Understanding the mechanics of myosin 5a through computational modelling and cryo-EM

Molly Sarah Collins Gravett

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
Astbury Centre for Structural Molecular Biology

December 2022

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 1 contains work from a jointly authored review paper:

Molly S. C. Gravett*, Ryan C. Cocking, Alistair P. Curd, Oliver Harlen, Joanna Leng, Stephen P. Muench, Michelle Peckham, Daniel J. Read, Jarvellis F. Rogers, Robert C. Welch, Sarah A. Harris. Moving in the mesoscale: Understanding the mechanics of cytoskeletal molecular motors by combining mesoscale simulations with imaging. *WIREs Comput Mol Sci*. 2022; 12:e1570. https://doi.org/10.1002/wcms.1570.

Molly S. C. Gravett: Writing - original draft (lead). Ryan C. Cocking: Writing - original draft (equal); writing-review & editing (equal). Alistair P. Curd: Writing - original draft (equal); writing-review & editing (equal). Oliver Harlen: Writing - original draft (equal); writing-review & editing (equal). Joanna Leng: Writing - original draft (equal); writing-review & editing (equal). Stephen P. Muench: Writing - original draft (equal); writing-review & editing (equal). Michelle Peckham: Writing - original draft (equal); writing-review & editing (equal). Daniel J. Read: Writing - original draft (equal); writing-review & editing (equal). Jarvelllis F. Rogers: Writing - original draft (equal); writing-review & editing (equal). Robert C. Welch: Writing - original draft (equal); writing-review & editing (equal). Sarah A. Harris: Conceptualization (lead); writing - original draft (equal); writing-review & editing (equal).

Appendix C contains data which has been published in two papers and one cover:

David P. Klebl, **Molly S. C. Gravett**, Dimitrios Kontziampasis, David J. Wright, Robin S. Bon, Diana C.F. Monteiro, Martin Trebbin, Frank Sobott, Howard D. White, Michele C. Darrow, Rebecca F. Thompson, Stephen P. Muench. Need for Speed: Examining Protein Behaviour during CryoEM Grid Preparation at Different Timescales. *Structure*. 2020; 28: 1238–1248. https://doi.org/10.1016/j.str.2020.07.018

Conceptualization and experimental design: D.P.K., M.S.C.G., D.C.F.M., D.K., M.T., F.S., H.D.W., M.C.D., R.F.T., and S.P.M. Performed experiments: D.P.K., M.S.C.G., D.J.W., M.C.D., and R.F.T. Analysis of data: D.P.K., M.S.C.G., F.S., M.C.D., R.F.T., and S.P.M. Supervision: F.S., R.S.B., H.D.W., R.F.T., and S.P.M. Discussed the data and wrote the manuscript: All authors.

James A. Poulter*, **Molly S. C. Gravett*,** Rachel L. Taylor,Kaoru Fujinami, Julie De Zaeytijd, James Bellingham, Atta Ur Rehman, Takaaki Hayashi, Mineo Kondo, Abdur Rehman, Muhammad Ansar, Dan Donnelly, Carmel Toomes, Manir Ali, UK Inherited Retinal Disease Consortium, Genomics England Research Consortium, Elfride De Baere, Bart P. Leroy, Nigel P. Davies, Robert H. Henderson, Andrew R. Webster, Carlo Rivolta, Christina Zeitz,Omar A. Mahroo, Gavin Arno, Graeme C. M. Black, Martin McKibbin, Sarah A. Harris, Kamron N. Khan, Chris F. Inglehearn. New variants and in silico analyses in GRK1 associated Oguchi disease. *Human Mutation*. 2021; 42: 164-176. https://doi.org/10.1002/humu.24140

James A. Poulter and Molly S. C. Gravett should be considered joint first author (*). Molly S. C. Gravett was responsible for computational modelling and analysis. Sarah A. Harris, Kamron N. Khan, and Chris F. Inglehearn should be considered joint senior author.

James A. Poulter*, **Molly S. C. Gravett*,** Rachel L. Taylor,Kaoru Fujinami, Julie De Zaeytijd, James Bellingham, Atta Ur Rehman, Takaaki Hayashi, Mineo Kondo, Abdur Rehman, Muhammad Ansar, Dan Donnelly, Carmel Toomes, Manir Ali, UK Inherited Retinal Disease Consortium, Genomics England Research Consortium, Elfride De Baere, Bart P. Leroy, Nigel P. Davies, Robert H. Henderson, Andrew R. Webster, Carlo Rivolta, Christina Zeitz,Omar A. Mahroo, Gavin Arno, Graeme C. M. Black, Martin McKibbin, Sarah A. Harris, Kamron N. Khan, Chris F. Inglehearn. Cover, Volume 42, Issue 2. *Human Mutation*. 2021; 42: ii-ii. https://doi.org/10.1002/humu.24169

Cover design and creation Molly S. C. Gravett

Acknowledgements

I would like to thank my supervisory team, Sarah Harris, Ste Muench and Michelle Peckham. Thank you for making the experience so fun, I have felt super supported and extremely encouraged throughout. It has truly been a pleasure.

Thank you to everyone in the FFEA team, Oliver Harlen, Jo Leng, Daniel Read, Ryan Cocking, Llis Rogers, Rob Welch, and Ben Hanson. Our Friday mornings of cake and gossip are cherished.

Thank you to the Muench group, especially David Klebl for teaching me EM, fun discussions, and the liquid ethane song.

Thank you to the Peckham group, in particular Glenn Carrington and David Casas-Mao for making the lab less scary after a long break.

A special thank you to Peter Knight, Charlie Scarff, Harry Takagi, and Howard White for invaluable myosin chats.

I would like to also thank the Astbury Biostructure Laboratory team for all their help and support over the years.

Thanks to my loves, Franny Chandler, Vronnie DeJesus, Benji Lane, Cross Smith, and Ammy Turner, there has been a great comfort in the shared pain. Also, thanks to the rest of the lunchtime crunchtime gang, Alex Flynn and Laura Marr.

Thank you to Robbie Cook, Harry Coulter and Ellen (The Generous) Dixon for making lockdown bearable and letting me practice presentations on you, even when you had no clue what I was saying.

Thank you to Nikkita Beghi for encouraging me to do this at all, Adam Hanss for a roof over my head, and Chloe Peglau for keeping me zen.

Finally, thank you to Ollie Debski-Antoniak for being a constant from both near and afar, and my number 1 cheerleader. I can do it, WOO WOO WOO!

Abstract

Myosin 5a (Myo5a) is responsible for transporting cellular cargos, such as the endoplasmic reticulum, along filamentous actin (F-actin) to the periphery of cells. The lever has a key role in the function of Myo5a, however little is known about its properties in the active state, and how this contributes to Myo5a walking mechanics. In conjunction with this, neither experimental techniques, nor atomistic molecular dynamics simulations, have the appropriate spatiotemporal resolution to visualise the large-scale dynamics this motor performs (too low and too high, respectively). Fluctuating Finite Element Analysis (FFEA) is a tool that can bridge this gap. FFEA simulates protein dynamics across large spatiotemporal scales (10 nm to 500 nm, and 1 μs to 1 s) with little computational expense.

Here I present a cryo-EM structure of Myo5a-S1 (monomeric) in the rigor state bound to F-actin, with sufficient resolution to visualise detail within the lever, and reveal novel properties. Through integration of this data into an FFEA model of Myo5a-HMM (dimeric), it was possible to determine the function of these features in the context of the dimer. Additionally, I describe the requirements for building a FFEA model of Myo5a-HMM bound to F-actin and under strain (out of equilibrium) that with minor improvements would be capable of walking. Finally, this work presents a pipeline to go from cryo-EM data to a non-equilibrium FFEA model.

Cytoskeletal motors have a broad range of functions and are therefore implicated in a multitude of disease pathways. The pipeline described is broadly applicable to all cytoskeletal motors. Furthermore, this work sheds light on how nanoscale motors function in a highly viscous environment. Therefore, this research is of broad interest to medicine, fundamental biology, physics and engineering alike.

Table of Contents

List of Abbreviations

°	Degrees
°C	Degrees celsius
1D	1-dimensional
2D	2-dimensional
3D	3-dimensional
Å	Angstrom
A_{280}	Absorbance at 280 nm
A_{320}	Absorbance at 320 nm
ADP	Adenosine diphosphate
ADP-P_i	Hydrolysed ATP
AFM	Atomic force microscopy
apo-CaM	Ca^{2+}-free calmodulin
ARC4	Advanced research computing 4
ATP	Adenosine triphosphate
BeF_3	Beryllium trifluoride
blob	BioLOgical Body
Ca^{2+}	Calcium ion
Calm2	Calmodulin 2
CaM	Calmodulin
CaMn	Calmodulin bound to IQn
CBD	Cargo binding domain
C-lobe	C-terminal lobe
CPU	Central processing unit
Cryo-EM	Cryo-electron microscopy
Cryo-ET	Cryo-electron tomography
CTF	Contrast transfer function
ctforce	Constant force
D-loop	DNase I binding loop
DNA	Deoxyribonucleic acid

dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGTA	Egtazic acid
ELC	Essential light chain
EM	Electron microscopy
EMDB	Electron Microscopy Data Bank
ENM	Elastic network model
F-actin	Filamentous actin
FFEA	Fluctuating Finite Element Analysis
FSC	Fourier shell correlation
Gctf	GPU accelerated CTF
GPU	Graphics processing unit
GRK1	G-Protein coupled receptor kinase 1
GTP	Guanosine triphosphate
HCM	Hypertrophic cardiomyopathy
HF	High-fidelity
HLH	Helix-loop-helix
HMM	Heavy meromyosin
hs-AFM	High-speed atomic force microscopy
HSP90	Heat shock protein 90
HT	Head-tail junction
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IQ	Isoleucine-glutamine
iSCAT	Interferometric light scattering microscopy
k_B	Boltzmann constant
kbp	Kilobase pair
kDa	Kilodalton
KOBRA	KirchOff Biological Rod Algorithm
L	Litre

L50	Lower 50 kDa
LCBD	Light chain binding domain
LN_2	Liquid nitrogen
M	Molar
MD	Molecular dynamics
MEDYAN	Mechanochemical Dynamics of Active Networks
Mg^{2+}	Magnesium ion
mins	Minutes
MLC1P	Myosin light chain 1P
MOI	Multiplicity of infection
MOPS	3-(N-morpholino)propanesulfonic acid
MRC	Medical research council
Myo5a	Myosin 5a
Myo5a-S1-6IQ	Myo5a S1 with all 6 IQ domains
N	Newton
N-lobe	N-terminal lobe
NMR	Nuclear magnetic resonance
ns-EM	Negative stain electron microscopy
Pa	Pascal
PBS	Phosphate-buffered saline
PC	Principal component
PCA	Principal component analysis
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PFU	Plaque forming units
P_i	Inorganic phosphate
pmemd	Particle Mesh Ewald Molecular Dynamics
PMSF	Phenylmethylsulfonyl fluoride
post-PS	Post-powerstroke
pre-PS	Pre-powerstroke

res	Residue
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
rpm	Revolutions per minute
S.O.C Medium	Super Optimal broth with Catabolite repression
S1	Subfragment-1
SAH	Single alpha helix
SD	Standard deviation
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SEM	Scanning electron micrograph
SH3	SRC homology 3
SPOC	Statistical Processing of Cryo-EM maps
STORM	Super-resolution stochastic optical reconstruction microscopy
T	Temperature
TAE	Tris-acetate-EDTA
TB	Terabyte
TIP3P	Transferable intermolecular potential 3P
TIRF	Total internal reflection fluorescence
U50	Upper 50 kDa
UCSF	University of California San Francisco
UK	United Kingdom
Unc45B	Unc-45 Myosin Chaperone B
USA	United States of America
VMD	Visual molecular dynamics
VO_4	Vanadate
VTK	Visualization Toolkit
x g	Times gravity
σ^2	Variance

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Cytoskeletal motors

Cytoskeletal motors are responsible for essential cellular processes, such as transport, motility, division, and growth (Geeves and Holmes, 1999; Houdusse and Sweeney, 2016; Sweeney and Holzbaur, 2018). They are mechanoenzymes that convert chemical energy, in the form of ATP, to unidirectional mechanical motion. However, there is little understanding of how small-scale changes at the active site are translated to their large scale coordinated movement. The mechanics of motors operating at the nanoscale are largely different to macroscopic motors, due to the viscosity of their environment, and the fact they are subject to thermal fluctuations and Brownian motion (Purcell, 1977; Bustamante et al., 2001). Because of their ability to endure and navigate this distinct environment, and the integral role of cytoskeletal motors in cellular function, understanding the relationship between their mechanical properties and their function is of interest to the fields of biology, physics and engineering alike (Bustamante et al., 2001; Nelson et al., 2010; Hess, 2011; Robert-Paganin et al., 2020).

There are 3 superfamilies of cytoskeletal motor: myosins which bind to filamentous actin (F-actin), and kinesins and dyneins which bind microtubules (Sweeney and Holzbaur, 2018). Myosins are a large class of motor proteins, and are one of the most well understood motor and track systems (Robert-Paganin et al., 2020).

1.1.1 Myosin

The myosin superfamily is made up of approximately 80 classes with a diverse array of functions, such as vesicle transport, muscle contraction and membrane tethering (Foth et al., 2006; Hartman and Spudich, 2012; Kollmar and Mühlhausen, 2017). The human genome contains almost 40 myosin genes organised into 12 classes (Figure 1.1) (Berg et al., 2001; Peckham, 2016). Myosins are therefore implicated in a multitude of disease states (Hartman and Spudich, 2012; Yotti et al., 2019; Trivedi et al., 2020), such as heart disease (Barrick et al., 2021), muscle and bleeding disorders (Parker and Peckham, 2020), deafness (Cirilo et al., 2021), and cancer (Peckham, 2016). Thus, studying the myosin superfamily has broad medical significance.

1.1.1.1 Myosin blueprint

Myosins are multimeric protein complexes consisting of heavy and light chains. The heavy chain is composed of an N-terminal motor domain, followed by a light chain binding domain (LCBD), and a C-terminal tail (Figure 1.1). The motor domain is highly conserved between myosins, and contains both the nucleotide and F-actin binding sites. It can be subdivided into 4 mobile domains that move relative to each other: the N-terminal domain, the upper 50 kDa (U50) domain, the lower 50 kDa (L50) domain, and the converter (Figure 1.2A). The nucleotide binding site is located between the U50 and L50 domains, which also form the actin binding cleft. When the U50 and L50 domains come together, closing the cleft between them, a high-affinity actin binding site is formed (Coureux et al., 2003) (Figure 1.2A).

The lever is made up of the converter, the LCBD and bound light chains, whilst the motor and the lever together comprise the myosin head (Figure 1.2B). Small-scale conformational changes in the motor are amplified to large-scale motions of the lever via the converter. The movement of the lever on F-actin is known as the powerstroke (Geeves and Holmes, 1999; Tyska and Warshaw, 2002). The length of the LCBD, and the number and type of light chains bound, varies from myosin to myosin (Heissler and Sellers, 2015). The characteristic length of the powerstroke and/or step size of each myosin is directly related to the length of the lever (Uyeda et al., 1996; Sakamoto et al., 2003; Sakamoto et al., 2005; Oke et al., 2010).

The LCBD contains a variable number of IQ (isoleucine-glutamine) motifs which bind myosin light chains (Figure 1.2D&E) (Heissler and Sellers, 2015). IQ motifs are α-helical and have the consensus sequence IQXXXRGXXXR (Mooseker and Cheney, 1995). The pattern and spacing of IQ motifs in the LCBD varies between myosins, they typically appear as multiple tandem repeats alternating between 22-29 residue spacing (counting from Ile) (Bähler and Rhoads, 2002).

Myosin light chains reinforce the lever domain, and are involved in protein-protein interactions and regulation (Heissler and Sellers, 2015). They are members of the calmodulin (CaM) family or are CaM-related proteins (Heissler and Sellers, 2015), and have a conserved structure. CaM-related light chains generally consist of 2 globular regions, the N-lobe (N-terminal) and C-lobe (C-terminal), joined by a flexible linker. The C-lobe forms contacts with the N-terminal region of the IQ motif, and the N-lobe forms contacts with the C-terminal region of the IQ motif. Light chains bind IQ motifs either in a compacted conformation, with both lobes forming contacts with the IQ motif (Figure 1.2E), or an extended conformation, in which only the C-lobe binds the IQ motif (Terrak et al., 2003; Terrak et al., 2005; Houdusse et al., 2006).

Finally, the tail is responsible for the cellular function of myosin through its binding to other proteins. In dimeric myosins (myosin-2, -5 and -18), the tail contains a coiled-coil forming region. In non-muscle myosins the tail can contain one or more domains, such as a cargo-binding domain (CBD) (Hartman and Spudich, 2012).

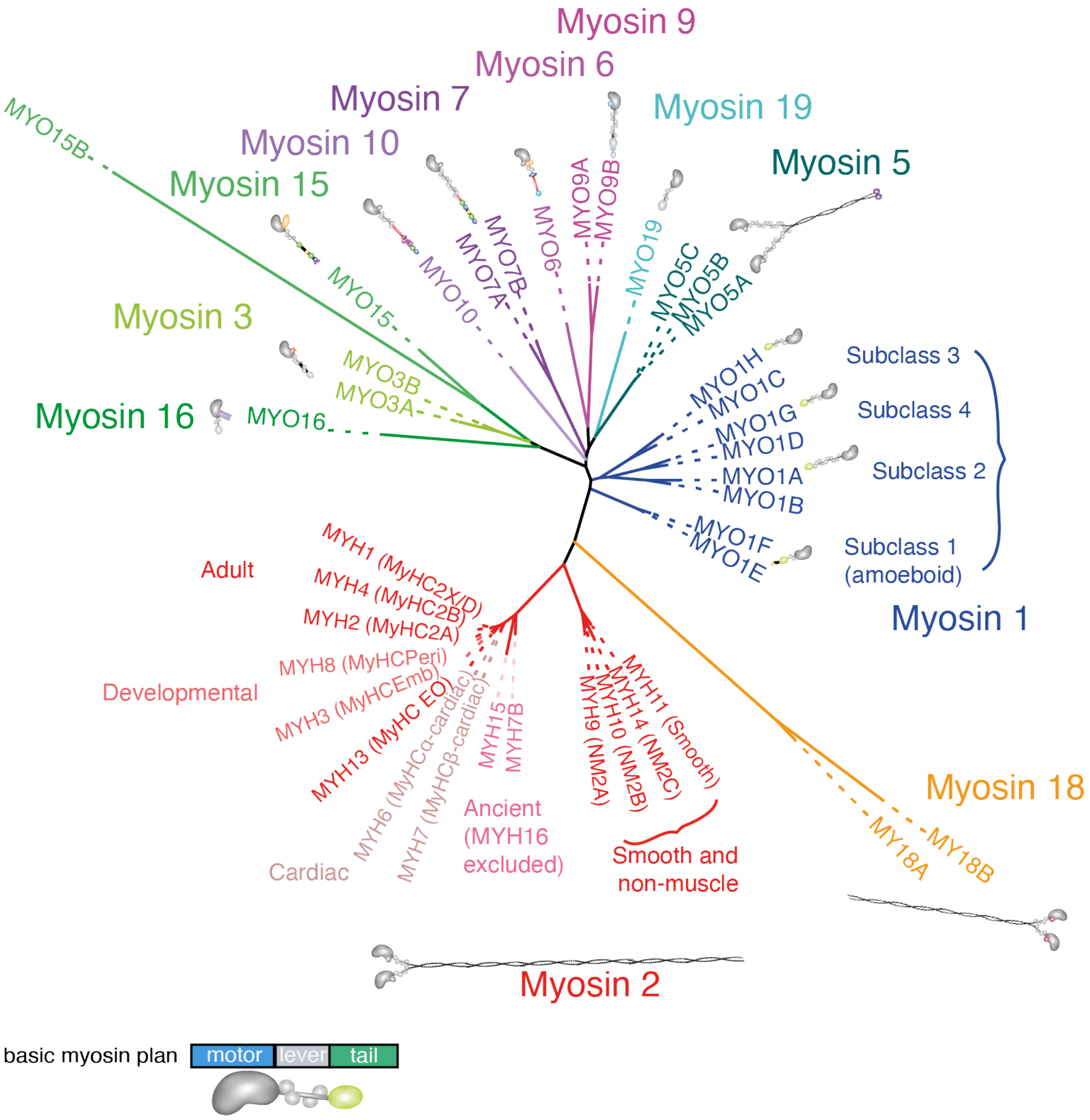


Figure 1.1: The myosin family tree for humans

Figure created by Michelle Peckham. Trees were constructed using Phyml (http://www.phylogeny.fr). Sequences for human myosins downloaded from Uniprot (Consortium et al., 2022). Tail domains shown as described in Uniprot (Consortium et al., 2022). Insert is a schematic of the basic myosin plan.

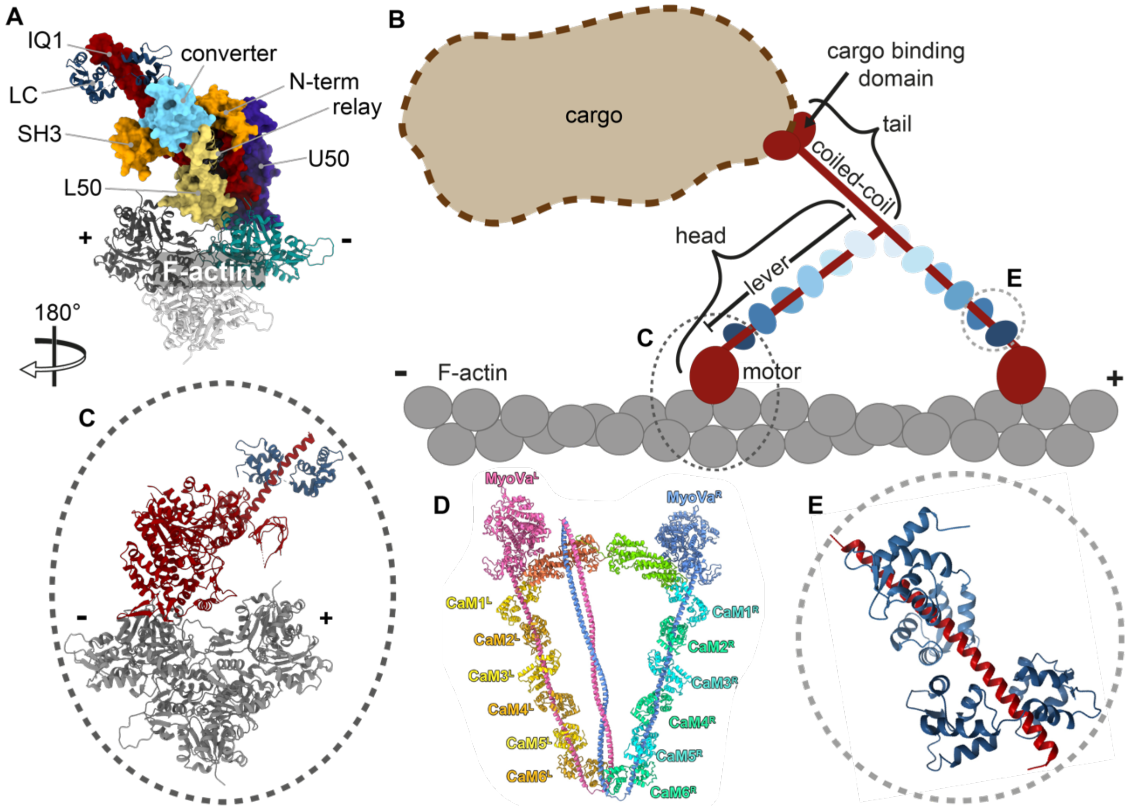


Figure 1.2: The structure of Myo5a

A: *Cryo-EM structure of Myo5a-S1-1IQ bound to F-actin (PDB: 7PM6 (Pospich et al., 2021)) with the mobile subdomains of the motor domain highlighted. In Myo5a the motor domain contains an N-terminal SH3-like fold (SH3). The relay helix (relay) translates conformational changes in the actin-binding cleft to the converter. The N-terminal domain (N-term) is displayed in orange, the converter in light blue and the U50 (Upper 50 kDa) in purple and the L50 (Lower 50 kDa) in yellow. The bound light chain is shown as a cartoon in dark blue. The F-actin subunit closest to the minus end (-) is coloured in teal and the plus end dark grey.* ***B:*** *Schematic of Myo5a walking on F-actin with domains highlighted. The heavy chain is coloured in red and light chains are coloured in blue.* ***C: A*** *rotated 180° about the y-axis and coloured as in* ***B. D:*** *Structure of full-length Myo5a in the shutdown state, taken from* (Niu et al., 2022). *CaMn refers to the CaM bound to IQn. Heads distinguished as left (L) and right (R).* ***E:*** *X-ray crystallography structure of IQ1-2 of Myo5a bound to apo-CaM. Coloured as in* ***B****.*

1.1.2 Unconventional Myo5a

1.1.2.1 Function

Unconventional myosin 5a (Myo5a) is one of three vertebrate myosin 5 isoforms (a, b & c) (Hammer and Sellers, 2012). It is expressed in pigment producing cells, found in the skin, brain, ears, eyes and hair, called melanocytes (Provance and Mercer, 1999), and Purkinje neurons (Wagner et al., 2011). Myo5a is a non-muscle myosin primarily involved in the transport of cellular cargos along cytoskeletal F-actin to the periphery of cells (Hammer and Sellers, 2012). Such cargos include pigment-containing melanosomes, for protection from ultraviolet radiation (Reck-Peterson et al., 2000), and the endoplasmic reticulum in the dendritic spines of Purkinje neurons, important for motor learning (Wagner et al., 2011). Mutations in Myo5a lead to the pigmentation and neuronal defects seen in the rare disease Griscelli Syndrome (Mercer et al., 1991; Pastural et al., 1997; Van Gele et al., 2009). Myo5a has also been implicated in cancer progression (Lan et al., 2010; Alves et al., 2013). Myo5a is one of the most well-characterised myosins, making it an ideal system to study structure-property-function relationships.

1.1.2.2 Structure of divergent regions in Myo5a

In Myo5a, the LCBD contains 6 IQ motifs, each of which is bound to a CaM light chain, or member of the CaM family (Cheney et al., 1993). In the Myo5a LCBD the IQ motifs are organised into a repeating pattern of 23 then 25 residues in length, which is thought to be important for encoding alternating regions of stiffness and flexibility (Terrak et al., 2005). In mouse Myo5a, all bound light chains are CaM, however in chicken Myo5a, essential light chain binds the first IQ (Espindola et al., 2000; Wang et al., 2000).

CaM is Ca^{2+} sensitive and binds Ca^{2+} ions through the 2 EF-hands in each lobe. Ca^{2+} binding can either change the conformation of CaM bound to an IQ motif or cause it to dissociate, which is thought to play a role in the regulation of myosin (Watanabe et al., 2005; Koide et al., 2006; Trybus et al., 2007; Wada et al., 2016; Shen et al., 2016). Micromolar levels of Ca^{2+} have been shown to activate Mg.ATPase activity through disrupting the shutdown state (Wang et al., 2004; Niu et al., 2022). However, Ca^{2+} can also inhibit motor function by inducing dissociation of CaM from the LCBD leading to a floppy non-functional lever domain (Nguyen and Higuchi, 2005; Trybus et al., 2007). Myo5a is therefore a Ca^{2+} sensor and involved in Ca^{2+}-regulated vesicle transport and exocytosis, for example Ca^{2+}-CaM dissociation from the LCBD exposes IQ motifs for binding to the exocytosis mediator Syntaxin-1A (Watanabe et al., 2005; Wada et al., 2016).

The tail of Myo5a is composed of a coiled-coil domain and a CBD (Figure 1.2B). Myo5a monomers dimerise at the tail domain through formation of a coiled-coil. The CBD forms complexes with vesicle associated proteins such as Rab27a and melanophilin for melanosome transport (Strom et al., 2002), and interacts with the motor domains in the inactive shutdown state (Thirumurugan et al., 2006; Liu et al., 2006a; Zhang et al., 2018; Niu et al., 2022).

For *in vitro* experiments, Myo5a is often expressed in the form of monomeric subfragment-1 (S1) or dimeric heavy meromyosin (HMM). S1 is a single myosin head (residues 1-907), and HMM includes the motor, lever and the coiled-coil forming domain to preserve dimerization (residues 1-1091) (Heissler and Sellers, 2016). HMM is constitutively active as it lacks the CBD meaning it cannot form the shutdown state.

Currently there are a number of high-resolution structures of individual components of Myo5a (Figure 1.2A,C&E, Table 1.1), and the full-length molecule in the shutdown state (Figure 1.2D) (Niu et al., 2022). However, only low-resolution data from negative stain electron microscopy (ns-EM) is available for constructs containing the full-length lever (S1/HMM) and in the active state, such as Myo5a-HMM bound to actin (Walker et al., 2000; Oke et al., 2010). Despite these advances, higher resolution structures that contain the full-length lever in the active state are necessary. Such structures would reveal how each of the previously resolved components contribute to the overall mechanics and function of Myo5a.

Table 1.1 - Deposited Myo5a structures (or relevant homologues)

Structure	State	Nucleotide	Light chain	PDB	Resolution (Å)	Method	Reference
Myosin 5c, motor	Pre-powerstroke	ADP-VO$_4$ (ADP-P$_i$ analogue)	N/A	4ZG4	2.4	X-ray diffraction	(Wulf et al., 2016)
Actomyosin-5a, motor + 1IQ	Post-powerstroke	ADP	Myosin light chain 6B	7PM6-9, 7PMA-C	3.0/3.1	Cryo-EM	(Pospich et al., 2021)
Actomyosin-5a, motor + 1IQ	Rigor	N/A	Myosin light chain 6B	7PM0-2, 7PLT-Z	3.2/3.3	Cryo-EM	(Pospich et al., 2021)
Actomyosin-5a, motor + 1IQ	Post-rigor transition state?	AppNhp ADP-BeF$_3$ (ATP analogue)	Myosin light chain 6B	7PMD-J, 7PML	2.9	Cryo-EM	(Pospich et al., 2021)
Myo5a + 1IQ	Post-rigor	ADP-BeF$_3$ (ATP analogue)	Myosin light chain 6B	1W7J		X-ray diffraction	(Coureux et al., 2004)
Myo5a, IQ1 + IQ2	N/A	N/A	CaM	2IX7	2.5	X-ray diffraction	(Houdusse et al., 2006)
Myo2p (yeast class 5), IQ2 + IQ3	N/A	N/A	Myosin light chain 1P	1N2D	2.0	X-ray diffraction	(Terrak et al., 2005)
Myo5a, full-length	Autoinhibited shutdown state	N/A	CaM	2DFS, 7YV9	24 4.8	Cryo-ET, Cryo-EM	(Liu et al., 2006a; Niu et al., 2022)
Myo5a, CBD	Apo	N/A	N/A	4J5L, 4LLI, 4LX1, 3WB8	2.2, 2.2, 1.87, 2.5	X-ray diffraction	(Wei et al., 2013; Nascimento et al., 2013; Velvarska and Niessing, 2013; Pylypenko et al., 2013)

1.1.3 ATPase cycle

Myosin is thought to have 4 conformational energy minima, which are stabilised by the 4 nucleotide states of the ATPase cycle (Pospich et al., 2021). These include the: pre-powerstroke (pre-PS) state (ADP-P_i bound), post-powerstroke (post-PS) state (ADP bound), rigor state (nucleotide free), and the post-rigor state (ATP bound). Each conformation is associated with an F-actin binding affinity and lever position. Cryo-EM and x-ray crystallography structures of the motor domain highlight the critical differences between these states in Myo5a.

The pre-PS state, also known as the primed state, has ADP-P_i (hydrolysed ATP) in the nucleotide binding pocket, weak F-actin binding affinity and the lever in the primed position (Figure 1.3, Figure 1.4E & Figure 1.5E). In the primed or pre-PS position the lever points towards the minus end of F-actin. A crystal structure of myosin-5c, containing the ADP-P_i analogue ADP-VO_4, shows partial actin binding cleft closure near the nucleotide binding pocket, but not near the F-actin binding site (Wulf et al., 2016). This partial closure is thought to allow weak initial F-actin binding. Although there is currently no published structure of the pre-PS state bound to F-actin, it is thought initial non-specific interactions with F-actin are formed through loop 2, and then a stronger interaction forms between L50 domain and the actin subunit (Holmes et al., 2003; Yengo and Sweeney, 2004). It is likely that this interaction forms between the helix-loop-helix (HLH) and loop 3 of the L50 domain and the D-loop of the actin subunit, based upon the L50-actin interactions in the ADP bound actomyosin 5a structure (Pospich et al., 2021). In the ADP-VO_4 bound myosin-5c structure the relay helix is bent and the converter is in the primed position (Wulf et al., 2016).

Actin binding accelerates P_i release (>250 s^{-1}) and the formation of the ADP bound post-PS state (De La Cruz et al., 1999). Based on data from myosin 6 it is thought P_i exits through a tunnel in the inner cleft (back door) (Llinas et al., 2015). When P_i is released the actin binding cleft closes and the lever arm completes the powerstroke by swinging from the pre-PS, primed state towards the plus end of F-actin (Figure 1.3, Figure 1.4A & Figure 1.5A). Cleft closure creates the high affinity actin interface, by bringing the U50 domain down towards F-actin. The hypertrophic cardiomyopathy (HCM) loop, loop 4 and loop 2 can then form stable interactions with the neighbouring F-actin subunit (on the minus end side), and the interactions through the L50 domain are maintained (Pospich et al., 2021). In Myo5a, ADP release is the rate limiting step (~80 ms) meaning most of the ATPase cycle is spent in the ADP state, which has high F-actin affinity. Therefore, Myo5a is described as having a high duty ratio, as a large proportion of its ATPase cycle is spent bound to F-actin in the high affinity state (67 %) (De La Cruz et al., 1999).

The nucleotide free, rigor state is formed when Mg^{2+} is released followed by ADP (Rosenfeld and Sweeney, 2004; Coureux et al., 2004; Hannemann et al., 2005). The actin binding interface is maintained between the ADP and rigor states, and the cleft remains closed (Figure 1.3, Figure 1.4B &Figure 1.5B) (Pospich et al., 2021). However, there is pivoting of the U50 domain, which is caused by rearrangements in the active site that prevent the coordination of Mg^{2+}, and binding of the nucleotide with high affinity. This pivoting results in a 9° rotation of the lever domain from the ADP state to the rigor state towards the plus end of F-actin (Pospich et al., 2021).

ATP binds the rigor state resulting in conformational changes in the head (~900 μs at 1 mM ATP (saturating)) (De La Cruz et al., 1999; Coureux et al., 2004). From a crystal structure of Myo5a-S1-1IQ bound to ATP analogue ADP-BeF_3, it was determined that coordination of the γ- and β-phosphates of ATP and the Mg^{2+} ion in the active site is coupled with cleft opening (Figure 1.3, Figure 1.4C & Figure 1.5C). Cleft opening destabilises F-actin binding leading to detachment (Figure 1.3, Figure 1.4D & Figure 1.5D). It is thought the ATP bound, post-rigor state is flexible, and there is no coupling between the subdomains. However, there is minimal movement of converter and lever between the rigor and post-rigor state, thus ATP binding does not reverse lever arm movement (Coureux et al., 2004). Re-priming into the pre-PS state occurs in a process known as the recovery stroke, this rearranges the catalytic site so that ATP can then be hydrolysed (~1 ms) (Geeves and Holmes, 1999; De La Cruz et al., 1999; Shiroguchi et al., 2011).

There is an additional ADP bound state in the dimer when both heads are bound to F-actin (Figure 1.6.1). The trail head is in the post-PS conformation, as expected for the F-actin bound, ADP bound state. However, the trail head prevents the powerstroke occurring in the lead head after P_i release, holding it in the pre-PS position despite the molecule being F-actin and ADP bound (Walker et al., 2000; Burgess et al., 2002; Oke et al., 2010). In this double head bound state, the strain on the lead head means ADP release is ~250-fold slower from the lead head than the trail head (ADP release in the trail head is the same as S1) (Veigel et al., 2002; Forgacs et al., 2008). There may be a variety of conformations for this state dependent on the distance between the two heads. ns-EM has shown that the lead lever can be: sharply bent where it joins the motor, gradually bent throughout, or a small portion are even in the post-PS conformation (Oke et al., 2010). The ratio of particles in each of these conformations is dependent upon the number of actin subunits between each head (11, 13 or 15). However, on average the lead lever is angled back at ~115° relative to F-actin reminiscent of a pre-PS state (Walker et al., 2000). An F-actin and ADP bound structure of myosin-1b with the lever pointing towards the minus end of the actin filament has been reported (PDB: 6C1D) (Mentes et al., 2018). However, it is unclear whether this structure is representative of a strained state.

It is not well understood how the energy from ATP hydrolysis is stored by myosin. Despite the pre-PS state being called the primed state, the recovery stroke (entrance into the pre-PS state) has to occur before hydrolysis (Geeves and Holmes, 1999). This is because switch 2 in the active site must be closed to bring the catalytic residues into the correct position for hydrolysis and this closure occurs during the recovery stroke (Geeves and Holmes, 1999). In the pre-PS state hydrolysis is reversible, and Pi release is also reversible provided the powerstroke has not yet happened (Bowater et al., 1990; Holmes et al., 2004). It is only once the powerstroke has occurred that the process becomes irreversible, as the energy has been dissipated (Geeves, 2016). There are two theories for energy storage I am aware of, one involving structural distortions such as lever-arm bending and twisting of the central β-sheet in the motor domain (Geeves, 2016; Robert-Paganin et al., 2020), which is a way of spring loading the molecule, and another that relies on thermodynamics (Zhang and Feng, 2016). In the latter they suggest instead of the energy of ATP being released in hydrolysis, that it is released upon ATP binding to myosin, which is required to break the strong interaction between F-actin and myosin. In this theory the powerstroke is instead driven by the energy released when myosin binds F-actin, which is only available because ATP broke this interaction in the first place.

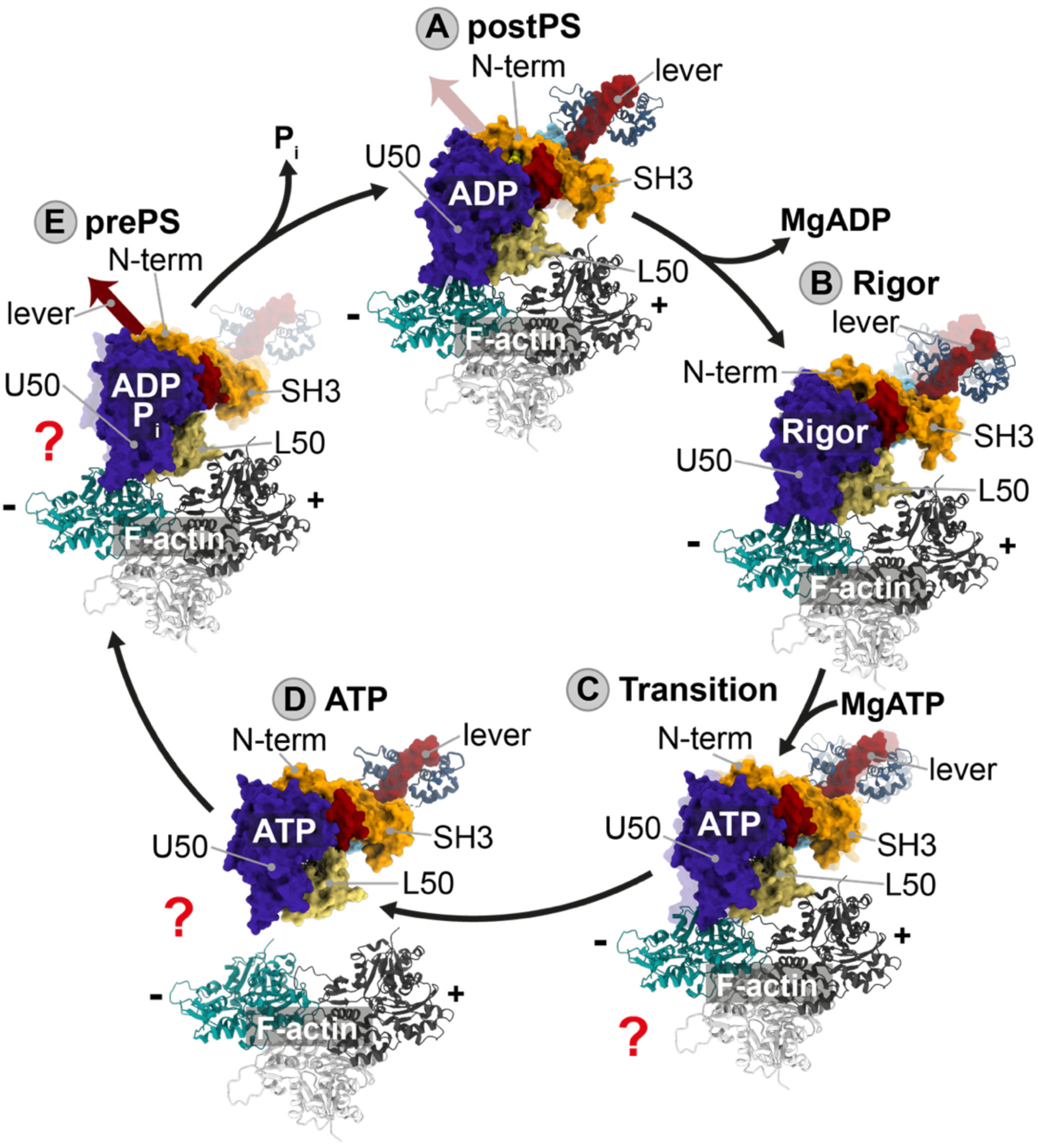


Figure 1.3: Conformations associated with ATPase cycle (front view)
ATPase cycle of Myo5a with the current state of each step in the foreground, overlaying the previous state in the background (faded). The front view highlights lever motion. Question marks indicate unknown states. The motor mobile domains are highlighted, the N-terminal domain is displayed in orange, the U50 in purple and the L50 in yellow. The F-actin subunit closest to the minus end is coloured in teal and the plus end dark grey. All other colours are as in Figure 1.2B. ***A:*** *The post-PS (post-powerstoke) ADP bound actomyosin-5a structure in the foreground (PDB: 7PM6) (Pospich et al., 2021). The lever points towards the plus end of F-actin.* ***B:*** *The rigor actomyosin-5a structure in the foreground (PDB: 7PLU) (Pospich et al., 2021). There is a small change in lever position towards the plus end of F-actin from the ADP state (9 °).* ***C:*** *Post-rigor Myo5a structure with ATP analogue ADP-BeF$_3$ bound (PDB: 1W7J) docked on actin (PDB: 7PLU) via the L50 domain* (Coureux et al., 2003). ***D:*** *Post-rigor Myo5a structure with ATP analogue ADP-BeF$_3$ bound off actin (PDB: 1W7J)* (Coureux et al., 2003). ***E:*** *Pre-PS structure with ADP-P$_i$ analogue ADP-VO$_4$ bound dock on actin (PDB: 4ZG4)* (Wulf et al., 2016). *The recovery stroke into the pre-PS state occurs before ATP can be hydrolysed (Geeves and Holmes, 1999).*

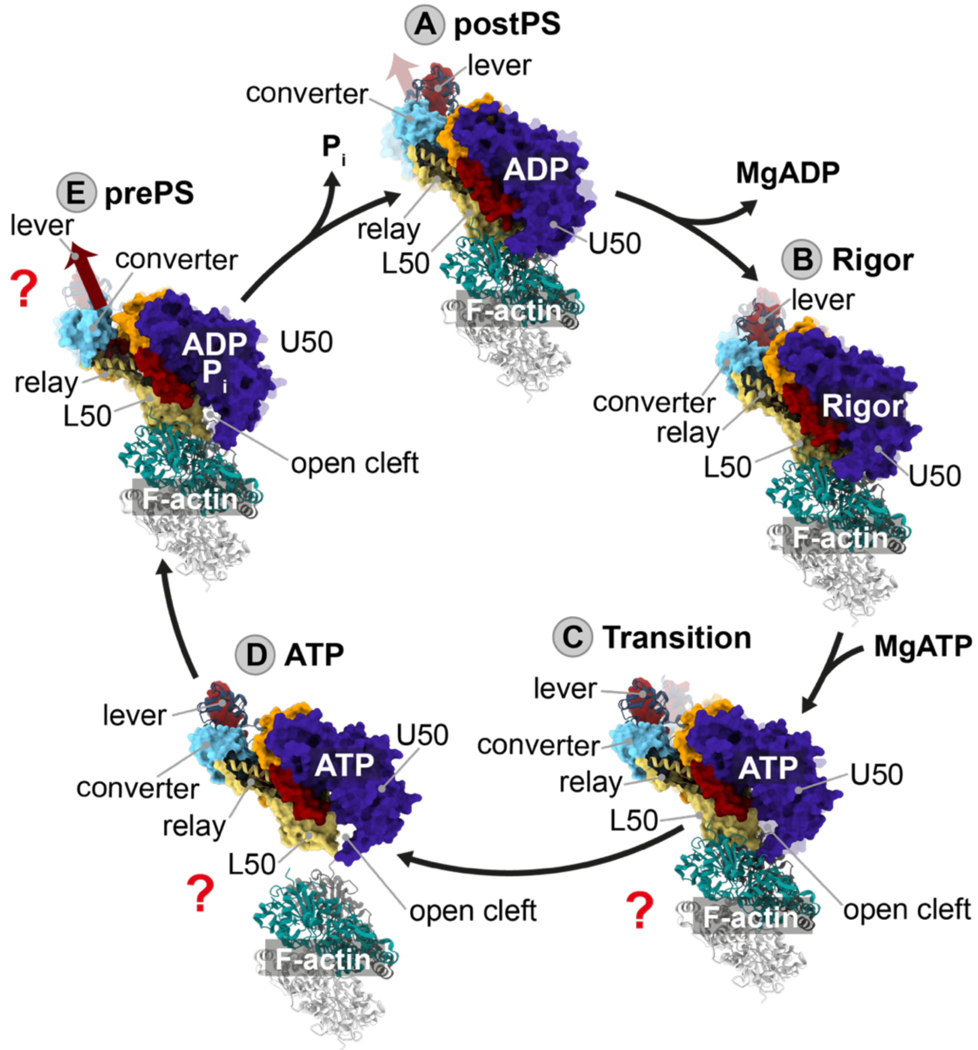


Figure 1.4: Conformations associated with ATPase cycle (side view)
ATPase cycle of Myo5a with the current state of each step in the foreground, overlaying the previous state in the background (faded). The side view highlights actin binding cleft opening and closing. Question marks indicate unknown states. The motor mobile domains are highlighted, the N-terminal domain is displayed in orange, the U50 in purple and the L50 in yellow. The F-actin subunit closest to the minus end is coloured in teal and the plus end dark grey. All other colours are as in Figure 1.2B. ***A:*** *The post-PS (post-powerstoke) ADP bound actomyosin-5a structure in the foreground (PDB: 7PM6) (Pospich et al., 2021). The relay helix is straight and the actin binding cleft is closed.* ***B:*** *The rigor actomyosin-5a structure in the foreground (PDB: 7PLU) (Pospich et al., 2021). The relay helix is straight and the actin binding cleft is closed.* ***C:*** *Post-rigor Myo5a structure with ATP analogue ADP-BeF$_3$ bound (PDB: 1W7J) docked on actin (PDB: 7PLU) via the L50 domain* (Coureux et al., 2003)*. The relay helix is straight and the actin binding cleft is open.* ***D:*** *Post-rigor Myo5a structure with ATP analogue ADP-BeF$_3$ bound off actin (PDB: 1W7J)* (Coureux et al., 2003)*. The relay helix is straight and the actin binding cleft is open.* ***E:*** *Pre-PS structure with ADP-P$_i$ analogue ADP-VO$_4$ bound dock on actin (PDB: 4ZG4)* (Wulf et al., 2016)*. The recovery stroke into the pre-PS state occurs before ATP can be hydrolysed (Geeves and Holmes, 1999). The relay helix is bent and the actin binding cleft is open.*

