Structural and mechanistic insights into the function of the unconventional class-XIV myosin, MyoA, from *Toxoplasma gondii*

Cameron J. Powell[a], Raghavendran Ramaswamy[a], Anne Kelsen[b], David J. Hamelin[a], David M. Warshaw[c], Jürgen Bosch[d,e], John E. Burke[a], Gary E. Ward[b], and Martin J. Boulanger*[a]

[a]Department of Biochemistry and Microbiology, University of Victoria, Victoria, BC, V8P 5C2, Canada

[b]Department of Microbiology and Molecular Genetics and [c]Department of Molecular Physiology and Biophysics, University of Vermont Larner College of Medicine, Burlington, VT 05405, USA

[d]Division of Pulmonology and Allergy/Immunology, Department of Pediatrics, Case Western Reserve University School of Medicine, Cleveland, OH

[e]InterRayBio, LLC, Baltimore MD

SI Methods and Materials

Cloning, protein production and purification. Twin-strep-II (TST)-tagged TgMyoA (1-831) (ToxoDB ID TGGT1_070410/GenBank™ accession number EPR60343.1) wild-type and phosphomimetic mutants were co-expressed with 6xhis-tagged TgELC1 (ToxoDB TGME49_269442, GenBank™ XP_002365635.1) and TgMLC1ΔN66 (ToxoDB TGME49_257680, GenBank™ KYK62927.1) in Hi5 cells. Protein was first purified by nickel-

affinity, followed by streptactin affinity, and finally SEC, using a final buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM $CaCl_2$, and 1 mM DTT. Purified, concentrated complex was used for motility assays, HDX-MS, and DSF. TgMyoA (1-778) was expressed and purified as described previously (1) into a final buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl and 1 mM DTT. Purified, concentrated TgMyoA (1-778) was diluted with buffer containing $MgCl_2$, ADP, NaF, and $Al(NO_3)_3$, to final concentrations of 2 mM $MgCl_2$, 2 mM ADP, 6 mM NaF, 1.75 mM $Al(NO_3)_3$.

Actin co-sedimentation assay. Purified TgMyoA was dialyzed over-night into "buffer A" containing: 150 mM KCl; 20 mM HEPES, pH 7.5; 30 mM imidazole; 1 mM $MgCl_2$; 0.2 mM ADP; 5 mM BME. TgMyoA was then diluted to a concentration of 2.8 μM and combined with 6.0 μM rabbit skeletal muscle actin and spun down at 340,000 x *g*; supernatant containing globular actin and non-actin-binding myosin was removed. The pellet was then resuspended in "buffer B", which is identical to "buffer A" with the exception of ADP being replaced with 4 mM ATP, to elute functional myosin from actin filaments. The protein was again spun down at 400,000 x *g* and the supernatant containing functional myosin was removed. The pellet was resuspended in "buffer A" that now included only non-functional myosin, and left-over actin.

Crystallization and data collection. Crystals of TgMyoA (1-778) were grown at 4 °C by mixing protein in a 1:1 ratio with reservoir solution containing 22% PEG 8000, 200 mM $NH_4(SO_4)$, and 100 mM MES, pH 6.5. Crystals were cryoprotected in reservoir solution supplemented with 15% glycerol and flash cooled in liquid nitrogen. Diffraction data was collected on beamline 14-1 at the Stanford Synchrotron Radiation Lightsource (SSRL).

Data processing, structure solution and refinement. Diffraction data for TgMyoA (1-778) were processed to 2.6 Å resolution using Imosflm (2) and Aimless (3) and initial phases were

obtained by molecular replacement in Phaser (4). The search model consisted of the motor domain of human myosin 1c (PDB ID 4BYF, chain A) with the N-terminus and converter subdomain removed and loops trimmed using CHAINSAW (5). Iterative rounds of refinement first allowed for building of the converter subdomain and several surface loops, after which sufficient electron density emerged to enable tracing of the N-terminal region using *D. discoideum* myosin II (PDB ID 1FMW) as the model. Despite relatively higher B-factors in this region, electron density maps were sufficient for building main chain loops and secondary structure elements of the SH3 sub-domain. The electron-dense side chains of residues W39 and W90, which flank the SH3-domain at its N- and C-termini, displayed clear electron density enabling their accurate placement. However, high B-factors limit confidence in precise sequence register and side chain positions in the intervening region. COOT (6) was used for model building and selection of solvent atoms and the model was refined in Phenix.refine (7). Structural validation was performed with Molprobity (8), including analysis of the Ramachandran plots, with 97.5% of residues in the most favored conformations. Five percent of reflections were set aside for calculation of R_{free}. Data collection and refinement statistics are presented in Table 1. Atomic coordinates and structure factors have been deposited in the PDB with ID 6DUE.

