Flexibility in the Heads of Myosin 5

Derek John Revill

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

The University of Leeds
Faculty of Biological Sciences
School of Molecular and Cellular Biology

February 2013

The candidate confirms that the work submitted is his/her own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Acknowledgements

Firstly, I must thank my principal supervisor Professor Peter Knight for all his effort in getting me to the end of this PhD. Without his scientific input, encyclopaedic knowledge and many hours of discussions I would not have made it. Thanks also to my other supervisors, to Professor Peter Olmsted for helping me with physics and mathematics, and to Dr Stan Burgess for introducing me to image processing.

For part of this project I was helped by members of Dr James Sellers' Laboratory of Molecular Physiology at the National Institutes of Health in Bethesda, USA. I thank Jim and his team for welcoming me to their lab, letting me use their resources and for making my visits enjoyable. In particular I must acknowledge and thank Attila Nagy for teaching me cloning and the dark arts of myosin 5 expression and purification. I thank him also for the many hours spent continuing the cloning I began, and for expressing and purifying protein samples for me. I'm sorry there were so many. I thank Neil Billington for the EM work at NIH and Harry Takagi for help in all areas of the lab (and for always buying the beer and keeping us laughing).

Thanks to Dr Matt Walker for EM training, and for his superb datasets that form the basis of the first part of this project. Thanks to Dr Kavitha Thirumurugan for having picked over 20,000 molecules before I even started! I thank members of the lab at Leeds both past and present, especially Melanie Colegrave who guided me through the calmodulin work. Thanks to Stephen Meunch, Anthony Roberts and, very sadly, the now late Chris Kennaway. Thanks to Mara Rusu for being a fun desk neighbour.

Thanks to Professor Alan Berry and The Astbury Centre for allowing me to undertake the research, and to Hazel Jones for administering everything and being supportive. A special thank you to Dr Jo Rushworth for the friendship, support, proofreading and many laughs.

I am grateful to the Wellcome Trust for financial support.

Lastly, and most importantly, thanks to Lucie and to George. I know I have not been the best husband and father whilst I have struggled to complete this work, and I know you have sacrificed a lot to keep me afloat. I am very grateful for your love and look forward to spending enjoyable time with you again. I hope, in time, to be able to look back and be proud of this work.

Abstract

Class 5 myosins are motor proteins involved in the transport and localisation of diverse cargoes in eukaryotic cells. Molecules have two heads and a tail and can move along actin filaments, by making step-like movements driven by MgATP hydrolysis. Each head comprises an actin-binding motor and a lever, an extended α-helix carrying six light chains bound to IQ motifs. The lever is an essential mechanical element in myosin 5 function, but an understanding of its mechanical properties and how these derive from its substructure is lacking. To address this, the first aim of this study was to look for and characterise flexibility within the heads of myosin 5. A second aim was to test the effect of altering the IQ motif spacing on the properties and behaviour of the molecule.

The structure and flexibility of the head region of mammalian myosin 5a was studied using single-particle image processing of images from negative stain electron microscopy (EM). Image averaging revealed features of the motor and lever and provided evidence of independent rotational freedom of the motor about the lever axis, and thermally-driven flexing at the motor-lever junction. A stiffness of 32-51 pN·nm·rad^{-2} for the flexing was inferred, which represents a significant compliance in the unattached head. Variation in light chain orientations at each IQ motif suggested torsional flexibility. Lever bending analysis, by decomposition of average lever shapes into Fourier modes, yielded a persistence length of 50.9 ± 18.6 nm. All these results indicate that, under EM conditions, the unattached head is more flexible than has previously been measured for the actin-bound complex in optical trap experiments.

Lever mutants with altered IQ motif spacings were constructed and initial characterisations by negative stain EM, MgATPase assay and single-molecule fluorescence microscopy were made. For the latter, two calmodulin point mutants were made, for site-specific labelling and attachment to the lever as a fluorescent probe. One lever mutant, with closer-spaced IQ motifs, resembled wild-type and had similar actin-activated MgATPase activity. However, its motile properties remain undetermined. Another mutant, with wider-spaced IQ motifs did not bind calmodulin and it had no actin-activated MgATPase. The molecule appeared misfolded but retained residual actin-binding properties. These contrasting results highlight the importance of the lever's structure in the function of the molecule.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Units and constants

Å	10^{-10} metres
bp	Base pair(s)
°C	Degrees Celsius
Da	Dalton(s)
e^-	Electron(s)
g	Acceleration due to gravity = 9.8 ms^{-2}
hr(s)	Hour(s)
J	Joule(s)
k_b	Boltzmann constant = 1.38×10^{-23} JK^{-1}
kb(p)	10^3 base (pairs)
kDa	10^3 Daltons
kV	10^3 volts
K	Kelvin
mbar	10^{-3} bar
min(s)	Minute(s)
mL	10^{-3} litre(s)
mm	10^{-3} metre(s)
mM	10^{-3} molar
μL	10^{-6} litre(s)
μm	10^{-6} metre(s)
μM	10^{-6} molar
nm	10^{-9} metre(s)
nM	10^{-9} molar
pH	Potential of hydrogen
pN	10^{-12} newtons
rad	radian(s)
s	Second(s)

Other abbreviations

2D	Two dimensional
3D	Three dimensional
ADP	Adenosine-5'-diphosphate
AFM	Atomic Force Microscopy
AN	Dr Attila Nagy (LMP-NIH)
ATP	Adenosine-5'-triphosphate
ATPase	Adenosin 5'-triphosphatase
BFA	Biotinylated rhodamine phalloidin F-actin
C-	Carboxyl-
cDNA	Complementary deoxyribonucleic acid
Cryo-EM	Cryo-electron tomography
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
ELC	Essential light chain
EM	Electron Microscopy
F-actin	Fibrous/filamentous actin
FLAG	FLAG epitope sequence = DYKDDDDK
FoV	Field of View
G-actin	Globular actin
GFP	Green Fluorescent Protein
GTD	Globular tail domain
HC	Heavy chain
HMM	Heavy Meromyosin
IQ	IQ motif
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LaB$_6$	Lanthanum hexaboride
LB	Lysogeny Broth
LC	Light chain
LMM	Light Meromyosin
LMP	Laboratory of molecular physiology
MgATP	Magnesium adenosine-5'-triphosphate
MOPS	3-(N-morpholino)propanesulfonic acid
MW	Molecular weight

MWCO	Molecular weight cut off
N-	Amino-
NB	Dr Neil Billington (LMP-NIH)
NIH	National Institutes of Health
OD	Optical Density (absorbance)
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
P_i	Inorganic phosphate
PMSF	Phenylmethylsulfonyl Fluoride
RhPh	Rhodamine phalloidin
RLC	Regulatory light chain
S1	Myosin subfragment 1
S2	Myosin subfragment 2
SAH	Stable Single alpha helix
S.D.	Standard deviation
SDS	Sodium Dodecyl Sulphate
PAGE	Polyacrylamide Gel Electrophoresis
SFM	Serum Free Medium
SH3	Src-homology region type 3
S.O.C.	Super Optimal broth with Catabolite repression
SPIDER	System for Processing Image Data in Electron Microscopy and Related fields
TAE	Tris-Acetate-EDTA
TBS	Tris-buffered saline
TEM	Transmission electron microscopy
TEMED	N,N,N',N'-tetramethylethylenediamine
TIFF	Tagged Information File Format
TIRF	Total Internal Reflection Fluorescence (microscopy)
Tris	Tris(hydroxymethyl)aminomethane
TRITC	Tetramethylrhodamine-5-(and 6)-isothiocyanate
UV	Ultraviolet
v/v	Volume / volume
w/v	Weight / volume

Chapter 1: Introduction

1.1 Molecular machines, motor proteins and movement

Life depends upon a myriad of dynamic microscopic processes that constantly occur within the cell (Alberts et al., 2002). The central actors in these processes can be viewed as molecular machines; assemblies of proteins, nucleic acids and other molecular components able to produce mechanical actions in response to specific external stimuli (Alberts, 1998; Ballardini et al., 2001). The cell has evolved many types of molecular machine to carry out its essential functions. These include cytoskeletal motor proteins (myosins, kinesins and dyneins) to produce directed movement along actin filaments and microtubules; rotary motors, pumps and translocation pores (*e.g.* the bacterial flagella motor, F- and V-ATPases) to generate rotational movement and move atoms across membrane partitions; construction machines (*e.g.* the ribosome and polymerases) to build proteins and nucleic acid polymers; degradation machines (*e.g.* proteosomes and Clp proteases) to breakup proteins, as well as machines to separate, uncoil or pack nucleic acid strands (helicases, topoisomerases and condensins).

Many of the cell's molecular machines are also called molecular motors or, if formed from proteins or protein complexes, motor proteins. A molecular motor or motor protein can be most generally defined as a mechanochemical enzyme; a molecule that is able to convert different forms of chemical energy into mechanical work, through conformational changes and structural displacements (Bustamante et al., 2004). However, the name motor protein is often used more specifically to refer to proteins or protein complexes that convert the chemical energy from ATP hydrolysis into force used to drive various types of cellular motility (Howard, 2001; Schliwa and Woehlke, 2003).

The ability to generate and control movement is a defining characteristic of life at all structural levels, from the co-ordinated motions of many proteins that produce macroscopic scale movements such as muscle contraction or sperm propulsion, to the intracellular transport of cargoes along biopolymer tracks (Vale and Milligan, 2000). Motor proteins are the force-generating molecules that participate in all these processes. Therefore, understanding their structure, their mechanical properties, and how these molecular machines have evolved to function are important research goals for biology and biophysics.

1.2   Subject of this thesis: structure and flexibility in myosin 5a

In this thesis, the relationship between structure and function of a particular motor protein, myosin 5a, is investigated. Class 5 myosins (collectively termed myosin 5) form a class of proteins from the myosin superfamily, a large group of force-producing motor proteins associated with actin filaments in eukaryotic cells (Coluccio et al., 2008; Sellers, 1999). Certain class 5 myosins, such as vertebrate myosin 5a (the subject of this thesis) have the ability to move as individual molecules along actin filaments, making step-like movements driven by MgATP hydrolysis (Hammer and Sellers, 2012). They achieve this feat by alternately advancing one of two heads in a molecule from a trailing to a leading position, while the other head remains attached to the actin 'track'. Each head comprises a catalytic motor domain that binds the actin filament and hydrolyses MgATP, and an extended α-helical domain, referred to throughout this thesis as the lever, which binds six accessory light chains at sequences called IQ motifs (see 1.6.1). The myosin 5a lever has a unique structure and is unusually long among myosins, comprising six successive IQ motifs which occur at a conserved amino acid spacing pattern.

The basic motor-lever structure is common to all myosins, though there is diversity in the precise substructure and functional properties across the family. The lever is referred to as such because it acts as a distance multiplying mechanical element, transforming ångström-sized nucleotide-dependent movements in the motor domain into a nanometre-sized movement of its distal end. In myosin 5a, the lever plays a particularly important role in the protein's function. Its long length allows the two-headed molecule to span the psuedo-helical repeat length of the actin filament and so enables the molecule to walk in a straight path as it moves along the filament. In addition, there is evidence that the lever can bend significantly under load and occupy functionally-important strained states thought to be important in regulating the alternating movement of the heads.

Despite its functional significance, a understanding of the mechanical properties of the myosin 5a lever is lacking. To address this, the work presented in this thesis seeks to explore the relationship between the structure and mechanical properties of this domain.

Following the section below listing the more precise aims and overall structure of the thesis, the remainder of this chapter offers a broader introduction to cytoskeletal motor proteins, myosins, class 5 myosins and myosin 5a. Many of the terms already introduced above are defined in more detail. Brief introductions to the techniques of negative stain EM and single-particle image processing used in this project are also included.

1.3   Thesis aims and structure

The principal aim of this project was to investigate the relationship between structure and function of the lever of mammalian myosin 5a. The work undertaken sought to address the following research questions which, to date, have not been investigated elsewhere:

1. How does the unique structure of the myosin 5a lever contribute to its ability to function as a processive motor protein?

2. What degrees of mechanical freedom does the myosin 5a lever possess? In particular, what bending and twisting movements can it make?

3. Is bending and twisting in the lever isotropic, *i.e.* similar in all directions, and uniform along the lever's length?

4. Is the unique, alternating spacing of light chain-binding sites in the lever necessary for myosin 5a to function and make processive movements along actin filaments?

5. How does altering the spacing of light chain-binding sites in the lever effect its structure and mechanical properties, as well as the overall function of the protein?

The work described in this thesis divides into two parts.

Part I: Structure and flexibility in the head region of myosin 5a

In chapters 3 and 4, the structure and flexibility of the myosin 5a head (its motor and lever domains) in free myosin 5a molecules is examined in detail, using single-particle image processing of images obtained by negative stain EM. In chapter 3, existing atomic models of the head are evaluated and adapted according to the structure observed in new negative stain images. Flexibility seen at the motor-lever junction is measured, and unexpected variation in the orientations of calmodulin (CaM) light chains along the length of the lever is described in relation to torsional flexibility. In chapter 4, the bending flexibility of the lever is examined by single-particle image analysis of the same negative stain EM data. This chapter seeks to answer the question of whether the lever exhibits isotropic bending flexibility, and to quantitate the flexibility seen.

Part II: Investigation of the significance of the spacing of light chain binding sites in the myosin 5a lever

This section of the thesis covers studies investigating the influence of spacing of the light chain binding sequences on myosin 5a structure and function. Chapter 5 describes the design, expression and purification of a series of heavy meromyosin (HMM)-like myosin

5a mutants, with altered spacings between light chain binding sites. Results of initial structural, biochemical and mechanical characterisations are described. Chapter 6 describes a related but standalone project to create a functionalised CaM mutant, which could be site-specifically labelled with a fluorescent dye and exchanged onto myosin 5a molecules to act as a probe for the position of the lever in microscopy experiments. Cloning, expression, the results of exchange of the labelled CaMs into myosin 5a and observation by total internal reflection (TIRF) microscopy are reported. The creation of the CaM mutant was designed with general uses in mind; for exchange onto other classes of myosin for tracking and microscopy studies, and to allow attachment of other types of functional label such as gold nanoparticles for identification in EM or other microscopy studies.

Following this introductory chapter, a materials and methods chapter is presented, before the four results-based chapters and a final conclusion. The materials and methods chapter contains only generic materials and laboratory methods that are common to the work in more than one of the results chapters. Where methods are specific to work in only one chapter, these are presented at the outset of the particular chapter, for convenience of reference.

1.4   Cytoskeletal motor proteins

The cytoskeletal motor proteins are mechanochemical-enzymes capable of converting the chemical energy from ATP hydrolysis into force for directed movement along cytoskeletal filaments. There are three distinct families: myosins, kinesins and dyneins. The cytoskeletal 'tracks' these motors bind are actin filaments and microtubules. Both tracks have a polar structure, with faster- and slower-polymerising ends known as plus and minus ends respectively. Myosins bind to and make movements on actin filaments, generally directed to the plus end, with the exception of myosin 6 which moves towards the minus end (Wells et al., 1999). Kinesins and dyneins bind and move on microtubules. Most kinesins move towards the microtubule plus end (with exceptions such as the ncd kinesins), while dyneins move to the minus end.

Despite moving on different cytoskeletal filaments, comparison between myosin and kinesin structures reveals certain similarities in domain arrangement and sequence of mechanical movements, as well as a very similar structure around the ATP-binding site at the core of their motor domains (Howard, 2001; Vale and Milligan, 2000). The structure around the active site is also found in G-proteins and suggests that these enzymes share a common evolutionary origin (Vale and Milligan, 2000). Dynein is set somewhat apart

from myosin and kinesin, owing to its much greater mass and the particular structure of its motor domain. Dynein's motor domain contains four ATP binding sites (as opposed to myosin and kinesin's one) positioned around a hexameric ring formed from six AAA+ modules. This structure clearly places dynein in the AAA+ family of ring-shaped ATPases (Neuwald et al., 1999).

Myosins have a long research history and are perhaps best known for their role in muscle contraction (Geeves and Holmes, 2005). The many other classes of myosin since discovered are involved in numerous other cellular functions, from short-range cargo transport (typically close to the plasma membrane) to cytokinesis (Kieke and Titus, 2003; Krendel and Mooseker, 2005). The myosin family is discussed in more detail in section 1.5.

The kinesins form another superfamily of motor proteins. They are involved in intracellular transport and a wide range of cellular processes. These include organelle and vesicle transport in neurons, regulation of microtubules (length, number and dynamics), and formation of the spindle and chromosomes during mitosis (Hirokawa et al., 2009; Hirokawa and Takemura, 2007). Cargoes include the Golgi apparatus, endoplasmic reticulum, endosomes, lysosomes, mitochondria, tubulin dimers, various types of vesicle and other protein complexes (Hirokawa et al., 2009). Kinesin superfamily proteins (KIFs) are encoded by 45 genes in mammals. The many isoforms (produced by alternative mRNA splicing) are divided into fifteen separate families (kinesins 1-14B), according to phylogenetic sequence analysis. Generally, KIFs have a conserved ~350 residue motor domain which contains ATP- and microtubule-binding sites, a ~40-residue neck linker domain and a coiled coil sequence that dimerises KIFs into two-headed molecules (though there are monomeric exceptions). The mechanism for conventional kinesin 1's processive movement along microtubules is based on a hand-over-hand alternation of the motor domains (Gennerich and Vale, 2009). The motor domains alternately bind and hydrolyse ATP. ATP binding to a lead head causes the neck linker region to dock onto kinesin's core, which, in turn, strongly biases the trail head forward toward the next binding site on the microtubule. The trail head, now in the lead position, completes its forward movement by diffusive search and binds to the next microtubule binding site to complete the stepping cycle (Vale and Milligan, 2000).

Dyneins are the most conserved of the cytoskeletal motor families, there being only two principal classes: axonemal and cytoplasmic. The coordinated movement of arrays of different dyneins cause the sliding of microtubules against one another, driving the beating motion of cilia and the undulating motion of flagella, for example in a swimming

sperm. Cytoplasmic dyneins are responsible for the majority of minus end-directed intracellular transport on microtubules. They are important in many forms of cargo trafficking, moving various vesicles and organelle cargoes, and they play important roles in mitosis, cell polarisation and directed cell movement (Höök and Vallee, 2006). Compared to myosins and kinesins, dynein is a much larger protein complex. It comprises one to three heavy chains, each > 500 kDa, and variable numbers of associated intermediate and light chains that bind as accessory subunits. Each heavy chain consists of four structurally and functionally distinct domains: head, tail, stalk and microtubule binding domains. The head contains six AAA+ modules (four of which bind ATP) arranged in a hexameric ring. It is responsible for the motor activity. The ~160 kDa N-terminal tail emerges from one side of the ring and binds most of the accessory subunits and cargoes. It functions to dimerise (or trimerise) the heavy chains. The stalk, formed from a coiled coil, emerges from the opposite side of the ring and the microtubule binding domain is at its tip. The precise method by which dynein produces directed motion from its mechanochemical cycle is still less well understood compared to kinesin and myosin. However, recent EM and structural studies have shed light on this process, revealing the intramolecular interactions different parts of the motor make and the parts of the complex that undergo conformational changes linked to the different nucleotide states (Burgess et al., 2003; Burgess, Walker, Sakakibara, et al., 2004; Kon et al., 2011; Roberts et al., 2009; Roberts, 2009).

All three cytoskeletal motor families share the ability make directed movement along a filamentous track. The functional characteristic common to each class of motor is that there are coordinated cycles of ATP hydrolysis, track binding/unbinding and conformational changes. Each motor goes through a cycle of track attachment and detachment, which is coupled to ATP hydrolysis and product release. While attached to the track, the motor moves a mechanical element in the direction of travel and while detached from the track, the motor reverses this movement ready for the next cycle (Howard, 2001). In kinesins this mechanical element is a 13 amino acid neck linker region which on docking to a lead head swings the detached trail head forward and allows it to reach the next binding site (Rice et al., 1999). In dynein (cytoplasmic form) the mechanical element appears to be a much larger domain also called the linker. During dynein's mechanochemical cycle the linker swings across the AAA+ ring acting as a distance multiplying lever (Roberts et al., 2009). Its N-terminus tip moves ~17 nm during its swinging movement. In myosin, the mechanical element is called the lever and is generally an extended α-helix that emerges from the motor domain and which carries bound light chains. Small nucleotide-dependent movements within the motor are amplified

6

into a nm-sized swing of the lever's N-terminus, the tip moving between $10 - 25$ nm depending on the myosin. In the following section the features of myosin motor proteins are described in more detail.

1.5 Myosins

Myosins form a superfamily of actin-based motor proteins that participate in a host of biological processes, including muscle contraction, hearing, vision, cell migration, cell division and intracellular transport (Coluccio et al., 2008; Kieke and Titus, 2003; Krendel and Mooseker, 2005; Mermall et al., 1998). All myosins can reversibly bind to actin filaments and use the energy from Mg^{2+}-ATP (MgATP) hydrolysis to generate force. Frequently, the force generated is used to produce directed movement along actin filaments, but in other cases myosins may act as passive strain sensors, as force-generating anchors that maintain tension in the cell, or as actin assembly regulators (Bloemink and Geeves, 2011; Coluccio, 2007).

Myosins share and are identified by a conserved catalytic motor domain that interacts with fibrous actin (F-actin) and hydrolyses MgATP. To date, the myosin family has been separated into at least 35 classes, based on a phylogenetic analysis of motor domain sequences (Odronitz and Kollmar, 2007). In humans, some 39 genes encode for myosin heavy chains (Berg et al., 2001), which fall into 12 of the different classes (Peckham and Knight, 2009). Testifying to their importance in many eukaryotic cellular processes, myosin mutations are associated with diseases, including cardiomyopathies, neurological disorders, immunodeficiencies, as well as hearing and vision loss (Kieke and Titus, 2003; Sellers, 1999).

Skeletal muscle myosin, first described in 1864 by Willy Kühne and co-workers, was the first myosin to be discovered. It is the most abundant and well studied of the myosins, and its involvement in muscle contraction has been extensively studied (Geeves and Holmes, 2005). The first non-muscle myosin to be identified, was discovered over one hundred years later Pollard and Korn (1973). Being the first myosin to be discovered (and for many years the only type known), skeletal muscle myosin is often referred to as 'conventional' myosin, and the many other classes subsequently discovered are termed 'unconventional' myosins (Cheney and Mooseker, 1992). The most frequently used nomenclature for myosin classes uses a Roman numeral, *e.g.* myosin I, myosin II, myosin V etc. (Cheney, Riley, et al., 1993). The numeral indicates the order in which the myosin classes were discovered following the initial classifications of the first non-muscle and original skeletal muscle myosins as class I and II respectively. However, in this thesis, for simplicity,

Arabic numerals will be used for class numbering instead *i.e.* myosin 1, myosin 2, myosin 5 etc.

1.5.1 Basic structure

Myosins are composed of one or two long polypeptide chains, termed heavy chains (HCs), and one or more light chains (LCs). Typically, each HC divides into three regions corresponding to three structurally and functionally distinct domains. Beginning at the N-terminus is the motor domain, followed by the LC-binding regulatory domain (alternatively termed neck, light chain-binding domain or lever), followed by the tail, Figure 1.1.

In myosin 2, the region of the molecule encompassing both the motor and LCs bound to the regulatory domain form a distinct actin-binding 'head'. Historically, the term 'head', or alternatively 'cross-bridge', was applied to what later became known as the motor and regulatory domain separately. In this thesis, the term head will be used in its original way for all myosins discussed *i.e.* as a label for the motor domain *and* regulatory domain plus associated LCs. (In some studies authors now refer to the motor domain alone as the head.)

In some myosins types, HCs can self-associate via their tails (see 1.5.2.3) to form two-headed molecules. Such myosins are sometimes referred to as 'dimers' and single-headed molecules are termed 'monomers', although strictly each species are often made from more than two polypeptide subunits *e.g.* myosin 2 has two HCs and four LCs and is technically a hexamer, while myosin 5a has two HCs, twelve lever-bound LCs and two tail-bound LCs, making it a hexadecamer.
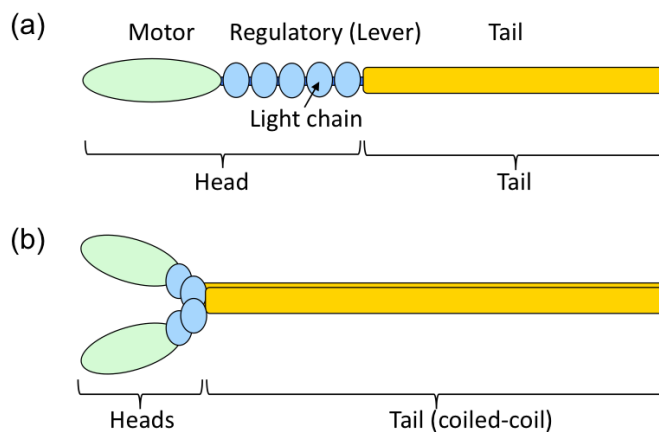
Figure 1.1: Cartoon of general domain structure of myosin.

