PNAS
www.pnas.org

Supplementary Information for
The Myosin II Coiled-Coil Domain Atomic Structure in its Native Environment

Hamidreza Rahmani, Wen Ma, Zhongjun Hu, Nadia Daneshparvar, Dianne W. Taylor, J. Andrew McCammon, Thomas C. Irving, Robert J. Edwards and Kenneth A. Taylor

Kenneth A. Taylor
Email: taylor@bio.fsu.edu

This PDF file includes:

Supplemental Figures S1 & S2. Sequence alignment of representative striated and smooth muscle myosins.
Supplemental Figure S3. PairCoil analysis
Supplemental Figure S4. Coiled-coil azimuthal and radial changes
Supplemental Figure S5. Expanded comparison between the atomic model and the crystal structure
Supplemental Figure S6. Symmetry of the two heavy chains
Supplemental Figure S7. Atomic model quality analysis and Ramachandran plot
Molecular Dynamics simulations
Supplemental Figure S8. Comparison of coiled coil parameters at the skip residues
Supplemental Figure S9. C-terminal non-helical extension
Supplemental Figure S10. Conservation of human cardiac and Lethocerus flight muscle myosin tails
Conservation of the striated muscle KGGKK motif
Sample Preparation
Supplemental Figures S11 and S12. Data Collection/Preprocessing
Supplemental Figure S13 and S14. Classification and Refinement
Supplemental Figure S15. Resolution and Sharpening
Helical parameters and extension
Supplemental Figure S16. Building and validating the Atomic Model
Coiled-coil Analysis
Fiber Diffraction
Movie Legends
References

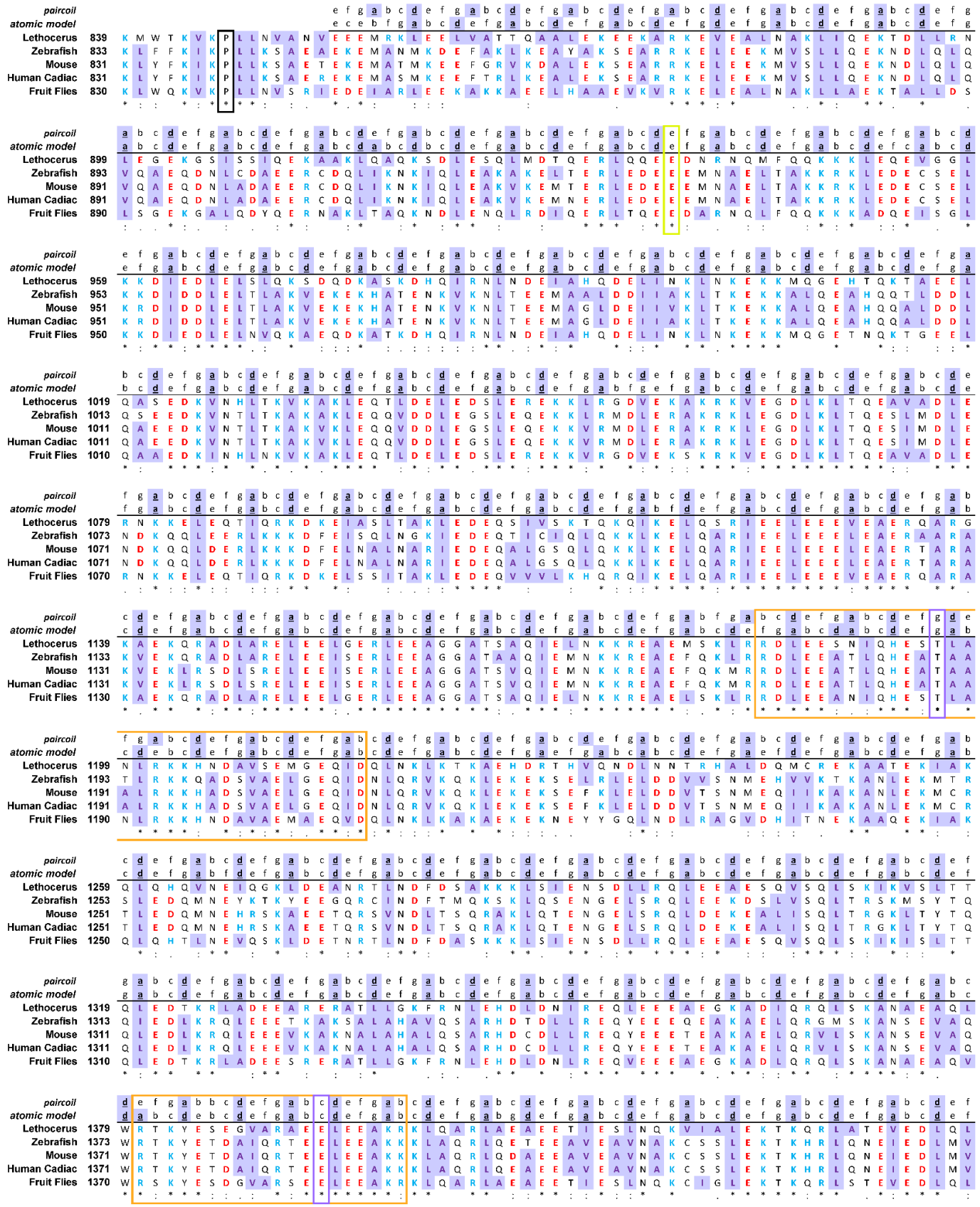
Other supplementary materials for this manuscript include the following:

Movies S1 to S8

Sequence alignment of representative striated and smooth muscle myosins

Sequences were obtained from the NCBI server to show the high level of conservation between myosin tail sequences. The heptad pattern with 4 skips is conserved and makes the myosin tail sensitive to insertion or deletions. Consequently, differences between vastly different species are limited to single amino acid substitutions. The alignment was performed in ClustaX 2.1 and the output was prepared in the online interface of Clustal Omega (1). However, the coloring and formatting and annotation was done manually in Microsoft Excel to emphasize structurally relevant sections (with color-coding consistent with our figures) and the location of hydrophobic residues (in purple) and the conservation of charged residues (colored in red and blue). The top row shows the heptad profile predicted based on the sequence by the software PairCoil 2 (2) and below that are the observed position from the atomic model. There are few instances where the heptad position is not the same in two chains after fitting the atomic model. This is because the real coiled coil is often a little different when compared to an idealized one and the C_α atoms can be quite close to the border between two theoretical positions. For simplicity only positions for the free head heavy chains are shown here (Fig. S1).

Alignment of the *Lethocerus* and human cardiac sequences with chicken smooth and human non-muscle myosin II (Fig. S2) shows a single residue deletion prior to the conventional Skip 2 location effectively negating the effect of Skip 2. Smooth and non-muscle myosin also lack the KGGKK motif at Skip 4 replacing it with two unrelated sequences lacking any glycine. Although chicken smooth and human non-muscle myosin have long non-helical C-terminal extensions, they lack the strongly acidic character of *Lethocerus* and human cardiac myosin.


paircoil
atomic model
Lethocerus
Zebrafish
Mouse
Human Cadiac
Fruit Flies

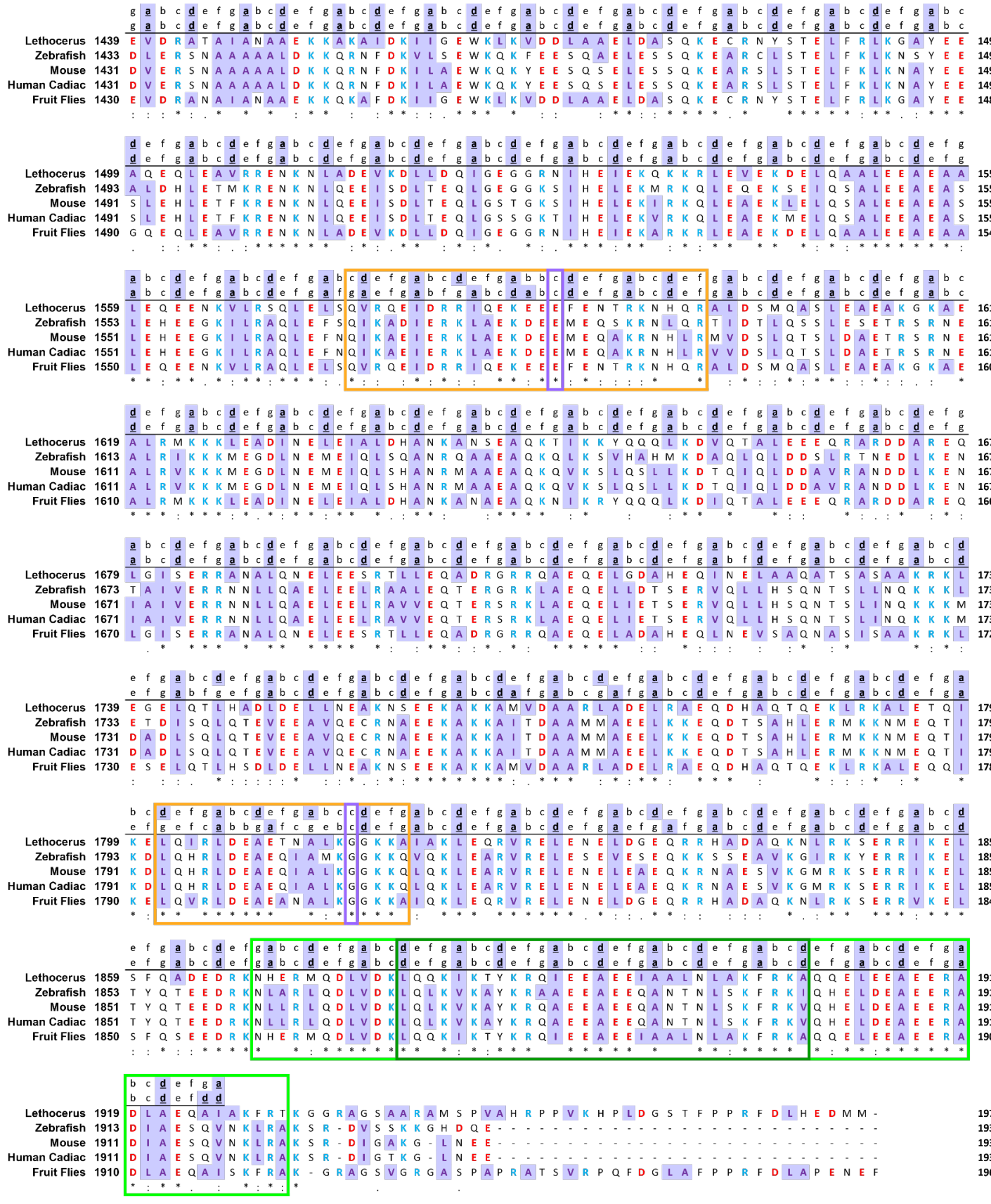
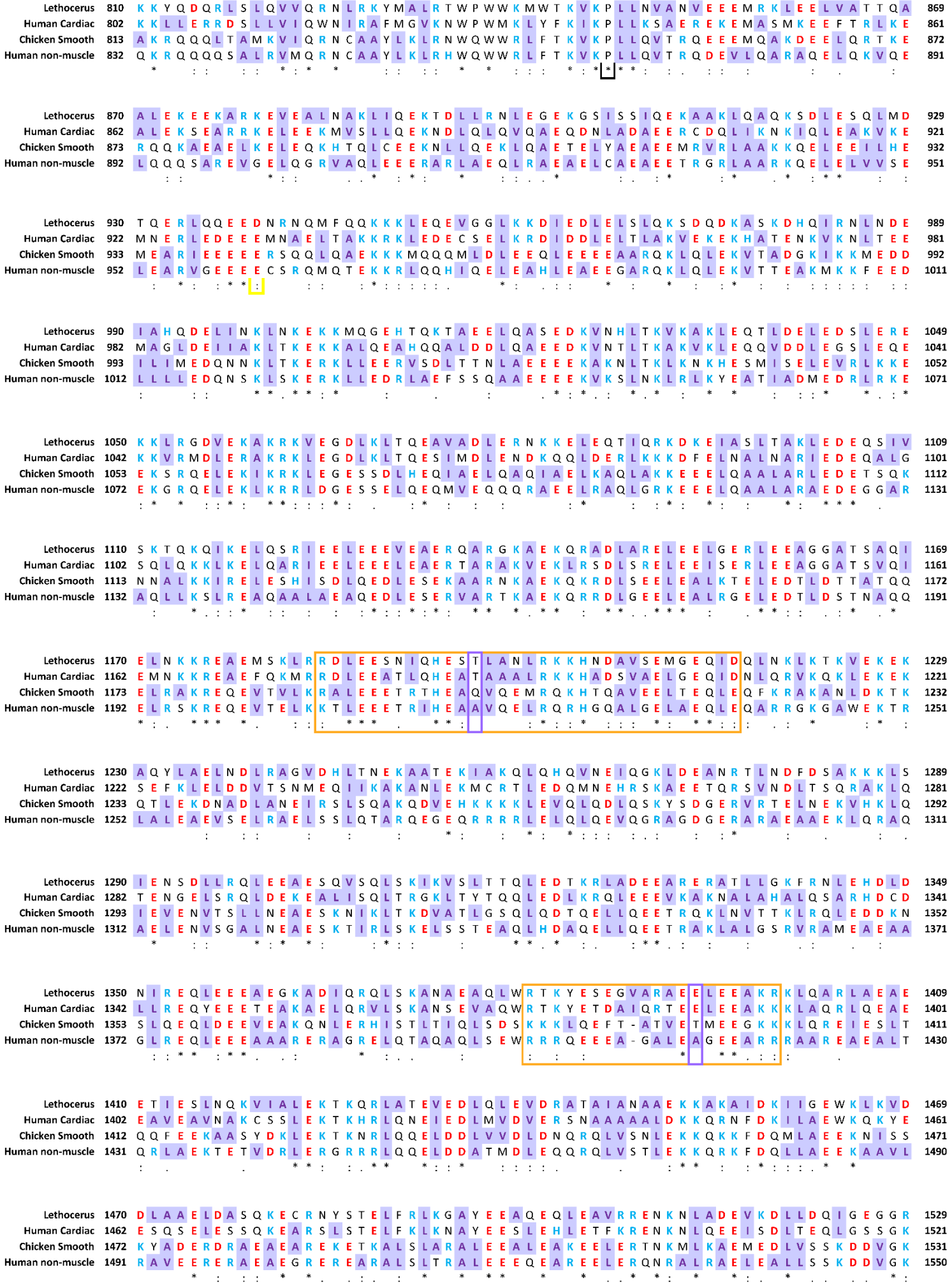