Homology Modeling. Using our crystal structure of TgMyoA (1-778) as a template, high confidence homology models of other class XIV myosins were generated using MODELLER v 9.18 (9). Models were generated for MyoA from *P. falciparum*, *C. parvum*, and MyoH from *T. gondii* (TgMyoH), which respectively share 63%, 54%, and 34% sequence identity with TgMyoA (1-778), and harbor no major insertions or deletions (Fig. S1). The best models were chosen based on the lowest discrete optimized protein energy (DOPE) score and GA341 score closest to 1, which were -90,376/1.0 for PfMyoA, -89,210/1.0 for CpMyoA, and -89,376/1.0 for

TgMyoH (Fig.S4 and S1). Comparing these scores to those of models generated without using TgMyoA as template (~-50,000/0.45) supports the conclusion that these models are more accurate than previously possible.

HDX-MS. HDX reactions were conducted with TST-tagged TgMyoA (1-831), wild-type and phosphomimetic mutants, co-expressed with 6xhis-tagged TgELC1 and TgMLC1ΔN66 in 50 μL reactions with a final concentration of 400 nM of each protein per sample. Protein was saturated with Mg^{2+} and ADP to lock TgMyoA in near-rigor state. Reactions were initiated by the addition of 49 μL of D_2O buffer (10 mM HEPES pH 7.5, 100 mM NaCl, 97% D_2O) to 1 μL of protein solution to give a final concentration of 94% D_2O. Exchange was carried out for 3 s and 300 s at 23 °C. Exchange was terminated by the addition of guanidine-HCl (final 0.6 M, pH 2.5) and 0.8% formic acid. Experiments were carried out in triplicate, with all replicate samples obtained from a single preparation of MyoA complex. Samples were immediately frozen in liquid nitrogen and stored at -80 °C until mass analysis. Protein samples were rapidly thawed and injected onto an ultra-performance liquid chromatography (UPLC) system at 2 °C. The protein was run over two immobilized pepsin columns (Applied Biosystems; Porosyme, 2-3131-00) at 10 °C and 2 °C at 200 μL/min for 3 min and the peptides collected onto a VanGuard Precolumn trap (Waters). The trap was subsequently eluted in line with an ACQUITY 1.7 μm particle, 100 x 1 mm^2 C18 UPLC column (Waters), using a gradient of 5%–36% B (buffer A 0.1% formic acid, buffer B 100% acetonitrile) over 16 min. Mass spectrometry experiments were performed on an Impact II QTOF (Bruker) acquiring over a mass range from 150 to 2,200 m/z using an electrospray ionization source operated at a temperature of 200 °C, and a spray voltage of 4.5 kV. Peptides were identified using data-dependent acquisition methods following tandem MS/MS experiments (0.5 s precursor scan from 150-2200 m/z; twelve 0.25 s fragment scans

from 150-2200 m/z). MS/MS datasets were analyzed using PEAKS7 (PEAKS), and a false discovery rate was set at 1% using a database of purified proteins and known contaminants. Notably, no phosphorylation was detected for any peptides from TgMyoA, TgELC1, or TgMLC1, even when reducing the false discovery rate to 0.1%, following purification from Hi5 cells. The parameters for MS/MS searches had a fragment mass error of 0.02 Da, and a precursor mass error of 20 ppm, with phosphorylation searched on all Ser, Thr, and Tyr residues.

HDExaminer Software (Sierra Analytics) was used to calculate peptide deuterium incorporation. All peptides were manually inspected for correct charge state and presence of overlapping peptides. Deuteration levels were calculated using the centroid of the experimental isotope clusters. The average error of all time points and conditions for each HDX project was less than 0.2 Da.