(a) cartoon example of a single-headed myosin molecule with motor (green)-lever (blue)-tail (yellow) domain structure. In this particular example, the lever binds five light chains (light blue ovals), as is the case for myosin 7a. **(b)** cartoon example a two-headed myosin dimerised via a coiled coil tail. Each heavy chain binds two light chains. This is the general structure of myosin 2 molecules.

In the case of myosin 2, the two-headed molecules can also assemble (at physiological ionic strength or below) into larger filamentous structures known as thick filaments. In vertebrate striated muscle, bipolar thick filaments form through the parallel packing together of tails of ~300 molecules, to form a backbone with the myosin heads projecting outwards at either end, Figure 1.2(a),(b). These 'thick filaments' interdigitate with actin 'thin filaments' to form the sarcomere, the contractile apparatus in muscle cells. In smooth muscle tissue, thick filaments of different heavy and light chain composition form, with a different tail packing structure. These are known as side polar thick filaments, Figure 1.2(c)
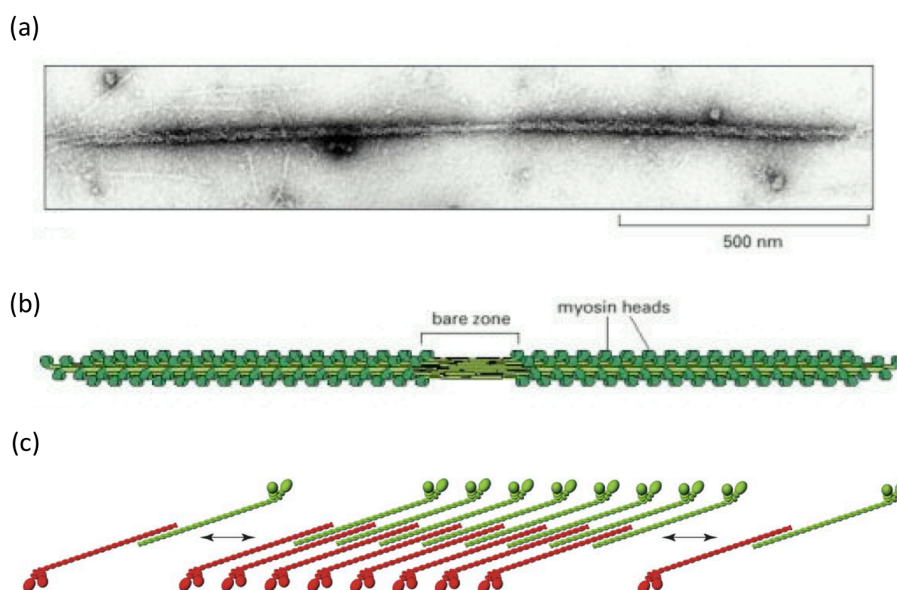
Figure 1.2: Myosin 2 thick filament structures

(a) electron micrograph of myosin 2 bipolar thick filament from frog muscle. **(b)** schematic diagram (not to scale) of bipolar filament where myosin tails have antiparallel interactions in the centre creating a 'bare zone' and parallel arrangement at the filament ends where the myosin heads project outwards. (a) and (b) taken from Alberts et al. (2002) **(c)** schematic diagram of the formation of side polar filaments formed in smooth muscle, taken from Craig and Woodhead (2006).

1.5.1.1   Proteolytic fragmentation of myosin 2

The myosin 2 heavy chain is susceptible to proteolysis, including at two sites in the tail. Though originally applying to myosin 2, the names of the fragments produced have become common labels used across myosin classes, and also as names for equivalent recombinant constructs. The protease chymotrypsin cleaves the myosin 2 HC at a site ~130 kDa from the N-terminus (about halfway along the tail), producing two fragments. The heavier fragment contains the head and first half of the tail and is termed heavy meromyosin (HMM). The remaining tail fragment is known as light meromyosin (LMM), the names originating from the original proteolysis study (Szent-Györgyi, 1953). HMM retains enough coiled-coil tail sequence for two HCs to remain associated, forming a two-

headed molecule (with LCs) (~340 kDa) like its parent, but with a truncated tail. The heads still bind actin and exhibit similar activity to the full length molecule. The HMM fragment can be further digested at the head-tail junction (C-terminal to the second LC binding site) using papain (Lowey et al., 1969). This produces an enzymatically active intact head with bound LCs (~120 kDa) called papain subfragment-1 (S1). The remaining short tail is called subfragment-2 (S2) (Knight and Trinick, 1984). Alternatively, digestion of HMM with chymotrypsin cleaves the head *within* the lever, C-terminal to the first LC, yielding S1 but with only the first LC.
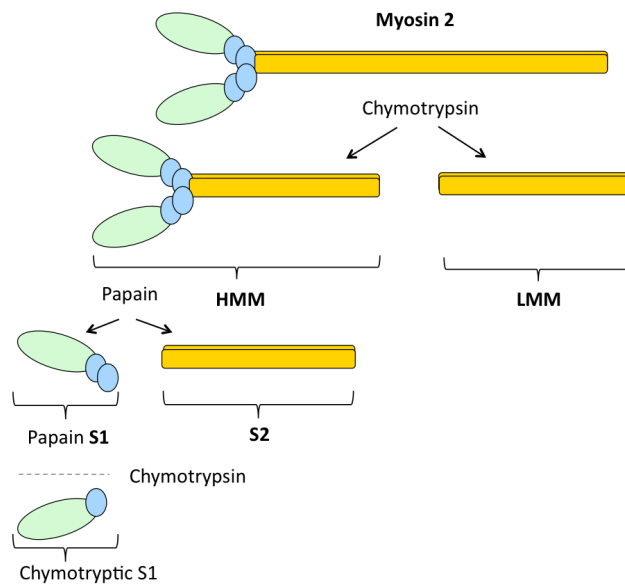
Figure 1.3: Proteolytic fragmentation of myosin 2.

The terms S1 to describe a single head and HMM to describe a two-headed molecule with a truncated tail, have become common myosin parlance. They are frequently applied to myosins other than skeletal muscle myosin and are often used when describing experiments with recombinant proteins of equivalent structure. HMM and S1 will be used later in this thesis when referring to recombinant myosin 5a constructs.

1.5.2 Principal domains

1.5.2.1 Motor

The ~80 kDa N-terminus motor domain is the most conserved myosin domain and contains the actin- and nucleotide-binding sites. It is chiefly responsible for the catalytic and motor activity of the protein, undergoing structural rearrangements on MgATP hydrolysis and product release.

Crystal structures of heads for a range of myosins, *e.g.* (Coureux et al., 2004; 2003; Dominguez et al., 1998; Kollmar et al., 2002; Ménétrey et al., 2005; Rayment, Rypniewski, et al., 1993), show a similar essential structure for the motor, Figure 1.4. A

7-stranded β-sheet forms the core of the motor and is surrounded by α-helices. A deep cleft divides one side of the motor into the so-called upper and lower 50K domains (the 50K domain being a 50 kDa motor fragment produced by proteolysis of surface loops). Both upper and lower 50K domains are involved in actin-binding. The cleft opens and closes in response to nucleotide binding, which is coupled to weak (cleft open) or strong (cleft closed) actin-binding states. A small pocket near the apex of the 50K cleft forms the nucleotide-binding site and contains a phosphate-binding 'P-loop' flanked by two catalytic loops named Switch 1 and 2, similar to those found in G-proteins (Smith and Rayment, 1996a). Switches 1 and 2 are positioned to allow communication between the nucleotide and actin-binding sites. Switch 2 also connects to the 'relay helix' which is in turn positioned to make interactions with the so-called 'converter domain', a compact structure at the C-terminus end of the motor. The converter domain anchors the N-terminus end of the lever helix in the motor.
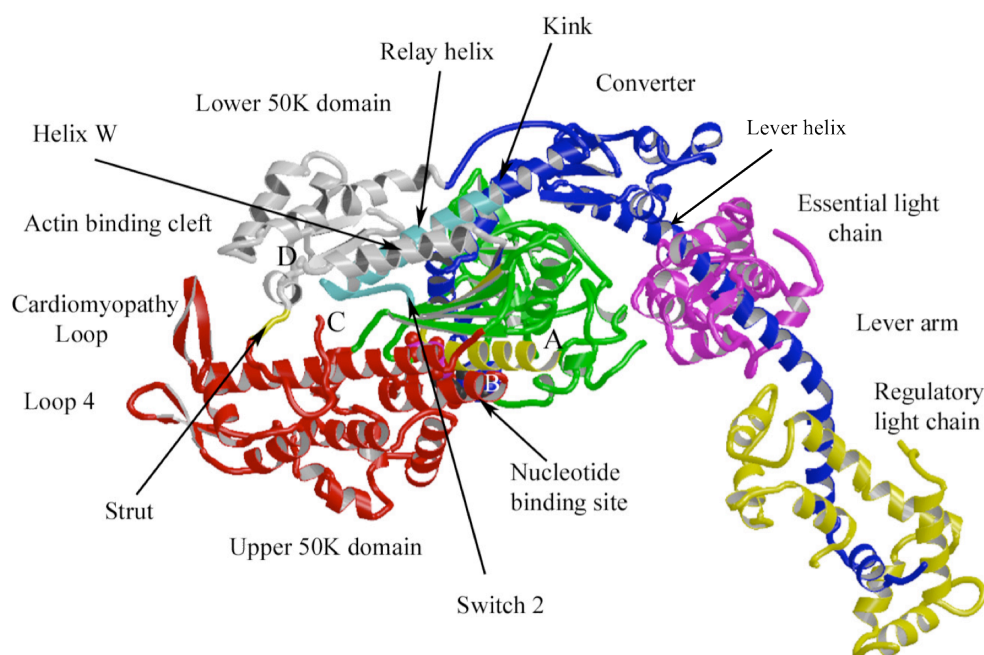
Figure 1.4 Structure of the myosin motor head.

Ribbon diagram of chicken myosin 2 head crystal structure (Rayment, Rypniewski, et al., 1993), adapted from Geeves and Holmes (2005). The head appears in the orientation as if bound to actin (at the actin-binding cleft) and viewed from the pointed end of the filament. Subdomains of the motor and the two LCs bound to the lever are shown. N-terminus region (green), nucleotide-binding P-loop and adjoining helix (yellow), upper 50K (red), lower 50K (grey) with relay helix (light blue). Converter and lever helix (blue) with Essential light chain (pink) and Regulatory light chain (yellow). In this conformation, the relay helix is straight and the converter and lever have rotated ~60° (about the end of the SH2 helix in the motor, not seen in this view) such that the lever is post-powerstroke in a 'down' position but the actin-binding cleft is open. The state is also termed 'post-rigor'.

1.5.2.1.1   Molecular basis of the powerstroke

Structural rearrangement in the motor (twisting of the core β-sheet), triggered by the strong binding to actin, allows a kink in the relay helix to straighten. The straightening of the relay helix rotates the converter by ~60°, which in turn rotates the lever helix by the same amount (Geeves and Holmes, 2005). The rotation of the lever helix constitutes the myosin 'powerstroke' (or working stroke), the mechanical action capable of transmitting force and generating directed movement of F-actin. The converse action, the creation of a kink in the relay helix (which occurs when myosin is detached from actin and hydrolyses MgATP), reverses the converter-lever rotation, priming the head for another cycle (Fischer et al., 2005). This is known as the recovery stroke.

1.5.2.2   Regulatory domain – Lever

The regulatory domain extends from the motor and is typically a long α-helix (the lever helix) bearing one or more copies of a characteristic LC-binding sequence, known as an 'IQ motif' (see 1.6.1). Calmodulin or camodulin-like LCs bind to the IQ motifs and are involved in regulating the activity of the motor (hence the domain's name), frequently in response to phosphorylation or altered calcium levels, (1.6.1). The extended regulatory domain is also known as the lever(-arm), because of its role in force transmission. The nucleotide-dependent ångström-sized structural changes within the motor are amplified into a nm-sized movement of its C-terminal end. The bound LCs are thought to stabilise and rigidify the lever, thus providing a mechanism for the regulation of force transmission. In this thesis, which focuses on the structure and mechanics of this domain, the term lever (rather than neck, regulatory or light chain-binding domain) will be used throughout. The importance of the lever in myosin function is discussed further in 1.5.3.

1.5.2.3   Tail

The C-terminal tail domain follows the lever. It is the most variable myosin domain in terms of sequence, length, subdomain composition and organization. It can target the myosin to a particular cellular location and determine its function *e.g.* through binding to different cargo or binding partners, or associating with other myosin tails to form filaments such as for myosin 2. The tail can also be involved in regulation, for example in myosin 5a, where the tail folds back and contacts the motor domain in absence of cargo, shutting the molecule down in a stable folded conformation (Thirumurugan et al., 2006).

Another key property of the tail is that it determines whether one HC can associate with another to form a two-headed molecule. Some myosins such as myosin 2 and myosin 5 contain long α-helical tail sequences featuring heptad repeats. These are 7-residue

sequences labelled *a-b-c-d-e-f-g*, in which hydrophobic residues consistently occur at positions *a* and *d*. The *a* and *d* residues form a hydrophobic stripe that slowly winds around the surface of the α-helix. When two such helices meet, the hydrophobic stripes pack against one to avoid solvent exposure and the helices form a left-handed superhelix known as a coiled-coil (reviewed in Parry et al. (2008)). This is the chief 'dimerisation' mechanism seen in myosins. Coiled-coil tail sequences cause HC dimerisation in myosins 2, 5 and 18. Computationally predicted coiled-coil sequences are also found in myosins 6, 7 and 10 (Krendel and Mooseker, 2005), but there is still controversy over whether these really function to dimerise two HCs. Knight et al. (2005) and Peckham and Knight (2009) have presented strong evidence to suggest that these sequences are actually stable single α-helices (SAHs) that do not dimerise, but rather may extend the mechanical length of the lever.

1.5.3 Actomyosin interaction and the swinging lever arm model

The binding of myosin to actin and the conformational changes associated with steps in the ATP hydrolysis cycle were first elucidated in the context of muscle contraction. Electron microscopy showed linkages, 'cross-bridges', between the thick and thin filaments of the sarcomere, which were subsequently shown to originate from the thick filaments and identified as the heads (corresponding to the S1 region) of individual myosin molecules (Huxley, 1963; 1957). Later it was proposed that the heads make cyclic interactions with the actin and undergo a conformational change which 'rows' the thin filament along relative to the thick filament, the process being fuelled by ATP hydrolysis (Huxley, 1969). This became known as the swinging-cross-bridge theory. Shortly after, the Lymn-Taylor cycle was suggested as a model of how the ATP hydrolysis cycle might be linked to the cross-bridge conformational changes (Lymn and Taylor, 1971). The events in the cycle, thought to be common to all myosins, are summarised in Figure 1.5. (See also Figure 1.7 for a description of the mechanochemical steps in the cycle.)

With the solution of the first crystal structure of skeletal myosin S1 (Rayment, Rypniewski, et al., 1993) and an atomic model of F-actin (Holmes et al., 1990), an atomic model of the actomyosin complex could be constructed (Rayment, Holden, et al., 1993). This showed that a more globular motor domain was likely bound to the actin while the extended light chain-binding domain projected away from the filament, and could act as a distance multiplying lever in the swinging cross-bridge theory. This idea was consistent with earlier spectroscopic evidence that had indicated that most of the cross-bridge mass did not actually move during muscle contraction (Cooke, 1986). A new crystal structure of the myosin head (in a different nucleotide state) with its lever projecting in a different

direction (rotated 60° to the first structure) (Smith and Rayment, 1996b), and a key experiment by Uyeda et al. (1996), confirmed the idea that the motor domain likely adopts a fixed geometry on actin while the C-terminal light chain-carrying lever swings through a large angle during the cross-bridge cycle. The Uyeda et al. experiment involved the creation of *Dictyostelium* myosin heads with genetically engineered levers of different lengths, and the speed at which they propelled actin filaments across a surface covered in myosin molecules (the actin gliding motility assay of Kron and Spudich (1986)) was found to be a linear function of the length of the lever. Hence, the swinging cross-bridge theory was refined and became the swinging lever arm model (Holmes, 1997; Spudich, 1994).
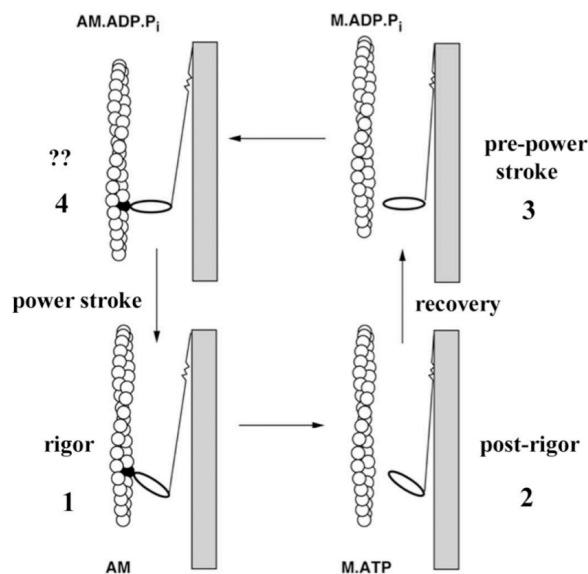
Figure 1.5: The Lymn-Taylor cross-bridge cycle.

Figure taken from Geeves and Holmes (2005). A = actin, M = myosin, AM = actomyosin. The myosin head (*i.e.* motor and lever) is depicted as a single lozenge-shaped domain, connected via its tail (a line) into the thick filament backbone (grey). The head cyclically interacts with the actin filament (spheres) and undergoes a powerstroke (a rotation about its contact point with actin) on P_i release. This is the force-generating step that causes filament sliding in the sarcomere. ATP binding frees the head from actin and hydrolysis is associated with a recovery stroke, where the head rotates back so that it is primed to undergo another cycle. See also Figure 1.7.

In the swinging lever arm model, the motor maintains a fixed position on actin while internal nucleotide-dependent rearrangements (1.5.2.1.1) cause the converter subdomain to rotate through ~60° with respect to the main body of the motor. Since the start of the lever helix is firmly anchored to the converter it too rotates through 60°, carrying the light chains with it. This movement of the lever as a semi-rigid body constitutes the myosin powerstroke, Figure 1.6. The lever acts as a distance multiplier, amplifying the ångström-sized movements within the motor into a nm-sized movement of its C-terminal end. The distance moved by lever tip in the direction of the filament axis is known as the working

distance (Howard, 2001) (or more loosely the powerstroke or working stroke). For myosin 2 the working distance is ~10 nm.
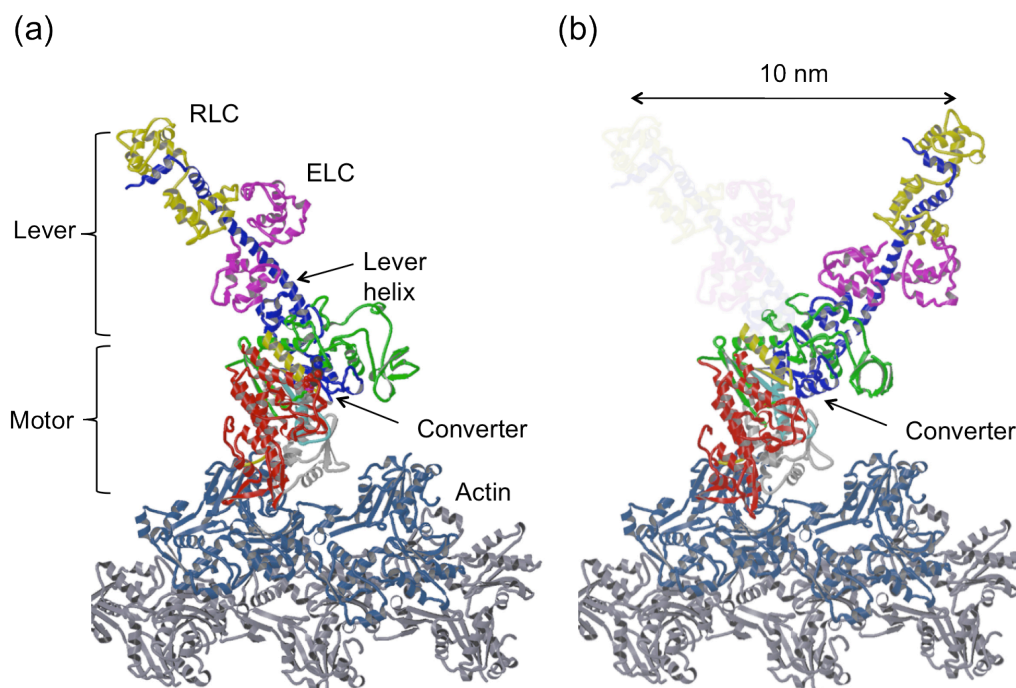
Figure 1.6: The myosin powerstroke.

Atomic models of myosin 2 in **(a)** the strongly bound conformation before the powerstroke and **(b)** the 'rigor' state conformation at the end of the powerstroke (the lever position of the pre-powerstroke state is shown faintly in (b) to illustrate the overall movement of the lever). Image adapted from Geeves and Holmes (2005). The model in (b) is based on EM images of decorated actin and the crystal structures of actin and myosin 5. The model in (a) is extrapolated from comparison with a myosin 5 pre-powerstroke structure. In both cases the lever of chicken skeletal myosin has been used for the lever. Note the movement of the converter between the states. It carries the entire lever and LCs with it as it rotates ~60° about the motor centre.

For the swinging lever arm model to work, clearly the lever needs to maintain its structure as it moves and when under load. In the case of muscle myosin, this load can be thought of as the force exerted by the actin filament being moved by the bound motor domain (if the myosin tail is assumed to be 'anchored' in the thick filament). In the case of a processive cargo-carrying myosin, like myosin 5a, the force on the lever is felt at its C-terminal end, as a tension transmitted through the coiled-coil tail due to the resistance of the cargo to movement (due to the viscosity or obstacles in the cytoplasm). In either of these cases, if the lever was too pliant and buckled under load during its swing, then it would not fulfil its role as a distance multiplying lever and the force generated in the motor domain would not be efficiently transmitted. For this reason, the flexural rigidity (bending stiffness) and the specific mechanical properties of the lever are crucial to myosin function.

In contrast, it can also be advantageous for the lever to bend or deform to some extent. In the general 'powerstroke model' for cytoskeletal motor proteins, the motor needs an elastic

element (like a spring) which can store mechanical energy (Howard, 2001). When the motor undergoes its conformational change, strain is stored in this 'internal spring' and the restoring force is the motor force. Relief of the strain is the driving force for the motion, the powerstroke. In myosin the 'spring' has not been definitively located in the structure of the protein. Indeed, elastic energy may be stored in more than one structural element. Flexing of the lever was suggested as one possible source of elasticity (Howard and Spudich, 1996). Alternatively, the elasticity may be due to deformation of the converter domain, or due to bending at a pliant region at the junction of the lever helix and converter. Evidence for the latter has been seen in crystal structures of scallop myosin 2 (Houdusse et al., 2000). In this case the internal spring would be more analogous to a torsion spring, which provides a restoring torque proportional to the angle it is rotated away from its equilibrium position. EM images of myosin 5a suggest that both lever bending and distortion at the motor-lever junction may play a role in this myosin's function (Burgess et al., 2002; Oke et al., 2010; Walker et al., 2000).

Given the importance of the mechanical properties of myosin levers to function, it is interesting that variety in lever structure is found across different myosin types. This diversity suggests possible evolution of structure for different mechanical functions. One obvious difference is the different myosin lever lengths found. Extending the length of the lever, *e.g.* by addition of more LC-binding IQ motifs, will clearly extend the achievable working distance. While class 2 myosins possess only 2 IQ motifs, myosin 5a and 5b have one of the longest levers with 6. The linear relationship between step size and lever length has been shown in many studies (Köhler et al., 2003; Moore et al., 2004; Purcell et al., 2002; Ruff et al., 2001; Sakamoto et al., 2003; 2005; Uyeda et al., 1996; Warshaw et al., 2000). Interestingly, the myosin 1b lever is alternatively spliced in different tissues to give 3 isoforms with 4, 5 and 6 IQ motifs respectively. It has been suggested that the differently spliced levers confer different mechanical properties on the protein, important in different force-sensing roles (Laakso et al., 2010).

As well as the lever length, the spacing at which the IQ motifs appear along the lever also varies across myosins, ranging from 22 to 36 amino acid repeats (Houdusse et al., 1996). Precise IQ motif spacing is important as it dictates the relative orientation and so potential interactions between neighbouring LCs along the lever. This is discussed further in 1.6.1. More recent work has also shown that for some unconventional myosins the typical LC-IQ motif building block of the lever is varied by other structural components. For example, myosin 10 has a stable single α-helix (SAH) directly following its 3 IQ motifs which is thought to extend its lever and the working distance (Knight et al., 2005), though perhaps

the SAH has a different flexural rigidity to the LC-IQ motif section. Myosin 6 has perhaps the most unconventional lever, beginning with a unique bent insert sequence that binds a CaM with Ca^{2+} bound, followed by a conventional LC-IQ motif subunit followed by a 3-helix bundle (that may unfold as a lever extension (Mukherjea et al., 2009)) and then a SAH which may further augment the lever (Spink et al., 2008).