Lethocerus 1439 EVDRATAIANAAEKKAKAIDKIIGEWKLKVDDLAAELDASQKECRNYSTELFRLKGAYEE 1498
Zebrafish 1433 DLERSNAAAAAALDKKQRNFDKVLSEWKQKFEESQAELESSQKEARCLSTELFKLKNSYEE 1492
Mouse 1431 DVERSNAAAAAALDKKQRNFDKILAEWKQKYEESQSELESSQKEARSLSTELFKLKNAYEE 1490
Human Cadiac 1431 DVERSNAAAAAALDKKQRNFDKILAEWKQKYEESQSELESSQKEARSLSTELFKLKNAYEE 1490
Fruit Flies 1430 EVDRANAIANAAEKKQKAFDKIIGEWKLKVDDLAAELDASQKECRNYSTELFRLKGAYEE 1489
Lethocerus 1499 AQEQLEAVRRENKNLADEVKDLLDQIGEGGRNIHEIEKQKKRLEVEKDELQAALEEAEAA 1558
Zebrafish 1493 ALDHLETMKRENKNLQEEISDLTEQLGEGGKSIHELEKMRKQLEQEKSEIQSALEEAEAS 1552
Mouse 1491 SLEHLETFKRENKNLQEEISDLTEQLGSTGKSIHELEKIRKQLEAEKLELQSALEEAEAS 1550
Human Cadiac 1491 SLEHLETFKRENKNLQEEISDLTEQLGSSGKTIHELEKVRKQLEAEKMELQSALEEAEAS 1550
Fruit Flies 1490 GQEQLEAVRRENKNLADEVKDLLDQIGEGGRNIHEIEKARKRLEAEKDELQAALEEAEAA 1549
Lethocerus 1559 LEQEENKVLRSQLELSQVRQEIDRRIQEKEEEFENTRKNHQRALDSMQASLEEAEAKGKAE 1618
Zebrafish 1553 LEHEEGKILRAQLEFSQIKADIERKLAEKDEEMEQSKRNLQRTIDTLQSSLESETRSRNE 1612
Mouse 1551 LEHEEGKILRAQLEFNQIKAEIERKLAEKDEEMEQAKRNHLRMVDSLQTSLDAETRSRNE 1610
Human Cadiac 1551 LEHEEGKILRAQLEFNQIKAEIERKLAEKDEEMEQAKRNHLRVVDSLQTSLDAETRSRNE 1610
Fruit Flies 1550 LEQEENKVLRAQLELSQVRQEIDRRIQEKEEEFENTRKNHQRALDSMQASLEEAEAKGKAE 1609
Lethocerus 1619 ALRMKKKLEADINELEIALDHANKANSEAQKTIKKYQQQLKDVQTALEEEQRARDDAREQ 1678
Zebrafish 1613 ALRIKKKMEGDLNEMEIQLSQANRQAAEAQKQLKSVHAHMKDAQLQLDDSLRTNEDLKEN 1672
Mouse 1611 ALRVKKKMEGDLNEMEIQLSHANRMAAEAQKQVKSLQSLLKDTQIQLDDAVRANDDLKEN 1670
Human Cadiac 1611 ALRVKKKMEGDLNEMEIQLSHANRMAAEAQKQVKSLQSLLKDTQIQLDDAVRANDDLKEN 1670
Fruit Flies 1610 ALRMKKKLEADINELEIALDHANKANAEAQKNIKRYQQQLKDIQTALEEEQRARDDAREQ 1669
Lethocerus 1679 LGISERRANALQNELEESRTLLEQADRGRRQAEQELGDAHEQINELAAQATSASAAKRKL 1738
Zebrafish 1673 TAIVERRNNLLQAELEELRAALEQTERGRKLAEQELLDTSERVQLLHSQNTSLLNQKKKL 1732
Mouse 1671 IAIVERRNNLLQAELEELRAVVEQTERSRKLAEQELIETSERVQLLHSQNTSLINQKKKM 1730
Human Cadiac 1671 IAIVERRNNLLQAELEELRAVVEQTERSRKLAEQELIETSERVQLLHSQNTSLINQKKKM 1730
Fruit Flies 1670 LGISERRANALQNELEESRTLLEQADRGRRQAEQELADAHEQLNEVSAQNASISAAKRKL 1729
Lethocerus 1739 EGELQTLHADLDELLNEAKNSEEKAKKAMVDAARLADELRAEQDHAQTQEKLRKALETQI 1798
Zebrafish 1733 ETDISQLQTEVEEAVQECRNAEEKAKKAITDAAMMAEELKKEQDTSAHLERMKKNMEQTI 1792
Mouse 1731 DADLSQLQTEVEEAVQECRNAEEKAKKAITDAAMMAEELKKEQDTSAHLERMKKNMEQTI 1790
Human Cadiac 1731 DADLSQLQTEVEEAVQECRNAEEKAKKAITDAAMMAEELKKEQDTSAHLERMKKNMEQTI 1790
Fruit Flies 1730 ESELQTLHSDLDELLNEAKNSEEKAKKAMVDAARLADELRAEQDHAQTQEKLRKALEQQI 1789
Lethocerus 1799 KELQIRLDEAETNALKGGKKAIAKLEQRVRELENELDGEQRRHADAQKNLRKSERRIKEL 1858
Zebrafish 1793 KDLQHRLDEAEQIAMKGGKKQVQKLEARVRELESEVESEQKKSSEAVKGIRKYERRIKEL 1852
Mouse 1791 KDLQHRLDEAEQIALKGGKKQLQKLEARVRELENELEAEQKRNAESVKGMRKSERRIKEL 1850
Human Cadiac 1791 KDLQHRLDEAEQIALKGGKKQLQKLEARVRELENELEAEQKRNAESVKGMRKSERRIKEL 1850
Fruit Flies 1790 KELQVRLDEAEANALKGGKKAIQKLEQRVRELENELDGEQRRHADAQKNLRKSERRVKEL 1849
Lethocerus 1859 SFQADEDRKNHERMQDLVDKLQQKIKTYKRQIEEAEEIAALNLAKFRKAQQELEEAEERA 1918
Zebrafish 1853 TYQTEEDRKNLARLQDLVDKLQLKVKAYKRAAEEAEEQANTNLSKFRKIQHELDEAEERA 1912
Mouse 1851 TYQTEEDRKNLLRLQDLVDKLQLKVKAYKRQAEEAEEQANTNLSKFRKVQHELDEAEERA 1910
Human Cadiac 1851 TYQTEEDRKNLLRLQDLVDKLQLKVKAYKRQAEEAEEQANTNLSKFRKVQHELDEAEERA 1910
Fruit Flies 1850 SFQSEEDRKNHERMQDLVDKLQQKIKTYKRQIEEAEEIAALNLAKFRKAQQELEEAEERA 1909
Lethocerus 1919 DLAEQAIAKFRTKGGRAGSAARAMSPVAHRPPVKHPLDGSTFPPRFDLHEDMM- 1971
Zebrafish 1913 DIAESQVNKLRAKSR-DVSSKKGHDQE---------- 1938
Mouse 1911 DIAESQVNKLRAKSR-DIGAKG-LNEE---------- 1935
Human Cadiac 1911 DIAESQVNKLRAKSR-DIGTKG-LNEE---------- 1935
Fruit Flies 1910 DLAEQAISKFRAK-GRAGSVGRGASPAPRATSVRPQFDGLAFPPRFDLAPENEF 1962

Fig. S1. Sequence alignment between *Lethocerus indicus* (ASP18627.1) and other landmark species: Human (NP_000248.2), Mouse(NP_542766.1), Fruit flies (*Drosophila melanogaster* NP_001246050.1), and Zebrafish (*Danio rerio*; NP_001106204.1). Boxed areas are colored as follows: Accommodation Regions are orange, the Assembly Competence Domain is dark green, the extended domain is light green, skip residues are purple, yellow denotes where the S2 bending starts and the invariant proline is black.