For full HDX-MS data sets, see supplementary document "Dataset S1.xls"

Motility Assays. *in vitro* motility assays were performed with TST-tagged TgMyoA (1-831), wild-type and phosphomimetic mutants, co-expressed with 6xhis-tagged TgELC1 and TgMLC1ΔN66, as previously described (10) with the following modifications: (*a*) mouse Anti Strep-tag Immo antibody (0.2 ug/ml: BioRad Cat# MCA2488) in buffer B was added to the flow cells first, followed by three washes with BSA (in buffer B) and three washes with buffer B; (*b*) 180- 250 ng of purified TgMyoA was used per flow cell; (*c*) buffer C was in all cases supplemented with oxygen scavengers (3 mg/ml glucose, 0.125 mg/ml glucose oxidase, and 0.05 mg/ml catalase); (*d*) all washes with buffer C were done three times; and (*e*) bacterially expressed ELC1 and MLC1 were each added to the flow cells at a concentration of 25 μg/ml. The data were collected on an Eclipse Ti-U inverted microscope (Nikon) equipped with a 100× Plan Apo objective lens (1.49 NA) and a XR/Turbo-Z camera (Stanford Photonics) running Piper

Control software (v2.3.39). The data were collected at 10 frames/s. Actin filament tracking and analysis was done using Imaris x64 v 7.6.4 (Bitplane). Presented mean velocities and error calculations (S.D.) represent five independent experiments with all replicate samples obtained from a single preparation of MyoA protein.

Differential Scanning Fluorimetry. Purified TgMyoA (1-778) and TgMyoA (1-831)/ELC1/MLC1 were diluted to 0.5 mg/mL in buffer containing 150 mM NaCl; 20mM HEPES, pH 7.5; 1 mM $MgCl_2$; 1 mM ADP; 1 mM DTT; and SYPRO® orange dye at a final concentration of 5X. 20 µL reactions were tested in triplicate using the StepOnePlus™ Real-Time PCR System from ThermoFisher, with all replicate samples obtained from a single preparation of MyoA protein.

Pocket Analysis. Pocket analysis was carried out using the OpenEye software OEDocking 3.2 (11). The TgMyoA crystal structure was completely submerged in a box to identify potential pockets using the slower but more precise molecular detection algorithm. Four pockets were readily identified using this approach. The P1 pocket between the converter domain and the L50 domain with a volume of 2700 $Å^3$, the ATP-binding pocket (2200 A^3), the P2 pocket (1070 $Å^3$), and the P3 pocket (550 $Å^3$). Notably, the P2 pocket was not detected until the missing residues of Loop 2 (M632-K639) were modelled *in silico* using I-TASSER (12), followed by energy minimization in Openeye, and manual correction of stereochemistry in Coot (6, 13). The presence or absence of Loop 2 did not affect any other pockets. Representative animal myosin crystal structures* were analysed using the same parameters, for comparison of pocket volumes. Visualization and inspection of each pocket of TgMyoA was carried out with OpenEye's visualization software VIDA 4.3. Residues contacting corresponding pockets were manually selected and are indicated in the structure-guided sequence alignment (Fig S3).

*PPS Myosin Crystal structures used for comparison:

- Class-I: PDB 4BYF (Human); 1LKX, 1VOM (*D. discoideum*)

- Class-II: PDB 5N6A (Bovine); 1BR1, 1BR2 (Chicken); 1DFL, 1QVI (Scallop)

- Class-V: PDB 5HMP, 4ZG4 (Human)

- Class-VI: PDB 2V26, 4E7Z (Porcine)

- Class-X: PDB 5IOI (Human)

Structure-guided sequence alignment. Structure-guided sequence alignment was carried out using the PROMALS3D multiple sequence and structure alignment server (14). Crystal structures were included as input for TgMyoA, HsMyo-Ic, DdMyo-Ie, and DdMyo-II. Additionally, homology models, using the TgMyoA crystal structure as template, were included as input for PfMyoA_3D7 and CpMyoA_IowaII. A full list of sequence accession numbers and PDB entries (where applicable) are as follows:

