prediction of secondary structure using the programs PROFsec,[33] JPRED,[34] PSIPRED,[35] HNN,[36] SAM,[37] SOPMA,[38] Sspro,[39] HMMSTR,[40] NNPred,[41] 3D_PSSM,[42] and PHYRE.[43] All programs predicted 10 helices, with A–H corresponding to four EF-hand motifs[10] and L and P located in the NTF. Helix L, rich in positively charged Lys residues, is connected through a Pro–Ala repeat *coil linker* to the *phosphorylation* helix P, containing the two phosphorylatable serines (in red: Ser35 and Ser45), each located in an MLCK consensus sequence[44] (highlighted in yellow). H indicates helix; E, strand; C, coil; G, 3/10 helix; T, H-bonded turn; and S, bend.

myosin S2 tail. For the present work, we used an initial model for atomic fitting comprising the refined chicken smooth muscle structure as determined by cryo-EM[16] (Fig. 5a) and the human cardiac S2 crystal structure[45] (Fig. 5c) (see Materials and Methods). The intervening amino acids connecting them were energy minimized using molecular dynamics (MD; Fig. 5a; see Materials and Methods). We replaced the RLC with the tarantula RLC-predicted atomic structure including the NTF (Fig. 4). The complete initial hybrid model used for atomic fitting is shown in Fig. 5a. The sequence numbering used in this work is that corresponding to chicken[16] for both HMM heavy chains, including in this numbering the S2 sequence (which is from humans[45]) and the tarantula (Fig. 2) for the RLC.

Atomic fitting of the interacting-head atomic model to the three-dimensional reconstruction

The atomic model described above was flexibly fitted to the volume comprising the interacting heads (Fig. 1) of the three-dimensional map using Situs[47] (see Materials and Methods). The best fitting achieved using 31 positional markers (Fig. 5a) is shown in Fig. 5b. The fitting reveals details of the two new interactions described above (in addition to those described previously[18]), an intramolecular interaction between the free-head motor domain and its S2 ("a" in Figs. 5b and 6a and b) and an intermolecular interaction between the RLC of the blocked head and the motor domain of the axially adjacent free head ("b" in Figs. 5b and 6a and b). The

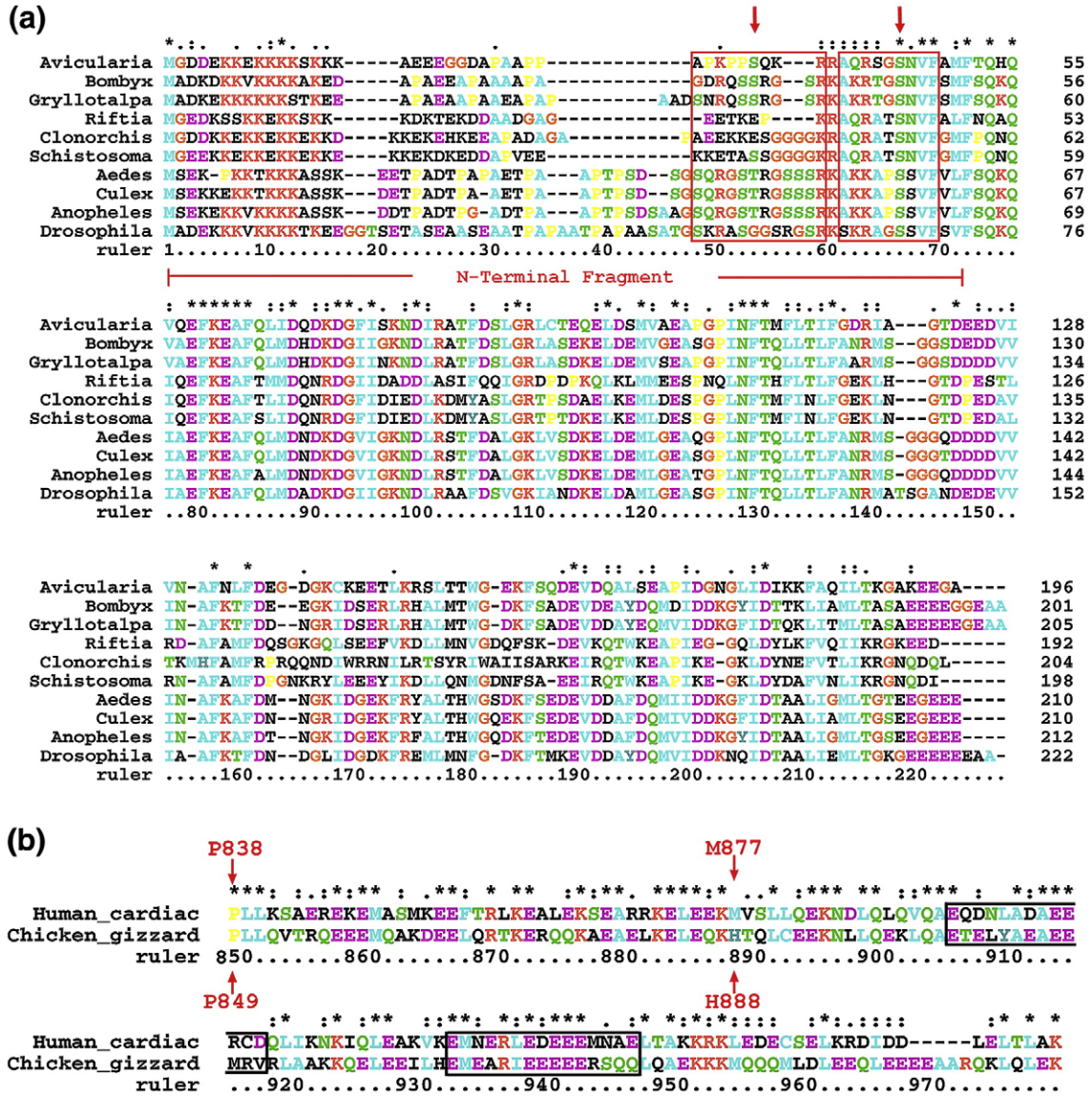


Fig. 3. (a) Sequence alignment of RLCs for the 10 reported species with long NTFs. The two tarantula phosphorylatable serines (Ser35 and Ser45) are marked with arrows, and their MLCK consensus sequences[44] are marked with red boxes. UniProtKB ID code sequences (striated): *Bombyx mori*, Q1HPS0; *Gryllotalpa orientalis*, Q49M29; *Riftia pachyptila*, Q9GUA2; *Clonorchis sinensis*, Q2YHG1; *Schistosoma japonicum*, Q5DB55; *Aedes aegypti*, Q17HX1; *Culex pipiens*, Q45FA2; *Anopheles gambae*, Q7PUV3; and *Drosophila melanogaster*, P18432. (b) Sequence alignment of S2 from human cardiac[45] and chicken smooth[32] muscles. The S2 kink positions are indicated with red arrows: Met877 (human[45]) and His888 (chicken[32]). The black boxes show the two negatively charged rings,[45] 1 (Glu905–Asp917) and 2 (Glu932–Gln946), of S2.

spatial arrangement of these interactions is shown in the stereoviews of Fig. 6a and b, along with the interaction previously described between the ELC of the blocked head and the motor domain of the neighboring free head ("c" in Figs. 5b and 6a and b). After flexible fitting (Fig. 5b), both the blocked- and free-head structures remained similar to the myosin "closed" conformation seen in the smooth muscle model[16] and in the original tarantula fitting[18] (Fig. 6c); however, the part of the free-head heavy-chain α-helix associated with the regulatory domain had an orientation midway between the closed and transition state pre-power stroke structures (1BR1 and 1DFL; Fig. 6c). A comparison of this new model (Fig. 5b) with the refined smooth muscle model[16] [derived from Protein Data Bank (PDB) entry 1I84; without S2] and the atomic model published previously[18] (with S2) is shown in Supplementary Fig. 2.

The flexible fitting modifies the orientation of the cardiomyopathy loop (CM loop; Lys407–Val415) and the cardiac loop (C loop; Gln361–Asp380) of the free head by displacing them such that both loops fit better into the intramolecular tubes of density ("a" and "c" in Fig. 6b). Thus, the motor domain of the free head interacts with the negatively

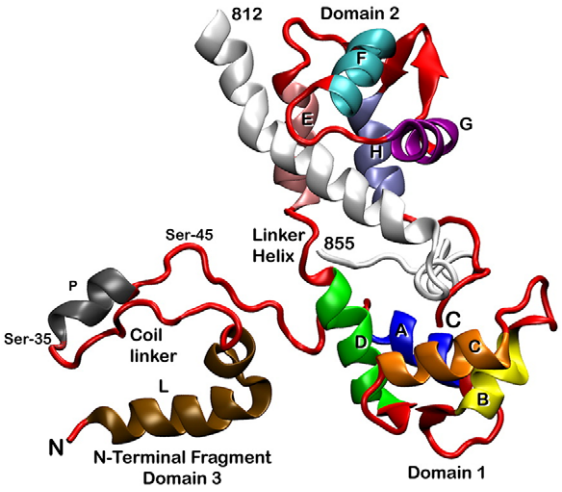


Fig. 4. Predicted atomic structure for the tarantula RLC obtained using the PredictProtein server.[33] The structure has three domains: domain 1 (helices A–D), domain 2 (helices E–H), and domain 3 (helices L and P). The predicted secondary structure for the 52-aa NTF (domain 3 or "phosphorylation domain"[13]) is formed by the positively charged helices L and P (with phosphorylatable Ser35 and Ser45) connected by a Pro–Ala coil linker. The 8.5-nm IQ motif helix of the myosin heavy chain (Glu812–Phe855) is shown in light gray.

charged ring 2[45] (Glu932–Gln946; see Fig. 3b) of its own S2 (interaction "a"). The flexible fitting also modifies the RLC NTF from its initial predicted structure (Figs. 4 and 5a) to a final structure that better fits the intermolecular tube of density ("b" in Figs. 5b and 6b). An intermolecular interaction occurs between the Leu305–Gln327 interaction loop (I loop) of the free head and Met1–Asp3 of NTF helix L and Pro25–Ser35 of NTF helix P of the adjacent blocked-head RLC ("b" in Fig. 6b). The new fitting also suggests that the intermolecular interaction previously described[18] is between the ELC of the blocked head and the C loop of the adjacent free head ("c" in Figs. 5b and 6a and b).