Gaining a better understanding of how the mechanical properties of the myosin 5a lever derive from its substructure is one of the goals of part I of this thesis. In part II, the influence of IQ motif spacing on the properties and function of the molecule is studied.

1.5.4 Actomyosin ATPase cycle

All myosin isoforms studied to date are ATPases – enzymes that catalyse the breakdown of ATP into ADP and P_i. In myosins, the substrate is MgATP and the ATPase activity is strongly stimulated ('activated') by binding to actin. Binding of myosin to actin catalyses the release of P_i, whereas the release of myosin from actin is catalysed by the binding of ATP. In this way, myosin's hydrolysis cycle is tightly coupled to its cycle of attachment to and detachment from actin. The events in the actomyosin ATPase cycle appear to be common to all myosins known, though the rates and the equilibria between steps in the cycle vary widely depending on the specific isoform. The common hydrolysis pathway and the accompanying conformational changes are depicted in Figure 1.7 and the events in the cycle described in the caption. *NB*: the figure is an oversimplification of the exact details concerning the coupling between the chemical and mechanical cycles (see Geeves and Holmes (2005)) but conveys the essential mechanism.

Though all myosins appear to possess a common ATPase mechanism within the motor (the catalytic site being strongly conserved), specific sequence variations in the motor still give rise to contrasting kinetic properties. There are important variations in absolute and relative rate constants suggesting that the biochemistry and mechanics of F-actin interaction are 'tuned' according to the myosin's physiological function (La Cruz and Ostap, 2004). The duty-ratio of a given myosin (defined as proportion of time a myosin spends strongly bound actin during its hydrolysis cycle (La Cruz et al., 1999)) can vary greatly, and is an indicator of whether the myosin is processive (meaning it can make multiple interactions with the track per encounter, and thus move distances greater than a single ATP-driven step). For example, mammalian myosin 5a, the subject of this thesis, is a highly processive two-headed molecule, capable of taking multiple 'steps' along actin before detaching, see 1.8.2.3 (Mehta et al., 1999). This is partly achieved by having a very high duty-ratio, spending 70 % of its cycle attached to actin. In contrast, non-processive myosins undergo only a single cycle before detaching and cannot support movement over

long distances. Skeletal muscle myosin 2 has a low duty-ratio, and has evolved to spend very little time in its cycle bound to actin (El-Mezgueldi and Bagshaw, 2007). This allows myosin 2 to function properly in large asynchronous ensembles of thick filaments. As an isolated molecule myosin 2 is not processive (it would readily diffuse away from the actin between steps). However muscle contraction (*i.e.* thick-thin filament sliding in the sarcomere) is made possible because, in the sarcomere structure, the actin filaments are continually held in close proximity to the ~300 myosin 2 molecules arrayed in each thick filament. In general, for myosins with low duty-ratios (mostly detached from actin) the release of P_i is the rate limiting step in the hydrolysis cycle, while for high duty-ratio myosins (mostly attached to actin) the rate limiting step is ADP release, Figure 1.7.
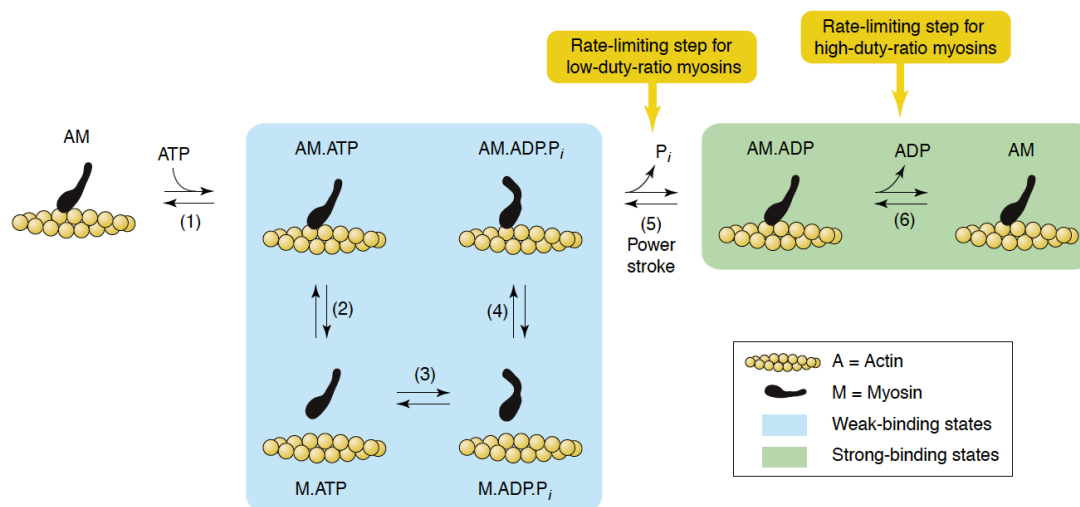
Figure 1.7: The actomyosin ATPase cycle.

> Figure taken from La Cruz and Ostap (2004). M = myosin, AM = actomyosin. In the absence of ATP myosin binds tightly to actin forming the 'strong' or 'rigor' complex. ATP binding to the myosin head (step 1) weakens the affinity and causes the head to rapidly detach from actin (step 2). Myosin then hydrolyses ATP to form the stable products complex (M.ADP.P_i) and there is an accompanying conformational change, the 'recovery stroke', where the lever swings and the head is primed ready to deliver its powerstroke. Myosin then re-binds actin with products bound (step 4). P_i is released from the binding pocket accompanied by the force-generating powerstroke (the lever swinging back again) (step 5). Myosin now occupies a strongly bound state again. ADP is released (step 6) and the cycle can repeat.

1.5.5 Diversity in structure and function

Despite having a relatively conserved motor domain, myosins are divergent in their lever and tail domains. The tails contain a host of protein- and membrane-binding domains such as SH3, TH1-4, FERM, PDZ, and PH domains, Figure 1.8 (Krendel and Mooseker, 2005). These serve to target the myosins, allow them to bind different cargoes or mediate protein-protein interactions. The purpose of many of the motifs identified remain to be determined (Foth et al., 2006; Mooseker and Foth, 2007). Variation in the length and composition of the lever must also play a role in differentiated functions, see 1.5.3 above. Variable

18

numbers of LC-binding IQ motifs are found across the classes, dictating a range of lever lengths. Some classes such as myosins 6, 7 and 10 may also have lever extensions formed by stable single α-helices, see 1.5.3. These alternate lever components might serve to alter the mechanical response of the lever under load according to function.

The structural diversity in myosins reflects the many functions that myosins have evolved to perform. These include participation in force and tension generation, cell movement and adhesion, endocytosis, exocytosis, organelle movement and localization, mRNA transport and localisation, cytoskeletal cross-talk, signal transduction, as well as roles in cell polarity determination and nuclear function (Mooseker and Foth, 2007).
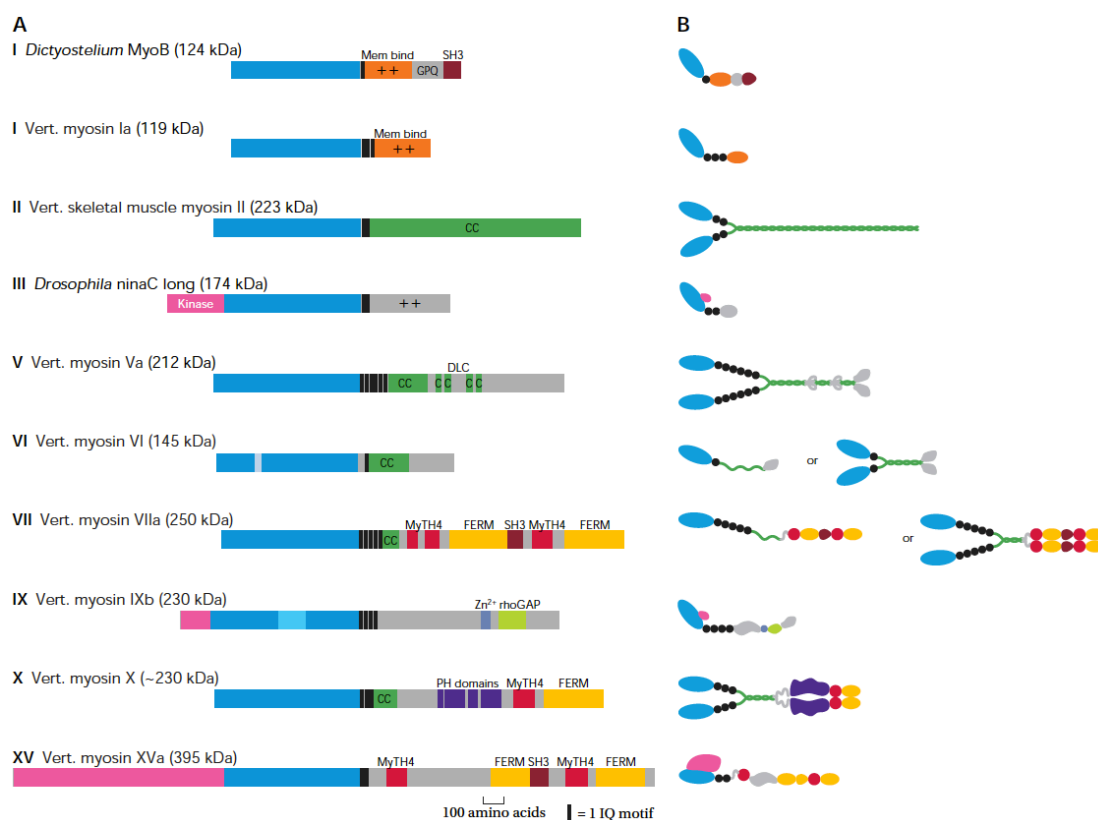
Figure 1.8: Diverse myosin domain structures.

> Image from Krendel and Mooseker (2005). *A*: bar depiction of domain structures found in examples from various myosin classes. Blue = motor, Green = coiled coil (predicted), Black bar = IQ motif. Other colours represent known domain motifs. *B*: schematic depiction of myosin structures (some putative). Note, since this figure was published domain structures in several of the classes have been revised. In particular myosins 6, 7a and 10 contain stable single α-helices at there is debate over whether they form dimers. See also Peckham and Knight (2009) for a more recent depiction of myosin structures.

1.6   Calmodulin and calmodulin-like light chains

Calmodulin (CaM) is a highly conserved calcium-binding protein expressed in all eukaryotic cells. It acts as a messenger protein; responding to changes in Ca^{2+} levels by binding or releasing Ca^{2+} cations and modifying its conformation and interaction with a wide variety of target proteins, including myosins, which it binds to as a light chain (LC). The amino acid sequence of CaM is identical in higher vertebrates, with only a single amino acid change in the sequence of the electric eel protein in fish (Lagacé et al., 1983).

CaM is formed from a 148 residue (16.7 kDa) polypeptide chain that folds into a dumbbell-shaped molecule with two homologous globular domains, known as N- and C-lobes according to their position relative to the N-terminus, Figure 1.9. The lobes are joined by a flexible inter-lobe loop that allows the protein to 'wrap around' its target. CaM can bind up to four Ca^{2+} cations via four helix-loop-helix motifs known as EF-hands, two in each lobe (each separated by linkers). The EF-hand motif consists of two roughly perpendicular α-helices joined by a 11-14-residue loop (Grabarek, 2006). Atoms in six residues of the loop create an electronegative environment to coordinate the Ca^{2+} cation. Residues at positions 1, 3, 5, 7, 9 and 12 in the loop can all be involved in cation binding (Grabarek, 2006).
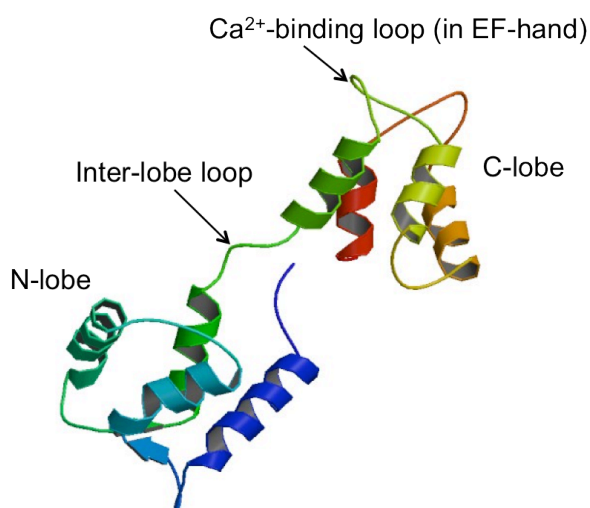
Figure 1.9: Crystal structure of calmodulin without calcium bound

Ribbon depiction of calmodulin (CaM) from *Xenopus laevis* in apo conformation (without Ca^{2+} bound). PDB accession: 1CFD. The two lobe regions separated by the flexible inter-lobe linker are shown. The N-terminus is before the dark blue helix. In each lobe two helix-loop-helix EF-hand motifs occur (each joined by a linker). The loops in the EF-hands can bind Ca^{2+} cations.

CaM is a member of a the larger EF-hand family of proteins (Persechini et al., 1989), which includes the myosin light chains and troponin C (Collins, 1991; Kawasaki et al., 1998). The myosin light chains include the Essential Light Chain (ELC) and Regulatory

Light Chain (RLC) that bind the myosin 2 heavy chain at the lever helix adjacent to the motor. These and other myosin-binding LCs all share the same two-lobe dumbbell-like structure with EF-hands (Nelson and Chazin, 1998), though the ELC, RLC and other specific LCs lack most or all cation-binding capacity (Collins, 1991; Kawasaki et al., 1998).

NMR studies, (Finn et al., 1995; Kuboniwa et al., 1995; Zhang et al., 1995), showed that Ca^{2+} binding to CaM or other CaM-like LCs initiates a key conformational change; a reorientation of the EF-hand α-helices, opening the lobe regions from the so-called 'closed' conformation (where hydrophobic residues are buried in a tight helix bundle) to an 'open' conformation where hydrophobic residues (*e.g.* methyl groups on methionine residues) are solvent-exposed, Figure 1.10. These present a binding-site for CaM-binding regions on target proteins, which are typically basic amphipathic α-helices of variable sequence, ~20-25 residues long, with positive residues interspersed with hydrophobic ones (O'Neil and DeGrado, 1990; Rhoads and Friedberg, 1997). The hydrophobic residues of the target helix complement those in the exposed hydrophobic pocket of each of the open Ca^{2+}-CaM lobes. The target peptide can thus insert into the hydrophobic cavities of the CaM, becoming 'gripped' by the two lobes which wrap around the target, the inter-lobe loop acting as a hinge.
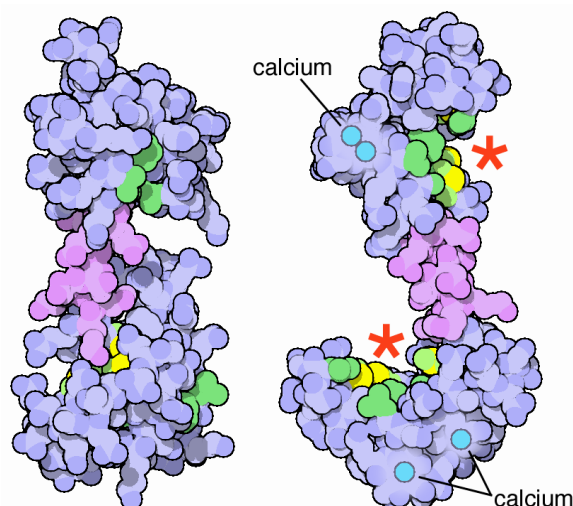
Figure 1.10: Ca^{2+}-induced conformational change in calmodulin

Cartoon of Ca^{2+}-free (left) and Ca^{2+}-bound (right) conformations of CaM. Image by David Goodsell from Protein Data Bank August 2003 Molecule of the Month article. The structure on the left is from *Xenopus laevis* PDB: 1CFD (also depicted in Figure 1.9), while the structure on the right is from *Homo sapiens* PDB: 1CLL. Lobes regions are coloured purple with the inter-lobe loop pink. The non-polar target-binding regions are coloured with carbon atoms green and sulphur atoms of methionine residues in yellow. On Ca^{2+} binding, the bundled lobes open out, exposing the non-polar amino acids in two grooves (red stars) which are ready to grip the target protein.

Although Ca^{2+}-binding is the chief mechanism that initiates CaM binding to a target, CaM and other LCs can also bind targets in a Ca^{2+}-free conformation, known as 'apocalmodulin' (apo-CaM). At low Ca^{2+} concentrations, apo-CaM binds tightly to proteins such as neuromodulin, neurogranin, and to many of the unconventional myosin classes, including myosin 5 (Rhoads and Friedberg, 1997; Wang et al., 2000; Wolenski, 1995). In these cases, an increase in Ca^{2+} can trigger a functional change in the target protein, frequently through dissociation or conformational change of one or more of the LCs (Bähler and Rhoads, 2002; Nguyen and Higuchi, 2005; Wolenski, 1995), hence the term 'regulatory domain' as an alternative name for the lever. (See also regulation by phosphorylation of the RLC in 1.6.1).

Modelling work and analysis of subsequently-obtained crystals of apo-CaMs bound to part of the mouse myosin 5a lever, showed how apo-CaM can bind to a target by adopting a so-called 'semi-open' conformation, Figure 1.11(a)-(b) (Houdusse et al., 1996; Houdusse and Cohen, 1995; Houdusse et al., 2006). Here, despite the lack of Ca^{2+}, the C-lobe is partially open, forming a shallow groove with hydrophobic residues sufficient to accommodate the target peptide, typically the residues in first half of the binding sequence. The N-lobe remains closed and interacts less with the remainder of the target sequence. Crystal structures of apo-Mlc1p (a LC for the yeast class 5 myosin myo2p) bound to fragments of the myo2p lever also exhibited the 'semi-open' C-lobe conformation but, in addition, at particular binding sequences, N-lobes appeared completely detached from the target helix (with no apparent interactions), the LCs adopting a so-called 'extended' conformation, Figure 1.11(c)-(d) (Terrak et al., 2005; 2003).

In cases of apo-CaM binding, the target recognition sequence is far more restricted than for Ca^{2+}-CaM binding (Rhoads and Friedberg, 1997). In myosins, a particular form of sequence, located in the myosin lever region and often found in tandem repeats, was recognised as a LC-binding sequence. This ~23-residue sequence was named an 'IQ motif', after the first two amino acids in the consensus sequence of the conserved core (Cheney and Mooseker, 1992).
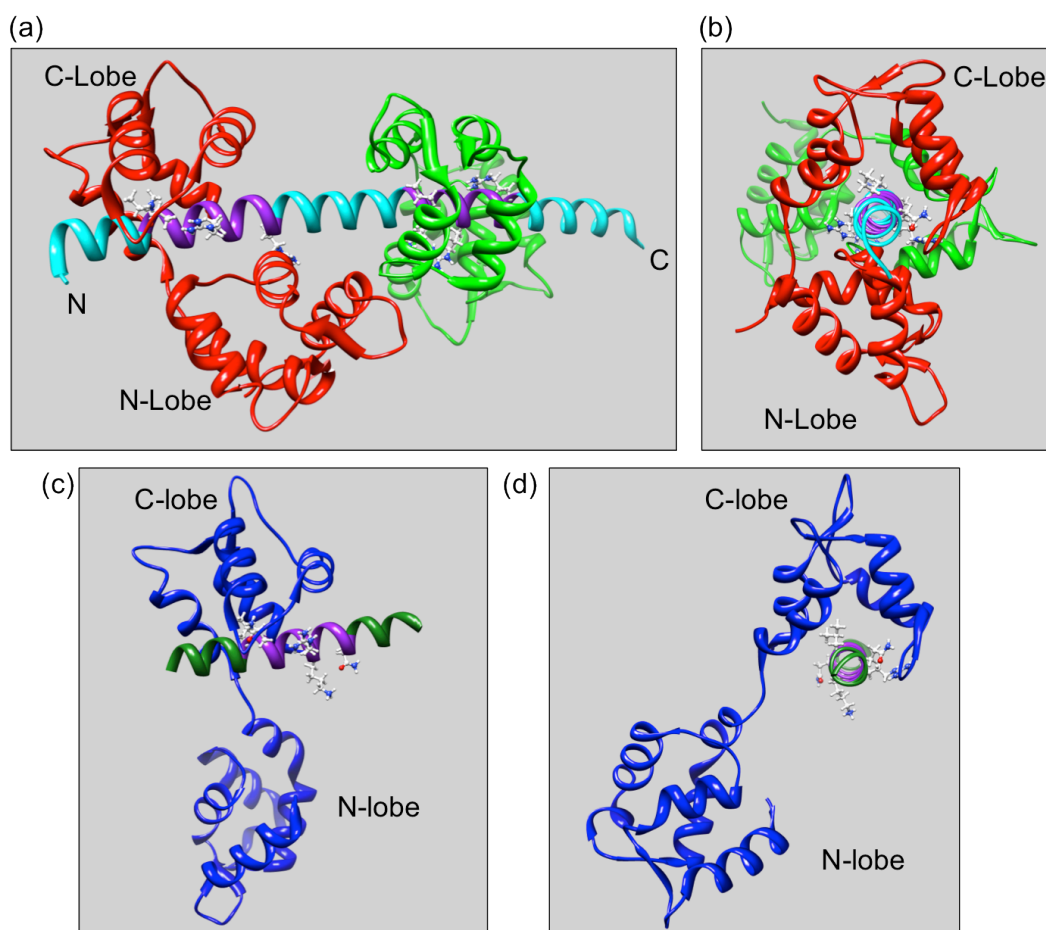
Figure 1.11: Apo-CaM bound to class 5 myosin lever fragments.

(a) ribbon depiction of apo-CaMs (red and green) bound to a fragment of mouse myosin 5a lever encompassing IQs 1-2. PDB: 2IX7. Heavy chain is shown in cyan with specific target sequences (IQ-cores, see 1.6.1) in purple and the conserved residues in ball and stick representation. **(b)** view of the same structure as in (a) looking end-on down the lever helix. **(c)** Mlc1p light chain (blue) bound to the fourth target sequence (IQ 4) of the myo2p lever (green). PDB: 1M46. Heavy chain in green, IQ-core in purple with ball and stick representation of conserved residues. **(d)** same as (c) but viewed looking end on down the lever helix. Due to a bulky lysine at position 7 in the IQ-core and lack of a positively charged amino acid at position 11 (normally arginine) the N-lobe in the lower structures is detached from the target peptide. In the upper structures it makes interactions with the lever helix (albeit fewer than the C-lobe).

1.6.1 IQ motifs in myosins

The IQ motif appearing in the myosin lever, was originally identified as a ~23-25 amino acid unit, with a conserved 11-residue core fitting the consensus sequence Ile-Gln-x-x-x-Arg-Gly-x-x-x-Arg (IQxxxRGxxxR) (where x is any amino acid) (Cheney and Mooseker, 1992). The terminology has subsequently become less precise, and in some cases the term 'IQ motif' has become synonymous with just the 11-residue core. In this thesis, in which the recurrence pattern of the core is investigated for myosin 5 (chapter 5), the term 'core' will be used to refer specifically to the 11-residue consensus sequence, and the term 'IQ motif' or 'IQ sequence' will refer to the encompassing ~23-25-residue sequence that can also interact with the LC.

Most myosins contain between one and six IQ motifs (Sellers, 2000). They usually appear directly after the motor domain, arranged in tandem to form the lever helix. Structural studies show that IQ sequences form α-helices, which are believed to be stabilised by the binding of CaM or a CaM-like LC. By enveloping the amphipathic IQ sequences, bound LCs are assumed to prevent collapse of the lever helix, rigidifying this extended domain, so that it may act as a semirigid body and thus transmit force or sustain load. The shortening of myosin heads on removal of LCs, as observed in early and recent EM studies (Flicker et al., 1983; Tominaga et al., 2012; Walker and Trinick, 1989), is consistent with this mechanical role of lever reinforcement. LC dissociation (or conformational change), rendering the lever pliant and non functional, may thus be one mechanism of myosin regulation.