TgMyoA: PDB 6DUE - Accession, O00934

PfMyoA_3D7: Accession, XP_001350147

CpMyoA_IowaII: Accession, XP_628541

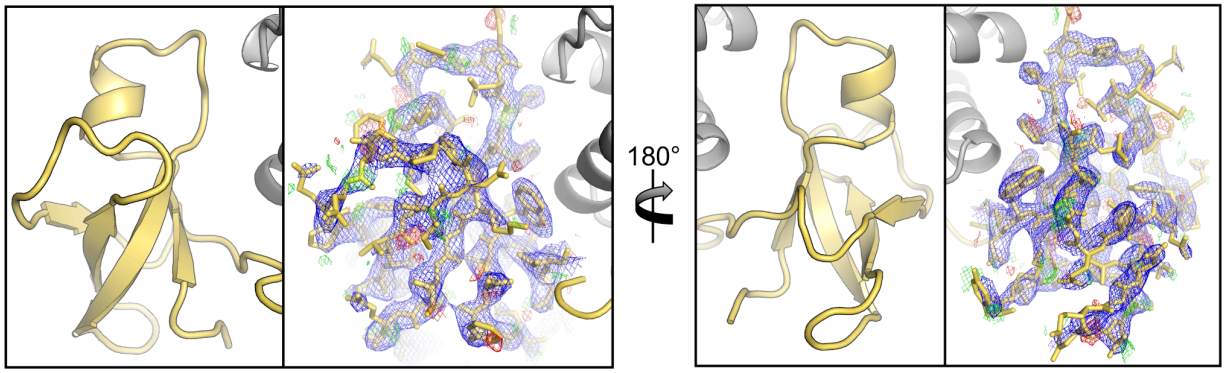
HsMyo-Ic: PDB 4BYF, Accession: NP_001074419

DdMyo-Ie: PDB 1LKX, Accession: XP_636580

DdMyo-II: PDB 1VOM, Accession: XP_637740

A

SH3 Subdomain (TgMyoA Y33-L95)