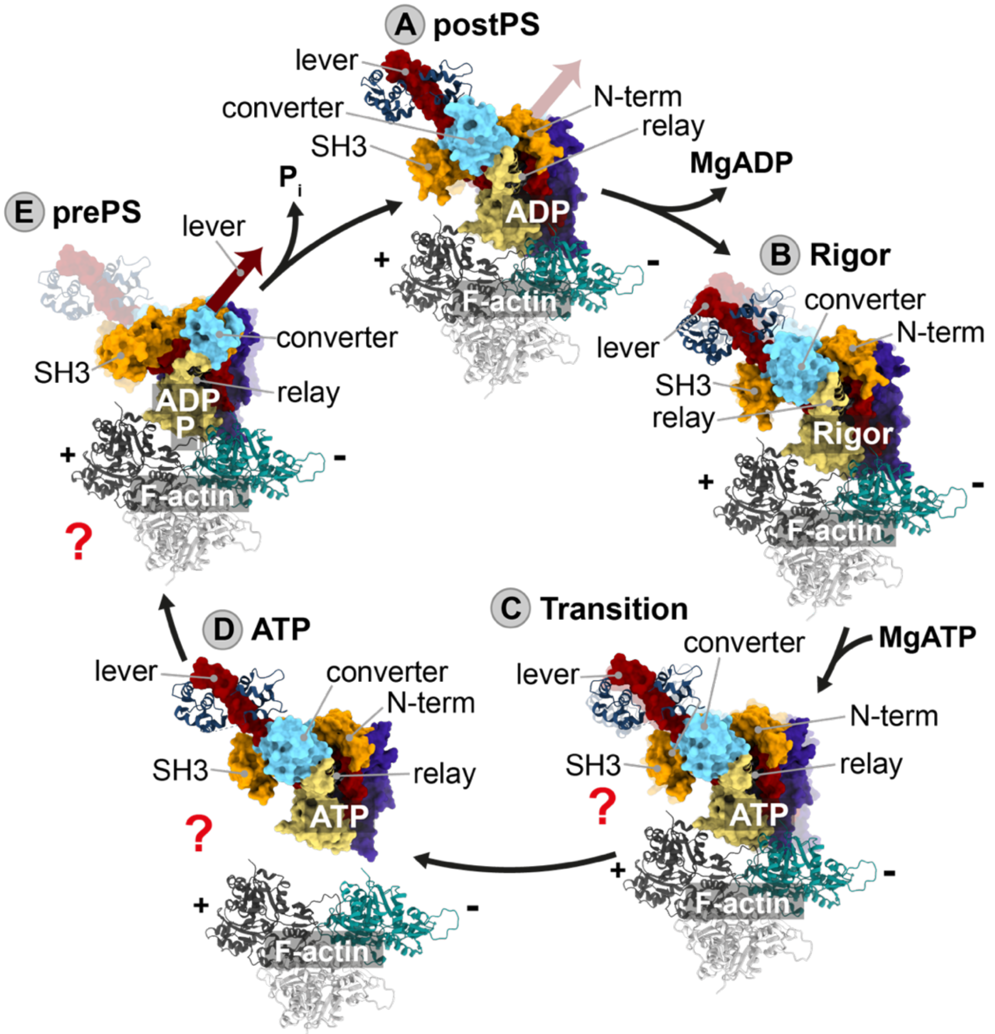


Figure 1.5: Conformations associated with ATPase cycle (back view)
ATPase cycle of Myo5a with the current state of each step in the foreground, overlaying the previous state in the background (faded). The back view highlights the change in the converter. Question marks indicate unknown states. The motor mobile domains are highlighted, the N-terminal domain is displayed in orange, the U50 in purple and the L50 in yellow. The F-actin subunit closest to the minus end is coloured in teal and the plus end dark grey. All other colours are as in Figure 1.2B. ***A:*** *The post-PS (post-powerstoke) ADP bound actomyosin-5a structure in the foreground (PDB: 7PM6) (Pospich et al., 2021). The lever points towards the plus end of F-actin. The relay helix is straight.* ***B:*** *The rigor actomyosin-5a structure in the foreground (PDB: 7PLU) (Pospich et al., 2021). There is a small change in lever position towards the plus end of F-actin from the ADP state (9 °). The relay helix is straight.* ***C:*** *Post-rigor Myo5a structure with ATP analogue ADP-BeF$_3$ bound (PDB: 1W7J) docked on actin (PDB: 7PLU) via the L50 domain* (Coureux et al., 2003)*.* ***D:*** *Post-rigor Myo5a structure with ATP analogue ADP-BeF$_3$ bound off actin (PDB: 1W7J)* (Coureux et al., 2003). *The relay helix is straight.* ***E:*** *Pre-PS structure with ADP-P$_i$ analogue ADP-VO$_4$ bound dock on actin (PDB: 4ZG4)* (Wulf et al., 2016)*. The recovery stroke into the pre-PS state occurs before ATP can be hydrolysed (Geeves and Holmes, 1999). The relay helix is bent.*

1.1.2 Myo5a walking

Myo5a has evolved to walk along an F-actin track. F-actin has a double-stranded, right-handed helix with a 166° helical twist, and assumes a 13/6 helical symmetry (Hanson, 1967). Typically, both Myo5a heads bind on the same face of the actin filament 13 subunits apart, which is equal to 36 nm (Figure 1.6.1) (Walker et al., 2000; Burgess et al., 2002; Oke et al., 2010). This means the step size of Myo5a matches the F-actin pseudo repeat allowing the molecule to walk in a straight line along the F-actin longitudinal axis (Mehta et al., 1999; Walker et al., 2000; Veigel et al., 2002; Burgess et al., 2002; Uemura et al., 2004; Veigel et al., 2005; Sakamoto et al., 2005; Kodera et al., 2010; Andrecka et al., 2015). However, some exceptions are seen with heads binding 11 or 15 subunits apart as opposed to the common 13, which may instead lead to a spiralling path of Myo5a around the actin filament (Walker et al., 2000; Ali et al., 2002; Beausang et al., 2008; Oke et al., 2010; Lombardo et al., 2017).

Actin filaments are polar with a distinct barbed (plus) and pointed (minus) end. Each head of dimeric Myo5a binds F-actin in the same stereospecific manner, which means the dimer is also polar (Walker et al., 2000; Burgess et al., 2002; Pospich et al., 2021). This enables a directional strain to be built across the molecule when both heads are bound. In the strong actin binding ADP state, the lead lever as a monomer would be in the post-PS position, however in the dimer it is prevented from completing the powerstroke by the actin bound trail head. Intramolecular strain is generated as the lead head pulls the trail head towards the plus end of F-actin, and is pulled back towards the minus end by the trail head (Figure 1.6.1). This strain further polarises Myo5a, both in terms of binding strength and kinetics (Veigel et al., 2002; Forgacs et al., 2006; Oguchi et al., 2008). The lead head imposes load in the direction of the powerstroke on the trail head, which has been shown to cause heads to detach from F-actin more easily. Whereas the trail head imposes resistive load on the lead head, which has been shown to cause heads to require greater forces to detach from F-actin (Oguchi et al., 2008). Direction of load also influences ADP release. Intramolecular strain inhibits ADP release from the lead head by ~250-fold compared to the trail head (a process known as gating) (Veigel et al., 2002; Forgacs et al., 2008). This ensures the trail head detaches before the lead head biasing for forward walking. These characteristics have all evolved to bias Myo5a to walk along an F-actin track towards the plus end.

Gating gives Myo5a direction, but the combination of being a dimer and a high duty ratio gives Myo5a processivity. This means that when one head detaches, the other head remains attached for a sufficient length of time that the detached head can re-attach. As a result, the motor can take multiple steps along its track without detaching completely (Sakamoto et al., 2000; Rief et al., 2000). Myo5a walks in a hand-over-hand compass-like motion. As it steps the trailing head swings out off-axis and around the old leading head to become the new leading head. Once started, this off-axis stepping always occurs along the same side of F-actin, with a slight preference for right-handed walking over left (Andrecka et al., 2015). The ability to walk along either side indicates that whatever the angle of approach between the head and F-actin, the molecule generates directional strain towards the plus end of F-actin. This is important to consider when forming a hypothesis for intramolecular strain generation and walking mechanics. Additionally, this compass-like motion has led to the idea that there is a constrained angle between the two levers at the head-tail junction. However, this needs to be further explored in 3D (3-dimensions) as in negative stain it is unclear what angle the molecules are being viewed from (Walker et al., 2000; Takagi et al., 2014; Andrecka et al., 2015; Hathcock et al., 2020).

Walking of the dimer along F-actin involves complex choreography of the ATPase cycles within the two heads. When both heads are actin bound the trail head is in the post-PS position with ADP bound (Figure 1.6.1). The lead head is ADP bound, but the lever is held under strain in the pre-PS position by the trail head. Priming of the lead lever through intramolecular strain is a way of storing the chemical energy produced by ATP hydrolysis as elastic potential energy. The rate of ADP release in the lead head is reduced, and in the trail head is accelerated, when both heads are bound (Veigel et al., 2002; Rosenfeld and Sweeney, 2004; Forgacs et al., 2008). This means ATP is more likely to bind the trail head first, causing it to dissociate from actin (Figure 1.6.3). The strain on the lead lever is released and it then completes the powerstroke. It has been proposed that the energy for the lever swing is from generation of tension and not chemical input (De La Cruz et al., 1999). Optical trap experiments have shown that in the dimer the lever swing accounts for ~25 nm of the working stroke, and the further 11 nm is thought to be a Brownian search for the next binding site, resulting in a 36 nm overall step size (Veigel et al., 2002). The transition time from unbinding of the trail head to rebinding as the lead head is approximately 18 ms ((Figure 1.6.3-4) (Andrecka et al., 2015). During this time the transient head undergoes ATP hydrolysis (~1 ms), and is able to search along actin in its low affinity ADP-P_i state (Figure 1.6.4). When the appropriate site is found, P_i is quickly released (~4 ms) and the head transitions into the high affinity ADP bound state (Figure 1.6.1) (De La Cruz et al., 1999). What was the trail head is now a bound lead head under strain, and the cycle starts again.

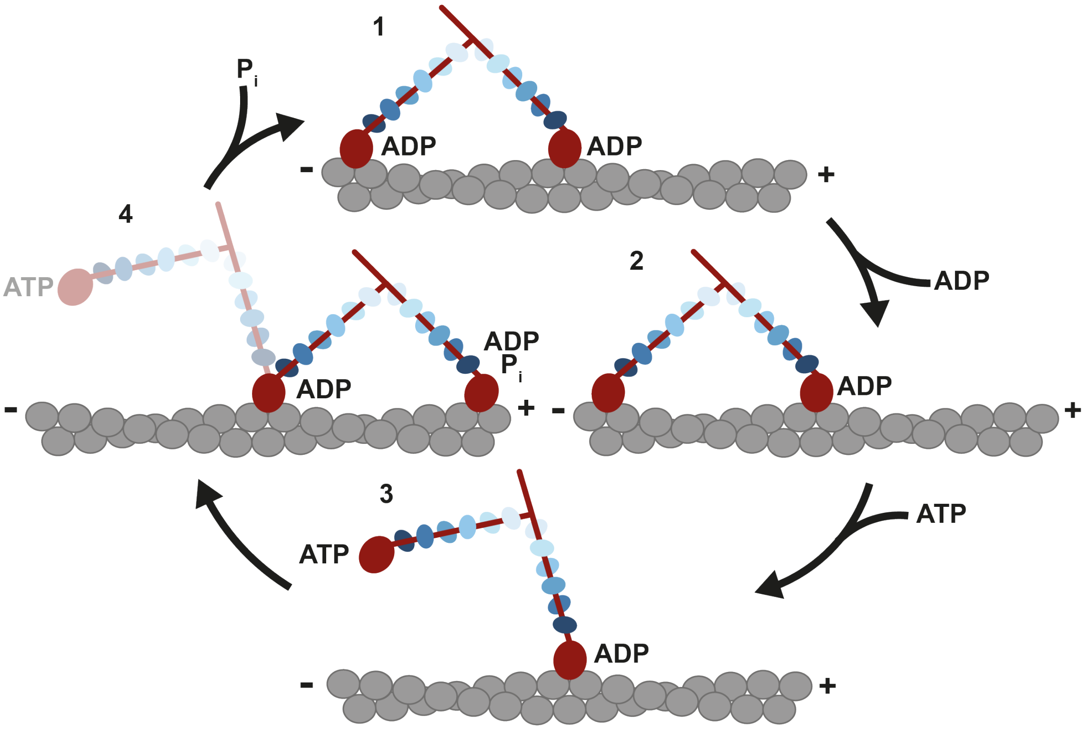


Figure 1.6: Schematic of ATPase cycle in Myo5a as it walks along F-actin
1: *both heads are bound to F-actin and have ADP in their active sites. The trail head (proximal to the minus end) is in the post-powerstroke (post-PS) conformation (Figure 1.3A, Figure 1.4A & Figure 1.5A). The lead head (proximal to the plus end) is in the pre-powerstroke (pre-PS) conformation, as it is being restrained by the trail head* (Walker et al., 2000; Burgess et al., 2004)*. Intramolecular strain generated by this restraint accelerates ADP release from the trail head by ~250-fold, ensuring ADP is preferentially released from the trail head* (Veigel et al., 2002; Forgacs et al., 2008)*.* ***2:*** *the trail head is in the rigor conformation with no nucleotide bound (Figure 1.3B, Figure 1.4B & Figure 1.5B), and the strain on the lead head is maintained.* ***3:*** *ATP binds the trail head causing it to detach from actin (Figure 1.3C-D, Figure 1.4C-D & Figure 1.5C-D) and releasing the strain on the lead head. The lead head undergoes the powerstoke.* ***4:*** *The old lead head becomes the new trail head in the post-PS conformation (Figure 1.3A, Figure 1.4A & Figure 1.5A). ATP in the transient head is hydrolysed to ADP-P_i (~1 ms)* (De La Cruz et al., 1999)*. The ADP-P_i state has weak affinity for actin and is able to bind the proximal binding site becoming the new lead head (Figure 1.3E, Figure 1.4E & Figure 1.5E). The transition from state* ***3-4*** *takes ~18 ms* (Andrecka et al., 2015)*.* ***4-1:*** *P_i is released from the new lead head (~4 ms) generating the strained ADP state* (De La Cruz et al., 1999)*. The cycle starts again. Colours are as in Figure 1.2B.*

1.2 Cryo-EM

Since 2012, with the advent of improved electron detectors and enhanced computational image processing power, the number of high-resolution structures (≤4 Å) resolved via cryo-EM has increased ~1000-fold (Scheres, 2012; Kühlbrandt, 2014a; Kühlbrandt, 2014b; Cheng, 2015; Fernandez-Leiro and Scheres, 2016). One of many advantages of using cryo-EM is that it enables visualisation of proteins in an aqueous environment, and therefore potentially closer to their native state (Taylor and Glaeser, 1974). Additionally, it is possible to capture different dynamic states as molecules are able to move freely up to the point of freezing the specimen (Clare et al., 2012; Murata and Wolf, 2018; Dong et al., 2019; Zhong et al., 2021).

Myosin and F-actin are too large for crystallisation (Robert-Paganin et al., 2020). Therefore, it wasn’t until 2016, after the cryo-EM “resolution revolution”, that a high-resolution structure of actomyosin was published (Von Der Ecken et al., 2016). Since then, 32 high-resolution structures of actomyosin have been deposited in the EMDB (Electron Microscopy Data Bank) (McMillan and Scarff, 2022).

Atomistic molecular dynamics (MD) simulations are increasingly being used to fit models into cryo-EM density maps (Croll, 2018; Kidmose et al., 2019; Scarff et al., 2020). In MD simulations, atoms are represented as classical balls connected by harmonic springs (covalent bonds), whilst non-bonded atoms interact through Van der Waals forces and electrostatics. The system acts under a force-field, which is a set of parameters based on experimental data and quantum mechanics. Newtonian mechanics is used to calculate the change in atom position at each timestep, as each one responds to the forcefield and surrounding atoms (McCammon et al., 1977). Simulations can be used to flexibly fit a model of one conformation into the map of another, by providing a chemically and physically possible path from one state to the other (Trabuco et al., 2009; Croll, 2018; Kidmose et al., 2019). They can also be used to assign chemically and physically relevant side-chain positions where cryo-EM resolution isn’t sufficient (Scarff et al., 2018). Myosins are highly dynamic which makes high-resolution reconstructions of flexible domains challenging. Therefore, using MD is likely to become increasingly relevant for building models of this class of protein as large flexible domains become more accessible with cryo-EM.

1.2 Computational modelling

Computational modelling is an invaluable tool in the pursuit of uncovering protein dynamics, where experimental tools alone are unable to provide the spatial and temporal resolution necessary to capture the features of interest (Figure 1.7). Proteins are not static objects and fluctuation in their conformational state, due to thermal noise, can dictate their function. Cytoskeletal motor proteins in particular depend on large-scale transitions driven by active processes, on timescales we struggle to probe experimentally (Gravett et al., 2022). Models provide a means of integrating experimental data, and developing *in silico* model systems permits the visualisation and prediction of conformational transitions, trialling of hypotheses, and can inform experimental approach (Gray et al., 2015). Myo5a has often been the subject of computational modelling and a variety of mesoscale (10 nm to 500 nm, and 1 μs to 1 s) modelling approaches have contributed to our understanding of this complex system (Kolomeisky and Fisher, 2003; Vilfan, 2005; Cecchini et al., 2008; Craig and Linke, 2009; Zheng, 2011; Hinczewski et al., 2013; Mukherjee et al., 2017; Hathcock et al., 2020).

Fluctuating finite element analysis (FFEA) is a tool designed for simulating the dynamics of complex biomolecular systems across the mesoscale, and has been used previously to model motor and track systems (Hanson et al., 2021). Its unique ability to explore longer spatiotemporal scales than traditional atomistic MD simulation techniques, means it is ideal for modelling proteins with large-scale conformational changes intrinsic to their function, such as motor proteins. It is a low-resolution, non-atomistic, coarse-grained approach, that represents globular biomolecules as viscoelastic continuum solids, using 3D tetrahedral finite element meshes (Oliver et al., 2013; Solernou et al., 2018), and elongated biomolecules as coarsened 1-dimensional (1D) Kirchhoff rods (Welch et al., 2020). It simulates the dynamics of a system in response to thermal noise, which is dependent on the shape of the molecule, and the user defined elastic moduli and viscosities. Trajectories are calculated by solving a set of Langevin equations that describe the motion (Oliver et al., 2013; Solernou et al., 2018).

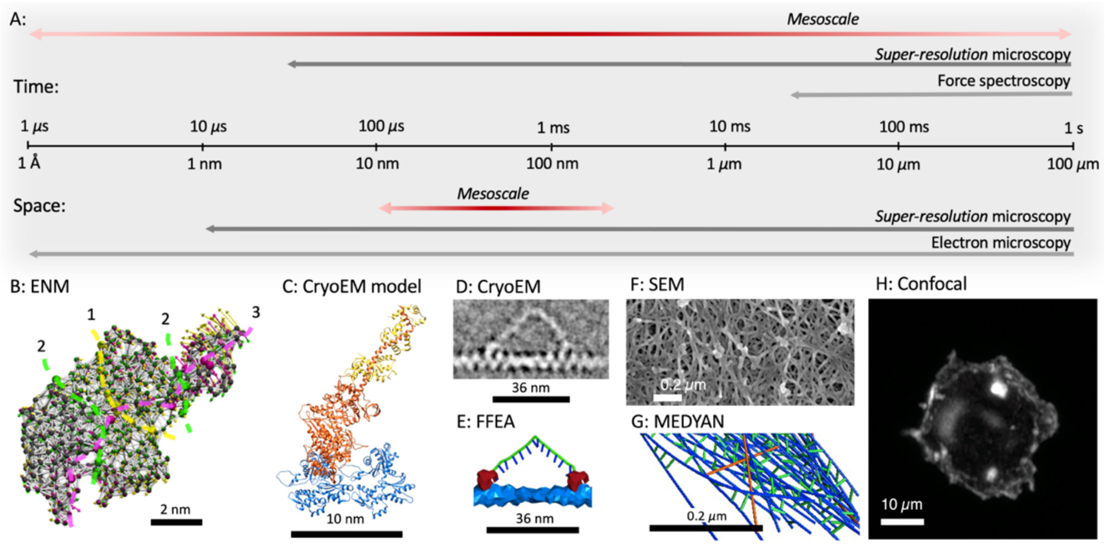


Figure 1.7: An indication of scale (taken from (Gravett et al., 2022))
A: *Scale describing the temporal and spatial resolution of experimental techniques and the "biological mesoscale" (10 nm-500 nm, 1 µs-1 s). Highest temporal resolution of super-resolution microscopy = 50 µs (interferometric light scattering microscopy* (Taylor and Sandoghdar, 2019)*), and force spectroscopy = 67 ms (high speed atomic force microscopy* (Ando, 2019)*). Highest spatial resolution of super-resolution microscopy = 1 nm (single molecular light microscopy* (Sahl et al., 2017)*), and electron microscopy (EM) = 1.2 Å (cryo-EM* (Nakane et al., 2020)*). Arrowless heads indicate the resolution range goes beyond the scale displayed.* ***B:*** *"Bead and spring" elastic network model (ENM) of Myo5a motor domain (PDB: 1OE9* (Coureux et al., 2003)*) using ProDy* (Bakan et al., 2011; Bakan et al., 2014)*. Arrows show anisotropic network model modes of motion. Dotted lines indicate division of motions as domains and larger beads show "hinge" regions motions predicted from gaussian network model. Yellow = mode 1, green = mode 2, pink = mode 3.* ***C:*** *Cryo-EM fitted model of rabbit skeletal muscle actomyosin rigor complex (PDB: 5H53* (Fujii and Namba, 2017)*). Blue = 2 subunits of actin, orange = myosin and yellow = myosin light chain.* ***D:*** *Cryo-EM image of dimeric Myo5a walking along actin, kindly donated by Kavitha Thirumurugan.* ***E:*** *Fluctuating finite element analysis (FFEA) combined mesh (red) and rod (green, blue lines help display twist) model of dimeric Myo5a on F-actin (blue).* ***F:*** *Scanning electron micrograph (SEM) of F-actin in a cell.* ***G:*** *MEchanochemical DYnamics of Active Networks (MEDYAN) model of actin filaments bundling after 2000 s due to contractile forces from motors and cross-linkers. Simulation box volume is 1x1x1 µm^3. Blue = actin filament, orange = myosin 2 mini-filaments, green = α-actinin cross-linker. Data displayed from* (Popov et al., 2016)*.* ***H:*** *Confocal image of F-actin in a cell.*

1.3 Aims/intentions

Whilst high-resolution structures of components of Myo5a have enabled an in-depth understanding of local structure-function relationships (Table 1.1), we lack a full understanding of how the individual components contribute to the properties of the entire molecule. Despite the lever domain being essential for Myo5a to walk, there is still no high-resolution structure of the full-length lever in the active state, nor do we have an in-depth understanding of its physical properties. Further understanding of this region is required for understanding Myo5a walking mechanics.

Characterising flexibility within the lever domain is critical for understanding how conformational changes in the motor domain are translated across the molecule, and how strain is generated. Optical trapping and ns-EM have previously been used to characterise lever flexibility (Veigel et al., 2002; Moore et al., 2004; Oke et al., 2010). However, these approaches are not without their disadvantages. In previous optical trap assays, the stiffness of Myo5a was measured as a change in stiffness of F-actin suspended between 2 optically-trapped beads, when Myo5a bound to a third surface-bound bead binds the F-actin (Veigel et al., 2002). This is an indirect way of measuring the properties of the lever. It is therefore also unclear where along myosin the third bead is attached, and the influence on flexibility that the bead attachment has. The connection stiffness of the F-actin bead attachment has however been explored (Dupuis et al., 1997; Mehta et al., 1997; Veigel et al., 1998). When using ns-EM micrographs to calculate flexibility (Oke et al., 2010), the particles are a projection into 2D space. In the presence of F-actin, F-actin acts as a reference by which it is possible to calculate stiffness of the myosin molecule as a function of the distance between the heads bound to F-actin. However, in 2D without F-actin as a reference, it is impossible to determine whether the emergent angles and distances are true, or simply appear acute or obtuse due to the angle of viewing. Additionally, without 3D reconstruction it is impossible to determine torsional stiffness from ns-EM. Finally, neither approach captures sufficient detail to observe intra-lever domain flexibility. Therefore, direct measurements of lever properties in 3D are required to improve the accuracy of estimates and help determine structure-function relationships.

There is a wealth of literature contributing to our understanding of Myo5a walking dynamics (Mehta et al., 1999; Walker et al., 2000; Veigel et al., 2002; Burgess et al., 2002; Uemura et al., 2004; Veigel et al., 2005; Sakamoto et al., 2005; Kodera et al., 2010; Andrecka et al., 2015). However, techniques that offer the resolution required to capture ms timescale but nm resolution are limited (Gravett et al., 2022). This means thus far it has been impossible to directly visualise (i.e. without a label) transitional dynamics using experimental techniques (~1 ms ATP hydrolysis, ~4 ms P_i release, ~80 ms ADP release, ~18 ms step) (De La Cruz et al., 1999; Andrecka et al., 2015). Visualising transitional dynamics is important for understanding what pathways exist between states, and what modes of motion are favoured, all of which help to understand how strain can be generated, e.g. if it is torsional or flexural, and how stepping occurs. Computational modelling and simulation techniques can bridge these gaps.

Therefore, questions this thesis seeks to answer are:

1. What is the structure of the full-length lever domain in the active state?
2. What is the flexibility of the lever domain?
3. Can a representative coarse-grained model be generated in FFEA? How do we know our model is reliable?
4. Can the strained state be simulated in FFEA? What is the strained state and how is strain generated?

More broadly, I seek to use a combination of cryo-EM and computational modelling techniques in order to develop methods of probing motor protein mechanics, by directly looking at structure-property-function relationships.

Chapter 2 Structure of the full-length lever domain of rigor actomyosin-5a

2.1 Introduction

The long lever of Myo5a is important for its stepping behaviour and yet we currently lack a high-resolution structure of the Myo5a motor domain bound to actin with a full-length lever. High resolution crystal structures of short segments of the lever include IQ1-2 (23 and 25 residues) of murine Myo5a bound to CaM (Houdusse et al., 2006) and IQ2-3 (25 and 23 residues) of yeast myosin 5 (Myo2p) bound to myosin light chain 1 (Mlc1p) (Terrak et al., 2005) (PDB: 2IX7 and 1N2D, respectively). A recently published high-resolution cryo-EM structure shows the structure of the full-length lever in the shutdown state (Niu et al., 2022). However, only low-resolution structures the full-length Myo5a lever in the active state are available (Walker et al., 2000; Oke et al., 2010). In contrast, the structure of the Myo5a motor domain bound to actin was recently solved to high resolution (approx. 3 Å) for constructs containing the motor domain and the first IQ bound to ELC in both rigor (nucleotide free) and ADP states (Pospich et al., 2021).

To better understand how Myo5a walks along F-actin, and how the lever contributes to its walking mechanics, I have obtained a high-resolution cryo-EM structure of murine Myo5a comprising the motor domain and a 6IQ lever (Myo5a-S1), in the rigor state, bound to F-actin. Each IQ motif in the lever is occupied by a CaM. Using MD simulations, I have generated a pseudoatomic model of the lever, predicted contacts along the lever, and demonstrated increased flexibility of the CaM bound to IQ6 (CaM6). I have also resolved 9 distinct lever conformations, which allowed me to directly assess lever stiffness, and calculate intra-lever stiffnesses, for the first time. Additionally, this collection of conformations enabled me to determine how the inherent flexibility of the lever contributes to Myo5a walking.

2.2 Materials and Methods

2.2.1 Sample preparation

Murine Myo5a-S1-6IQ (1-907) and CaM proteins were purified as described in (Wang et al., 2000; Forgacs et al., 2006), and stored in liquid nitrogen (LN_2). Both were kindly provided by Howard White. After thawing stored Myo5a-S1-6IQ, additional CaM was added in a 2:1 ratio to ensure all IQ motifs were fully occupied. Before grid preparation rabbit skeletal G-actin was polymerised on ice, by first exchanging Ca^{2+} for Mg^{2+} in exchange buffer (final concentration: 85 μM actin, 1 mM EGTA, 0.27 mM $MgCl_2$) followed by polymerisation in polymerisation buffer (final concentration: 78 μM actin, 25 mM KCl, 1.0 mM $MgCl_2$, 1.0 mM EGTA, 10 mM MOPS, pH 7.0).

2.2.2 Grid preparation and cryo-EM data acquisition

Quantifoil R2/2 carbon Cu 300 mesh grids (Agar Scientific, Stansted, UK) were glow discharged in an amylamine vapour for 30 s (GloCube, Quorum Technologies Ltd, Laughton, UK). Directly before application, F-actin was sheared by being repeatedly drawn up and ejected with a gel loading pipette tip to shorten filaments, to increase the amount of F-actin observed occupying grid holes. 1 μL of sheared F-actin (0.5 μM) was applied to the grid and incubated for 2 mins. Next, 3 μL of Myo5a-S1-6IQ together with CaM (3 μM + 6 μM, respectively) was added to the grid, followed by a second incubation of 2 mins. Final concentrations of proteins were: 0.125 μM actin, 2.5 μM Myo5a-S1, 5 μM CaM. Final solution conditions were: 100 mM KCl, 0.1 mM EGTA, 1 mM $MgCl_2$, 10 mM MOPS, pH 7.0. Grids were then blotted with Whatman no. 42 Ashless filter paper (Agar Scientific, Stansted, UK) for 3 s at force -25, 8 °C and 80 % humidity, drained for 0.5 s and flash-frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher, Altrincham, UK). Data was recorded on the FEI Titan Krios I (Astbury Biostructure Laboratory, University of Leeds, Leeds, UK) equipped with a FEI Falcon III detector operating in linear mode. Data collection parameters can be found in Table 2.1.

Table 2.1 - Microscope parameters

Microscope	Titan Krios I
Magnification	75000
Voltage (kV)	300
Electron dose per image ($e^{-}/Å^{2}$)	62.66
Exposure time (s)	1.5
Number of fractions	59
Defocus range (µm)	-1.8 to -3.6 (0.3 steps)
Pixel size (Å)	1.065

2.2.3 Cryo-EM image processing

First, MotionCor2 (Zheng et al., 2017) was used to correct for beam-induced motion, and the contrast transfer function (CTF) was estimated using Gctf (GPU accelerated CTF) (Zhang, 2016), before subsequent processing steps (Figure 2.1). The start-end coordinates of F-actin filaments were manually picked using RELION 3.1 (Zivanov et al., 2018). Particles were extracted in RELION 3.1 using helical parameters (box size 608 px, helical rise 27.5 Å, tube diameter 250 Å). The data was initially binned 2 x to 2.13 Å. Helical 3D refinement was used to produce an initial model (tube outer diameter 140 Å, angular search range tilt 15° and psi 10°, initial twist -166.15°, helical rise 27.5 Å, twist search -165° to -167° with a 0.1° step, rise search 26.5° to 28.5° with a 0.1° step). A known structure of the Myo5a motor (PDB: 7PLU, (Pospich et al., 2021)) was rigid fit into the helical refinement map at a motor domain that was the most well resolved and positioned at the centre of the box, using Chimera (Pettersen et al., 2004). A map of the fitted PDB structure was generated in Chimera (Pettersen et al., 2004), and used as input to generate a wide mask of the motor domain in RELION 3.1 for masked 3D classification (Zivanov et al., 2018). Masked 3D classification was performed in RELION 3.1 to classify out undecorated actin and to only include particles with myosin bound at the centre of the box (5 classes, no image alignment, regularisation parameter T = 4). This dataset was then un-binned (1.065 Å) as initial reconstructions were reaching the Nyquist limit. 3D helical refinement followed by masked post-processing of this subset of particles produced a map with a 3.8 Å global resolution, but with limited detail across the lever. All global resolutions were determined using the gold standard Fourier Shell Correlation (FSC) reported to FSC = 0.143 using RELION 3.1. To compare the motor domain to the previously published chicken actomyosin-5a rigor structure (PDB: 7PLU (Pospich et al., 2021)), particle subtraction was performed subtracting all density outside of a mask comprising of the motor, the first 2 CaMs and 3 F-actin subunits (2IQ). 3D refinement followed by post-processing, produced a map with a global resolution of 4.2 Å according to the $FSC_{0.143}$-criterion, which was highly similar to that previously published (3.2 Å, PDB: 7PLU (Pospich et al., 2021)). The 2IQ map was locally sharpened using DeepEMhancer (Sanchez-Garcia et al., 2021).

To improve lever resolution, particle subtraction was required to aid particle alignment and centralise the lever domain within the box. It was noted from initial 3D helical reconstructions that there was smearing of the lever density, so a wide mask containing 1 actin subunit, the motor and a cone shape for the lever to accommodate flexibility was used for subtraction (Figure 2.1). A map was generated in Chimera comprising of a single actin subunit, a motor (both from PDB: 7PLU, (Pospich et al., 2021)), and multiple copies of a Myo5a lever model (PDB: 2DFS (Liu et al., 2006b)) arranged in a cone shape tapering at the motor and splaying towards the lever end. These were positioned so that the boundaries of the cone met the density of the neighbouring actin bound heads in the helical map. A wide cone-shaped mask was generated in RELION 3.1 using the cone-shaped map. Signal outside of this mask was subtracted from the binned (to 2.13 Å) 2D images, and particles were re-centred bringing the lever to the focal point of the box. 3D refinement of the subtracted particles produced a map (cone subtracted map) with 4.3 Å global resolution according to the $FSC_{0.143}$-criterion. This map was locally sharpened using DeepEMhancer (Sanchez-Garcia et al., 2021). The cone subtracted map had improved resolution across the lever, however defined density for CaMs 5 and 6 could not be seen (Figure 2.1).

In order to resolve CaMs 5 and 6, 3D classification of the lever domain using a cone shaped lever mask was performed. This mask was generated in the same was as previously, but excluding the actin subunit and motor domain in order to classify based on lever conformation only. Classification into 10 or 20 classes both produced ~9 reasonable classes. However, the conformations reconstructed by 3D classification were not distinct conformational states but continuous thermally driven motion, as the conformation of the classes produced differed in both modes of division (10 or 20). The significance of 9 classes is instead likely to be the number of groups the particles could be reasonably divided into to have sufficient particles to produce a reconstruction. The maps were locally sharpened using DeepEMhancer (Sanchez-Garcia et al., 2021). Though classification led to a reduction in overall resolution due to loss of particles in the reconstruction, for the first time we are able to see across the full length of the lever while the motor is actin bound (Figure 2.1).

The cryo-EM 3D class with the best global resolution (7.5 Å) was selected to build a pseudoatomic lever model. The local resolution of this map was calculated using SPOC (statistical post-processing of cryo-EM maps), as the local resolution calculations in RELION are unreliable at resolutions lower than 10 Å (Beckers and Sachse, 2020).

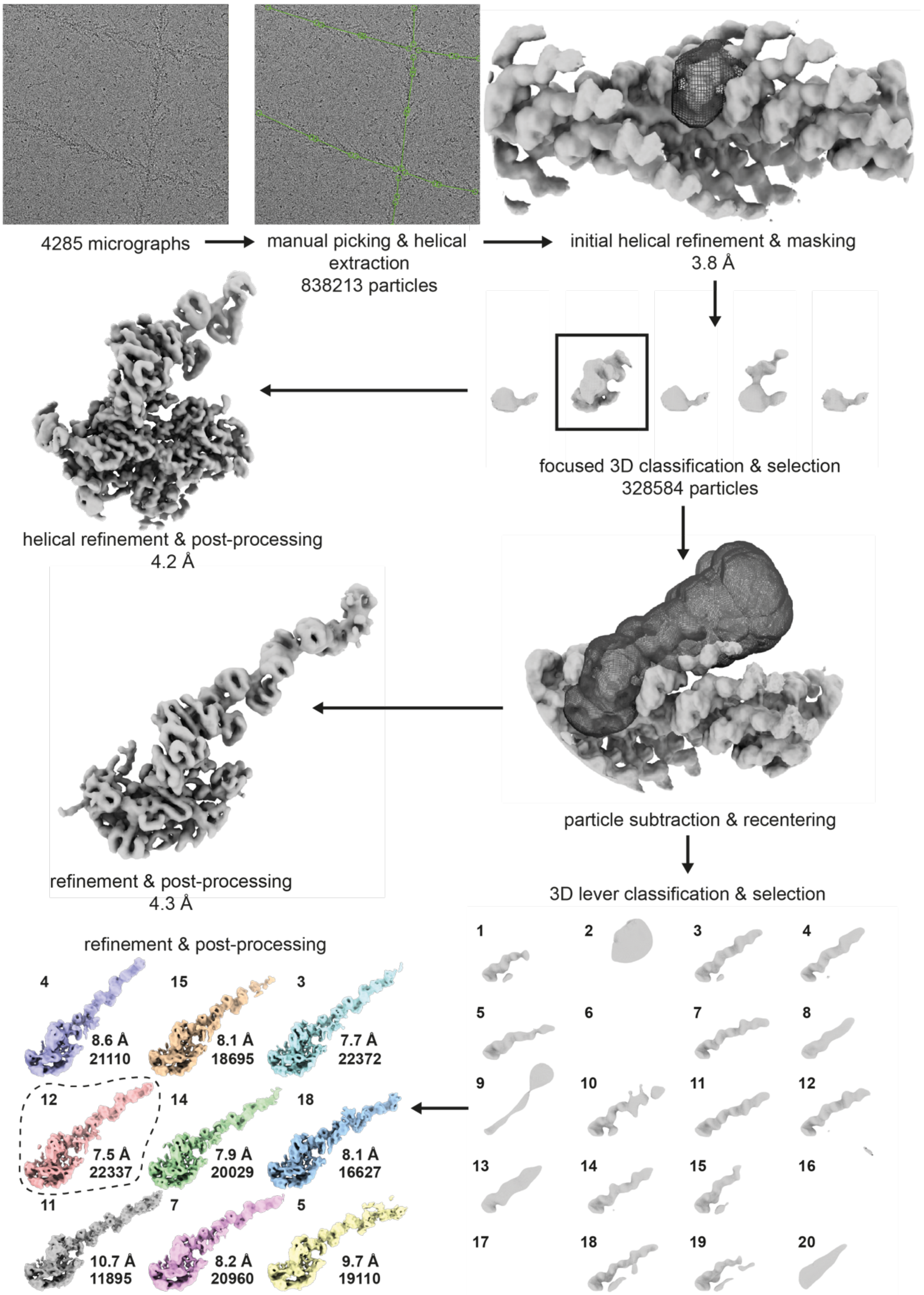


Figure 2.1: **Processing workflow**

cryo-EM data processing pipeline for Myo5a-S1-6IQ. Showing the number of micrographs used, the particle numbers, corresponding density maps, and masks for particle subtraction. All resolutions quoted are based on global resolution at 0.143 FSC (Fourier Shell Correlation). All resultant density maps displayed were post-processed in DeepEMhancer (Sanchez-Garcia et al., 2021).

2.2.4 Cryo-EM model building and refinement

An initial lever model including the converter + 6IQ motifs sequence (698-907) + 8 residue FLAG-tag, and 6 CaM sequences, was built in Alphafold 2.0 using collabfold Google collab notebooks (Mirdita et al., 2022). This model was then flexibly fit into the density of the best 3D class using Isolde (Croll, 2018), applying distance and torsional restraints taken from murine IQ1-2 and CaM1-CaM2 structure (PDB: 2IX7, (Houdusse et al., 2006)) to each CaM pair. Distance and torsional restraints from the Myo2p 25-residue spaced pair structure (PDB: 1N2D, (Terrak et al., 2005)) were applied to interacting residues in 25-residue spaced pairs (Glu14-Arg91, in CaM2-3 and CaM4-5), as density corresponding to these interactions could be seen in the selected class map. Only the C-lobe of CaM6 was included in fitting as only density for this half of the molecule could be seen in the map. To model the N-lobe of CaM6 a homology model of CaM in an extended state, based on a Myo2p-MLC1P structure (PDB: 1M46, (Terrak et al., 2003)), was built using Swiss-model (Waterhouse et al., 2018). The C-lobe of the homology model was superimposed onto the C-lobe of the initial CaM6 model. The N-lobe (residues 3-84) of the homology model was then joined to the C-lobe of the initial CaM6 model, and minimised in Isolde without map weighting.

A pseudoatomic lever model was generated by performing all-atom MD simulations of the fitted lever model to gain side-chain conformations. All simulations were performed using Amber20 (Case et al., 2020) with the FF19SB forcefield (Tian et al., 2020). The lever model was protonated according to the Amber residue templates and then solvated with TIP3P water molecules in an octahedral box that extended 14 Å from the protein. K^+ ions were added to neutralise the system, then KCl was added to a final concentration of 100 mM. After initial energy minimisation the system was heated to 300 K as positional restraints were decreased from 100 to 0 kcal/mol/$Å^2$, except for across the lever heavy chain. A restraint of 1 kcal/mol/$Å^2$ was applied to the backbone of the lever heavy chain throughout the minimisation, equilibration and production runs, to permit CaM motion and interaction with the heavy chain side-chains, but maintain the lever position seen in the cryo-EM map. Minimisation and equilibration steps were performed on the ARC4 standard nodes. NMR distance restraints were also applied between interactions visible in the cryo-EM density (Glu14-Arg90). To do this, distance restraints were applied between the Cδ atom of Glu, and Cζ of Arg, in order not to dictate which N and O interact. These were weighted at 20 kcal/mol/Å^2 within 1.9 Å of the bounds of the flat well restraint (3.4-5.3 Å), and a 20 kcal/mol/Å^2 harmonic potential was applied outside of this range. The MD production runs used the pmemd.cuda module from Amber20 and were run on Bede using Tesla V100 GPUs. MD was performed for 300 ns in triplicate repeat. The Berendsen Thermostat was used for maximising GPU performance.

Following simulation, the average conformation was calculated in VMD (Visual Molecular Dynamics) (Humphrey et al., 1996), and the frame of the trajectory with the lowest global RMSD (Root Mean Square Deviation) with the average conformation was selected as a model for each repeat. A further round of minimisation was then performed on the ARC4 general nodes to restore side-chains to low energy conformations. After minimisation, each model was scored in MolProbity (Chen et al., 2009), and the repeat with the lowest MolProbity score and the largest proportion within the cryo-EM density was selected as the final lever model.

To compare the flexibility of each CaM, per residue root mean square fluctuation (RMSF) analysis was performed using CPPTRAJ on each CaM in the restrained simulation and for each repeat (Roe and Cheatham, 2013).

Unrestrained MD simulations were run in conjunction with the restrained simulations to look at all possible residue contacts. Unrestrained MD simulations were performed under the same conditions as previously, however without restraints applied. Equation (**2.1**) was used to estimate the order of magnitude of timescale required to capture the thermal fluctuations of the lever in explicit water,

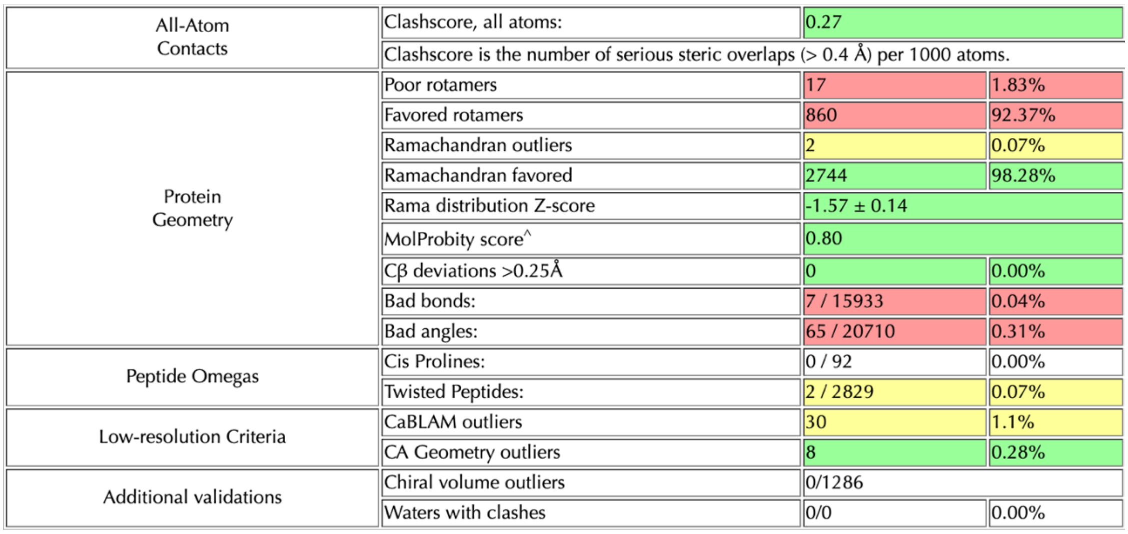
$$\text{timescale of motion} = \frac{\mu L}{\kappa}, \qquad \textbf{(2.1)}$$

where μ = 1 pN nm^{-2} ns (the viscosity of water), L = 20 nm (approximate lever length), $\kappa_{\text{cantilever}}$ = 0.76 pN nm^{-1} (lever cantilever spring constant). This gives a relaxation timescale of ~15 ns. Therefore, a 300 ns simulation should contain

~20 independent observations of the slowest motion. Although ~1000 observations would be required for robust statistical analysis, ~20 independent observations in triplicate repeat should be sufficient to identify prevalent interactions.