Atomic model of tarantula thick filament

Figure 6a shows two levels of interacting heads of a relaxed tarantula filament. Four helices[52] of interacting heads occur on the filament surface. The interacting heads appear to be kept in the helically ordered arrangement by the establishment of the nine interactions shown in Fig. 5d. Five interactions are intramolecular: between the CM loop of the free head and its own S2 ("a" in Fig. 5d); between the motor domains of the free and blocked heads[15–18,30] ("d" in Fig. 5d); between the motor domain of the blocked head and the ELC of the free head[16,18] ("e" in Fig. 5d); between the motor domain of the blocked head and its own S2[45] ("f" in Fig. 5d); and between the ELC of the blocked head and its own S2 ("g" in Fig. 5d). Four interactions are intermolecular: between the RLC NTF of the blocked head and the I loop of the adjacent free head ("b" and " b′ " in Fig. 5d); between the ELC of the blocked head and the C loop of the adjacent free head[18] ("c" and " c′ " in Fig. 5d); between the SH3 domain of the blocked head and the S2 of the adjacent myosin molecule[18] ("h" in Fig. 5d); and between the ELC of the blocked head and the backbone (not shown). The five intramolecular interactions presumably stabilize each interacting-head motif as a rigid unit, while the helical organization of motifs is established by formation of the four intermolecular interactions. Intramolecular interaction "a" and the two intermolecular interactions ["h" and the one between the blocked-head ELC and the backbone (not shown) in Fig. 5d] would help anchor each interacting-head motif to the backbone. Taken together, these interactions appear to keep the myosin heads close to the thick filament surface and to minimize actin interaction and ATPase activity in the relaxed state.

Discussion

The three-dimensional reconstruction

The three-dimensional map together with flexible atomic fitting reveals a new intramolecular interaction between the motor domain of the blocked head and its S2 ("a" in Fig. 1) and a new intermolecular interaction between the blocked-head RLC and the free-head motor domain ("b" in Fig. 1), in addition to the interactions previously described[16–18,45] (Fig. 1). The map shows specific density protrusions that match well with the fitted atomic structure (Fig. 5b). The interaction between the blocked-head RLC and the free-head motor domain ("b" in Figs. 1, b and d, and 6a and b) may be a key element in the phosphorylation-mediated release of the myosin heads from the filament surface that occurs on activation.[8]

The density in the RLC region of the reconstruction ("b" in Fig. 1) is explained in part by the greater mass of the tarantula (21,679.6 Da) compared with the smooth muscle RLC. This difference is included in the 5671.4 Da NTF, representing 26% of the RLC mass, which is absent from crystal structures of the myosin head. The tarantula RLC (196 aa) is one of the longest RLCs sequenced so far (111 species; range = 153–222 aa), with its 52-aa NTF belonging to the class of long NTFs (43–61 aa; Fig. 3a; Supplementary Fig. 1). The functional significance of the diversity in NTF structures and lengths (especially between the long and short classes) and its relation to regulation are unknown.

The structure of the interacting-head motif

In the absence of a crystallographic structure for tarantula HMM, we built a hybrid HMM model (Fig. 5a) for atomic fitting by including the crystallographic structure of human cardiac S2,[45] the refined chicken smooth muscle interacting-head structure,[16] and the tarantula RLC-predicted atomic structure (Fig. 4). Flexible fitting of this model produced a final pseudo-

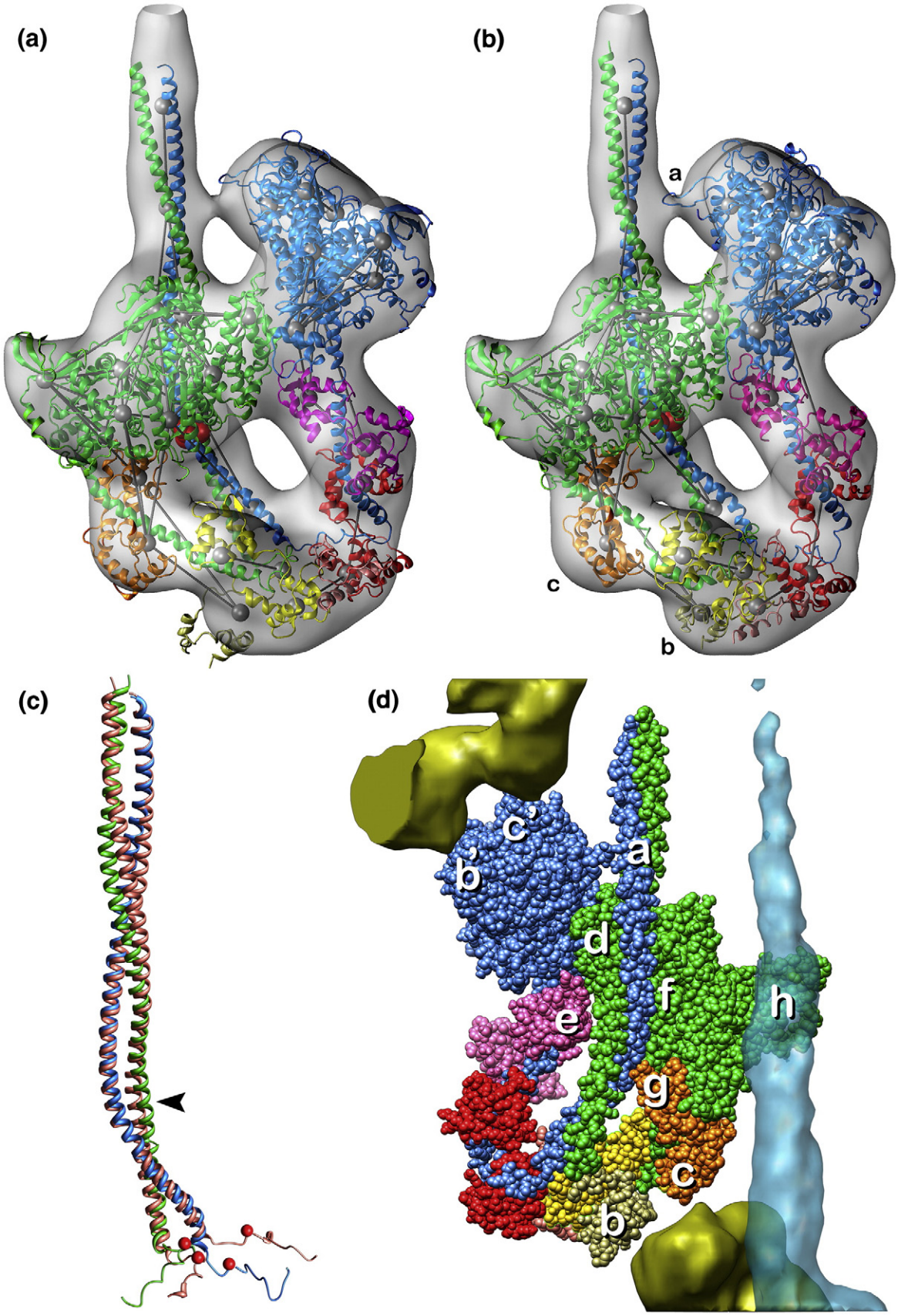


Fig. 5 (*legend on next page*)

atomic model (Fig. 5b) that matched well with the three-dimensional map. The new model (Fig. 5b) agreed well with the general shape of the previous interacting-head atomic models[16,18] (Supplementary Fig. 2). The main differences were in the precise S2 position, in the orientation of the CM loop, and in the RLC regions of both regulatory domains. Because we do not have a complete tarantula crystallographic structure available yet, the precise amino acids involved in the interactions shown in Fig. 5d cannot be deduced from our current atomic fitting, although the loops involved can in principle be identified when their densities protrude in a characteristic way, as discussed below. We are sequencing the tarantula ELC and myosin heavy chain to improve the initial atomic model and the flexible atomic fitting (Zhu *et al.* and Patiño *et al.*, unpublished results).