B

Poly-Ala omit maps

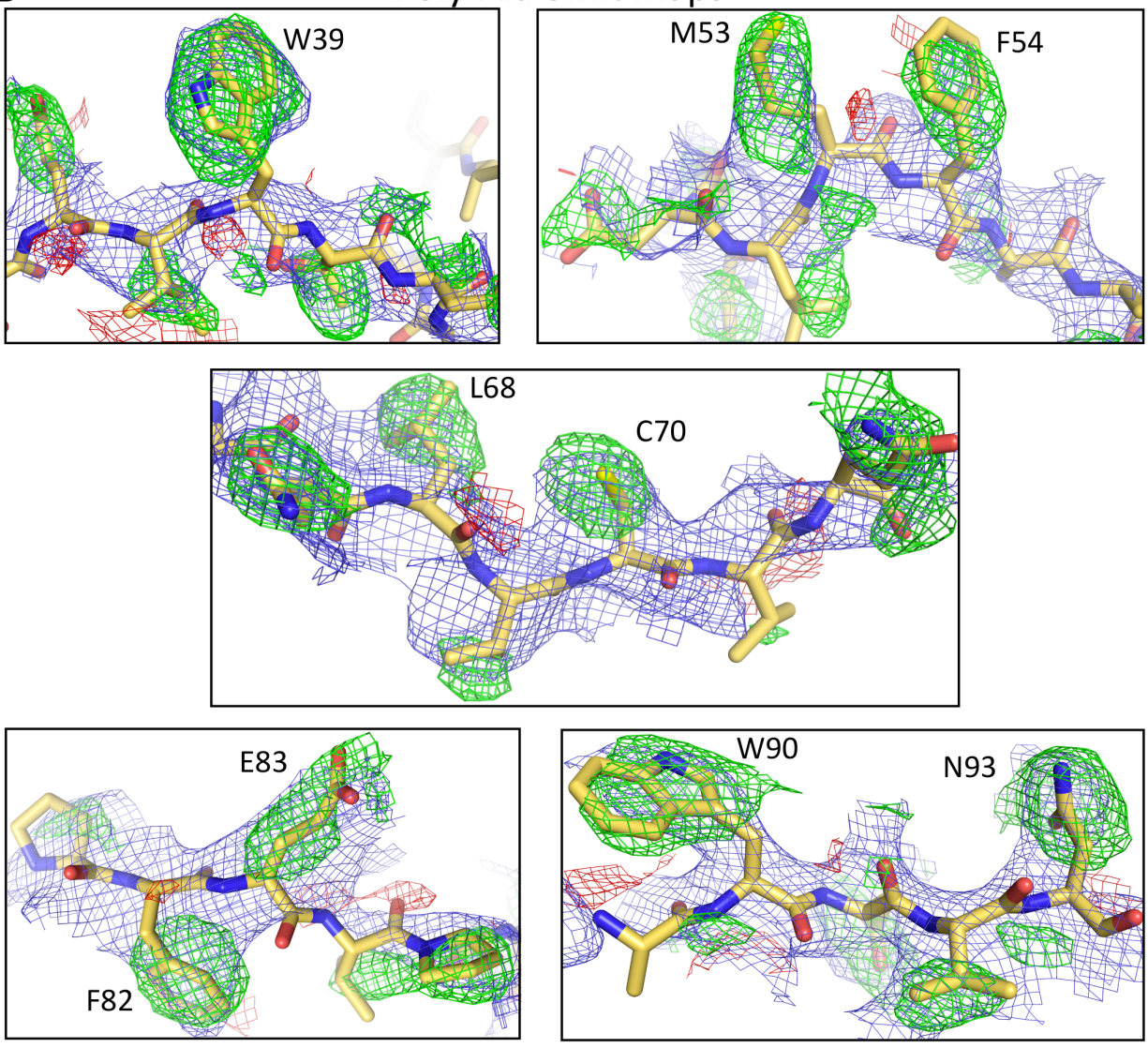


Figure S1. TgMyoA SH3 subdomain. (*A*) Overview of the TgMyoA SH3 subdomain. $2F_o$-F_c electron density map is shown in blue, contoured at 1.0 σ. F_o-F_c difference map is shown in green (+ve) and red (-ve) mesh, contoured at 2.0 σ. (*B*) Close-ups of electron-dense side-chains used to anchor sequence register of the SH3 subdomain. Density maps coloured and contoured same as in (*A*), with the exception that maps were obtained from refinement of a poly-ala model, omiting side-chains from density calculations.