As well as acting in a purely mechanical role, LCs are involved in myosin regulation in other ways. In crystal structures of myosin subfragment-1 (S1) (Fisher et al., 1995; Rayment, Rypniewski, et al., 1993), regions of the ELC are seen interacting directly with the converter within the motor, making stabilizing interactions that may be essential for the converter to function properly (and hence the motor and lever overall). In structures of smooth muscle myosin (Dominguez et al., 1998), the ELC is observed to contact a motor domain loop important in ADP release. The RLC, which binds the second IQ motif (C-terminal to the ELC), is also important in regulation (hence its name). In many myosins, specific RLC sites (*e.g.* Ser 19) must be phosphorylated to activate the motor (Sellers, 1999; 1991; Trybus et al., 1994), though the mechanisms of allosteric communication between the distal RLC and motor are not fully understood. Many myosins, both conventional and unconventional, are sensitive to Ca^{2+} levels, and this sensitivity is conferred via the LCs. Conformational changes in LCs associated with Ca^{2+}-binding/dissociation (*e.g.* lobes opening, closing or detaching from the heavy chain), can give rise to altered interactions between LC and motor, LC and lever helix, as well as between adjacent LCs.

Sequence alignments of myosin IQ motifs reveal the conservation and properties of residues at different positions in the IQ sequence Figure 1.12 (Houdusse et al., 1996; Rhoads and Friedberg, 1997). The first residue of the core is variable, mostly between Ile, Leu, or Val (with Met, Phe or Thr in some cases). Core position 11 varies between Arg, Lys or Gln. The Gly at position 7 is less well conserved. In addition, a hydrophobic residue (Φ) is frequently found at position 14, three residues beyond the core, Figure 1.12. A more general consensus sequence for the core, as seen in myosins, could therefore be written as:

[I,L,V]QxxxRxxxx[R,K,Q]xxΦ.
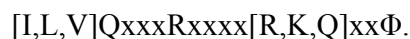
A general pattern exhibited through the core of myosin IQ motifs is that two hydrophobic 'patches' occur. One is located in the N-terminus half of the helix, usually beginning four or five residues before the core start, and the other (shorter) patch (~4-residues) occurs in the C-terminus half of the sequence, beginning at position 7. The hydrophobic residues are interspersed with positive residues, giving rise to a helix with hydrophobic and polar faces. The hydrophobic face is the target for CaM lobe binding.

Crystal structures of apo-CaM and other LCs bound to IQ motifs (Houdusse et al., 2006; Mukherjea et al., 2009; Terrak et al., 2005; 2003), indicate that the two hydrophobic patches in the heavy chain delineate the binding of the C- and N-lobes of the LC respectively. The C-lobe always binds to the N-terminus half of the IQ-core (IQxxxR), while the N-lobe interacts with the C-terminus part (GxxxR). The C-lobe, which adopts the semi-open 'gripping' conformation, interacts most strongly. Its hydrophobic pocket interacts with apolar heavy chain residues, and certain loops (*e.g.* linker 3 between helices F and G) can make van der Waals interactions and hydrogen bonds with the heavy chain, such as the conserved Gln. in position 2 in the IQ-core (Houdusse et al., 2006). The N-lobe is normally in a closed non-gripping conformation, interacting more weakly with the second half of the core, on the opposing face of the helix. The precise orientation of and interactions with the N-lobe is more variable, depending on the residues in the heavy chain sequence.

Inspection of sequences in Figure 1.12 reveals that there is variation not only in the number of IQ motifs different myosin types possess, but also in the amino-acid spacing at which the IQ motifs cores reoccur. Counting inclusively from position 1 in the core to the residue before the next core begins, the two IQ motifs in myosin 2 are always separated by 26 residues. In contrast, in the unconventional myosins, which carry from one to six IQ motifs, spacings ranging from 22 (class 1 myosin Myr4 (rat), GenBank acession: x71997) to 36 residues are seen (the latter separating the two IQ motifs in the myosin 3 sequence of NinaC (*Drosophila melanogaster*), Uniprot accession: P10676), Figure 1.12. The most common IQ-spacing in the unconventional myosins is 23 residues, for example in the separation of the three IQ motifs in bovine adrenal myosin 3, IQs 1-5 of myosin 7 (human) and IQs 1-3 of myosin 10 (frog). 22-, 26- and 25-residue spacings also occur, though less often and, intriguingly, always flanked by 23-residue spacings. For example, myosin 5a and 5b (and the plant homologue myosin 11) follow the spacing pattern 23-25-23-25-23, while myosin 8 (plants) follows the pattern 23-26-23. The class 9 myosin, Myr5 (rat), has the alternating pattern 23-22-23, Table 1.1.

Figure 1.12: Sequence alignments of IQ motifs in different myosin types.

Figure from Houdusse et al. (1996). *NB*: continues overleaf.

```
13)LAYLEKLRSNKMHNSIVM   IQ KKI RAKYY  R KQYL              a2 myo2, Yst. (773-935)p19524SP
          QISQAIKY      LQ NNI KGFII  R QRVNDE            a1
          MKVNCATL      LQ AAY RGHSI  R ANVF              a
          SVLRTITN      LQ KKI RKELK  Q RQLKQE            c2
          HEYNAAVT      IQ SKV RTFEP  R SRFL              a2
          RTKKDTVV      VQ SLI RRRAA  Q RKLKQLKADAKSVNH   c2
```

Myosin VI

```
14)    LRNRIIYRNKCVLI   AQ RIA RGFLA  R KQHRPRYQGIGKINK  a  95F Dr.    (801-840)q01989SP
```

Myosin VII

```
15)DMLLEVERDKAITDRVIL  LQ KVI RGFKD  R SNFL             a  Human     (733-871) C Petit PC
          KLKNAATL      IQ RHW RGHNC  R KNYG             a
          LMRLGFLR      LQ ALH RSRKL  H QQYR             c2
          LARQRIIQ      FQ ARC RAYLV  R KAFR             a2
          HRLWAVLT      VQ AYA RGMYS  P AGCTIRLRAEYLWRL  c2/d
```

Myosin VIII (Plant myosins)

```
16)QIGVLEDTRNRTLHGILR  VQ SSF RGYQA  R CLLK             a  ATM1      (828-943)s33812GB
          ELKRRISI      LQ SFV RGEKI  R KEFAELR          a
          RRHKAAAT      IQ SQV KSKIA  R IQYK             a2
          GIADASVV      IQ SAI RGWLV  R RCSGDIGWLKSGGAK  a
```

Myosin IX

```
17)RQALQERLHGEVLRRILL  LQ SWF RMVLE  R RHFV             b1 myr 5, Rat (946-1057)x77609GB
          QMKHAALT      IQ ACW RSYRV  R RTL              a2
          ERTRAAVY      LQ AAW RGYLQ  R QAYH             a
          HQRHSIIR      LQ SLC RGHLQ  R RSFSQMMLEKQKAEQ  a
```

Myosin X

```
18)MAELDARRAEVLGNAARV  IR AHV LGYLA  R KQYK             c/d Frog                 D Corey PC
          KVLDCVVI      IQ KNY RAFLL  R RRFL             a2
          HLKKAAVV      FQ KQL RGQIA  R RVYRQKKAEK       a
```

Myosin XI (Plant myosins)

```
19)MAELDARRAEVLGNAARV  IQ RQF RTCMA  R KNYR             a2 Mya1     (721-884)z28389GB
          SIRNAAIV      LQ SFL RGEIA  R AVHKKL           a
          RIEAAALR      VQ KNF RRYVD  R KSFV             b1
          TTRSSTIV      LQ TGL RAMIA  R SEFRLR           a2
          RQRKAAIV      LQ AHW RGRQA  F SYYT             c2
          LRQKAAIV      TQ CAW RCRLA  R ELRMLKMAARDTGAL  a2
```

Figure 1.12 continued.

The conserved residues in the core of the IQ motif are identified in red. Residue numbers for the sequences in each protein are given in parentheses. Accession numbers for GeneBank (GB) or Uniprot/SwissProt (SP) are given thereafter (PC implies personal communication to author). The letter code in purple after each sequence relates to the article text for the figure. Species abbreviations: Ac, Acanthamoeba; ATM1, Arabidopsis thaliana myosin 1; CdbH, human cardiac beta; Ck, chicken; Dct, Dictyostelium; Dr, Drosophila; HMWMI, high molecular weight myosin I; MM, mammalian myosin; Nm, non-muscle; Nom, neuromodulin; Ng, neurogranin; Sc, scallop; Sk, skeletal; Sm, smooth; and Yst, yeast.

Table 1.1: Example IQ motif spacings in different myosin types

Number of IQ motifs	IQ spacing pattern	Myosin type
6	23-25-23-25-23	Myosin 5a, 5b (animals)
	23-25-23-25-23	Myosin 11 (plants)
	23-23-29-29-29	Myosin 1a (Myr1a, Rat)
		NB: the 29 residue motifs have non-canonical sequences.
5	23-23-23-23	Myosin 7
4	23-26-23	Myosin 8
	23-22-23	Myosin 9 (Myr5, Rat)
3	23-23	Myosin 10
	23-23	Myosin 1 (bovine adrenal)
2	22	Myosin 1 (Myr4, Rat)
	26	Myosin 2
	36	Myosin 3 (NinaC, *Drosophila*)
1	N.A	Myosin 1 (amoeba)
	N.A	Myosin 4
	N.A	Myosin 6

The IQ motif spacing pattern has important consequences for the arrangement and the interactions the light chains can make, both with the heavy chain and with neighbouring light chains. Since the IQ motifs in the myosin heavy chain are presented to the light chains in an α-helix, the number of residues separating binding sites dictates not only the axial spacing of the light chains along the helix (~1.5 Å rise per residue for a straight helix), but also the *azimuth* at which the light chains bind around the helical axis. In this way, the IQ-spacing can have implications on the three-dimensional structure of the lever at the nm-scale, and so could dictate the lever's mechanical properties as well as regulatory mechanisms.

Assuming a straight α-helix with 3.6 residues per 360° turn, each residue traversed makes a rotation of ~100° about the helical axis, thereby rotating the orientation of any bound LC by a similar amount. Where neighbouring LCs are separated by 23 amino acids (6.39 turns), a relative rotation between LCs of 140° about the helical axis is expected, while in a 25 spacing (6.94 turns), a 340° (−20°) relative rotation would result. The two spacings, though differing by only 3 Å in axial spacing along the helix, can therefore result in quite different (200° rotated) azimuthal orientations of neighbouring light chains bound to the lever. In turn, this is likely to influence the possible interactions neighbouring LCs can

make with one another. This is demonstrated in two crystal structures obtained of 23- and 25-spaced light chains binding to fragments of the lever domain of myosin 5a, discussed further in 1.8.2.1.2.2.

Chapter 5 of this thesis focuses on a mutation study, undertaken to explore the structural and/or functional significance of myosin 5's alternating IQ spacing pattern. Results of initial characterisations of mutant myosin 5 molecules with altered IQ spacing patterns are presented.

1.7 Actin

Actin is an essential protein ubiquitous in all eukaryotic cells. It forms the major component of the thin filaments in muscle cells (Holmes and Kabsch, 1991), and in non-muscle cells it forms the most abundant part of the cytoskeletal matrix. As such, it is essential for processes such as muscle contraction, maintenance and dynamics of the cell structure, formation of lamellipodia and filopodia, formation of pathways for intracellular cargo transport, as well as a host of other biological processes. Actin filaments provide the cytoskeletal 'track' that all myosins bind via their motor domain.

1.7.1 G-actin

At low (nonphysiological) ionic strength, actin exists in its monomeric form, known as globular actin (G-actin). G-actin is formed from a single 375 residue (42 kDa) polypeptide that folds to form a compact globular shape. Crystal structures of G-actin complexed with other molecules that inhibit polymerisation (such as deoxyribonuclease I or tetramethyl-rhodamine-5-maleimide (TMR)) have been solved (Kabsch et al., 1990; Otterbein et al., 2001), and show that the monomer structure comprises two principal domains, each divided into two subdomains, separated by a deep cleft into which ATP or ADP and the cations Mg^{2+} or Ca^{2+} bind, Figure 1.13. The base of the cleft acts as a hinge, while the bound nucleotide and ion act as a 'latch' that hold the two domains together. Further ions can bind at the surface of the subdomains.

1.7.2 F-actin

Addition of salt to physiological concentration polymerises G-actin into fibrous (filamentous) form (F-actin) (Holmes and Kabsch, 1991), the dominant form of the protein *in vivo*. Polymerisation involves subtle conformational changes in each of the subunit subdomains (Fujii et al., 2010; Oda et al., 2009), allowing the actin molecules to pack together to form an extended lattice. On assembly into the filament, the bound ATP is hydrolysed to ADP which remains trapped in the binding site.
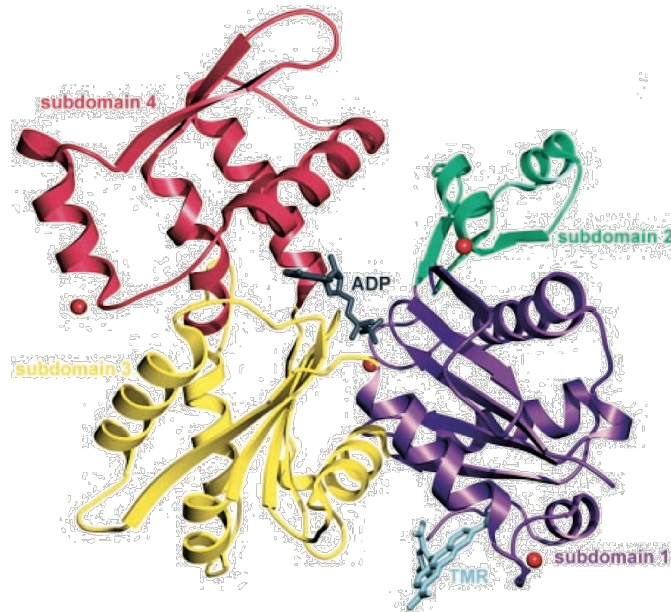
Figure 1.13: Crystal structure of G-actin with ADP and TMR bound.

The four subdomains here are shown in different colours. ADP (stick, grey) is shown bound in the cleft. TMR which inhibits polymerisation is bound to subdomain 1 (stick, light blue). Four ions are shown as red spheres, one complexed with the ADP in the cleft. Adapted from Otterbein et al. (2001).
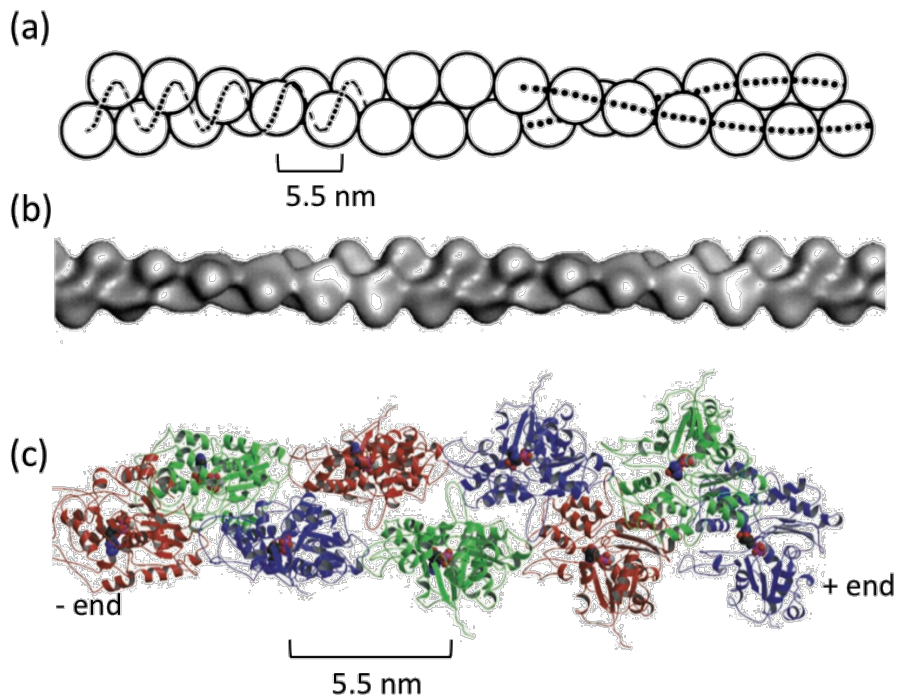
Figure 1.14: Structure of F-actin.

(a) lattice structure of F-actin with actin subunits as spheres. Adapted from O'Brien et al. (1975). The dotted lines indicate the helical paths of two interpretations of the structure. On the left, the left-handed single helix with period 26 subunits (2.75 nm rise per subunit). On the right, the right-handed double-stranded helix interpretation, with rise per strand subunit of 5.5 nm and a 13-subunit crossover length. (b) reconstruction of F-actin (McGough et al., 1997). (c) atomic model of F-actin showing polarity of actin subunits giving rise to a + and − end (Geeves and Holmes, 1999).

The lattice structure of the filament, depicted in Figure 1.14(a), was first elucidated from EM studies (Hanson and Lowy, 1963; Trinick et al., 1986). The actin subunits assemble into a left-handed helical polymer. The full period contains 26 subunits which make 12 complete turns – making the rotation per subunit −166.15° and the rise per subunit 2.75 nm (Holmes and Kabsch, 1991). Because the rotation per subunit is close to 180°, the filament can also be viewed as a two right-handed helices, each helix formed from a string of subunits that slowly wrap around one another. Projected onto the plane, the helical repeat of the two-stranded helix appears as 36 nm, as this is the distance over which the strands crossover one another, Figure 1.14(a). Each crossover is accomplished in 13 subunits (36 nm), through 6 left-handed subunit turns. The rise per subunit along each strand is 5.5 nm. The 36 nm/13-subunit crossover distance is also known as the 'pseudo helical repeat length'.

Atomic models of F-actin have been built by modelling crystal structures into structures obtained by EM (Holmes et al., 1990) and from analysis of X-ray diffraction pattern from gels of oriented filaments (Lorenz et al., 1993; Oda et al., 2009). Recently, a 6.6 Å resolution structure of the filament has been obtained by cryo-EM (Fujii et al., 2010). These structures show that the actin subunits pack into the filament lattice in only one orientation, Figure 1.14. Because of the asymmetry of each subunit, this gives the filament a structural polarity. One consequence of the polarity is that polymerisation occurs faster at one end, known as the plus (+) end, with the slower-growing end termed the minus (−) end (Neuhaus et al., 1983). A means of visualizing the polarity in EM images is through stoichiometric binding of myosin 2 S1 fragments. The F-actin becomes 'decorated' with myosin heads, their long axis making make an ~45° angle with the filament longitudinal axis, giving rise to an 'arrowhead' appearance (Huxley, 1963). The arrowheads point towards the so-called 'pointed' end of the filament and away from the 'barbed' end. The barbed end corresponds to the fast-polymerising + end, the pointed end to the slower-growing − end (Pollard and Craig, 1982). Most myosins move towards the plus end of F-actin, except for myosin 6 that moves towards the minus end (Sweeney and Houdusse, 2010; Wells et al., 1999).

The structure of the actin filament is crucial in determining the binding interaction with myosin heads. The helical arrangement of the F-actin subunits, whether in the context of thin and thick filament sliding in the muscle sarcomere, or as a cytoskeletal track in non-muscle cells, should make the myosin binding sites more or less accessible depending on the aspect of the actin monomer presented to the myosin head. Indeed, EM studies of insect flight muscle show that thick filament myosin heads bind preferentially to 'target zones' on the thin filaments (Taylor et al., 1999). Single-molecule optical trap studies of

surface-tethered myosin heads have elegantly demonstrated the same modulation in actin-binding affinity with filament position (Steffen et al., 2001; Veigel et al., 2002). The latter experiments resolve a periodicity in the actin-binding affinity of ~36 nm, equal to the F-actin crossover period, and show further that the heads preferentially bind at this spacing to 'target zones' of 3 subunit width, interacting with only one of the two strands in the actin double-helix (Steffen et al., 2001).

For processive myosins, the helical structure of F-actin has implications for the path the myosin takes along the filament. To function, the two heads of a processive myosin must be able to simultaneously bind two subunits of a filament. If the two heads can span the 36 nm crossover distance then the molecule can effectively take 'straight' steps equal to the F-actin pseudo helical repeat. Owing to its elongated lever, myosin 5a is thought to be one of the only myosins adapted for 'walking straight' in this manner, preferentially taking 13-subunit steps (though 11 and 15 subunit steps also occur) (discussed further below). If the myosin cannot span the pseudo-helical repeat, or if it takes shorter or longer steps, then it will spiral around the filament as it progresses. A slight tendency for this behaviour has been seen directly for myosin 5 and 6 *in vitro* when moving along suspended actin filaments (Ali et al., 2004; 2002). Another assay that reports on spiralling is the so-called 'twirling filament' assay, where myosins are tethered to a surface and upward-facing heads can capture and propel F-actin around. If the filament is seen to twirl as it glides, then off-axis myosin-stepping is inferred. Such twirling has been reported for myosins 2, 5 and 6 (Beausang et al., 2008; Lewis et al., 2012; Sun et al., 2007).

1.8 Myosin 5

1.8.1 Overview

Class 5 myosins are perhaps the best characterised of all the unconventional myosins. Since their identification in the early 1990s, a wealth of biochemical and biophysical studies have established a great deal of information concerning the protein's structure and function. See Hammer and Sellers (2012) and Sellers and Weisman (2007) for reviews.

Myosin 5 was first identified as a CaM-binding protein with ATPase activity in actomyosin preparations of vertebrate brain (Espindola, 1992; Larson et al., 1990; 1988) and was later determined to be a new myosin class (Espreafico et al., 1992) with sequence similarities to two unconventional myosins identified in separate genetic studies of the yeast *Saccaromyces cerevisiae MYO2* gene (Johnston et al., 1991), and the mouse *dilute* coat colour gene (Mercer et al., 1991). Genes encoding myosin 5 have subsequently been found in almost all eukaryotic cells (Odronitz and Kollmar, 2007), and often multiple

myosin 5 genes are found in a particular organism. Three myosin 5 genes are found in mammals (*myo5a*, *myo5b* and *myo5c*), which encode three separate myosin 5 subtypes; myosin 5a, myosin 5b and myosin 5c (Berg et al., 2001; Reck-Peterson et al., 2000). Of these, myosin 5a has been the most extensively studied, and is the myosin 5 type studied in this thesis. *Saccaromyces cerevisiae* has two myosin 5 genes (*MYO4* and *MYO2*), while *Drosophila melanogaster* has just one (Berg et al., 2001). Class 11 myosins found in plants are a structural homologue of myosin 5, with similar enzymatic properties as well as structural features to myosin 5a (including a 6-IQ lever) (Yamamoto, 2007). This suggests that an ancient class 5-type myosin, from which both myosin 5 and 11 derive, existed before the division of plants and animals.

Early studies of myosin 5a (Cheney, O'Shea, et al., 1993) correctly identified the key structural features of many class 5 myosins; namely that the molecule is a two-headed multimeric protein, composed of two ~216 kDa heavy chains (HCs) and twelve CaM-family LCs (six per heavy chain), Figure 1.15. Later it was found that myosin 5a also binds two '8 kDa' dynein light chains (actual MW 10 kDa per LC) as a dimer in the tail region (Benashski et al., 1997; Espindola et al., 2000). At the N-terminus, each myosin 5 HC begins with a typical myosin motor domain, that binds actin and hydrolyses ATP. Extending from the motor is the lever (known also as the 'neck' or 'regulatory' domain), formed from a single α-helix bearing six IQ motifs, to which CaM or CaM-like LCs bind. Together, the motor domain and lever are often called the 'head' (as opposed to the term labelling just the motor) and this is the nomenclature adopted in this thesis. The lever is followed by a tail region which has long stretches of predicted coiled-coil sequence (periodically interrupted by non-helical sections) that facilitate dimerisation of two HCs into the two-headed molecule (Cheney, O'Shea, et al., 1993). Each HC ends in a globular tail domain (GTD) that binds cargo. Unlike myosin 2, myosin 5 tails do not assemble to form filaments (Cheney, O'Shea, et al., 1993).

Not all class 5 myosins have the same structure and properties. For example, in the yeast *Saccaromyces cerevisiae* the class 5 myosin myo4p has a tail that forms a heterodimeric coiled-coil with the Swi5-dependent HO expression 3 protein to form a single-headed molecule (Dunn et al., 2007). The myosin is non-processive in *in vitro* single-molecule assays, but if artificially dimerised or aggregated (*e.g.* on a polymer bead) short movements can result (Hodges et al., 2008).