To generate a model of the motor domain, a homology model of murine Myo5a was made based on the chicken actomyosin-5a rigor structure (PDB:7PLU, (Pospich et al., 2021)) using Swiss-model (Waterhouse et al., 2018). This was then fit into the cryo-EM density using Isolde (Croll, 2018), applying distance and torsional restraints based on the template. Residues 1-698 where then joined to the pseudoatomic lever model (699-915). The F-actin subunits from the chicken actomyosin-5a rigor structure (PDB:7PLU, (Pospich et al., 2021)), with phalloidin removed, were also fit into the density corresponding to actin using Isolde (Croll, 2018). As the resolution of F-actin and the motor domain was insufficient to fit side-chains, and their structures have already been published at high-resolution, only the backbone is included for these domains in our model (Pospich et al., 2021). Side-chain orientations are included for the lever from the simulated pseudoatomic model. The quality of the final model was assessed using MolProbity (Chen et al., 2009) (Table 2.2).

Table 2.2: ***MolProbity statistics for the pseudoatomic model of rigor Myo5a-S1-6IQ bound to F-actin***



All-Atom Contacts	Clashscore, all atoms:	0.27	
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Protein Geometry	Poor rotamers	17	1.83%
	Favored rotamers	860	92.37%
	Ramachandran outliers	2	0.07%
	Ramachandran favored	2744	98.28%
	Rama distribution Z-score	-1.57 ± 0.14	
	MolProbity score^	0.80	
	Cβ deviations >0.25Å	0	0.00%
	Bad bonds:	7 / 15933	0.04%
	Bad angles:	65 / 20710	0.31%
Peptide Omegas	Cis Prolines:	0 / 92	0.00%
	Twisted Peptides:	2 / 2829	0.07%
Low-resolution Criteria	CaBLAM outliers	30	1.1%
	CA Geometry outliers	8	0.28%
Additional validations	Chiral volume outliers	0/1286	
	Waters with clashes	0/0	0.00%

2.2.5 Assessing conformational differences between classes

To calculate angles and distances between lever subdomains, the Isolde generated lever model (pre-simulation, Section 2.2.4) and the motor model (Section 2.2.4) were flexibly fit in to each cryo-EM class with torsional and distance restraints applied using Isolde (Croll, 2018). To calculate the angles, the molecule was subdivided into 7 mobile domains: base (actin binding interface), converter, and CaMs 1-6 (Figure 2.2). Vectors were used to represent each domain, drawn between Cα atoms for the base (res 384 and 543), the converter (res 721 and 760), and CaMs (res 136 and 63). Local uvw axes were then calculated for each subdomain pair. The u axis is the vector between the midpoints of each vector in the pair. The v axis is the vector that is orthogonal to u, which sits in the plane of u and the vector representing the first subdomain of the pair. The w axis is the vector orthogonal to u and v. v' and w' were calculated to improve the contrast between vectors, as the first subdomain of the pair is dimensionless in w. The v' axis is the v axis rotated 45 ° about the u axis. The w' axis is the w axis rotated 45 ° about the u axis. The angles between the subdomain vector pairs were then calculated in each plane (vw, uv' and uw'). $\theta_{vw} = \theta$, $\theta_{uv'} = \theta_1$-θ_2, $\theta_{uw'} = \theta_3$-θ_4 (Figure 2.2). Distances were calculated between the Cα atoms of known interacting residues in the N- and C-lobe of consecutive CaMs (Ser17 and Asn111, respectively).

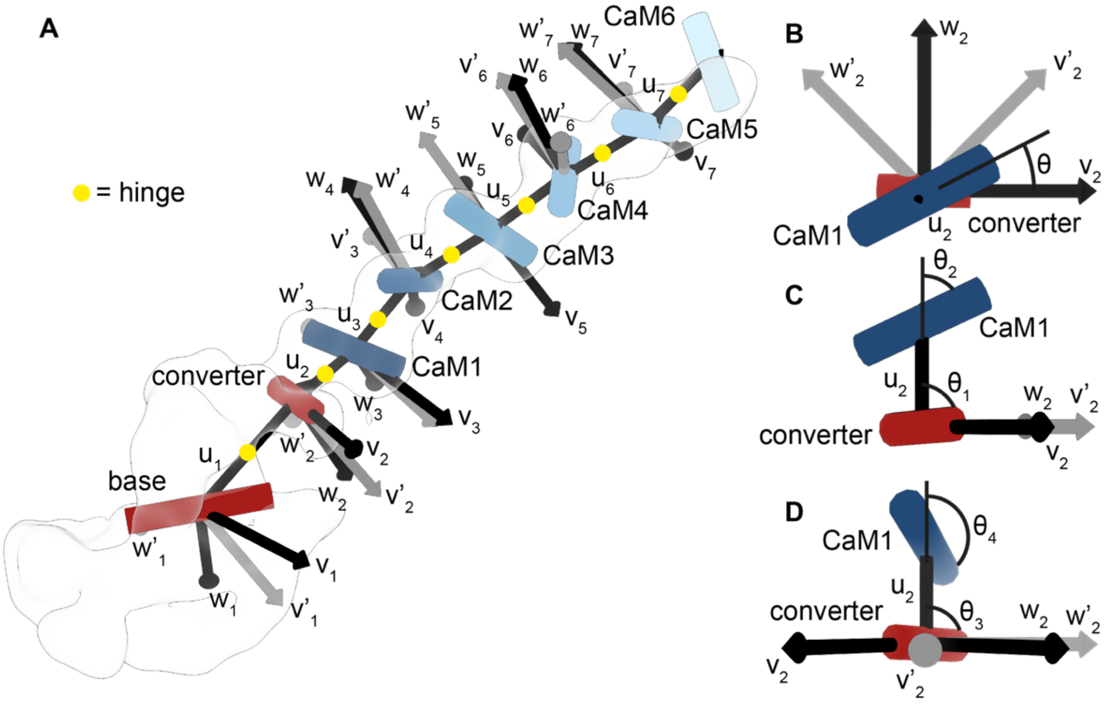


Figure 2.2: Setting up coordinate systems for calculating lever subdomain angles

A: *Example division of subdomains for class D, displayed within cryo-EM density map for class D. Vectors representing the subdomains are visualised as cylinders and coloured as in Figure 2.4. Local subdomain pair axes are numbered by the pair they represent, uvw axes displayed as black arrows, and v'w' axes as grey arrows. Schematics of hinge points are shown as yellow circles.* ***B:*** *Example of calculating subdomain pair angles in vw.* $\theta_{vw} = \theta$. ***C:*** *example of calculating subdomain pair angles in uv'.* $\theta_{uv'} = \theta_1 - \theta_2$. ***D:*** *Example of calculating subdomain pair angles in uw'.* $\theta_{uw'} = \theta_3 - \theta_4$.

2.2.6 Flexibility of the whole lever

To calculate the cantilever bending stiffness of the whole lever, a vector representing the lever was drawn between the midpoint of the vector describing the converter, and the midpoint of the vector describing CaM6 for all classes (Section 2.2.5). To determine the variance in lever displacement, a new coordinate system was calculated using the mean vector as the z-axis (Figure 2.10C-D). The y-axis was calculated as the vector that is orthogonal to z, which sits in the plane of z and the vector representing F-actin (res 145 of the 1st and 13th subunit of a model F-actin). The model of F-actin was made by superimposing the first and last subunit of multiple copies of the actin from 7PLU (Pospich et al., 2021).The x-axis was calculated as the vector orthogonal to z and y. The distance between the end point of the lever unit vector and the end point of the mean unit vector was calculated for each class (d_c) in xy. The variance in lever conformation was calculated using Equation (**2.2**),

$$\text{lever variance} = \frac{\sum_{c=1}^{a} n_c d_c^2}{\sum_{c=1}^{a} n_c}, \qquad \textbf{(2.2)}$$

where n_c is the number of particles in the class, and d_c is distance between the class unit vector and the mean unit vector. This was done for the overall displacement (Figure 2.3A-B) and also for displacement in a particular direction (tilt/slew) (Figure 2.3A & C), to calculate the overall bending stiffness and the bending stiffness in both directions (tilt/slew).

The cantilever spring constant ($k_{\text{cantilever}}$) was calculated using Equation (**2.3**),

$$\text{spring constant } (k_{\text{cantilever}}) = \frac{2k_B TL}{\sigma^2}, \qquad \textbf{(2.3)}$$

where $k_B = 1.380649 \times 10^{-23}$ N m K^{-1} (Boltzmann constant), T = 281.15 K (temperature of grid making in Kelvin), L is the mean length of the lever vectors, and σ^2 = is the variance from Equation (**2.2**). The spring constant was then divided by L^3 to convert into the units pN/nm.

2.2.7 Flexibility within the lever

To calculate the spring constants at the connections between lever subdomains ($k_{\text{subdomain}}$), the variance (σ^2) in angle between subdomain pairs within the classes (weighted by the number of particles in each class) (Section 2.2.5), was calculated in 3 orthogonal planes (vw, uv’, vw’) (Figure 2.2). The vw plane describes torsional stiffness, and the uv’ and uw’ plane describe bending stiffness in 2 directions (Figure 2.2). $k_{subdomain}$ in each plane was calculated using Equation (**2.4**),

$$\text{spring constant } (\mathrm{k_{subdomain}}) = \frac{k_B T}{\sigma^2}, \tag{2.4}$$

where $k_B = 1.380649 \times 10^{-23}$ N m K^{-1} (Boltzmann constant), T = 281.15 K (temperature of grid making in Kelvin), and σ^2 is the variance in angle between subdomain pairs within the classes. The overall bending stiffness was calculated by averaging the values in the uv’ and uw’ plane.

The error for the bending stiffness calculation at each subdomain was estimated by predicting the standard deviation (SD) of bending stiffnesses calculated by randomly assigning angles between the subdomains. For each subdomain pair, the angles calculated between the subdomain pair in each class were sorted into ascending order. A random value for the angle was generated for each class between the angle values either side of it in the sorted list. The ranges for values at the start and end of the list were calculated using the value ± the distance from the single neighbour. The variance (σ^2) was calculated for these random angles, and used to calculate a bending stiffness using Equation (**2.4**). This was repeated 1000 times and the SD of the random bending stiffnesses was quoted as the error.

2.2.8 Predicting the working stroke

To determine what distance the working stroke can travel along the F-actin longitudinal axis, the lever models generated from each class were superimposed onto the converter of a pre- (PDB: 4ZG4, (Wulf et al., 2016)), post-powerstroke (PDB: 7PM6, (Pospich et al., 2021)) and rigor structure (PDB: 7PLU, (Pospich et al., 2021)). As 4ZG4 is not actin bound, to model the motor bound in the ADP-P_i (pre-PS) state the L50 domain of 4ZG4 was superimposed with the actin interacting region of the L50 domain of 7PM6, as the L50 domain is thought to bind actin first (Holmes et al., 2003). The distance from the end (res 914) of each class model of the ADP-P_i state to every other class model of the ADP state, was calculated along the vector defining the actin filament (Section 2.2.6). The same was done from ADP-P_i to rigor and ADP to rigor. The probability of this combination of classes being paired was calculated by multiplying the fraction of particles in the classes, out of the total number of particles.

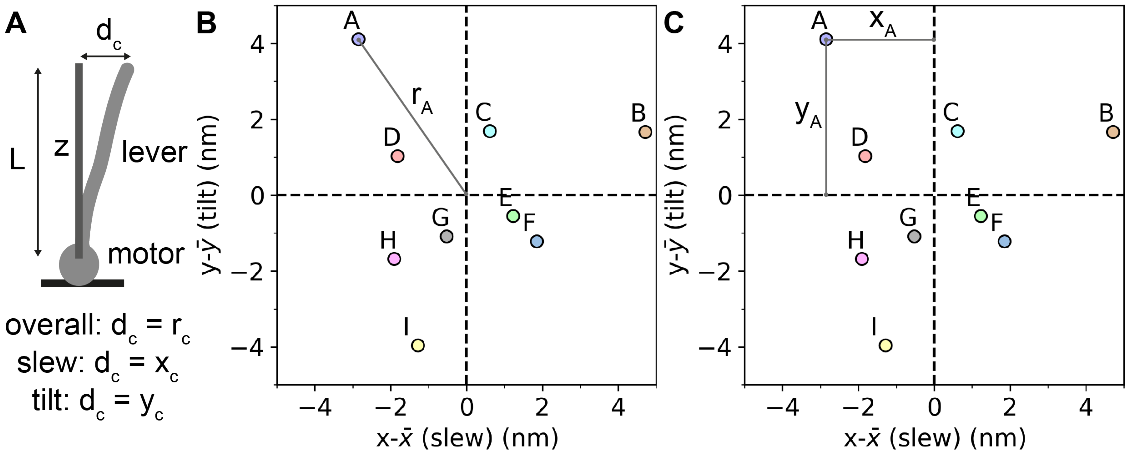


Figure 2.3: Calculating lever displacement for cantilever bending stiffness

A: *schematic of cantilever-type bending of the lever of Myo5a. d_c is the displacement of the lever from the mean position (z) for each class (c). For the overall bending stiffness the displacement (d_c) was calculated using r_c (**B**). To calculate the bending stiffness in each direction, tilt and slew, the displacement (d_c) was calculated using x_c and y_c, respectively (**C**). L is the mean length of the lever. z is the mean vector of the class lever vectors (converter to CaM6) used as the z axis in **B** and **C**. **B:** demonstration of how r_c were calculated for each cryo-EM class using class A (r_A) and the distribution of end points from Figure 2.10D. **C:** demonstration of how x_c and y_c were calculated for each cryo-EM class using class A (x_A and y_A, respectively) and the distribution of end points from Figure 2.10D.*

2.3 Results and Discussion

2.3.1 Structure of the full length actomyosin-5a lever by cryo-EM

To determine the properties of the lever in the active state, I obtained a sub-nanometre cryo-EM structure of Myo5a-S1 bound to F-actin in the rigor state (7.5 Å global resolution) (Figure 2.4A&D). The flexible nature of the lever results in a high level of particle heterogeneity within the cryo-EM data. It was therefore necessary to use focused 3D classification on the lever domain to obtain a high-resolution structure of the lever. This approach, which grouped particles based on lever conformation, resulted in 9 separate classes. The lever conformation varies between the classes, which represent snapshots of continuous motion (see Methods for details). Despite being approximately equally populated, the resolution of each class varied from 7.5 Å-10.7 Å (Figure 2.5). The 3D class with the best resolution for the lever was selected for model fitting (Figure 2.4B & C, Figure 2.5D). In this class the resolution along the lever, from the converter to IQ6, gradually declined from ~7 to 25 Å (Figure 2.4E). The reduced resolution towards the end of the lever is likely a result of conformational heterogeneity even within a single class. Of note, focused refinement, where everything outside of 3 F-actin subunits, the motor and the first 2 IQs was excluded, generated a structure that was highly similar to that generated for the motor domain of chicken Myo5a bound to F-actin (Pospich et al., 2021) and of similar resolution (4.2 Å) (Figure 2.6). Though the conformation of individual subunits is the same (motor and each actin) (Figure 2.6A), small changes in the motor relative to actin (Figure 2.6B) may be due to flexibility, species difference, or their use of phalloidin to stabilise F-actin.

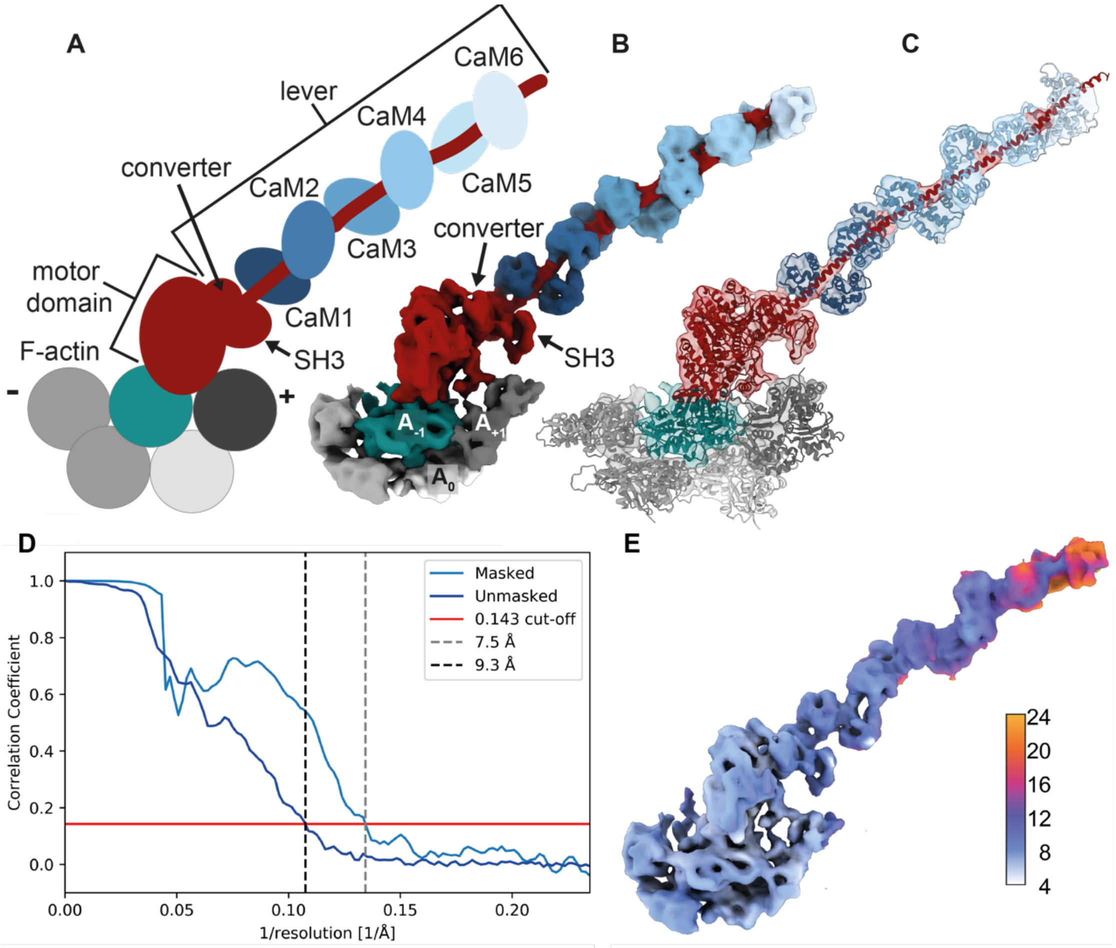


Figure 2.4: **Structure of Myo5a-S1-6IQ bound to actin**

A: *Schematic of Myo5a-S1 bound to filamentous actin (F-actin). The motor domain is shown in red, with the position of its N-terminal SH3-like fold domain and C-terminal converter domain indicated. The lever is comprised of 6 IQ motifs bound to calmodulin (CaM) from 1-6. The polarity of the F-actin filament is shown (+ = plus end, and - = minus end). Two adjacent actin monomers that interact with the motor domain are shown where green indicates the actin subunit closest to the minus end* A_{-1}*, and dark grey indicates the actin subunit closest to the plus end* A_{+1}*.* ***B:*** *Cryo-EM split map (contour level: 0.25) of a single Myo5a-S1-6IQ class with the full-length lever (global resolution of 7.5 Å* $FSC_{0.143}$*).* ***C:*** *Shows the cryo-EM density map (contour level: 0.25) of Myo5a-S1-6IQ with the pseudoatomic model fitted****.*** ***D:*** *Fourier shell correlation (FSC) curve illustrating masked 7.5 Å resolution and unmasked 9.3 Å resolution at 0.143 FSC.* ***E:*** *Local resolution calculated with SPOC (Statistical Processing of Cryo-EM maps) (Beckers and Sachse, 2020) displayed on the actin bound Myo5a-S1-6IQ cryo-EM map (contour level: 0.25). Colour bar shows the resolution in Å.*

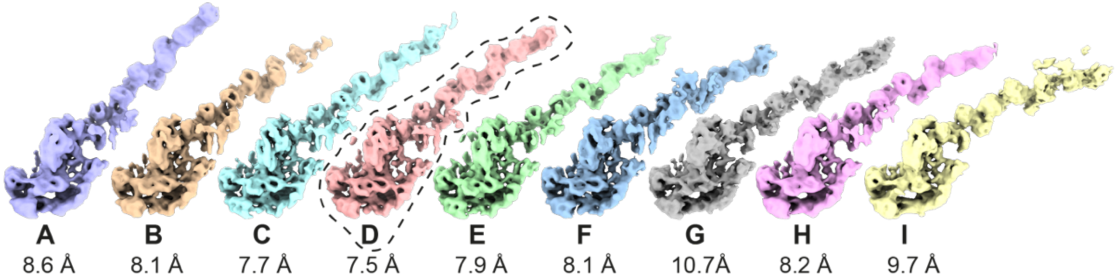


Figure 2.5: **9 cryo-EM 3D classes of Myo5a-S1-6IQ bound to F-actin**
Focused cryo-EM 3D classification on the lever domain reveals snapshots of continuous lever motion. Classes are ordered by the proximity of the lever end to F-actin. ***A-I*** *show cryo-EM maps for each 3D class, processed using DeepEMhancer (Sanchez-Garcia et al., 2021), and with dust hidden (contour level: 0.28). The estimate of the masked global resolution at 0.143 FSC (Fourier Shell Correlation) is displayed beneath each class. The circled class (*__D__*) shows the class displayed in the main figure used for model building (Figure 2.4).*

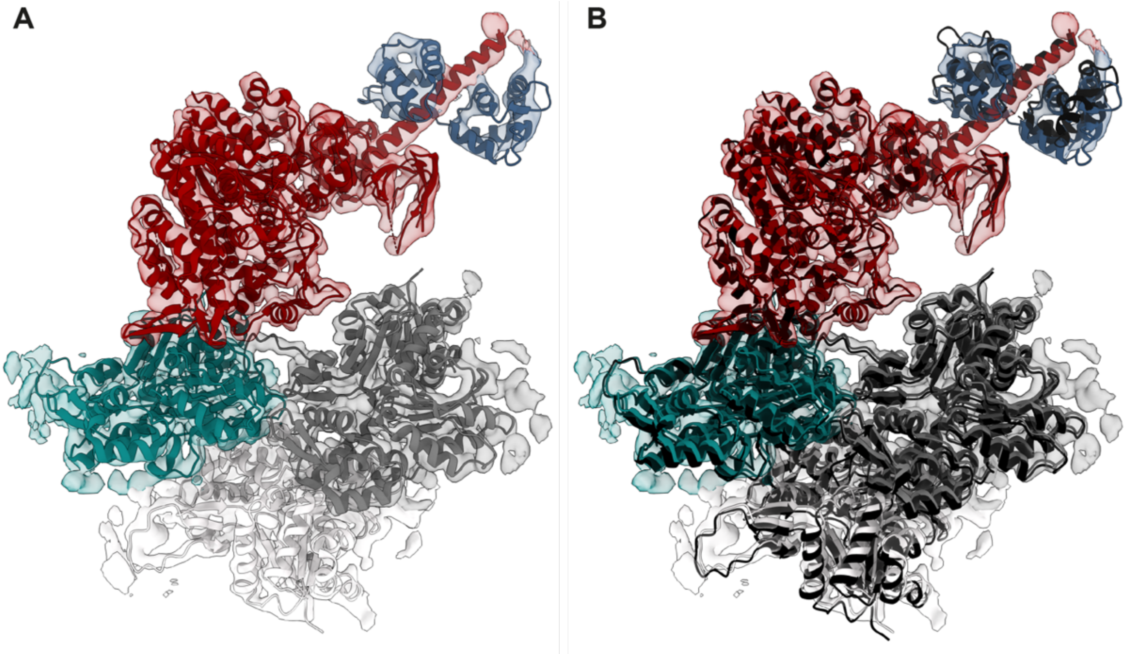


Figure 2.6: Fitting of published rigor motor and CaM structure into the cryo-EM motor map obtained in this study
A: *cryo-EM map of murine Myo5a motor + 1IQ and 3 actin subunits, into which individual chains of the previously solved chicken actomyosin-5a rigor model (PDB: 7PLU, (Pospich et al., 2021)) were rigid fit into the EM density map. Essential light chain (as in 7PLU) was substituted with the CaM1 structure from the CaM1 bound to IQ1 crystal structure (PDB: 2IX7, (Houdusse et al., 2006)). Colours as in* Figure 2.4. ***B:*** *As in* ***A****, but with the original structure (7PLU) overlayed in black. The motor domain of the original structure is rigid fit into the density (displayed in black). All rigid fitting was performed in ChimeraX* (Pettersen et al., 2021).

2.3.2 Conformation of CaM pairs within cryo-EM classes

CaM pairs bound to IQ domains spaced 23 residues apart (CaM1-2, 3-4, 5-6), hereafter described as 23-residue pairs, exhibited conserved mean conformations that matched with a previous crystal structure of CaM bound IQ1-2 (Figure 2.7) (Houdusse et al., 2006). Conformations were determined by the relative orientation of the CaMs in 3 orthogonal planes (vw, uv', uw') (Figure 2.2), and by the distance of known interacting residues in the N- and C-lobe of consecutive CaMs (Ser17 and Asn111, respectively). The orientation in the vw plane describes the angle of rotation between CaMs in a pair about the heavy chain axis (Figure 2.2). The average angle for each 23-residue pair was between 88 ± 3° to 96 ± 9° (mean ± SD), in the vw plane (Figure 2.7C). The orientation angles in the uv' and uw' planes describe bending of the lever (Figure 2.2). The average angle for each 23-residue pair was between 129 ± 2° to 134 ± 7° (mean ± SD) and 10 ± 4° to 21 ± 10° (mean ± SD) in the uv' and uw' planes, respectively (Figure 2.7A-B). The average distance between each 23-residue pair was between 0.9 to 1.3 nm (Figure 2.7D). In all 3 planes, the angle between the CaM pair in the crystal structure differed by a modest ~10° from the mean angle between 23-residue pairs, and 0.6 nm from the mean distance. These slight differences are likely to be a result of the flexibility allowed in cryo-EM that is not allowed in a packed crystal (Ravikumar et al., 2022). It should be noted that CaM5-6 is more variable, especially in the uw' plane. Therefore, if it is excluded from the averages, smaller differences are apparent. It is difficult to determine whether the variability in CaM5-6 is related to the reduced resolution in this region or is truly a reflection of the properties of the lever.

In contrast, CaM pairs bound to IQ domains spaced 25 residues apart (CaM2-3, 4-5), hereafter described as 25-residue pairs, varied more in their relative orientations. The average angle for each 25-residue pair was between -53 ± 2° to -38 ± 13°, in the vw plane (Figure 2.7C). The average angle for each 25-residue pair was between -9 ± 3° to -5 ± 7° and 44 ± 23° to 73 ± 4° in the uv' and uw' planes, respectively (Figure 2.7A-B). This suggests that 25-residue pairs do not have conserved mean conformations in the uw' plane.

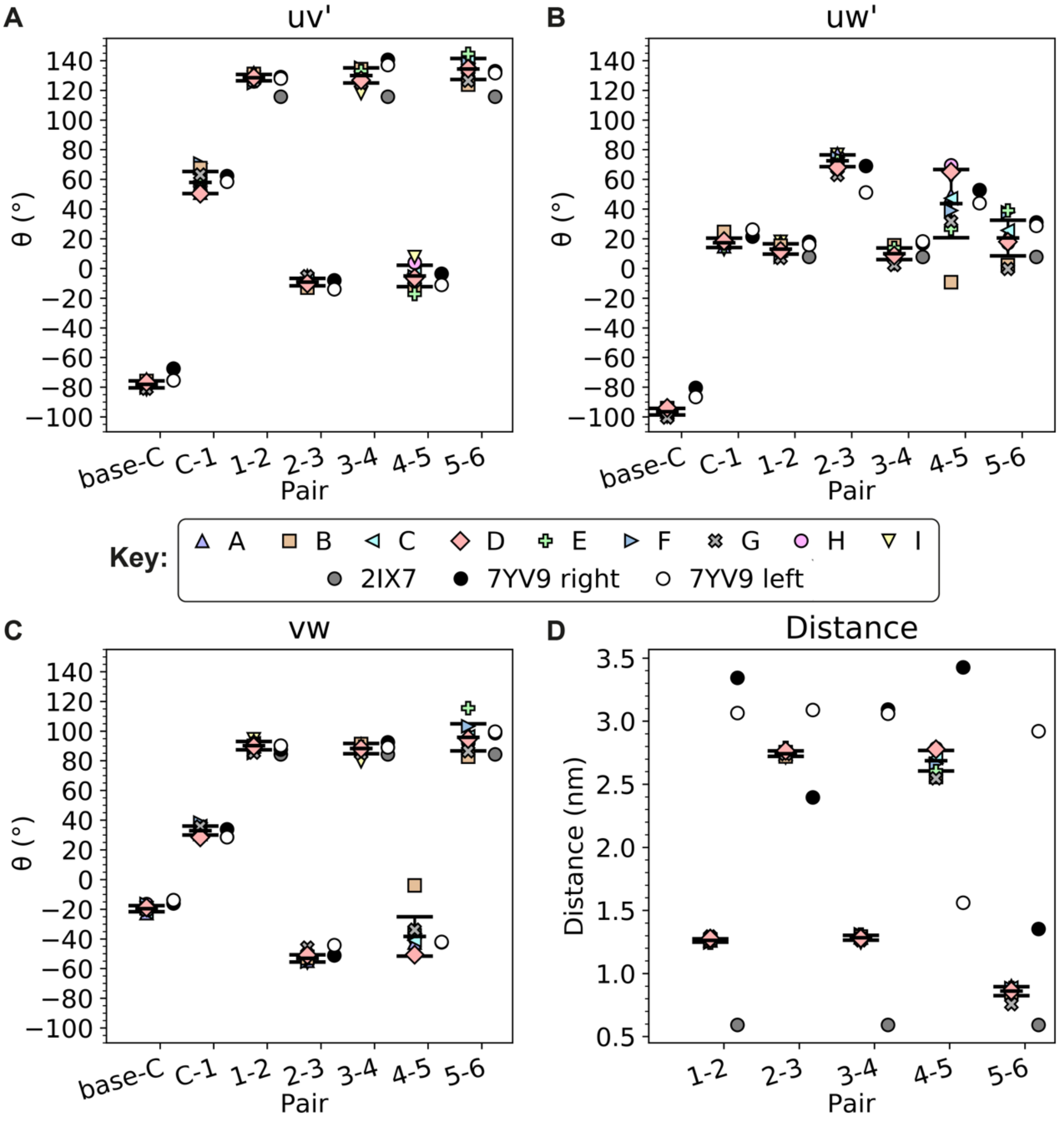


Figure 2.7: Angles and distances between lever subdomain pairs within cryo-EM 3D classes

Classes A-H are named and coloured as in Figure 2.5. Error bars show the mean and standard deviation of the classes. Conformations of known structures (PDB: 2IX7 and 7YV9 (Houdusse et al., 2006; Niu et al., 2022)) are displayed to the right of our data. Grey circles indicate values for the CaM pair bound to IQ1-2 crystal structure (PDB: 2IX7, (Houdusse et al., 2006)). Black circles indicate conformations of subdomain pairs in the structure of the shutdown state in the head furthest from the C-terminal region of the coiled-coil (right) (PDB: 7YV9 (Niu et al., 2022)*). White circles indicate conformations of subdomain pairs in the structure of the shutdown state in the head closest to the C-terminal region of the coiled-coil (left) (PDB: 7YV9* (Niu et al., 2022)*).* ***A:*** *Angle between lever subdomain pairs in their local uw' plane (Figure 2.2).* ***B:*** *Angle between lever subdomain pairs in their local uv' plane (Figure 2.2).* ***C:*** *Angle between lever subdomain pairs in their local vw plane (Figure 2.2).* ***D:*** *Distance known interacting residues in the N- and C-lobes of consecutive CaM pairs (measured between the Cα of Ser17 and Asn111, respectively). Base = actin binding interface, C = converter, 1 = CaM1, 2 = CaM2, 3 = CaM4, 5 = CaM5, 6 = CaM6 (Figure 2.2).*

2.3.3 Predicting side-chain interactions with atomistic MD simulations

The cryo-EM 3D classes revealed possible interactions between all neighbouring CaMs, and the converter and CaM1. As there was insufficient resolution for de-novo model building, all-atom MD was used to simulate a pseudoatomic model based on class D (Figure 2.5). This enabled identification of possible side-chain interactions that fit into the low-resolution density (Appendix A.1). A limited number of interactions exist between the converter and CaM1, which could explain the proposed “pliant point” (Houdusse et al., 2000; Burgess et al., 2002). The pseudoatomic model shows two conserved interactions seen in previous structures: a hydrophobic interaction Val712-Val92, and an ionic interaction Lys715-Glu87 (Appendix A.1) (Coureux et al., 2003; Pospich et al., 2021). These interactions appear >90 % of the time in unrestrained MD simulations (Appendix A.3).

The pseudoatomic model also reveals conserved CaM-CaM contacts between 23-residue pairs that are observed in the crystal structure of CaM bound to IQ1-2 (Houdusse et al., 2006). Further transient and novel contacts were seen in unrestrained MD simulations but are not discussed here (Appendix A.3). However, this highlights the dynamic nature of the molecule and necessity to capture multiple conformational states for a true understanding of how CaM-CaM interactions play a role in lever mechanics.

The extended density indicative of an interaction between 25-residue pairs is not present in all maps (present in classes A, D and H, Figure 2.5). The interface between 25-residue pairs corresponds with a potential electrostatic interaction between Glu14-Arg90. This interaction is equivalent to that previously reported between adjacent MLC1P light chains (Asp9-Lys88) observed in a crystal structure of MLC1P bound to IQ2 and IQ3 (25-residue pair) for Myo2p (yeast Myo5a homologue (Terrak et al., 2005)). The interaction between Glu14-Arg90 was therefore restrained in our pseudoatomic model of class D. In the unrestrained MD simulations, the presence of this interaction is limited to 2.6% in CaM2-3 and 4.9% in CaM4-5 (Appendix A.3). The transient nature of this interaction could explain why an interface cannot be seen in the remaining 6 classes, as they reflect conformational states when the interaction is not present (~95%). The presence of these interactions for ~5% of the MD simulations compared to ~30% in the cryo-EM classes may be reflective of the averaging of particles performed in the 3D reconstruction of classes. Further transient and novel contacts were seen in unrestrained MD simulations but are not discussed here (Appendix A.3).

Side-chains can be highly dynamic, therefore it is unlikely there is a single answer for the conformation of all side-chains in a structure (Miao and Cao, 2016). It has also been shown that experimental technique can influence side-chain conformations (Ravikumar et al., 2022; Bock and Grubmüller, 2022). Using atomistic MD simulations to predict side-chain orientations and interactions, allows chemically realistic alternative conformations to be explored. This approach is also a way of determining possible interactions, working out the probability a specific interaction will take place, quantifying their importance, and determining if they are just a product of the experiment. It has previously been shown that in a given molecule approximately 16-17% of residues are flexible, and that there is a need for methods that predict which side-chain orientations are stable and which are flexible (Miao and Cao, 2016). The percentage of time an orientation appears in MD simulations could be used to predict stable and flexible orientations, in the same way it has been used to predict stable interactions here. An understanding of what side-chain orientations and interactions are relevant would facilitate protein docking, protein design and computer-aided drug design.

2.3.4 The N-lobe of CaM6 is highly dynamic

The cryo-EM density map shows weak density for the N-lobe of CaM6 (Figure 2.4 & Figure 2.9). This could be accounted for if the N-lobe of CaM6 is highly mobile, arising from reduced interactions with the IQ domain and the adoption of an extended conformation, as previously predicted (Terrak et al., 2003; Terrak et al., 2005; Black and Persechini, 2010). We therefore used a crystal structure for MLC1P bound to IQ4 of the yeast Myo5 isoform Myo2P, in which the N-lobe of MLC1P does not interact with IQ4 (PDB:1M46, (Terrak et al., 2003)) to build our pseudoatomic model of CaM6. MD simulations demonstrated that N-lobe of CaM6 does have increased mobility in the extended state compared to the compact state seen in CaMs 1-5 (Figure 2.8). It is therefore plausible that reduced interactions between the N-lobe of CaM6 and IQ6, and the increased flexibility of CaM6 in an extended state, are responsible for the weak of density in the cryo-EM map for the N-lobe of CaM6.

Amino acid substitutions in IQ6 may help to account for the weaker interactions between IQ6 and the N-lobe of CaM (Figure 2.9). In both Myo5a and Myo2p, the conserved Gly residue at position 7 in IQ6 is replaced by a bulky, charged residue (Arg in Myo5a and Lys in Myo2P). In addition, Arg at position 11 is replaced with Lys (Myo5a) or Gln (Myo2p). It is thought that amino acid substitutions at positions 7 and 11 are responsible for light chains forming an extended state (Terrak et al., 2003). Replacement of Gly7 with Arg in a reference IQ motif weakens its interaction with Ca^{2+} free CaM (apo-CaM) by ~2-fold and is expected to promote the extended state of the N-lobe of CaM, although compensatory electrostatic interactions between Arg and the IQ may occur (Black and Persechini, 2010). In support of this, our pseudoatomic model, and unrestrained MD simulations, show electrostatic contacts do form between Arg7 and Lys11 with the charged N-terminal residues (Glu7 and Asp2, respectively) (Appendix A.1, 2 & 3).

In IQ3 of Myo5a, Gly7 is replaced by Met, a bulky and hydrophobic residue. Replacement of Gly7 with Met in a single reference IQ motif weakens its interaction with apo-CaM ~10-fold (Black and Persechini, 2010). Based on this, the N-lobe would be expected to adopt an extended conformation. However, CaM3 adopts a compact conformation in our cryo-EM map (Figure 2.9C, Figure 2.4). The presence of neighbouring CaMs may reduce the flexibility of the N-lobe and stabilise interactions between the N-lobe and the IQ motif, suggesting CaMs bind the heavy chain cooperatively. It is possible that under some conditions this lobe could become extended, as a potential strategy for regulating myosin activity, by weakening the lever.

The flexibility seen in CaM6, and its reduced interactions with IQ6, could be important for motor mechanics as plasticity in this region is likely to be necessary for switching between the active and shutdown state. A recent cryo-EM structure of the shutdown inhibited state of Myo5a showed the N-lobes of each CaM6 interact with the coiled-coil to stabilise the sharp bend at the head-tail junction (Niu et al., 2022). Therefore weakened interactions between the N-lobe and IQ6 are likely to be necessary for Myo5a to readily move from the active to the shutdown state. Flexibility at CaM6 may also prevent steric clashes between the two levers at the head-tail junction of dimeric Myo5a. Based on the different behaviour Met substitution has on the CaM bound to a single IQ compared to CaM3 in our lever, the behaviour of CaM6 might be quite different in the dimeric molecule compared to S1 where there is no downstream sequence. High-resolution structures are needed of the head-tail junction in the active dimeric state in order to truly understand the role CaM6 plays.

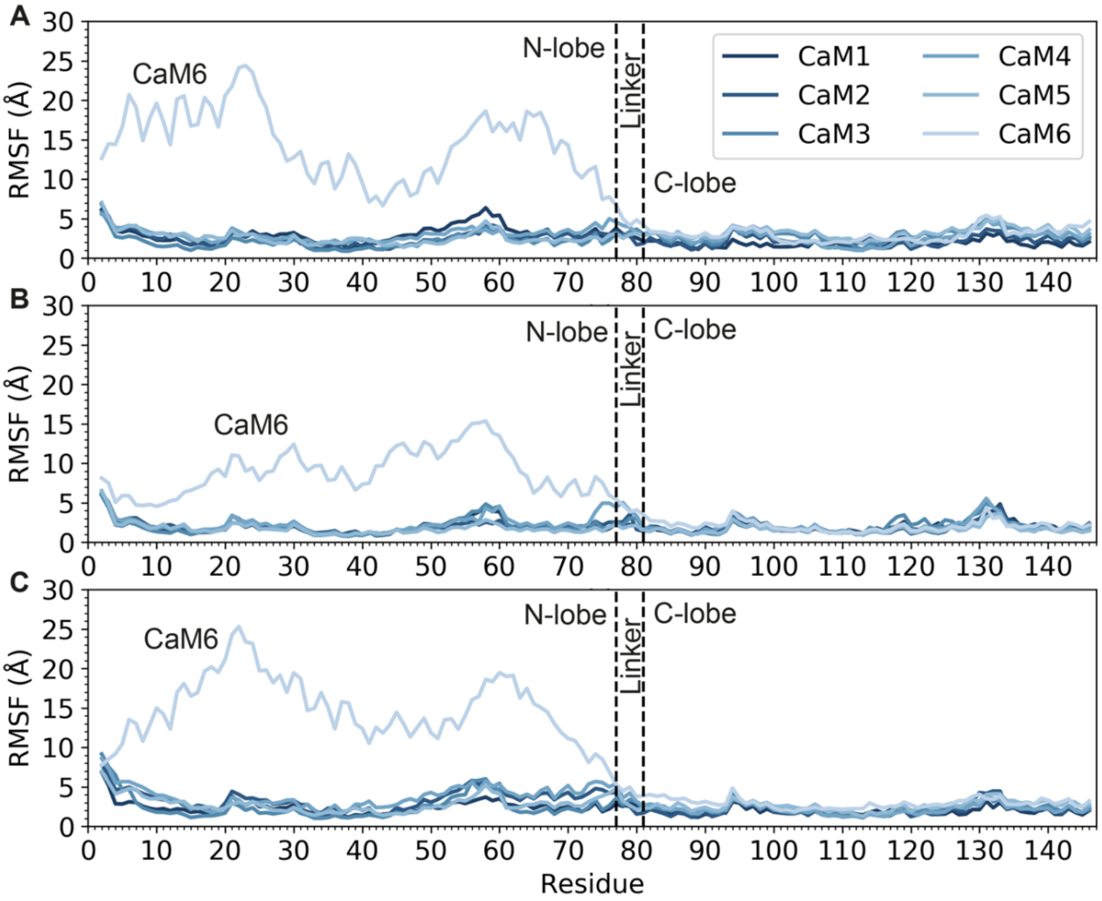
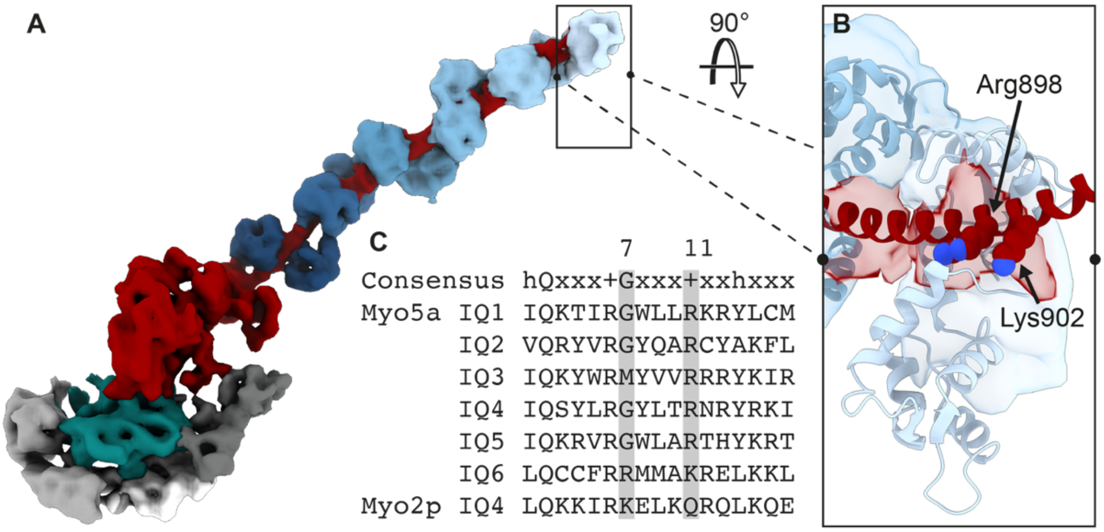


Figure 2.8: Per residue RMSF of CaM for triplicate repeats
***A-C** plots of per residue RMSF (Root Mean Square Fluctuation, Å) of CaMs in pseudoatomic model simulations in which the heavy chain is restrained and CaMs are unrestrained. The locations of the N-lobe, linker and C-lobe are indicated. Subfigures are individual repeats.*



```
Consensus hQxxx+Gxxx+xxhxxx
Myo5a IQ1 IQKTIRGWLLRKRYLCM
      IQ2 VQRYVRGYQARCYAKFL
      IQ3 IQKYWRMYVVRRRYKIR
      IQ4 IQSYLRGYLTRNRYRKI
      IQ5 IQKRVRGWLARTHYKRT
      IQ6 LQCCFRRMMAKRELKKL
Myo2p IQ4 LQKKIRKELKQRQLKQE
```

Figure 2.9: CaM6 N-lobe forms fewer interactions with heavy chain
***A:** cryo-EM map of Myo5a-S1-6IQ, CaM bound to IQ6 (CaM6) boxed. **B:** 90° rotation of boxed region in **A,** with a pseudoatomic model fitted to the cryo-EM density. Residues at positions 7 and 11 of IQ6 are indicated (Arg898 and Lys902 respectively). **C:** Sequence alignment of all 6 IQ motifs of murine Myo5a together with IQ4 of Myo2p (PDB:1M46* (Terrak et al., 2003)*).* *Positions 7 and 11 of each IQ motif are highlighted.*

2.3.5 Overall lever flexibility is directionally isotropic

Using the variance in lever position within the cryo-EM 3D classes, treating the lever as a cantilever, I calculated the overall stiffness of the lever to be 0.78 pN/nm for slew motions (across F-actin short axis) and 0.74 pN/nm for tilt motions (bending towards and away from F-actin) (Figure 2.10). This indicates flexibility is directionally isotropic. Additionally, the change in tilt and slew is uncorrelated, with a correlation coefficient of 0.004, which implies the motion of the lever in these directions is not coordinated. Both of these attributes may be important for strain generation in multiple directions. Equal rigidity in all directions would allow intramolecular strain to be generated between Myo5a heads regardless of their orientation. This may be required in order for Myo5a to walk in either a left- or right-handed manner as seen in interferometric scattering microscopy (iSCAT) and across filaments in an F-actin network as seen in super-resolution stochastic optical reconstruction microscopy (STORM) (Andrecka et al., 2015; Lombardo et al., 2019).