Lethocerus
Human Cardiac
Chicken Smooth
Human non-muscle

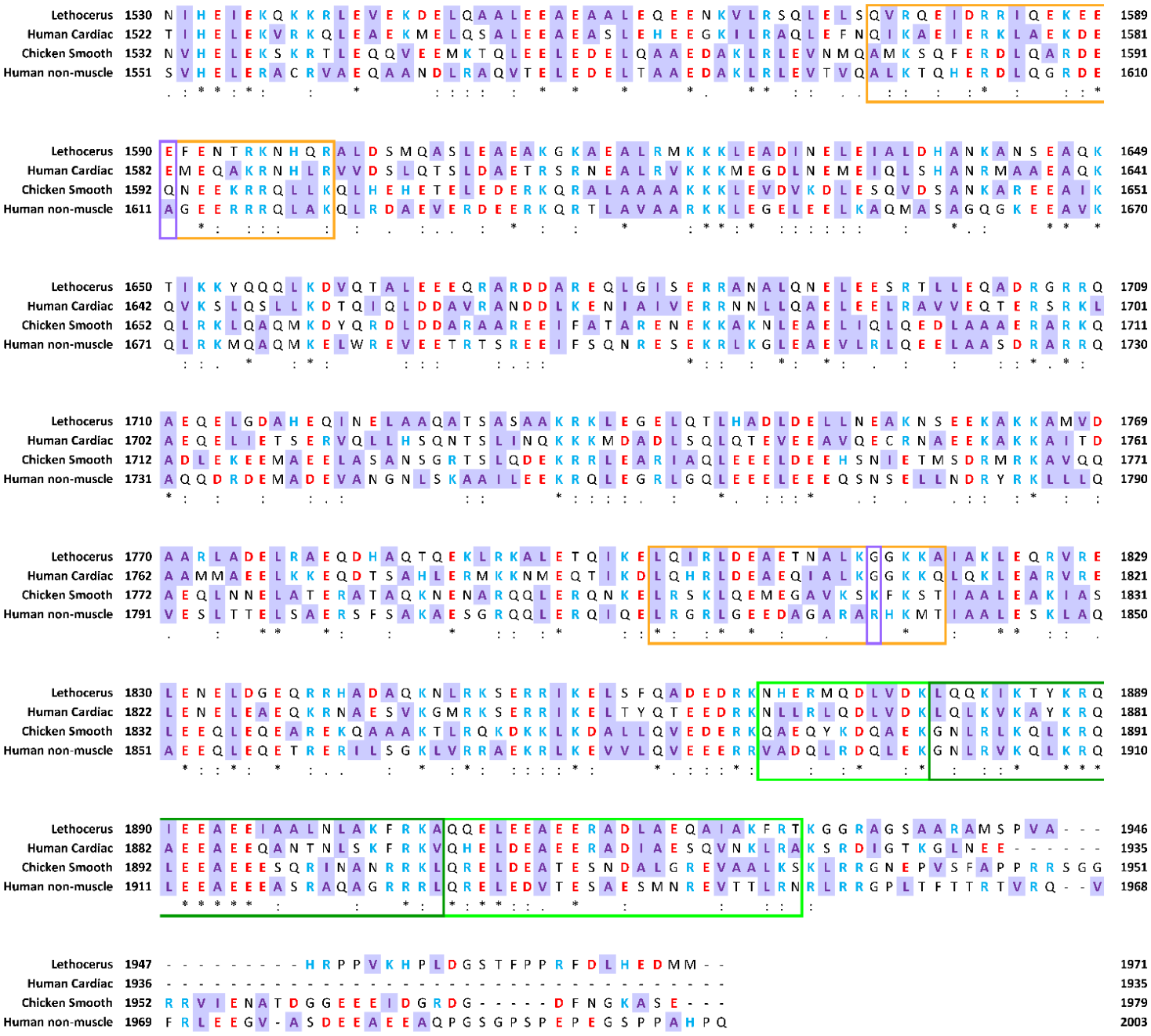


Fig. S2. Alignment between *Lethocerus* myosin (ASP18627.1), Human (NP_000248.2), Chicken Smooth Muscle (NP_990605.2), and Human non-muscle myosin (NP_001070654). Striated myosins have an insertion at 1388 which introduces Skip 2 to the heptad of cardiac, skeletal and insect flight muscle. The predicted Skips are highlighted with purple boxes. The accommodation regions are boxed for the four skip residues (gold) as well as the ACD (light and dark green).

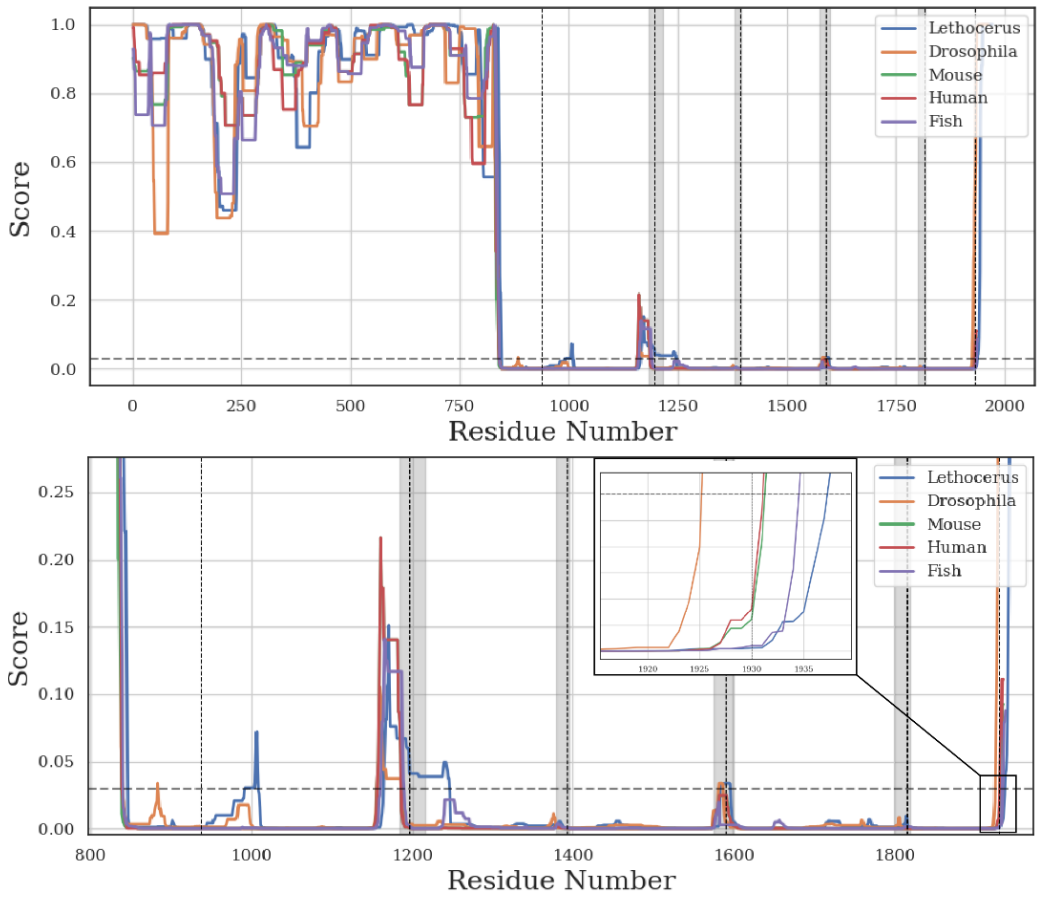


Fig. S3. PairCoil scores for myosin sequences from landmark species. Top panel shows the score for the entire sequence. Bottom panel is an enlarged version showing only the tail scores. (A) The Paircoil score shows the coiled-coil starting at the tail region. This score is calculated in a moving average with a window of 28 residues. The 0.03 threshold corresponds to sensitivity 0.730 and specificity of 0.998 (2). The predicted profile shows the 4 Skips. (B) The score increases above the threshold especially close to Skips 1 and 3, however the coiled-coil is not broken in the structure. Note presence of a skip residue often means a worse coiled-coil prediction but does not necessarily change the PairCoil score significantly since the score is calculated over a 28 residue window. Vertical dotted lines show in order: end of the proximal S2, Skips 1-4 and the end of the coiled-coil. The first vertical solid line from 1.0 to 0.0 is the head-tail junction at residue 850. The last vertical solid line is the end of the tail α-helix at residue 1931. The inset shows how the score declines starting at 1930 but the probability of the coiled-coil is still high at residue 1937. This kind of discrepancy is expected because of the 28 residue window for fitting.