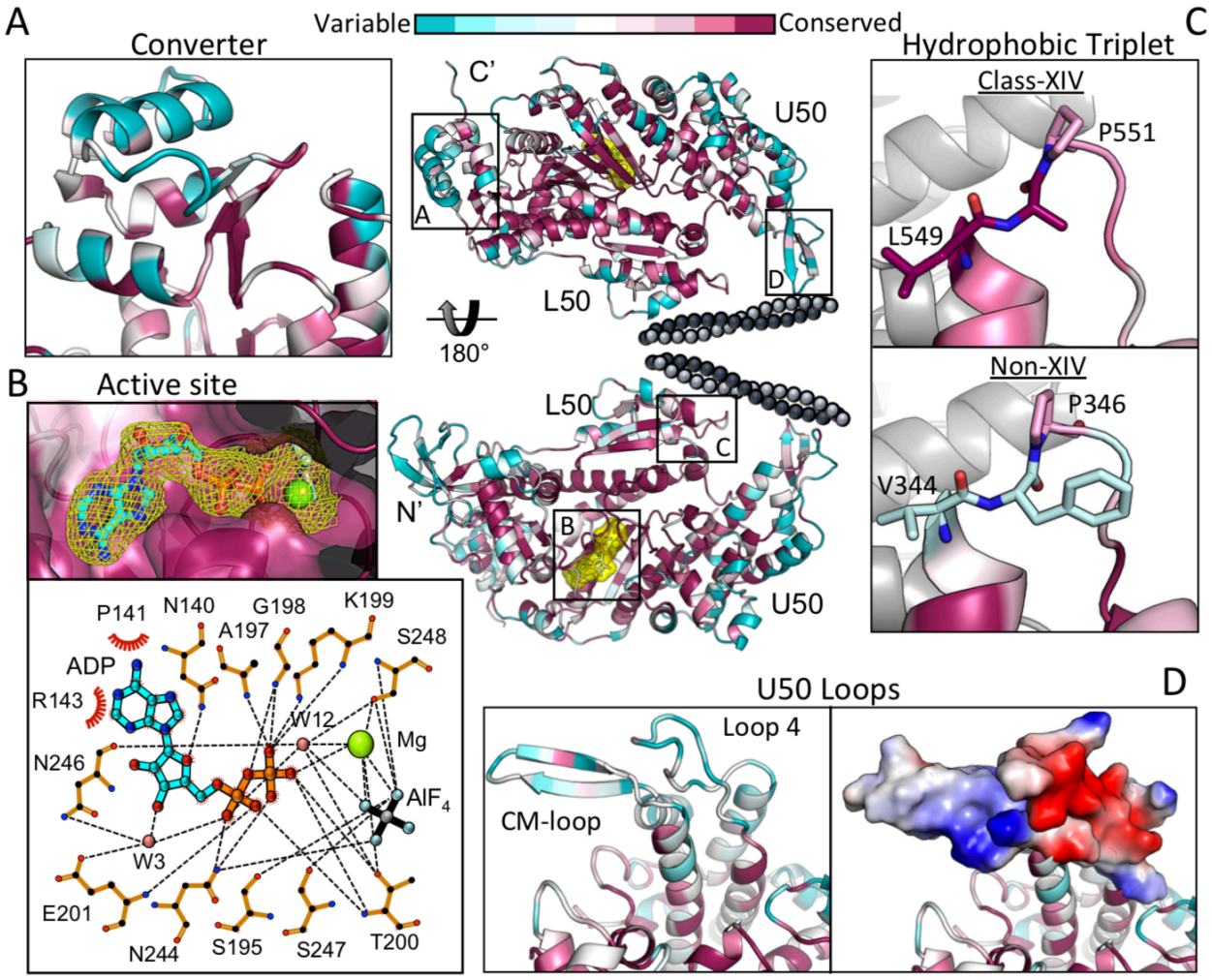


Figure S2. TgMyoA maintains a conserved core, with divergent surface elements. Center: Consurf presentation of the TgMyoA motor domain, with nucleotide surface in yellow. (*A*)

Enlarged and rotated consurf presentation of the TgMyoA converter domain, showing conserved hydrophobic core and variable surface. (*B*) Top, cut-away consurf presentation of nucleotide binding pocket with bound MgADP-AlF_4 (cyan). $2F_o$-F_c electron density map is shown in yellow mesh, contoured at 1.0 σ. Bottom, LigPlot presentation highlighting key interactions in the active site, including hydrogen bonds (black dashed lines) and hydrophobic interactions (red lines). (*C*) Separate consurf analyses of the hydrophobic triplet motifs for class-XIV myosins, mapped on to TgMyoA (Top), and non-class-XIV myosins, mapped onto class-II myosin from *D. discoideum* (bottom) (PDB 1VOM). (*D*) TgMyoA U50 subdomain enlarged and rotated view (Left) highlighting electrostatics of the CM-loop and loop 4 (Right).

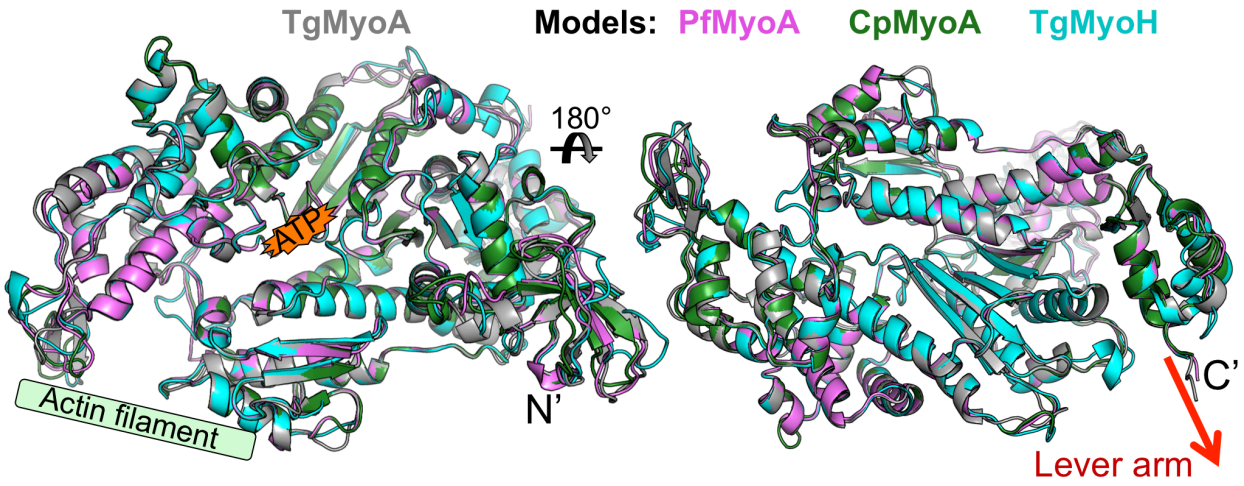


Figure S3. Conservation of Class XIV motors allows for high confidence homology modeling. Crystal structure of TgMyoA motor domain (grey) aligned with homology models of PfMyoA (pink), CpMyoA (green), and TgMyoH (Cyan). Models were generated using TgMyoA crystal structure for initial threading template.