The overall cantilever bending stiffness (0.76 pN/nm) is similar, but somewhat higher than the stiffness of Myo5a-S1 determined by optical trap measurements (0.2 pN/nm, (Veigel et al., 2002)), dimeric actomyosin-5a by negative stain EM (0.26 pN/nm, (Oke et al., 2010)), the unbound 2IQ Myo2 lever (0.37 pN/nm, (Billington et al., 2014)), and an SAH (0.46 pN/nm, (Sivaramakrishnan et al., 2008)). Our approach is the first direct analysis of Myo5a lever stiffness in 3D, which could account for some of the differences in values between our measurements and other Myo5a-S1 data. It should also be taken into consideration that grouping the cryo-EM data into conformational classes during 3D classification will have a reduced variance compared to looking at individual particles.

Furthermore, the conformations of the lever seen could be somewhat constrained by the packing of Myo5a-S1 on F-actin. Additional density corresponding to neighbouring motors can be observed in most classes (Figure 2.5A-F, and H). This could also be a barrier to capturing the full motion of the lever as interactions or clashes between adjacent levers may occur. However, not all filaments in the micrographs were fully decorated. Future work would involve classifying motors with and without neighbours, to determine the true influence of packing on F-actin on the emergent bending stiffness.

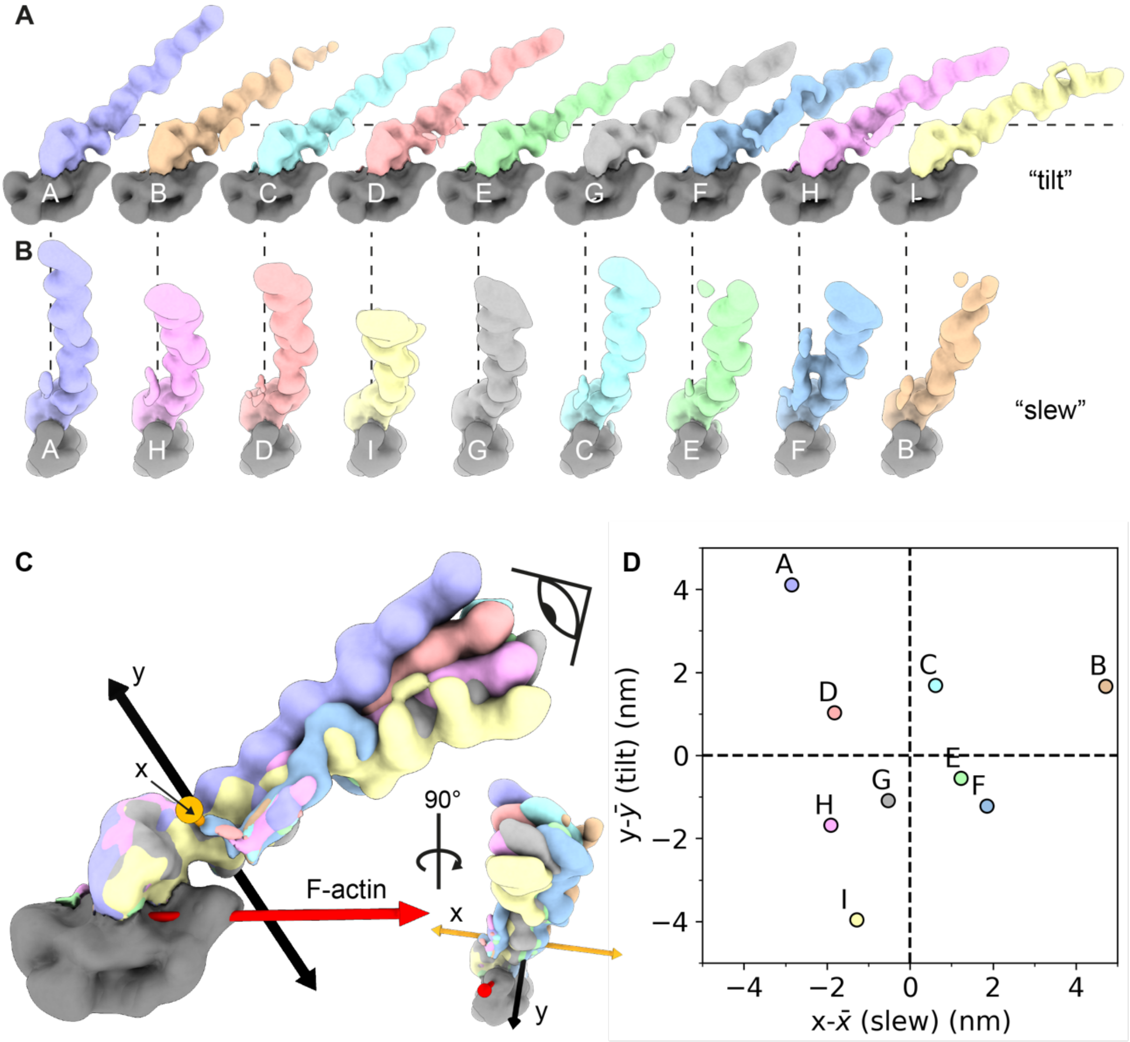


Figure 2.10: Lever flexibility is directionally isotropic in 3D classes
Classes are the result of focused cryo-EM 3D classification on the lever domain and reveal snapshots of continuous lever motion. ***A:*** *Post-processed maps of 3D classes in order of lever bend along the F-actin longitudinal axis (tilt).* ***B:*** *post-processed maps of 3D classes in order of lever bend along the F-actin short axis (slew). Maps are Gaussian smoothed (SD 5 Å) (contour level: 0.15).* ***C:*** *Post-processed maps of 3D classes (gaussian filtered SD 5 Å, contour level: 0.15). The coordinate system the stiffness measurements were taken from are displayed as 3D arrows (see Materials and Methods). x-axis = orange arrows, y-axis = black arrows, z = axis mean lever vector (see Materials and Methods). Eye shows viewpoint of* ***D****.* ***D:*** *Plot of the displacement of the end lever in each class from the mean (z) (see Materials and Methods). Changes in x represent motion across the F-actin short axis (slew), changes in y represent motion along the F-actin longitudinal axis (tilt).*

2.3.6 Flexibility is not focused at the pliant point

Using the ensemble of lever conformations in the cryo-EM 3D classes, we found that the lever of Myo5a has variable stiffness at points along its length. Spring constants for bend and twist at regions between the motor and converter and all 6 CaMs were determined (Table 2.3 and Figure 2.11) assuming bend and twist are focused at single hinge points between lever subdomains (Figure 2.2). The stiffest connections are between CaM1-2 and CaM2-3. The latter is somewhat surprising as there are limited contacts between CaM2-3 (Appendix A.3), and it was previously thought this would be a region of increased flexibility (Terrak et al., 2005). This suggests that the distance between CaM pairs (Figure 2.7D) and CaM-CaM interactions do not dictate the stiffness at this region. The connections between the converter and CaM1, and CaM3-4 are about half as stiff as CaM1-2 and CaM2-3, while CaM4-5 and CaM5-6 are the most flexible connections. This demonstrates regions of pliancy are distributed along the whole length of the lever rather than being localised at a 'pliant point' between the converter and CaM1 as previously thought (Houdusse et al., 2000; Burgess et al., 2002), although this is more pliant than its neighbouring connections. No correlation could be found between bending stiffness and subdomain pair conformation, distance or interdomain interactions (Figure 2.7, Appendix A.3). Interestingly, in performing the same analysis of subdomain conformations for previous published structures, I found previous structures fall within the range of subdomain conformations seen within the 3D classes (Figure 2.7).

Variable stiffness along the length of the lever is likely to be important for Myo5a mechanics. The lever must be sufficiently stiff to generate intramolecular strain, and to withstand load from cargos without collapsing. The lever must also be sufficiently flexible to accommodate small changes in step size, stereospecific binding orientations, and swap filament track in F-actin networks (Oke et al., 2010; Lombardo et al., 2019). Constructs that have 2 IQ domains but the rest of the lever replaced with a single alpha helix (SAH) can still walk processively, but do not produce intramolecular strain (Baboolal et al., 2009). The stiffness of a SAH is ~50 pN nm, which is ~1/3 the stiffness at hinges between CaM4-6. Taken together this could mean that: some rigidity at the start of the lever is required for translating the powerstroke, some rigidity (at least 3 x SAH stiffness) is required to communicate intramolecular strain between heads, and that some flexibility exists towards the end of the lever to accommodate for stepping errors.

Despite 3D classification and reconstruction reducing the conformational variance that can be observed, in comparison to looking at individual molecules, it is only through this approach that sufficient resolution can be acquired to resolve differences in subdomain conformations and have information in 3D. Therefore, in performing bending stiffness calculations, there is a compromise between the requirement to average particles to allow 3D reconstruction, and the reduction in the conformational variance that comes with it. As the molecule is so flexible, more data is required in order to generate more classes. If, for example, 1000 classes are necessary to perform proper statistical analysis, and ~20000 particles are required for a reconstruction with sufficient resolution to resolve CaMs. 20 million particles would be required to gain the data necessary. This is ~60 times more particles than we acquired, which would mean 60 times more micrographs (~300,000). Collecting this number of micrographs is starting to become achievable on reasonable timescales with the new fast detectors and automation software, but, together with the large amount of processing required, may be a significant investment that is likely to only modestly improve our estimates.

Table 2.3: Spring constants for stiffness at hinge points within the lever (2.s.f), ± SD of random error (see Methods for details).

Hinge	Combined spring constant bend (pN nm)	uv' spring constant bend (pN nm)	uw' spring constant bend (pN nm)	vw spring constant twist (pN nm)
Base-converter	2500 ± 560	2400 ± 420	2600 ± 710	2900 ± 1000
converter-CaM1	770 ± 410	230 ± 39	1300 ± 770	1400 ± 390
CaM1-2	2000 ± 470	2900 ± 680	1100 ± 260	1600 ± 470
CaM2-3	1400 ± 420	2000 ± 460	820 ± 380	2100 ± 910
CaM3-4	660 ± 330	480 ± 340	840 ± 320	1100 ± 1500
CaM4-5	140 ± 54	250 ± 94	24 ± 13	73 ± 57
CaM5-6	170 ± 36	250 ± 45	89 ± 27	150 ± 66

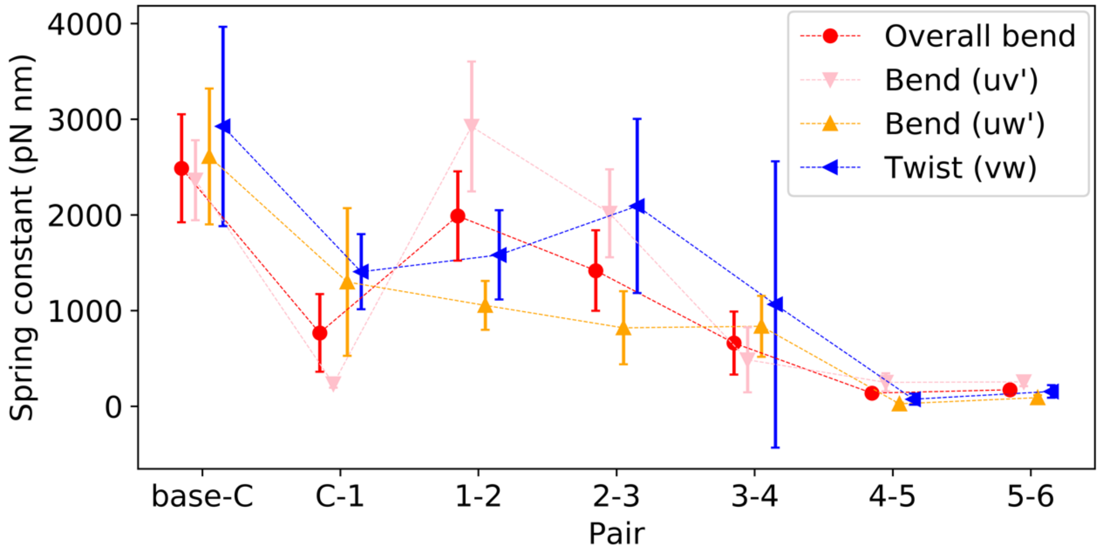


Figure 2.11: Spring constants for stiffness at hinge points within the lever

Spring constants for hinge points at subdomain connections (Figure 2.2). Overall bend is the average bend constant in the uv' and uw' planes. Twist is the spring constant in the vw plane. Base = actin binding interface, C = converter, 1 = CaM1, 2 = CaM2, 3 = CaM4, 5 = CaM5, 6 = CaM6 (Figure 2.2). Error bars show SD of random error (see Materials and Methods for details).

2.3.7 Lever flexibility accommodates 33 nm working stroke

Estimations of the Myo5a-S1 working stroke using our cryo-EM 3D classes revealed a 33 nm working stroke (Figure 2.12). To model lever conformations in the ADP-P_i and ADP state, the modelled lever domains of each 3D class were superimposed onto the converter domains of a known ADP-P_i and a known ADP structure. The lever end points (res 914) translate 31 nm along the F-actin longitudinal axis on average from the ADP-P_i (PDB:4ZG4, (Wulf et al., 2016)) to the ADP (PDB: 7PM5, (Pospich et al., 2021)) state. There was an additional translation of 2 nm along the F-actin longitudinal axis, on average, between the ADP and the rigor state. The total translation of 33 nm along the actin filament is larger than an earlier estimate of the working stroke (28.5 nm), which used a similar approach but looked at lever end-point distances between the states rather than change along F-actin, and was limited to a single lever conformation from the shutdown state (Wulf et al., 2016). Our estimate is also larger than the experimentally measured working stroke for Myo5a-S1 (21 nm) using an optical trap. In the optical trap data the working stroke was comprised of a 16 nm step as Myo5a transitioned from ADP-P_i to ADP, and a further 5 nm step as Myo5a transitioned from ADP to rigor, although the working stroke measured for dimeric Myo5a was longer (25 nm) (Veigel et al., 2002). Previously the differences seen between S1 and the dimer were suggested to be due to differences in how the molecules attach to the nitrocellulose bead in the optical trap assay. However, the difference between S1 and the dimer could be attributed to the leading head in the dimer starting its working stroke from a strained ADP conformation, and not the ADP-P_i state. If the 4 nm increase in the working stroke length of the dimer compared to S1 is considered in our measurements, a working stroke of 37 nm could be extrapolated for the dimer. This is much closer to the Myo5a step size (36 nm) (Walker et al., 2000; Veigel et al., 2002; Andrecka et al., 2015).

In the 3-bead optical trap, restrictions to the movement of the end of the S1 lever through its attachment to a nitrocellulose bead, together with some uncertainty as to how the lever is attached to the bead, may contribute to an underestimate of the working stroke. Importantly, the previous estimates of the working stroke are much shorter than the step size of 36 nm. This has led to the idea that when the rear head of Myo5a detaches, it has to perform a ~10 nm diffusive search to rebind to the actin filament at the next binding site (Veigel et al., 2002). However, our estimate of the working stroke suggests that this diffusive search would only be a few nm, if any (as 3 nm would be within 10% error). This suggests strain across the molecule is not generated during binding of the new lead head, but only after P_i release, as the lead lever is prevented from entering the post-powerstroke conformation as it is restrained by the trail head.

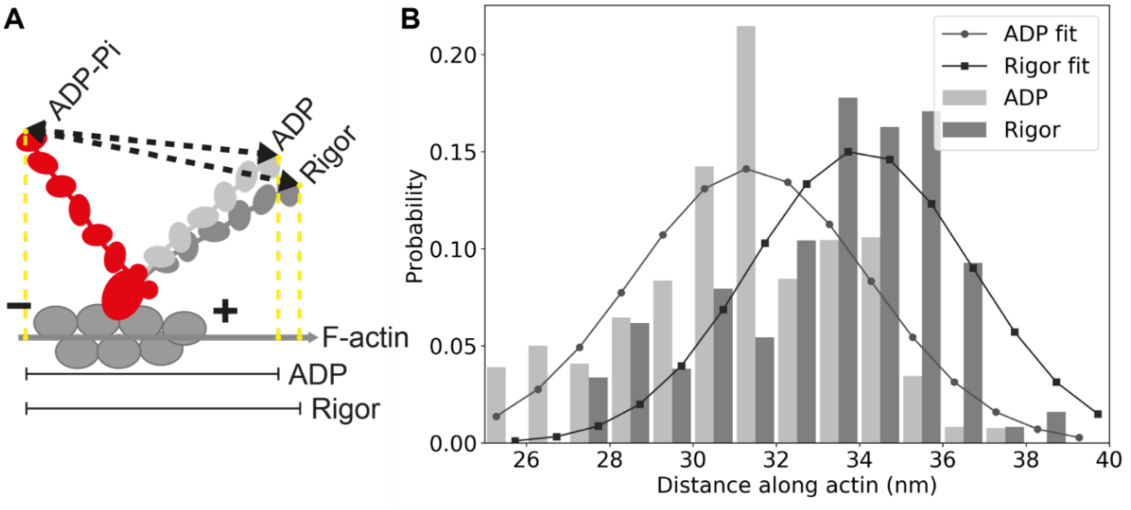


Figure 2.12: Predicting the working stroke of Myo5a-S1
A: *schematic demonstrating how distances for* ***B*** *were calculated. Distances between lever ends in the modelled ADP-Pi, ADP and rigor conformations are shown as black dashed arrows. Yellow dashed lines show these distances as a translation along the F-actin vector (1st to 13th subunit, grey arrow, see Materials and Methods for details). + indicates the plus end of F-actin, - indicates the minus end of F-actin.* ***B:*** *Histogram of the translation of Myo5a lever ends along F-actin vector from the ADP-P_i conformation to ADP and to rigor (see Materials and Methods for details), and the fitted Gaussian distributions. The probability was calculated using the proportion of particles in each of the cryo-EM 3D classes paired.*

2.4 Conclusion

Here, for the first time, the structure of Myo5a bound to actin with the full-length lever has been resolved with sufficient resolution to distinguish neighbouring CaMs, and their individual lobes. Using MD simulations, a pseudoatomic model of the lever domain was generated to predict side-chain orientations, and to identify novel and conserved interactions along the length of the lever. Further characterisation of the relative orientation of neighbouring CaMs in 3D, demonstrated that the conformation of 23-residue pairs is well conserved, however the conformation of 25-residue pairs is not. Most interestingly no single characteristic, i.e. the orientation of the CaMs in a single axis, or the number of CaM-CaM interactions, seems to correlate with the bending stiffness at connections between connecting regions. This suggests that either the combined influence of these features encodes stiffness at subdomain connections, or features as yet unexplored. Only improved resolution of the CaMs is likely to reveal the answers to this.

3D reconstruction of the full-length Myo5a lever domain in 9 different conformations has revealed properties of the lever that contribute to its mechanics. Absence of the N-lobe of CaM6 in the cryo-EM maps, and increased dynamics in the MD simulations, have provided further evidence that the CaM bound to IQ6 exists in an extended conformation, and has increased flexibility. This may be important for dimer formation, and pliancy at the head-tail junction to aid formation of the shutdown state. Structural resolution of the head-tail junction in the active state is required to further understand the role of CaM at this location.

Analysing lever flexibility in 3D, has demonstrated it is directionally isotropic, which may be important for strain generation in multiple directions. Analysing intra-lever flexibility in 3D, has further supported evidence for a hinge region between the converter and CaM1, and revealed a block of rigidity from CaM1-3, followed by a flexible end. This challenges our view of the lever as a rigid continuous beam. It is possible that 2D ns-EM images of the dimer with both heads bound to F-actin under strain have led to this, as they show a taut lever as opposed to one that can flex and bend (e.g. a tightrope is made of string), which prevents different properties along the length of the lever from being revealed. Whilst, 2D images of the free molecules make it difficult to determine the viewpoint or distortion due to stain, without F-actin as a reference. This highlights why 3D reconstruction, despite its shortcomings, is necessary.

Finally, estimates of the length of the working stroke from the variability in lever position observed in our structures, suggests that it is longer (~33 nm) than previously described (28.5 and 21 nm) (Veigel et al., 2002; Wulf et al., 2016). A working stroke closer to the length of the F-actin helical pseudo-repeat suggests stepping is more precise than previously thought, and that strain is generated after the motor has bound to F-actin, and P_i has been released. However, for a true understanding of the full working stroke, structural resolution of the strained state is required.

Through this analysis I have revealed a more complete picture of what is happening within the lever, and how features of the lever contribute to the function of the whole molecule. Many molecular motors beyond Myo5a use light chain bound lever-like domains (Bähler and Rhoads, 2002; Heissler and Sellers, 2015), so not only does this work contribute to our knowledge of the mechanism of Myo5a but may also shed light on how other motor proteins that share these domains function.

Chapter 3 Simulating Myo5a dynamics

3.1 Introduction

Myo5a is large (431 kDa) and has long-range dynamics (~17.5 ms to take a single step), therefore coarse-grained MD approaches are necessary to capture its key modes of motion. FFEA has previously been used as a coarse-grained MD simulation approach to study the cytoskeletal motor dynein, and parametrised using ns-EM data or atomistic MD simulations (Richardson et al., 2020; Hanson et al., 2021). Here I demonstrate FFEA can be used to generate suitable models of Myo5a dynamics and parametrised using material parameters extracted from cryo-EM data (Chapter 2).

FFEA uses principles from finite element analysis (FEA) which models how macroscale objects respond to stress. In FEA objects are modelled as continua as atomistic detail has little effect on the structure of the object at this scale. In order to simulate continuum objects, objects are represented as a network of elements referred to as a finite element mesh. Partial differential equations for each element can then be obtained. FFEA extends FEA to include a thermal noise term as mesoscale objects are sufficiently small that thermal fluctuations drive their ability to explore conformational space (Oliver et al., 2013).

In FFEA globular biomolecules are modelled as 3D viscoelastic continuum solids (described as a blob model), therefore the mechanical response to stress not only depends on the strain (as in purely elastic bodies) but also on the strain rate (as in viscous fluids). In FFEA viscoelasticity is modelled through the Kelvin-Voigt constitutive model. At the mesoscale, dynamics are overdamped at the timescale of interest and therefore not effected by mass (inertialess), thus density is excluded (Solernou et al., 2018; Hanson et al., 2021). Trajectories are calculated using the Cauchy momentum equation of motion (equivalent to Newton’s 2nd law for a continuum) where inertia is neglected,

$$\nabla \cdot (\sigma^{v} + \sigma^{e} + \pi) + f = 0, \tag{3.1}$$

where σ^v is the viscous stress, σ^e is the elastic stress, π is the stochastic thermal stress, and f is the external force density. The viscous stress is proportional to velocity gradients in the system so the above equation is a differential equation from which the velocity of each node in the finite element mesh can be deduced. This leads to a set of differential equations for the positions of the nodes of tetrahedral elements.

Elongated biomolecules are modelled instead as 1D elastic Kirchoff rods (described as a rod model) because 3D continuum methods are less well-suited to representing long, thin objects (Welch et al., 2020). The elastic energy of rods is dependent on extension, torsion and bend. As above, the system is overdamped so that inertia can be neglected, thus velocities of the rod ends are found from a Brownian equation of motion (overdamped Langevin equation),

$$\frac{dr}{dt} = M \cdot (F + f), \qquad \textbf{(3.2)}$$

where $\mathrm{dr/dt}$ is the velocity of the rod node, M is the mobility tensor of the rod segment in the fluid medium, F is the internal elastic force, and f is the random force from thermal noise. In FFEA the mobility tensor is assumed to be isotropic and includes a viscosity term (Welch et al., 2020).

In order to model a protein in FFEA, the equilibrium molecular shape is obtained by cryo-EM or cryo-ET. Globular proteins or protein domains are represented as 3D tetrahedral finite element meshes (blobs), and elongated domains are represented as 1D elements connected by nodes (rods) (Oliver et al., 2013; Solernou et al., 2018; Welch et al., 2020). Material parameters must be applied to the system to define the elastic moduli as, unlike in atomistic simulations, these are not implicit. The bulk modulus (changes in volume) and shear modulus (shearing deformations) are input into blob simulations, and stiffness values for bend, stretch and twist are input into rod simulations. Material parameters can be inferred from experimental results, e.g. bending stiffness from ns-EM, cryo-EM data, or atomistic simulations of small protein domains (Richardson et al., 2014; Welch et al., 2020; Hanson et al., 2021). FFEA blobs and rods are anisotropic and inhomogeneous meaning different parameters can be applied to different elements to, for example, reflect different levels of stiffness throughout a molecule.

In the previous chapter I determined the structure of rigor Myo5a-S1 bound to F-actin and derived the material parameters of the lever from cryo-EM 3D classes. Here I used that data (Chapter 2) to generate both a FFEA blob model and rod model, to perform coarse-grained simulations of Myo5a equilibrium dynamics. I successfully simulated Myo5a-S1 lever conformations, using a FFEA blob model generated from my cryo-EM structure (Chapter 2). Additionally, in order to contextualise the cryo-EM derived lever properties within the dimer, I used a FFEA rod model to successfully simulate a dimeric construct of Myo5a (HMM).

Characterising the properties of the head-tail junction in the active state is essential for understanding walking mechanics (Andrecka et al., 2015; Hathcock et al., 2020). As there is currently no structure of the active dimeric molecule, I expressed Myo5a-HMM and used ns-EM to predict the stiffness at the head-tail junction in free molecule (off actin). Through simulation I was able to determine that changes in bending stiffness at the head-tail junction have little to no effect on the dimer. Finally, I further demonstrated that FFEA is a suitable method for simulating cytoskeletal motors and determined what is required to generate optimal and reliable coarse-grained models using FFEA. Overall, this work highlights how simulations provide a means to integrate disparate pieces of experimental data, in order to understand how they function together, and demonstrates a pipeline by which this can be done.

3.2 Materials and Methods

3.2.1 Myo5a-HMM production

3.2.1.1 Baculovirus production

The murine Myo5a-HMM plasmid, kindly provided by Derek Revill, was used for baculovirus generation. The plasmid consisted of the pFastBac1 baculovirus shuttle vector containing the murine Myo5a-HMM sequence (residues 1-1090), linked to a C-terminal FLAG tag (DYKDDDDK) for purification (Revill, 2014). This was transformed into MAX Efficiency DH10Bac Competent Cells according to the manufacturer’s instructions. Briefly, competent cells were thawed and 5 µL was transferred into a sterile 1.5 mL Eppendorf that had been pre-chilled on ice. 20 ng of DNA was added to the cells and incubated on ice for 10 mins. Cells were heat-shocked for 45 seconds in a 42 °C water bath, then placed on ice for 2 mins. 0.5 mL of room temperature S.O.C Medium (Super Optimal broth with Catabolite repression) was added and shaken at 225 rpm for 4 hrs at 37 °C. 50 µL was spread on Luria Agar plates containing 50 µg/mL kanamycin sulfate, 10 µg/mL tetracycline, 7 µg/mL gentamycin, 100 µg/mL bluo-gal and 40 µg/mL isopropyl β-D-1-thiogalactopyranoside (IPTG), and incubated for 24 hrs at 37 °C. Ten white colonies were selected and screened again via the same process to confirm their colour. Single white colonies were then selected for plasmid preparations.

Colony PCR was performed to confirm transposition of the construct into the bacmid DNA. 250 µL of 10 x Phusion® (New England Biolabs) master mix was prepared on to a final concentration of 1 x Phusion HF buffer, 10 µM dNTPs, 0.5 µM M13 forward primer, 0.5 µM M13 reverse primer, and 1.00 units/50 µL Phusion High-Fidelity DNA polymerase in nuclease-free MilliQ water. The master mix was divided into PCR tubes and a single colony was dipped into a corresponding tube before being used to inoculate an overnight culture. The initial denaturation of DNA was at 98 °C for 3 mins, followed by 30 cycles of denaturation at 98 °C for 10 s, annealing at 55 °C for 20 s, and extension at 72 °C for 4 mins. This was followed by a final extension at 72 °C for 10 mins. PCR products were then resolved by gel electrophoresis on a 1% agarose, 1 x TAE (0.4 M tris acetate, 0.01 M EDTA, pH 8.3), 1 µg/mL ethidium bromide gel. The presence of the Myo5a-HMM sequence was confirmed by the appearance of bands at ~3.3 kbp, using a 1 kb Plus DNA Ladder (New England BioLabs).

White colonies were grown overnight at 37 °C in 5 mL of Luria Broth containing 50 µg/mL kanamycin, 7 µg/mL gentamycin, and 10 µg/mL tetracycline. Bacmid DNA successfully verified with colony PCR, was isolated using the microcentrifuge protocol from as QIAprep spin Miniprep Kit (Qiagen). Briefly, the starter culture was centrifuged at 3000 x g for 15 mins at 4 °C. The supernatant was discarded and the pellet was then resuspended in 250 µL of Qiagen P1 resuspension buffer, followed by the addition of 250 µL of Qiagen P2 lysis buffer and mixing by inversion 4-6 times. 350 µL of N3 neutralisation buffer was then added and mixed by inversion 4-6 times. This was then centrifuged for 10 mins at 17000 x g in a MICRO STAR 17R (VWR) centrifuge. The supernatant was transferred to a new Eppendorf where an equal volume of pre-chilled isopropanol was added followed by mixing by inversion 4-6 times and incubation on ice for 15 mins. This was then centrifuged for 10 mins at 17000 x g in a MICRO STAR 17R (VWR) centrifuge. The supernatant was discarded and the pellet was washed once by resuspension in 70 % ethanol, followed by centrifugation for 10 mins at 17000 x g in a MICRO STAR 17R (VWR) centrifuge. The supernatant was removed via pipette and the DNA pellet left to air dry for 10 mins, followed by resuspension in 30 µL nuclease-free MilliQ water (boiled and 0.22 µm filtered). The concentration was determined using a NanoDrop (Thermo).

Bacmid DNA was transfected into Sf9 insect cells. 300 µL of 0.083 mg/mL bacmid DNA in sterile PBS (Phosphate-buffered saline) was mixed with 300 µL polyethylenimine (1 mg/mL) (Sigma Aldrich) and incubated at room temperature for 10 mins. The mixture was then added to 50 mL of Insect-Xpress serum free medium containing L-glutamine (SFX media) (Lonza) containing 0.5 x 10^6 cells/mL and incubated for 5 days (27 °C, 210 rpm). To harvest the baculovirus, the cell suspension was centrifuged using a Beckman Coulter Avanti J25 centrifuge at 3000 x g for 15 mins. The supernatant was then transferred to a light resistant 50 mL falcon tube and stored at 4 °C (P1 stock).

To determine the viral titre, purified virus was diluted 1/500 in SFX medium and 125 µL was added to each well of the first column of a 96 well plate. Consecutive 5 times serial dilutions in SFX media were then performed across the 96 well plate from columns 1-10, such that each well contained a total of 100 µL, and each column was 1/5th the concentration of the previous column. 100 µL of SFX medium was added to columns 11-12 as a control. 100 µL of Sf9 insect cells (7.4 x 10^4 cells per well) was then added to each well across all 12 columns and left to incubate for 5 days at 27°C. The viral titre was then determined using the end-point dilution method (Hopkins and Esposito, 2009), and used to calculate the number of plaque forming units per mL (PFU/mL).

Recombinant baculovirus was amplified from P1 stage to P2 stage according to the Bac-to-Bac protocol (Thermo). To obtain P2 stock, 8 mL (Equation (**3.3**)) of the initial P1 stock (5.6 x 10^6 PFU/mL) was used to infect Sf9 cells (1.5 x 10^6 cells/mL) in 300 mL SFX media in a 500 mL Erlenmeyer flask with a MOI (multiplicity of infection) of 0.1.

$$\text{Inoculum (mL)} = \frac{MOI \times cell\ count\ (cells/mL) \times culture\ volume\ (mL)}{viral\ titre\ (PFU/mL)} \quad \textbf{(3.3)}$$

The viral amplification was then incubated for 14 days at 27 °C with 210 rpm shaking. The cell suspension was centrifuged using a Beckman Coulter Avanti J25 centrifuge at 3000 x g for 15 mins. The supernatant was harvested and titred as previously, giving a final titre of 3.3 x 10^8 PFU/mL (P2).

3.2.1.2 Sf9 expression

Myo5a-HMM was co-expressed with Calm2, Unc45B and HSP90. A Calm2 baculovirus (3.8 x 10^8 PFU/mL), and a Unc45B/Hsp90 co-chaperone baculovirus (1.1 x 10^9 PFU/mL) produced by Glenn Carrington were used. 23 mL Myo5a-HMM baculovirus, 20 mL Calm2 baculovirus and 7 mL Unc45B/Hsp90 co-chaperone baculovirus were used to infect Sf9 cells (1.5 x 10^6 cells/mL) in 950 mL SFX media in a 3 L Erlenmeyer flask with MOIs of 5 (Equation (**3.3**)). This was then incubated for 72 hrs at 27 °C with 120 rpm shaking. Following expression, the cells were pelleted by centrifugation at 3000 x g for 15 mins using a MegaFuge 16R centrifuge (Thermo) and the supernatant discarded. The pellet was then frozen at -80 °C for future use.

3.2.1.3 Purification

The cell pellet was thawed and mechanically homogenised in extraction buffer (10 mM MOPS, 200 mM NaCl, 10 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT, 0.1 mM PMSF, cOmplete™ EDTA-free Protease Inhibitor Cocktail (1 tablet/50 mL) (Roche), 10 µM leupeptin, pH 7.3) to total volume of total volume of 120 mL, using a pre-chilled ground glass homogeniser. The homogenised mixture was then sonicated using a Soniprep 150 MSE (SANYO) sonicator using a cycle of 10 s on/10 s off for 3 mins at 50% power. 2 mM ATP was then gradually added to the homogenate on ice whilst stirring for ~20 mins. The cellular lysate was centrifuged at 30000 x g for 30 mins at 4 °C using a Beckman Coulter Avanti J25 centrifuge. A sample of both the soluble and insoluble fraction was collected for SDS-PAGE. The supernatant was then split into 50 mL falcons, and FLAG ANTI-FLAG M2 Affinity Gel (Sigma) equilibrated in PBS was added

(1-3 mL resin/100 mL lysate). The mixture was then incubated at 4 °C for ~12 hours on a rocker.

To pellet the resin the mixture was centrifuged at 400 x g for 2 mins at 4 °C using a MegaFuge 16R centrifuge (Thermo). The supernatant was collected as a sample for SDS-PAGE. Each resin was then washed with 50 mL Mary’s buffer (10 mM MOPS, 0.1 mM EGTA, 0.5 M NaCl, 1 mM DTT, pH 7.2) with 1mM ATP and 5 mM $MgCl_2$ by centrifugation at 400 x g for 2 mins at 4 °C using a MegaFuge 16R centrifuge (Thermo), and again with 50 mL Mary’s buffer, followed by washing twice with 50 mL HMM buffer (10 mM MOPS, 0.1 mM EGTA, 1 mM DTT, pH 7.2), and resuspension in 3-5mL of HMM buffer and loaded onto a column. The supernatant was collected as samples for SDS-PAGE after each wash.

To elute the protein from the FLAG affinity column, 1 mL fractions of buffer (10 mM MOPS, 0.1 mM EGTA, 0.1 M NaCl, pH 7.2) containing 0.5 mg/mL FLAG-peptide (Cohesion Biosciences) were applied to each resin. 10 x 1 mL elution fractions were collected followed by a final elution of 20 mL saved for further analysis. The FLAG resin was then washed in an equal volume of 0.1 M glycine (pH 2.5) ~5 times, then with ~300 mL 1 x PBS with 3 mM NaN_3, and finally stored in 1 x PBS with 3 mM NaN_3. Gel electrophoresis was used to determine protein containing fractions. For gel electrophoresis, samples were denatured at 95 °C for 5 mins before loading. 10 µL of BLUeye Pre-Stained Protein Ladder (Geneflow) was loaded into the first well of a precast 4-20% gradient polyacrylamide (BioRad) and 10 µL of protein sample were loaded into the following wells. Gels were run at a constant voltage of 100 V for 45 min in running buffer (3 g/L Tris base, 14.4 g/L Glycine, 1% SDS). The gel was stained using Instant Blue Coomassie stain (Expedeon) on a rocker for > 30 mins.

Fractions that contained Myo5a-HMM were then combined and dialysed using dialysis tubing with a molecular weight cut off of 3.5 kDa (Gebaflex Maxi, Generon). This was left for ~12 hours in 2 L of HMM buffer with 0.5 M KCl at 4 °C, with continual stirring using a magnetic stirrer. The concentration of the purified HMM was determined as 7.2 µM from the absorbance at 280 nm and 320 nm measured using a NanoPhotometer NP80 spectrophotometer with a 1 cm path length. Values were 1.749 and 0.012 respectively (Equation (**3.4**)). 1.5

x A_{320} is used to correct for light scattering. The molar extinction coefficient (120000) at A_{280} was taken from (Forgacs et al., 2009).

$$\text{HMM concentration (M)}$$
$$= 0.5 \times \frac{A_{280} - (1.5 \times A_{320})}{molar\ extinction\ coefficient\ per\ head\ (M^{-1}cm^{-1}) \times\ path\ length\ (cm)} \quad \textbf{(3.4)}$$
$$= 0.5 \times \frac{1.749 - (1.5 \times 0.012)}{120000 \times 1} \approx 7.2\ \mu M$$

The protein product was then drop frozen in LN_2 (~20 µL beads) and stored in a cryo-vial in LN_2 for later use.

3.2.2 Generating head-tail junction parameters from ns-EM

Carbon-coated grids, produced in-house by Martin Fuller (Astbury Biostructure Laboratory, University of Leeds, UK), were negatively glow discharged for 30 s at 0.38 mBar and 15 mAmp (PELCO easiGlow discharge unit, Ted Pella Inc., USA). Myo5a-HMM was mixed with an equal volume of F-actin (polymerisation methods in Chapter 2) and ATP, to give a final concentration of 125 nM, 12 nM and 150 nM, respectively, in 100 mM KCl, 0.1 mM EGTA, 0.25 mM $MgCl_2$, and 10 mM MOPs at pH 7.0. This was immediately applied to the glow-discharged, carbon-coated grids and stained using 1% uranyl acetate using the flicking method (Scarff et al., 2018). Images were recorded on an FEI Tecnai F20 at 29000 x magnification.

To calculate the stiffness at the head-tail junction, the local angles between levers at the head-tail junction, in particles identified as free HMM in the ns-EM micrograph, were measured using ImageJ (Figure 3.1) (Schindelin et al., 2012). The variance (σ^2) in lever angles at the head-tail junction was then used to calculate the spring constant, assuming bending occurs at a single hinge point, using Equation **(3.5)**,

$$\text{spring constant} = \frac{k_B T}{\sigma^2}, \quad \textbf{(3.5)}$$

where $k_B = 1.380649 \times 10^{-23}$ N m K^{-1} (Boltzmann constant), T = 293 K (temperature of grid making in Kelvin), and σ^2 is the variance of the angle at the head-tail junction.

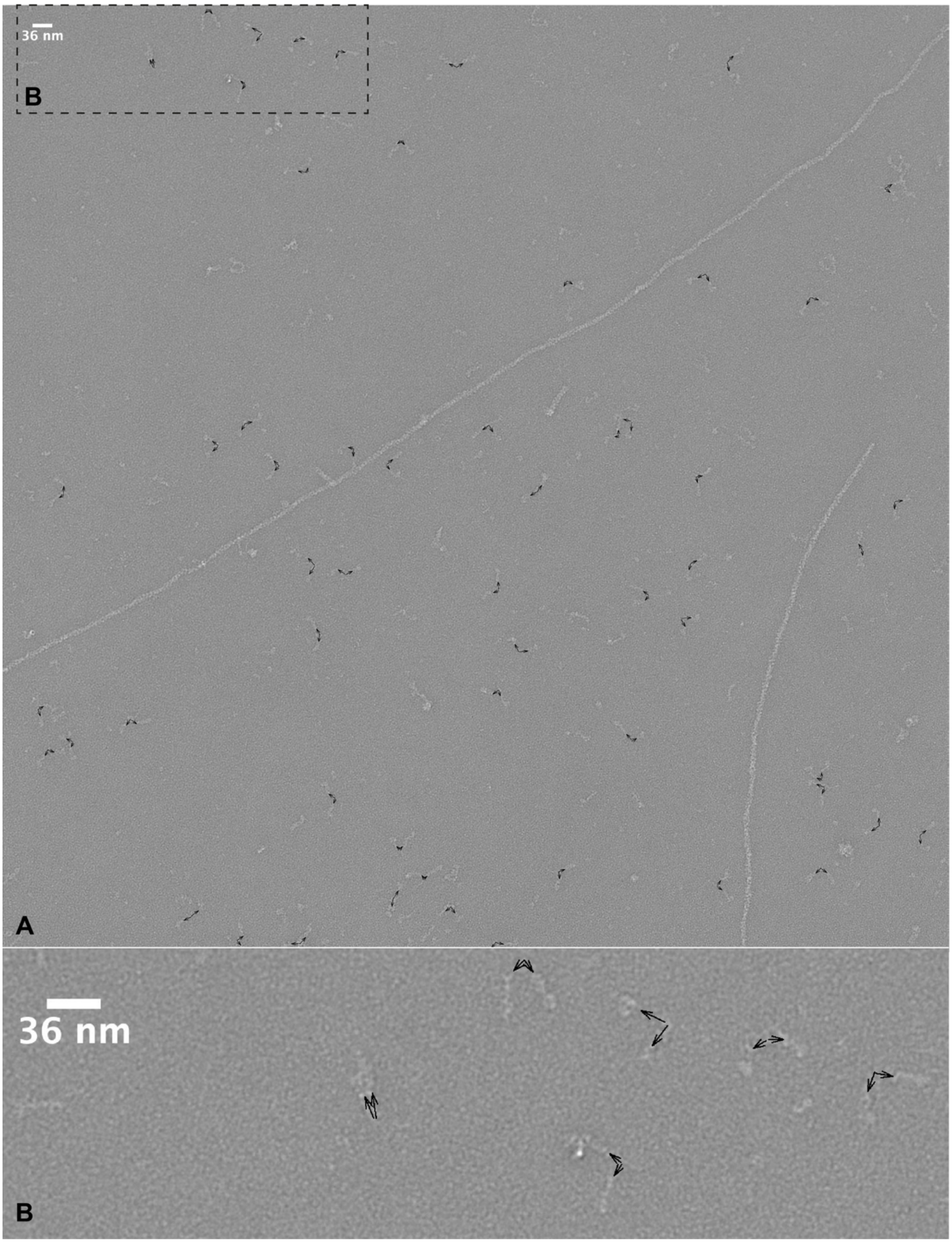


Figure 3.1: Calculating angle between levers in HMM

***A:** ns-EM micrograph of free murine Myo5a-HMM. Filaments are rabbit skeletal Factin. Scale bar 36 nm. Arrows showing the angle between levers were drawn with ImageJ (Schindelin et al., 2012). The angles between the arrows were also calculated using ImageJ (Schindelin et al., 2012). **B:** zoomed section.*

3.2.3 Myo5a-S1 blob model mesh generation and parametrisation

To generate 3D finite element meshes for FFEA blob models, the Myo5a-S1 cryo-EM map (Chapter 2) was pre-processed before mesh generation in 2 ways. The first was by gaussian filtering the map in ChimeraX (5 SDs) (Pettersen et al., 2021) (Figure 3.2A) to smooth the surface. The second was by rescaling the map to a voxel size of 15 Å using Relion3.1 (Zivanov et al., 2018) (Figure 3.2C) to coarsen the map, and then gaussian filtering in ChimeraX (5 SDs) (Pettersen et al., 2021) to smooth the surface. A triangular surface mesh was then generated for each pre-processed map, and coarsened to 15 Å using FFEA tools (Solernou et al., 2018). TetGen (Si, 2015) was then used to generate a volumetric tetrahedral mesh (Figure 3.2B & D). The TetGen mesh was then converted into FFEA format using FFEA tools (Solernou et al., 2018). Details of the FFEA tools commands can be found in the FFEA tutorial (https://ffea.readthedocs.io/en/latest/Tutorial.html).

FFEA also requires material parameters to be assigned. Molecules are modelled as a viscoelastic material characterised by two elastic moduli (bulk and shear modulus), and two viscosities for bulk and shear deformations. Values for the elastic material parameters were calculated by assuming that the lever flexes uniformly, and so can be approximated as a cantilever. The Young’s modulus can therefore be calculated from the bending stiffness derived via cryo-EM in Chapter 2 (0.76 pN/nm) by rearranging Equation (**3.6**),

$$\text{bending stiffness } (\kappa) = \frac{3EI}{L^3}, \qquad \textbf{(3.6)}$$

where L is cantilever length (20.3 nm, the mean lever length in the cryo-EM classes), E the Young’s modulus (Equation (**3.10**)) and I is the moment of inertia (Equation (**3.9**)). Hence by measuring the lever length and radius, first the Young’s modulus can be obtained, and then the bulk and shear moduli by assuming a value for the Poisson’s ratio.

First, the measured radii were averaged so that thin regions had an increased weighting, as the flexibility at thin regions dominates, using Equations (**3.7**) and (**3.8**),

$$\text{weighted average (A)} = \frac{1}{n} \times \sum \frac{1}{r^4}\,, \tag{3.7}$$

$$\text{radius (R)} = {}^{1}\!/_{A}^{\frac{1}{4}} \approx 1.4\,nm\,, \tag{3.8}$$

where n is the number of distances measured and r is half the distance measured across the short axis of the lever (Figure 3.3).

The area moment of inertia was then calculated using Equation (**3.9**) with the radius from Equation (3.9), approximating the lever as a solid cylinder,

$$\text{Area moment of inertia (I)} = \frac{\pi R^4}{4} \approx 3.0\,nm^4. \tag{3.9}$$

Equation (**3.6**) can then be rearranged to give the Young's modulus using Equation (**3.10**),

$$\text{Young's Modulus (E)} = \frac{\kappa L^3}{3I} \approx 700\,MPa\,, \tag{3.10}$$

where κ is the spring constant (0.76 pN/nm), L is the mean length (20.3 nm) of the lever in the cryo-EM 3D classes (Chapter 2), and I is the area moment of inertia.