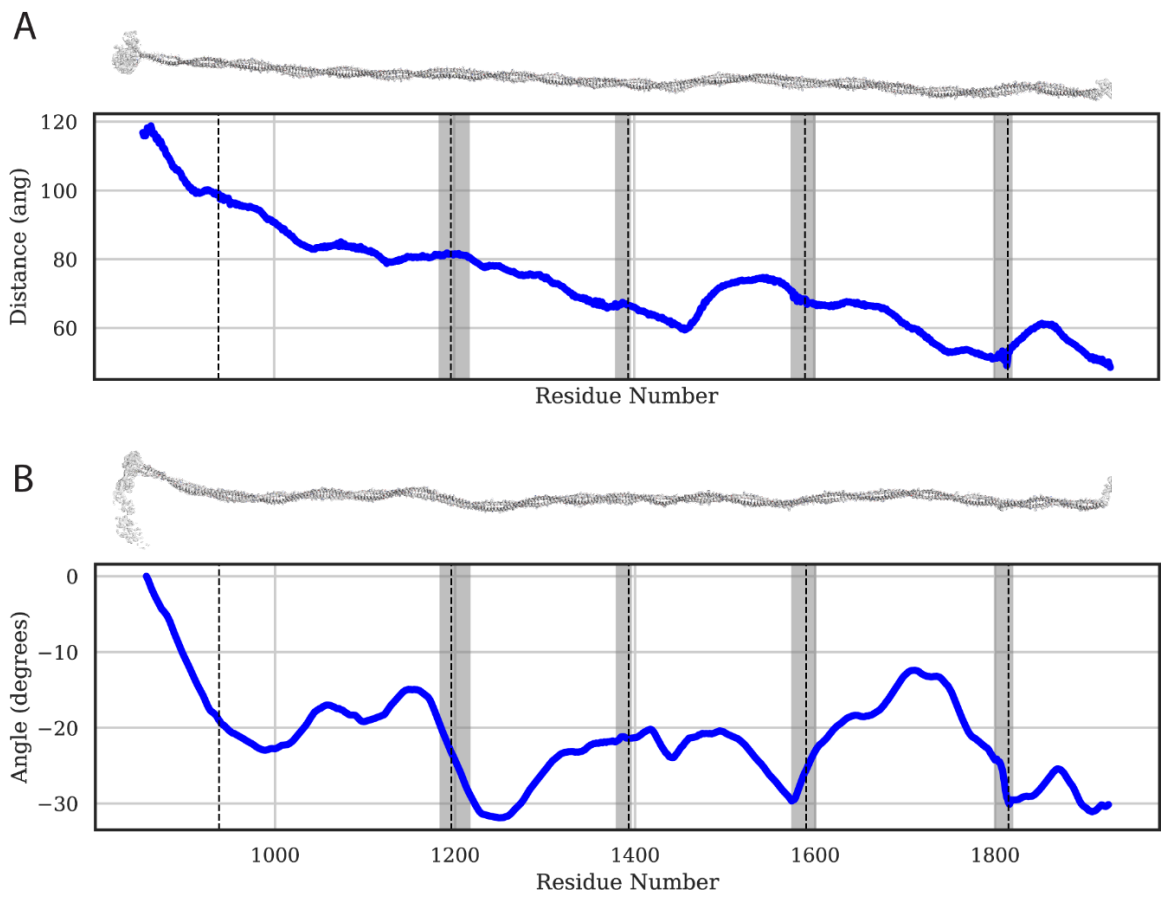


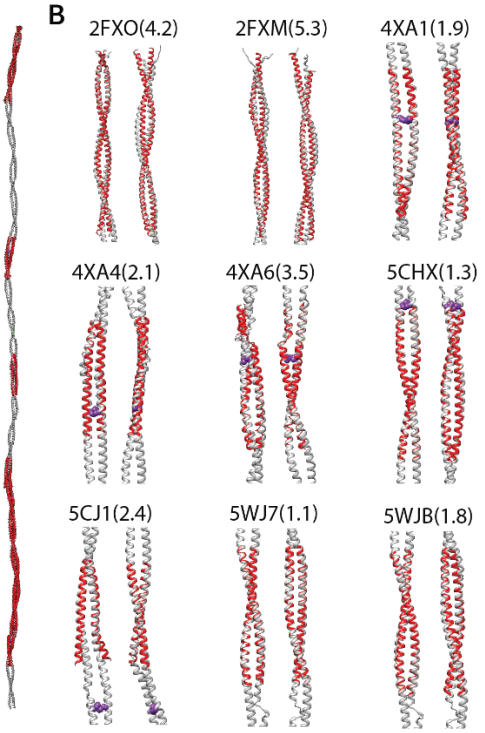
Fig. S4. Radial and azimuthal changes in the myosin tail coiled coil. (A) Plot of the radial position of the myosin tail from N- to C-terminus. Above the graph is an image of the myosin tail scaled to the same length and viewed perpendicular to the radial vector approximately in the middle of the graph. The plot shows the distance from the center of the filament (i.e. half of the distance from the myosin tail on the opposite side). As expected, the distance decreases from ~120Å at the head-tail junction to ~50Å at the C-terminus. However, the distance does not decrease steadily, but shows sharp increases after Skips 2 and 4. The coiled coil reaches the filament core twice: once at Skip 4 and again at the end of the coiled coil. (B) The azimuthal angle of the myosin showing the tangential movements of the coiled coil. Above the graph is an image of the myosin tail scaled to the same length and viewed along a radial vector from roughly the middle of the graph. Sharp changes in direction occur at eight locations; the bend of the proximal S2, before and after Skip 1, after Skip 2, before the Skip 3 Accommodation Region, midway between Skips 3 and 4 and after Skip 4.

Expanded comparison between the atomic model and the crystal structure

Myosin coiled coil atomic model from Lethocerus was compared to segments of this coiled coil from human cardiac muscle myosin that were crystallized and solved at high resolution (3-5). We used the MMTSB package (6) to fit the crystal structures and our atomic model after shifting the residue numbers and selecting the Cα atoms. As expected, the proximal S2 (PDBs 2FXO and 2FXM) had a very high RMSD due to the difference in structure caused by the presence and arrangement of myosin heads in the Lethocerus thick filament as previously observed (7). The comparison between the crystal structures and our atomic model in the Skip regions is discussed in the main text. Here we show all 13 relevant crystal structures and their RMSDs after fitting to the thick filament atomic model.

A

B



C

PDB Name	First Residue	Last Residue	Ca RMSD (Å)	Zone
2FXM	838	963	5.3	S2
2FXO	838	961	4.2	S2
4XA1	1173	1238	1.9	Skip 1
4XA1	1173	1238	1.8	Skip 1
4XA3	1361	1425	3.7	Skip 2
5CJ1	1526	1571	2.4	
4XA4	1551	1609	2.1	Skip 3
5CJ4	1562	1622	2.3	Skip 3
5CJ4	1562	1622	2.3	Skip 3
5CHX	1590	1657	1.3	Skip 3
5CJ0	1631	1692	1.7	
5WLZ	1677	1758	1.9	
5WLZ	1677	1758	2.0	
5WJ7	1733	1797	1.1	
5WJB	1733	1797	1.8	
5WJB	1733	1797	1.8	
4XA6	1777	1855	3.5	Skip 4

Fig. S5. Comparison of all crystal structures of myosin tail segments with the *Lethocerus* atomic model. (A) An overall view of crystal structures fitted on our atomic model. (B) Closer views of fits before and after 90º rotation. (C) Information of all the crystal structures including their residue range (in cardiac sequence) and the RMSD after fitting them on the *Lethocerus* atomic model.

Symmetry of the two heavy chains

Myosin has two identical heavy chains that differ in structure in situ because their interactions with nearest neighbor myosin tails within the filament backbone involve axial shifts of multiple crowns (several hundred residues). When the two chains are fitted based on their C_α atoms, the backbones have a very small RMSD (this is done with a moving window of 7 and 14 residues with the same results) but the side chains do not fit beyond 2Å for the most part. The proximal S2 in the filament shows more asymmetry than observed in the crystal structure. In general the side-chains of the crystal structure are more symmetrically arranged than our atomic model reflecting the significant effects of their environment in crystals and filaments. An ideal coiled coil would have a perfect symmetry in the backbone.

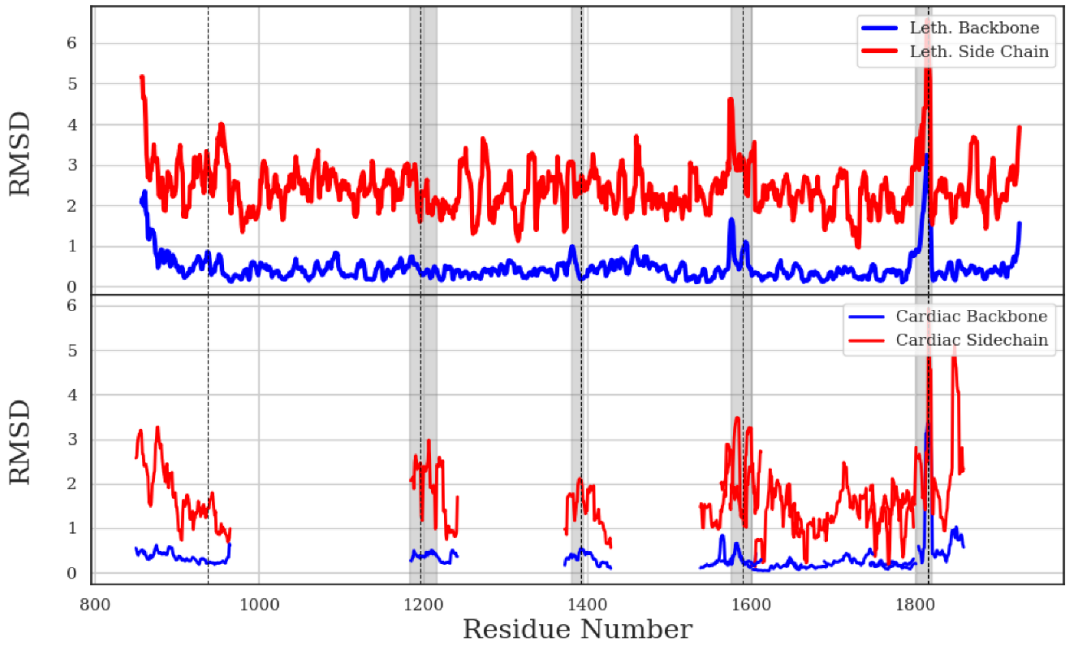


Fig. S6. Coiled coil asymmetry measured by the RMSD between the two heavy chains of myosin when aligned on the backbone atoms. When the coiled coil is examined one heptad at a time the RMSD between the backbone atoms is very low. Side chains show higher levels of asymmetry since they interact with other myosin tails. Large differences can be detected at the proximal S2 and Skips 3 and 4.

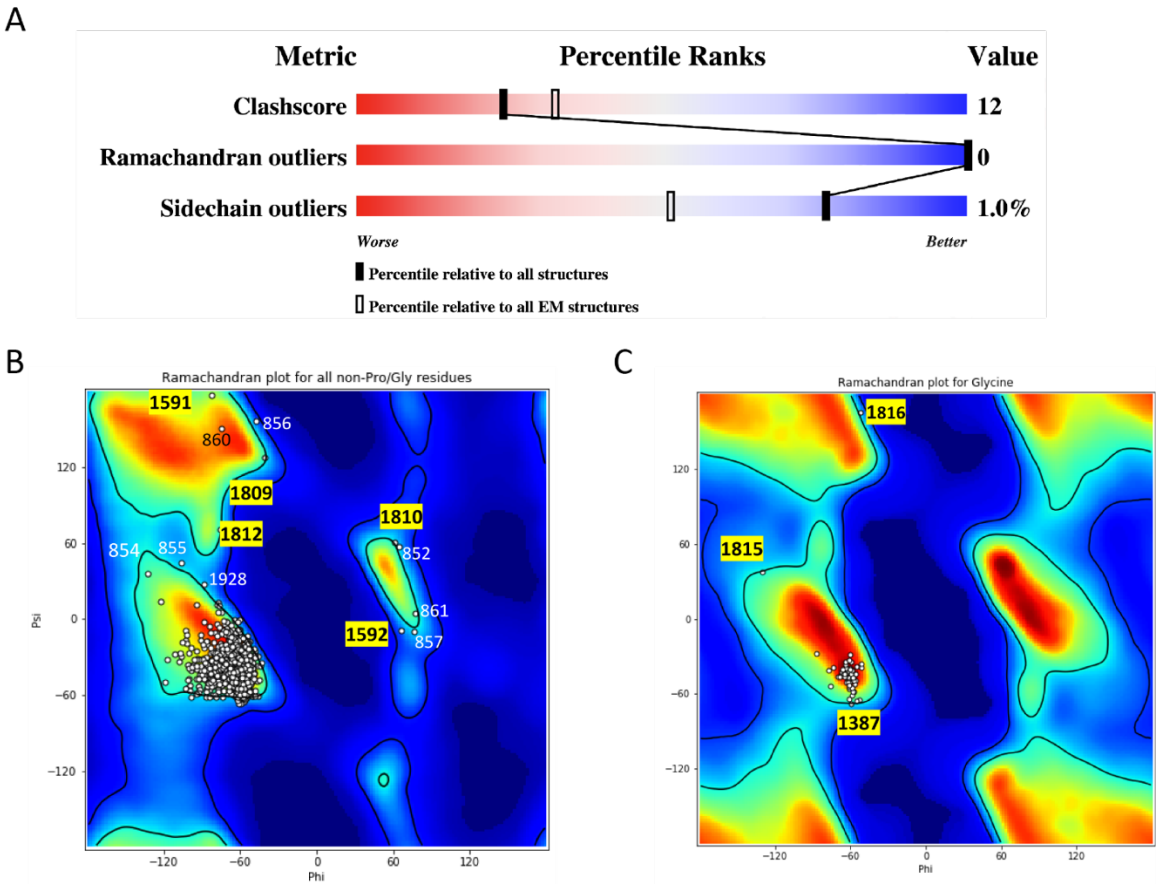


Fig. S7. Atomic model validation. (A) Validation scores are given and compared to all structures and EM structures in the PDB database. (B) Ramachandran plots for non-glycine residues shows all the angles are concentrated in the alpha-helical region with some exceptions: the N-terminal part of the S2 (residues 852-861) where the free-head α-helix is unwound, the uncoiled region of the Skip 4 (residues 1815-1816), and two residues close to Skip 3 where the coiled coil is bent (residues 1591-1592). Right: Ramachandran plot for glycine residues. Two glycine residues of the KGGKK motif, G1815-G1816, and G1387 in the Skip 2 region which fall just within the allowed α-helical region.