Both α-helices of the S2 atomic model were connected to the respective heavy-chain α-helices of the regulatory domains of the interacting heads (Fig. 5a and c). The connection of the S2 N-terminal Ser853 (Figs. 3b and 5c) with the C-terminal Gln852 of the head structure shown in Fig. 5a required an unbending and torsion in the 6 and 8 intervening amino acids of the free- and blocked-head S2 N-terminals, respectively (Fig. 5c), as shown in the structure of the head–tail junction (Fig. 5b). The flexing changed the relative position of the coil-coiled α-helices but not the kink position (Met888; Figs. 3b and 5a–c). The final structure showed an excellent fit of the coiled-coil α-helix into the tube of density (Fig. 5a and b). These unbending and torsion are in agreement with the results from crystallographic studies on the 10–14 N-terminal residues of the regulated scallop myosin S2 peptide construct,[53] which showed they were disordered. In contrast, the 126 N-terminal residues of the unregulated human cardiac myosin S2 peptide were found by crystallographic studies to be a straight parallel dimeric coiled coil.[45] Recent crystallographic and thermal studies on scallop myosin S2 peptide constructs[54] suggest that the N-terminal region of the S2 in regulated myosins has less stability than unregulated myosin S2, indicating that the uncoiling of the N-terminal region of the S2 may be key to its function in regulated myosins.[54]

It has been shown previously that helical ordering of the myosin heads in thick filaments requires the "closed" conformation of the switch 2 element of the nucleotide pocket, preventing phosphate release.[55,56] It has also been shown that blebbistatin stabilizes helical ordering by promoting the switch 2 closed state.[57–59] The model shown in Fig. 5b offers further insights into the specific conformation of the blocked and free heads in the relaxed filament. After flexible fitting (Fig. 5b), both head structures remain similar to the closed conformation of the original HMM atomic model used, in agreement with these earlier results (Fig. 6c) and with our previous fitting,[18] showing that the blocked head was in the closed conformation (i.e., MgADP-AlF4, PDB entry 1BR1)[16] (Fig. 6c). However, the free head, also in the closed conformation, shows a difference in the orientation of its heavy-chain regulatory domain, which is midway between the closed 1BR1 and transition 1DFL structures (Fig. 6c). Fitting of the CM loop of the free head to the three-dimensional map required some rearrangement of the 50K domain of the head, including some closure of the cleft between the upper and lower parts of the 50K domain. These structural differences between the free- and blocked-head structures, derived from the three-dimensional map, may have functional implications. One possibility is that the apparent cleft closure and the specific CM loop orientation are required to guide the free head, detached after the power stroke, to establish the precise electrostatic docking interaction "a" (Figs. 1, 5b, and 6a and b) onto ring 2[45] of S2, a

Fig. 5. (a) Starting HMM atomic model used for flexible fitting built using the predicted structure of the tarantula RLC (Fig. 4), the structure of human S2,[45] and the chicken smooth muscle HMM atomic model (without the S2).[16] The kink (red spheres) in the S2 was located at Met888, in agreement with Blankenfeldt *et al.*[45] [see Fig. 3b and panel (c)]. The connected skeleton of 31 positional markers used for flexible fitting, located inside the density subset of the interacting heads, is shown with gray spheres and rods. (b) Final HMM atomic model after flexible fitting. The RLCs of the blocked and free heads are shown in yellow and red, with their NTFs in tan and pink, respectively. The ELCs of the blocked and the free heads are shown in orange and purple, respectively. The heavy chains of the blocked and free heads are shown in green and blue, respectively. The three-dimensional map is shown as a pale gray surface in (a) and (b). (c) Comparison of the crystal structure of the human S2[45] before (red α-helices) and after (blue and green α-helices) the flexible fitting, which only changed the position of the coil-coiled α-helices. The arrowhead indicates that the position of the kink (Met888) did not change before and after the flexible fitting. C-termini are Lys974 and Leu972 (top). N-terminus is Pro849 (red spheres, bottom). The connection of the two α-helices of the S2 with the two α-helices of the interacting-head regulatory domains was done between Ser853 and Gln852. The connection of the S2 N-terminal structure with the C-terminal structure of the heads required some unbending and torsion in the 6 (blocked head) and 8 (free head) intervening amino acids (see Results). (d) The nine interactions of the interacting-head motif in a relaxed thick filament, as viewed toward the filament surface [cf., panels (a) and (b) from above the filament surface]. The five intramolecular interactions ("a," "d," "e," "f," and "g") are between (a) the CM loop of the free-head motor domain (blue) and the S2 of the interacting-head motif (blue), (d) the blocked-head motor domain (green) and the free-head motor domain[16] (blue), (e) the blocked-head motor domain (green) and the free-head ELC[16] (purple), (f) the blocked-head motor domain (green) and the S2 of the interacting-head motif[45] (blue and green), and (g) the blocked-head ELC (orange) and the S2 of the interacting-head motif (blue). Four intermolecular interactions (b/b′, c/c′, h) are between (b/b′) the blocked-head RLC NTF (b, tan; b′, gold) and the I loop of the motor domain of the neighbor free head (b, gold; b′, blue); (c/c′) the blocked-head ELC[18] (c, orange; c′, gold) and the C loop of the neighboring free head (c, gold; c′, blue); and (h) the SH3 of the blocked-head motor domain[18] (green) and the S2 of the neighbor interacting head (pale blue backbone surface). An interaction between the blocked-head ELC (orange) and the backbone is not shown. The three-dimensional map is shown as a pale blue (backbone S2) or gold (up and low interacting-head pairs) surface.

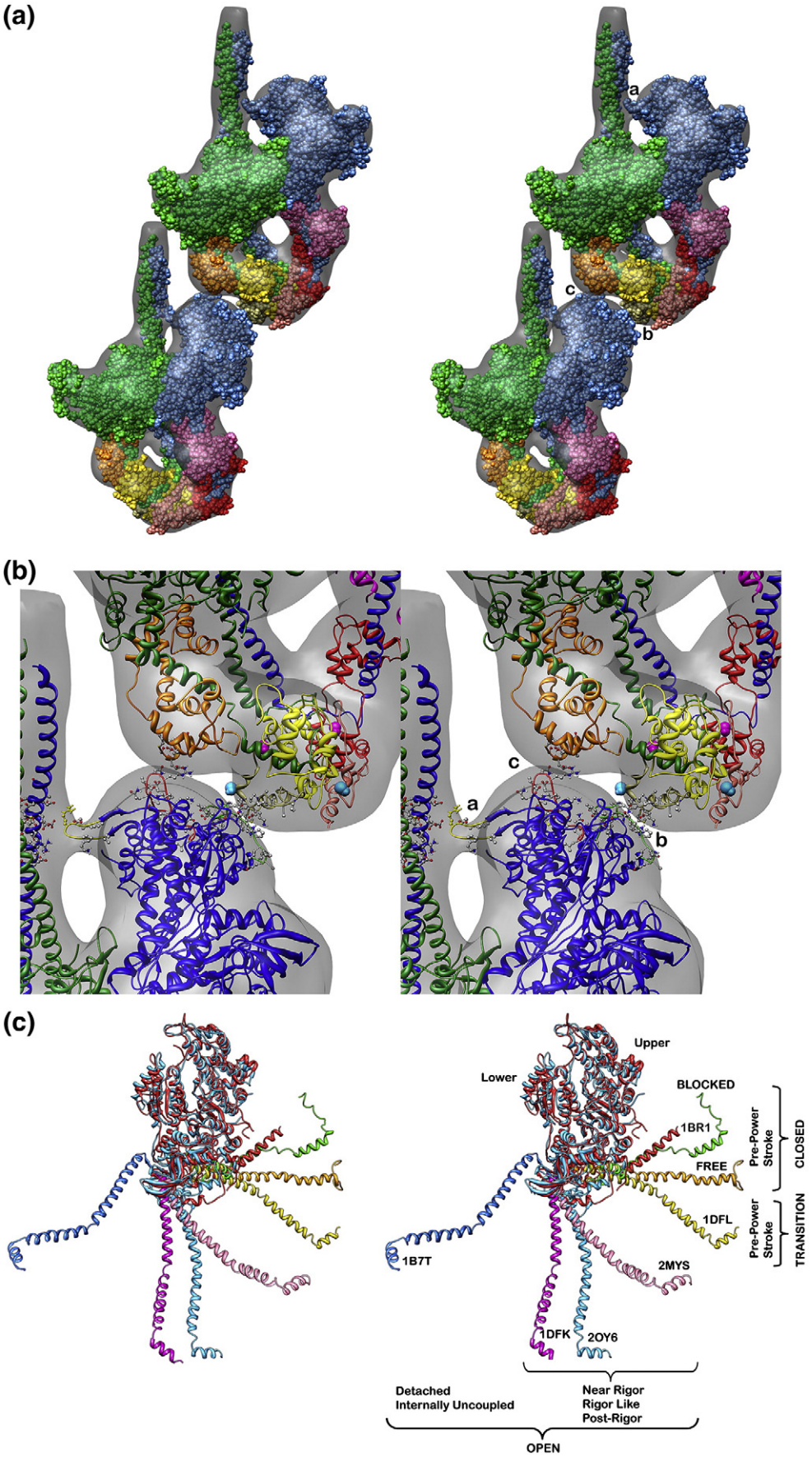


Fig. 6 (*legend on next page*)

requisite for reforming the helices of interacting heads. These structural differences could also preset the order in which the two heads are released on phosphorylation. The free head would be the one conformationally and properly located to be released first to interact with actin, followed by the blocked head (if required).

The structure of the blocked- and free-head regulatory domains

The predicted NTF tertiary structure was positioned in an arbitrary initial configuration (Figs. 4 and 5a) for the two heads and became rearranged during the flexible fitting and the associated MD energy minimization (Figs. 5b, 6a and b, and 7a–c). The linker lengths for both heads increased after the flexing, while the lengths of helices L and P decreased (Table 1). The accuracy of the secondary structure prediction for the NTF is only 80%; therefore, the beginnings and endings of the two α-helices may vary by ±3 aa. As a consequence of the flexing, the tertiary structure of the NTF of the blocked-head RLC appears to be more compact ("a" in Fig. 7) than the free-head NTF ("c" in Fig. 7). The final atomic fitting (Fig. 5b) shows that the positioning of both RLCs on the free and blocked heads is not symmetric, in agreement with the asymmetry of the two-headed motif as a whole, including the asymmetric positioning of Ser35 and Ser45 on the two RLCs (Fig. 6b).