The shear and bulk moduli could then be obtained using Equations (**3.11**) and (**3.12**),

$$Shear\ modulus\ (G) = \frac{E}{2(1+v)} \approx 2.6\ \times\ 10^2\ MPa, \tag{3.11}$$

$$Bulk\ modulus\ (K) = \frac{E}{3(1-2v)} \approx 7.7\ \times\ 10^3\ MPa, \tag{3.12}$$

where E is the Young's modulus (Equation (3.10)) and v is the Poisson's ratio (0.35). The Poisson's ratio relates the bulk and shear modulus, it is the ratio of compression in transverse axes (yz) to the compression in the x-axis, typically materials have a value between 0-0.5. Materials with a Poisson's ratio of 0.5 are incompressible, e.g. rubber. It is hard to determine the exact Poisson ratio for proteins as it is hard to measure their volumes as they undergo deformation. However, it is thought proteins have a Poisson ratio between 0.3-0.4 similar to glassy polymers (Kharakoz, 2000), which has also found to be a suitable value in previous FFEA studies (Oliver, 2013). The viscosity was set to the viscosity of water (10^{-3} Pa s). Note that the value of viscosity does not affect the range of motion, only the rate at which this is explored. Therefore if the viscosity was increased, to for instance that of the cytoplasm, it would not affect the conformational space the lever explores but how quickly it explores it.

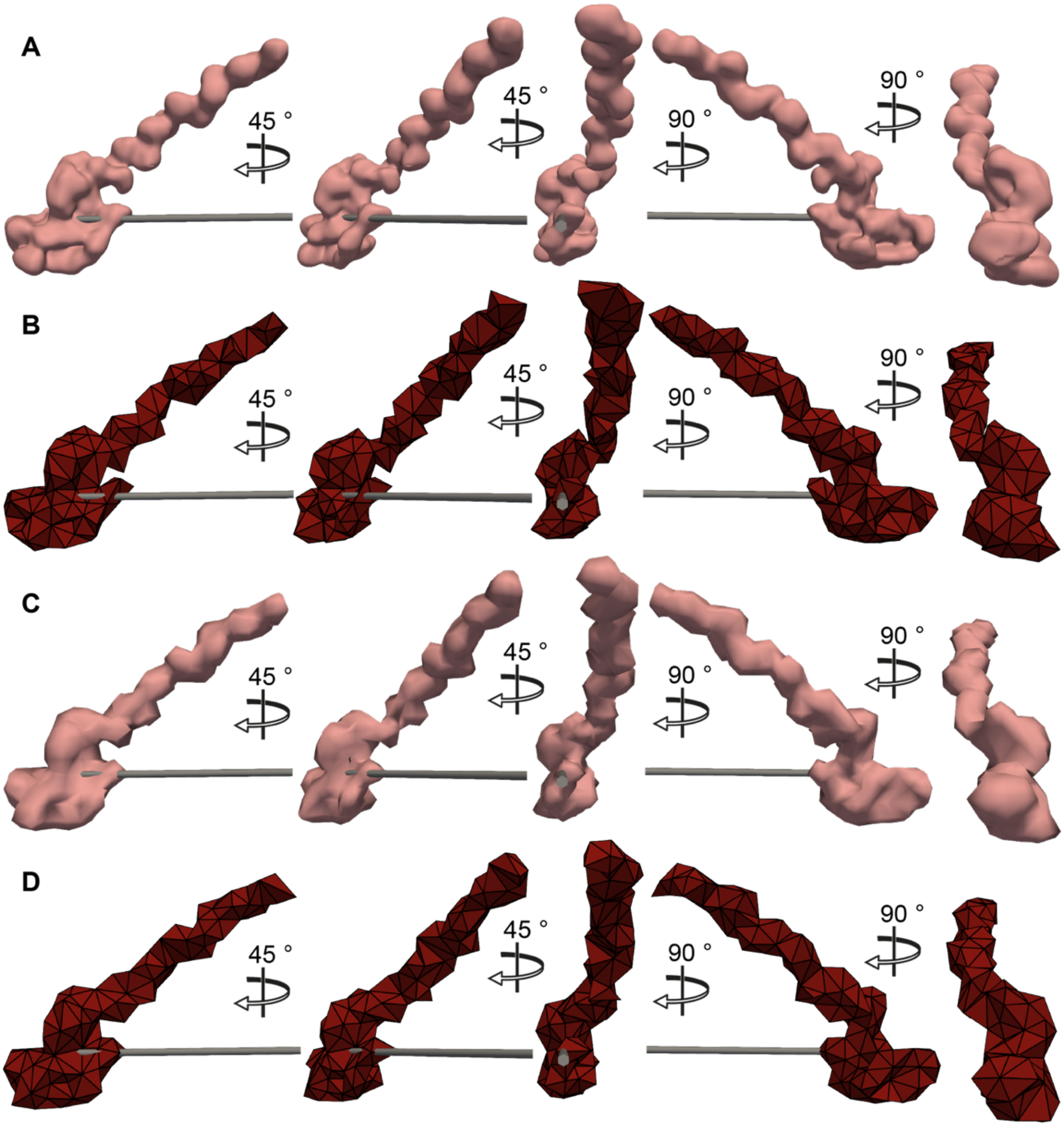


Figure 3.2: Pre-processed cryo-EM maps and 3D meshes

A: *Gaussian filtered (SD 5) cryo-EM map of rigor Myo5a-S1 bound to F-actin.* ***B:*** *3D tetrahedral finite element mesh of* ***A****.* ***C:*** *Cryo-EM map of rigor Myo5a-S1 bound to F-actin, rescaled to a voxel size of 15 Å and gaussian filtered (SD 5).* ***D:*** *3D tetrahedral finite element mesh of* ***C****.*

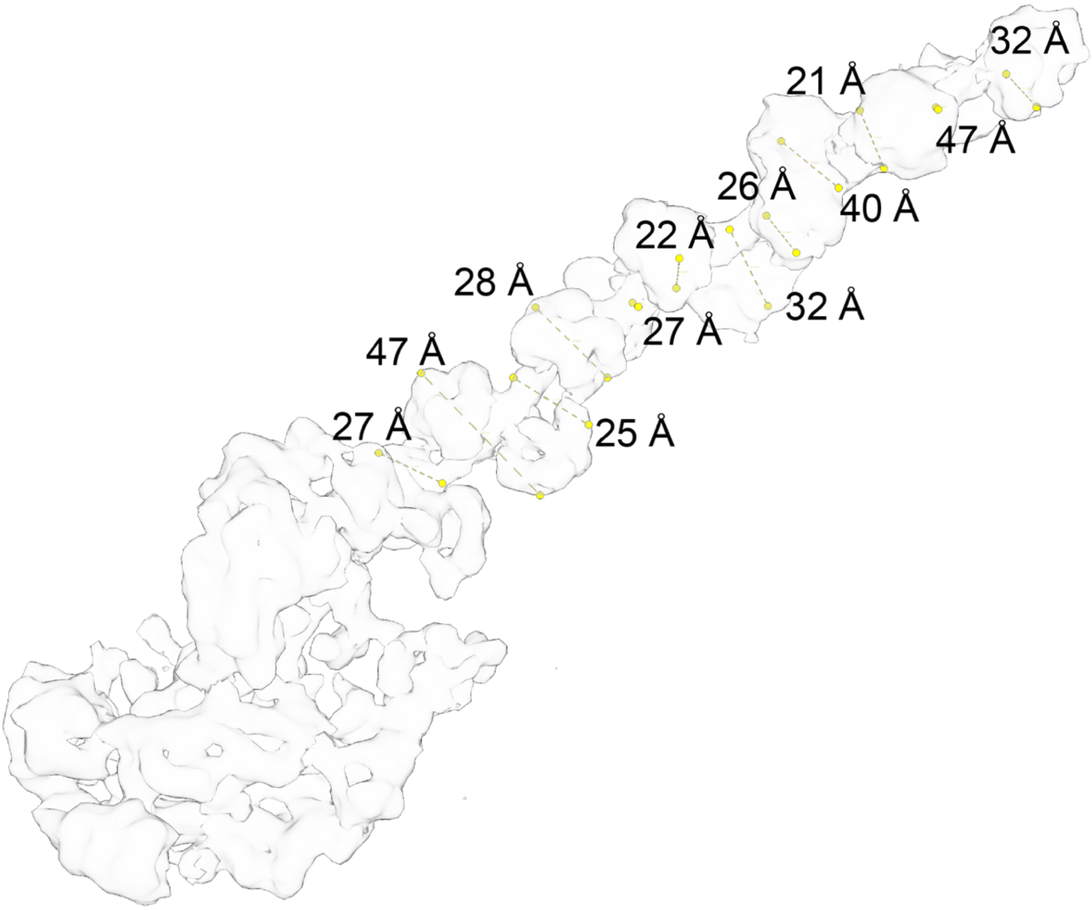


Figure 3.3: Measuring the diameter of the lever

The diameter of the lever in the cryo-EM map (contour level: 0.25) was measured at intervals along its length in ChimeraX (Pettersen et al., 2021).

3.2.4 HMM rod model rod generation and parametrisation

To generate a FFEA rod model, the equilibrium rod model from Chapter 4 was used. The cryo-EM derived material parameters (Chapter 2) for twist and bend were converted into the correct units (N m^2) for FFEA simulation. This was done by multiplying the spring constant for the subdomain connection (Chapter 2) by the average length of the rod elements either side of the corresponding node. These were then applied to the equivalent nodes in the rod model (Table 3.1).

3.2.5 Simulations

FFEA simulations were performed using FFEA version 2.7.2 (Solernou et al., 2018), on 4 CPUs (details in Table 3.1), using the CG_nomass (inertialess) solver. For blob simulations the shear and bulk viscosities were set to 1 mPa.s. Simulations were run at 24.5 °C ($K_bT = 4.11 \times 10^{-21}$ J).

3.2.6 Calculating model emergent material properties

For the FFEA blob model of Myo5a-S1, the emergent material properties were calculated as in Chapter 2. However, the coordinates for the Cα atoms, were replaced with the coordinates of the nodes closest to the Cα atoms in the modelled class (Figure 3.4). To calculate the cantilever bending stiffness, the displacement from the mean was calculated for every frame of the trajectory (Figure 3.5). The variance (σ^2) was then used in Equation **(2.3)** to calculate the cantilever bending stiffness. Calculation of the bending stiffnesses at hinges between subdomain pairs was completed as in Chapter 2 using Equation **(3.5)**, however instead of comparing conformational variation between cryo-EM classes, the conformational variation between frames within the FFEA trajectory was compared.

For the FFEA rod model of Myo5a-HMM the bending stiffness at the subdomain connections was calculated using Equation **(3.5)**. Where σ^2 was the variance in the angle between the elements either side of the node across every frame of the trajectory.

The error for the bending stiffness calculation at each subdomain connection in the rod model was estimated using the SD of bending stiffnesses calculated from randomly assigned subdomain angles. First the angles calculated at a subdomain in each frame were sorted into ascending order. A random value for the angle was then generated for each frame between the angle values either side of it in the sorted list. The ranges for values at the start and end of the list were calculated using the value ± the distance from its single neighbour. The variance (σ^2) was calculated for these random angles, and used to calculate a bending stiffness using Equation **(3.5)**. This was repeated 1000 times and the SD of the random bending stiffnesses was quoted as the error.

The overall bending stiffness of the rod model was calculated using Equation **(2.3**). Where σ^2 was the variance in the distance between the start and end nodes of the rod.

To obtain the overall bending stiffness of a single lever from the distance between the rod model start and end nodes, it was assumed that each lever contributes to the variance equally. As the levers can bend in all directions the distance recorded is a projection of their motion along the end-to-end vector. Therefore, the measured distance can be equated to the mean conformation (L_{base}) plus the change in distance projected along the end-to-end vector using the Equation (**3.13**),

$$\text{measured distance (L)} \approx L_{base} + x_2 \cos\theta - x_1 \cos\theta \, , \qquad \textbf{(3.13)}$$

where L_{base} is the theoretical mean conformation, x_n is the displacement of the end of lever n from the mean, and θ is the angle between the direction of motion and the end-to-end vector (Figure 3.6).

The SD is therefore the difference between the measured distance and the theoretical mean, shown in Equation (**3.14**),

$$\text{change in distance } (\Delta\text{L}) = L - L_{base} = x_2 \cos\theta - x_1 \cos\theta \, , \qquad \textbf{(3.14)}$$

The variance is the SD squared, shown in Equation (**3.15**),

$$\text{variance both levers } (\langle \Delta\text{L}^2 \rangle) = \cos^2\theta [\langle x_2^2 \rangle + \langle x_1^2 \rangle] \, , \qquad \textbf{(3.15)}$$

assuming the x_2 and x_1 are varying independently.

Therefore, the variance of a single lever assuming they have equal variance can be calculated by rearranging Equation (**3.15**) to Equation (**3.16**),

$$\text{variance single lever } (\sigma^2) = \langle x_2^2 \rangle = \langle x_1^2 \rangle = \frac{\Delta L^2}{2\cos^2\theta}, \qquad \textbf{(3.16)}$$

where ΔL^2 is the variance in the distance between the start and end nodes of the rod, and θ is half the average angle between the levers (vectors drawn from head-tail junction node to 1st, and last node) (Figure 3.6D). This variance (σ^2) can then be used to obtain the bending stiffness using Equation **(2.3**).

3.2.7 Calculating the angle between levers in rod simulations

To determine if the flexibility from CaM4-6 dampened constraint at the head-tail junction, the angle was measured between vectors drawn from the head-tail junction to CaM6 in each lever (rod nodes 7-6 and 7-8), and also from the head-tail junction to CaM4 in each lever (rod nodes 7-4 and 7-10) for every frame of the trajectory. This was also done in the xy plane only, to compare the effect of using a 2D projection, as opposed to a 3D reconstruction, to calculate bending stiffnesses.

3.2.8 Principal component analysis

Principal component analysis (PCA) was used to assess the differences in behaviour between the gaussian filtered and rescaled then gaussian filtered meshes in the FFEA blob simulations. PCA divides the dynamics of a system into its constituent dominant modes of motion. PCA was performed using pyPcazip (Shkurti et al., 2016). In order to use pyPcazip in conjunction with FFEA, trajectories had to be converted into pseudo atomic trajectory files by converting the FFEA nodes into Cα atoms using the FFEA tools makepseudopdb tool (Solernou et al., 2018). pyPcazip calculates the averaged correlations between node coordinates in the trajectory to create a co-variance matrix, which is diagonalised into a set of eigenvectors and eigenvalues. The eigenvectors describe the direction of each mode of motion, and the eigenvalues describe the proportion each mode of motion contributes to the overall dynamics.

Table 3.1 - FFEA simulation conditions. HT = head-tail junction.

Simulation	Length (s)	Timestep (s)	Shear modulus (N·m^{-2})	Bulk modulus (N·m^{-2})	Stretch constant (N)	Radius (m)	Twist constant (N·m^{2})	Bend constant (N·m^{2})	Node
Myo5a-S1 blob, gaussian	1.5×10^{-4}	1×10^{-13}	2.6×10^{8}	7.7×10^{8}					
Myo5a-S1 blob, rescaled + gaussian	1.5×10^{-4}	1×10^{-13}	2.6×10^{8}	7.7×10^{8}					
Myo5a-HMM rod, unconstrained head-tail junction	5×10^{-4}	1×10^{-11}			3.5×10^{-11}	1.85×10^{-9}	1.37×10^{-26}	1.17×10^{-26}	0 & 14
							5.84×10^{-27}	3.19×10^{-27}	1 & 13
							5.71×10^{-27}	7.20×10^{-27}	2 & 12
							7.48×10^{-27}	5.08×10^{-27}	3 & 11
							3.77×10^{-27}	2.35×10^{-27}	4 & 10
							2.57×10^{-28}	4.82×10^{-28}	5 & 9
							5.51×10^{-28}	6.16×10^{-28}	6 & 8
							1.07×10^{-26}	7.04×10^{-29}	7 (HT)
Myo5a-HMM rod, constrained head-tail junction	5×10^{-4}	1×10^{-11}			3.5×10^{-11}	1.85×10^{-9}	1.37×10^{-26}	1.17×10^{-26}	0 & 14
							5.84×10^{-27}	3.19×10^{-27}	1 & 13
							5.71×10^{-27}	7.20×10^{-27}	2 & 12
							7.48×10^{-27}	5.08×10^{-27}	3 & 11
							3.77×10^{-27}	2.35×10^{-27}	4 & 10
							2.57×10^{-28}	4.82×10^{-28}	5 & 9
							5.51×10^{-28}	6.16×10^{-28}	6 & 8
							1.07×10^{-26}	9.09×10^{-27}	7 (HT)

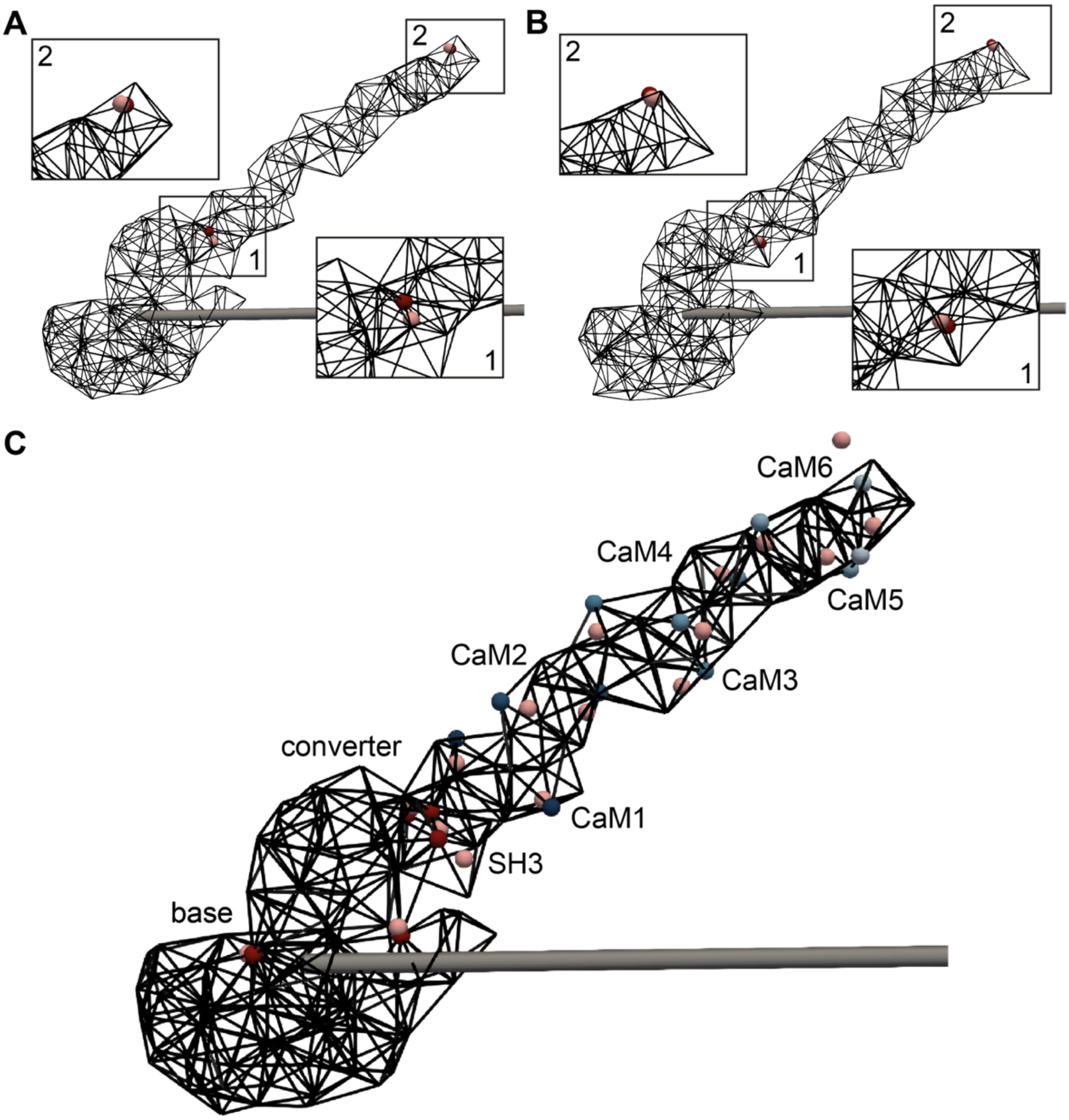


Figure 3.4: Nodes in mesh closest to vector coordinates in cryo-EM model

A-B: *Gaussian filtered (**A**) and rescaled then gaussian filtered (**B**) meshes with lever start (**1**) and end (**2**) coordinates highlighted. Red spheres = lever start and end nodes. Pink spheres = cryo-EM model lever vector start and end coordinates (Chapter 2).* ***C:*** *gaussian filtered mesh with lever subdomain node coordinates equivalent to lever subdomain coordinates in the cryo-EM model shown as spheres and coloured by domain as in Chapter 2 (**Figure 2.2**). Pink spheres show the cryo-EM model lever subdomain coordinates Chapter 2 (**Figure 2.2**). Grey cylinder = F-actin vector (1st to 13th subunit).*

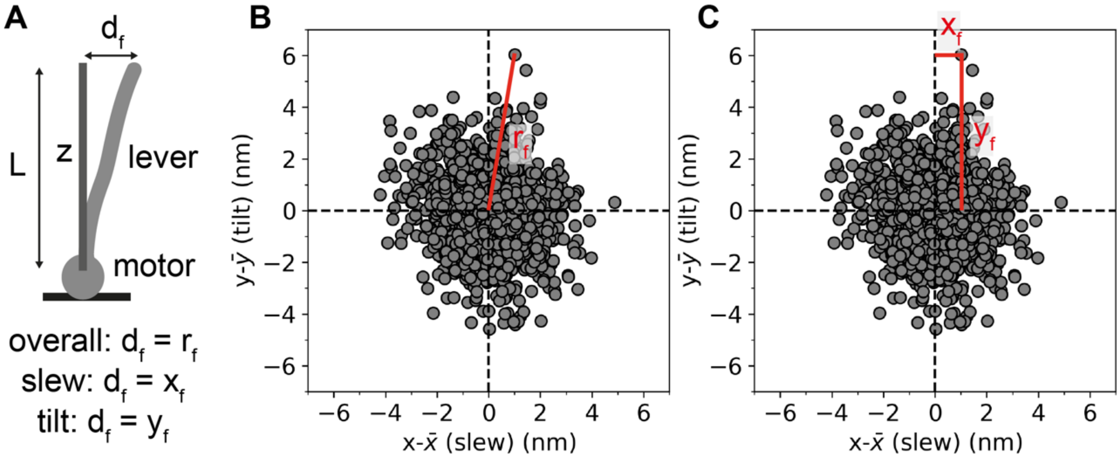


Figure 3.5: Calculating lever displacement for cantilever bending stiffness

A: *schematic of cantilever-type bending of the lever of Myo5a. d_f is the displacement of the lever from the mean position (z) in each frame (f) of the FFEA blob simulation. For the overall emergent bending stiffness the displacement (d_f) was calculated using r_f (**B**). To calculate the bending stiffness in each direction, tilt and slew, the displacement (d_f) was calculated using x_f and y_f, respectively (**C**). L is the mean length of the lever. z is the mean of the lever vectors across the trajectory (converter to CaM6) used as the z axis in **B** and **C**. **B:** demonstration of how r_f were calculated for each FFEA blob trajectory frame (r_f) and the distribution of end points from Figure 3.7D. **C:** demonstration of how x_f and y_f were calculated for each FFEA blob trajectory frame (x_f and y_f, respectively) and the distribution of end points from Figure 3.7D.*

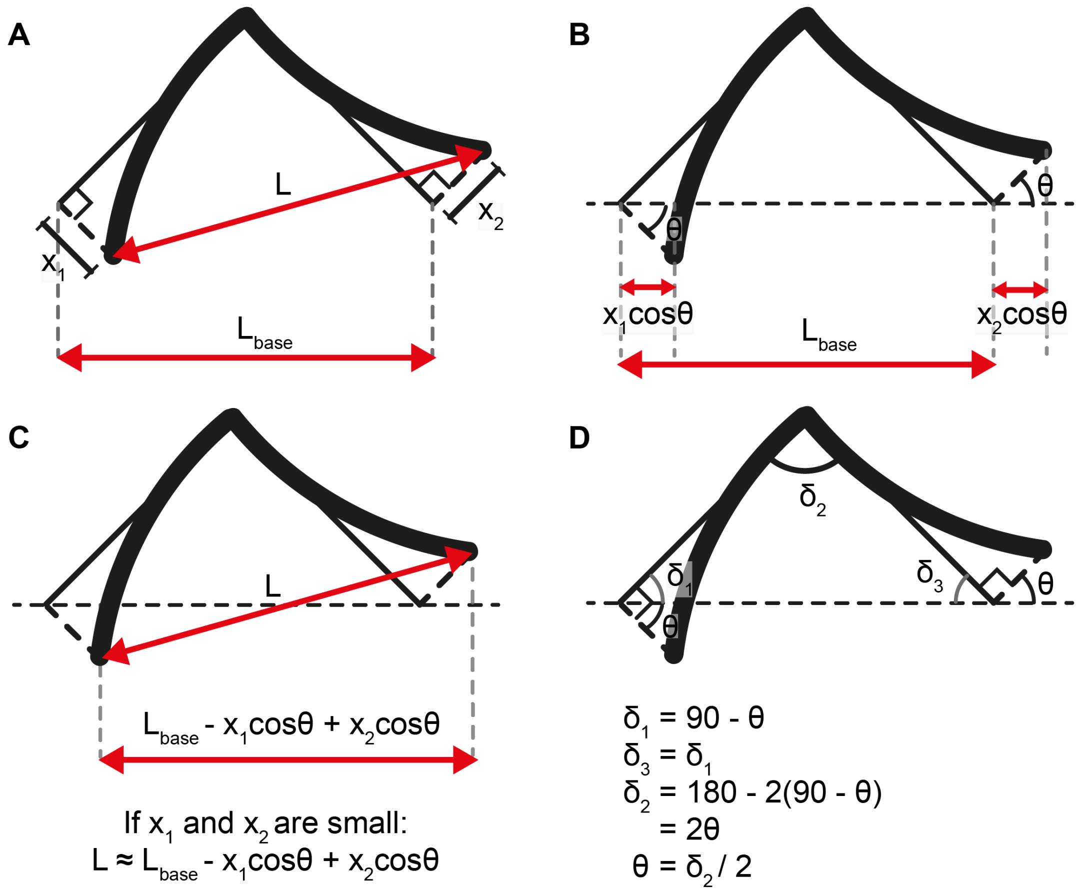


Figure 3.6: Calculating the emergent lever bending stiffness from the FFEA rod simulation of the dimer

A: *Schematic representation of the FFEA rod model of the dimer (thick curved line). Assuming the head-tail junction is stable and the lever flexes as a cantilever from the head-tail junction, the measured distance (L) between the converter domains of each head is a projection of the true variance (x_1 and x_2) of the each lever from the mean (thin solid line). L_{base} is the distance between the converter domains of each head in the mean conformation.* ***B-C:*** *If the displacement of the levers from the mean position is small, using trigonometry the measured distance (L) can be approximately equated to $L_{base} - x_1\cos\theta + x_2\cos\theta$, where θ is the angle between the vector connecting the converter domains in the mean conformation, and the vector connecting the converter domain in the current lever conformation and the mean conformation.* ***D:*** *Using trigonometry θ can be equated to $\delta_2/2$. δ_1 and δ_3 are the angles at the base of the triangle representing the mean conformation, as the levers are equal length the triangle is an isosceles therefore δ_1 and δ_3 are equal. θ is half the average angle between the levers at the head-tail junction.*

3.3 Results

3.3.1 FFEA simulation of Myo5a-S1 reproduces cryo-EM data

The FFEA blob model of actin bound Myo5a-S1 in the rigor state reproduced global lever properties derived by cryo-EM (Chapter 2). Firstly, the range of distances in lever position from the mean in the simulation was approximately equivalent to that seen in the cryo-EM 3D classes, -5 to 6 and -4 to 4 for tilt, and -5 to 5 and -3 to 5 for slew, respectively (Figure 3.7A&D). Secondly, the lever in the FFEA simulation fluctuated about the same conformational space as the cryo-EM 3D classes. When the vectors representing the lever conformations in the simulation were plotted in the cryo-EM coordinate system, they occupied a similar conformational space (Figure 3.7A&B). This demonstrates the FFEA blob model has similar global lever properties to the cryo-EM data, which gives confidence in the shape of the mesh and the material parameters assigned.

However, despite these conformational similarities there are key differences in the emergent material properties derived from the FFEA blob model compared with those derived the cryo-EM 3D classes. Firstly, the emergent cantilever bending stiffness of the lever derived from the FFEA blob simulation is much stiffer than the cantilever bending stiffness derived by cryo-EM used to parametrised it (1.7 pN/nm and 0.76 pN/nm, respectively). This is a product of the coarseness of the mesh. As the elements are big (~15 Å maximum) there are a small number of linear elements across the width of the lever, which causes stiffer emergent material properties than those that were assigned (Oliver et al., 2013). Doubling the number of elements has been shown to half the error, therefore if the ratio κ_{sim}/κ_{EM} is 2.24 (where κ is the cantilever bending stiffness (pN/nm)), a maximum element size of ~7 Å would be necessary to reduce the error sufficiently to reproduce the assigned stiffness.

Secondly, unlike that observed for the cryo-EM data, the global bending stiffness of the lever in the FFEA blob model is mildly anisotropic. In the FFEA blob model the bending stiffness in the direction of the F-actin longitudinal axis (tilt) is 20% more flexible than across the F-actin short axis (slew), 1.6 pN/nm and 1.9 pN/nm respectively. This anisotropy could also be related to the error created by the coarseness of the mesh (124% percentage error between 0.76

(expected) and 1.7 pN/nm (actual)), for instance the lever mesh may be coarser in one direction than the other. Finally, the material properties at the intra-lever connections are not representative of the cryo-EM data (Figure 3.8A&C). The overall stiffness is ~1000 pN nm at all connections, which is to be expected as the material parameters were applied isotropically across the mesh. These properties demonstrate further complexity is required to more accurately model Myo5a-S1 in FFEA.

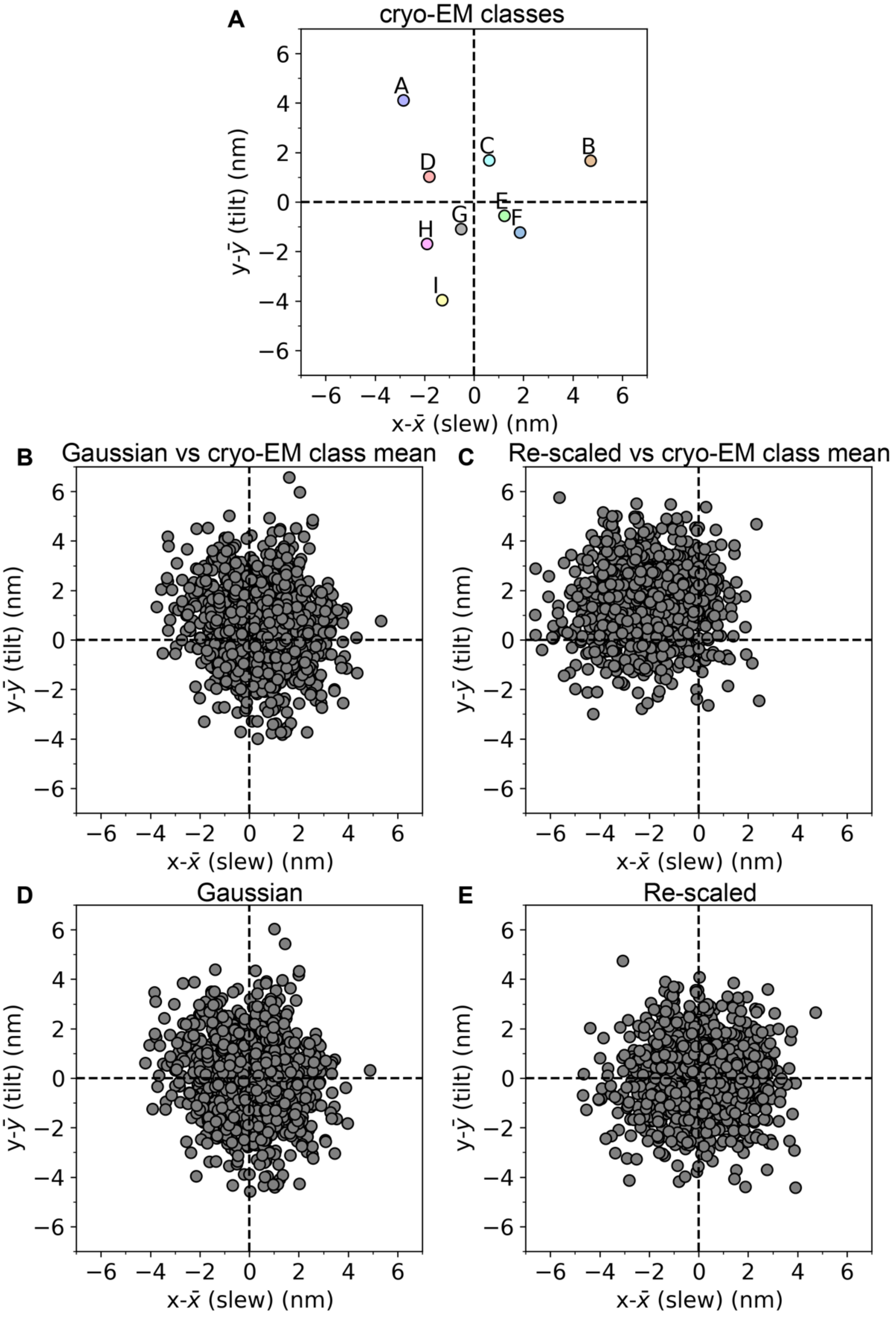


Figure 3.7: Lever end point analysis

A: *Plot of displacement of lever end in cryo-EM 3D classes from the mean (z) (see Chapter 2).* ***B-C:*** *Plots of displacement of lever end in the gaussian filtered (**B**) and re-scaled then gaussian filtered (**C**) FFEA blob simulations from the cryo-EM mean (z).* ***D-E:*** *Plots of displacement in lever end in the gaussian filtered (**D**) and re-scaled then gaussian filtered (**E**) FFEA blob simulations from their means (z). Changes in x represent motion across the F-actin short axis (slew), changes in y represent motion along the F-actin longitudinal axis (tilt).*

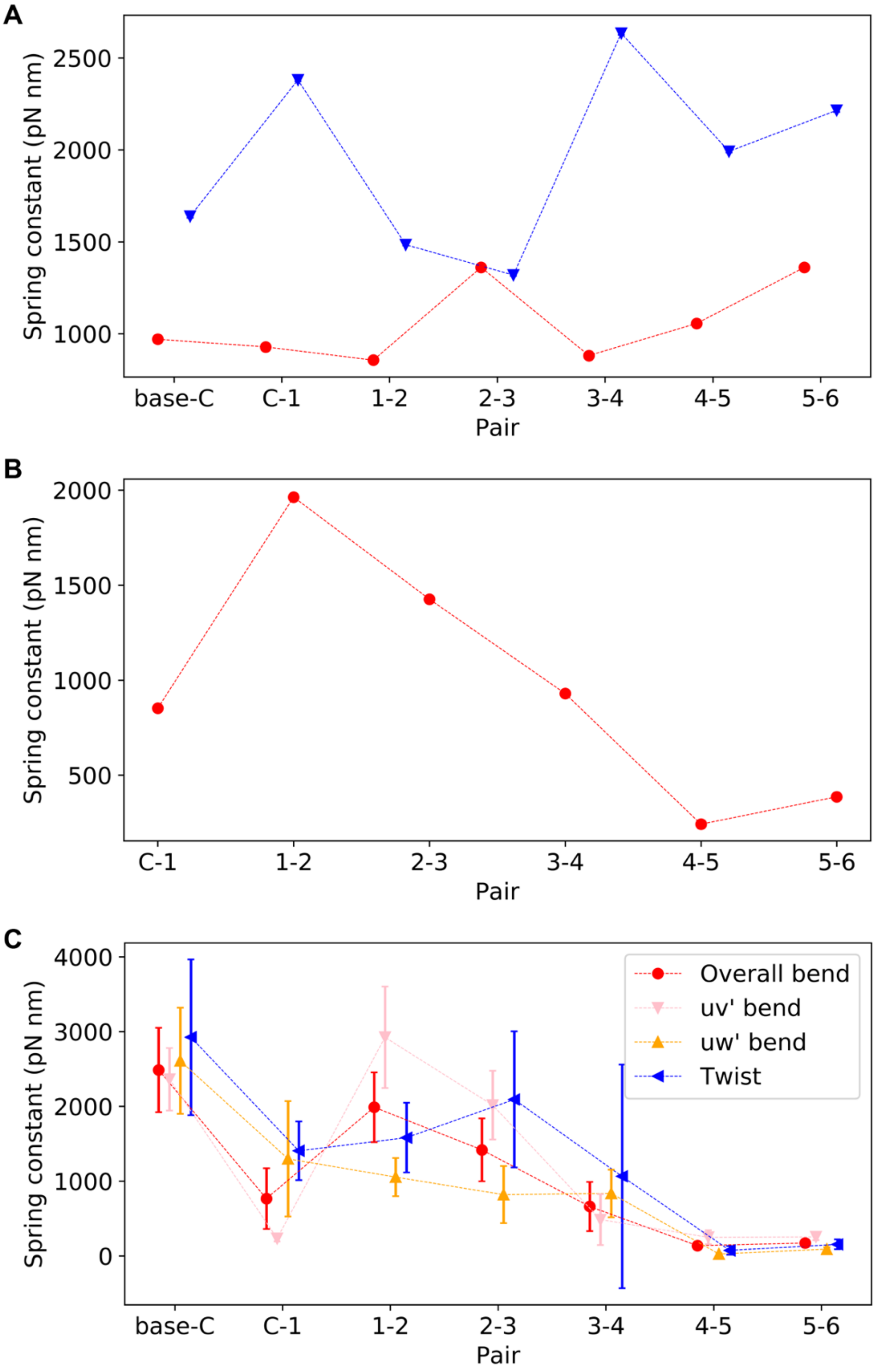


Figure 3.8: Spring constants for hinge points at subdomain connections ***A:*** *gaussian filtered FFEA blob model.* ***B:*** *unconstrained FFEA rod model.* ***C:*** *cryo-EM data (Chapter 2). Overall bend is the average bend constant in the uv' and uw' axes (see Chapter 2). Twist is the spring constant in the vw axes (see Chapter 2). Base = actin binding interface, C = converter, 1 = CaM1, 2 = CaM2, 3 = CaM4, 5 = CaM5, 6 = CaM6 (see Materials and Methods). Error bars show SD of random bending stiffnesses (see* ***Materials and Methods****).*

3.3.2 Coarse-graining cryo-EM maps can create unwanted features in the model

Cryo-EM map pre-processing affected the shape of the tetrahedral mesh, and therefore had downstream effects on the behaviour of the model. FFEA models from meshes generated from a cryo-EM map pre-processed either by gaussian filtering or rescaled to a voxel size of 15 Å and then gaussian filtered, occupied different conformational space in the simulations. When the vectors representing the lever conformations in both simulations were plotted in the cryo-EM map coordinate system, they occupied different conformational spaces (Figure 3.7B-C). The model generated from the solely gaussian filtered cryo-EM map matched the lever motion seen in the experimental data and was therefore selected for further analysis (Section 3.3.1). However, the model generated from the rescaled then gaussian filtered cryo-EM map was skewed positively in y (tilt) and negatively in x (slew), despite having a similar range (-5 to 5 for both tilt and slew) to both the gaussian filtered model and cryo-EM data.

The cause for the skew in the rescaled then gaussian filtered model is likely to be due to detail being lost in the map through coarse-graining, which then persists in the model. Clear differences can be seen within both the maps and meshes at the junction between the motor and lever domain (Figure 3.9). Rescaling the voxel size to 15 Å then gaussian filtering the cryo-EM map smooths over the motor-lever junction and joins the N-terminal domain of the motor to CaM1, which is an artificial feature. PCA analysis directly showed the effect of this artificial feature on the modes of motion (Figure 3.10, Appendix B.1 Movie 3.1). For the dominant mode of motion (PC1, Table 3.2), the direction of the rescaled then gaussian filtered model has a diagonal skew relative to the side-to-side (along F-actin longitudinal axis) motion of the gaussian filtered model (Figure 3.10). In both PC1 and PC2, hinging occurs between CaM1 and 2 in the rescaled then gaussian filtered model, but between the converter and CaM1 in the gaussian filtered model (Figure 3.10). This directly shows how coarse-graining of the map led to an artificial connection between the motor and lever in the mesh, which then masked the hinge point between the converter and CaM1 in the rescaled and gaussian filtered model.

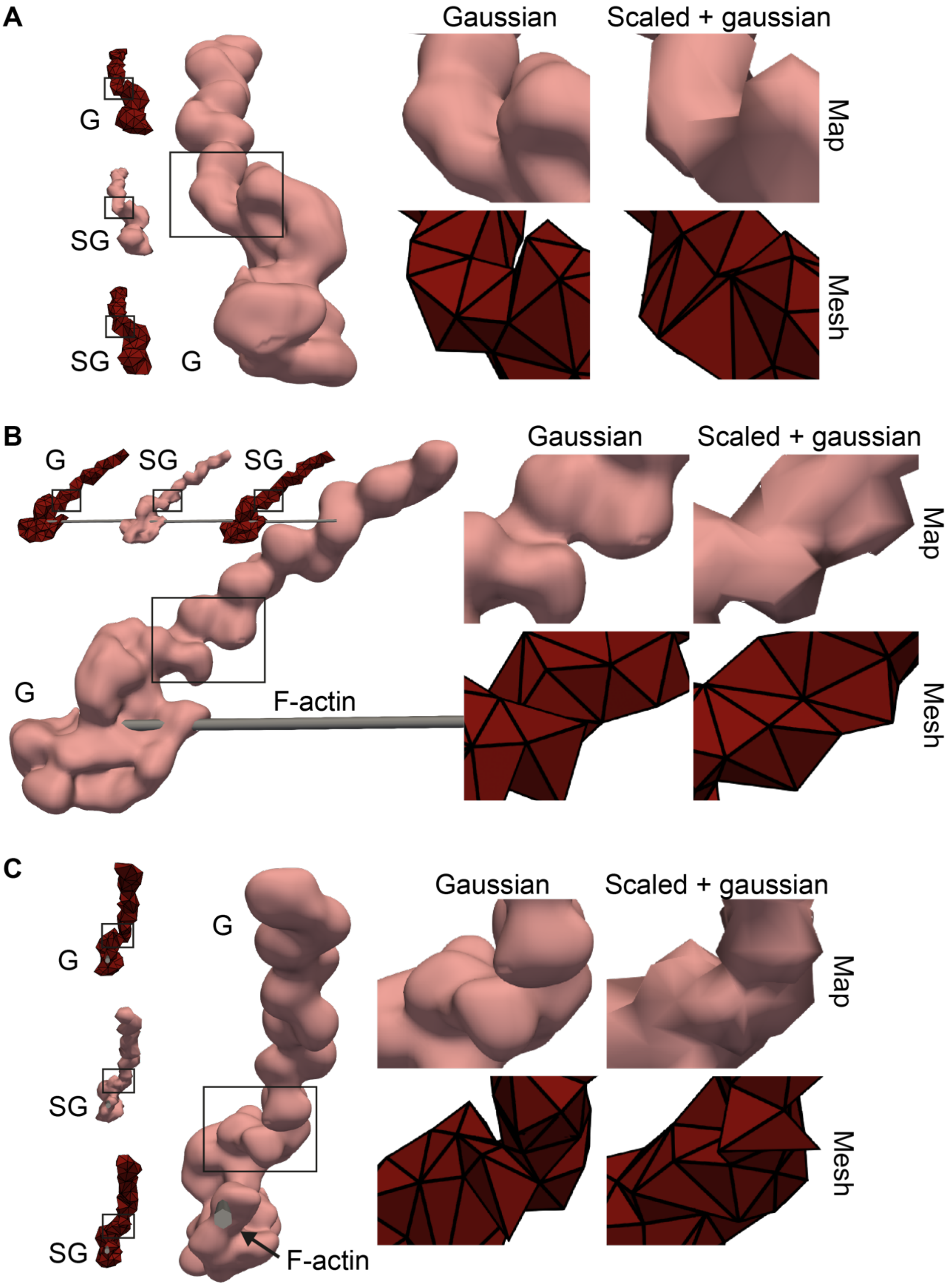


Figure 3.9: Maps pre-processed in 2 ways produce different meshes
*Rotated views (**A-C**, 90° about y-axis) of pre-processed cryo-EM maps and FFEA blob meshes. Boxed region highlights where coarse-graining has merged the motor-lever junction in the rescaled and gaussian filtered model (scaled + gaussian). Gaussian = cryo-EM map pre-processed by gaussian filtering (5 SDs). Scaled + gaussian = cryo-EM map pre-processed by rescaling to a voxel size of 15 Å and gaussian filtering (5 SDs). Map = cryo-EM map. Mesh = 3D tetrahedral finite element mesh. F-actin = vector from 1st to 13th subunit.*

Table 3.2: Eigenvalues for PCA modes

	Eigenvalues (Å^2)	
Mode	**Gaussian**	**Scaled + gaussian**
PC1	3.85×10^3	3.40×10^3
PC2	2.61×10^3	2.68×10^3
PC3	4.67×10^2	4.65×10^2
PC4	4.07×10^2	4.36×10^2
PC5	2.79×10^2	3.41×10^2

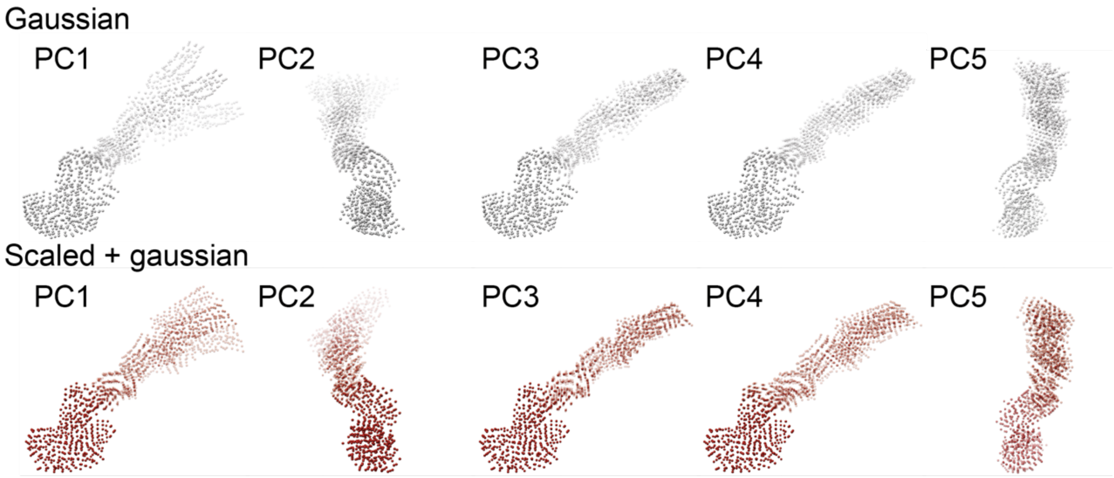


Figure 3.10: Principal modes of motion

Principal modes of motion for the gaussian filtered (Gaussian, grey), and re-scaled then gaussian filtered (Scaled + gaussian, red), FFEA blob simulations. PCA was performed using pyPcazip (Shkurti et al., 2016)*. The nodes of the mesh are displayed. Chimera was used to visualise the modes and hold the nodes representing F-actin* (Pettersen et al., 2004)*. PC = principal component.*

3.3.3 FFEA rod simulation of Myo5a-HMM reproduces experimental data

The FFEA rod simulation of Myo5a-HMM reproduced experimentally derived features of Myo5a. Firstly, the emergent subdomain connection stiffnesses in the model (Figure 3.8B) matched the parameters they were assigned (Figure 3.8C), showing no artificial features were introduced in the model. The stiffnesses of the connections between CaM4-5 and CaM5-6 were moderately stiffer in the HMM model than the cryo-EM data for S1 (CaM4-5: 240 ± 1 pN nm and 140 ± 54 pN nm, and CaM5-6: 390 ± 1 pN nm and 170 ± 36 pN nm, respectively). This is likely to be due to reduction in variability of the lever end positions created by dimerization.