Molecular Dynamics simulations

The four isolated skip regions, which were obtained from the atomic reconstruction for the coiled coil, were solvated in TIP3P water with 150 mM NaCl. The MD simulations for the four systems were performed with the GPU-accelerated version of Amber18 using the ff14SB force field (8, 9). Firstly, energy minimization was carried out for the systems while keeping the protein positions fixed with 1 kcal/(mol Å2) harmonic constraints. Applying the same constraints, a 1 ns simulation in the isothermal-isobaric ensemble (NPT) was performed at 1 bar and 300K. The subsequent production simulation without any constraints for each Skip region was performed for 100 ns in the NPT ensemble. Langevin dynamics with a friction coefficient of 1 ps-1 was applied to maintain the temperature and MC barostat (10) was coupled to the system for pressure control. Bonds involving hydrogen atoms were constrained with the SHAKE algorithm (11). Particle Mesh Ewald (PME) (12) was used for full-system periodic electrostatics and a 9 Å cutoff was applied to Lennard-Jones interactions. The MD movies for each Skip region were rendered with VMD (13).

Comparison of coiled coil parameters at the skip residues

The MD simulation shows that the RMSD between filament atomic model and the crystal structures decreases over time. However, a closer look at coiled coil parameters reveals more about the Skip accommodation. At Skip 2 coiled coil parameters (ω_0 and α) are closer to zero in both crystal structure and the MD results than in the filament atomic model. Skips 3 and 1 had similar parameter values already and Skip 4 the Accommodation Index becomes un-interpretable as a coiled coil because of the double-hinge even though these hinges are not as large as the crystal structure.

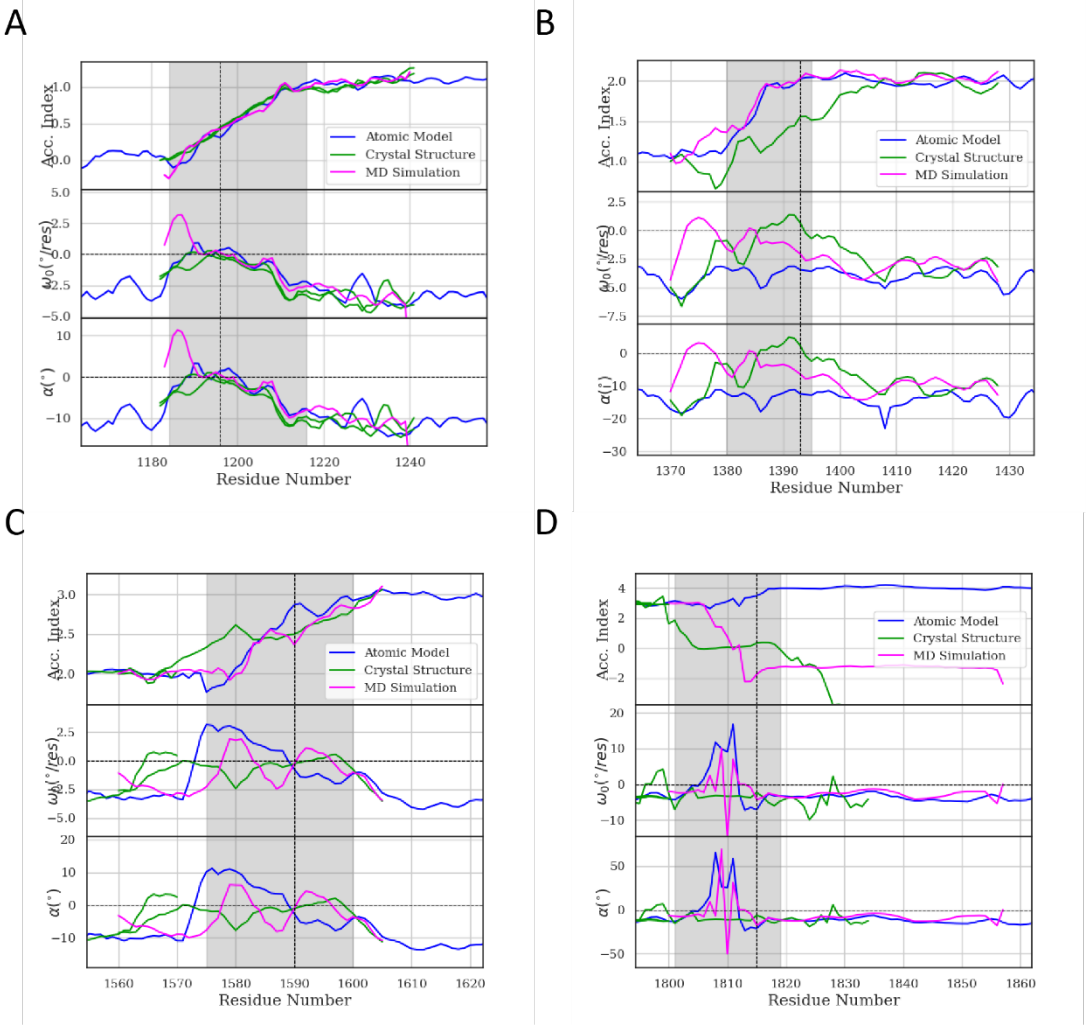


Fig. S8. Important geometrical parameters for Skips 1-4, (A-D) respectively. These parameters explain how the filament atomic model agrees with the crystal structure in Skip 1 and not in other Skips. ω0 and α are both going to zero in Skips 1-3 in crystal structure. In the filament atomic model at Skip 2, both parameters remain unchanged while at Skips 3 and 4 both parameters become positive, meaning that the coiled coil becomes right handed for few residues. In both places this seems to be cause by proteins present in the thick filament that are not part of the coiled coil.

C-terminal non-helical extension

The same sequences that were examined for conservation of the KGGKK motif were also examined for the length of their C-terminus extension. We assume that the C-terminal extension starts after T1930 based on our structure even though PairCoil predicted that the coiled-coil could go further (S1937). This discrepancy cannot be attributed to an error in our structure since our resolution is high in that region and the resolved lower resolution extension ends at the C-terminus (M1971). The length of C-terminal extension is on average 38 residues for invertebrates and 13 residues for vertebrates (See Fig. S3). We also observed that almost all vertebrate C-termini end with one or more glutamic acid residues.

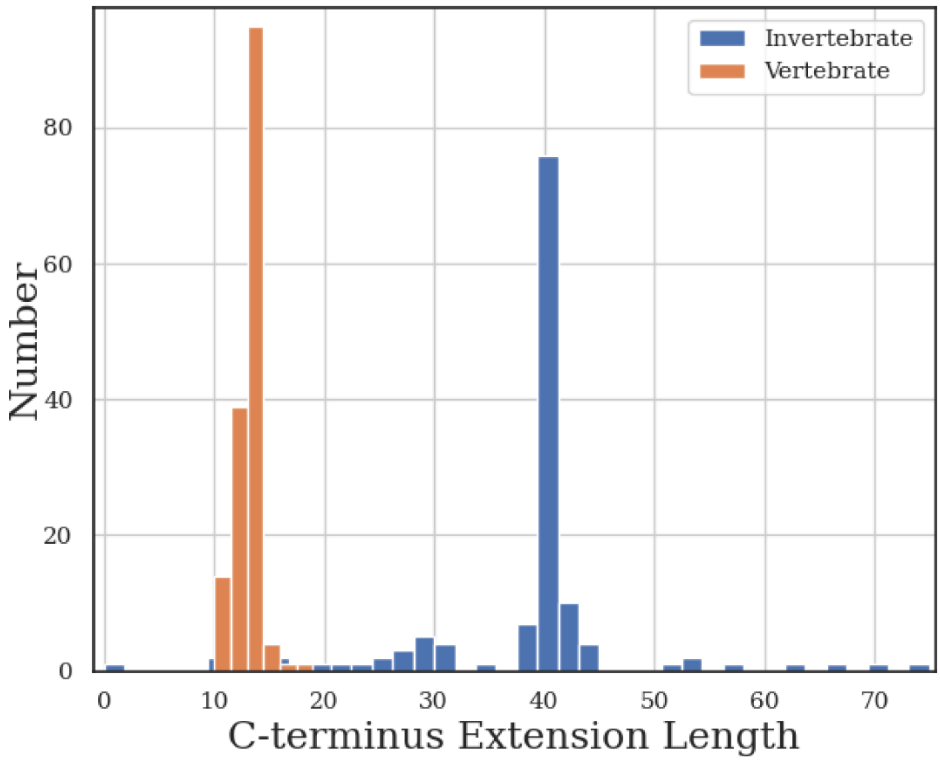


Fig. S9. Histogram showing that C-terminal extension for vertebrates is shorter than invertebrates by almost 20 residues. The sequences are from the NCBI server (one per species).

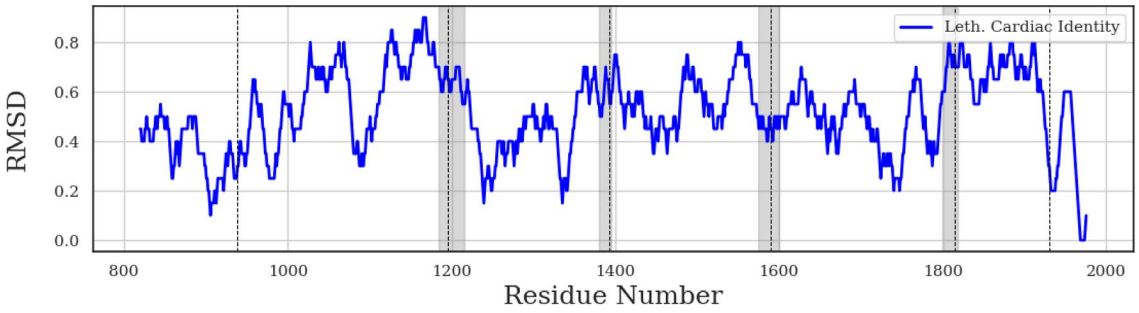


Fig. S10. Conservation (identity) between Lethocerus and human cardiac myosin sequence averaged over a window of 20 residues. Vertical dotted lines mark residues 938, the point of entry of the proximal S2 into the thick filament backbone, the four skip residues and the end of the coiled coil at residue 1930. Average identity is 52%.