Both phosphorylatable serines of the free-head RLC seem to be freely accessible to the MLCK as they are located in an exposed surface of helix P (Fig. 6b). The corresponding serines on the blocked head are located in the more hidden region of the intermolecular interaction with the neighboring motor domain I loop (Fig. 6b) and a close free-head NTF facing it (Fig. 7b), which could interfere with the MLCK access (Fig. 6b). Differential MLCK access to the free and blocked serines may determine the phosphorylation sequence of the two heads. Two MLCKs have been reported in mouse skeletal muscle:[61] an skMLCK present in adult skeletal muscle and an smMLCK present in smooth, nonmuscle, cardiac, and skeletal muscles. Mass spectrometry results[61] suggest that skMLCK and smMLCK differentially phosphorylate the Ser16 or Ser15 in mouse skeletal muscle, favoring the possibility that two MLCKs may also exist in tarantula striated muscle (Zhu *et al.*, personal communication): one phosphorylating the homologous Ser35 (possibly presetting the basal phosphorylation level) and another phosphorylating the homologous Ser45 (enabling the release of the heads when required after repetitive stimulation).

The new interactions

Flexible fitting of the atomic structure suggests the nature of the two new interactions, an intramolecular interaction between the CM loop of the free head and its own S2 ("a" in Fig. 6b) and an intermolecular interaction between the I loop of the free head and part of helices L and P of the NTF of the adjacent blocked head ("b" in Fig. 6). It also reveals that the interaction between the blocked-head ELC and the adjacent free head[18] occurs specifically between the C loop of the free head and the adjacent ELC ("c" in Fig. 6).

The tube of density between the motor domain of the free head and its own S2 ("a" in Fig. 5d) is well fitted with the CM loop. The positively charged CM loop is located in the tube of density such that it faces ring 2 of the closest α-helix of the coiled coil located ~0.5 nm away. Ring 2 is one of the three negatively charged rings described in the human S2 crystal.[45] This intramolecular interaction could help electrostatically tack the free head to the closest position in the backbone of the thick filament by docking onto its own tail. This intramolecular interaction has not been seen in isolated myosin molecules,[17,19,20] suggesting that its formation requires the presence of a positioning backbone framework. R403 mutants were the first mutations to be shown to cause familial hypertrophic CM by linkage analysis.[62] This finding led to the adoption of the nomenclature "cardiomyopathy loop" to designate the loop on which the residue occurs. The close proximity and possible interaction of four S2 mutations (E924K, E927K, E930K, and E935K) to this CM loop (Fig. 8) may be helpful in understanding how mutations in the rod region of human cardiac myosin lead to disease. One intramolecular interaction ("f" in Fig. 5d) has been reported[45] between ring 1 (Glu905–Asp917; Fig. 3b) of the S2

Fig. 6. (a) Stereo pair of flexible atomic fitting of two adjacent interacting heads on one helical track of the tarantula filament. Three interactions are labeled as intramolecular interaction "a" and intermolecular interactions "b" and "c."[18] The three-dimensional map is shown as a pale gray surface. (b) Stereo pair of intramolecular interaction "a" between the CM or CM loop (yellow ribbon) of the free head (Ile407–Val425 highlighted using balls and sticks; Arg411 residue is in yellow) and its own S2 (blue α-helix), intermolecular interaction "b" between the NTF (tan) of the RLC (yellow) of the blocked head and the neighbor motor domain free-head I loop (Leu305–Gln327 shown in balls and sticks; green ribbon), and intermolecular interaction "c" between the ELC of the blocked head (orange) and the C loop (Gln361–Asp380 shown in balls and sticks; red ribbon) neighbor motor domain free head (blue).[18] Glu932–Glu946 of the negatively charged ring 2[45] of S2 (Fig. 3b) is shown in a ball-and-stick representation. Ser35 and Ser45 are shown as pale blue and pale magenta spheres, respectively. (c) Stereo pair of a comparison of the blocked (green) and free (yellow) heads (Fig. 5b) from tarantula myosin, with the published crystal structures known as the pre-power stroke closed (1BR1)[48] (red) and transition (1DFL)[49] (pale yellow) structures and the near (1DFK)[49]/post (2MYS)[10]/like (2OY6)[50] rigor (pink/light blue/purple) or detached internally uncoupled (1B7T)[51] (dark blue) open structures. The structures of the blocked and free heads used for the flexible fitting were assumed initially to be in the closed conformation[15] (Fig. 5a). For clarity, the ELC and RLC have been removed. Note that the regulatory domain α-helices for 1BR1, 1DFL, and the blocked and free heads are in the same plane. The difference between the blocked- and free-head regulatory domain α-helices is mostly a change in angular orientation.

and loop 2 Trp625–Ser647 of the actin binding interface of the blocked-head motor domain, suggesting that electrostatic interactions play an important role in stabilizing the off state. The intermolecular interactions between the motor domain of the free head and the ELC ("c" and "c′" in Fig. 5d) and RLC ("b" and "b′" in Fig. 5d) of the neighboring blocked head are well fitted with the C loop and the I loop of the free head (Fig. 5d). A third intermolecular interaction ("h" in Fig. 5d)[18] probably helps anchor

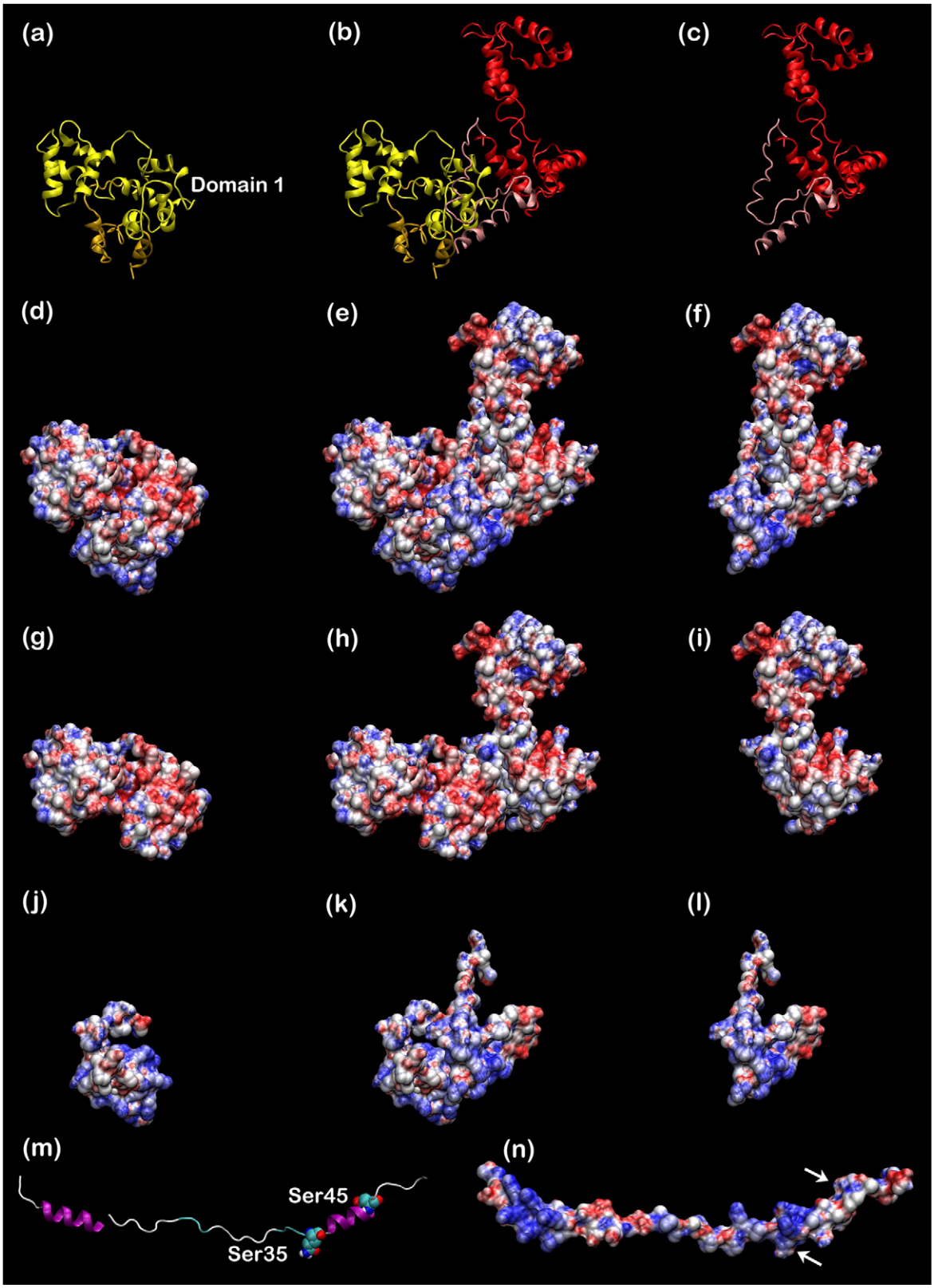


Fig. 7 (*legend on next page*)

the SH3 domain of the blocked head onto the neighboring S2 as it descends toward the thick filament backbone.

Ordering of the heads in relaxed thick filaments

Myosin heads protrude from their helical origins on the filament backbone every 14.5 nm axially.[45,63–65] Once interacting-head motifs have formed at these origins during relaxation, how do these flexibly attached motifs form into the ordered helices of the relaxed state? This presumably involves the intermolecular interactions we have described. Each motif appears to bind to an adjacent S2 via the SH3 motif of its blocked head[18] ("h" in Fig. 5d), while axially adjacent motifs interact via the interactions between the ELC of the blocked head and the C loop of the free head[18] ("c" and "c′" in Fig. 5d) and the NTF of the RLC of the blocked head and the I loop of the adjacent free head ("b" and "b′" in Fig. 5d). These three intermolecular interactions may serve to tack down the flexibly attached heads in regular positions reflecting their origins in the filament backbone,[66] at the same time keeping the heads away from the actin filaments and inhibiting their ATPase activity. This stable conformation of the relaxed thick filament would help maintain the shutdown state.