Secondly, the distribution of distances between the start and end nodes (converter-converter) in the FFEA rod simulations, for both constrained and unconstrained head-tail junctions, matched experimentally derived distances between the motors in Myo5a-HMM (Figure 3.11). For the unconstrained head-tail junction simulation the mean distance was 41.0 ± 5.7 nm (mean ± SD), and for the constrained head-tail junction simulation the mean distance was 39.9 ± 4.8 nm (mean ± SD) (Figure 3.11A). In optical trap data the stepping distance is 34.5 ± 10.4 nm (mean ± SD) (Figure 3.11B), which suggests the simulated distances are within an acceptable range (Veigel et al., 2002). For distances resolved previously by ns-EM, 77% of the data is within the 13-subunit peak (35.8 nm), however there was no quantitative assessment of the distribution (Walker et al., 2000), but on visual inspection the histograms from the model are similar (Figure 3.11C). Some differences with ns-EM data are to be expected as the molecule can appear distorted due to flattening (Zhang et al., 2010). It should also be noted that the distribution of the rod model converter-converter distances does not peak at 36 nm. This could be because the rod model equilibrium structure has an converter-converter distance of 38 nm.

Thirdly, the cantilever bending stiffness for a single lever calculated from the FFEA rod simulation with a constrained head-tail junction was 0.16 pN/nm, which matches closely with optical trap data (0.2 ± 0.01 pN/nm for S1 and 0.18 pN/nm for a single HMM head bound to F-actin). This however differs from the

0.76 pN/nm calculated from my cryo-EM data. In both the optical trap method and the FFEA rod model, the static end of the cantilever may have some variability, which could partially account for the similarity in values derived by these two approaches compared with the cryo-EM data.

Finally, the bending stiffness for HMM derived from the FFEA constrained rod simulation was 0.18 pN/nm, and from the unconstrained simulation was 0.12 pN/nm. These values are similar to the stiffnesses calculated from optical trap data, 0.35 pN/nm, and ns-EM data, 0.26 pN/nm, for the dimer. Increased flexibility in the simulations for free HMM is to be expected, as it is not confined to actin-binding sites as it is in the other approaches (Walker et al., 2000; Veigel et al., 2002).

3.3.4 Constrained and unconstrained head-tail junctions are indistinguishable

Constraints at the head-tail junction had little to no effect on the overall features of Myo5a-HMM. In order to determine the impact of a constrained head-tail junction, FFEA rod model simulations were run where the node representing the head-tail junction either had the ns-EM derived bending stiffness value (7.04×10^{-29} N m^2) or the bending stiffness of a globular protein domain, e.g. the motor domain (9.09×10^{-27} N m^2) (Table 3.1). The converter-converter distances for models constrained and unconstrained at the head-tail junction were approximately equal. For the unconstrained head-tail junction simulation the mean distance was 41.0 ± 5.7 nm (mean ± SD), and for the constrained head-tail junction simulation the mean distance was 39.9 ± 4.8 nm (mean ± SD) (Figure 3.11A). Additionally, although differences between the constrained and unconstrained head-tail junction could be seen in the distributions of the angle at the head-tail junction locally (to CaM6) (Figure 3.12C), the effects are dampened if the angle is measured further along the lever further along the lever (to CaM4) (Figure 3.12D). Therefore, it can be hypothesised that the flexibility at the lever subdomain connections (CaM4-6) dampen any effects a constrained head-tail junction would have on the overall properties of the dimeric molecule.

No difference could be seen in the 2D projections between the constrained and unconstrained head-tail junction simulations, even locally (Figure 3.12A&B, Figure 3.13 and Figure 3.14). This shows the use of 2D projections introduces artificial variance as the angle of view cannot be determined. Therefore, determining parameters for the free molecule using 2D ns-EM data is not appropriate. This is less of a problem for the actin-bound molecule as the actin filament can be used to determine the angle of view (Oke et al., 2010).

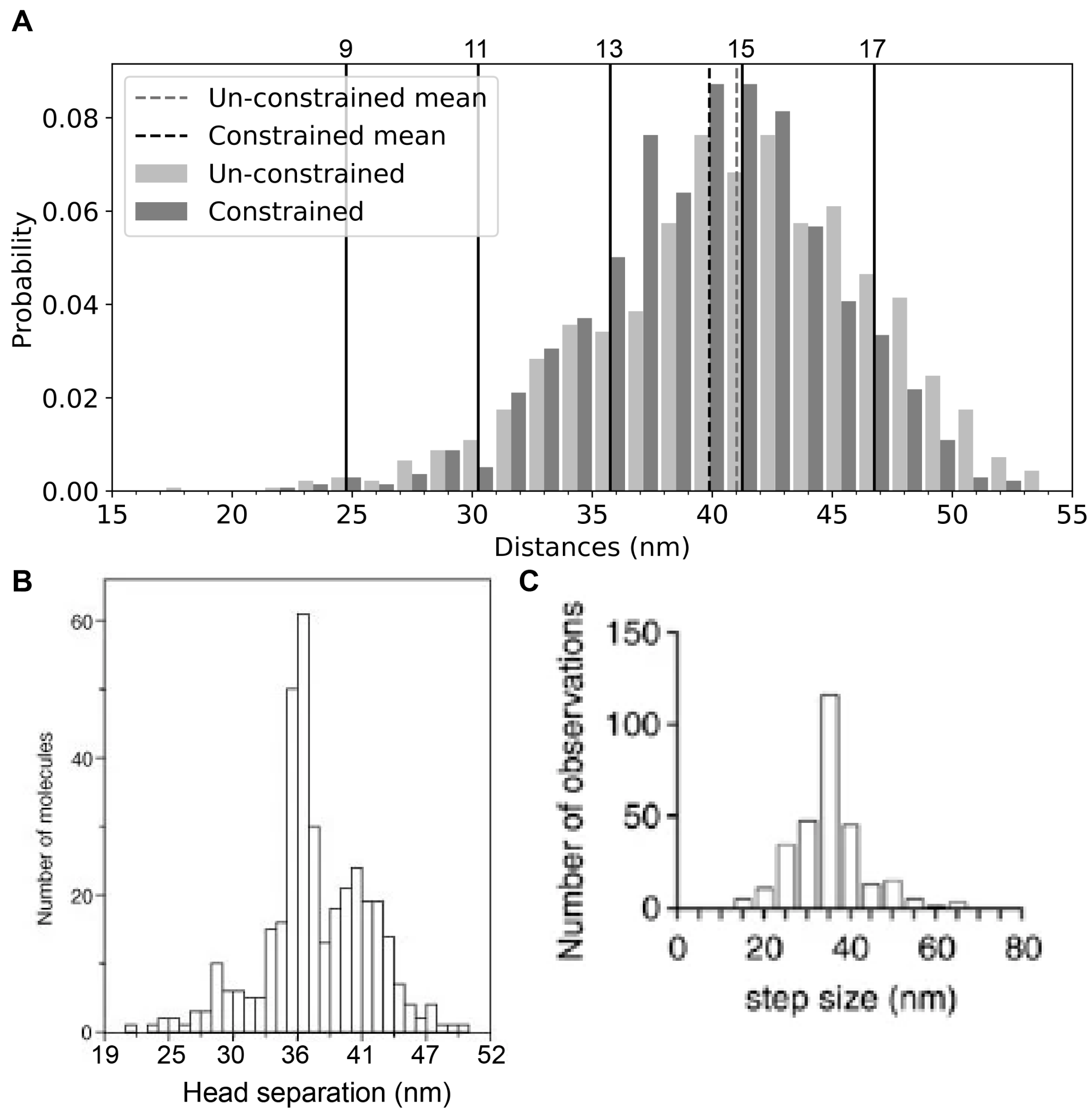


Figure 3.11: Distribution of distances between the motor in HMM
A: *histogram of distances measured between the nodes representing the start of the converter domain (converter-converter distances) in the constrained and unconstrained FFEA rod models. Dashed lines indicate the mean. Solid lines show distances between actin bound heads seen experimentally* (Walker et al., 2000; Oke et al., 2010)*, with the number of actin subunits indicated on top.* ***B:*** *histogram adapted from (Walker et al., 2000), indicating distribution of head separation distance seen using ns-EM.* ***C:*** *histogram taken from (Veigel et al., 2002), indicating distribution of step sizes determined with an optical trap.*

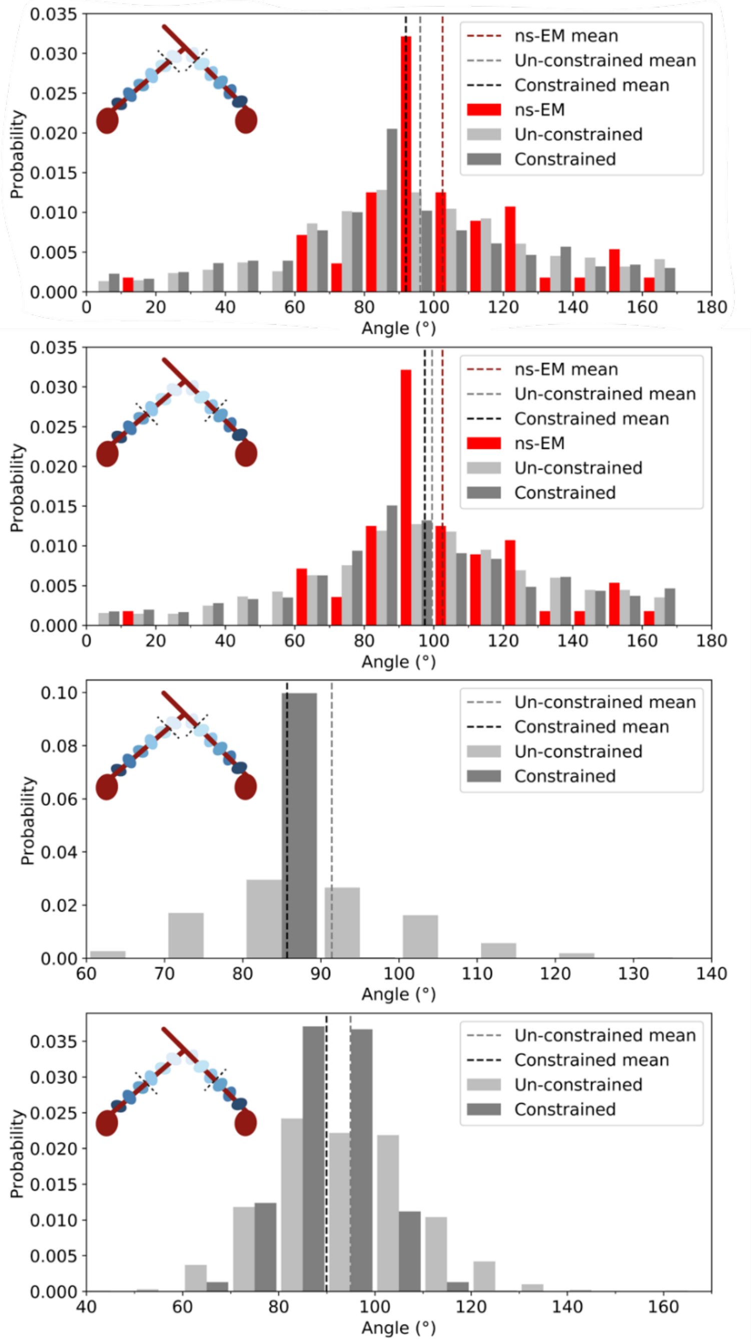


Figure 3.12: Distribution of lever angles measured at the head-tail junction

A-B: Histograms of the distribution of angles between the levers at head-tail junction, in the constrained and unconstrained FFEA rod models, in the xy plane (2D projections). Measured to CaM6 (**A**) (vectors between nodes 7-6 and 7-8), and CaM4 (**B**) (vectors between nodes 7-4 and 7-10). The distribution of angles measured between levers at the head-tail junction in ns-EM (2D projection) are plotted alongside the models (the number of CaMs included in the line representing the lever has not been taken into consideration). **C-D:** Histograms of the distribution of angles between the levers at head-tail junction, in the constrained and unconstrained FFEA rod models, in 3D. Measured to the CaM6 (**C**) and CaM4 (**D**). Inset shows schematic representation of HMM, dotted line shows where along the lever the angle is measured between.

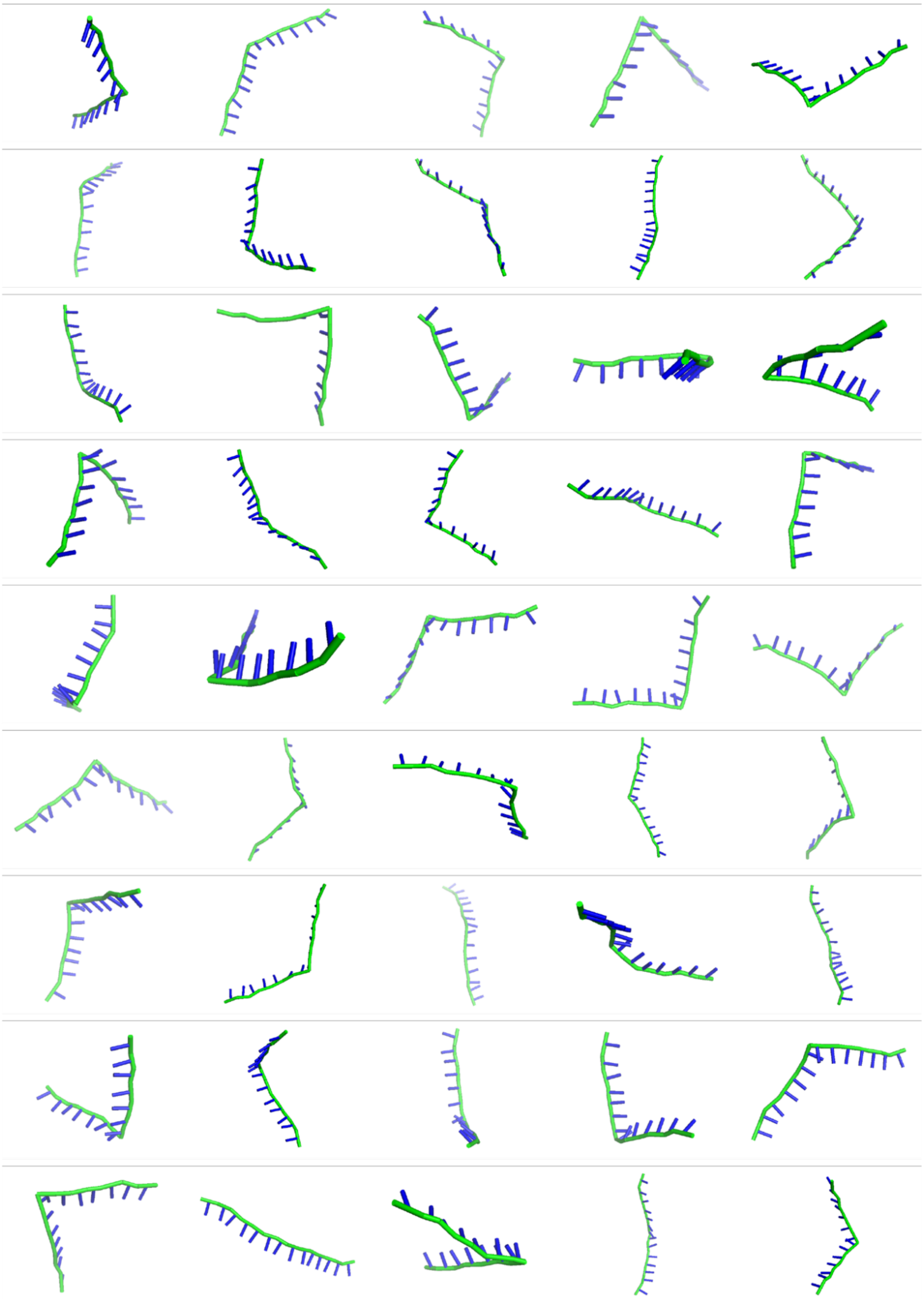


Figure 3.13: Snapshots of unconstrained head-tail junction FFEA rod simulation

Randomly selected snapshots from the xy plane of the unconstrained head-tail junction FFEA rod simulation.

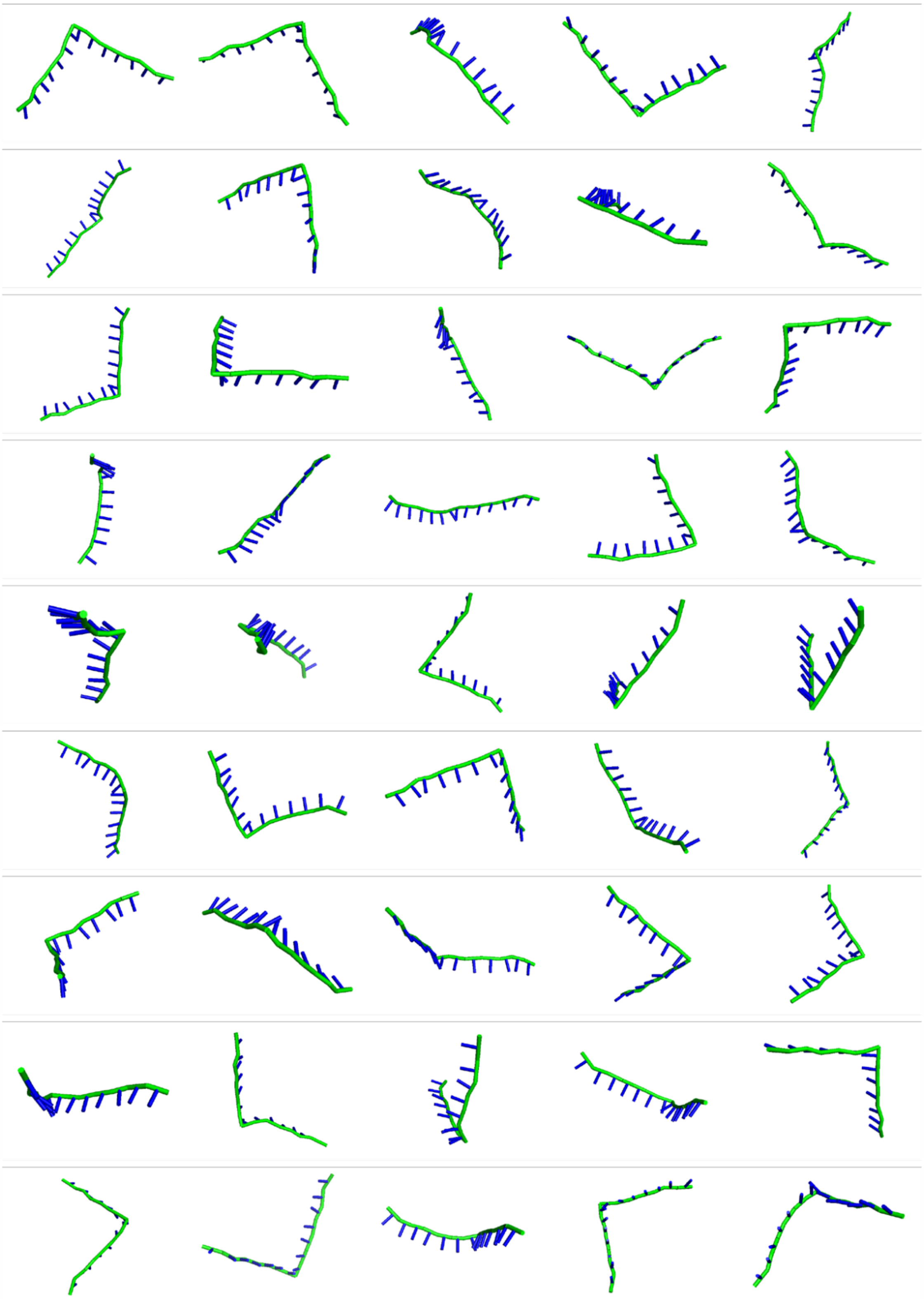


Figure 3.14: Snapshots of constrained head-tail junction FFEA rod simulation

Randomly selected snapshots from the xy plane of the constrained head-tail junction FFEA rod simulation.

3.4 Discussion

In summary, I have demonstrated FFEA models of Myo5a, generated in two ways (blob and rod) and parametrised with cryo-EM data, were able to reproduce experimentally derived features, such as lever conformations and stepping distances. However, further complexity needs to be added in order to generate more representative models. The Myo5a-S1 blob model requires material parameters to be applied anisotropically to include the subdomain hinge regions. The Myo5a-HMM rod model is based on a hypothetical equilibrium model (Chapter 4), and therefore would be greatly improved by the resolution of a structure of the active dimeric molecule (either HMM or full-length). This is also required to accurately model the head-tail junction, as characterisation of the stiffness from 2D projections leads to artificially low bending stiffnesses, as I have demonstrated here.

The FFEA rod model of Myo5a-HMM enabled interpretation of how the properties of the lever influence the dimer. By varying bending stiffness at the head-tail junction I demonstrated that constraint at the head-tail junction had little to no effect on the overall properties of Myo5a, due to the flexibility of the lever. This challenges previous theories that a constrained angle at the head-tail junction may be required for compass-like walking (Andrecka et al., 2015; Hathcock et al., 2020). Even with a flexible head-tail junction, I found there to be a relatively tight range of distances explored (SD 5.7 nm, 4.8 nm for constrained). In both models the rotational freedom at the head-tail junction was constrained by a high torsional stiffness, the impact of rotation at this joint therefore needs to be determined. FFEA rod models provide a suitable means to do this. 3D classes of the head-tail junction will be necessary to calculate the torsional stiffness, further emphasising the need for more 3D structures of the dimer.

Methods of coarse-grained modelling using FFEA are still being optimised. Through this process I have determined factors important for consideration when building a FFEA model. Firstly, calculating parameters from ns-EM using 2D projections can introduce artificial flexibility. Therefore, if the stiffness is being calculated from EM data, reconstruction in 3D is preferable. Secondly, the mesh must be assessed to determine that its shape and/or coarseness has not introduced artificial features. Therefore, the model must be validated against experimental data before simulating unknowns. I.e., does the model reproduce experimentally derived material parameters? Does the model explore the same conformational space?

This study further demonstrates FFEA as a suitable tool for modelling cytoskeletal motor dynamics and highlights the value in using simulations to integrate experimental data. 3D structures of the active dimer, and models with further complexity, are required in order to ultimately generate a continuum model of Myo5a walking. It is likely that a more complex, combined rod and blob model would be required to model the strained state (with the motor modelled as a blob, and the lever as a rod), as the rod model is more stable under strain (Chapter 4).

Chapter 4 Building the strained state of Myo5a and development of an algorithm for mesh generation

4.1 Introduction

Current experimental techniques do not have sufficient temporal or spatial resolution to capture Myo5a taking a step (reviewed in (Gravett et al., 2022)). The transient period of the step, as the free head moves through space to find the next binding site, takes approximately 17.5 ms, and the gait of the step is 36 nm. Therefore, a time resolution of less than 17.5 ms and spatial resolution of less than 36 nm would be necessary to distinguish anything occurring during this period. It is possible to directly visualise the dwell state using techniques such as ns-EM (Walker et al., 2000; Oke et al., 2010) and high-speed atomic force microscopy (hs-AFM) (Kodera et al., 2010)). However, current data on the transient state comes from tracking the labelled molecule, which doesn’t allow direct visualisation of its structural dynamics (interferometric scattering microscopy (iSCAT) (Andrecka et al., 2015)).

Computational simulations bridge these gaps in temporal and spatial resolutions, and provide a means to model the transient state and uncover Myo5a walking mechanics. The timescale (tens of milliseconds) together with the large number of atoms required to model Myo5a walking along the F-actin track (Myo5a-HMM comprises 62415 atoms without F-actin or solvent) is prohibitive for all-atom MD simulations (typically 10^5 atoms and 1 μs). However, this system is a good candidate for FFEA simulation. Here I construct and optimise the starting point of a theoretical model of a Myo5a step using FFEA. This is described as a toy model as it is a simplistic attempt to model Myo5a walking based on experimentally informed but hypothetical states (e.g. there is no 3D structure of active Myo5a), and an experimentally informed but hypothetical mode of walking (there is limited information on the transient state).

The motion of cytoskeletal motors is directed (i.e. the probability of forward motion exceeds that of reverse motion), therefore they function outside of equilibrium (Brown and Sivak, 2020). Typically simulations show protein equilibrium dynamics, however non-equilibrium simulations can be carried out by time dependent swapping between distinct equilibrium states (Richardson et al., 2020). This is complicated in FFEA as the structure of the mesh of each state is distinct, which means the elements and nodes are not the same in each state, therefore mapping from one mesh to the next is difficult. Whereas in atomistic MD the structure of each state has the same atoms (node equivalent), so mapping is easier. Alternatively, non-equilibrium simulations can be carried out through imposing force on the motor to create directed motion and prevent reversal (Ovchinnikov et al., 2010), which was our method of choice. To model the starting point of Myo5a walking it was necessary to generate a model that incorporated directional intramolecular strain. I opted for a two-state model where I generated a hypothetical unrestrained “equilibrium” state, and used this to generate a hypothetical strained state, where torque was applied to the lead head to create directed motion.

Here I demonstrate the construction of a FFEA combined rod-blob model of Myo5a-HMM under torsional strain, and the optimisation of finite element mesh generation for the purposes of FFEA simulation.

4.2 Materials and Methods

4.2.1 Equilibrium model

First, a 16-subunit F-actin was made by superimposition of the first and last subunit of multiple copies of the actin filament in PDB: 6C1H (Mentes et al., 2018). To model the trail head, a motor domain in the post-PS ADP bound state (PDB: 1W7I (Coureux et al., 2004)) was superposed onto an actomyosin structure (PDB: 6BIH (Banerjee et al., 2017)). This model was made before the Pospich et al. structure was published (Pospich et al., 2021), however the modelled actomyosin is only moderately different to the actomyosin cryo-EM structure, and as they are later coarsened the differences are negligible. The lever domain was excised from PDB: 2DFS (Liu et al., 2006a), then aligned with the single IQ domain in the ADP-bound structure (this model was made before I resolved the lever structure Chapter 2). To model the lead head a copy of the trail head was rotated 180° about axis 1 (Figure 4.2) and then translated so the lever ends joined in Chimera (Pettersen et al., 2004).

From this initial PDB model, a single rod model of the dimerised levers was generated and discretised so that there was an element representing each CaM bound IQ and each converter domain of the dimer (14 elements, 15 nodes) (Welch et al., 2020). A low resolution (30 Å) electron density map of the motor domain, excluding the SH3-like fold for simplification, was generated from the PDB model using Chimera (Pettersen et al., 2004). For the TetGen meshes, a surface mesh was then generated and coarsened to 15 Å using FFEA tools, and TetGen (Si, 2015) was used to generate a volumetric mesh (see 4.2.3 for Marching Tetrahedra meshes). Finally the relative location and orientation of each domain for initialisation in the simulation box was extracted, using Chimera (Pettersen et al., 2004) to find the centre of mass of each object and the rotation matrix of the trail motor to the lead.

4.2.2 FFEA simulations

FFEA simulations were performed using FFEA version 2.7.2 (Solernou et al., 2018), on 4 CPUs, using specific conditions (Table 4.1), using the CG_nomass (inertialess) solver. For blob models the shear and bulk viscosities were set to 1 mPa.s. calc_noise was set to 0 in FFEA for the strained state setup simulations to disable the mesh thermal noise. Simulations were run at 24.5 °C ($K_bT = 4.11 \times 10^{-21}$ J).

Table 4.1 - FFEA simulation conditions. App. = Appendix B.2 Movie

				Mesh			**Rod**			
Simulation	**Length (s)**	**Timestep (s)**	**Positioning mode**	**Shear modulus (N m^{-2})**	**Bulk modulus (N m^{-2})**	**Type**	**Stretch constant (N)**	**Twist constant (N m^2)**	**Radius (m)**	**Bend constant (N m^2)**
Strain setup, flexible rod (Figure 4.3, App. 4.1)	1.005 x 10^{-9}	3 x 10^{-13}	Ctforce torque (1 x 10^{-11} N)	5.5 x 10^8	2.2 x 10^9	TetGen	3.5 x 10^{-12}	0	5 x 10^{-9}	3.5 x 10^{-29}
	7.05 x 10^{-10}	3 x 10^{-13}	Ctforce linear (1 x 10^{-11} N)							
Strain setup, stiffer rod (Figure 4.4, App. 4.2)	7.35 x 10^{-11} (inverts)	3 x 10^{-13}	Ctforce Torque (1 x 10^{-11} N)	5.5 x 10^8	2.2 x 10^9	TetGen	3.5 x 10^{-11}	5 x 10^{-30}	5 x 10^{-9}	3 x 10^{-29}
15 Å (Figure 4.6, App. 4.3)	5 x 10^{-7} (in 2 hrs for max.)	Max. 3 x 10^{-13}	N/A	9.2 x 10^6	2.8 x 10^7	Marching Tet	N/A			
20 Å (Figure 4.6, App. 4.4)	5 x 10^{-7} (in 40 s for max.)	Max. 3 x 10^{-11}	N/A	9.2 x 10^6	2.8 x 10^7	Marching Tet	N/A			
30 Å (Figure 4.6, App. 4.5)	5 x 10^{-7} (in 4 s for max.)	Max. 1 x 10^{-10}	N/A	9.2 x 10^6	2.8 x 10^7	Marching Tet	N/A			
Strain setup, marching tetrahedra (Figure 4.8, App. 4.6-4.8)	1 x 10^{-5}	1 x 10^{-12}	Spring (1 N/m, equilibrium: 1 Å)	5.5 x 10^8	2.2 x 10^9	Marching Tet	1 x 10^{-11}	5 x 10^{-26}	5 x 10^{-9}	3 x 10^{-27}
	2.18 x 10^{-5}	1 x 10^{-12}	Pin axis, ctforce torque (1 x 10^{-11} N)							
	5.5 x 10^{-5}	1 x 10^{-12}	Pin base							

4.2.3 Marching tetrahedra meshing algorithm

The equilibrium state was used as a test case. An electron density map was generated from the PDB file using Chimera (Pettersen et al., 2004). To determine the best method of coarsening, the map was then coarsened to the desired voxel size using either the scipy.ndimage.zoom function in SciPy (Virtanen et al., 2020), the volume voxelSize function in ChimeraX (Pettersen et al., 2021) or the relion_image_handler --rescale_angpix function in RELION 3.1 (Zivanov et al., 2018). Based on the preservation of the surface of the molecule, RELION was selected as the best tool (Figure 4.7). Briefly, the meshing tool then works via the following process (for more in-depth logic see Figure 4.1):

1. Use threshold/isosurface to decide which voxels will be included
2. Generate cubic mesh from voxel data
3. Divide cubic mesh into 5 tetrahedra
4. Write out as TetGen style output files
5. Generate FFEA mesh with TetGen style output files

The stability of the meshes in FFEA was then tested by running 500 ns simulations with timesteps ranging from 1×10^{-13} to 1×10^{-10}. Simulation settings were left as default in all cases, for further conditions see Table 3.1.

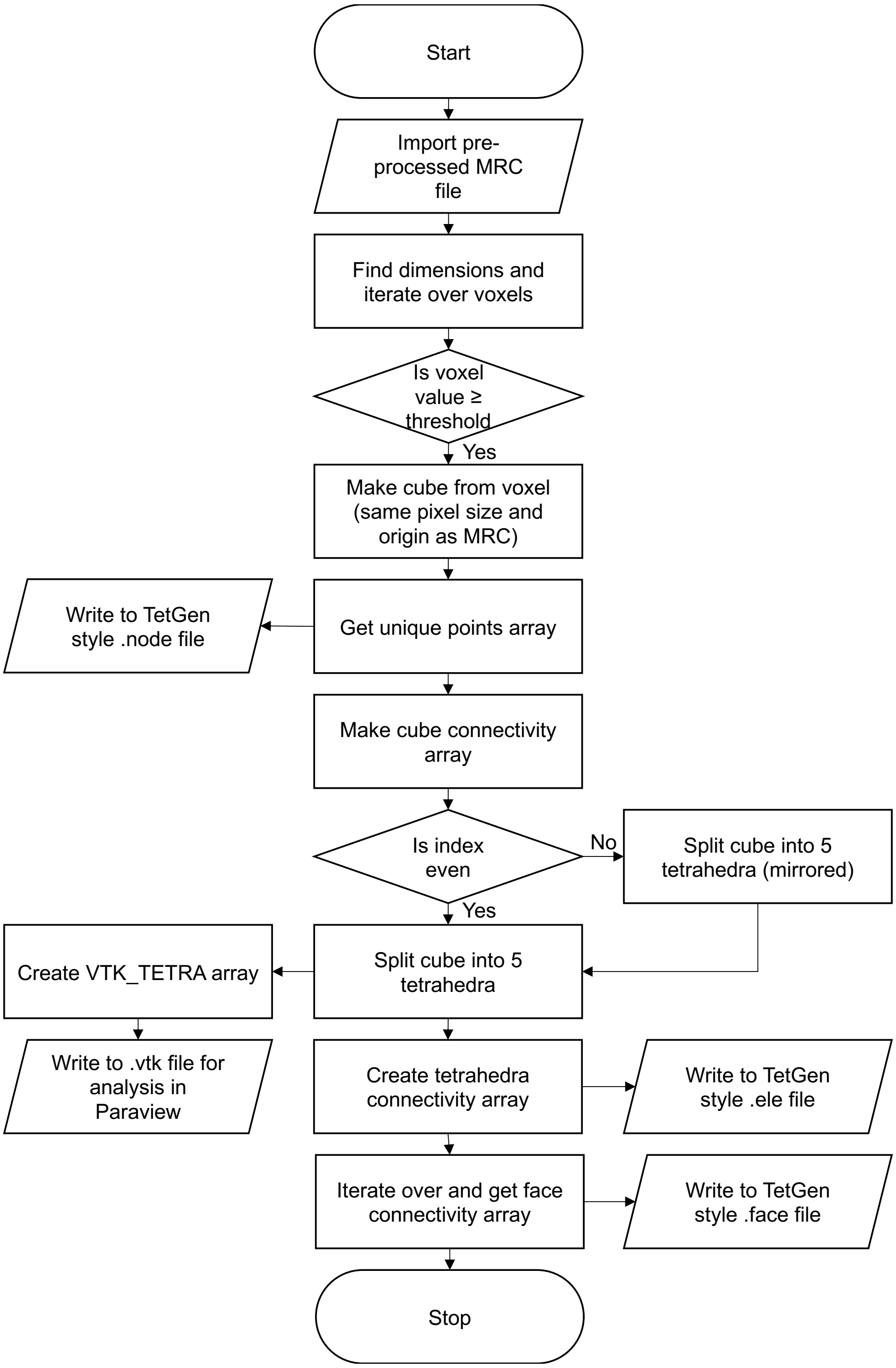


Figure 4.1: Marching tetrahedra algorithm logic flowchart
Logic flowchart describing each step of the marching tetrahedra algorithm. Cryo-EM maps in the MRC file format were pre-processed in SciPy (Virtanen et al., 2020), ChimeraX (Pettersen et al., 2021) or RELION 3.1 (Zivanov et al., 2018) beforehand. Pill shape = start/stop, rhomboid/parallelogram = data input/output, rectangle = process, diamond = decision/if statement. MRC = Medical Research Council cryo-EM map file format (Cheng et al., 2015). .ele, .node and .face files were written in the same format as TetGen (Si, 2015). .vtk files were written in the VTK format (Schroeder et al., 2006). The success of the mesh was assessed by visualisation in ParaView (Ayachit, 2015).

4.3 Results and Discussion

4.3.1 Equilibrium model

A hypothetical, “equilibrium” model was generated in order to represent directional intramolecular strain at the mesoscale. For the majority of the dwell state, when both heads of Myo5a are bound to F-actin, both motor domains contain ADP. In the ADP bound state, the lever points towards the plus end of F-actin. However, when both heads are bound the lead head lever is restrained to point towards the minus end of F-actin. Modelling nucleotide and actin binding associated conformational changes within FFEA is not trivial, therefore I opted for a simplistic representation of this mechanism. In this model I assume the intramolecular strain generated is torsional to induce a compass-like motion as shown previously (Andrecka et al., 2015).

I generated an “equilibrium” model of HMM with both heads in the ADP bound post-PS state, as that is the equilibrium state of a single head when bound to actin. Each head is a 180° rotation about the first axis (Figure 4.2), in order to position the motor domain of the leading head at the 13th subunit, followed by a translation to connect the lever ends. To replicate compass-like walking, directional torsional strain can then later be incorporated by rotating the lead motor 180° about the second axis (Figure 4.2). This will place the motor in the correct orientation for F-actin binding and will force the lead head into the pre-PS state. When the trail head is released the torsional strain across the lead head should then be released creating a lever swing, as continuum models are elastic and thus return to their equilibrium state (Solernou et al., 2018).

In order to capture torsional strain a combined rod-blob model was used (Figure 4.3A) where the levers were modelled as rods. Rods were chosen to model the levers as they are more suitable for modelling elongated molecules. Additionally, meshes are unstable when twisted, whereas the rod model better withstands twisting (Welch et al., 2020). The motor domain was modelled as a blob to allow application of force during the simulation.

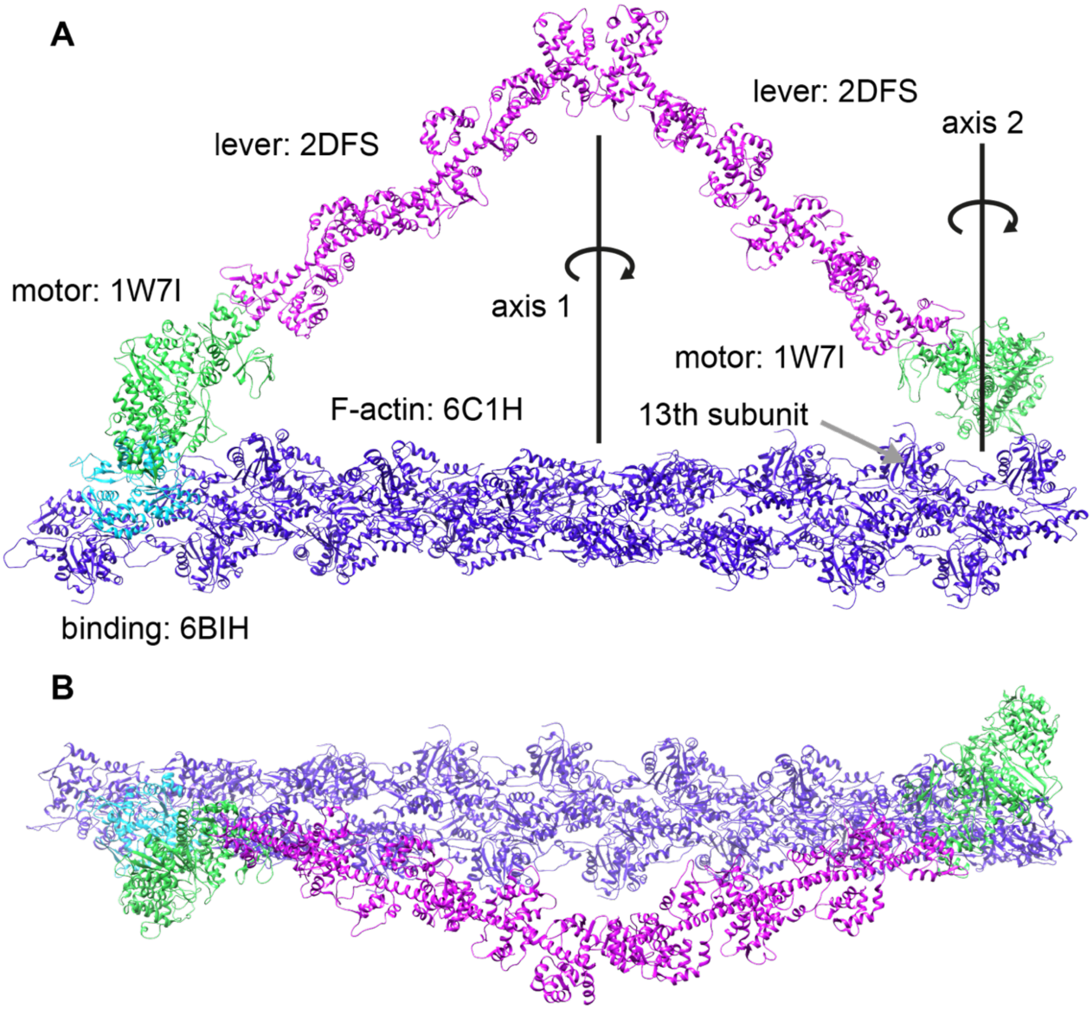


Figure 4.2: Building equilibrium model
A: *F-actin created by superimposing the first and last subunit of PDB: 6C1H repeats (Mentes et al., 2018). Binding pose of the trail head was found by superimposing an ADP bound motor structure (PDB: 1W7I* (Coureux et al., 2004)*) onto an actomyosin structure (PDB: 6BIH* (Banerjee et al., 2017)*). Levers excised from shutdown state structure (PDB: 2DFS* (Liu et al., 2006a)*) were superimposed onto first IQ of the motor structure. The lead head was generated by rotating a copy of the trail head about axis 1 and then translating it to join the lever ends. Axis 2 shows the rotation axis the lead motor needs to be rotated about to generate the strained state.* ***B:*** *a 90° rotation of* ***A*** *about x axis.*

4.3.2 Imposing forces on rod-blob connections causes element inversion

Forces can be used to successfully position blob models in FFEA in a directed manner. The FFEA ctforce function, which imposes a constant force on defined nodes within the FFEA mesh, was used to position the motor of the lead head during simulation. Ctforces were used to rotate the lead head from its orientation in the equilibrium model to its actin bound orientation at the actin binding site of the 13th subunit, in order to model strained Myo5a-HMM with both heads F-actin bound. The axes and location parameters for this function were extracted from Chimera (Pettersen et al., 2004) (see Materials and Methods) to enable informed and targeted positioning. Using these parameters, the lead motor domain was successfully manoeuvred by ctforces to the desired location during simulation (Figure 4.3, Appendix B.2 Movie 4.1). This is the first time this function has been used for positioning molecules during a simulation, and the first time in which a FFEA simulation has been steered using force. This approach has prospects for steering mesoscale simulations in a similar way to that of MD simulations in order to: simulate non-equilibrium dynamics, speed up simulation timescales, encourage conformational changes, binding and unbinding events, and simulate force-based single-molecule experiments, such as AFM (Isralewitz et al., 2001).

However, when material parameters were increased 10-fold for the rod stretch constant, and from 0 to 5 x 10^{-30} N m^2 for the rod twist constant (Table 3.1), positioning using ctforces resulted in element inversion. Element inversion describes when a node moves so far in one direction that it passes through the face of an associated element, turning itself inside out. For initial proof of function testing, low material parameters were used. When rod stiffness was increased element inversion occurred at a neighbouring element to where the rod-blob connection occurs in the TetGen (Si, 2015) generated meshes (Figure 4.4B&C, Appendix B.2 Movie 4.2). This is likely to be due to large forces being translated at the rod-blob connection. Two features need to be investigated in order to prevent inversion happening, mesh stability and the rod-blob connection.