Conservation of the striated muscle KGGKK motif

A subsequent search of the NCBI's protein database for striated muscle myosin heavy chains resulted in 3675 invertebrate and 590 vertebrate sequences. The biopython package was used to select one sequence per species (14) with the longest chain length. All of the sequences were then aligned to the *Lethocerus* sequence and the Skip 4 KGGKK (starting with K1814) motif was examined. Among vertebrates every sequence with a gap or different amino acids in this region turned out to be a predicted sequence. Of the 154 remaining sequences, all had the KGGKK motif conserved (2 with the K > R conservative substitution). Among the 134 invertebrate sequences remaining, the KGGKK motif was ubiquitous with one exception (AXA20421.1). AXA20421.1 from *Ptychodera flava* (a genus of acorn worm) shows very unusual characteristics including a >100 residue insertion before Skip 1. Note that the sequences having the KGGKK domain were not examined one by one so there could be more predicted sequences in the final list.

Sample Preparation

Lethocerus indicus dorsal longitudinal indirect flight muscle previously glycerinated for long term storage (15) was the starting material. Myofibrils for the filament preparation were washed free of glycerol and treated with calpain to dissolve the Z-disk followed by gelsolin to dissolve the thin filaments (7). Thick filaments were separated by drawing the solution 10X through a 1 ml syringe with 26G needle. Quantifoil carbon grids were frozen manually using the back-blotting technique (16) followed by rapid immersion into liquid ethane.

Data Collection/Preprocessing

We improved on the previously reported backbone resolution of 5.5Å by collecting a larger data set on a new specimen grid using the FSU Titan Krios at equipped with a DE64. The 8,192 x 8,192 frame size recorded longer lengths of individual filaments with a frame size 3.4x larger than the previous dataset (7). The collection was done in integration mode using leginon/appion pipeline software (17), with the dose of 60 $e^-/Å^2$. The sample was diluted to reduce the number of crossing filaments and this caused a large number of micrographs to end up empty. Motioncor2 (18) was used for frame alignment and dose weighting with GCTF (19) utilized for global CTF estimation (Fig. S11).

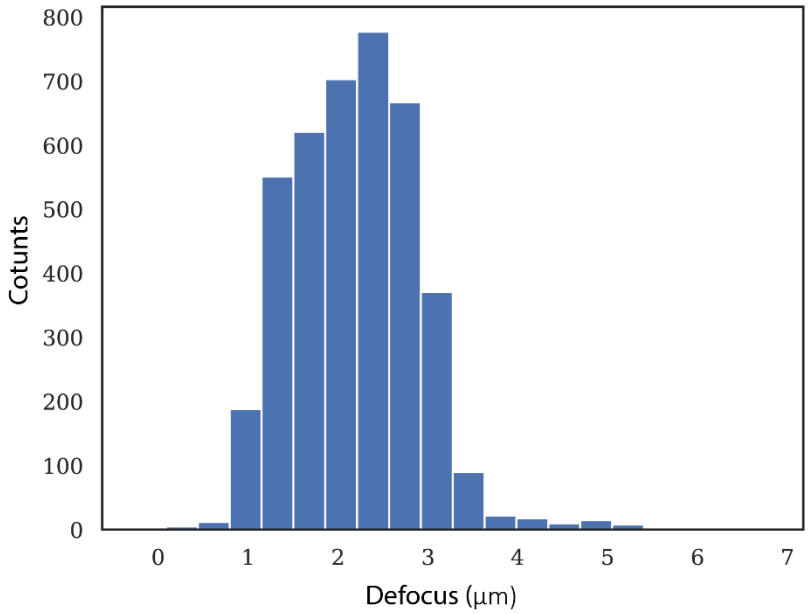


Fig. S11. Distribution of defocus estimation of for all micrographs calculated using GCTF. Local values were also calculated after manual picking determined particle positions.

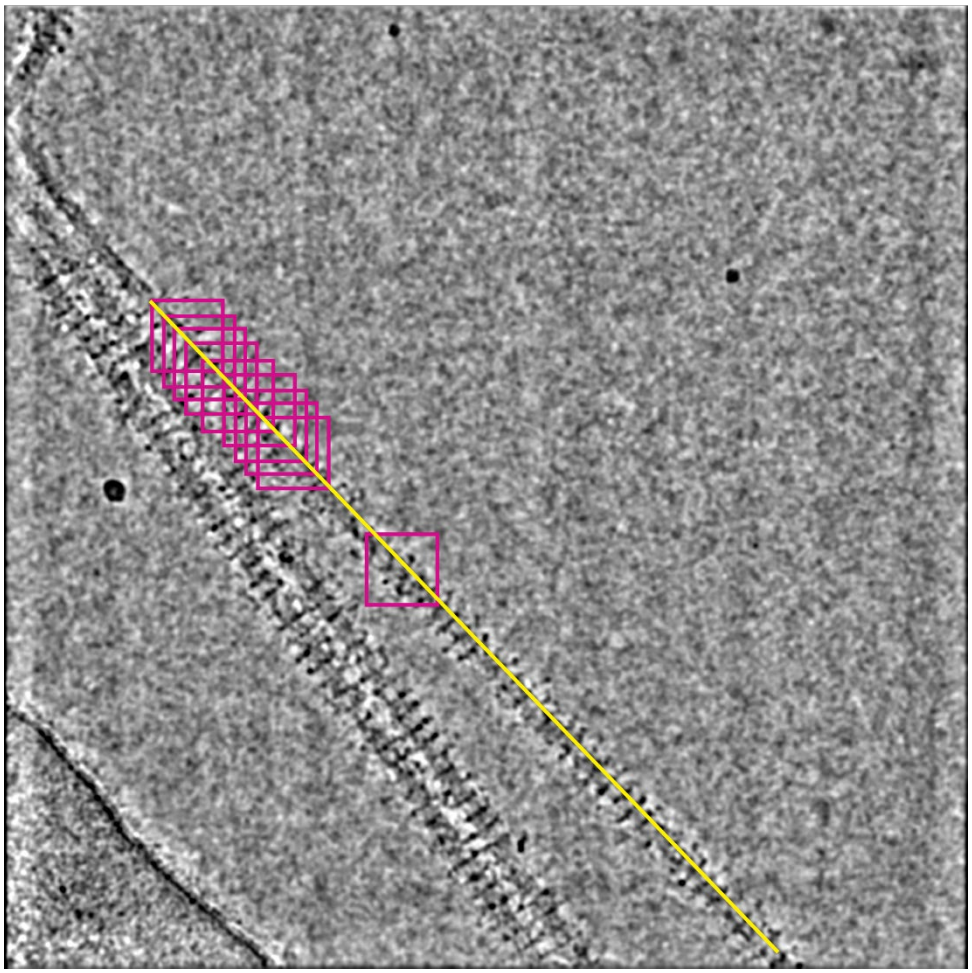


Fig. S12. An example micrograph is shown with manual picking and extraction: a line defined with 2 points starting at the M-band towards Z-disk and overlapping boxes are extracted from the filament. Not all the M-bands were visible and the manual picking was from M-to-Z when possible.

Classification and Refinement

Unfortunately, no automatic particle picking software tried worked reliably in picking our filaments. Consequently, filament segments were picked manually and then extracted in relion (20). An example micrograph for manual picking is shown (Fig. S12). Two points are selected to define the filament axis starting at the M-band and ending towards Z-disk (when either the M-band or the tapered end were visible). These two points also define the filament polarity. Overlapping segments are later extracted along these filaments. From 2993 collected micrographs 7650 filaments were picked which were cut into 216,045 helical segments displaced by one crown (145 Å axial repeat) unit per segment. The helical segments were first extracted in 145 Å distances but later re-extracted with 150 Å distances to account for magnification error and to preserve the prior information of the segments for later calculations. Initially the box size was 512 pixels because of the GPU memory limit and this caused the spherical mask of the reconstruction to clash with myosin heads and limit the resolution. The box size of 768 was then used and the CPU-based program cisTEM produced a high-resolution reconstruction (21). cisTEM was also used to perform 2D classification (Fig. S13). Those classes with pronounced head appearance were selected for the final stack. The final refinement was done on 173,515 segments and yielded a 4.25 Å structure. The structure was reconstructed under C4 rotational symmetry since cisTEM does not perform iterative helical reconstruction. It is worth noting that there is an implicit helical symmetry imposed because of overlapping segments.

Afterwards, Relion3 was used on larger box size but the resolution did not improve whether the helical symmetry was imposed or only the rotational symmetry was used.