(a)
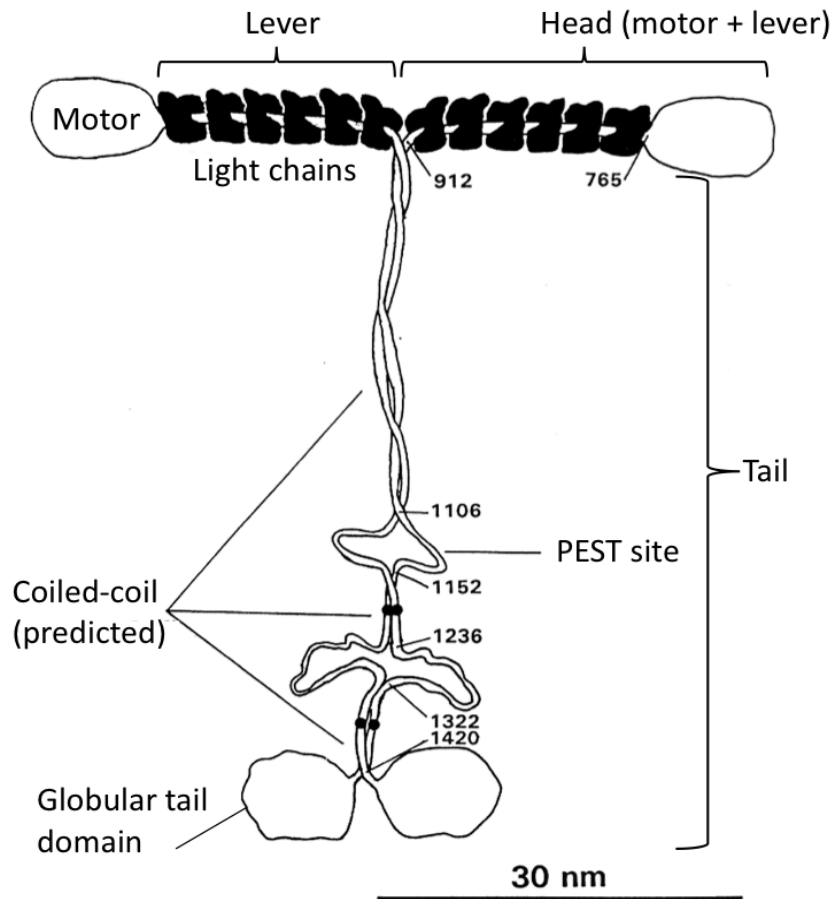
(b)
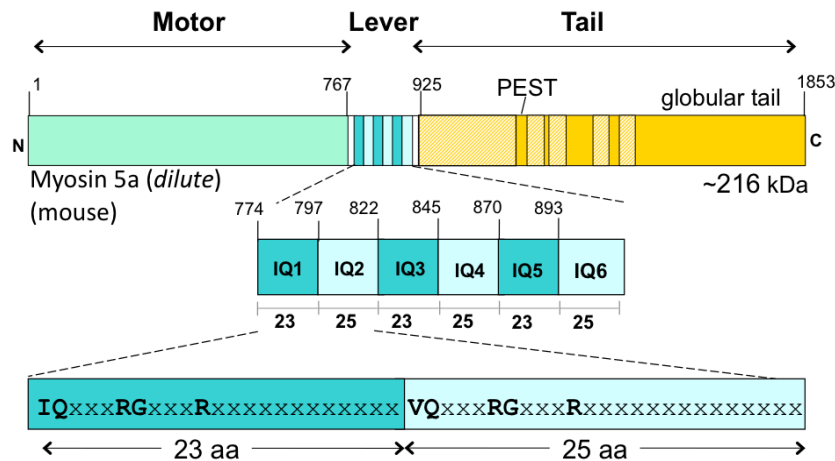
Figure 1.15: Cartoon of myosin 5a structure and domain map

(a) cartoon of two-headed myosin 5a molecule showing domain structure, adapted from Cheney, O'Shea, et al. (1993). The light chains are shown in black, two heavy chains white. *NB:* the residue number annotations apply to the chicken myosin 5a **(b)** domain map of myosin 5a mouse sequence, showing motor (green), lever (cyan, blue) and tail (yellow). Predicted coiled-coil regions in the tail are hatched. The IQ region of the lever is enlarged to show the alternating 23-25-23-25-23 residue spacing pattern.

Class 5 myosins have been implicated in a range of intracellular organelle transport and dynamic tethering across various cell types. In yeast, the Myo2p class 5 myosin transports various vesicular cargoes such as secretory vesicles, vacuoles and peroximsomes from the mother cell into the bud, while Myo4p transports specific mRNA as cargo (Hammer and Sellers, 2012; Reck-Peterson et al., 2000). In mammals, myosin 5a is particularly abundant in the neurons, forming ~0.3 % of total brain protein (Cheney, O'Shea, et al., 1993). Here, the molecule is also implicated in organelle transport. In particular, one isoform of myosin 5a has been demonstrated to transport smooth endoplasmic reticulum tubules into the dendritic spines of cerebellar Pukinje neurons (Wagner et al., 2011; Wagner and Hammer, 2003). Another myosin 5a isoform found in melanocyte cells of amphibians and mammals, captures and moves pigment granules (melanosomes) in the actin cortex at the cell periphery. Mutations in myosin 5a have been associated with neurological and pigmentary disorders in mice and humans (Reck-Peterson et al., 2000; Sellers, 1999).

Less characterisation studies have been performed on mammalian myosin 5b and 5c types. Myosin 5b has similar structural and kinetic properties to those of myosin 5a and is widely expressed in all tissues, though northern blot studies have shown that it is found chiefly in the testes, kidney, liver, lung and heart (Zhao et al., 1996). It has been found to be involved in recycling of proteins that cycle between the plasma membrane and endosomes (Sellers and Weisman, 2007). In hippocampal neurons, myosin 5b is thought to directly transport the recycling endosomes into the dendritic spine (Wang et al., 2008). Myosin 5c is mainly expressed in epilthelial cells in the pancreas, prostate, mammary, stomach, colon, and lung, where it has been implicated in trafficking of transferrin (Rodriguez and Cheney, 2002). Though thought to be structurally similar to myosin 5a and 5b it has different kinetic properties. ADP release is not rate limiting and it has a low duty-ratio (*i.e.* it spends most of its cycle in weak actin-binding state) (Takagi et al., 2008). In single-molecule assays it was found to be non-processive (Takagi et al., 2008).

1.8.2 Myosin 5a

Vertebrate myosin 5a has become one of the most extensively characterised non-muscle myosins. It was the first myosin to be shown *in vitro* to be able to move along actin filaments (Mehta et al., 1999; Sakamoto et al., 2000). Its large size and its ability to move long distances on F-actin before detaching have made it an attractive motor protein to study. Its kinetics, structure and processive mechanism have all been investigated. For reviews see Sellers and Veigel (2006) and Sellers and Weisman (2007). Over the past 20 years, these studies have established that myosin 5a is a cargo-transporting protein, capable of moving as a single molecule along F-actin, making consecutive step-like

movements driven by MgATP hydrolysis. Owing to certain structural and kinetic adaptations, myosin 5a has evolved to 'walk straight' along the 13-subunit F-actin pseudo-helical repeat, taking consecutive 'hand-over-hand' steps without detaching – a behaviour termed processive movement. The main structural and kinetic features of the molecule that enable it to walk straight and move processively are described in more detail below, with particular emphasis on the lever domain, the main focus of this thesis.

1.8.2.1 Structure

1.8.2.1.1   Motor

As in all myosin classes, myosin 5a's motor domain contains the nucleotide-binding and actin-binding sites. The motor's sequence is highly conserved and shares ~40 % identity with the myosin 2 motor (Cheney, O'Shea, et al., 1993). Crystal structures of the myosin 5a motor (Coureux et al., 2004; 2003) show that its 3D structure is also very similar to structures of myosins 1, 2 and 6, with the most notable difference being a smaller cleft between the upper and lower 50K domains, forming a structure most akin to that predicted for myosin 2 in the strong actin-binding state. This may account for the fast binding and increased affinity for F-actin when ADP.Pi is bound in myosin 5a compared to other myosins (La Cruz et al., 1999).

1.8.2.1.2   Lever

Extending from the motor and before the coiled-coil tail is the lever. Its length makes it one of the longest levers in the myosin superfamily, and it is perhaps the most distinctive feature of myosin 5a. The lever is formed from a single α-helix bearing six tandem IQ motifs, which are the binding sites of CaM or CaM-family LCs. In mouse, all the LCs are CaM (Wang et al., 2000), whereas in chicken, at least one essential LC (the 23- or 17-kDa LCs encoded by genes *LC23* and *LC17*) occupies one or more of the IQ motif binding sites (Espindola et al., 2000; Koide et al., 2006; La Cruz, Wells, et al., 2000). In yeast, the LCs are Mlc1p, another member of the CaM EF-hand family (Stevens and Davis, 1998).

LC binding to the myosin 5a lever is thought to stabilize the lever helix (its bare IQ motifs having hydrophobic regions) and increase the flexural rigidity of this extended region of the molecule – allowing it to function as a mechanical lever that can sustain load and move the molecule forward on F-actin via the powerstroke. Raising Ca^{2+} levels (to ~μM concentration) leads to dissociation of one or more CaMs from the lever (Cameron et al., 1998; Homma et al., 2000; Nguyen and Higuchi, 2005; Trybus et al., 2007), and under these conditions myosin 5a motility is abolished or significantly reduced (Krementsov, 2004; Lu et al., 2006; Nguyen and Higuchi, 2005). Though not believed to be the main

physiological mechanism of myosin 5a deactivation (which relies instead on the presence/absence of cargo, see 1.8.2.1.3), the calcium-induced loss of CaMs is thought to render the lever mechanically compliant, and thus unable to sustain load or transmit force efficiently (Tominaga et al., 2012).

1.8.2.1.2.1   *Lever length*

Being formed from six consecutive IQ motifs, at ~24 nm in length (discussed in detail below), the myosin 5a lever it is approximately three times the length of the myosin 2 lever that has only two IQ motifs. Its long lever is the key structural feature that allows myosin 5a to take steps equal to the 36 nm pseudo-helical repeat of the F-actin. For this project, establishing an accurate length for the lever is important. In the literature there is some discrepancy over this due in part to ambiguity in definition of the lever's precise start and end points. In his elastic lever-arm modelling study, Vilfan used a lever length of 26 nm which, via a diagram, is defined as a distance from *within* the motor domain (presumably including the converter region) to the head-tail junction (the point at which the coiled-coil tail begins) (Vilfan, 2005). In their optical trap study, Moore and co-workers assumed a length of 4 nm per IQ motif, implying a 24 nm lever (though it is unclear if this also includes the converter region) (Moore et al., 2004). In EM studies, Oke et al. measured head lengths from negatively-stained EM images of lever length mutants (Oke et al., 2010). From the slope of a graph of head lengths, a length of 8.75 nm was obtained for the motor domain long axis, and 3.805 nm per IQ motif, making the length of the section of lever *beyond* the motor 22.83 nm (Oke, 2004).

For this thesis, a working length for the lever was made from measurements of an atomic model. No crystals of the entire motor-lever structure have to date been obtained. However, Liu et al. built a model of the head, Figure 1.16, by fitting motor and CaM structures into a 3D density map obtained from cryo-EM tomography of myosin 5a in shutdown state (Liu et al., 2006). In this model, which follows the chicken sequence (UniProt: Q02440), the following measurements were found (using UCSF Chimera v1.4 software to measure atom-atom distances).

The lever helix is 156 residues (running G754-L909) and is 23.0 nm long (backbone C_α-C_α). It extends from the end of the converter within the motor, to Leu. 909, just beyond the sixth IQ motif, after which the EM density map is too ill-defined to model-in an α-helix. The portion of the lever helix that lies outside the motor domain (*i.e.* excluding the region near the converter) encompasses L767-L909 and is 21.4 nm (compares with Oke et al. 22.8 nm).

<u>Lever start</u>

Detailed crystallographic and simulation studies of the myosin 2 recovery stroke indicate that during ATP hydrolysis the converter and lever helix rotate together as one semi-rigid body and that this 'mechanical lever' (*i.e.* helix-plus-converter) rotates about the axis of the SH1 helix, with a fulcrum located at its C-terminus end (Fischer et al., 2005). More properly then, the entire 'mechanical lever' in myosin 5a should be considered to encompass the lever helix (plus LCs) *and* the converter region, up to the fulcrum within the motor domain. In myosin 5a, the equivalent fulcrum (the end of the SH1 helix) is located at A696 (chicken). The length from this residue to the end of the defined lever helix in the Liu model measures 23.9 nm, Figure 1.16.

<u>Lever end</u>

There is most ambiguity in defining the C-terminus end of the lever, due to the unknown structure between the end of the lever helix and the start of the coiled-coil tail. In the Liu et al. model the section between L909 and M951 is absent, presumably due to its flexible nature. The *Paircoil2* algorithm can be applied to the sequence to define a likely starting residue for the coiled-coil formation, see section 1.8.2.1.3. Requiring a high probability coiled-coil score (0.8) defines the start of the tail at I924 in chicken (M925 in mouse see Figure 1.20). There is thus a ~23 residue section of polypeptide of undefined structure between lever helix and coiled-coil tail. For this to function as part of the lever it must form a secondary structure with a flexural rigidity sufficient to transmit force and sustain load. One possibility is that the polypeptide continues the lever α-helix through to the coiled-coil tail. For a straight helix (3.6 residues per turn, ~1.5 Å rise per residue) this would add an additional 6.4 turns of helix and 3.45 nm of length to the lever. In this case this would result in a lever length of $23.9 + 3.45 = 27.35$ nm. A contrasting view is that the additional peptide length is unstructured and does not contribute to the effective length of the lever, which remains 23.9 nm. For this thesis, the latter view is preferred, given that the Liu et al. cryo-EM tomographic density map indicates that the lever α-helix is broken after L909. Such a section of unstructured polypeptide may in fact be important myosin 5a function, allowing an unattached head to freely rotate about the lever axis in searching for an actin binding site (Dunn and Spudich, 2007).
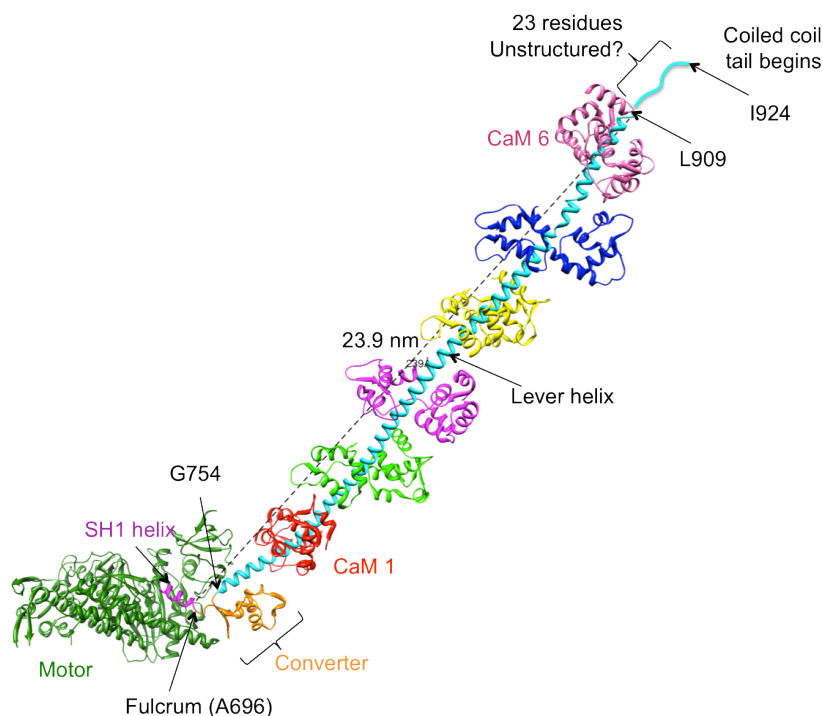
Figure 1.16: Atomic model of myosin 5a head encompassing motor domain and lever.

The model shown is that in Liu et al. (2006). The lever helix (cyan) emerges from within the motor domain (dark green). The converter region (orange) is formally part of the motor domain. It rotates along with the lever as a single semi-rigid body during the powerstroke. The SH1 helix within the motor is highlighted (magenta). The lever-converter rotate about the axis of the SH1 helix during the working stroke. The end of the SH1 helix can be considered a fulcrum for the entire 'mechanical lever' formed by the converter, lever helix and six CaMs. The straight line distance from the fulcrum to the end of the lever helix (A696-L909) (dotted line) is 23.9 nm.

Table 1.2: Summary of different lengths related to the myosin 5a lever

Measurements made on Liu et al. (2006) atomic model, using UCSF Chimera v1.4. For this study, the preferred length of the mechanical lever in myosin 5a is 23.9 nm (bold).

Region	Residues (chicken)	Length (nm)
Lever α-helix	G754-L909	23.0
Lever α-helix outside motor domain	L767-L909	21.4
Fulcrum to end of lever α-helix	**A696-L909**	**23.9**
Fulcrum to start of coiled coil tail (assuming continuous α-helix)	A696-I924	27.4

1.8.2.1.2.2   IQ motif spacing pattern

Inspection of the amino acid sequence in the lever section reveals that the consecutive IQ motif cores are separated from each other by 23 or 25 amino acids. The IQ-core spacing alternates with the pattern 23-25-23-25-23, so that IQ cores 1-2, 3-4 an 5-6 are separated by 23 residues, while cores 2-3 and 4-5 are separated by 25. Although the amino acids in

each motif vary, Figure 1.17, the 23-25-23-25-23 spacing is conserved across all myosin 5a and 5b sequences, and in the plant homologue myosin 11.
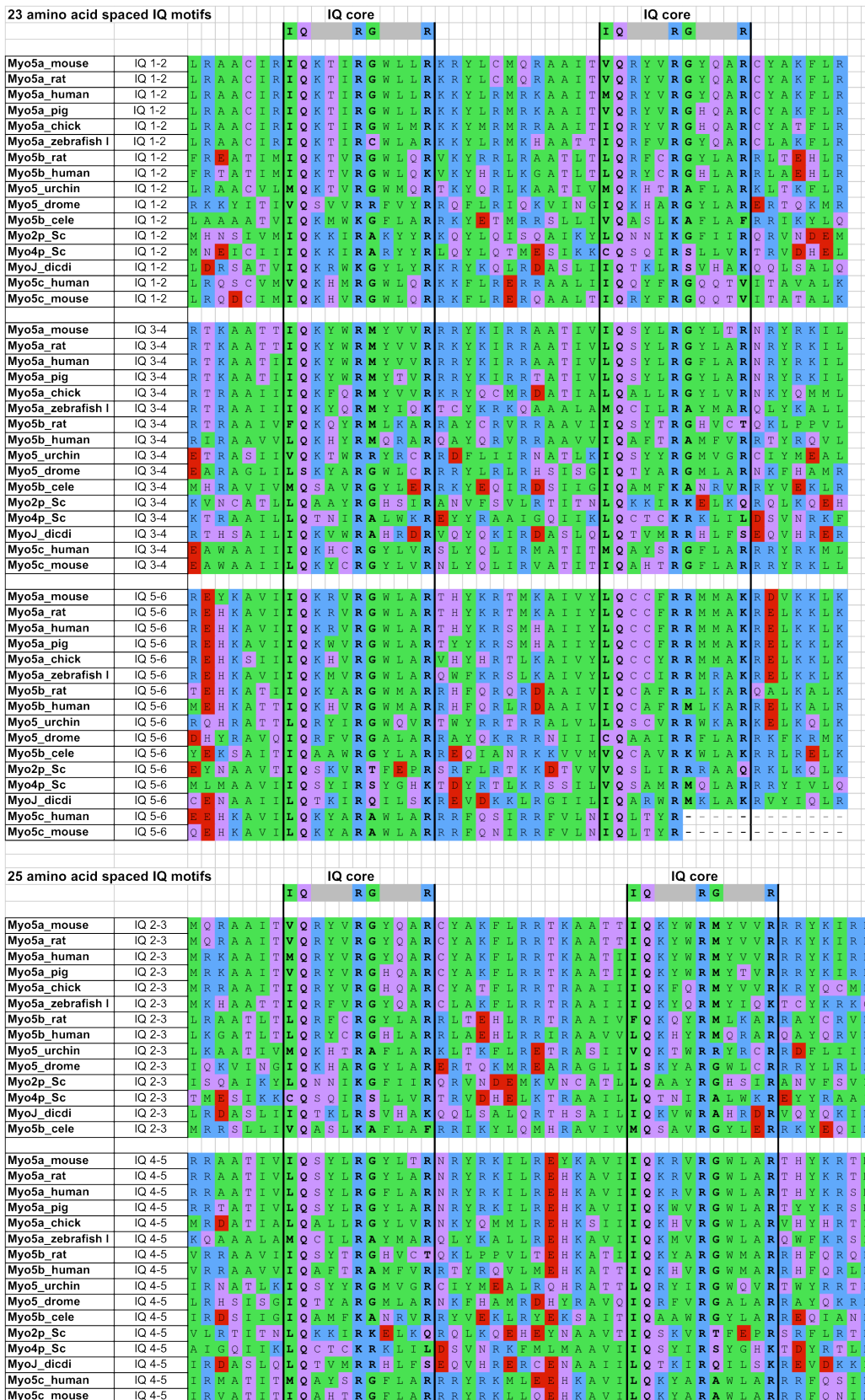
Figure 1.17: Sequence alignments of IQ motifs in selected myosin 5 sequences.

Colours indicate amino acid properties. Green = hydrophobic, pink = polar, blue = positively charged, red = negatively charged.

Since other myosin classes have different IQ-spacing patterns (see Table 1.1), the question as to the structural and/or functional importance of myosin 5a's conserved IQ-spacing is raised. This question has, to date, not been systematically studied. There have been many studies examining the influence of truncating or extending the lever by deletion or addition of IQ motifs (Moore et al., 2004; Purcell et al., 2002; Sakamoto et al., 2005; 2003), but there is only one precedent for a construct where an IQ spacing was altered. In their lever-length studies, Sakamoto et al. (2005; 2003) characterised a separate mutant, 'HMM-2Ala-6IQ', with two alanine residues inserted between IQs 3-4 - turning a 23-spaced IQ-pair into a 25-spaced one. Further EM and kinetic characterisations of this lever mutant was then made by Oke et al. (2010).

Characterisation of the HMM-2Ala-6IQ mutant across these studies showed a number of differences with a wild-type control. Firstly, its working distance as measured by optical trapping was approximately half that of wild-type (Sakamoto et al., 2003), suggesting a shortening of the effective lever length under load. The step-size distribution measured by fluorescent imaging was broader than wild-type (Sakamoto et al., 2005), though the EM studies did not reproduce this result. Rather, in EM image averages, the mutant appeared to bind F-actin with its motors spaced similarly to wild-type, at mostly 13 then 11 actin subunits apart, though the mutant's distribution was slightly more skewed towards the 11 subunit spacing (Oke et al., 2010). The EM analysis also showed that the levers in the heads of the mutant molecules were intact, had a similar length to wild-type, and apparently bound all twelve CaMs. Stopped-flow kinetic analysis showed that the ADP dissociation rate from the mutant's lead head increased slightly over wild-type (to 0.04 s^{-1} from 0.015 s^{-1}) (Oke et al., 2010). In wild-type, lead head ADP dissociation is ~30-fold reduced compared to the trail head (Forgacs et al., 2008), due, it is thought, to the lead lever occupying a restrained pre-powerstroke state when both heads are bound to F-actin. The increase in lead head ADP dissociation rate in the HMM-2Ala-6IQ mutant is therefore suggestive of increased flexibility (*i.e.* reduced capacity to sustain intramolecular strain) in its lever. Taken together, the results suggested that the mutated lever was unbroken and bound all its CaMs, but that it was mechanically weakened by the double residue insertion halfway along its length.

The idea that a 25-spaced structure might be mechanically weaker than a 23-spaced one is also consistent with a hypothesis based on crystal structures. In structures of 23- and 25-spaced lever fragments obtained for mouse and yeast myosin 5a respectively (Houdusse et al., 2006; Terrak et al., 2005), differing relative azimuthal orientations between adjacent LCs are seen, Figure 1.18. Relative angles of rotation about the lever axis between LCs measure ~93° and ~−48° for the 23- and 25-spaced structures respectively. It has been

suggested that in the 23-spaced case, residues in adjacent N- and C-lobes of the CaMs are favourably positioned to form stabilising non-covalent interactions (Houdusse et al., 2006), perhaps stiffening this section of lever, while in the 25-spaced case, the orientation and extra axial displacement along the helix preclude such interactions and may create more flexible lever regions, Figure 1.18 (Craig and Linke, 2009; Terrak et al., 2005).
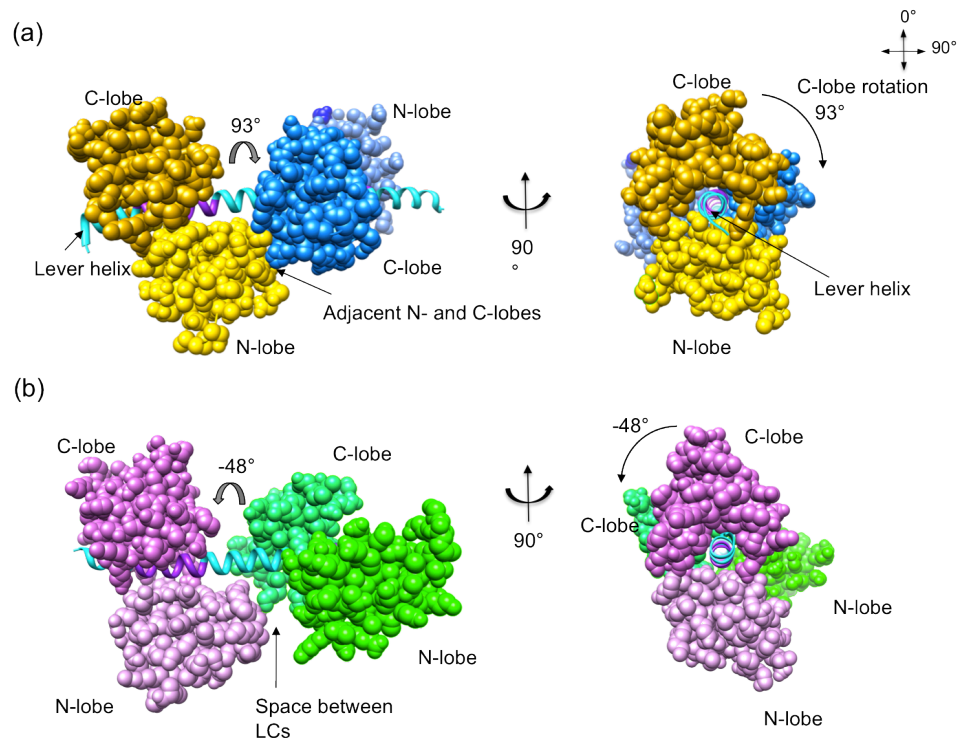
Figure 1.18: Crystal structures of different myosin 5 lever sections with light chains bound to fragments of lever helix.