The structural mechanism of thick filament activation

The structural basis of myosin activation via RLC phosphorylation, leading to release of myosin heads from their helices in the thick filament, remains unknown.[66] The effect of RLC phosphorylation in striated muscle has been explained as a loosening of the myosin heads from the thick filament backbone,[8,25,26,60,67,68] causing disordering. It has been suggested that the restricted mobility of the relaxed state may be due to electrostatic interactions between positively charged amino acids near the phosphorylatable serine of the RLC and an adjacent structure[60] and that reduction of the net charge by phosphorylation would be the simplest mechanism to weaken this effect.[60] The positively charged amino acids in the NTF have been shown to be important for the catalytic activity of MLCK.[69]

The 52-aa-long tarantula NTF belongs to the long N-terminal class (Fig. 3a; Supplementary Fig. 1), which contains many Lys and Arg residues near the N-terminus (Fig. 3a). There are 14 positively charged amino acids in the NTF of tarantula, which, together with the 8 negatively charged amino acids, give a total net positive charge of 6. Phosphorylation of Ser35 could reduce this to 4 positive charges; additional phosphorylation of Ser45, to only 2 positive charges. Phosphorylation could induce the formation of two Arg–phospho-Ser salt bridges in helix P similar to the Arg16–phospho-Ser19 for the smooth muscle RLC,[27] one possibly between Arg42 and phospho-Ser45, and another — less strong — between Lys32 and phospho-Ser35. MD simulations of the de- and phosphorylated tarantula NTFs are in progress to test this point. The structural model proposed by Nelson *et al.*[13] suggests that phosphorylation of the smooth muscle NTF increases its helical order, internal dynamics, and solvent accessibility, causing a decrease of head–head interactions and activating myosin. MD simulations[27] have helped refine this model. It was found that phosphorylation causes a disorder-to-order transition in residues Lys11 to Ala17, converting a dynamically unstructured region to a stable α-helix, with formation of a salt bridge between phospho-Ser19 and Arg16. Atomic fitting of the tarantula reconstruction shows that the dephosphorylated blocked-head NTF fits well into the density, corresponding to the new intermolecular interaction ("b" in Fig. 6b) in the tarantula filament switched-off state. In this case, phosphorylation of the RLC NTF at the homologous Ser45 could also induce a disorder-to-order transition, weakening the intermolecular interaction, as with smooth muscle myosin.

The atomic structure shown in Fig. 5b suggests that the 52-aa blocked-head NTF (27 aa longer than the

Table 1. Extension of the NTF helix L, helix P, and linker before and after flexible fitting

	Helix L	Linker	Helix P
Before flexing	Asp3–Gly21	Gly22–Pro33	Pro34–Gln41
Both heads	18 aa	11 aa	7 aa
After flexing	Glu5–Lys11	Lys12–Pro33	Pro34–Ala40
Blocked head	6 aa	21 aa	6 aa
After flexing	Asp4–Lys15	Lys16–Ser35	Glu36–Arg39
Free head	11 aa	19 aa	3 aa

The table shows starting and ending amino acids and the amino acid lengths.

Fig. 7. The structure of the tarantula RLC in the atomic model following flexible fitting. The top row (a–c) shows the flexed RLCs (in ribbons) from the blocked (a; yellow/orange), blocked and free (b), and free (c; red/pink) heads. The NTF structures are asymmetric: the blocked-head NTF (a; orange) exhibits a more compact conformation than the free head (c; pink). The second, third, and fourth rows show the RLC electrostatic surface charges (blue as positive and red as negative) for the whole RLCs (d–f), the RLC without their NTFs (g–i), and the NTFs alone (j–l). The bottom row (m and n) shows the elongated NTF (m, ribbon; n, surface charge; arrows: Ser35 and Ser45). In relaxed muscle, the NTF of the dephosphorylated blocked-head RLC (a,d) exhibits a compact conformation (j) that enables the interactions shown in Fig. 5d, while the NTF of the basal Ser35 phosphorylated free-head RLC (c and f) is partially elongated (l). Both positively charged L helices are packed—due to complementary charges (see Supplementary Movie 1)—against the blocked-head domain 1 (cf., Ref. 60; see Discussion). If phosphorylation in Ser45 of the free-head RLC produces elongation of its NTF (m and n), this could release the free head, exposing the blocked-head NTF, which in turn could be phosphorylated at Ser45 and elongated (m and n), and the blocked head could be released. This mechanism could explain how RLC phosphorylation weakens the head–head interactions, releasing them so they can interact with the thin filament.

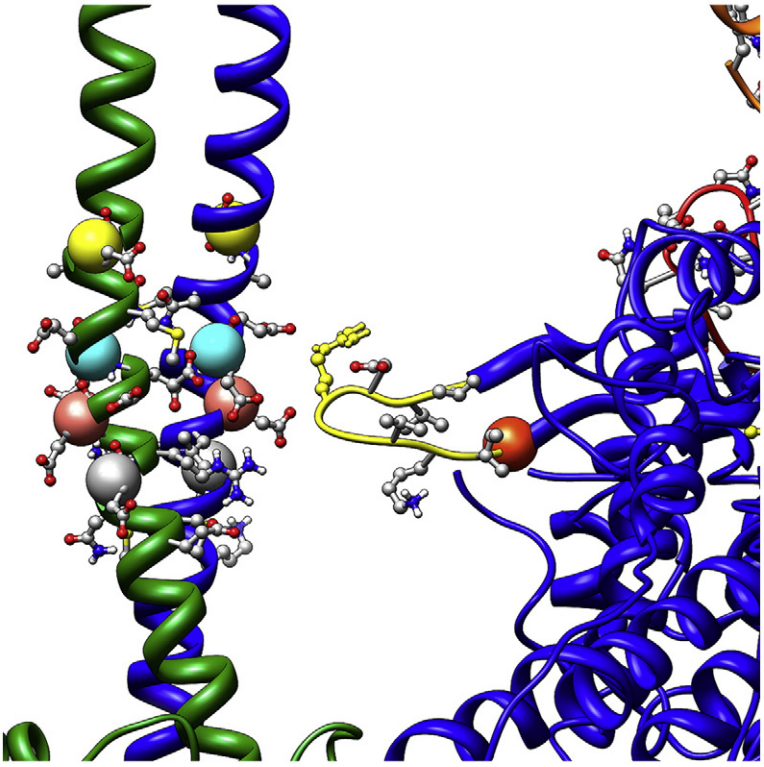


Fig. 8. Five of the reported mutations in the myosin head heavy chain and S2 associated with familial hypertrophic CM. Myosin heavy chain: R403Q, R403L, or R403W (orange sphere) in the CM loop (yellow) and E924K (grey spheres), E927K (pink spheres), E930K (cyan spheres), and E935K (yellow spheres) in the negatively charged ring 2 of S2,[45] facing the CM loop. The side-chain orientations at the interface of the loop may be unreliable as this is not a structure determination down to atomic detail but a prediction at an amino acid level of detail. See the legend to Fig. 6b.

short 25-aa-long smooth muscle NTF) compacts itself (Fig. 7a) in the small intermolecular interaction density "b" shown in Fig. 6b. In contrast, the free-head NTF, which is not engaged in any intermolecular interaction, adopts the structure shown in Fig. 7c, which is less compact than that in the blocked head (Fig. 7a), but not the fully elongated structure expected when it is Ser45 phosphorylated (Fig. 7m and n). Under relaxing conditions, the heads are ~35%–50% monophosphorylated,[46] and liquid chromatography–tandem mass spectrometry shows that this is on Ser35 (Brito *et al.*, unpublished results). It seems likely that this would represent mainly phosphorylation of the free head, based on the greater accessibility of its phosphorylation sites (Fig. 6b). If so, this would imply that 70%–100% of the free heads are basally phosphorylated. This structural information (Figs. 6b and 7a–c), together with EPR[13] and MD evidence,[27] suggests that in tarantula striated muscle, the compact NTF (Fig. 7a) of the unphosphorylated blocked head and the less compact NTF (Fig. 7c) of the Ser35 monophosphorylated free head could elongate when they are Ser45 phosphorylated (Fig. 7m and n), weakening the interactions and releasing the heads.

Putting together biochemical,[8,46] *in vitro* motility (Guerrero *et al.*, unpublished results), and liquid chromatography–tandem mass spectrometry results (Brito *et al.*, unpublished results), we suggest that in the relaxed state, tarantula striated muscles exhibit a degree of basal Ser35 phosphorylation in their free heads (presetting the basal level of force production), which are able to interact rapidly with actin when the troponin–tropomyosin switch on the thin filaments is activated by Ca^{2+}. Phosphorylation of Ser45 of the free and blocked RLCs on longer-term activation could recruit additional heads, increasing force.

Materials and Methods

Sequencing

Total RNA was isolated from the flexor metatarsus longus striated muscle of the legs of the tarantula *A. avicularia* in the presence of Trizol® (Invitrogen) to clone the RLC gene (rlc_{Aa}). An initial cDNA was obtained using oligo d(T) and an internal degenerate primer, designed against the uniquely conserved region of published RLCs. Nucleotide sequence analysis of the 315-bp fragment amplified allowed the design of two gene-specific primers over the carboxy-terminus (GSP2: 5′-GGTGCAAAAGAAGAGGGAGCT-3′). The cDNA for the entire rlc_{Aa} transcript was obtained using these gene-specific reverse primers and a 5′ RACE kit (Invitrogen). The amplified 591-bp product was sequenced and verified with available RLC sequences. The RLC sequences from tarantula striated muscle and other species with long NTFs were aligned using ClustalX[70] (Fig. 3a).