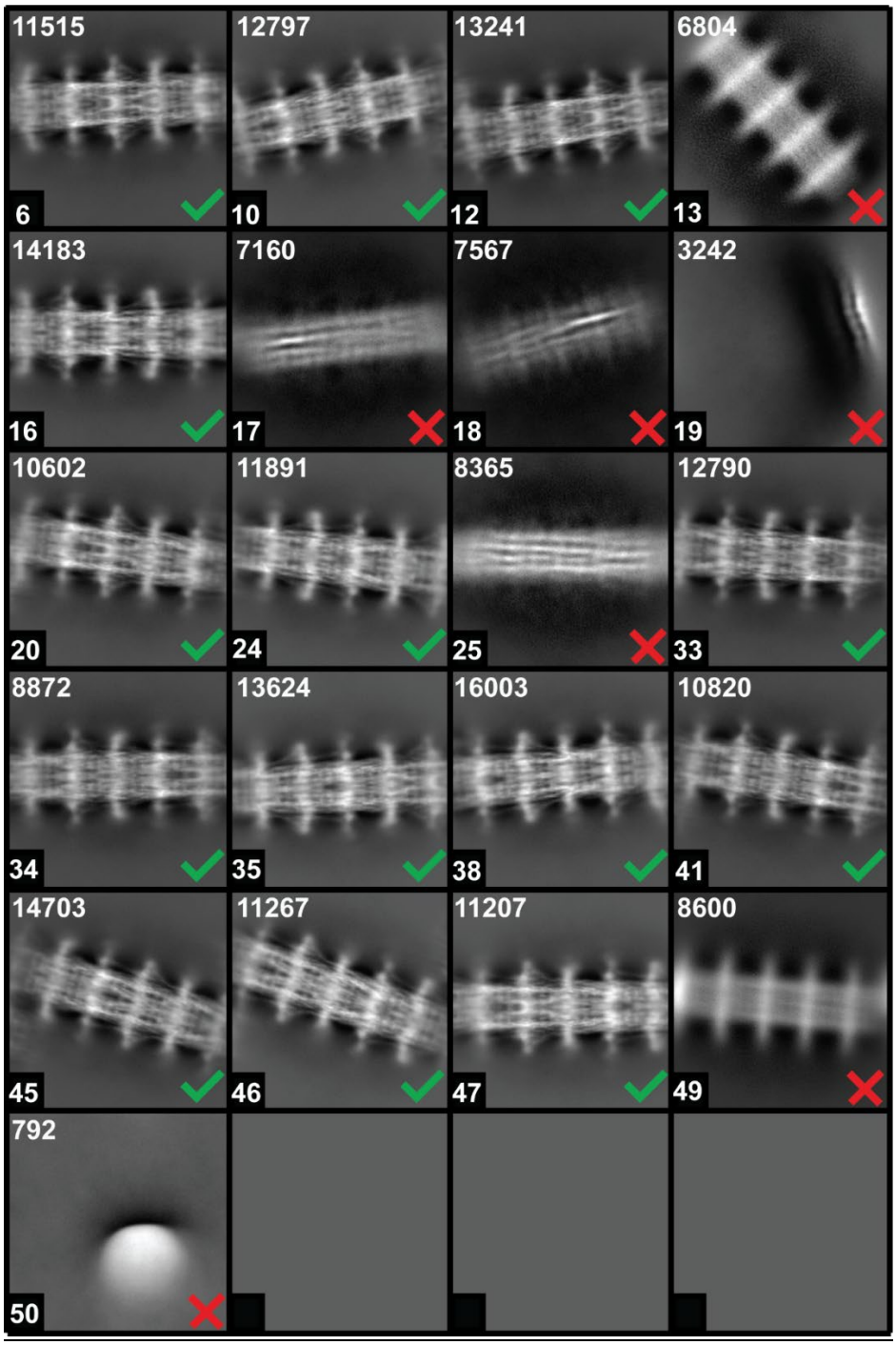


Fig. S13. 2D Classification results in cisTEM. All non-empty classes are shown with their populations and the rejected classes are denoted with red X marks.

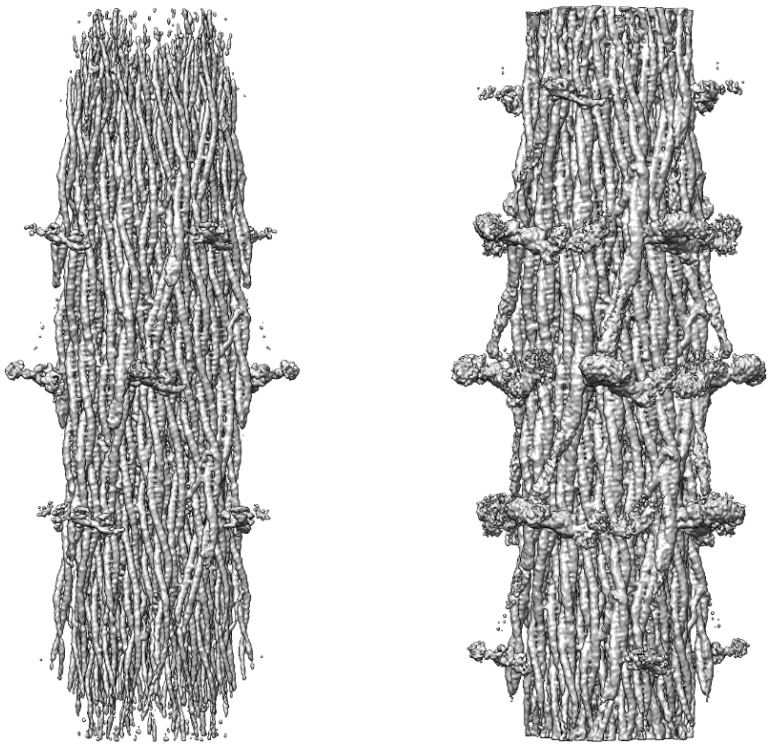


Fig. S14. cisTEM reconstruction before sharpening. We can see the quality of the map is lower on the top and bottom where the map goes beyond the background mask. The left picture has a higher threshold.

Resolution and Sharpening

The resolution reported by cisTEM is 4.26 Å which is re-scaled from the unmasked FSC resolution based on the molecular weight of the sample. To validate this resolution, we generated the half-maps and calculated the FSC with and without soft edge masks. As expected, the resolution changes if we include the poorly resolved myosin heads in the mask and a tighter mask gives us an FSC of 4.1 Å (Fig. S15A). The peak in the FSC plot at ~5 Å resolution is due to the high α-helix content of the structure.

Local resolution estimation showed that except for the S2 segment close to the myosin heads on the surface the resolution was uniform (Fig. S15B). In the backbone the resolution was in the range of 4.2-8 Å, the highest inside the backbone and the lowest being close the S2 region. Applying only C4 symmetry the resolution was not uniform along the helical axis but improved after applying helical symmetry. Running monores (22) on the map after imposing helical symmetry caused over-estimation of the resolution. Myosin heads were not resolved beyond 20 Å because they are poorly ordered compared to the backbone. Sharpening was done using local-Deblur (23) and embfactor (24) and both gave us similar results, but local-Deblur was more effective in resolving the large side chains (Fig. S15E-G).

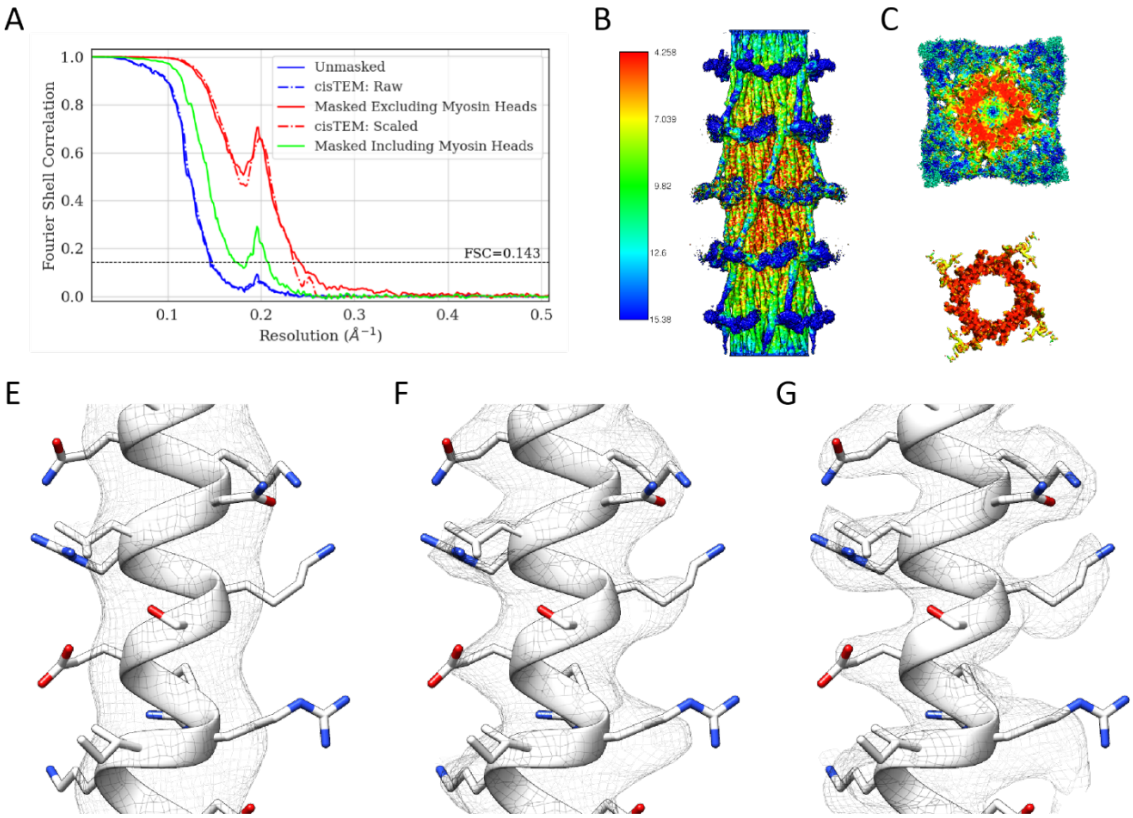


Fig. S15. Resolution determination. (A) FSC curves from cisTEM compared to FSC curves produced in relion using the half maps. The un-masked cisTEM FSC (blue dashed line) compares very well to the un-masked relion FSC (blue solid line) and the same is true for the cisTEM scaled FSC (red dashed line) and relion FSC (red solid line) when masking out the myosin heads. The FSC peaks at around 5.1Å is due to the pronounced presence of the coiled-coil in our reconstruction. As expected, leaving the heads inside the mask will result a lower FSC resolution (solid green line). (B) Local monores resolution calculation shows that the backbone coiled-coils are uniformly resolved to the highest resolution. The lower resolution parts on the top and bottom improve after imposing helical symmetry. (C) Cross-sections of the reconstruction shows the resolution in the backbone is uniformly 4.25Å. Side chain quality for residues 1845-1854 is compared in a non-sharpened map (E), after global B-factor sharpening (F), and after local deblur sharpening (G). Well-resolved side chains are seen when the map is sharpened using local deblur.

Helical parameters and extension

Helical parameters calculated using relion_helix_toolbox (25) yielded a helical rise of 149.56 Å and helical twist of 34.08°. The pixel size was then reset from 1.009 to 0.978 to scale the helical rise to the known value of 145Å from X-ray studies (26). Using relion_helix_toolbox, we expanded the structure to a large (2048^3) box sufficient to contain a full myosin molecule, which was subsequently segmented out by the Segger algorithm in chimera (27). The complete myosin tail ~1600Å long was used to build the atomic model in a single piece.

Building and validating the Atomic Model

The complete myosin volume was imported into COOT (28) in which the atomic model of the ~1,070 residue coiled coil was built manually. The Cα positions were located on the skeletonized map in the baton-mode and then converted to a poly-alanine with α-helical constraints (except on the Skip-4 loop segment). We then used the "mutate" tool to replace the alanines with the residues from the sequence, starting from the C-terminus because the resolution is higher than at the N-terminus (proximal S2). The Lethocerus indicus flight muscle myosin sequence (29) was obtained from Genbank accession number MF07800. We used the protein sequence isoform Mhc_X1 (ASP18627.1), which is 1971 residues long and agrees with our structure better in the alternatively spliced region (29). Isoforms X1, X4, and X5 are identical in the tail region, except for a single splice site. The alternatively spliced region is resolved and the atomic model has sufficient resolution and definition to differentiate between isoforms (Fig. S16). The larger and positive side chains like tryptophan, histidine, arginine, and lysine were better resolved and acted as guides to make sure the atomic model did not deviate from the volume. We used Phenix to solve clashes between the two chains. A few side chains were not fitted well in the density. The atomic model was fixed in COOT to make the side chains fit better and the final model was refined for the last time in Phenix (30). Ramachandran plots show that the structure is almost entirely α-helical except the KGGKK region in Skip 4, the head-tail junction and the non-helical C-terminus. At residue 1592 one of the chains is pulled far enough to be out of the allowed α-helical region.