(a) side and end-on views of 23 residue-spaced CaMs for IQs 1-2 of mouse myosin 5a (Houdusse et al., 2006). **(b)** similar views of 25 residue-spaced LCs (Mlc1p) binding lever helix of yeast Myo2p (Terrak et al., 2005). Lever helix is displayed in ribbon view (cyan) with IQ-cores highlighted in purple. LCs are shown in spacefill in various colours. In the 23-spaced arrangement a ~93° rotation of the CaMs places adjacent N- and C-lobes in proximity, potentially stiffening the lever by stabilising non-covalent interactions or by steric clash when the lever bends. In the 25-spaced structure the LCs make a −48° relative rotation angle about the helix axis and are not as favourably positioned for interaction or to clash if the lever helix bends.

Whether or not an alternating rigid/soft lever is created by the 23-25 IQ-spacing in myosin 5a, inspection of the different orientations of the light chains pairs in Figure 1.18 suggests that the 23-spaced and 25-spaced lever sections would likely permit different degrees of bending in different directions, due to the differing steric clashes neighbouring light chains would make. The 25-spaced structure appears as if it would be freer for bending (either in the plane of the page or perpendicular to it) than the 23-spaced structure, which appears inhibited for bending in certain directions by the proximity of adjacent N- and C-lobes of the neighbouring CaMs.

However, some caution must be exercised in over interpreting the structures in Figure 1.18. Since they are obtained from crystals (in crystallisation conditions and under the constraints of crystal packing), they may not reflect the true orientations the CaMs adopt in free molecules in physiological conditions, or indeed when the lever bends under load. It is also possible that there is mobility in the CaM positioning in the molecules, which may differ depending on conditions. Indeed, a cryo-EM and fluorescence spectroscopy study has shown that in the presence of calcium there can be significant rearrangement of the positions of bound CaMs (with propagated changes in the motor domain) (Trybus et al., 2007).

An exploration of the bending flexibility of the myosin 5a lever and its relation to the IQ configuration is one of the main focuses of this thesis. Chapter 4 undertakes an analysis of the shapes of levers subject to thermal bending, while chapter 5 continues the work begun by Sakamoto et al. and Oke et al. with the HMM-2Ala-6IQ construct. For the latter chapter, a series of HMM-like constructs were made with IQ spacings in the lever changed progressively to become 'All-23' or 'All-25'. Chapter 3 in this thesis examines the issue of CaM rotational freedom, and the implications for torsion in the lever.

1.8.2.1.2.3   Lever bending

Bending and distortion of the lever is thought to play an important part in myosin 5a's processive movement cycle. In EM studies (Burgess et al., 2002; Oke et al., 2010; Walker et al., 2000), lead levers in prepowerstroke state (and in rare cases postpowerstroke state) were seen to bend backwards along their length, or flex at motor-lever junction (presumably by melting of the helix at a pliant point located there), Figure 1.19. Striking high-speed AFM video imaging has also observed lead levers fluctuating between straight and curved conformations (Kodera et al., 2010).

Backward lead lever-bending is attributed to trail head-tethering under the geometrical constraints of the doubly-bound molecule that preferentially binds F-actin at a motor-motor spacing of 13 actin subunits (Burgess et al., 2002). The strained lead lever is believed to be prevented from undergoing its powerstroke by the trail head tethering, and this has in turn been posited as the chief mechanism by which ADP release is inhibited in the lead head compared to the trail head (Forgacs et al., 2008; Rosenfeld and Sweeney, 2004; Sakamoto et al., 2008; Veigel et al., 2002). Kinetic studies indicate that ADP release in the lead head is inhibited by a factor of ~30 when tethered by the trail head (Forgacs et al., 2008; Oke et al., 2010; Sakamoto et al., 2008). This kinetic 'gating' between the heads helps keep the molecule continuously attached to F-actin during walking, coordinates

hand-over-hand stepping (the trail head is more likely to detach first from F-actin) and thus limit futile step cycles.
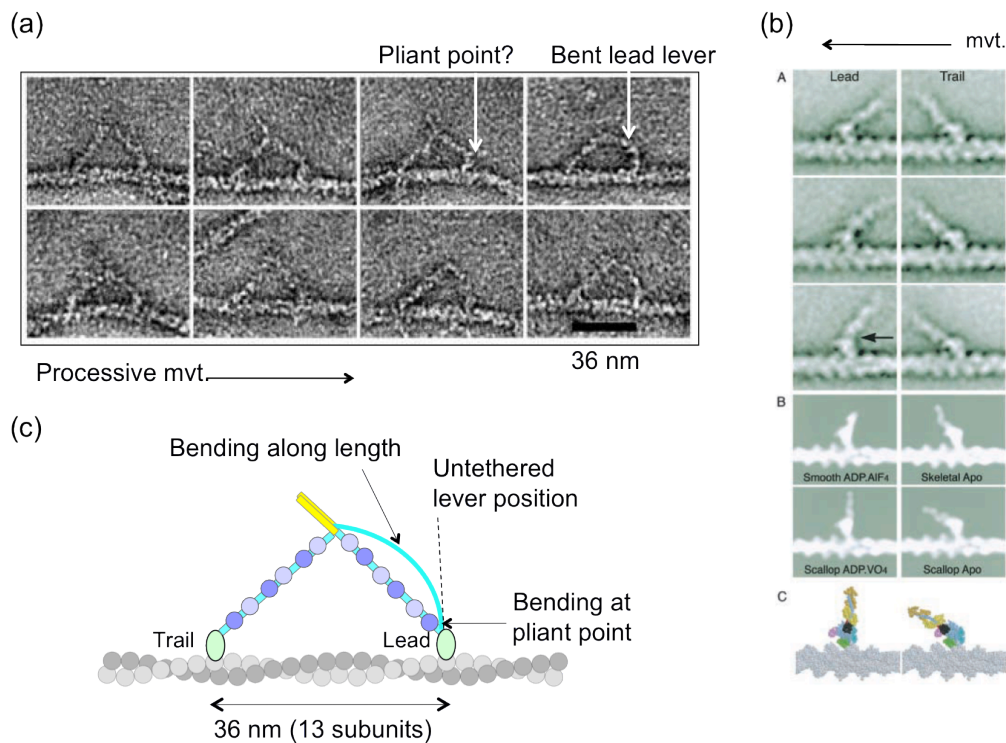
Figure 1.19: EM evidence of lever bending

(a) Negative stain EM images of myosin 5a HMM molecules bound by both heads to F-actin. From Walker et al. (2000). Some lead levers clearly curve along their length. Others are straight. (b) comparison of (A) image averages of lead and trail heads of myosin 5a HMM bound to F-actin and (B) the profiles of atomic models of scallop myosin 2 structures (C). Note the difference in the position of the lead head compared to the averages. (c) cartoon of myosin 5a HMM showing the two types of lever deformation (bowing or pliant point bending) inferred from the images. Deformation in the lead head is attributed to trail head tethering.

Thus, in myosin 5a, it is clear that the lever must not only be rigid enough to move as a semi-rigid body and sustain load in its role as a lever, but also that in certain circumstances it must be able to deform to occupy the functionally important strained states.

1.8.2.1.3   Tail

The myosin 5a tail follows the lever and performs a number of functions. The main extended part of the tail dimerises the two HCs via long sections of coiled-coil-forming sequence interspersed with non-helical regions, Figure 1.15(a) and Figure 1.20. The non-helical regions may be important in determining the extensibility of the tail, which must function as an elastic tether between the cargo and head domains, as the motor moves through a crowded and viscous cytoplasm (Nagy et al., 2009; Schilstra and Martin, 2006). The HC is differentially spliced at three exons located in the tail, giving rise to specific isoforms that localise to different tissue types and/or cellular locations (Sellers and

Weisman, 2007; Seperack et al., 1995). A PEST sequence, a proteolytic site rich in proline, glutamic acid, serine and threonine has also been identified in one of the non-helical sections of the tail (Espreafico et al., 1992), Figure 1.15. Chicken and mouse brain myosin 5a isoforms also co-purify with two '8 kDa' dynein light chains (actual MW 10 kDa), which have been shown to bind, possibly as a dimer, C-terminus to the PEST site in the tail (Benashshki et al., 1997; Espindola et al., 2000).

Locating the start of the coiled-coil tail is important, as it delineates the end of the lever. Using the sequence analysis algorithm *Paircoil2* (Massachusetts Institute of Technology, Cambridge, USA) (McDonnell et al., 2006), a coiled-coil-forming probability can be calculated per residue. Applying the algorithm to the mouse sequence, and requiring a minimum probability of 0.8, identifies the start of the coiled-coil at Met. 925 (score 0.98) (or Ile 924 in chicken sequence, score 0.90), Figure 1.20.

(a)

(b)

Residue		Coiled-coil register	Probability score
900	M	g	0
901	A	a	0
902	K	b	0
903	R	c	0
904	E	d	0
905	L	e	0
906	K	f	0
907	K	g	0
908	L	a	0
909	K	b	0
910	I	c	0
911	E	d	0.11588
912	A	e	0.22313
913	R	f	0.28004
914	S	g	0.28004
915	V	a	0.34405
916	E	b	0.34405
917	R	c	0.34405
918	Y	d	0.34405
919	K	e	0.34405
920	K	f	0.34405
921	L	g	0.34405
922	H	a	0.34405
923	I	b	0.34405
924	G	c	0.72285
925	M	d	0.98382
926	E	e	0.98382
927	N	f	0.98382
928	K	g	0.98382
929	I	a	0.98382
930	M	b	0.98382

Figure 1.20: Paircoil coiled-coil probablities for mouse myosin 5a sequence.

(a) plot of *Paircoil2* coiled-coil probability scores (0-1) by residue number for the mouse myosin 5a sequence. The dashed line shows the 0.8 probability cut-off that identifies Met 925 as the first certain coiled-coil residue. **(b)** per residue probabilities for coiled-coil formation in the 900-930 residue region. Values start to increase at Glu 911, but do not rise above the 0.8 threshold until Met 925.

For this thesis, which uses the mouse sequence, the Met. 925 position is preferred as the defined start of the coiled-coil tail. However, the choice of cut-off is interpretative. Other studies place the start of the coiled-coil earlier, *e.g.* at residue 912 (chicken) (Cheney, O'Shea, et al., 1993). In fact, if the cut-off is relaxed further, the hydrophobic seam of the coiled coil can be traced back upstream as far as Leu 908 (mouse). However, it is unlikely

that the coiled-coil would run this far upstream as the sixth CaM of each head would then be too crowded. More likely, there is some non-helical 'breathing space' sequence between the end of the lever helix and the start of the coiled-coil tail. Moreover, this may actually be important for the molecule in allowing the head-tail junction to act as a universal joint, allowing a detached head to freely explore the actin filament for its next binding site or to switch to a new filament (Dunn and Spudich, 2007). Adopting the M925 residue start position and the 0.8 coiled-coil cut-off threshold, the calculated length of total coiled-coil in the tail (excluding intervening non-helical regions) is 41 nm (Nagy et al., 2009). The section of coiled-coil up to the first non-helical section is predicted to be ~30 nm (Baboolal et al., 2009; Oke et al., 2010; Rodriguez and Cheney, 2002; Sakamoto et al., 2005; 2003).

At the C-terminus of the myosin 5a tail is the globular tail domain which performs two main functions; it binds cargo (via various adaptor proteins), and, in the absence of cargo, it binds to the motor, folding the molecule into a stable inhibited state (Liu et al., 2006; Thirumurugan et al., 2006; Wang et al., 2004). Structures of the molecule in its folded shutdown state were obtained by separate negative stain and cryo-electron tomography studies (Liu et al., 2006; Thirumurugan et al., 2006). These revealed not only details of the folded-back tail, but also the position/profile of the motor domain and lever, Figure 1.21.
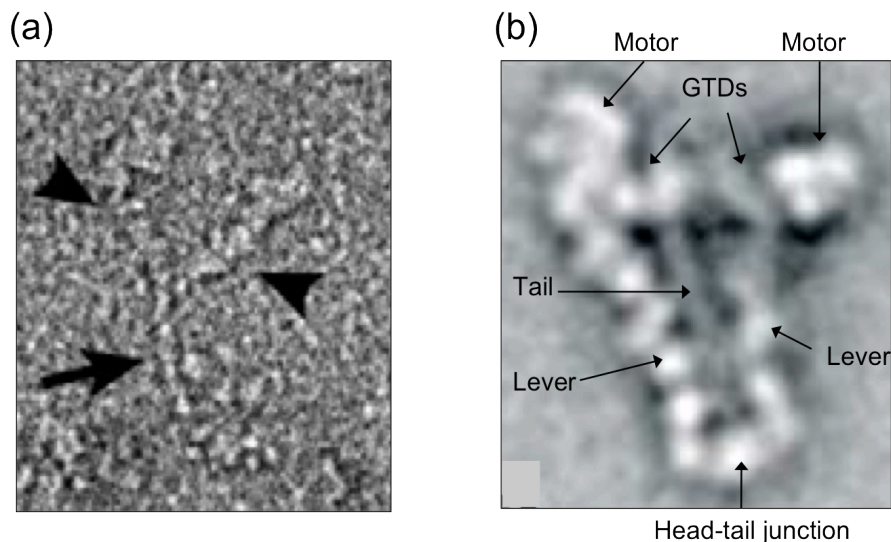
Figure 1.21: Negative stain EM images of full length myosin 5a molecules in open and folded conformations.

(a) negative stain image of a single full length myosin 5a molecule, opened out in high salt conditions such that the coiled-coil tail (arrow) and heads (arrowheads) are seen. From Sellers et al. (2008). **(b)** image average of negative stain images of full length molecule in compact shutdown conformation. The two motor domains are the globular regions top left and right in the image with the two levers extending downwards to a vertex (six individual CaMs are seen in the left lever). The tail folds back between the two levers such that the GTDs contact the motors. From Thirumurugan et al. (2006).

1.8.2.2 Kinetics

Myosin 5a's kinetic cycle of actin-activated ATP hydrolysis follows the same pathway as most myosins, 1.5.4. Heads cycle between states with a low affinity for F-actin (M.ATP and M.ADP.P$_i$) and high affinity for F-actin (M.ADP, M.Apo). The key distinguishing feature of myosin 5a's enzymatic cycle compared to muscle myosin 2 is that the rate limiting step is ADP release (12-16 s^{-1}) from the actomyosin complex (A.M.ADP) (La Cruz et al., 1999). All other steps in the cycle, including ATP hydrolysis and P$_i$ release are fast (La Cruz et al., 1999). These adaptations mean that a single myosin 5a head spends most (>70%) of its kinetic cycle strongly bound to F-actin, in the A.M.ADP state, which is important for processivity. Furthermore, the kinetic cycles of the two heads in a molecule appear to be altered by conformation and intramolecular strain, according to whether a head is a trail head in a post-powerstroke position, or is a lead head held in a (strained) pre-powerstroke position. In saturating ATP conditions it was found that ADP release is strongly (~30 fold) inhibited in an attached lead head compared to an attached trail head (Forgacs et al., 2008; Oke et al., 2010; Purcell et al., 2005; Rosenfeld and Sweeney, 2004; Sakamoto et al., 2008; Veigel et al., 2005). EM images of myosin 5a bound to F-actin suggest that the lead head cannot easily transition from pre- to post-powerstroke while it is tethered by the attached trail head (Burgess et al., 2002; Oke et al., 2010; Walker et al., 2000) (see Figure 1.19). These effects combine to ensure that when the molecule is 'waiting' on F-actin (both heads in A.M.ADP state) the trail head is the most likely of the two to release ADP, quickly bind ATP and detach from F-actin. The attached lead head is then freed to undergo its powerstroke, moving the detached trail head forward along the actin filament. Since ATP hydrolysis and F-actin reattachment are fast, the unattached trail head is also likely to undergo its recovery stroke and rebind strongly to F-actin (now in the lead position) before the other attached head (formerly the lead head) loses ADP and detaches.

1.8.2.3 Processive stepping cycle

From the evidence accumulated from many biochemical, structural and single-molecule studies, a consensus for myosin 5a's mechanochemical cycle and processive stepping mechanism have emerged. Remaining issues still debated chiefly concern the exact nucleotide states and actin affinities of the two heads when bound to F-actin. The main events in the consensus cycle are depicted in Figure 1.22 and outlined in the text thereafter, with references to the main evidence supporting each event included.
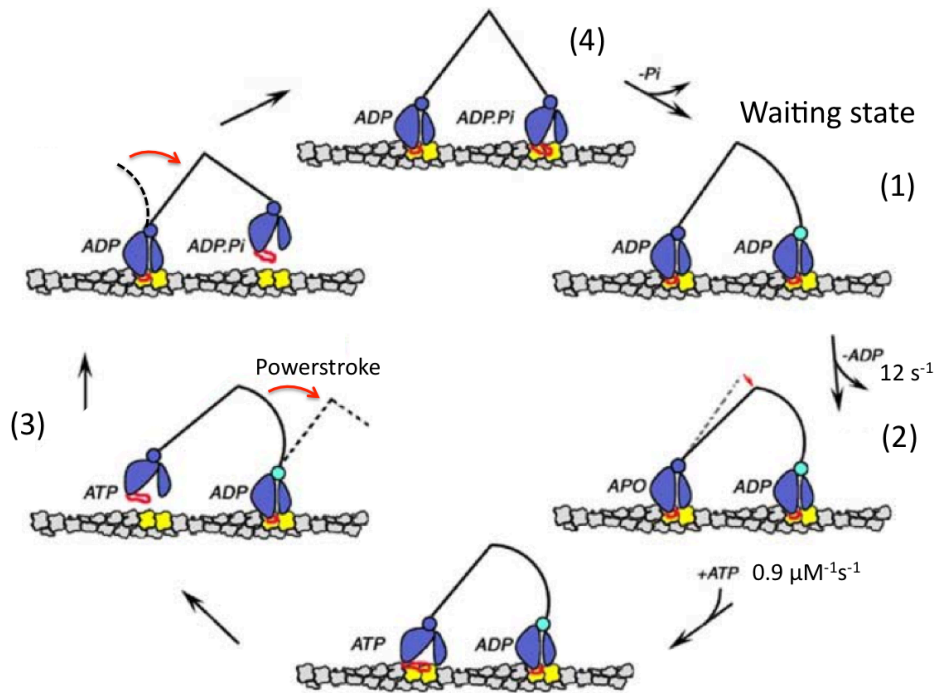
Figure 1.22: Consensus model for myosin 5a processive stepping cycle

For simplicity only the two heads of a myosin 5a molecule are depicted, the motor domain in blue, stylized as the upper and lower 50K domains split by the actin-binding cleft which opens and closes depending on nucleotide state. The levers are shown as lines, without bound light chains. The converter is shown as a small circle at the top of the motor. Intramolecular strain is depicted by curved lever shapes and a light blue converter. Loop 2 from the motor, important in actin binding, is shown in red (see reference for more information). The F-actin subunits are depicted in grey with the favoured binding site subunits 13 subunits apart in yellow. The reference for this figure includes a discussion of myosin 5a-F-actin interactions extending over two actin subunits, hence two are highlighted at each F-actin crossover point. Figure adapted from Volkmann et al. (2005).

(1) Two heads bound in 'waiting' state:

Optical trap and gold label tracking studies suggest that at saturating ATP, myosin 5a dwells in a state in which both heads are strongly bound to F-actin (A.M.ADP or A.M.Apo) (Dunn and Spudich, 2007; Veigel et al., 2002). Kinetic studies are consistent with this and suggest that this dwelling or 'waiting' state is likely dominated by both heads having ADP bound (A.M.ADP) (Forgacs et al., 2008; Rosenfeld and Sweeney, 2004). In this state, the two heads are likely under strain due to the mismatch between the measured working stroke, ~20-25 nm (Dunn and Spudich, 2007; Moore et al., 2001; Sakamoto et al., 2003; Veigel et al., 2002), and the molecule's preference for binding at a head-head spacing of 36 nm (Oke et al., 2010; Walker et al., 2000).

The lever emergence position in EM images indicates that, in waiting state, the lead head is usually held in pre-powerstroke conformation (due to the constraints of trail head tethering), while the trail head is in post-powerstroke conformation (Burgess et al., 2002; Oke et al., 2010). Furthermore, the tethered lead head appears pulled further 'backwards'

48

along the actin filament compared to the orientation a free pre-powerstroke head would adopt (Burgess et al., 2002). The strain in the lead head manifests itself either as a curvature along the lever's length, or as flexing at the motor lever junction (Burgess et al., 2002; Oke et al., 2010; Walker et al., 2000).

The intramolecular strain and conformation of the heads determine the kinetics of ADP dissociation and thus the likely next step in the cycle. ADP release is inhibited ~30 fold in the lead head over the trail head (Forgacs et al., 2008; Oke et al., 2010; Sakamoto et al., 2008).

(2) Trail head ADP dissociation:

ADP dissociation from the trail head is therefore the most likely next event. Optical trap measurements and cryo-EM structures indicate that this may be coupled to an additional 'kick' (forward rotation) of the trail head lever, advancing the head-tail junction by ~5 nm (Veigel et al., 2005; 2002; Volkmann et al., 2005). This sub-step may be associated with ADP release or may be an isomerisation that precedes it. The movement would serve to further increase the deformation/strain in levers of both heads.

(3) Trail head detachment, lead head powerstroke:

Following ADP dissociation from the trail head, ATP binding, head detachment and ATP hydrolysis occur quickly in the trail head (La Cruz et al., 1999). EM images indicate that coupled to the hydrolysis, a free head motor rotates relative to its lever by ~90°, executing the recovery (priming) stroke (Burgess et al., 2002).

Concurrent to the trail head detachment, the attached lead head, now unrestrained, can relieve any elastic strain energy from trail head tethering (which may create an initial forward displacement) (Dunn and Spudich, 2007) and execute its powerstroke. In the powerstroke, the lead lever rotates (in a forward direction along the actin filament) through an angle in the range ~60-105°, as evidenced by EM, structural and various polarized TIRF microscopy studies (Burgess et al., 2002; Coureux et al., 2004; Forkey et al., 2003; Lewis et al., 2012; Sun and Goldman, 2011; Syed et al., 2006; Volkmann et al., 2005; Walker et al., 2000). The attached head powerstroke moves the head tail junction ~20-25 nm forward, carrying the detached free head with it.

(4) Diffusive search by detached head, reattachment to F-actin:

Evidence from EM, optical trapping, TIRF microscopy and gold label tracking suggests that the detached head preferably rebinds to F-actin at a distance of 36nm or 13 actin subunits in front of the attached head, with some variation to 11 or 15 subunits also possible (Cappello et al., 2007; Dunn and Spudich, 2007; Mehta et al., 1999; Moore et al.,

2004; Oke et al., 2010; Purcell et al., 2002; Rief et al., 2000; Sakamoto et al., 2005; Veigel et al., 2002; Walker et al., 2000). To achieve rebinding at the 36 nm step size, the free head must undergo a thermally-driven diffusive search, since the powerstroke distance alone is insufficient to bring the free head to the next binding site. The dynamics of this diffusive search have been characterised by dark field imaging of a gold particle-labelled molecules (Dunn and Spudich, 2007), and 'travelling wave tracking' using an optical trap (Cappello et al., 2007). That the free head prefers to rebind F-actin at the 13th subunit, at the same azimuth as the other head, or at 11 or 15 subunits distances at neighbouring ±28° azimuthual orientations (Oke et al., 2010; Walker et al., 2000), suggests that the powerstroke's main component is directed along the actin filament and/or that there is a strong energetic cost (high stiffness) associated with bending the lever perpendicular to the filament axis to find more azimuthally distant binding sites.