Image processing and three-dimensional reconstruction

Low-dose electron micrographs of frozen–hydrated thick filaments[18] (1008 filament halves) were digitized at 0.248 nm/pixel using a Nikon Super Coolscan 8000 ED scanner. Filaments were aligned with the bare zone at the top to ensure correct polarity in subsequent steps. A total of 15,504 segments, each 62 nm long, with an overlap of 55.8 nm and containing ~40,000 unique pairs of interacting heads went into the reconstruction. Three-dimensional single-particle reconstruction was carried out by a

modification of the IHRSR method[29–31,71,72] using SPIDER.[73] We used the tarantula negatively stained three-dimensional map as an initial reference model,[52,74–76] which was axially rotated, axially shifted, and tilted out of plane[30,31] up to ±12° for projection matching, giving a total of 45×7×13=4095 projections (13 tilted projections ±12° every 2°), 45 reference rotated projections (0°–90°, 2° rotation angle), and seven image axial shifts of 2.2 nm. The new three-dimensional map shown in Fig. 1 combines about 10,700 of 15,504 filament segments, a yield of 69% of included segments. The three-dimensional map was Fourier filtered to 2-nm resolution (see Supplementary Fig. 3a and Methods).

Atomic model for three-dimensional fitting

We created a hybrid atomic model of HMM using the refined chicken smooth muscle HMM atomic model[16] (derived from PDB entry 1I84;[15] atomic coordinates were kindly provided by Dr. Kenneth Taylor) in which we replaced the human cardiac myosin S2 crystallographic structure[45] and both RLC structures with the predicted RLC structure from tarantula striated muscle, including a predicted NTF secondary structure (Fig. 4). In this model, the two heads consist of the high-resolution X-ray structure of the smooth muscle motor domain plus ELC (called MDE) with bound MgADP.AlF^{4-} (PDB entry 1BR1),[48] together with the skeletal muscle RLC and its associated heavy chain (PDB entry 2MYS).[10] The motor domain conformation therefore corresponds to the structure of the myosin head in which the γ-phosphate pocket is closed and the lever arm is "up," approximately perpendicular to the thin filament axis, as reported for the ADP.AlF^{4-}, ADP.Vi, and ADP.BeF$_x$ states.[49,77–79] The HMM model consists of two of these S1s with their COOH termini moved closer together to achieve a better fit, accomplished by an ~20° rotation of the light-chain binding domain.[15,16,18]

S2 atomic model

Wendt *et al.*[15] were uncertain of the position of S2 in their atomic model, while in the thick filament reconstruction,[18] a rod-like volume of density about 30 nm long was observed running from the junction of the light-chain domains toward the backbone and therefore identified as the first portion of S2. The human cardiac S2 crystallographic structure (human cardiac β-myosin II S2-Δ, 126 aa, P838–K963)[45] has been reported, and we have used it (atomic coordinates were kindly provided by Drs. Ilme Schlichting and Wulf Blankenfeldt) to match this rod-like density. This density in the three-dimensional map is not straight (Fig. 1) but has a bend of ~23° about 5.7 nm from the head–tail junction. We made a preliminary visual adjustment of the S2 atomic structure with a kink at Met888[45] to match this bend and the head–tail junction in the three-dimensional map. We then connected the two α-helices of the S2 atomic model to the two α-helices of the interacting heads as shown in Fig. 5a and performed energy minimization of the intervening amino acids by MD using the Chimera MD module.[80]

RLC secondary structure prediction

We predicted the secondary structure of the tarantula myosin RLC from its amino acid sequence (Fig. 2) using 11 program packages: PROFsec,[33] JPRED,[34] PSIPRED,[35] HNN,[36] SAM,[37] SOPMA,[38] Sspro,[39] HMMSTR,[40] NNPred,[41] 3D_PSSM,[42] and PHYRE.[43] The predicted structures were similar in all cases. For tarantula, the myosin RLC included 10 helices: 8 (A–H, Fig. 2) correspond to the two helices of each of four EF-hand motifs as determined by X-ray crystallography[10] and 2, which we have called L and P, are located in the NTF (Fig. 2). The prediction of these new helices, L and P, is in agreement with other secondary structure predictions[27] and MD simulations.[27]

RLC-predicted atomic model

We used the tarantula RLC sequence to obtain the predicted atomic model (Fig. 4) from Tyr53 to Ala196, using the PredictProtein server,[33] choosing the scallop RLC (PDB entry 1B7T)[51] X-ray crystallographic structure. The predicted secondary structure of the 52-aa NTF was used to predict the tertiary structure of this region using *de novo* and template modeling methods.[33] The complete 196-aa tarantula RLC atomic model was assembled bonding both the homology RLC-predicted atomic structure and the *de novo*-predicted structure model of the NTF, as shown in Fig. 4. The NTF model was positioned in an arbitrary initial configuration to allow a further rearrangement during the flexing fitting procedure (Fig. 5a).

Flexible atomic fitting

We first attempted a rigid-body fitting of the atomic model shown in Fig. 5a into the three-dimensional map using the "colores" tool of Situs.[81] This allowed us to create a mask for a single HMM unit by low-pass filtering of the docked atomic structure (yellow in Fig. 1). The mask was used with Situs tools to segment and subtract densities from neighboring symmetry-related subunits[52] to obtain the density of a single HMM from the helical three-dimensional map. This single HMM map can then be compared with the atomic structure. Rigid-body docking of the atomic model to the single HMM map was not satisfactory with respect to the position and direction of the S2 subunit, the RLC conformations of both heads, and the conformation of the 50K domain of the free head, even when performed independently for each structural subunit. Therefore, we subjected the predicted atomic model (Fig. 5b) to flexible docking to characterize the observed changes. The flexible docking procedure is based on a connected "motion capture" network of identified features within the atomic model.[82] The atomic model is allowed to move according to displacements tracked by 31 control points defined by the network in order to find the best match to the cryo-EM map (Fig. 5b). The number of control points was judged to be sufficient for capturing the shape details of the single HMM map that occupies a volume of 620 nm^3 at the isocontour level shown (Fig. 5b). The number of independent pieces of information contained in the three-dimensional map is then $620/2^3=77.5$. This number comprises an upper bound for the number of recognizable features in this particular volume. Our conservative choice of 31 points (corresponding to a spatial resolution of 2.7 nm in our reduced network) was slightly below this upper bound to avoid overfitting of the data.

The longitudinal distance constraints in the motion capture network were assigned manually[83,84] by following the connectivity of the polypeptide chain to ensure robustness of the control points during the shape change. We found by trial and error that robustness was best achieved by affording more flexibility to the RLC regions and to the 50K domain in the free head by eliminating some constraints on the motion in these regions. The

final network used for the automated flexing is shown in Fig. 5b. We performed the flexing by adding a constraint energy function to the Hamiltonian of a MD simulation that penalizes global shape differences between the data sets.[82] In the MD run, we added 116 buried water molecules predicted by DOWSER[85] to the system, resulting in a total system size of 25,715 atoms. We expect that at 2-nm resolution (see Supplementary Fig. 3a), the flexing faithfully reproduces conformational differences with a precision of 0.3 nm if atomic structures are locally conserved.[82] Side chains are rearranged automatically to accommodate global conformational changes. Otherwise, the algorithm leaves the initial structure intact on the local level. Whether this assumption holds depends on the extent of the conformational rearrangement, which is not known *a priori*. However, it has been shown that only about 7% of protein domain rearrangements documented in the PDB are irregular motions where the tertiary structure is significantly perturbed.[86] Therefore, it is plausible, at least for predominantly hinge- and shear-type domain motions exhibited by HMM, that the low-resolution flexible fitting is performed with a precision of single amino acid residues. Inspection of the computer-generated flexed model (not shown) revealed slight discrepancies between a low-pass-filtered atomic model and the full helical three-dimensional map due to the omission of symmetry-related contacts. The flexing is limited to single molecules[47,82,83] and required the above mentioned masking and editing of subunits. Also, the modeled N-terminal RLC region of the blocked head exhibited an unsatisfactory overlap with the corresponding EM density before flexing due to structure prediction (see above) that did not consider shape information. Since the motion capture assumes that the shape of features remains similar during flexing (a requirement for finding a robust representation for the deformation), the flexing initially overcompensated the shape discrepancy with an unrealistic compression of the N-terminal region of the blocked RLC. We therefore slightly modified the original approach for the present case by manually accounting for the initial discrepancies due to symmetry and modeling. Discrepancies were first computed and visualized based on difference maps between low-pass-filtered atomic structures and the three-dimensional map. Subsequently, the network distance constraints were relaxed and control points in the most flexible RLC regions, in the free head, and in the S2 region were moved manually to minimize observed discrepancies. This manual "touch-up" of the computer-generated model was judged independently by three authors (W.W., L.A., and R.P.). The manual difference map-assisted movements of control points were on the order of the docking precision (RMSD = 0.38 nm) and small compared with the overall motion. The final atomic model is shown in Fig. 5b, and the final flexing-induced RMSD in this atomic model was 0.74 nm. The atomic models were displayed using Chimera[80] and Visual Molecular Dynamics.[87] The surface charge was calculated using the program APBS (Adaptative Poisson–Boltzmann Solver)[88] (a VMD plug-in).

Accession numbers

The tarantula myosin RLC DNA sequence has been deposited into the GenBank (accession no. EU090070), the frozen–hydrated tarantula thick filament three-dimensional map has been deposited into the Electron Microscopy Data Bank (code EMD-1535), and the atomic coordinates of the tarantula myosin II interacting heads have been deposited into the Research Collaboratory for Structural Bioinformatics PDB (code 3DTP).