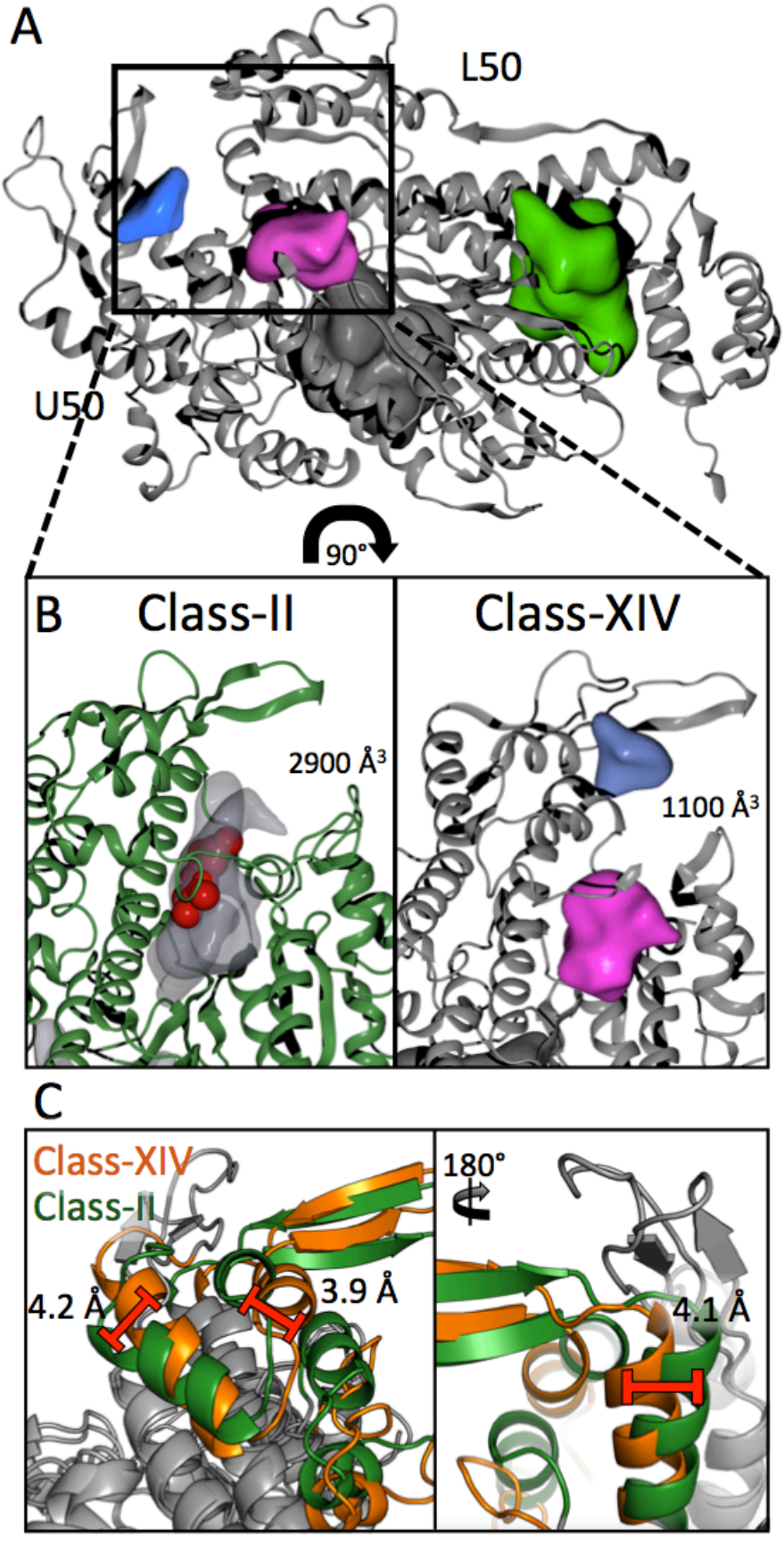

A
L50
U50
90°
B
Class-II
Class-XIV
2900 Å3
1100 Å3
C
Class-XIV
Class-II
4.2 Å
3.9 Å
180°
4.1 Å