That it is definitely the detached head of the two that rebinds to F-actin in the new lead position is confirmed by TIRF microscopy assays capable of tracking differently labelled heads of individual molecules (Churchman et al., 2005; Warshaw et al., 2005). The new lead head likely rebinds to F-actin first in the weak $ADP.P_i$ nucleotide state, before P_i is released and the motor achieves strong F-actin binding. The new trail head is assumed to remain in the A.M.ADP (or A.M.Apo) state while the new lead head achieves strong binding to F-actin.

Thus, a stepping cycle is completed. In a 'hand-over-hand' fashion, the trail head has detached F-actin, been propelled forward by the lead head powerstroke and, following a diffusive search, rebound F-actin in a new forward position. One ATP molecule is hydrolysed and released in the process (Sakamoto et al., 2008). Most frequently, the trail head steps to the 13th subunit along from the lead head position, maintaining the same azimuthal orientation around the filament. In a step, the net forward movement of the trail head is 72 nm, while the head-tail junction (and attached cargo) advance 36 nm.

The steps described above represent only the dominant pathway myosin 5a follows in a stepping cycle. Since the cycle is governed by the probabilities of nucleotide binding, conformational changes and diffusive motion, many (less probable) variations are possible. One example is a futile cycle, where a lead head releases ADP and F-actin before the trail head has rebound, such that the myosin 5a molecule becomes detached from the actin filament.

1.8.2.4 Importance of the myosin 5a lever

In outlining the structural and kinetic features of myosin 5a, and describing the stepping cycle, the important mechanical role played by the lever has been highlighted. The lever is important to myosin 5a function because:

- It acts as a mechanical amplification element, transforming the ångström-sized nucleotide-dependent movements in the motor domain into a ~20-25 nm movement of its tip.

- Its 6-IQ/24 nm length allows the two heads of myosin 5a to span and make steps equal to the 13-subunit F-actin crossover distance.

- It has sufficient flexural rigidity that it can rotate as a semi-rigid body capable of sustaining load required for both cargo transport and movement of a detached head.

- It has sufficiently elasticity to deform (bending along its length or deforming at pliant regions) to allow the molecule to distort when bound to F-actin at its preferred head-head spacing. The distortion is posited to play a key role in regulating the kinetics of the two motors, biasing the molecule towards processive unidirectional movement.

Allied to its important mechanical function, the myosin 5a lever also has a unique structure among myosins, comprising six LC-binding IQ motifs, with the conserved spacing pattern of 23-25-23-25-23 residues between IQ-cores. The IQ-core spacing dictates the azimuthal orientation as well as axial spacing of LCs bound to the lever, and thus may influence the mechanical properties of the lever and therefore the ability of the motor protein to walk straight along the F-actin pseudo-helical repeat.

Myosin 5a's long levers, large powerstroke, strong binding to F-actin for the majority of its mechano-chemical cycle, and its ability to move processively, make it especially amenable to study by various forms of microscopy (EM, TIRF, AFM etc.), as well as to biochemical and single-molecule characterisation. For this reason, and because of the specific structure and functional importance of the lever, myosin 5a was chosen as the model system for this thesis. Moreover, gaining an understanding of the relationship between structure and function of myosin 5a is important given the protein's role in many important biological processes, and because it may answer more general questions about the structure, mechanical properties and function of other lever types across the myosin family.

1.9   Transmission electron microscopy

In this PhD project, a central technique used was transmission electron microscopy of protein samples stained with heavy metal salt solutions, a technique referred to as negative stain EM. The technique has proved to be a powerful tool for the analysis of biological specimens and offers particular advantages for the study of biological macromolecules with continuously flexible domains, or which exhibit interdomain flexing (Burgess, Walker, Thirumurugan, et al., 2004; Tskhovrebova et al., 2010). It has also been particularly successful in imaging individual myosin molecules and their complexes with actin (Burgess et al., 1997; Song et al., 2010; Walker et al., 1991; 1985; Walker and Trinick, 1989; Walker et al., 2000; Yang et al., 2009).

Although the resolution of negative stain EM is lower than that obtainable in X-ray crystallography, it offers the advantage that crystals do not need to be formed, so different samples can be prepared and analysed relatively rapidly. Also, unlike in crystallography, where only the conformation(s) favourable to crystallisation are selected, negative stain EM can image whole populations of molecules in a range (sometimes a continuum) of conformations. In this way, EM (both negative stain and cryogenic techniques) can provide a better representation of the full dynamic range of a protein's conformations.

In the transmission electron microscope, a beam of accelerated electrons is directed under vacuum via a series of electromagnetic lenses at a sample on an EM grid, typically a metal mesh supporting a thin film substrate to which sample molecules are adsorbed. Electrons in the beam can undergo elastic and inelastic scattering events as they interact with the atoms in the sample. Contrast is formed by different amounts of electron scatter from different regions of the sample. Electrons scattered through high angles are excluded by a lens aperture. Phase-contrast (interference between scattered an unscattered waves) is the other contrast mechanism which can be enhanced by controlled defocus.

Application of a heavy metal solution, such as uranyl acetate, to a sample grid can enhance image contrast and provide specimen stabilisation. Typically, stain solution is applied to the sample on the EM grid. As the stain dries it forms a thin layer of crystallites across the grid. In areas of optimum stain depth, substrate-adsorbed molecules protrude through and exclude the stain, which tends to accumulate around the edge of the molecule and, to an extent, enters into any cavities in the structure. The stain can also act as a support and protect the molecules from collapse in the microscope vacuum, though distortions from the adsorption and grid drying processes are still significant concerns (see 1.9.1 below). The heavy metal atoms of the stain scatter electrons much more strongly than the light atoms of

biological molecules, and hence the stain envelope around the molecules is imaged with high contrast. The molecule image is thus formed from the absence of stain, hence the term 'negative stain'. In the case of uranyl acetate stain, which has one of the smallest grain sizes, the size of the stain crystallites limits the achievable resolution to ~2 nm (Ohi et al., 2004).

1.9.1 Potential artefacts of negative stain EM

Although negative stain EM offers the advantages of rapid sample preparation and high contrast imaging of heterogeneous molecule populations, it is also important to be aware of both the technique's limitations and the processes involved that have the potential to produce artefacts. These processes, discussed below, can be chemical or physical in nature.

Chemically induced artefacts concern the possible interactions between the elements of the stain solution and the sample, and also of changes to the concentration of buffer components and stain as the grid dries. Firstly, stain solutions of heavy metal salts such as uranyl formide or uranyl acetate are acidic (Hayat, 2000). For example, the 1% (w/v) aqueous uranyl acetate solution as prepared for negative staining in this PhD, is unbuffered and has a pH of 4.2-4.5. Clearly, when mixed with a protein sample, a solution with such a low pH has the potential to be detrimental to molecules that are unstable in acidic conditions, or to induce non-physiological changes in structure and appearance. To an extent, these problems may sometimes be mitigated by the rapid (< 10 ms timescale) fixative properties of uranyl acetate, which have been demonstrated (in particular for actin and myosin 2 filaments) to rapidly stabilise ionic and hydrophobic interactions before collapse or significant conformational change can occur (Zhao and Craig, 2003).

A second potential chemical artefact concerns the pattern of stain deposition around and within cavities of the protein. Though uranyl acetate in aqueous solution, in the pH range used in negative staining, exists mostly in un-ionised form (*i.e.* the uranyl ion is mainly associated with the acetate ion), some charged uranyl ions are present (Hayat, 2000). In fact, these ions exist as a series of complexes of both anionic and cationic form and are thus capable of binding to both negatively and positively charged side chains of a protein. The relative binding of the different species is unpredictable and depends on stain concentration, pH, duration of staining and other components present in the solution (Hayat, 2000). The presence of particular charged regions on a protein molecule can thus influence the particular deposition pattern of stain, and, as a result, areas of a given molecule may be particularly stain-attracting or stain-excluding depending purely on the pH and particular type of stain used.

A third type of chemical artefact concerns the potential behaviour of the sample as the stain and buffer solution dries on the grid. As the aqueous phase evaporates, so the stain and salt concentrations around the surface-adsorbed protein molecules dramatically increase (theoretically to infinity!), presumably bringing about a large change in the ionic strength. Again, the rapid fixative properties of uranyl acetate may help stabilise proteins against related conformational changes or collapse (Zhao and Craig, 2003), but clearly the potential for this form of chemically-induced artefact should still be borne in mind.

Physical mechanisms that can create artefacts in EM and the staining process include the deleterious effects of radiation damage from the electron beam, surface tension forces acting on molecules during the drying process, interactions between the substrate and adsorbing molecules and variability in imaging depending on stain depth.

Radiation damage to a sample as it is exposed to a beam of accelerated electrons is the primary limiting factor in gaining high resolution images of biological macromolecules in transmission EM. The chief mechanism of radiation damage arises from ionisation or excitation of specimen atoms caused by inelastic scattering of incident electrons with orbital electrons in the sample. Ions or radicals that result from these scattering events in turn lead to bond ruptures and other molecular damage, as well as the formation of new cross-links between fragments as they recombine. Since these molecular rearrangements occur very rapidly and at a much lower electron dose (< 500 e$^-$/nm^2) than is required for high contrast imaging, the actual chemical composition of the protein sample imaged is probably significantly altered from the original molecule, and has likely reached an amorphous quasi steady state, whereby further incident energy is being dissipated non-destructively (Unwin, 1974). The degree to which this steady state retains the morphology of the originally adsorbed macromolecule determines the accuracy with which the EM image depicts the protein structure. Again, the presence of a supporting stain may help to prevent gross morphological changes from occurring, perhaps by constraining the protein in its original conformation while extensive cross-linking can occur (Zhao and Craig, 2003).

Similarly, promotion of interactions between a protein and the grid substrate, *e.g.* by pre-treating carbon with UV radiation to increase hydrophilicity, has been shown to be beneficial in stabilising certain proteins (particularly myosins – as individual molecules or filaments), against collapse or damage from the stain itself (Knight and Trinick, 1984; Trinick and Elliott, 1979; Trinick, 1981; Walker et al., 1985). However, strong molecule-substrate interactions can also bias the orientation of adsorbing molecules and induce shape distortions, particularly in elongated and flexible macromolecules. Adsorption of

protein molecules (including various myosins) in preferred orientations on EM grids, has long been reported in negative stain EM studies (Burgess et al., 2002; 2003; Knight and Trinick, 1984; Ohi et al., 2004; Ungewickell and Branton, 1981; Walker et al., 1985). This tendency presumably arises due to favourable interactions between charged and/or hydrophilic surfaces on one side the protein interacting with complementary charged/hydrophilic regions of the substrate. Related to this, is the phenomenon of molecule-substrate interaction effecting the overall shape of adsorbed molecules. This is particularly relevant to long string-like molecules such as titin or vimentin intermediate filaments (Mücke et al., 2004; Tskhovrebova and Trinick, 2001). If the molecule-substrate interaction strength is weak ($\leq$ thermal energy) molecules can more readily re-equilibrate in 2D as they adsorb to the surface (Tskhovrebova and Trinick, 2001). In contrast, if the interaction strength is significantly greater than thermal energy, such that the molecule binds irreversibly to the surface at one or more contact points (from where adsorption is then propagated), then molecules can become trapped in a 'frustrated' conformations. In the latter scenario, trapped molecules appear more condensed with more pronounced curvatures and kinks than the 2D re-equilibrated case (Mücke et al., 2009). In 'ideal' trapping, the molecule's un-equilibrated conformation can be modelled as a projection of the molecule's 3D conformation onto the plane of the substrate (Frontali et al., 1979; Mücke et al., 2009; Rivetti and Guthold, 1996). In reality, an intermediate regime between 2D re-equilibration and ideal trapping may apply and this can vary significantly depending on the substrate material, any preparatory treatments and the type of sample being examined (Mücke et al., 2009; 2004). A further example of a structural difference induced by substrate interaction is the flared structure and preservation of the projecting heads of negatively stained myosin 2 thick filaments. In this case, only when the carbon substrate is pre-treated with UV radiation to make it highly hydrophilic, are the myosin heads visible and splayed outwards from the filament backbone when stained, likely because of strong interactions with the hydrophilic carbon film prior to addition of the stain (Knight and Trinick, 1984).

While the effects of radiation damage described above would argue for use of as low an electron dose as possible, use of moderate to high dose has, somewhat surprisingly, been shown to *increase* clarity and contrast in certain cases (Walker et al., 1991; 1985). An explanation for this counter-intuitive phenomenon is that a higher electron dose can cause active stain redistribution on the grid. Unwin (1974) showed that, when irradiated, stain can initially behave in a liquid-like manner, with stain pockets apparently contracting about their centres of mass, as if with surface tension properties. By increasing the electron dose, *e.g.* to $\sim 10^5$ e^-/nm^2, studies have shown that this stain redistribution in the electron

beam can significantly increase the clarity of certain fine structures not normally visible under lower dose conditions, *e.g.* the coiled-coil tail of myosin 2 (Walker et al., 1991; 1985). Here, it is supposed that, in the electron beam, an initially even layer of stain retracts and progressively crystallizes around the perimeter of surface-adsorbed structures. There is thus a balance between limiting the deleterious effects of radiation damage and increasing contrast due to beneficial stain migration. In this PhD, the experiences of the EM group at the University of Leeds were built upon, and higher dose (~50,000 e^-/nm^2) imaging was preferred for the negative stain EM undertaken (Burgess, Walker, Thirumurugan, et al., 2004).

Distortion or collapse of macromolecules due to surface tension forces acting during grid preparation is another source for potential artefacts. The potential for macromolecules to appear flattened or elongated compared to molecular mass expectations and crystal structures is well known in negative stain EM studies (Boisset et al., 1990; Cheng et al., 2006; Knight and Trinick, 1984; Radermacher et al., 1994). Different processes associated with the drying buffer/stain solution can cause this. Firstly, in some studies, more commonly in metal shadowing performed on a mica substrate, e.g. Tskhovrebova (2001), distortion or elongation of macromolecules due to patchy dewetting and surface tension forces associated with receding droplet menisci, can align, elongate and distort fields of molecules – a process termed 'molecular combing' (Bensimon et al., 1994; 1995; Tskhovrebova and Trinick, 2001; 1997). More generally, as the stain solution evaporates across the entire grid and the surface-adsorbed molecules break the liquid surface, surface tension forces with components parallel and perpendicular to (towards) the substrate must act on the molecules (*cf.* the downward force produced by surface tension on a semi-submerged Wilhelmy plate). The magnitude of each force component will depend on the contact angle the liquid makes with the molecule (*i.e.* the 'wettability' of the molecule) and the surface tension at the air-liquid interface of the meniscus formed around the molecule's perimeter. Previous estimates of the forces associated with menisci formed around DNA and titin cross sections are relatively high, of the order of hundreds of pN (Bensimon et al., 1995; Tskhovrebova and Trinick, 2001; 1997), and so may well account for the kinds of flattening or structural collapse artefacts seen frequently in negative stain studies.

Finally, variation in stain depth can be added to the list of processes that can influence the appearance of negatively stained macromolecules. The degree to which molecules protrude through the layer of stain influences the shape of and extent to which the stain accumulates around the molecule. For this reason, it is important to be consistent in

collecting images in areas of similar stain depth. A medium depth (usually found in areas exhibiting a gradation in stain depth) is found to be optimal, since the accumulation of stain around the molecule perimeter is maximised, leading to highest contrast imaging. If the stain is too deep, the background is darker and there is little or no molecule outlining. If too shallow, the molecules are easily damaged by irradiation in the electron beam (Burgess, Walker, Thirumurugan, et al., 2004).

Given the above descriptions of the harsh chemical environment, radiation damage and the potential distorting forces associated with the grid drying, it is perhaps surprising that negative stain EM can produce any useful structural information on biological macromolecules. However, as the high degree of felicity between many negative stain EM structures and corresponding crystal structures testifies (*e.g.* the myosin 2 S1 structure by EM in Burgess et al. (1997) *cf.* that in Rayment, Rypniewski, et al. (1993), or the dynein structure by EM in Roberts et al. (2009) *cf.* that of Kon et al. (2011)), the technique nevertheless often produces excellent results. This is especially the case when negative stain EM is allied with image alignment and classification methods (Ohi et al., 2004), techniques that are discussed in the following section.

1.10 Image processing

A powerful technique that has been used in conjunction with EM and is used in this study is single-particle image processing (Frank, 2006). This technique involves the computational alignment and averaging of digital images of individual macromolecules from micrographs. By averaging, the random noise inherent in the individual images cancel out, while consistent structural features reinforce and the signal to noise ratio is greatly improved.

In image alignment, a cross-correlation function is used to find the best resemblance of one image to another as one is subjected to iterative translation and rotation. To remove bias from aligning noisy images with a predetermined reference, a reference-free alignment strategy can be employed (Penczek et al., 1992). In this strategy, two images are randomly picked, aligned and averaged. A third image is then picked and aligned to the average of the first two and all three are averaged, and so on until all images have been aligned. The first image is then aligned to the total average minus itself, the second to the total average minus itself, and similarly until there is no further change in the rotation or translation parameters.

Image classification is a method of separating the aligned images into homogeneous groups with similar features. The image groups are then averaged to produce class

averages. Computationally, this separation has been approached by multivariate statistical analysis (van Heel, 1984). An image of n pixels can be represented as a single point in an n-dimensional hyperspace (the point specified by an n-dimensional vector). Similar images will form clouds of such points in the hyperspace. Classification divides the cloud of points into appropriate clusters for averaging. One algorithm used is hierarchical ascendant classification (van Heel, 1984), another is K-means classification, which iteratively partitions the images based on a number of random seed images (Frank, 2006). For this study, both techniques were tested, but the K-means method was eventually adopted. Consistent with experience in the EM group at University of Leeds this method was found to perform better for the analysis of images of continuously flexible molecules (Burgess et al., 1997). The continuous variation in such images smears out the clouds in the hyperspace. Unlike hierarchical ascendant classification which can produce classes of widely varying size, the K-means algorithm subdivides these smeared clusters into subgroups of roughly similar size which can be more readily interpreted (Frank, 1990).

Chapter 2: General Materials and Methods

This is a general materials and methods chapter which only includes information on materials and protocols common to more than one of the results chapters. Methods information for procedures specific to a particular chapter are located in the chapter concerned. In particular, all information concerning the sequences, cloning, expression and purification of heavy meromyosin (HMM)-like myosin 5a constructs are given in chapter 5. Likewise, information on the construction, expression and purification of calmodulin (CaM) mutants for labelling studies are found in chapter 6. Detailed information on the use of *SPIDER* image processing software and the particular methods used to align and classify electron microscopy (EM) images are given in chapters 3 and 4 as appropriate.

All laboratory work carried out by the author of this thesis was undertaken at University of Leeds (UK) or at the Laboratory of Molecular Physiology (LMP) (Dr James Sellers group), National Institutes of Health (NIH) (Bethesda, MA, USA). Contributions to laboratory work and results obtained from co-workers are credited throughout as appropriate.

2.1 Materials

All chemicals used were analytical grade and, unless otherwise stated, were obtained from Sigma-Aldrich (UK or USA). For laboratory work at University of Leeds, the water source used for all buffers was from a Milli-Q purification system (Millipore UK Ltd.) which produces water with a resistivity of 18.2 MΩ·cm at 25 ºC, by successive filtration, adsorption and deionisation steps. For creation of Lennox L Broth (LB) (Invitrogen – Life Technologies, Grand Island, USA) for use with bacterial cells water from an Elix Purification System (Millipore UK Ltd.), with resistivity >5 MΩ•cm at 25 ºC was used, but this was always autoclaved before use. At LMP-NIH, bottled analytical grade purified and deionised water was used for use with DNA, and filtered water supplied to the laboratory was used for buffers.

2.2 Actin preparation

2.2.1 G-actin stocks

All G-actin sources used in this PhD were purified from rabbit skeletal muscle. At University of Leeds, G-actin was purified by Dr Gavin Allsop (University of Leeds, UK) from dried acetone powder following the method of Pardee and Spudich (1982). A ~200

µM G-actin solution was produced in 0.2 mM $CaCl_2$, 1.0 mM DTT, 0.2 mM ATP, 2.0 mM Tris-HCl, 0.5 mM Azide, pH 8.0 at 0 °C. 25 µL droplets were then drop-frozen in liquid nitrogen for storage. At LMP-NIH, G-actin stocks were prepared by Fang Zhang using similar methods and either drop-frozen for storage or kept at 4 °C in solution for use directly.

NB: to avoid inaccuracies in measuring and dispensing viscous G- or F-actin solutions, positive-displacement piston-operated pipettes (or alternatively a standard Gilson Pipetman with a cut-off tip) were used whenever possible.

2.2.2 Polymerisation of F-actin from G-actin

F-actin was polymerised from G-actin stocks using a method devised by Prof. Peter Knight (University of Leeds, UK). Frozen G-actin was transferred from liquid nitrogen storage to a tube and quickly hand-thawed before placing it on ice. For each 10 µL of G-actin stock, 1 µL of exchange buffer (3 mM $MgCl_2$, 11 mM EGTA, pH 7.0) was added and the mixture left on ice for 5 mins. This step is to allow Ca^{2+} ions in the G-actin to exchange with Mg^{2+} ions. This was followed by addition of 1 µL of polymerisation buffer (300 mM KCl, 12 mM $MgCl_2$, 1 mM EGTA, 120 mM MOPS, pH 7.0) per 10 µL of original G-actin stock, and the final mixture left on ice for at least one hour. The exchange-and-polymerisation protocol dilutes an initial concentration of 200 µM G-actin to 167 µM, with a final buffer composition of 25 mM KCl, 0.1 mM, 10 mM MOPS from the exchange and polymerisation buffers, plus a ×10/12 dilution of the G-actin storage buffer components.

If required, the F-actin mixture obtained from exchange-and-polymerisation was then dialysed for at least 16 hours against the buffer required for following work, typically against 1 L volumes with one buffer change after 2-4 hours. This dialysis step was deployed as a means of removing trace ATP from the F-actin mixture prior to further experiments.

The final F-actin subunit concentration was determined by spectrophotometric absorbance measurement, 2.6.2.

2.2.3 Preparation of biotinylated rhodamine-phalloidin F-actin

At LMP-NIH, 10% biotinylated rhodamine-phalloidin actin filaments (RhPh-BFA) were prepared for use in dual-colour total internal reflection fluorescence (TIRF) microscopy assays (chapters 5 and 6). A 20 µM stock of 10% biotinylated F-actin (BFA) was made by mixing 26 µL of 192 µM stock G-actin with 20 µg of biotinylated G-actin (Cytoskeleton Inc., Denver, CO, USA) dissolved in 24 µL of (filtered and deionised) H_2O. 200 µL of

polymerisation mix (50 mM KCl, 25 mM MOPS (pH 7.5), 2.5 mM $MgCl_2$, 10 mM NaN_3, 1 mM DTT, 2 mM ATP) was added to the G-actin/Biotinylated G-actin mix and the solution left overnight on ice. The following day, the BFA was labelled with tetramethylrhodamine (TRITC)-phalloidin (RhPh) by mixing 50 μL of 20 μM BFA with 5 μL of 200 μM TRITC-phalloidin (Invitrogen – Life Technologies) and 45 μL of motility buffer (20 mM MOPS (pH 7.4), 5 mM $MgCl_2$, 0.1 mM EGTA). This gave a final stock of 100 μL of 10 μM RhPh-BFA for use in experiments. The stock was covered in foil and left on ice for future use (usually within 1 week).

2.3   Calmodulin

Wild-type CaM, following the amino acid sequence for *Xenopus laevis* (UniprotKB accession: P62155), was kindly supplied by Dr Melanie Colegrave (formerly University of Leeds, UK) or by Dr Attila Nagy (LMP-NIH). The protein had been expressed in *E. coli* and purified using the calcium-dependent affinity chromatography method described in chapter 6. CaM stocks (typically ~100-200 μM) were available as frozen droplets or aliquots from a liquid nitrogen dewar or -80 °C freezer. Samples were thawed in small batches as required.

2.4   Protein dialysis

At University of Leeds, proteins were dialysed using D-tube™ dialysing tubes (Mini, Midi or Maxi sizes) (Novagen® – Merck KGaA, Darnstadt, Germany), while at LMP-NIH, dialysis cassettes (Thermo Fischer Scientific Inc., Glen Burnie, MA, USA) were used. Dialysis was routinely performed in a cold room at 4 °C, with continual stirring using a magnetic stir bar.

2.5   Polyacrylamide gel electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was carried out using standard methods (Laemmli, 1970).