Acknowledgements

This work was supported in part by the National Institutes of Health through grant R01GM62968, Alfred P. Sloan Foundation (BR-4297), and Human Frontier Science Program (RGP0026/2003; to W.W.); the National Institutes of Health through grant AR34711 (to R.C.); and FONACIT, Venezuela, and the Howard Hughes Medical Institute, USA (to R.P.). We thank Dr. Neal Epstein, Dr. Ulf Lunberg, Dr. Jose Reinaldo Guerrero, Reicy Brito, MSc, and Sol Patiño, MSc, for discussions; Dr. Kenneth Taylor for providing the unpublished atomic coordinates of the refined smooth muscle HMM;[16] Dr. Ilme Schlichting and Dr. Wulf Blankenfeldt for providing the atomic coordinates of the human cardiac S2 structure;[45] Dr. D. Thomas and Dr. M. Espinoza-Fonseca for helping with MD simulations of the tarantula RLC NTF; and Dr. Ulrich Meiβner for providing the solid model of the tilted-corrected tarantula thick filament three-dimensional map. We also thank Dr. Michael Schatz (Image Science Software, Berlin, Germany) for his kind help with the implementation of the IMAGIC program. EM was carried out at the Core Electron Microscopy Facility of the University of Massachusetts Medical School, supported in part by the Diabetes Endocrinology Research Center through grant DK32520.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2008.10.013

References

1. Craig, R. & Padrón, R. (2004). Molecular structure of the sarcomere. In (Engel, A. G. & Franzini-Armstrong, C., eds), pp. 129–166, McGraw-Hill, New York, NY.
2. Lehman, W. & Szent-Gyorgyi, A. G. (1975). Regulation of muscular contraction. Distribution of actin control and myosin control in the animal kingdom. *J. Gen. Physiol.* **66**, 1–30.
3. Vibert, P., Craig, R. & Lehman, W. (1997). Steric-model for activation of muscle thin filaments. *J. Mol. Biol.* **266**, 8–14.
4. Poole, K. J., Lorenz, M., Evans, G., Rosenbaum, G., Pirani, A., Craig, R. *et al.* (2006). A comparison of muscle thin filament models obtained from electron microscopy reconstructions and low-angle X-ray fibre diagrams from non-overlap muscle. *J. Struct. Biol.* **155**, 273–284.
5. Szent-Gyorgyi, A. G. (2007). Regulation by myosin: how calcium regulates some myosins, past and present. *Adv. Exp. Med. Biol.* **592**, 253–264.

6. Sweeney, H. L., Bowman, B. F. & Stull, J. T. (1993). Myosin light chain phosphorylation in vertebrate striated muscle: regulation and function. *Am. J. Physiol.* **264**, C1085–C1095.
7. Sellers, J. R. (1981). Phosphorylation-dependent regulation of *Limulus* myosin. *J. Biol. Chem.* **256**, 9274–9278.
8. Craig, R., Padrón, R. & Kendrick-Jones, J. (1987). Structural changes accompanying phosphorylation of tarantula muscle myosin filaments. *J. Cell Biol.* **105**, 1319–1327.
9. Chacko, S., Conti, M. A. & Adelstein, R. S. (1977). Effect of phosphorylation of smooth muscle myosin on actin activation and Ca^{2+} regulation. *Proc. Natl Acad. Sci. USA*, **74**, 129–133.
10. Rayment, I., Rypniewski, W. R., Schmidt-Base, K., Smith, R., Tomchick, D. R., Benning, M. M. *et al.* (1993). Three-dimensional structure of myosin subfragment-1: a molecular motor. *Science*, **261**, 50–58.
11. Sweeney, H. L. & Houdusse, A. (2004). Mammalian muscle myosin. In (Engel, A. G. & Franzini-Armstrong, C., eds), pp. 167–186, McGraw-Hill, New York, NY.
12. Kwon, H., Goodwin, E. B., Nyitray, L., Berliner, E., O'Neall-Hennessey, E., Melandri, F. D. & Szent-Gyorgyi, A. G. (1990). Isolation of the regulatory domain of scallop myosin: role of the essential light chain in calcium binding. *Proc. Natl Acad. Sci. USA*, **87**, 4771–4775.
13. Nelson, W. D., Blakely, S. E., Nesmelov, Y. E. & Thomas, D. D. (2005). Site-directed spin labeling reveals a conformational switch in the phosphorylation domain of smooth muscle myosin. *Proc. Natl Acad. Sci. USA*, **102**, 4000–4005.
14. Wendt, T., Taylor, D., Messier, T., Trybus, K. M. & Taylor, K. A. (1999). Visualization of head–head interactions in the inhibited state of smooth muscle myosin. *J. Cell Biol.* **147**, 1385–1390.
15. Wendt, T., Taylor, D., Trybus, K. M. & Taylor, K. (2001). Three-dimensional image reconstruction of dephosphorylated smooth muscle heavy meromyosin reveals asymmetry in the interaction between myosin heads and placement of subfragment 2. *Proc. Natl Acad. Sci. USA*, **98**, 4361–4366.
16. Liu, J., Wendt, T., Taylor, D. & Taylor, K. (2003). Refined model of the 10S conformation of smooth muscle myosin by cryo-electron microscopy 3D image reconstruction. *J. Mol. Biol.* **329**, 963–972.
17. Burgess, S. A., Yu, S., Walker, M. L., Hawkins, R. J., Chalovich, J. M. & Knight, P. J. (2007). Structures of smooth muscle myosin and heavy meromyosin in the folded, shutdown state. *J. Mol. Biol.* **372**, 1165–1178.
18. Woodhead, J. L., Zhao, F. Q., Craig, R., Egelman, E. H., Alamo, L. & Padrón, R. (2005). Atomic model of a myosin filament in the relaxed state. *Nature*, **436**, 1195–1199.
19. Jung, H. S., Burgess, S. A., Billington, N., Colegrave, M., Patel, H., Chalovich, J. M. *et al.* (2008). Conservation of the regulated structure of folded myosin 2 in species separated by at least 600 million years of independent evolution. *Proc. Natl Acad. Sci. USA*, **105**, 6022–6026.
20. Jung, H. S., Komatsu, S., Ikebe, M. & Craig, R. (2008). Head–head and head–tail interaction: a general mechanism for switching off myosin II activity in cells. *Mol. Biol. Cell*, **19**, 3234–3242.
21. Zoghbi, M. E., Woodhead, J. L., Moss, R. L. & Craig, R. (2008). Three-dimensional structure of vertebrate cardiac muscle myosin filaments. *Proc. Natl Acad. Sci. USA*, **105**, 2386–2390.
22. Zhao, F. Q., Woodhead, J. L. & Craig, R. (2008). Head–head interaction characterizes the relaxed state of scallop and *Limulus* muscle myosin filaments. *Biophys. J.* **94**, 630.
23. Craig, R. & Woodhead, J. L. (2006). Structure and function of myosin filaments. *Curr. Opin. Struct. Biol.* **16**, 204–212.
24. Sellers, J. R. & Knight, P. J. (2008). Folding and regulation in myosins II and V. *J. Muscle Res. Cell Motil.* **28**, 363–370.
25. Levine, R. J., Chantler, P. D., Kensler, R. W. & Woodhead, J. L. (1991). Effects of phosphorylation by myosin light chain kinase on the structure of *Limulus* thick filaments. *J. Cell Biol.* **113**, 563–572.
26. Levine, R. J., Kensler, R. W., Yang, Z., Stull, J. T. & Sweeney, H. L. (1996). Myosin light chain phosphorylation affects the structure of rabbit skeletal muscle thick filaments. *Biophys. J.* **71**, 898–907.
27. Espinoza-Fonseca, L. M., Kast, D. & Thomas, D. D. (2007). Molecular dynamics simulations reveal a disorder-to-order transition on phosphorylation of smooth muscle myosin. *Biophys. J.* **93**, 2083–2090.
28. Ikebe, M. & Morita, J. (1991). Identification of the sequence of the regulatory light chain required for the phosphorylation-dependent regulation of actomyosin. *J. Biol. Chem.* **266**, 21339–21342.
29. Egelman, E. H. (2000). A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy*, **85**, 225–234.
30. Burgess, S., Walker, M., Knight, P. J., Sparrow, J., Schmitz, S., Offer, G. *et al.* (2004). Structural studies of arthrin: monoubiquitinated actin. *J. Mol. Biol.* **341**, 1161–1173.
31. Pomfret, A. J., Rice, W. J. & Stokes, D. L. (2007). Application of the iterative helical real-space reconstruction method to large membranous tubular crystals of P-type ATPases. *J. Struct. Biol.* **157**, 106–116.
32. Maita, T., Chen, J. I. & Matsuda, G. (1981). Amino-acid sequence of the 20 000-molecular-weight light chain of chicken gizzard-muscle myosin. *Eur. J. Biochem.* **117**, 417–424.
33. Rost, B., Yachdav, G. & Liu, J. (2004). The PredictProtein server. *Nucleic Acids Res.* **32**, W321–W326.
34. Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M. & Barton, G. J. (1998). JPred: a consensus secondary structure prediction server. *Bioinformatics*, **14**, 892–893.
35. McGuffin, L. J., Bryson, K. & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics*, **16**, 404–405.
36. Combet, C., Blanchet, C., Geourjon, C. & Deleage, G. (2000). NPS@: network protein sequence analysis. *Trends Biochem. Sci.* **25**, 147–150.
37. Karplus, K., Karchin, R., Barrett, C., Tu, S., Cline, M., Diekhans, M. *et al.* (2001). What is the value added by human intervention in protein structure prediction? *Proteins*, **45(S5)**, 86–91.
38. Geourjon, C. & Deleage, G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput. Appl. Biosci.* **11**, 681–684.
39. Cheng, J., Randall, A. Z., Sweredoski, M. J. & Baldi, P. (2005). SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res.* **33**, W72–W76.
40. Bystroff, C., Thorsson, V. & Baker, D. (2000). HMMSTR: a hidden Markov model for local