Figure S4. Class-XIV myosins have a more closed actin-binding cleft, resulting in a smaller P2 pocket. (*A*) Surface pocket analysis of the TgMyoA crystal structure. Solvent accessible pockets, shown in Green (P1), Purple (P2), and Blue (P3). (*B*) Enlarged and rotated view of the actin-binding cleft, comparing P2 pockets of PPS DdMyo-II (Left) (PDB 1VOM), and PPS TgMyoA (Right), with the volume of the corresponding P2 pocket labeled. Pentabromopseudilin (red spheres) is aligned with 1VOM, based on it's position in the corresponding co-structure (PDB 2JHR). (*C*) U50 domains of TgMyoA (Orange) and DdMyo-II (Green)(PDB 1VOM). Motors were aligned via L50 domains only, in order to highlight ~4 Å closure of class-XIV actin-binding cleft, relative to class-II.

Table S1. TgMyoA (1-778) data collection and refinement statistics. Crystals contained one monomer in the asymmetric unit.

Data collection statistics	
Spacegroup	C121
a, b, c (Å)	99.25, 96.23, 92.05
α, β, γ (deg.)	90, 110.98, 90
Wavelength (Å)	1.19
Resolution range (Å)	85.95 – 2.60 (2.74 – 2.60)
Measured reflections	93,371 (13479)
Unique reflections	24,306 (3,495)
Redundancy	3.8 (3.9)
Completeness (%)	97.6 (96.7)
I/σ(I)	7.3 (1.6)
R_{merge}	0.124 (0.856)
Refinement statistics	
Resolution (Å)	48.11 – 2.6
R_{work}/R_{free}	0.219/0.261
No. of atoms	
Protein	5824
Solvent	25
Ligand	51
B-values (Å^2)	
Protein	68.92
Solvent	53.28
Ligand	56.69
r.m.s. deviation from ideality	
Bond lengths (Å)	0.004
Bond angles (deg.)	0.950
Ramachandran statistics (%)	
Most favoured	97.54
Allowed	2.32
Disallowed	0.14
Values in parentheses are for the highest resolution shell	
5% of reflections were set aside for calculation of R_{free}	

Table S2. All HDX peptide data for TgMyoA WT, S20/21D, S20/21/29D, S743D, and S20/21/743D. See supplemental file "**Table_S2.xlsx**". The residue start and residue end numbers, charge state (Z), retention time (RT), and sequence are displayed for every peptide. The two time points are labeled for the conditions tested. The relative level of HDX is coloured according to the amount of deuterium incorporated, on a blue to red continuum. Experiments were performed in triplicate.

Supplementary References

1. Powell CJ*, et al.* (2017) Dissecting the molecular assembly of the Toxoplasma gondii MyoA motility complex. *J Biol Chem* 292(47):19469-19477.
2. Battye TG, Kontogiannis L, Johnson O, Powell HR, & Leslie AG (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta crystallographica. Section D, Biological crystallography* 67(Pt 4):271-281.
3. Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta crystallographica. Section D, Biological crystallography* 69(Pt 7):1204-1214.
4. McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta crystallographica. Section D, Biological crystallography* 63(Pt 1):32-41.
5. Stein N (2008) CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *J. Appl. Crystallogr.* 41:641-643.
6. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 4):486-501.
7. Afonine PV*, et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta crystallographica. Section D, Biological crystallography* 68(Pt 4):352-367.
8. Chen VB*, et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 1):12-21.
9. Eswar N*, et al.* (2006) Comparative protein structure modeling using Modeller. *Current Protocols in Bioinformatics* Chapter 5:Unit 5 6.
10. Bookwalter CS, Kelsen A, Leung JM, Ward GE, & Trybus KM (2014) A Toxoplasma gondii class XIV myosin, expressed in Sf9 cells with a parasite co-chaperone, requires two light chains for fast motility. *J Biol Chem* 289(44):30832-30841.
11. McGann M (2012) FRED and HYBRID docking performance on standardized datasets. *Journal of computer-aided molecular design* 26(8):897-906.
12. Yang J*, et al.* (2015) The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods* 12(1):7-8.
13. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60(Pt 12 Pt 1):2126-2132.
14. Pei J, Kim B, & Grishin NV (2008) PROMALS3D: a tool for multiple sequence and structure alignment. *Nucleic Acids Res.* 36(7):2295-2300.