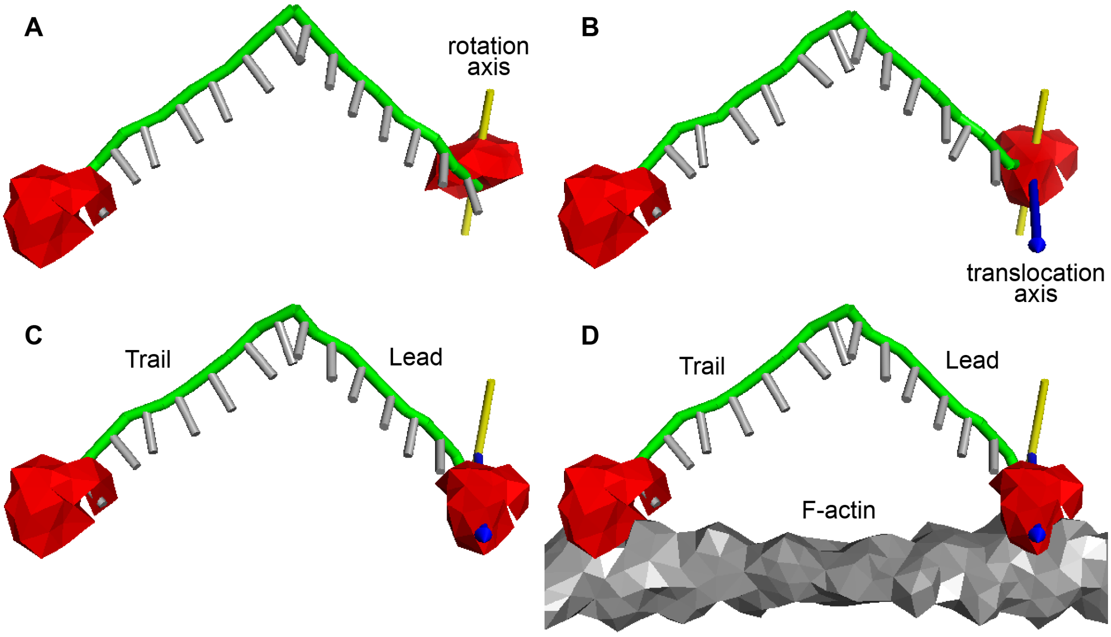


Figure 4.3: Positioning lead motor using forces

A: *FFEA combined rod-blob model of the equilibrium state (Figure 4.2) and starting pose for positioning simulation. Rotation axis about which the rotational ctforce is applied highlighted (yellow).* ***B:*** *end frame of rotation simulation where rotational forces (1 x 10^{-11} N) were applied to the lead motor about the rotation axis. The simulation was stopped when the lead motor was rotated to the point it matched the orientation of the trail motor. This is the starting frame for the translation simulation, the translocation axis along which linear forces (1 x 10^{-11} N) were applied is highlighted (blue arrow).* ***C:*** *end frame of the translation simulation. The simulation was stopped when the lead motor was at the actin-binding site.* ***D:*** *simulation end point with an F-actin mesh displayed to show lead head at binding site position. The grey bars represent the rod material axes, which are perpendicular to the rod elements, and are visualised to show twist in the rod.*

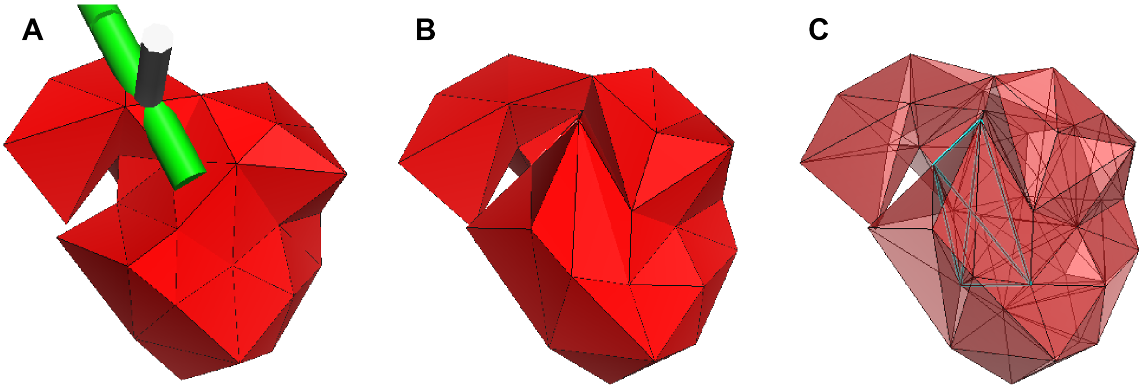


Figure 4.4: Element inversion occurs when rod stiffness increased
A: *start point of rotation simulation with stiffer rod parameters (Table 3.1).* ***B:*** *end frame showing element inversion.* ***C:*** *end frame with inverted element highlighted in cyan.*

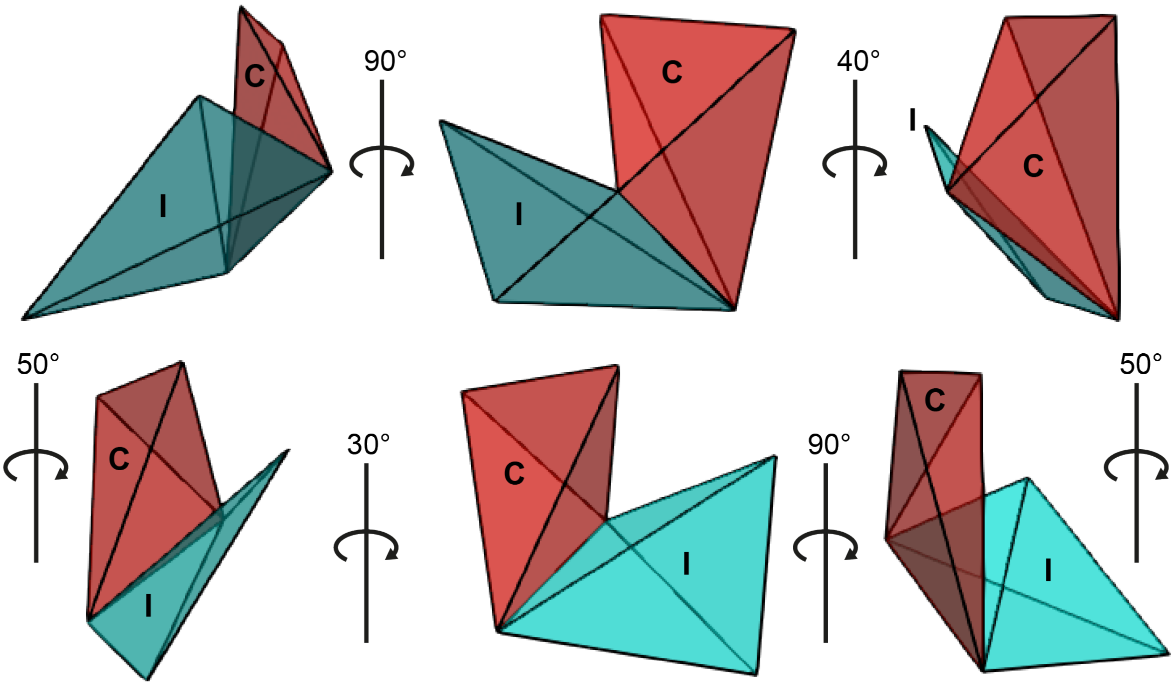


Figure 4.5: Connection and inverting element
*The element where rod-blob connection takes place (**C**) and the element which later inverts (**I**), visualised as a VTK* (Schroeder et al., 2006) *file in ParaView (Ayachit, 2015). The element where inversion takes place (**I**) has dihedral angles approaching 0°, which makes it unstable.*

4.3.2.1 Mesh stability

The first feature to address is the stability of the mesh. A common issue with mesh stability is the size and shape of the elements. Small elements are unstable in FFEA as at lengths below ~5 Å thermal fluctuations result in large strains in the elements (Oliver, 2013). Moreover, such elements are unrealistic as the continuum approximation breaks down below this length scale (Solernou et al., 2018). Geometry is also important, sliver elements, where all 4 vertices of the tetrahedron lie on approximately the same plane, lead to mesh instability. Geometrically unstable elements can be identified as tetrahedra with high aspect ratios (longest edge:shortest edge), or dihedral angles approaching 0° or 180°, which give increased errors in the finite element approximation (Si, 2015; Sukumar and Tupek, 2022).

In the motor mesh generated using TetGen (Si, 2015) there are small and geometrically unstable elements. The element where the rod-blob attachment takes place is small, and the neighbouring element that inverts has angles approaching 0° (Figure 4.5). When the rotation force (1×10^{-11} N) is applied to the blob, strain is created at the rod-blob connection as the rod resists the pull from the rotating blob. The connection element is small (the face is 1.86×10^{-18} m^2), therefore assuming the rod opposes the strain at an equal force of 1×10^{-11} N, the stress on the connection element is 5.38×10^6 Pa. The connection element has unstable neighbours and therefore cannot oppose the resisting force from the rod. The poor shape of the neighbouring element means it inverts when the strain from the connection element is transmitted. A method of creating more stable meshes from cryo-EM data is therefore required.

4.3.2.2 The rod-blob connection

The second feature that needs to be investigated is the rod-blob connection. Currently rods are connected to blobs at a single element. However, in Myo5a the connection of the lever to the motor is through multiple interactions, across a larger surface than the single rod-blob connection. Therefore, a single connecting element may not be appropriate (especially under force). As collision detection is being incorporated into the KOBRA code (Ryan Cocking, personal communication), the radius of the rod will become spatially resolved and therefore further interactions between the rod and the blob will take place. This could provide an opportunity for multiple rod-blob contacts to be incorporated, which may help with force distribution and reduce stress.

4.3.3 A new meshing algorithm improves mesh stability

An algorithm generating isotropic tetrahedral meshes from cryo-EM map voxel data was developed in order to improve mesh stability. Cryo-EM maps are structured as a 3D cartesian grid, with electron density values assigned to each element of the grid, which is termed a voxel. The algorithm generates a cubic mesh from the voxels, only including voxels with the electron density value greater than or equal to a specified threshold. The cubes are then subdivided into 5 tetrahedra (Figure 4.6A). The mode of subdivision has to alternate for neighbouring voxels as the faces of the tetrahedra must align. The algorithm is analogous to marching tetrahedra algorithms for graphic visualisation (Rajon and Bolch, 2003; Lu and Chen, 2012). Voxels can also be subdivided into 6 tetrahedra, which is advantageous in that alternating the mode of subdivision for adjacent voxels (as is necessary for 5 tetrahedra) is not necessary to match up adjoining faces. Alternating the mode of subdivision (Figure 4.6A) can lead to a different mesh structure depending on which mode of subdivision is used first (Rajon and Bolch, 2003). This might be problematic if a blob simulation was used to model the equilibrium state, as in the test case, as both heads need to behave in the same manner. However, to model the strained state two copies of the same mesh are used to model the motor so this is not a problem. Additionally, subdivision by 6 produces anisotropic meshes, which may introduce artificial anisotropy.

Because the marching tetrahedra algorithm generates meshes directly from voxel data, the size of the elements of the mesh are directly related to the resolution of the map. Thus, tools that change the voxel size of cryo-EM maps can be used to generate coarsened meshes, which alleviates issues with unstable small elements (Figure 4.6B-D, Appendix B.2 Movie 4.3-4.5). I trialled three ways of coarsening the cryo-EM map resolution (ChimeraX (Pettersen et al., 2004), SciPy (Virtanen et al., 2020), and RELION 3.1 (Zivanov et al., 2018)).

RELION performed the best at rescaling as it preserved the shape of the molecule without introducing jagged features to the surface (Figure 4.7). In RELION, the coarsened surfaces are smoothed by using a windowing function to coarsen the map in Fourier space and then inversing the Fourier transform. This is functionality that could be incorporated into the algorithm in future, so users do not have to download multiple programs. Another advantage of generating meshes in this way is that they are maintained within the same coordinate system and location as the data they are generated from, which allows for easy comparison of the mesh with the map or atomic model. This will prove useful for parametrisation, for example selecting faces that correspond with binding sites to apply interaction potentials to for protein-protein interaction simulations.

Meshes generated by the new algorithm had improved stability and increased the efficiency of the FFEA simulation. Typically, TetGen (Si, 2015) generated meshes are unstable in FFEA at timesteps larger than 100 fs as they invert. Meshes coarsened to 15 Å (Figure 4.6B, Appendix B.2 Movie 4.3), 20 Å (Figure 4.6C, Appendix B.2 Movie 4.4) and 30 Å (Figure 4.6D, Appendix B.2 Movie 4.5), could run timesteps of up to 300 fs, 30 ps and 100 ps, respectively, without crashing due to element inversion. The larger timesteps afforded by meshes built in this way will increase computational efficiency and allow access to longer timescales, which is what FFEA was built for.

The algorithm has since been adapted by Jarvellis Rogers, Jonathan Pickering and Joanna Leng for command line implementation, which can control the whole process from one script, and to smooth the shape of the mesh surface to more closely match the shape of the molecule.

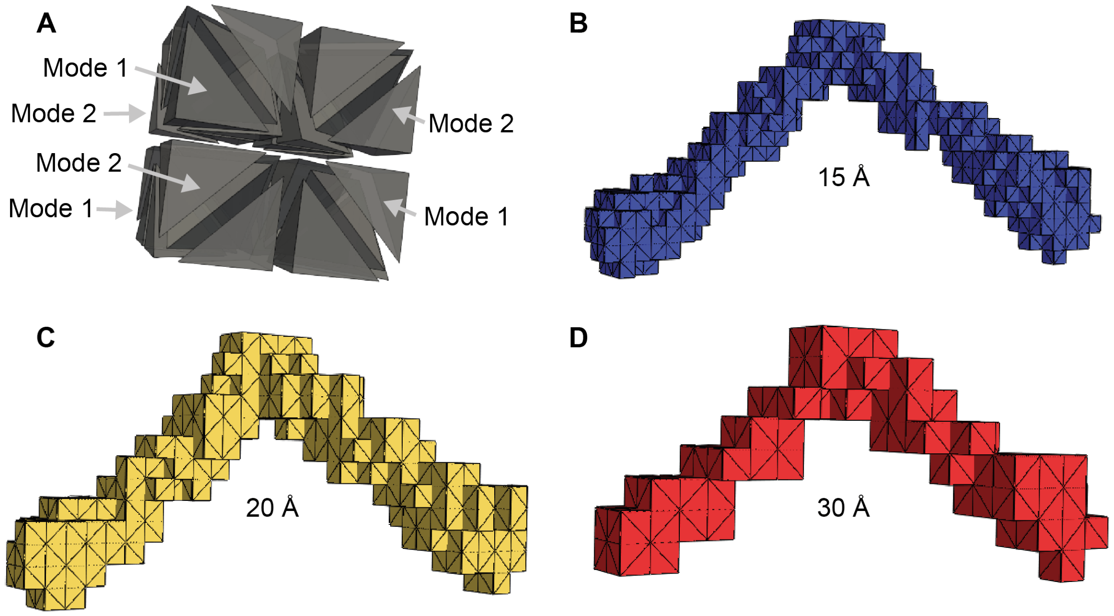


Figure 4.6: Marching tetrahedra meshes
A: *2x2x2 voxel cube marching tetrahedra mesh. Mesh has been blown out using Paraview* (Ayachit, 2015) *to highlight alternating modes of subdivision of neighbouring cubes into 5 tetrahedra.* ***B****,* ***C*** & ***D:*** *marching tetrahedra finite element meshes of the equilibrium model with increasing voxel size (15 Å, 20 Å & 30 Å, respectively) visualised using the FFEA visualiser* (Solernou et al., 2018) *PyMOL* (Schrödinger, 2015) *plugin.*

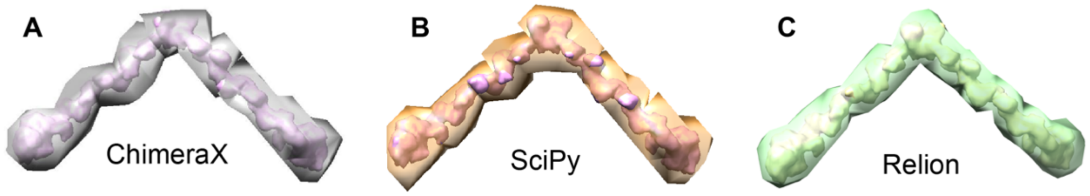


Figure 4.7: cryo-EM map coarsening approaches
*Map of equilibrium state coarsened to 30 Å voxels using volume voxelSize in ChimeraX (**A**), scipy.ndimage.zoom in SciPy (**B**) and relion_image_handler in Relion 3.1 (**C**). Relion provides the best surface interpolation.*

4.3.4 Model of HMM under strain

Following the creation of stable meshes, it was possible to generate a model of the strained state in FFEA (Figure 4.8, Appendix B.2 Movie 4.6-4.8). Due to the additional resistance created by increasing the rod stiffness, it was necessary to use a combination of springs and pin restraints to represent actin binding in addition to the external forces, to position the lead motor in the strained state. The strained simulation was run in 2 steps. First, to simulate initial actin binding, springs were applied between nodes of the actin binding domain of the lead motor mesh and nodes of a cube mesh representing the actin binding site (Figure 4.8A-B, Appendix B.2 Movie 4.6). Second, to simulate stereospecific binding of the lead motor and generate torsional strain across the molecule, a pin was applied to a node at the centre of the top of the lead motor mesh and a second pin was applied to a node at the centre of the bottom of the lead motor mesh. This maintained the motor at the actin binding site and created a stable rotation axis (Figure 4.8C-D, Appendix B.2 Movie 4.7). Ctforces were then used to rotate the lead head into its actin bound orientation, using the axis of the pinned nodes as the rotation axis (Figure 4.8C-D, Appendix B.2 Movie 4.7). Throughout both of these simulation steps nodes equivalent to the actin binding domain of the trail motor were pinned to simulate the actin bound trail motor. Using this approach, it was possible to generate a model of the strained state with the lead motor positioned at a location equivalent to a binding site 13 actin subunits away from the trail motor. Unlike the unrealistic flexible first proof model (Figure 4.3D), strain is successfully translated from the motor to the lever in this model. This can be seen by rotation of the rod material axes (shown as blue vectors Figure 4.8D, Appendix B.2 Movie 4.7), showing a twist across the length of the lever following positioning of the motor, and thereby the successful generation of the strained state within FFEA.

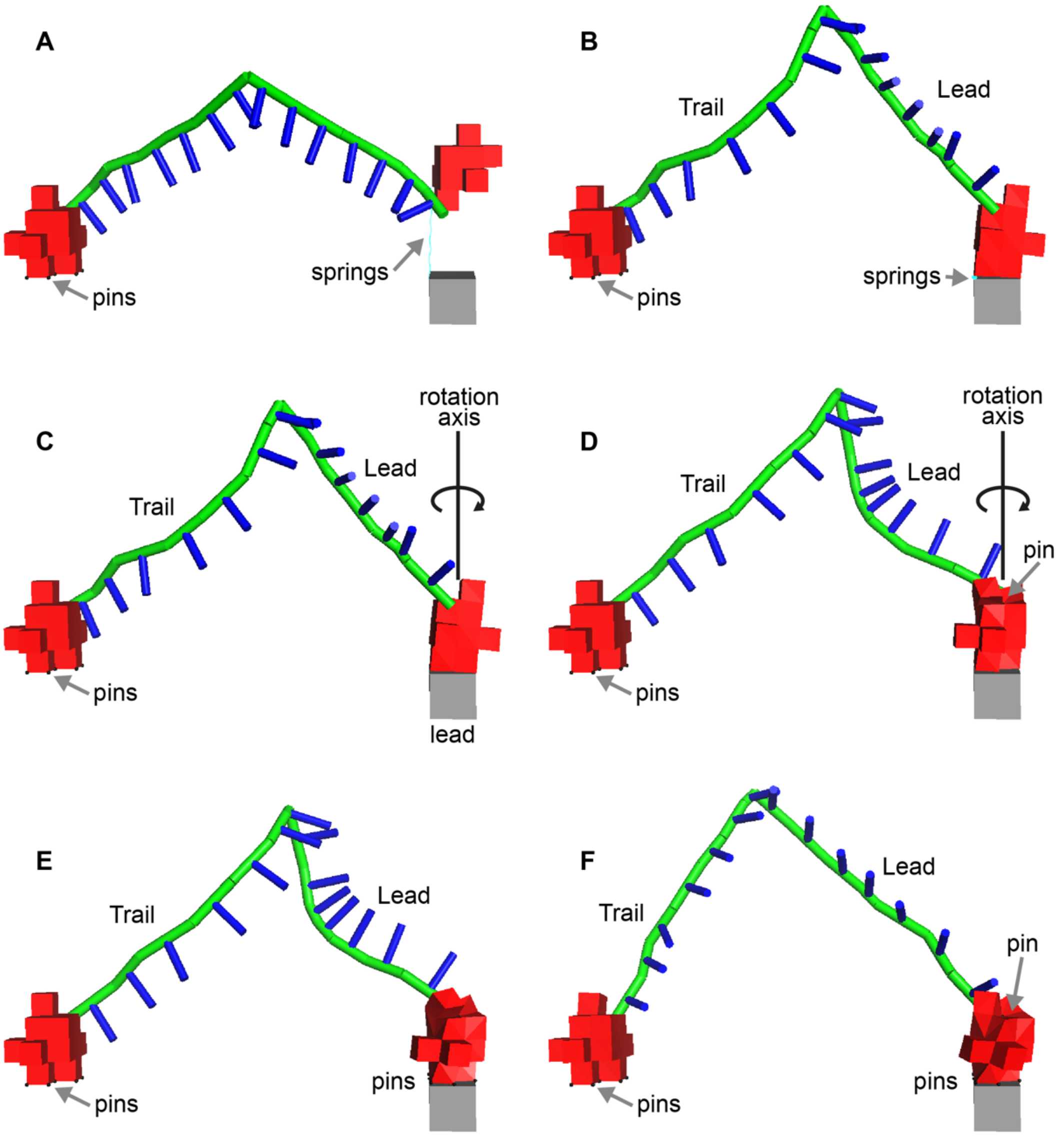


Figure 4.8: Setting up the strained state
A: *starting frame of the positioning simulation using the marching tetrahedra mesh. Springs (spring constant: 1 N/m, equilibrium distance: 1 Å) were applied between all the nodes at the base of the motor mesh (equivalent to the actin binding domain) and a cube mesh representing the actin binding site on the 13th subunit (only 1 spring can be displayed in the FFEA visualiser). The trail head (Trail) was pinned into place by pin restraints on all the nodes at the base of the motor mesh (equivalent to actin binding site). Rod indexing differs from blob indexing in the FFEA visualiser, therefore visualisation of the starting rod conformation was not possible (see* Figure 4.3*A for rod starting conformation).* ***B:*** *end frame of the spring simulation.* ***C:*** *start frame of the next step of the positioning simulation, the rotation simulation. Pins were applied to a node at the centre of the top and base of the lead (Lead) motor mesh to restrain the axis about which rotational ctforces (1 x 10^{-11} N) were applied to rotate the motor, to prevent motion of the motor away from the binding site.* ***D:*** *end frame of the rotation simulation.* ***E:*** *start frame of the next step in positioning, a pinned simulation. Pins were applied to the base of the lead motor (equivalent to actin binding site).* ***F:*** *end frame of pinned simulation.*

However, the rod-blob connection is too flexible to maintain torque in extended simulations (Figure 4.8E-F, Appendix B.2 Movie 4.8). Following positioning of the lead motor, simulations were run pinning the lead motor in place. In this simulation nodes equivalent to the actin binding domain and the rotation axis were pinned. During this simulation the material axes of the lead lever rod begin to return back to their original orientation showing torque is not maintained (Figure 4.8F, Appendix B.2 Movie 4.8). It is likely that this is a product of the rod-blob connection being too flexible. The converter domain is modelled as part of the lever rod model, however the coordination of the motor and the converter domain is important for translating conformational changes in the motor to the lever. Therefore, the connection between the motor and the converter must be sufficiently stiff that the converter moves in coordination with the motor. It will be necessary in future iterations to ensure the rod-blob connection is sufficiently stiff in order to maintain torque in the model.

In future work it would be necessary to use a more accurate rod model to improve generation of strain across the molecule. As a first attempt to model the strained state, the rod material parameters were left isotropic. To improve this model the anisotropic rod used in Chapter 3 should be used, which has different stiffness values at nodes along the lever. Only with this more accurate rod model can we be confident that the appropriate strain is being generated across the length of the lever and distributed appropriately. The final step in testing our model would be to simulate the binding of ATP to the trail head by releasing the pin restraints and seeing if the trail motor finds a path to the following actin binding site.

4.4 Conclusions

As experimental methods do not currently have the spatial or temporal resolution to visualise cytoskeletal motors taking a step, the development of *in silico* simulation approaches is necessary. Whilst all-atom MD simulation techniques are well established, mesoscale simulation tools such as FFEA are still in development. Here, for the first time, forces have been used to steer a FFEA simulation in order to simulate a molecule out of equilibrium and under strain. The methods demonstrated can be transferred to other systems, particularly cytoskeletal motors where strain is required for directional motion. This also provides the opportunity to think how else the FFEA ctforce function can be used to steer simulations in the future, what exciting energy barriers can we overcome?

FFEA presents new challenges divergent from traditional finite element analysis, due to the mesoscopic length scales and incorporation of thermal fluctuations. Coarse, geometrically favourable meshes become even more necessary, and therefore so has the development of new meshing approaches. The marching tetrahedra algorithm presented here demonstrates a way of moving directly from experimental data to simulation, and most importantly is fit for purpose, enabling access to timescales previously unobtainable.

Finally, through combining these developments I have generated a rod-blob model of Myo5a-HMM under directional strain. This further demonstrates the suitability of using a combined rod-blob modelling approach, modelling different domains in FFEA based on their properties. Using a blob model of the globular motor domain enabled controlled steering using ctforces, springs and pins, whilst the rod model of the lever enabled torsional strain to be stably induced across the molecule. Although there is much that needs to be done to improve upon this model of the strained state, it is a starting point from which Myo5a walking can be simulated. This model also offers a system through which hypotheses can be tested, such as flexural intramolecular strain instead of the torsional strain used here. Thanks to the improved meshing algorithm, stable long-scale dynamics provide the first opportunity to model structural dynamics on the timescale of Myo5a walking. Finally the methods developed here are broadly suitable for probing cytoskeletal motor mechanics in general.

Chapter 5 Summary, future perspectives and concluding remarks

5.1 Summary

This thesis identifies further properties of Myo5a that contribute to its walking mechanics, and highlights the benefits of using a combined computational and experimental approach. More generally, it describes a pipeline to go from cryo-EM data to a coarse-grained FFEA mesoscale model of protein dynamics (Figure 5.1).

Starting with the monomeric Myo5a-S1 construct (Chapter 2) I characterised features of the lever using cryo-EM. I resolved the structure of rigor Myo5a-S1 bound to actin in 9 different conformations, and demonstrated the benefit of using atomistic MD simulations to predict an ensemble of side chain orientations for model building. Comparison between the set of lever conformations in 3D allowed me to calculate the bending stiffness and showed that global lever flexibility was directionally isotropic (~0.76 pN/nm in both directions). Intra-lever flexibility could also be calculated as there was sufficient detail in the lever in the cryo-EM map to assign CaMs. This showed that there is varying flexibility within the lever (ranging from 136 pN nm to 1988 pN nm). Previous estimates of flexibility have not considered the full 3D space nor focused on the lever specifically. Cryo-EM is therefore a suitable tool as it has both high spatial resolution and 3D information. Using the lever conformations superimposed onto pre- and post-PS motor domain structures, I generated a distribution of working strokes and predicted a longer mean working stroke than previously recorded (~33 nm). This showed that Myo5a-S1 possesses all that is required to translate the lever C-terminal end along a single F-actin helical repeat. Thus, Myo5a is capable of accurately stepping along F-actin, with only a small diffusive search, if any.

In Chapter 3 I moved to coarse-grained FFEA modelling of Myo5a-S1 and the dimeric construct (Myo5a-HMM), in which I integrated the data derived by cryo-EM. In this chapter I demonstrated a workflow from cryo-EM data to an FFEA model, and discussed what is required to generate and validate a reliable FFEA model. Additionally, integration of cryo-EM data on Myo5a-S1 into a

FFEA rod model of the dimer showed the combined properties of the levers lead to a similar distribution of distances between the heads in the off-actin state as the F-actin bound states (Walker et al., 2000; Veigel et al., 2002; Oke et al., 2010). Using the rod model I explored the influence of varying flexibility at the head-tail junction in the Myo5a-HMM FFEA model, I showed that any stiffness at the head-tail junction is likely to be screened by flexibility in the lever.

Finally, within Chapter 4 are the first attempts at generating a FFEA model of Myo5a-HMM capable of walking, using a combined rod-blob FFEA model. More generally this chapter describes how to use forces in FFEA to generate non-equilibrium strained states, and presents a novel meshing algorithm. To generate a model of the strained state it was first necessary to improve mesh stability. Therefore, a novel marching tetrahedra meshing algorithm was developed, which generates tetrahedral meshes directly from cryo-EM map voxel data. This improved the timestep accessible in FFEA ~1000-fold. Following this, a model of Myo5a-HMM under torsional strain was generated, which provides a starting point for modelling Myo5a walking.

Taken together, the new insight into properties of Myo5a presented here offers a new perspective on walking mechanics. These challenge preconceived notions on lever behaviour, and consequently we require the generation of new theories on the walking mechanism. Previously hypotheses have been based on a uniform rod with flexibility only occurring at a pliant point, or localised between 25-residue spaced CaM pairs. Therefore, the distribution of flexibility along the length of the lever described here contradicts these. The appearance of compass-like walking has led to theories on a constrained head-tail junction. However, my data suggests this would have little to no effect in the overall molecule. This is not to say compass-like walking does not occur, but perhaps our macroscale preconceptions of motor mechanics are not applicable at this scale, and a stiff joint is not required in such a viscous environment. Ultimately further validation is required to properly interpret these findings, and it is likely that a combined experimental and computational approach is necessary.

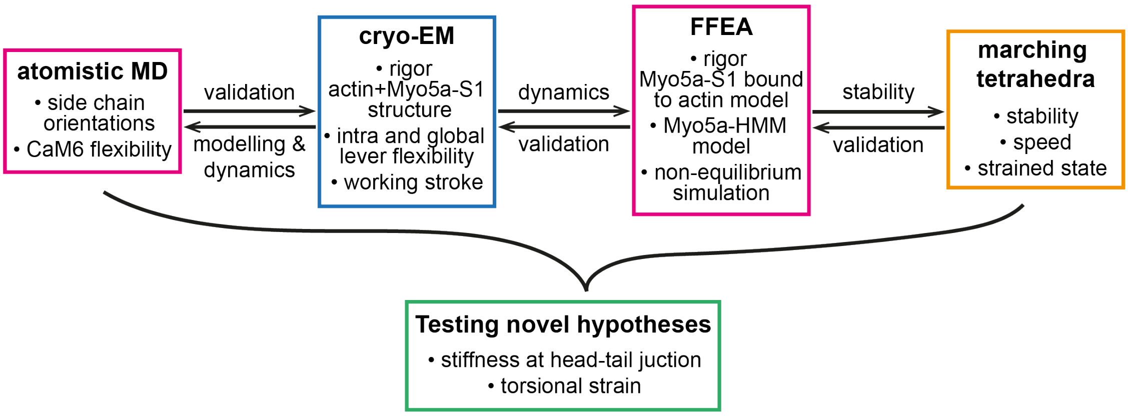


Figure 5.1: Pipeline of work presented here
Blue boxes indicate experimental techniques. Pink boxes highlight pre-existing computational tools. Orange boxes indicate tools developed here. Green boxes indicate an integrated approach.

5.2 Future perspectives

3D reconstruction of dimeric Myo5a in the active state would offer invaluable insights into the mechanics of this molecule. Particularly the structure of HMM bound to actin, which would allow direct interpretation of the strain the lead head is under (i.e. torsional or flexural), when compared with the ADP-bound post-PS state (Pospich et al., 2021). This will by no means be easy, due to the heterogeneity of the molecule, however the pipeline used here to resolve Myo5a-S1 would be applicable. If the resolution was low, focused classification of key regions of interest would still be informative. For example, solely the lead head motor domain should be sufficient to determine the type of strain that is induced. Additionally, resolution of the head-tail junction in the active state is of particular interest in order to understand how the 2 levers communicate, and how strain is transmitted across the molecule. The recent structure of Myo5a in the shutdown state resolves the head-tail junction to high-resolution, but the N-lobes of CaM6 in both heads interact with the folded coiled-coil domain, indicating the head-tail junction has a different conformation in the active state (Niu et al., 2022). Of course, the structure of the molecule off F-actin would also be interesting, in order to gain the ATP bound equilibrium state, although this is likely to be highly flexible and therefore challenging to resolve (Coureux et al., 2004; Pospich et al., 2021).

Furthermore, cryo-ET could provide the perfect tool to gain material parameters from single molecules in 3D, as it is possible to develop a 3D reconstruction from a single particle. This would resolve previous issues from calculating bending stiffnesses from cryo-EM 3D classes or 2D ns-EM micrographs. If single molecule 3D reconstructions from cryo-ET were used instead of cryo-EM 3D classes, the full ensemble of lever conformations would be captured whilst retaining 3D information. This would resolve the dependence on acquiring sufficient particles to reconstruct a class, which excludes extreme lever conformations. Additionally it would resolve the loss of conformational variance arising from the averaging of similar lever conformations in order to generate a 3D class. The caveat is that without particle averaging the resolution that is obtainable is limited, thus it may not be possible to differentiate lever subdomains. The combination of cryo-ET and FFEA could be really powerful. For example, the low-resolution single molecule reconstructions from cryo-ET could be used to generate a discretised shape, by fitting them with a specific number of pseudo atoms as in (Zhang et al., 2020). The conformations of the discretised shapes could then be directly compared to accurately gain material parameters. The pseudo atoms could then be used to generate an FFEA rod model with a specific number of nodes, providing a reliable and testable model. Additionally, as the marching tetrahedra mode of meshing generates regular shaped meshes, it may allow easier mapping between meshes for a similar approach to this, but with meshes.

Deriving parameters from cryo-EM data could be improved, even automated, by adapting cryo-EM flexibility analysis software that already exists, e.g. Relion Multi-body Refinement (Nakane et al., 2018), cryoDRGN (Zhong et al., 2021), or 3D Variability Analysis in CryoSPARC (Punjani and Fleet, 2021). All of these techniques return modes of motion, and unlike 3D classification, which shows distinct conformations, show a continuous array of conformations. Hinge regions could be identified from this data, the angles between mobile subdomains could be used to calculate bending stiffness values, and the magnitudes of the individual modes could be used to determine isotropy. Perhaps the magnitudes of the modes could even be correlated with the bending stiffness value?

There are a wealth of interesting simulations that could be run in continuation of this project. The most basic being the Myo5a-HMM rod model with a stiff lever and flexibility at head-tail junction. This would be useful to compare to the simulations presented here in order to directly interpret how flexibility in the lever influences the emergent features of the molecule. Like a mutagenesis study, this would simulate what happens when these features are removed to determine their influence. Future simulations of the strained state would be improved by incorporating the cryo-EM derived material parameters in the lever rod models. Also, testing the influence of a flexible head-tail junction in the strained state would be interesting. This would shed light on how the head-tail junction influences strain distribution across the molecule.

Ultimately the goal was to model Myo5a stepping, and gain insight into the transient state (as it takes a step). Therefore, the natural progression of this research would be to simulate the release of the trail head from the strained state model in order to visualise stepping. A stepping model would also provide insight on how motors operate at this scale, and within a highly viscous environment. This would allow direct interpretation of how different lever flexibilities behave in the moving molecule. For example, testing what the constrained head-tail junction means in the transient state as the molecule is in motion; determining how much flexibility causes it to deform under viscous drag; or looking at why SAHs are functional as levers (Baboolal et al., 2009). This would also give a better idea of how important stiffness is for mechanics in a viscous environment.

There are a multitude of functions that would be interesting to incorporate into the FFEA software based on my findings. Firstly, expanding the rod code so that functions currently only applicable to blobs can be applied to rods, and vice versa, would be beneficial. Such as pinning and ctforces for rods, and the ability to assign “equilibrium” and “current” conformations for blobs. This would aid the setup of non-equilibrium states. Additionally, as collision detection/steric repulsion is incorporated into the rod code by Ryan Cocking, rods will move from being 1D to 3D objects, thus how relevant is attachment to a single tetrahedra? The rod-blob connection will need re-thinking, and a connection instead to multiple tetrahedra is likely to be more stable for strained molecules.

In terms of developing the marching tetrahedra algorithm, this has been taken over by Joanna Leng and Jonathan Pickering to improve the surface of these meshes, and assign material parameters during mesh generation. However, adapting the algorithm to also generate rods, or combinations of rods and blobs depending on the shape of the cryo-EM map would be a useful feature. For this to be automated may require machine learning to predict the type of representation (blob or rod) based on the shape of the map.

5.3 Concluding remarks

In conclusion, I have revealed key features of Myo5a mechanics, through a combination of cryo-EM and computational modelling, and in the process developed a pipeline to go from cryo-EM data to FFEA coarse-grained models. Through this synergistic combination of techniques, I was able to move from a single snapshot towards a detailed model of how Myo5a functions.

Myo5a belongs to the wider myosin superfamily. The general structure of myosin is highly conserved, therefore comprehension of the mechanics of Myo5a is largely applicable across the superfamily. This work contributes towards understanding how each myosin has evolved different lever structures and properties to carry out their diverse array of functions. As the myosin superfamily has such a diverse array of functions, its members are implicated in a multitude of disease states. Thus, improving our understanding of myosin mechanics has comparable significance to the field of medicine as our understanding of fundamental biology.

In addition, this work further validates FFEA as a suitable tool for modelling cytoskeletal motors, and demonstrates new ways in which it can be used. The pipelines discussed here would be applicable to all cytoskeletal motors, and I look forward to seeing what FFEA is applied to next. Understanding how molecules behave on this scale is particularly interesting and has applications for engineering, for example designing nanobots that are required to operate in these viscous environments (Nelson et al., 2010).

Overall, this work highlights that only with a combined computational and experimental approach are we able to access the time and length scales of interest for cytoskeletal motor dynamics. Moreover, that considering dynamics alongside structural data is necessary. This is particularly important for motor proteins as dynamics are integral to their function, but also for proteins in general. Proteins are not static they are highly dynamic objects subject to thermal fluctuations. Now, with tools that can simulate dynamics over time and length scales of interest, we are able to access dynamic states previously unexplored.

Appendix A
Residue interaction tables for atomistic MD simulations

A.1 Residue contacts (within 5 Å) in the pseudoatomic model

https://www.dropbox.com/s/y4lrdrj0ql4kqav/A.1_table_of_contacts_pseudoatomic_model.pdf

*Average presence refers to the percentage of time the interaction is present in the trajectory, averaged over the 3 repeats. N/A refers to interactions not seen in all 3 repeats. The N-lobe of CaM6 is highlighted in grey for CaM6-IQ6 interactions. Interactions conserved in published structures labelled with ** (Coureux et al., 2003; Terrak et al., 2005; Houdusse et al., 2006; Pospich et al., 2021).

A.2 Residue contacts (within 5 Å) in the restrained MD simulations

https://www.dropbox.com/s/snew20nmyrijx5o/A.2_table_of_contacts_restrained_md.pdf

*Average presence refers to the percentage of time the interaction is present in the trajectory, averaged over the 3 repeats. The N-lobe of CaM6 is highlighted in grey for CaM6-IQ6 interactions. Interactions conserved in published structures labelled with ** (Coureux et al., 2003; Terrak et al., 2005; Houdusse et al., 2006; Pospich et al., 2021). *Contacts seen in the pseudoatomic model labelled with +.*

A.3 Residue contacts (within 5 Å) in the unrestrained MD simulations

https://www.dropbox.com/s/kw6cl2nwsl6i9zt/A.3_table_of_contacts_unrestrained_MD.pdf

*Average presence refers to the percentage of time the interaction is present in the trajectory, averaged over the 3 repeats. The N-lobe of CaM6 is highlighted in grey for CaM6-IQ6 interactions. Interactions conserved in published structures labelled with ** (Coureux et al., 2003; Terrak et al., 2005; Houdusse et al., 2006; Pospich et al., 2021)

Appendix B
Movie links

B.1 Chapter 3

Movie 3.1: Principal modes of motion

https://www.dropbox.com/s/5n8g7apom2ejf7a/Movie_3.1_PCA.mov

Principal modes of motion for the gaussian filtered (Gaussian, grey), and re-scaled then gaussian filtered (Scaled + gaussian, red), FFEA blob simulations. PCA was performed using pyPcazip (Shkurti et al., 2016). The pseudo atoms are displayed as a surface. VMD was used to visualise the modes (Humphrey et al., 1996). PC = principal component.

B.2 Chapter 4

Movie 4.1: Positioning lead motor using forces

https://www.dropbox.com/s/kj37dpofnknk67o/Movie_4.1_positioning_sim.mp4

Rotational forces (1 x 10^{-11} N) were applied to the lead motor about the rotation axis (yellow). The simulation was stopped when the lead motor was rotated to the point it matched the orientation of the trail motor. This is the starting frame for the translation simulation, the translocation axis along which linear forces (1 x 10^{-11} N) were applied is highlighted (blue arrow). The simulation was stopped when the lead motor was at the actin-binding site. The grey bars represent the rod material axes, which are perpendicular to the rod elements, and are visualised to show twist in the rod.

Movie 4.2: Element inversion occurs when rod stiffness increased

https://www.dropbox.com/s/eppdwitzbtl1uxi/Movie_4.2_inversion.mp4

Zoomed view of the lead motor mesh inverting. Rod displayed in green, material axes displayed in grey, mesh of motor domain shown in red with elements outlined in black.

Movie 4.3: 15 Å marching tetrahedra mesh

https://www.dropbox.com/s/vgqyas2n46xijq1/Movie_4.3_15A.mp4

FFEA simulation of the marching tetrahedra mesh of the equilibrium model, with a voxel size of 15 Å, shear modulus of 9.2 x 10^6 N m^{-2} and bulk modulus of 2.8 x 10^7 N m^{-2}. The simulation was visualised using the FFEA visualiser (Solernou et al., 2018) PyMOL (Schrödinger, 2015) plugin.

Movie 4.4: 20 Å marching tetrahedra mesh

https://www.dropbox.com/s/bazs4tv0fa3j2xc/Movie_4.4_20A.mp4

FFEA simulation of the marching tetrahedra mesh of the equilibrium model, with a voxel size of 20 Å, shear modulus of 9.2 x 10^6 N m^{-2} and bulk modulus of 2.8 x 10^7 N m^{-2}. The simulation was visualised using the FFEA visualiser (Solernou et al., 2018) PyMOL (Schrödinger, 2015) plugin.

Movie 4.5: 30 Å marching tetrahedra mesh

https://www.dropbox.com/s/t0couy71f9ywzt1/Movie_4.5_30A.mp4

FFEA simulation of the marching tetrahedra mesh of the equilibrium model, with a voxel size of 30 Å, shear modulus of 9.2 x 10^6 N m^{-2} and bulk modulus of 2.8 x 10^7 N m^{-2}. The simulation was visualised using the FFEA visualiser (Solernou et al., 2018) PyMOL (Schrödinger, 2015) plugin.

Movie 4.6: Step 1 of positioning the strained state using springs

https://www.dropbox.com/s/aa9k7ixlwajlpk9/Movie_4.6_spring.mp4

Positioning simulation of the combined blob-rod model using the marching tetrahedra mesh of the motor domain. To model initial actin binding springs (cyan), with a spring constant of 1 N/m and equilibrium distance of 1 Å, were applied between all the nodes at the base of the lead motor mesh (equivalent to the actin binding domain) and a cube mesh representing the actin binding site on the 13th subunit (only 1 spring can be displayed in the FFEA visualiser). The trail head was pinned into place by pin restraints (black spheres) on all the nodes at the base of the motor mesh (equivalent to actin binding site). The simulation was visualised using the FFEA visualiser (Solernou et al., 2018) PyMOL (Schrödinger, 2015) plugin. Rod displayed in green, material axes displayed in blue, mesh of motor domain shown in red.

Movie 4.7: Step 2 of positioning the strained state using pins

https://www.dropbox.com/s/c9sne7bwl702x2i/Movie_4.7_pin_twist.mp4

Positioning simulation of the combined blob-rod model using the marching tetrahedra mesh of the motor domain. To prevent motion of the motor domain away from the binding site, pins (black spheres) were applied to a node at the centre of the top and base of the lead motor domain mesh to restrain the axis about which rotational ctforces (1 x 10^{-11} N) were applied to rotate the motor. The simulation was visualised using the FFEA visualiser (Solernou et al., 2018) PyMOL (Schrödinger, 2015) plugin. Rod displayed in green, material axes displayed in blue, mesh of motor domain shown in red.

Movie 4.8: Simulation of the strained state pinning the motor domains

https://www.dropbox.com/s/gut5o4ljntiu0jc/Movie_4.8_pin_hold.mp4

Extended pinned simulation of the combined blob-rod model using the marching tetrahedra mesh of the motor domain. To assess whether torsion is maintain across the molecule pins (black spheres) were applied to the base of the trail and lead motor (equivalent to actin binding site). The simulation was visualised using the FFEA visualiser (Solernou et al., 2018) PyMOL (Schrödinger, 2015) plugin. Rod displayed in green, material axes displayed in blue, mesh of motor domain shown in red.

Appendix C
Publications

In the lead up to this project 3 manuscripts were published. The first (C.1) involved characterising the difference between disease and non-disease associated variants in G-Protein coupled receptor kinase 1 (GRK1) using *in silico* methods. By generating a homology model of GRK1, we were able to investigate the structural implications of variants, and assess scoring tools that were most capable of distinguishing between disease and non-disease associated variants. Of the scoring tools tested, Rhapsody, which includes dynamics in its prediction through elastic network modelling, performed best. This emphasises how important is to consider both structure and dynamics in protein function. In conjunction with this a back cover was designed for the issue (C.2).

The second (C.3) was a study investigating cryo-EM grid preparation techniques. We directly compared a custom-built high-speed grid preparation device with the commercially available Vitrobot Mk IV. This highlighted that increased grid preparation speeds reduced: preferred particle orientation, particle interaction with the air-water interface, and particle degradation. Although the high-speed grid preparation device was not used in later work, the effects of using the Vitrobot Mk IV for grid preparation are considered throughout.

Finally, the third (C.4) is a review discussing what is required to simulate cytoskeletal motors, informed by microscopy data (used in Chapter 1). We discuss both the experimental and simulation techniques that already exist, alongside what physics needs to be considered in order to combine them, and what is needed from the community. Cytoskeletal motors are large proteins with long-range dynamics, which makes them both interesting and challenging systems to research. Currently, dynamics of interest are too fast to capture experimentally, and too slow to capture via simulation. New techniques are required to integrate experimental data and simulate transitional dynamics.