sequence–structure correlations in proteins. *J. Mol. Biol.* **301**, 173–190.
41. Kneller, D. G., Cohen, F. E. & Langridge, R. (1990). Improvements in protein secondary structure prediction by an enhanced neural network. *J. Mol. Biol.* **214**, 171–182.
42. Kelley, L. A., MacCallum, R. M. & Sternberg, M. J. (2000). Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.* **299**, 499–520.
43. Bennett-Lovsey, R. M., Herbert, A. D., Sternberg, M. J. & Kelley, L. A. (2008). Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins*, **70**, 611–625.
44. Kemp, B. E. & Pearson, R. B. (1990). Protein kinase recognition sequence motifs. *Trends Biochem. Sci.* **15**, 342–346.
45. Blankenfeldt, W., Thoma, N. H., Wray, J. S., Gautel, M. & Schlichting, I. (2006). Crystal structures of human cardiac beta-myosin II S2-Delta provide insight into the functional role of the S2 subfragment. *Proc. Natl Acad. Sci. USA*, **103**, 17713–17717.
46. Hidalgo, C., Craig, R., Ikebe, M. & Padrón, R. (2001). Mechanism of phosphorylation of the regulatory light chain of myosin from tarantula striated muscle. *J. Muscle Res. Cell Motil.* **22**, 51–59.
47. Wriggers, W. & Chacon, P. (2001). Modeling tricks and fitting techniques for multiresolution structures. *Structure*, **9**, 779–788.
48. Dominguez, R., Freyzon, Y., Trybus, K. M. & Cohen, C. (1998). 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**, 559–571.
49. Houdusse, A., Szent-Gyorgyi, A. G. & Cohen, C. (2000). Three conformational states of scallop myosin S1. *Proc. Natl Acad. Sci. USA*, **97**, 11238–11243.
50. Yang, Y., Gourinath, S., Kovacs, M., Nyitray, L., Reutzel, R., Himmel, D. M. *et al.* (2007). Rigor-like structures from muscle myosins reveal key mechanical elements in the transduction pathways of this allosteric motor. *Structure*, **15**, 553–564.
51. Houdusse, A., Kalabokis, V. N., Himmel, D., Szent-Gyorgyi, A. G. & Cohen, C. (1999). Atomic structure of scallop myosin subfragment S1 complexed with MgADP: a novel conformation of the myosin head. *Cell*, **97**, 459–470.
52. Crowther, R. A., Padrón, R. & Craig, R. (1985). Arrangement of the heads of myosin in relaxed thick filaments from tarantula muscle. *J. Mol. Biol.* **184**, 429–439.
53. Li, Y., Brown, J. H., Reshetnikova, L., Blazsek, A., Farkas, L., Nyitray, L. & Cohen, C. (2003). Visualization of an unstable coiled coil from the scallop myosin rod. *Nature*, **424**, 341–345.
54. Brown, J. H., Yang, Y., Reshetnikova, L., Gourinath, S., Suveges, D., Kardos, J. *et al.* (2008). An unstable head–rod junction may promote folding into the compact off-state conformation of regulated myosins. *J. Mol. Biol.* **375**, 1434–1443.
55. Zoghbi, M. E., Woodhead, J. L., Craig, R. & Padrón, R. (2004). Helical order in tarantula thick filaments requires the "closed" conformation of the myosin head. *J. Mol. Biol.* **342**, 1223–1236.
56. Xu, S., Offer, G., Gu, J., White, H. D. & Yu, L. C. (2003). Temperature and ligand dependence of conformation and helical order in myosin filaments. *Biochemistry*, **42**, 390–401.
57. Xu, S., Gu, J., White, H., Offer, G. & Yu, L. (2005). Structural effects on myosin of three ATPase inhibitors. *Biophys. J.* **88**, 124a.
58. Xu, S., White, H. D., Offer, G. W. & Yu, L. C. (2008). Stabilisation of the helical order of myosin filaments by blebbistatin. *Biophys. J.* **94**, 624.
59. Zhao, F. Q., Padrón, R. & Craig, R. (2008). Blebbistatin stabilizes the helical order of myosin filaments by promoting the switch 2 closed state. *Biophys. J.* **95**, 3322–3329.
60. Sweeney, H. L., Yang, Z., Zhi, G., Stull, J. T. & Trybus, K. M. (1994). Charge replacement near the phosphorylatable serine of the myosin regulatory light chain mimics aspects of phosphorylation. *Proc. Natl Acad. Sci. USA*, **91**, 1490–1494.
61. Zhi, G., Ryder, J. W., Huang, J., Ding, P., Chen, Y., Zhao, Y. *et al.* (2005). Myosin light chain kinase and myosin phosphorylation effect frequency-dependent potentiation of skeletal muscle contraction. *Proc. Natl Acad. Sci. USA*, **102**, 17519–17524.
62. Geisterfer-Lowrance, A. A., Kass, S., Tanigawa, G., Vosberg, H. P., McKenna, W., Seidman, C. E. & Seidman, J. G. (1990). A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation. *Cell*, **62**, 999–1006.
63. Huxley, H. E. & Brown, W. (1967). The low-angle x-ray diagram of vertebrate striated muscle and its behaviour during contraction and rigor. *J. Mol. Biol.* **30**, 383–434.
64. McLachlan, A. D. & Karn, J. (1982). Periodic charge distributions in the myosin rod amino acid sequence match cross-bridge spacings in muscle. *Nature*, **299**, 226–231.
65. Decker, B. & Kellermayer, M. S. (2008). Periodically arranged interactions within the myosin filament backbone revealed by mechanical unzipping. *J. Mol. Biol.* **377**, 307–310.
66. Ikebe, M. (2008). Regulation of the function of mammalian myosin and its conformational change. *Biochem. Biophys. Res. Commun.* **369**, 157–164.
67. Metzger, J. M., Greaser, M. L. & Moss, R. L. (1989). Variations in cross-bridge attachment rate and tension with phosphorylation of myosin in mammalian skinned skeletal muscle fibers. Implications for twitch potentiation in intact muscle. *J. Gen. Physiol.* **93**, 855–883.
68. Barany, K., Barany, M., Gillis, J. M. & Kushmerick, M. J. (1979). Phosphorylation–dephosphorylation of the 18,000-dalton light chain of myosin during the contraction–relaxation cycle of frog muscle. *J. Biol. Chem.* **254**, 3617–3623.
69. Kemp, B. E. & Stull, J. T. (1990). Myosin light chain kinases. In (Kemp, B. E., ed.), pp. 115–133, CRC, Boca Raton, FL.
70. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882.
71. Egelman, E. H. (2007). The iterative helical real space reconstruction method: surmounting the problems posed by real polymers. *J. Struct. Biol.* **157**, 83–94.
72. Egelman, E. H. (2007). Single-particle reconstruction from EM images of helical filaments. *Curr. Opin. Struct. Biol.* **17**, 556–561.
73. Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M. & Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* **116**, 190–199.

74. Padrón, R., Alamo, L., Guerrero, J. R., Granados, M., Uman, P. & Craig, R. (1995). Three-dimensional reconstruction of thick filaments from rapidly frozen, freeze-substituted tarantula muscle. *J. Struct. Biol.* **115**, 250–257.
75. Padrón, R., Alamo, L., Murgich, J. & Craig, R. (1998). Towards an atomic model of the thick filaments of muscle. *J. Mol. Biol.* **275**, 35–41.
76. Offer, G., Knight, P. J., Burgess, S. A., Alamo, L. & Padrón, R. (2000). A new model for the surface arrangement of myosin molecules in tarantula thick filaments. *J. Mol. Biol.* **298**, 239–260.
77. Gulick, A. M., Bauer, C. B., Thoden, J. B. & Rayment, I. (1997). X-ray structures of the MgADP, MgATPgammaS, and MgAMPPNP complexes of the *Dictyostelium discoideum* myosin motor domain. *Biochemistry*, **36**, 11619–11628.
78. Fisher, A. J., Smith, C. A., Thoden, J. B., Smith, R., Sutoh, K., Holden, H. M. & Rayment, I. (1995). X-ray structures of the myosin motor domain of *Dictyostelium discoideum* complexed with MgADP.BeFx and MgADP.AlF^{4-}. *Biochemistry*, **34**, 8960–8972.
79. Bauer, C. B., Holden, H. M., Thoden, J. B., Smith, R. & Rayment, I. (2000). X-ray structures of the apo and MgATP-bound states of *Dictyostelium discoideum* myosin motor domain. *J. Biol. Chem.* **275**, 38494–38499.
80. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
81. Chacon, P. & Wriggers, W. (2002). Multi-resolution contour-based fitting of macromolecular structures. *J. Mol. Biol.* **317**, 375–384.
82. Wriggers, W., Chacon, P., Kovacs, J., Tama, T. & Birmanns, S. (2004). Topology representing neural networks reconcile biomolecular shape, structure, and dynamics. *Neurocomputing*, **56**, 365–379.
83. Darst, S. A., Opalka, N., Chacon, P., Polyakov, A., Richter, C., Zhang, G. & Wriggers, W. (2002). Conformational flexibility of bacterial RNA polymerase. *Proc. Natl Acad. Sci. USA*, **99**, 4296–4301.
84. Opalka, N., Chlenov, M., Chacon, P., Rice, W. J., Wriggers, W. & Darst, S. A. (2003). Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell*, **114**, 335–345.
85. Zhang, L. & Hermans, J. (1996). Hydrophilicity of cavities in proteins. *Proteins*, **24**, 433–438.
86. Gerstein, M. & Krebs, W. (1998). A database of macromolecular motions. *Nucleic Acids Res.* **26**, 4280–4290.
87. Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graphics*, **14**, 33–38.
88. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.

Note added in proof. It has come to our attention, after this paper was submitted, that a paper published by Espinoza-Fonseca, L. M., Kast, D. & Thomas, D. D. (2008). *J. Am. Chem. Soc.* **130**, 12208–12209, proposed that entropically balanced disorder-order transitions are a common theme in phosphorylation-induced conformational shifts involved in cell signaling. cf., Fig. 7.