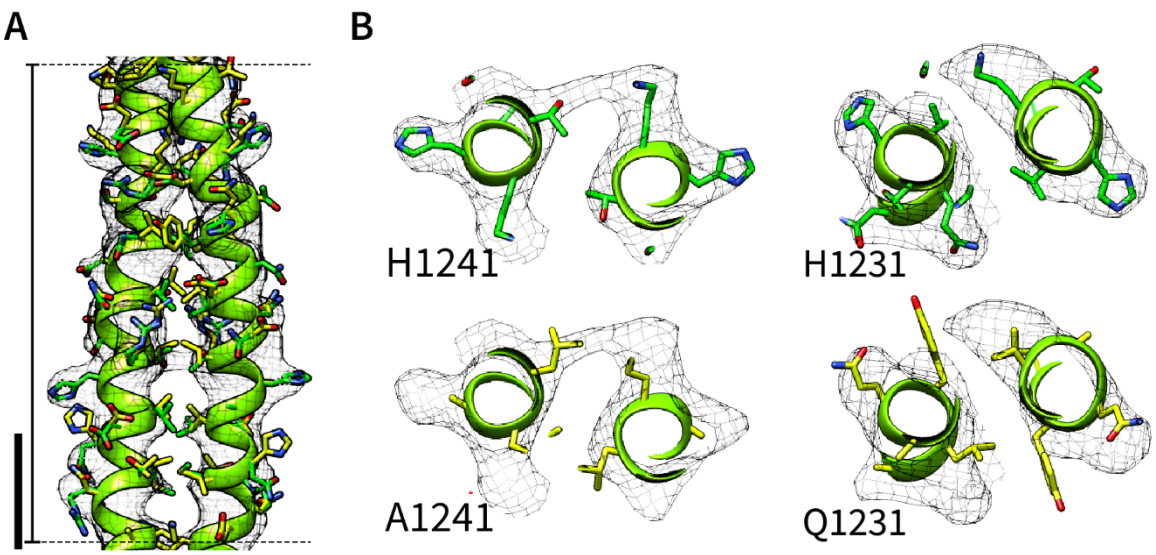


Fig. S16. (A) The alternatively spliced region in the myosin tail is shown along with the atomic model for isoform 1. (B) Large amino acid side chains are well resolved in the reconstruction. H1241 in isoform 1 is clearly resolved in the side chain density whereas A1241 present in isoform 2 is too small to account for the density. Similarly, for H1231 of isoform 1 vs. Q1231 of isoform 2.

Coiled-coil Analysis

As discussed in multiple figures throughout the manuscript, Crick's Coiled-Coil Parameterization (CCCP) server provided quantitative analysis of the entire myosin tail structure as well as quantitative comparison with corresponding crystal structures (20). The most significant calculation was the Accommodation Index , which provides a validation for our atomic model. The Accommodation Index provides a quantitative comparison between filament and crystal structures of the human cardiac myosin coiled-coil. Accommodation Index analysis is important because the filament resolution is insufficient to resolve all the side chains. The greatest challenge is to insure local errors of the structure do not propagate over longer ranges. Any such occurrence would have appeared as an unexpected change in Accommodation Index. Accommodation Index increased from zero to 4 along the coiled coil, which means an insignificant range of residues are off-set in their α-helical angle. Accommodation Index also provides a quantitative evaluation of the proximal S2 and its modified structure at the head-tail junction. The CCCP server also determined the *abcdefg* positions of the Cα atoms in the atomic model and compared it to predictions made by PairCoil.

Fiber Diffraction

X-ray fiber diffraction of relaxed IFM was done using the BioCAT beamline 18ID (31) at the Advanced Photon Source, Argonne National Lab with an X-ray energy of 12.0 keV and a specimen to detector distance of 0.2 m. Images were collected on a custom CCD detector with 48 μm pixels (32).

Movie S1-4 (separate files). Morphing movies from a canonical Crick coiled coil into the actual structure of the *Lethocerus* tail found in the thick filament backbone.

Movie S1, file Skip1.mov. Morphing movie of the Skip 1 region showing the change from canonical coiled to the *Lethocerus* tail structure which has the Skip 1 region with the pair of parallel helices typical of accommodation of a skip residue insertion. One chain is colored gray, the other white. The accommodation regions are gold, and the skip residue is purple. The structure is anchored at the top and swings azimuthally as it conforms to the rotated position of Skip 1 on the filament backbone.

Movie S2, file Skip2.mov. Morphing movie of the Skip 2 region showing the change from canonical coiled to the *Lethocerus* tail structure. One chain is colored gray, the other white. The accommodation regions are gold, and the skip residue is purple. The structure is anchored at the top. Only torsional movements of the separate helices are seen because the Skip 2 structure in the filament has normal coiled-coil pitch.

Movie S3, file Skip3.mov. Morphing movie of the Skip 3 region showing the change from canonical coiled to the *Lethocerus* tail structure. One chain is colored gray, the other white. The accommodation regions are gold, and the skip residue is purple. The structure is anchored at the bottom and swings azimuthally as it conforms to the Skip 3 orientation in the filament backbone.

Movie S4, file Skip4.mov). Morphing movie of the Skip 1 region showing the change from canonical coiled to the *Lethocerus* tail structure which has the Skip 1 region with the pair of parallel helices typical of accommodation of a skip residue insertion. One chain is colored gray, the other white. The accommodation regions are gold, and the skip residue is purple. The structure is anchored at the bottom and shows one of the helices unfolding and the other straightening as the canonical structure conforms to the Skip 4 structure in the filament backbone.

Movie S5-8 (separate files). Molecular dynamics trajectories of the skip regions

Movie S5, file Skip1-MD.mov. Molecular dynamics trajectories (orange) for the Skip 1 region displayed on the corresponding X-ray crystal structure (gray) showing minimal change during the time period of the calculation.

Movie S6, file Skip2-MD.mov. Molecular dynamics trajectories (orange) for the Skip 2 region displayed on the corresponding X-ray crystal structure (gray) showing the change from a predicted coiled coil pitch toward two parallel helices typical of a skip residue insertion during the time period of the calculation.

Movie S7, file Skip3-MD.mov. Molecular dynamics trajectories (orange) for the Skip 3 region displayed on the corresponding X-ray crystal structure (gray) showing minimal change during the time period calculation.

Movie S8, file Skip4-MD.mov. Molecular dynamics trajectories (orange) for the Skip 4 region displayed on the corresponding X-ray crystal structure (gray) showing unfolding of the free head tail helix during the time period calculation. The blocked head Skip 4 started unfolded and ended unfolded

SI References

1. M. A. Larkin *et al.*, Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).
2. A. V. McDonnell, T. Jiang, A. E. Keating, B. Berger, Paircoil2: improved prediction of coiled coils from sequence. *Bioinformatics* **22**, 356-358 (2006).
3. K. C. Taylor *et al.*, Skip residues modulate the structural properties of the myosin rod and guide thick filament assembly. *Proc Natl Acad Sci U S A* **112**, E3806-3815 (2015).
4. E. N. Korkmaz *et al.*, A composite approach towards a complete model of the myosin rod. *Proteins* **84**, 172-189 (2016).
5. M. P. Andreas, G. Ajay, J. A. Gellings, I. Rayment, Design considerations in coiled-coil fusion constructs for the structural determination of a problematic region of the human cardiac myosin rod. *J Struct Biol* **200**, 219-228 (2017).
6. M. Feig, J. Karanicolas, C. L. Brooks, 3rd, MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. *J Mol Graph Model* **22**, 377-395 (2004).
7. Z. Hu, D. W. Taylor, M. K. Reedy, R. J. Edwards, K. A. Taylor, Structure of myosin filaments from relaxed Lethocerus flight muscle by cryo-EM at 6 Å resolution. *Sci Adv* **2**, e1600058 (2016).
8. R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, R. C. Walker, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *Journal of chemical theory and computation* **9**, 3878-3888 (2013).
9. J. A. Maier *et al.*, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of chemical theory and computation* **11**, 3696-3713 (2015).
10. J. Åqvist, P. Wennerström, M. Nervall, S. Bjelic, B. O. Brandsdal, Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. *Chemical physics letters* **384**, 288-294 (2004).
11. S. Miyamoto, P. A. Kollman, Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comput Chem* **13**, 952-962 (1992).
12. T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *The Journal of chemical physics* **98**, 10089-10092 (1993).
13. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996).
14. P. J. A. Cock *et al.*, Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422-1423 (2009).
15. K. A. Taylor, S. Wu, M. C. Reedy, M. K. Reedy, Imaging actomyosin in situ. *Methods Cell Biol* **79**, 321-368 (2007).
16. C. Toyoshima, On the use of holey grids in electron crystallography. *Ultramicroscopy* **30**, 439-444 (1989).
17. C. Suloway *et al.*, Automated molecular microscopy: the new Leginon system. *J Struct Biol* **151**, 41-60 (2005).
18. S. Q. Zheng *et al.*, MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).
19. K. Zhang, Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1-12 (2016).
20. S. H. Scheres, A Bayesian view on cryo-EM structure determination. *Journal of molecular biology* **415**, 406-418 (2012).
21. T. Grant, A. Rohou, N. Grigorieff, cisTEM, user-friendly software for single-particle image processing. *Elife* **7** (2018).
22. J. L. Vilas *et al.*, MonoRes: Automatic and Accurate Estimation of Local Resolution for Electron Microscopy Maps. *Structure* **26**, 337-344 e334 (2018).
23. E. Ramírez-Aportela *et al.*, Automatic local resolution-based sharpening of cryo-EM maps. *Bioinformatics* **36**, 765-772 (2020).
24. J. J. Fernandez, D. Luque, J. R. Caston, J. L. Carrascosa, Sharpening high resolution information in single particle electron cryomicroscopy. *J. Struct. Biol.* **164**, 170-175 (2008).

25. S. He, S. H. Scheres, Helical reconstruction in RELION. *J. Struct. Biol.* **198**, 163-176 (2017).
26. R. J. Perz-Edwards *et al.*, X-ray diffraction evidence for myosin-troponin connections and tropomyosin movement during stretch activation of insect flight muscle. *Proc Natl Acad Sci U S A* **108**, 120-125 (2011).
27. G. D. Pintilie, J. Zhang, T. D. Goddard, W. Chiu, D. C. Gossard, Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. *J. Struct. Biol.* **170**, 427-438 (2010).
28. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
29. L. Fee, W. Lin, F. Qiu, R. J. Edwards, Myosin II sequences for Lethocerus indicus. *J Muscle Res Cell Motil* **38**, 193-200 (2017).
30. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66**, 213-221 (2010).
31. R. Fischetti *et al.*, The BioCAT undulator beamline 18ID: a facility for biological non-crystalline diffraction and X-ray absorption spectroscopy at the Advanced Photon Source. *Journal of synchrotron radiation* **11**, 399-405 (2004).
32. W. C. Phillips, A. Stewart, M. Stanton, I. Naday, C. Ingersoll, High-sensitivity CCD-based X-ray detector. *Journal of synchrotron radiation* **9**, 36-43 (2002).