C.1 New variants and in silico analyses in GRK1 associated Oguchi disease

Poulter, JA*, **Gravett, MSC***, Taylor, RL, et al. New variants and in silico analyses in GRK1 associated Oguchi disease. *Human Mutation*. 2021; 42: 164-176. https://doi.org/10.1002/humu.24140

C.2 Cover

Poulter, JA*, **Gravett, MSC***, Taylor, RL, et al. Cover, Volume 42, Issue 2. *Human Mutation*. 2021; 42: ii-ii. https://doi.org/10.1002/humu.24169

C.3 Need for Speed: Examining Protein Behaviour during Cryo-EM Grid Preparation at Different Timescales

Klebl, DP, **Gravett, MSC**, Kontziampasis, D, et al. Need for Speed: Examining Protein Behaviour during Cryo-EM Grid Preparation at Different Timescales. *Structure*. 2020; 28: 1238-1248.e4. https://doi.org/10.1016/j.str.2020.07.018

C.4 Moving in the mesoscale: Understanding the mechanics of cytoskeletal molecular motors by combining mesoscale simulations with imaging

Gravett, MSC, Cocking, RC, Curd, AP, Harlen, O, Leng, J, Muench, SP, et al. Moving in the mesoscale: Understanding the mechanics of cytoskeletal molecular motors by combining mesoscale simulations with imaging. *WIREs Comput Mol Sci*. 2022; 12:e1570. https://doi.org/10.1002/wcms.1570

Bibliography

Ali, M.Y., Uemura, S., Adachi, K., Itoh, H., Kinosita, K. and Ishiwata, S. 2002. Myosin v is a left-handed spiral motor on the right-handed actin helix. *Nature Structural Biology*. **9**(6), pp.464–467.

Alves, C.P., Moraes, M.H., Sousa, J.F., Pontes, C.L.S., Ramão, A., Yokoyama, S., Trindade, D.M., Fisher, D.E. and Espreafico, E.M. 2013. Myosin-Va contributes to manifestation of malignant-related properties in melanoma cells. *Journal of Investigative Dermatology*. **133**(12), pp.2809–2812.

Ando, T. 2019. High-speed atomic force microscopy. *Current Opinion in Chemical Biology*. **51**, pp.105–112.

Andrecka, J., Ortega Arroyo, J., Takagi, Y., de Wit, G., Fineberg, A., MacKinnon, L., Young, G., Sellers, J.R. and Kukura, P. 2015. Structural dynamics of myosin 5 during processive motion revealed by interferometric scattering microscopy. *eLife*. **4**, p.e05413.

Ayachit, U. 2015. *The ParaView Guide: A Parallel Visualization Application*. Clifton Park, NY, USA: Kitware, Inc.

Baboolal, T.G., Sakamoto, T., Forgacs, E., White, H.D., Jackson, S.M., Takagi, Y., Farrow, R.E., Molloy, D.E., Knight, P.J., Sellers, J.R. and Peckhama, M. 2009. The SAH domain extends the functional length of the myosin lever. *Proceedings of the National Academy of Sciences of the United States of America*. **106**(52), pp.22193–22198.

Bähler, M. and Rhoads, A. 2002. Calmodulin signaling via the IQ motif. *FEBS Letters*. **513**(1), pp.107–113.

Bakan, A., Dutta, A., Mao, W., Liu, Y., Chennubhotla, C., Lezon, T.R. and Bahar, I. 2014. Evol and ProDy for bridging protein sequence evolution and structural dynamics . *Bioinformatics*. **30**(18), pp.2681–2683.

Bakan, A., Meireles, L.M. and Bahar, I. 2011. ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics*. **27**(11), pp.1575–1577.

Banerjee, C., Hu, Z., Huang, Z., Warrington, J.A., Taylor, D.W., Trybus, K.M., Lowey, S. and Taylor, K.A. 2017. The structure of the actin-smooth muscle myosin motor domain complex in the rigor state. *Journal of Structural Biology*. **200**(3), pp.325–333.

Barrick, S.K., Greenberg, M.J., De, E. and Cruz, L. 2021. Cardiac myosin contraction and mechanotransduction in health and disease. *Journal of Biological Chemistry*. **297**(5), p.101297.

Beausang, J.F., Schroeder, H.W., Nelson, P.C. and Goldman, Y.E. 2008. Twirling of actin by myosins II and V observed via polarized TIRF in a modified gliding assay. *Biophysical Journal*. **95**(12), pp.5820–5831.

Beckers, M. and Sachse, C. 2020. Permutation testing of Fourier shell correlation for resolution estimation of cryo-EM maps. *Journal of Structural Biology*. **212**(1), p.107579.

Berg, J.S., Powell, B.C. and Cheney, R.E. 2001. A millennial myosin census.

Molecular Biology of the Cell. **12**(4), pp.780–794.

Billington, N., Revill, D.J., Burgess, S.A., Chantler, P.D. and Knight, P.J. 2014. Flexibility within the Heads of Muscle Myosin-2 Molecules. *Journal of Molecular Biology*. **426**(4), pp.894–907.

Black, D.J. and Persechini, A. 2010. Variations at the semiconserved glycine in the IQ domain consensus sequence have a major impact on Ca2+-dependent switching in calmodulin-IQ domain complexes. *Biochemistry*. **49**(1), pp.78–83.

Bock, L. V. and Grubmüller, H. 2022. Effects of cryo-EM cooling on structural ensembles. *Nature Communications 2022 13:1*. **13**(1), pp.1–13.

Bowater, R., Zimmerman, R.W. and Webb, M.R. 1990. Kinetics of ATP and inorganic phosphate release during hydrolysis of ATP by rabbit skeletal actomyosin subfragment 1. Oxygen exchange between water and ATP or phosphate. *Journal of Biological Chemistry*. **265**(1), pp.171–176.

Brown, A.I. and Sivak, D.A. 2020. Theory of Nonequilibrium Free Energy Transduction by Molecular Machines. *Chemical Reviews*. **120**(1), pp.434–459.

Burgess, S., Walker, M., Wang, F., Sellers, J.P., White, H.D., Knight, P.J. and Trinick, J. 2002. The prepower stroke conformation of myosin V. *Journal of Cell Biology*. **159**(6), pp.983–991.

Burgess, S.A., Walker, M.L., Thirumurugan, K., Trinick, J. and Knight, P.J. 2004. Use of negative stain and single-particle image processing to explore dynamic properties of flexible macromolecules. *Journal of Structural Biology*. **147**(3), pp.247–258.

Bustamante, C., Keller, D. and Oster, G. 2001. The physics of molecular motors. *Accounts of Chemical Research*. **34**(6), pp.412–420.

Case, D.A., Aktulga, H.M., Belfon, K., Ben-Shalom, I.Y., Berryman, J.T., Brozell, S.R., Cerutti, D.S., Cheatham, T.E., I., Cisneros, G.A., Cruzeiro, V.W.D., Darden, T.A., Duke, R.E., Giambasu, G., Gilson, M.K., Gohlke, H., Goetz, A.W., Harris, R., Izadi, S., Izmailov, S.A. and Kollman, P.A. 2020. Amber 2020.

Cecchini, M., Houdusse, A. and Karplus, M. 2008. Allosteric Communication in Myosin V: From Small Conformational Changes to Large Directed Movements M. P. Jacobson, ed. *PLoS Computational Biology*. **4**(8), p.e1000129.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. 2009. MolProbity: all-atom structure validation for macromolecular crystallography. *urn:issn:0907-4449*. **66**(1), pp.12–21.

Cheney, R.E., O'shea, M.K., Heuser, J.E., Coelho, M. V, Wolenski, J.S., Espraafico, E.M., Forecher, P., Larson, R.E. and Mooseker, M.S. 1993. Brain Myosin-V Is a Two-Headed Unconventional Myosin with Motor Activity. *Cell*. **75**(1), pp.13–23.

Cheng, A., Henderson, R., Mastronarde, D., Ludtke, S.J., Schoenmakers, R.H.M., Short, J., Marabini, R., Dallakyan, S., Agard, D. and Winn, M. 2015. MRC2014: Extensions to the MRC format header for electron cryo-microscopy and tomography. *Journal of Structural Biology*. **192**(2), pp.146–150.

Cheng, Y. 2015. Single-particle Cryo-EM at crystallographic resolution. *Cell*. **161**(3), pp.450–457.

Cirilo, J.A., Gunther, L.K. and Yengo, C.M. 2021. Functional Role of Class III Myosins in Hair Cells. *Frontiers in Cell and Developmental Biology*. **9**, p.285.

Clare, D.K., Vasishtan, D., Stagg, S., Quispe, J., Farr, G.W., Topf, M., Horwich, A.L. and Saibil, H.R. 2012. ATP-triggered conformational changes delineate substrate-binding and -folding mechanics of the GroEL chaperonin. *Cell*. **149**(1), pp.113–123.

Consortium, T.U., Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Ahmad, S., Alpi, E., Bowler-Barnett, E.H., Britto, R., Bye-A-Jee, H., Cukura, A., Denny, P., Dogan, T., Ebenezer, T., Fan, J., Garmiri, P., da Costa Gonzales, L.J., Hatton-Ellis, E., Hussein, A., Ignatchenko, A., Insana, G., Ishtiaq, R., Joshi, V., Jyothi, D., Kandasaamy, S., Lock, A., Luciani, A., Lugaric, M., Luo, J., Lussi, Y., MacDougall, A., Madeira, F., Mahmoudy, M., Mishra, A., Moulang, K., Nightingale, A., Pundir, S., Qi, G., Raj, S., Raposo, P., Rice, D.L., Saidi, R., Santos, R., Speretta, E., Stephenson, J., Totoo, P., Turner, E., Tyagi, N., Vasudev, P., Warner, K., Watkins, X., Zaru, R., Zellner, H., Bridge, A.J., Aimo, L., Argoud-Puy, G., Auchincloss, A.H., Axelsen, K.B., Bansal, P., Baratin, D., Batista Neto, T.M., Blatter, M.-C., Bolleman, J.T., Boutet, E., Breuza, L., Gil, B.C., Casals-Casas, C., Echioukh, K.C., Coudert, E., Cuche, B., de Castro, E., Estreicher, A., Famiglietti, M.L., Feuermann, M., Gasteiger, E., Gaudet, P., Gehant, S., Gerritsen, V., Gos, A., Gruaz, N., Hulo, C., Hyka-Nouspikel, N., Jungo, F., Kerhornou, A., Le Mercier, P., Lieberherr, D., Masson, P., Morgat, A., Muthukrishnan, V., Paesano, S., Pedruzzi, I., Pilbout, S., Pourcel, L., Poux, S., Pozzato, M., Pruess, M., Redaschi, N., Rivoire, C., Sigrist, C.J.A., Sonesson, K., Sundaram, S., Wu, C.H., Arighi, C.N., Arminski, L., Chen, C., Chen, Y., Huang, H., Laiho, K., McGarvey, P., Natale, D.A., Ross, K., Vinayaka, C.R., Wang, Q., Wang, Y. and Zhang, J. 2022. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*. **gkac1052**, pp.13–14.

Coureux, P.D., Sweeney, H.L. and Houdusse, A. 2004. Three myosin V structures delineate essential features of chemo-mechanical transduction. *EMBO Journal*. **23**(23), pp.4527–4537.

Coureux, P.D., Wells, A.L., Ménétrey, J., Yengo, C.M., Morris, C.A., Sweeney, H.L. and Houdusse, A. 2003. A structural state of the myosin V motor without bound nucleotide. *Nature*. **425**(6956), pp.419–423.

Craig, E.M. and Linke, H. 2009. Mechanochemical model for myosin V. *Proceedings of the National Academy of Sciences of the United States of America*. **106**(43), pp.18261–18266.

Croll, T.I. 2018. ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallographica Section D: Structural Biology*. **74**(6), pp.519–530.

Dong, Y., Zhang, S., Wu, Z., Li, X., Wang, W.L., Zhu, Y., Stoilova-McPhie, S., Lu, Y., Finley, D. and Mao, Y. 2019. Cryo-EM structures and dynamics of substrate-engaged human 26S proteasome. *Nature*. **565**(7737), pp.49–55.

Dupuis, D.E., Guilford, W.H., Wu, J. and Warshaw, D.M. 1997. Actin filament mechanics in the laser trap. *Journal of Muscle Research and Cell Motility*. **18**(1), pp.17–30.

Von Der Ecken, J., Heissler, S.M., Pathan-Chhatbar, S., Manstein, D.J. and Raunser, S. 2016. Cryo-EM structure of a human cytoplasmic actomyosin complex at near-atomic resolution. *Nature*. **534**(7609), pp.724–728.

Espindola, F.S., Suter, D.M., Partata, L.B.E., Cao, T., Wolenski, J.S., Cheney, R.E., King, S.M. and Mooseker, M.S. 2000. The Light Chain Composition of Chicken Brain Myosin-Va: Calmodulin, Myosin-II Essential Light Chains, and 8-kDa Dynein Light Chain/PIN. *Cell Motility Cytoskeleton*. **47**(4), pp.269–281.

Fernandez-Leiro, R. and Scheres, S.H.W. 2016. Unravelling biological macromolecules with cryo-electron microscopy. *Nature*. **537**(7620), pp.339–346.

Forgacs, E., Cartwright, S., Kovács, M., Sakamoto, T., Sellers, J.R., Corrie, J.E.T., Webb, M.R. and White, H.D. 2006. Kinetic mechanism of myosin V-Sl using a new fluorescent ATP analogue. *Biochemistry*. **45**(43), pp.13035–13045.

Forgacs, E., Cartwright, S., Sakamoto, T., Sellers, J.R., Corrie, J.E.T., Webb, M.R. and White, H.D. 2008. Kinetics of ADP dissociation from the trail and lead heads of actomyosin V following the power stroke. *Journal of Biological Chemistry*. **283**(2), pp.766–773.

Forgacs, E., Sakamoto, T., Cartwright, S., Belknap, B., Kovács, M., Tóth, J., Webb, M.R., Sellers, J.R. and White, H.D. 2009. Switch 1 mutation S217A converts myosin V into a low duty ratio motor. *Journal of Biological Chemistry*. **284**(4), pp.2138–2149.

Foth, B.J., Goedecke, M.C. and Soldati, D. 2006. New insights into myosin evolution and classification. *Proceedings of the National Academy of Sciences of the United States of America*. **103**(10), pp.3681–3686.

Fujii, T. and Namba, K. 2017. Structure of actomyosin rigour complex at 5.2 Å resolution and insights into the ATPase cycle mechanism. *Nature Communications*. **8**(1), pp.1–11.

Geeves, M.A. 2016. Review: The ATPase mechanism of myosin and actomyosin. *Biopolymers*. **105**(8), pp.483–491.

Geeves, M.A. and Holmes, K.C. 1999. Structural Mechanism of Muscle Contraction. *Annual Review of Biochemistry*. **68**(1), pp.687–728.

Van Gele, M., Dynoodt, P. and Lambert, J. 2009. Griscelli syndrome: a model system to study vesicular trafficking. *Pigment Cell & Melanoma Research*. **22**(3), pp.268–282.

Gravett, M.S.C., Cocking, R.C., Curd, A.P., Harlen, O., Leng, J., Muench, S.P., Peckham, M., Read, D.J., Rogers, J.F., Welch, R.C. and Harris, S.A. 2022. Moving in the mesoscale: Understanding the mechanics of cytoskeletal molecular motors by combining mesoscale simulations with imaging. *Wiley Interdisciplinary Reviews: Computational Molecular Science*. **12**(3), p.e1570.

Gray, A., Harlen, O.G., Harris, S.A., Khalid, S., Leung, Y.M., Lonsdale, R., Mulholland, A.J., Pearson, A.R., Read, D.J. and Richardson, R.A. 2015. In pursuit of an accurate spatial and temporal model of biomolecules at the atomistic level: A perspective on computer simulation. *Acta Crystallographica Section D: Biological Crystallography*. **71**(1), pp.162–172.

Hammer, J.A. and Sellers, J.R. 2012. Walking to work: Roles for class v myosins as cargo transporters. *Nature Reviews Molecular Cell Biology*. **13**(1), pp.13–26.

Hannemann, D.E., Cao, W., Olivares, A.O., Robblee, J.P. and De La Cruz, E.M. 2005. Magnesium, ADP, and actin binding linkage of myosin V: Evidence for multiple myosin V-ADP and actomyosin V-ADP states. *Biochemistry*. **44**(24), pp.8826–8840.

Hanson, B.S., Iida, S., Read, D.J., Harlen, O.G., Kurisu, G., Nakamura, H. and Harris, S.A. 2021. Continuum mechanical parameterisation of cytoplasmic dynein from atomistic simulation. *Methods*. **185**, pp.39–48.

Hanson, J. 1967. Axial period of actin filaments. *Nature*. **213**(5074), pp.353–356.

Hartman, M.A. and Spudich, J.A. 2012. The myosin superfamily at a glance. *Journal of Cell Science*. **125**(7), pp.1627–1632.

Hathcock, D., Tehver, R., Hinczewski, M. and Thirumalai, D. 2020. Myosin V executes steps of variable length via structurally constrained diffusion. *eLife*. **9**, p.e51569.

Heissler, S.M. and Sellers, J.R. 2016. Kinetic Adaptations of Myosins for Their Diverse Cellular Functions. *Traffic*. **17**(8), pp.839–859.

Heissler, S.M. and Sellers, J.R. 2015. Myosin light chains: Teaching old dogs new tricks. *Ceased*. **4**(6), pp.169–188.

Hess, H. 2011. Engineering Applications of Biomolecular Motors. *Annual Review of Biomedical Engineering*. **13**, pp.429–450.

Hinczewski, M., Tehver, R. and Thirumalai, D. 2013. Design principles governing the motility of myosin v. *Proceedings of the National Academy of Sciences of the United States of America*. **110**(43), pp.E4059–E4068.

Holmes, K.C., Angert, I., Kull, F.J., Jahn, W. and Schröder, R.R. 2003. Electron cryo-microscopy shows how strong binding of myosin to actin releases nucleotide. *Nature 2003 425:6956*. **425**(6956), pp.423–427.

Holmes, K.C., Trentham, D.R., Simmons, R., Takagi, Y., Shuman, H. and

Goldman, Y.E. 2004. Coupling between phosphate release and force generation in muscle actomyosin. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. **359**(1452), pp.1913–1920.

Hopkins, R.F. and Esposito, D. 2009. A rapid method for titrating baculovirus stocks using the Sf-9 Easy Titer cell line. *BioTechniques*. **47**(3), pp.785–788.

Houdusse, A., Gaucher, J.F., Krementsova, E., Mui, S., Trybus, K.M. and Cohen, C. 2006. Crystal structure of apo-calmodulin bound to the first two IQ motifs of myosin V reveals essential recognition features. *Proceedings of the National Academy of Sciences of the United States of America*. **103**(51), pp.19326–19331.

Houdusse, A. and Sweeney, H.L. 2016. How Myosin Generates Force on Actin Filaments. *Trends in Biochemical Sciences*. **41**(12), pp.989–997.

Houdusse, A., Szent-Györgyi, A.G. and Cohen, C. 2000. Three conformational states of scallop myosin S1. *Proceedings of the National Academy of Sciences of the United States of America*. **97**(21), pp.11238–11243.

Humphrey, W., Dalke, A. and Schulten, K. 1996. VMD: Visual molecular dynamics. *Journal of Molecular Graphics*. **14**(1), pp.33–38.

Isralewitz, B., Gao, M. and Schulten, K. 2001. Steered molecular dynamics and mechanical functions of proteins. *Current Opinion in Structural Biology*. **11**(2), pp.224–230.

Kharakoz, D.P. 2000. Protein Compressibility, Dynamics, and Pressure. *Biophysical Journal*. **79**(1), pp.511–525.

Kidmose, R.T., Juhl, J., Nissen, P., Boesen, T., Karlsen, J.L. and Pedersen, B.P. 2019. Namdinator - Automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCrJ*. **6**(4), pp.526–531.

Kodera, N., Yamamoto, D., Ishikawa, R. and Ando, T. 2010. Video imaging of walking myosin V by high-speed atomic force microscopy. *Nature*. **468**(7320), pp.72–76.

Koide, H., Kinoshita, T., Tanaka, Y., Tanaka, S., Nagura, N., Meyer Zu Hörste, G., Miyagi, A. and Ando, T. 2006. Identification of the single specific IQ motif of myosin V from which calmodulin dissociates in the presence of Ca2+. *Biochemistry*. **45**(38), pp.11598–11604.

Kollmar, M. and Mühlhausen, S. 2017. Myosin repertoire expansion coincides with eukaryotic diversification in the Mesoproterozoic era. *BMC Evolutionary Biology*. **17**(1), pp.1–18.

Kolomeisky, A.B. and Fisher, M.E. 2003. A simple kinetic model describes the processivity of myosin-V. *Biophysical Journal*. **84**(3), pp.1642–1650.

Kühlbrandt, W. 2014a. Cryo-EM enters a new era. *eLife*. **3**, p.e03678.

Kühlbrandt, W. 2014b. The resolution revolution. *Science*. **343**(6178), pp.1443–1444.

De La Cruz, E.M., Wells, A.L., Rosenfeld, S.S., Ostap, E.M. and Sweeney, H.L. 1999. The kinetic mechanism of myosin V. *Proceedings of the National Academy of Sciences of the United States of America*. **96**(24), pp.13726–13731.

Lan, L., Han, H., Zuo, H., Chen, Z., Du, Y., Zhao, W., Gu, J. and Zhang, Z. 2010. Upregulation of myosin Va by Snail is involved in cancer cell migration and metastasis. *International Journal of Cancer*. **126**(1), pp.53–64.

Liu, J., Taylor, D.W., Krementsova, E.B., Trybus, K.M. and Taylor, K.A. 2006a. Three-dimensional structure of the myosin V inhibited state by cryoelectron tomography. *Nature*. **442**(7099), pp.208–211.

Liu, J., Taylor, D.W., Krementsova, E.B., Trybus, K.M. and Taylor, K.A. 2006b. Three-dimensional structure of the myosin V inhibited state by cryoelectron tomography. *Nature*. **442**(7099), pp.208–211.

Llinas, P., Isabet, T., Song, L., Ropars, V., Zong, B., Benisty, H., Sirigu, S., Morris, C., Kikuti, C., Safer, D., Sweeney, H.L. and Houdusse, A. 2015. How Actin Initiates the Motor Activity of Myosin. *Developmental Cell*. **33**(4), pp.401–412.

Lombardo, A.T., Nelson, S.R., Ali, M.Y., Kennedy, G.G., Trybus, K.M., Walcott, S. and Warshaw, D.M. 2017. Myosin Va molecular motors manoeuvre liposome cargo through suspended actin filament intersections in vitro. *Nature Communications 2017 8:1*. **8**(1), pp.1–9.

Lombardo, A.T., Nelson, S.R., Kennedy, G.G., Trybus, K.M., Walcott, S. and Warshaw, D.M. 2019. Myosin Va transport of liposomes in three-dimensional actin networks is modulated by actin filament density, position, and polarity. *Proceedings of the National Academy of Sciences of the United States of America*. **116**(17), pp.8326–8335.

Lu, T. and Chen, F. 2012. Quantitative analysis of molecular surface based on improved Marching Tetrahedra algorithm. *Journal of Molecular Graphics and Modelling*. **38**, pp.314–323.

McCammon, J.A., Gelin, B.R. and Karplus, M. 1977. Dynamics of folded proteins. *Nature*. **267**(5612), pp.585–590.

McMillan, S.N. and Scarff, C.A. 2022. Cryo-electron microscopy analysis of myosin at work and at rest. *Current Opinion in Structural Biology*. **75**, p.102391.

Mehta, A.D., Finer, J.T. and Spudich, J.A. 1997. Detection of single-molecule interactions using correlated thermal diffusion. *Proceedings of the National Academy of Sciences*. **94**(15), pp.7927–7931.

Mehta, A.D., Rock, R.S., Rief, M., Spudich, J.A., Mooseker, M.S. and Cheney, R.E. 1999. Myosin-V is a processive actin-based motor. *Nature*. **400**(6744), pp.590–593.

Mentes, A., Huehn, A., Liu, X., Zwolak, A., Dominguez, R., Shuman, H., Ostap, E.M. and Sindelar, C. V. 2018. High-resolution cryo-EM structures of actin-bound myosin states reveal the mechanism of myosin force sensing.

Proceedings of the National Academy of Sciences of the United States of America. **115**(6), pp.1292–1297.

Mercer, J.A., Seperack, P.K., Strobel, M.C., Copeland, N.G. and Jenkins, N.A. 1991. Novel myosin heavy chain encoded by murine dilute coat colour locus. *Nature*. **349**(6311), pp.709–713.

Miao, Z. and Cao, Y. 2016. Quantifying side-chain conformational variations in protein structure. *Scientific Reports 2016 6:1*. **6**(1), pp.1–10.

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S. and Steinegger, M. 2022. ColabFold: making protein folding accessible to all. *Nature Methods 2022 19:6*. **19**(6), pp.679–682.

Moore, J.R., Krementsova, E.B., Trybus, K.M. and Warshaw, D.M. 2004. Does the myosin V neck region act as a lever? *Journal of Muscle Research and Cell Motility*. **25**(1), pp.29–35.

Mooseker, M.S. and Cheney, R.E. 1995. Unconventional myosins. *Annual review of cell and developmental biology*. **11**, pp.633–675.

Mukherjee, S., Alhadeff, R. and Warshel, A. 2017. Simulating the dynamics of the mechanochemical cycle of myosin-V. *Proceedings of the National Academy of Sciences of the United States of America*. **114**(9), pp.2259–2264.

Murata, K. and Wolf, M. 2018. Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochimica et Biophysica Acta - General Subjects*. **1862**(2), pp.324–334.

Nakane, T., Kimanius, D., Lindahl, E. and Scheres, S.H.W. 2018. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *eLife*. **7**, p.e36861.

Nakane, T., Kotecha, A., Sente, A., McMullan, G., Masiulis, S., Brown, P.M.G.E., Grigoras, I.T., Malinauskaite, L., Malinauskas, T., Miehling, J., Uchański, T., Yu, L., Karia, D., Pechnikova, E. V., de Jong, E., Keizer, J., Bischoff, M., McCormack, J., Tiemeijer, P., Hardwick, S.W., Chirgadze, D.Y., Murshudov, G., Aricescu, A.R. and Scheres, S.H.W. 2020. Single-particle cryo-EM at atomic resolution. *Nature*. **587**(7832), pp.152–156.

Nascimento, A.F.Z., Trindade, D.M., Tonoli, C.C.C., De Giuseppe, P.O., Assis, L.H.P., Honorato, R. V., Deoliveira, P.S.L., Mahajan, P., Burgess-Brown, N.A., Von Delft, F., Larson, R.E. and Murakami, M.T. 2013. Structural insights into functional overlapping and differentiation among myosin V motors. *Journal of Biological Chemistry*. **288**(47), pp.34131–34145.

Nelson, B.J., Kaliakatsos, I.K. and Abbott, J.J. 2010. Microrobots for Minimally Invasive Medicine. *Annual Review of Biomedical Engineering*. **12**, pp.55–85.

Nguyen, H.A. and Higuchi, H. 2005. Motility of myosin V regulated by the dissociation of single calmodulin. *Nature Structural & Molecular Biology 2005 12:2*. **12**(2), pp.127–132.

Niu, F., Liu, Y., Sun, K., Xu, S., Dong, J., Yu, C., Yan, K. and Wei, Z. 2022.

Autoinhibition and activation mechanisms revealed by the triangular-shaped structure of myosin Va. *Science Advances*. **8**(49).

Oguchi, Y., Mikhailenko, S. V., Ohki, T., Olivares, A.O., De La Cruz, E.M. and Ishiwata, S. 2008. Load-dependent ADP binding to myosins V and VI: Implications for subunit coordination and function. *Proceedings of the National Academy of Sciences of the United States of America*. **105**(22), pp.7714–7719.

Oke, O.A., Burgess, S.A., Forgacs, E., Knight, P.J., Sakamoto, T., Sellers, J.R., White, H. and Trinick, J. 2010. Influence of lever structure on myosin 5a walking. *Proceedings of the National Academy of Sciences of the United States of America*. **107**(6), pp.2509–2514.

Oliver, R. 2013. *A stochastic finite element model for the dynamics of globular proteins*. PhD Thesis, University of Leeds.

Oliver, R.C., Read, D.J., Harlen, O.G. and Harris, S.A. 2013. A stochastic finite element model for the dynamics of globular macromolecules. *Journal of Computational Physics*. **239**, pp.147–165.

Ovchinnikov, V., Trout, B.L. and Karplus, M. 2010. Mechanical Coupling in Myosin V: A Simulation Study. *Journal of Molecular Biology*. **395**(4), pp.815–833.

Parker, F. and Peckham, M. 2020. Disease mutations in striated muscle myosins. *Biophysical Reviews*. **12**(4), pp.887–894.

Pastural, E., Barrat, F.J., Dufourcq-Lagelouse, R., Certain, S., Sanal, O., Jabado, N., Seger, R., Griscelli, C., Fischer, A. and De Saint Basile, G. 1997. Griscelli disease maps to chromosome 15q21 and is associated with mutations in the myosin-Va gene. *Nature Genetics*. **16**(3), pp.289–292.

Peckham, M. 2016. How myosin organization of the actin cytoskeleton contributes to the cancer phenotype. *Biochemical Society Transactions*. **44**(4), pp.1026–1034.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. 2004. UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*. **25**(13), pp.1605–1612.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H. and Ferrin, T.E. 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*. **30**(1), pp.70–82.

Popov, K., Komianos, J. and Papoian, G.A. 2016. MEDYAN: Mechanochemical Simulations of Contraction and Polarity Alignment in Actomyosin Networks. *PLoS Computational Biology*. **12**(4), p.1004877.

Pospich, S., Sweeney, H.L., Houdusse, A. and Raunser, S. 2021. High-resolution structures of the actomyosin-v complex in three nucleotide states provide insights into the force generation mechanism. *eLife*. **10**, p.e73724.

Provance, D.W. and Mercer, J.A. 1999. Myosin-V: Head to tail. *Cellular and*

Molecular Life Sciences. **56**(3–4), pp.233–242.

Punjani, A. and Fleet, D.J. 2021. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *Journal of Structural Biology*. **213**(2), p.107702.

Purcell, E.M. 1977. Life at low Reynolds number. *American Journal of Physics*. **45**(1), pp.3–11.

Pylypenko, O., Attanda, W., Gauquelin, C., Lahmani, M., Coulibaly, D., Baron, B., Hoos, S., Titus, M.A., England, P. and Houdusse, A.M. 2013. Structural basis of myosin v Rab GTPase-dependent cargo recognition. *Proceedings of the National Academy of Sciences of the United States of America*. **110**(51), pp.20443–20448.

Rajon, D.A. and Bolch, W.E. 2003. Marching cube algorithm: review and trilinear interpolation adaptation for image-based dosimetric models. *Computerized Medical Imaging and Graphics*. **27**(5), pp.411–435.

Ravikumar, A., Gopnarayan, M.N., Subramaniam, S. and Srinivasan, N. 2022. Comparison of side-chain dispersion in protein structures determined by cryo-EM and X-ray crystallography. *IUCrJ*. **9**(1), pp.98–103.

Reck-Peterson, S.L., Provance, D.W., Mooseker, M.S. and Mercer, J.A. 2000. Class V myosins. *Biochimica et Biophysica Acta - Molecular Cell Research*. **1496**(1), pp.36–51.

Revill, D.J. 2014. *Flexibility in the heads of myosin 5*. PhD Thesis, University of Leeds.

Richardson, R.A., Hanson, B.S., Read, D.J., Harlen, O.G. and Harris, S.A. 2020. Exploring the dynamics of flagellar dynein within the axoneme with Fluctuating Finite Element Analysis. *Quarterly Reviews of Biophysics*. **53**, p.E9.

Richardson, R.A., Papachristos, K., Read, D.J., Harlen, O.G., Harrison, M., Paci, E., Muench, S.P. and Harris, S.A. 2014. Understanding the apparent stator-rotor connections in the rotary ATPase family using coarse-grained computer modeling. *Proteins: Structure, Function, and Bioinformatics*. **82**(12), pp.3298–3311.

Rief, M., Rock, R.S., Mehta, A.D., Mooseker, M.S., Cheney, R.E. and Spudich, J.A. 2000. Myosin-V stepping kinetics: A molecular model for processivity. *Proceedings of the National Academy of Sciences of the United States of America*. **97**(17), pp.9482–9486.

Robert-Paganin, J., Pylypenko, O., Kikuti, C., Sweeney, H.L. and Houdusse, A. 2020. Force Generation by Myosin Motors: A Structural Perspective. *Chemical Reviews*. **120**(1), pp.5–35.

Roe, D.R. and Cheatham, T.E. 2013. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation*. **9**(7), pp.3084–3095.

Rosenfeld, S.S. and Sweeney, H.L. 2004. A model of myosin V processivity. *Journal of Biological Chemistry*. **279**(38), pp.40100–40111.

Sahl, S.J., Hell, S.W. and Jakobs, S. 2017. Fluorescence nanoscopy in cell biology. *Nature Reviews Molecular Cell Biology*. **18**(11), pp.685–701.

Sakamoto, T., Amitani, I., Yokota, E. and Ando, T. 2000. Direct Observation of Processive Movement by Individual Myosin V Molecules. *Biochemical and Biophysical Research Communications*. **272**(2), pp.586–590.

Sakamoto, T., Wang, F., Schmitz, S., Xu, Y., Xu, Q., Molloy, J.E., Veigel, C. and Sellers, J.R. 2003. Neck length and processivity of myosin V. *Journal of Biological Chemistry*. **278**(31), pp.29201–29207.

Sakamoto, T., Yildez, A., Selvin, P.R. and Sellers, J.R. 2005. Step-size is determined by neck length in myosin V. *Biochemistry*. **44**(49), pp.16203–16210.

Sanchez-Garcia, R., Gomez-Blanco, J., Cuervo, A., Carazo, J.M., Sorzano, C.O.S. and Vargas, J. 2021. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Communications Biology 2021 4:1*. **4**(1), pp.1–8.

Scarff, C.A., Carrington, G., Casas-Mao, D., Chalovich, J.M., Knight, P.J., Ranson, N.A. and Peckham, M. 2020. Structure of the shutdown state of myosin-2. *Nature*. **588**(7838), pp.515–520.

Scarff, C.A., Fuller, M.J.G., Thompson, R.F. and Iadaza, M.G. 2018. Variations on negative stain electron microscopy methods: Tools for tackling challenging systems. *Journal of Visualized Experiments*. **2018**(132).

Scheres, S.H.W. 2012. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*. **180**(3), pp.519–530.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P. and Cardona, A. 2012. Fiji: an open-source platform for biological-image analysis. *Nature Methods 2012 9:7*. **9**(7), pp.676–682.

Schrödinger, L. 2015. The PyMOL Molecular Graphics System, Version 1.8.

Schroeder, W., Martin, K. and Lorensen, B. 2006. *The Visualization Toolkit: an object-oriented approach to 3D graphics*. Kitware.

Shen, M., Zhang, N., Zheng, S., Zhang, W.B., Zhang, H.M., Lu, Z., Su, Q.P., Sun, Y., Ye, K. and Li, X.D. 2016. Calmodulin in complex with the first IQ motif of myosin-5a functions as an intact calcium sensor. *Proceedings of the National Academy of Sciences of the United States of America*. **113**(40), pp.E5812–E5820.

Shiroguchi, K., Chin, H.F., Hannemann, D.E., Muneyuki, E., De La Cruz, E.M. and Kinosita, K. 2011. Direct Observation of the Myosin Va Recovery Stroke That Contributes to Unidirectional Stepping along Actin J. Spudich, ed. *PLoS Biology*. **9**(4), p.e1001031.

Shkurti, A., Goni, R., Andrio, P., Breitmoser, E., Bethune, I., Orozco, M. and Laughton, C.A. 2016. pyPcazip: A PCA-based toolkit for compression and analysis of molecular simulation data. *SoftwareX*. **5**, pp.44–50.

Si, H. 2015. TetGen, a Delaunay-Based Quality Tetrahedral Mesh Generator | ACM Transactions on Mathematical Software. *ACM Transactions on Mathematical Software*. **41**(2), p.11.

Sivaramakrishnan, S., Spink, B.J., Sim, A.Y.L., Doniach, S. and Spudich, J.A. 2008. Dynamic charge interactions create surprising rigidity in the ER/K α-helical protein motif. *Proceedings of the National Academy of Sciences of the United States of America*. **105**(36), pp.13356–13361.

Solernou, A., Hanson, B.S., Richardson, R.A., Welch, R., Read, D.J., Harlen, O.G. and Harris, S.A. 2018. Fluctuating Finite Element Analysis (FFEA): A continuum mechanics software tool for mesoscale simulation of biomolecules T. Poisot, ed. *PLOS Computational Biology*. **14**(3), p.e1005897.

Strom, M., Hume, A.N., Tarafder, A.K., Barkagianni, E. and Seabra, M.C. 2002. A Family of Rab27-binding Proteins. *Journal of Biological Chemistry*. **277**(28), pp.25423–25430.

Sukumar, N. and Tupek, M.R. 2022. Virtual elements on agglomerated finite elements to increase the critical time step in elastodynamic simulations. *International Journal for Numerical Methods in Engineering*. **123**(19), pp.4702–4725.

Sweeney, H. and Holzbaur, E.L.F. 2018. Motor proteins. *Cold Spring Harbor Perspectives in Biology*. **10**(5), p.a021931.

Takagi, Y., Farrow, R.E., Billington, N., Nagy, A., Batters, C., Yang, Y., Sellers, J.R. and Molloy, J.E. 2014. Myosin-10 produces its power-stroke in two phases and moves processively along a single actin filament under low load. *Proceedings of the National Academy of Sciences of the United States of America*. **111**(18), pp.E1833–E1842.

Taylor, K.A. and Glaeser, R.M. 1974. Electron diffraction of frozen, hydrated protein crystals. *Science*. **186**(4168), pp.1036–1037.

Taylor, R.W. and Sandoghdar, V. 2019. Interferometric Scattering Microscopy: Seeing Single Nanoparticles and Molecules via Rayleigh Scattering. *Nano Letters*. **19**(8), pp.4827–4835.

Terrak, M., Rebowski, G., Lu, R.C., Grabarek, Z. and Dominguez, R. 2005. Structure of the light chain-binding domain of myosin V. . **6**(36).

Terrak, M., Wu, G., Stafford, W.F., Lu, R.C. and Dominguez, R. 2003. Two distinct myosin light chain structures are induced by specific variations within the bound IQ motifs—functional implications. *The EMBO Journal*. **22**(3), pp.362–371.

Thirumurugan, K., Sakamoto, T., Hammer, J.A., Sellers, J.R. and Knight, P.J. 2006. The cargo-binding domain regulates structure and activity of myosin 5. *Nature*. **442**(7099), pp.212–215.

Tian, C., Kasavajhala, K., Belfon, K.A.A., Raguette, L., Huang, H., Migues, A.N., Bickel, J., Wang, Y., Pincay, J., Wu, Q. and Simmerling, C. 2020. Ff19SB:

Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *Journal of Chemical Theory and Computation*. **16**(1), pp.528–552.

Trabuco, L.G., Villa, E., Schreiner, E., Harrison, C.B. and Schulten, K. 2009. Molecular dynamics flexible fitting: A practical guide to combine cryo-electron microscopy and X-ray crystallography. *Methods*. **49**(2), pp.174–180.

Trivedi, D. V., Nag, S., Spudich, A., Ruppel, K.M. and Spudich, J.A. 2020. The Myosin Family of Mechanoenzymes: From Mechanisms to Therapeutic Approaches. *Annual Review of Biochemistry*. **89**(1), pp.667–693.

Trybus, K.M., Gushchin, M.I., Lui, H.J., Hazelwood, L., Krementsova, E.B., Volkmann, N. and Hanein, D. 2007. Effect of calcium on calmodulin bound to the IQ motifs of myosin V. *Journal of Biological Chemistry*. **282**(32), pp.23316–23325.

Tyska, M.J. and Warshaw, D.M. 2002. The myosin power stroke. *Cell Motility and the Cytoskeleton*. **51**(1), pp.1–15.

Uemura, S., Higuchi, H., Olivares, A.O., De La Cruz, E.M. and Ishiwata, S. 2004. Mechanochemical coupling of two substeps in a single myosin V motor. *Nature Structural and Molecular Biology*. **11**(9), pp.877–883.

Uyeda, T.Q.P., Abramson, P.D. and Spudich, J.A. 1996. The neck region of the myosin motor domain acts as a lever arm to generate movement. *Proceedings of the National Academy of Sciences*. **93**(9), pp.4459–4464.

Veigel, C., Bartoo, M.L., White, D.C.S., Sparrow, J.C. and Molloy, J.E. 1998. The stiffness of rabbit skeletal actomyosin cross-bridges determined with an optical tweezers transducer. *Biophysical Journal*. **75**(3), pp.1424–1438.

Veigel, C., Schmitz, S., Wang, F. and Sellers, J.R. 2005. Load-dependent kinetics of myosin-V can explain its high processivity. *Nature Cell Biology*. **7**(9), pp.861–869.

Veigel, C., Wang, F., Bartoo, M.L., Sellers, J.R. and Molloy, J.E. 2002. The gated gait of the processive molecular motor, myosin V. *Nature Cell Biology*. **4**(1), pp.59–65.

Velvarska, H. and Niessing, D. 2013. Structural Insights into the Globular Tails of the Human Type V Myosins Myo5a, Myo5b, and Myo5c. *PLOS ONE*. **8**(12), p.e82065.

Vilfan, A. 2005. Elastic lever-arm model for myosin V. *Biophysical Journal*. **88**(6), pp.3792–3805.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., Vijaykumar, A., Bardelli, A. Pietro, Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C.N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D.A., Hagen, D.R., Pasechnik, D. V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G.A., Ingold, G.L., Allen, G.E., Lee, G.R., Audren, H., Probst, I., Dietrich, J.P., Silterra, J., Webber, J.T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J.L., de Miranda Cardoso, J.V., Reimer, J., Harrington, J., Rodríguez, J.L.C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N.J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P.A., Lee, P., McGibbon, R.T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., Oshima, T., Pingel, T.J., Robitaille, T.P., Spura, T., Jones, T.R., Cera, T., Leslie, T., Zito, T., Krauss, T., Upadhyay, U., Halchenko, Y.O. and Vázquez-Baeza, Y. 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*. **17**(3), pp.261–272.

Wada, F., Nakata, A., Tatsu, Y., Ooashi, N., Fukuda, T., Nabetani, T. and Kamiguchi, H. 2016. Myosin Va and Endoplasmic Reticulum Calcium Channel Complex Regulates Membrane Export during Axon Guidance. *Cell Reports*. **15**(6), pp.1329–1344.

Wagner, W., Brenowitz, S.D. and Hammer, J.A. 2011. Myosin-Va transports the endoplasmic reticulum into the dendritic spines of Purkinje neurons. *Nature Cell Biology*. **13**(1), pp.40–47.

Walker, M.L., Burgess, S.A., Sellers, J.R., Wang, F., Hammer, J.A., Trinick, J. and Knight, P.J. 2000. Two-headed binding of a processive myosin to F-actin. *Nature*. **405**(6788), pp.804–807.

Wang, F., Chen, L., Arcucci, O., Harvey, E. V., Bowers, B., Xu, Y., Hammer, J.A. and Sellers, J.R. 2000. Effect of ADP and ionic strength on the kinetic and motile properties of recombinant mouse myosin V. *Journal of Biological Chemistry*. **275**(6), pp.4329–4335.

Wang, F., Thirumurugan, K., Stafford, W.F., Hammer, J.A., Knight, P.J. and Sellers, J.R. 2004. Regulated Conformation of Myosin V. *Journal of Biological Chemistry*. **279**(4), pp.2333–2336.

Watanabe, M., Nomura, K., Ohyama, A., Ishikawa, R., Komiya, Y., Hosaka, K., Yamauchi, E., Taniguchi, H., Sasakawa, N., Kumakura, K., Ushiki, T., Sato, O., Ikebe, M. and Igarashi, M. 2005. Myosin-Va regulates exocytosis through the submicromolar Ca 2+-dependent binding of syntaxin-1A. *Molecular Biology of the Cell*. **16**(10), pp.4519–4530.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R. and Schwede, T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research*. **46**(W1), pp.W296–W303.

Wei, Z., Liu, X., Yu, C. and Zhang, M. 2013. Structural basis of cargo recognitions for class V myosins. *Proceedings of the National Academy of Sciences of the United States of America*. **110**(28), pp.11314–11319.

Welch, R., Harris, S.A., Harlen, O.G. and Read, D.J. 2020. KOBRA: A fluctuating elastic rod model for slender biological macromolecules. *Soft Matter*. **16**(32), pp.7544–7555.

Wulf, S.F., Ropars, V., Fujita-Becker, S., Oster, M., Hofhaus, G., Trabuco, L.G., Pylypenko, O., Sweeney, H.L., Houdusse, A.M., Schröder, R.R. and Vale, R.D. 2016. Force-producing ADP state of myosin bound to actin. *Proceedings of the National Academy of Sciences of the United States of America*. **113**(13), pp.E1844–E1852.

Yengo, C.M. and Sweeney, H.L. 2004. Functional Role of Loop 2 in Myosin V. *Biochemistry*. **43**(9), pp.2605–2612.

Yotti, R., Seidman, C.E. and Seidman, J.G. 2019. Advances in the Genetic Basis and Pathogenesis of Sarcomere Cardiomyopathies. *Annual Review of Genomics and Human Genetics*. **20**(1), pp.129–153.

Zhang, K. 2016. Gctf: Real-time CTF determination and correction. *Journal of Structural Biology*. **193**(1), pp.1–12.

Zhang, L., Song, J., Newhouse, Y., Zhang, S., Weisgraber, K.H. and Ren, G. 2010. An optimized negative-staining protocol of electron microscopy for apoE4•POPC lipoprotein. *Journal of Lipid Research*. **51**(5), pp.1228–1236.

Zhang, N., Yao, L.L. and Li, X. dong 2018. Regulation of class V myosin. *Cellular and Molecular Life Sciences*. **75**(2), pp.261–273.

Zhang, X.C. and Feng, W. 2016. Thermodynamic aspects of ATP hydrolysis of actomyosin complex. *Biophysics Reports 2016 2:5*. **2**(5), pp.87–94.

Zhang, Y., Krieger, J., Mikulska-Ruminska, K., Kaynak, B., Sorzano, C.O.S., Carazo, J.M., Xing, J. and Bahar, I. 2020. State-dependent sequential allostery exhibited by chaperonin TRiC/CCT revealed by network analysis of Cryo-EM maps. *Progress in Biophysics and Molecular Biology*. **160**, pp.104–120.

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y. and Agard, D.A. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods*. **14**(4), pp.331–332.

Zheng, W. 2011. Coarse-grained modeling of conformational transitions underlying the processive stepping of myosin V dimer along filamentous actin. *Proteins: Structure, Function, and Bioinformatics*. **79**(7), pp.2291–2305.

Zhong, E.D., Bepler, T., Berger, B. and Davis, J.H. 2021. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nature Methods*. **18**(2), pp.176–185.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E. and Scheres, S.H. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife*. **7**, p.e42166.