For myosin 5a-HMM samples (chapter 5), 1 mm thick Tris-Glycine gels with of a 4-20% polyacrylamide gradient were used (Novex, Life Technologies). The use of the gradient allows resolution of both the myosin 5a heavy chain and the lower MW CaM. For specific details on myosin 5a-HMM SDS-PAGE see 5.3.5.1. For CaM only samples (chapter 6), 15% polyacrylamide gels (with 4% stacking gel sections) were made and a

staining/destaining protocol for small peptides was followed, see 6.3.4 for full details on CaM SDS-PAGE.

2.6   Protein concentration measurements

Concentrations of all proteins were determined by absorbance measurement using a Cary 50 Bio UV-VIS spectrophotometer (Agilent Technologies (formerly Varian Inc.), Santa Clara, CA, USA) applying the Beer-Lambert law (with corrections for light scattering). A cleaned quartz micro-cuvette of 1 cm path length was used and the spectrophotometer was first blanked with appropriate buffer to obtain a baseline spectrum, before zero-ing again just prior to reading sample absorbance. 70-100 μL protein sample (typically a stock dilution in appropriate buffer) was placed in the cleaned microcuvette, and an absorbance spectrum across the 200-600 nm range recorded.

2.6.1   Myosin 5a-HMM concentrations

Concentration estimates for all myosin 5a-HMM constructs were calculated from the 280 nm absorbance (A_{280}) minus 1.5×320 nm absorbance (A_{320}) (the latter a correction for light scattering) using a molar extinction coefficient per HMM head $\varepsilon_{\text{hmm-head},280}$ of 120,000 $M^{-1}cm^{-1}$, a number derived from the number of tryptophan and tyrosine residues in the HMM heavy-chain and six CaM sequences and their molar extinction coefficients at 280 nm (Forgacs et al., 2009). The HMM concentration (*i.e.* number of two-headed molecules) was then taken as $1/2 \times$ the head concentration:

$$[\text{myosin5a HMM (M)}] = \frac{A_{280} - (1.5 \cdot A_{320})}{2 \cdot \varepsilon_{\text{hmm-head},280} \cdot l},$$

$$\text{(Eq 2.1)}$$

where l is the path length in cm.

2.6.2   Actin concentrations

Concentration of actin (either G-actin or F-actin forms) was determined by measurement of 290 nm (A_{290}) and 320 nm (A_{320}) absorbance. The actin concentration (in mg/mL) was calculated using the 290 nm extinction coefficient, $\varepsilon_{act} = 0.63$ $(\text{mg/mL})^{-1}$ cm^{-1} (Lehrer, 1972) and the formula:

$$[\text{Actin (mg/mL)}] = \frac{A_{290} - (1.3 \cdot A_{320})}{\varepsilon_{act,290} \cdot l},$$

$$\text{(Eq 2.2)}$$

where l is the cuvette path length in cm. The A_{320} correction is for light scattering (Houk and Ue, 1974). The molar concentration of actin was calculated from the mg/mL value using the MW of 42 kDa.

2.6.3 CaM concentrations

CaM concentrations were determined from 276 nm absorbance, using a molar extinction coefficient of 3006 $M^{-1}cm^{-1}$ (Wallace et al., 1983). Corrections for light scattering were not made.

2.7 Actin-activated MgATPase assay

The steady-state MgATPase of different myosin 5a constructs were determined in the absence and presence of varying concentrations of F-actin, using an NADH-coupled assay. The assay works by monitoring the change in 340 nm absorbance that is due to the oxidation of NADH, which is linked to two coupled enzymatic reactions (La Cruz, 2009; Trentham et al., 1972). In the reaction system, Figure 2.1, the enzyme under study (here myosin 5a) hydrolyses a molecule of ATP and releases ADP and inorganic phosphate (P_i). ADP is required for the conversion of phospho(enol)pyruvate (PEP) to pyruvate, by the pyruvate kinase (PK). In turn, NADH is required for the conversion of pyruvate to lactate by lactate dehydrogenase (LDH). In the latter step NADH is oxidized to NAD+. Hence, one molecule of NADH is oxidized per ADP molecule produced by the ATPase enzyme. Since NADH absorbs strongly at 340 nm while NAD+ does not, the rate of reduction in 340 nm absorbance can be measured and converted to an ATPase rate. The NADH-coupled assay has the advantage over other methods that ATP is regenerated from the liberated ADP, as part of the PK reaction, so maintaining a constant ATP concentration and avoiding a build up of ADP, which is known to inhibit myosin 5a activity (La Cruz, Sweeney, et al., 2000).
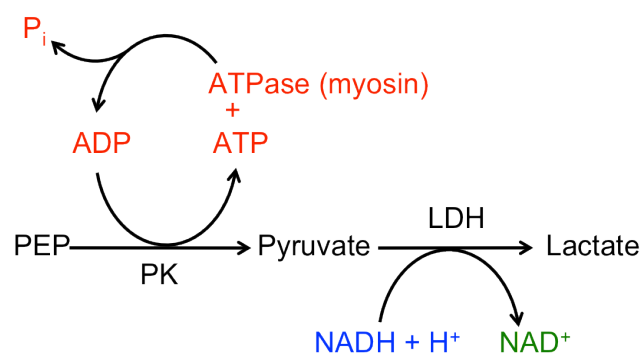
Figure 2.1: Reaction scheme of NADH-coupled ATPase assay

Assays were conducted by mixing myosin 5a protein (to ~40 nM head concentration) with various concentrations of F-actin (0-20 μM) in a 200 μL final assay mixture containing 10 mM MOPS (pH 7.0), 2 mM $MgCl_2$, 1 mM ATP, 50 mM KCl, 0.15 mM EGTA, 40 units/mL l-lactate dehydrogenase, 200 units/mL pyruvate kinase, 200 μM NADH, and 1 mM phospho(enol)pyruvate. The reaction mixture was first mixed on the bench in an

eppendorf tube, and activated by the addition of the myosin 5a last (so as to prevent build up of crosslinked F-actin aggregates), before quickly transferring the mixture to a quartz micro-cuvette (1 cm path length). The cuvette was then placed in a Cary 50 Bio UV-VIS spectrophotometer (Agilent Technologies) set-up to monitor 340 nm absorbance (A_{340}) over time, and equipped with a water circulator which (in later assays) was used to maintain a 25 °C temperature. Actin-only controls (*i.e.* myosin replaced by its storage buffer) were usually performed for each assay, and myosin-only (actin replaced by its storage buffer) were taken to calculate basal ATPase rates.

For each actin concentration, graphs of the falling A_{340} values were fit with a straight line. The gradient of the fitted line and the NADH extinction coefficient at 340 nm, $\varepsilon_{NADH,340} = 6220$ $M^{-1}cm^{-1}$, were used to calculate an ATPase rate V ($s^{-1}head^{-1}$) from the formula:

$$V = \frac{-[\nabla(A_{AM,340}) - \nabla(A_{A,340})]}{\varepsilon_{NADH,340} \cdot l \cdot [\text{myosin 5a}]},$$

(Eq 2.3)

where $\nabla(A_{AM,340})$ and $\nabla(A_{A,340})$ are the gradients (in a.u/s^{-1}) of the acto-myosin-ATP and actin-only A_{340} plots (*i.e.* the actin control subtracted as a rate correction), l is the path length (in cm) and [myosin 5a] the myosin 5a head concentration. Plots of the rates obtained were made in *OriginPro* (v8.6) graphing software and the parameters V_{max} (maximum rate) and K_{ATPase} (actin concentration at half maximum rate) were obtained by nonlinear least squares fitting of the data to the Michaelis-Menten equation $V = V_{max}[\text{actin}]/(K_{ATPase} + [\text{actin}])$.

2.8 Negative stain electron microscopy

In the negative stain EM method used in this thesis, protein molecules were allowed to adsorb from solution to a thin carbon film supported by a copper mesh grid. The heavy metal stain used to enhance contrast and support the molecules was 1% (w/v) unbuffered uranyl acetate in aqueous solution. Stain was briefly applied to the grid, before excess liquid was removed and the grid dried. The molecules adsorbed to the grid protrude through the thin layer of dried stain and, to some extent, the stain enters cavities in the molecule and accumulates around its edge forming a fringe. In so doing the stain also acts to support the molecules on the surface. Since incident electrons are scattered much more readily from the heavy metal atoms compared with the weakly interacting atoms of most biological macromolecules, image contrast is formed by the absence of stain, hence the term 'negative stain'.

2.8.1 Dilution Buffer

Unless otherwise noted, the standard dilution buffer used to dilute protein samples for EM in this study was: 25 mM KCl, 10 mM MOPS, 2 mM EGTA, 1 mM $MgCl_2$, 2 mM K-phosphate, pH 7.0 at 20 °C.

2.8.2 Preparation of carbon-coated copper grids

Carbon films were prepared by resistive evaporation of a 1.0 mm diameter carbon fibre (Agar Scientific, UK) onto freshly cleaved mica sheets (Agar Scientific). Evaporation was performed under vacuum (~10^{-6} mbar) in an Edwards E306A coating unit (Edwards, Crawley, UK). The carbon fibre was fixed between two electrodes and raised ~17 cm above the mica sheets to yield a carbon film of ~5-10 nm on the mica surface on evaporation of the fibre (Burgess, Walker, Thirumurugan, et al., 2004). Evaporation was initiated by passing a current rapidly through the fibre causing it to rapidly heat and evaporate. Carbon-coated mica sheets were removed from the coating unit and stored in a covered petri dish before transfer of the carbon (usually within 1 week of production) to 400 mesh copper grids (Agar Scientific).

Transfer of the carbon film to the grids was achieved by floating the carbon in a bath of milliQ H_2O and raising the grids arranged on a steel mesh support from under the water surface to collect the carbon film as it was raised out of the bath. Specifically, a freshly cleaned stainless steel mesh was placed on a support just beneath the water surface. The support was attached to a clamp stand through a rack and pinion mechanism, which allowed the height of the platform to be smoothly adjusted. The water surface was cleaned by dragging optical lens paper (96049, Ted Pella Inc., CA, USA) across the surface. Copper EM grids were dipped once in acetone and twice in milliQ H_2O and carefully placed, shiny side upwards, on the submerged mesh platform. The carbon-coated mica was then slowly introduced to the water bath, carbon side up, at an angle of ~45 °, such that carbon film floated free of the mica substrate which sank to the bottom of the bath. The grids were then gently raised using the rack and pinion through the surface of the water to collect the floating carbon film. The carbon-coated grids were then dried at ambient temperature in a semi-covered petri dish for at least 12 hrs before being ready for use.

2.8.3 Ultraviolet light treatment of carbon-coated grids

Prior to any staining work, carbon-coated EM grids were irradiated with ultraviolet (UV) light for 30-40 mins, using a low pressure mercury vapour lamp (type R51, UV products Inc., Pasadena, CA, USA). The lamp has the low pass (black light) filter removed and emits over a broad range of wavelengths. The radiation generates ozone, which may be involved in the modification of the carbon surface. Work in the University of Leeds EM

group has established that pre-irradiation of grids with UV light and/or exposure to ozone increases the hydrophilicity of the carbon surface, allowing the hydrophilic surfaces of proteins to readily adsorb. Generation of a hydrophilic surface is also found to promote a thin and even layer of stain on drying, and it may also promote interactions between the protein molecules and carbon, which stabilise the protein against damage from staining (Walker et al., 1985). Grids were placed ~5cm from the bulb, with their carbon side upwards and irradiated while enclosed in a cardboard box (in a fume cupboard) to retain ozone. Following UV treatment, grids were left ~20-30 mins before use.

2.8.4 Application of sample and stain

A grid was held in tweezers and 5 μL of sample applied to the carbon film side using a pipette. After ~10 s, a droplet of 1% (w/v) aqueous uranyl acetate (at ambient temperature) was applied using a Pasteur pipette. On contact the droplet was quickly flicked off the grid using the (still attached) pipette, to prevent wetting of the copper underside of the grid. A further 2-3 drops of stain were applied in the same way. Excess liquid was then wicked away from one side of the grid by touching the grid edge with torn filter paper (grade #1, Whatman). The grid was then allowed to dry fully for a few minutes before EM.

2.8.5 Electron microscopy

Negatively-stained grids were examined at University of Leeds using a JEOL 1200EX transmission electron microscope (Jeol Ltd., Tokyo, Japan) equipped with a LaB_6 or tungsten filament, operating at 80 kV accelerating voltage. Before use, the microscope apertures were aligned and a grid sample adjusted to the eucentric point using the goniometer. Corrections for objective astigmatism were made by examining the edges of holes in a holey carbon grid (Agar Scientific) and adjusting the objective stigmator controls to form a fringe of constant width around the edge of a hole, or, alternatively, by observing the carbon grain at 100,000× magnification and adjusting stigmators until shape changes in the grain were minimized as the image was under- and over-focused.

To record images, grids were surveyed at 4,000-10,000× magnification to locate optimal areas of negative staining – typically in areas of intermediate stain depth (not too deep, not too shallow), where molecules are delineated with an outline of stain (Burgess, Walker, Thirumurugan, et al., 2004). Once a suitable area was located, magnification was increased to 80,000× then decreased to 40,000× for imaging, to prevent hysteresis effects. Images were collected at 40,000× magnification at a defocus of ~600 nm (3 anticlockwise increments of the defocus control), using a 30 μm objective aperture. Micrographs were recorded on electron image film (type Kodak SO-163, Eastman Kodak, Rochester, USA) using a 1 s exposure time, giving an estimated electron dose (including focusing time) of

roughly ~500 e⁻/Å² (rendered as $\sim 500\ e^{-}/\text{Å}^2$). Experience in the University of Leeds EM group suggests that stain redistribution under this moderate radiation dose can improve clarity of images (Walker et al., 1991).

Micrograph films were developed in a dark room for 4 mins in Ilford ID-11 Developer (Ilford Ltd., Cheshire, UK) at 20 °C, rinsed for 1 min in water, fixed for 3 mins in Ilford Hypam rapid fixer (Ilford Ltd.) and rinsed under running water for 20 mins before drying at ambient temperature.

2.8.6 Micrograph digitisation

Micrograph films were digitised as 16-bit grey scale images in TIFF format, using an Imacon Flextight 848 scanner (Imacon A/S, Copenhagen, Denmark) at 1270 dpi (corresponding to a step size of 20 μm). The object scale was calibrated by scanning a micrograph of negatively-stained paramyosin filaments purified from oyster (*Crassostrea angulata*) (Dr Michael Murtagh, formerly University of Leeds), acquired under the same electron microscope conditions as the collected data (*i.e.* at 40,000× magnification), or as part of the same EM session, Figure 2.2.
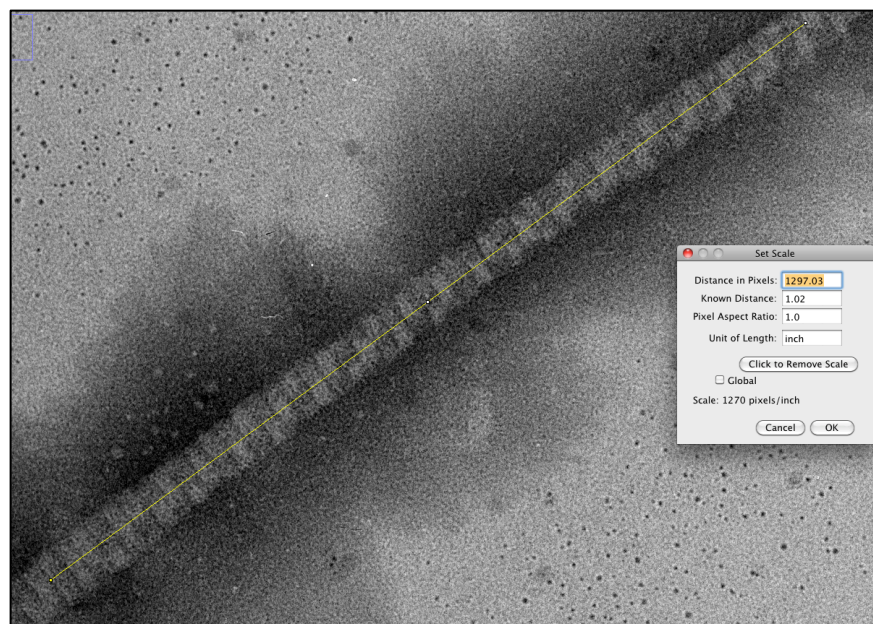
Figure 2.2: Image calibration measurement of a paramyosin filament.

Example of a negatively-stained paramyosin filament (imaged at 40,000× by JEOL 1200EX transmission electron microscope). The pixel distance over 47 axial repeats (yellow line) was measured using *ImageJ* (v1.43) software (NIH, USA). From the known 14.4 nm repeat the sampling scale in nm/pixel can be calculated. Calibration was carried out by averaging over a number of filament images. A value of ~0.52 nm/pixel was determined for micrograph images produced by the 1200EX operating at 40,000× magnification.

The object sampling scale was deduced from the 14.4 nm axial repeat of the paramyosin filaments, a distance which is unaltered by staining and dehydration (Elliott et al., 1976).

By measuring the distance in pixels across a number of repeats in a filament, the average number of pixels per 14.4 nm repeat was calculated, Figure 2.2. A number of filament calculations were averaged to produce a final object sampling scale. A value of ~0.52 nm/pixel was consistently determined.

2.9   Image processing

Scanned micrograph TIFF images were first converted to SPIDER format using the *em2em* programme (Image Science Software GmbH, Berlin, Germany).

All image processing work presented in this thesis was carried out using the *SPIDER* image processing system (versions 17.05 or 19.05) (Health Research, Inc., Rensselaer, NY, USA) running on linux (Frank et al., 1996). *SPIDER* commands were either issued at the command line during an interactive session or executed via *SPIDER* 'procedures', script files containing multiple commands and scripting logic, for frequently used sequences of operations. New script files were written as needed by the author of this thesis, while many of the procedures used to perform reference-free alignment and image classification were adaptations of scripts written by Dr Stanley Burgess of the University of Leeds EM group.

The general work-flow for image processing of negative stain EM images followed the approach and methods developed by Dr Stanley Burgess as described in Burgess, Walker, Thirumurugan, et al. (2004). In outline, the steps in the process are: (i) particle 'picking' and windowing-out from micrograph images, (ii) iterative reference-free rotational and translational alignment of images and (iii), classification and averaging of the images. Specific details concerning these steps are described in the results chapters of this thesis.

Chapter 3: Structure and appearance of the myosin 5a head

3.1   Introduction and aims

In this chapter, a detailed study of structural features of the head region of myosin 5a molecules is presented. The motivation for the work was to compare the appearance of the head with existing crystal structures and to look for evidence of flexibility. A large image dataset of free heavy meromyosin (HMM)-like molecules, obtained for a previous negative stain EM study, was used for the work. The dataset size and quality of staining were such that new structural details and conformational variety not previously described were revealed from single-particle image processing.

The specific aims of the work presented in this chapter were:

1. To align images of the heads of myosin 5a HMM molecules to reveal structural details of the motor and lever domains.
2. To compare the HMM head appearance with existing atomic models and to create a new head model if appropriate.
3. To use image classification to look for evidence of flexibility in different parts of the head.

Reflecting these aims, the chapter divides into three parts. First, details of the dataset used are described, followed by a description of how images were processed and aligned to produce an average of the head. Second, the average is compared to existing atomic models of the head. One of these is selected for modification to create a new model that better fits the HMM appearance. The new model is then used in further image analysis in this and the following chapter. Third, conformational flexibility evident from image classification is investigated. Three main types of flexibility are discussed: (i) motor rotation about the lever's axis, (ii) motor tilting about the motor-lever junction and (iii), rotation of discrete calmodulin (CaM)-IQ sections about the lever axis, the latter being evidence of torsional flexibility in the lever that has not previously been described. A detailed analysis of the lever's bending flexibility, also apparent from the images studied, is the subject of chapter 4.

3.2   Negative stain dataset used

For the structural studies in this and the following chapter, a pre-existing negatively-stained dataset of recombinant myosin 5a molecules was used. The data was collected

previously by members of the EM group at University of Leeds as part of a collaboration with the group of Dr James Sellers at the Laboratory of Molecular Physiology, National Institutes of Health (LMP-NIH) (Bethesda, MA, USA). Originally collected for a CaM labelling study, the EM images acquired were not subsequently used, the level of CaM labelling being too low. Due to the high quality of staining and the large dataset acquired, it was decided to make use of the existing images for the studies of the head in this thesis. The provenance, details of the sample and information on the original EM data collection is detailed below.

The protein sample used was a recombinant myosin 5a fragment, containing amino acids 1-1090 of the mouse heavy chain sequence (NCBI GenBank accession: NM_010864), with a FLAG tag epitope (DYKDDDDK) added at the C-terminus to facilitate purification. On expression, two such heavy chains form a two-headed molecule with a truncated tail that ends after the first section of coiled-coil (Wang et al., 2000). Being analogous to the proteolytically-produced HMM fragment of myosin 2 (see 1.5.1.1), the myosin 5a construct is hereafter referred to as myosin 5a-HMM, or simply as an HMM molecule.

The myosin 5a-HMM construct was cloned and co-expressed with CaM by Dr Takeshi Sakamoto at LMP-NIH, using a Baculovirus/Sf9 cell expression system (Wang et al., 2000). Purification was by FLAG-affinity chromatography (Sakamoto et al., 2003; Wang et al., 2000). Prior to shipment to University of Leeds, the protein was prepared for a labelling study. GFP-CaM (CaM with Green Fluorescent Protein sequence added at the N-terminus) was exchanged with endogenous CaMs on the heavy chain, using calcium-induced dissociation and reassociation (Forkey et al., 2003; Sakamoto et al., 2000), and the protein re-purified by Dr Sakamoto. However, the degree of GFP-CaM exchange was subsequently estimated to be very low (< 0.9 per head) (Prof. Peter Knight, University of Leeds, after personal communication by Dr Sakamoto), a level too low to be seen by EM and image-processing studies. No evidence for the GFP-labelled CaM over endogenous CaM has been seen in EM images subsequently analysed at The University of Leeds. Consequently, for purposes of this thesis, the large majority of HMM molecules are assumed to bear only endogenous CaMs.

The HMM molecules were imaged by Dr Matthew Walker (MLW consulting, Launceston, UK) at ×40,000 magnification using a JEOL 1200EX electron microscope (equipped with a LaB_6 filament) at The University of Leeds (see 2.8.5 for the general EM methods used). The HMM molecules were in apo state (*i.e.* no ATP present), free from actin, on carbon-coated 300- or 400-mesh copper grids and stained with 1% (w/v) unbuffered uranyl acetate (see 2.8.4 for staining methods). Final buffer conditions of the sample on the grid prior to

stain addition were: 100 mM KCl, 5 mM MOPS, 3 mM MgCl$_2$, 1 mM EGTA, pH 7.5. Protein samples and dilution buffer were kept on ice until application to the grid and staining with room temperature stain, a process encompassed in ~10-15 s. The temperature of the sample at the point of staining was not measured, but is estimated at between 10-23 °C. 103 micrographs of fields of negatively-stained HMM molecules were captured on Kodak S0-183 film (Eastman Kodak, Rochester, USA) using a 1 s exposure time. Films were developed by Dr Walker in-house (2.8.5).

The images acquired, which exhibit exceptionally good staining of a pure protein sample Figure 3.1(a), are consistent with earlier observations of free myosin 5a-HMM molecules made by the EM group at University of Leeds (Burgess et al., 2002; Oke, 2004). V-shaped molecules are seen, each arm of which is a head of the two-headed HMM. Head-head angles range widely, from ~30-180°, but appear most frequently at ~100-120°. This is roughly consistent with the Gaussian angle distribution, centred at 128.6 ± 40.6° (mean ± SD, n = 117), measured previously from rotary-shadowed EM images (Baboolal et al., 2009). Motor domains are visible as a bulbous tip at the end of each head (arrowheads Figure 3.1(a)). The CaM substructure in the lever is suggested by the rough, lumpy appearance of the elongated domains extending from each motor. Significant flexibility in the lever is suggested by the range of differently curved or sometimes kinked shapes, see Figure 3.1. The levers come together at a vertex (the head-tail junction) but the HMM tail (estimated 27 nm long) is not seen, presumably because its coiled-coil structure is too thin to accumulate significant stain around it.

The 103 micrograph films were digitized as 16-bit grey scale TIFF images by Dr Kavitha Thirumurugan (formerly University of Leeds, UK), using an Imacon 848 scanner (Imacon/Hasselblad) at 1270 ppi. Object sampling in the image was calibrated by measurement of a scanned paramyosin micrograph, see 2.8.6, and calculated as 5.2 Å/pixel for 40,000× nominal magnification.

The micrograph TIFF files were converted to SPIDER format using the *em2em* programme (Image Science Software GmbH, Berlin, Germany). Pixel coordinates of the motors of 23,158 individual heads from the HMM molecules were recorded from the images by Dr Thirumurugan, by clicking with a mouse on the centre of the globular motor domains (a process termed 'particle picking'), using the 'markers' function in *Web* (*SPIDER*'s image-viewing software) (Wadsworth Center, New York State Department of Health, Albany NY, USA) (Frank et al., 1996).

(a)
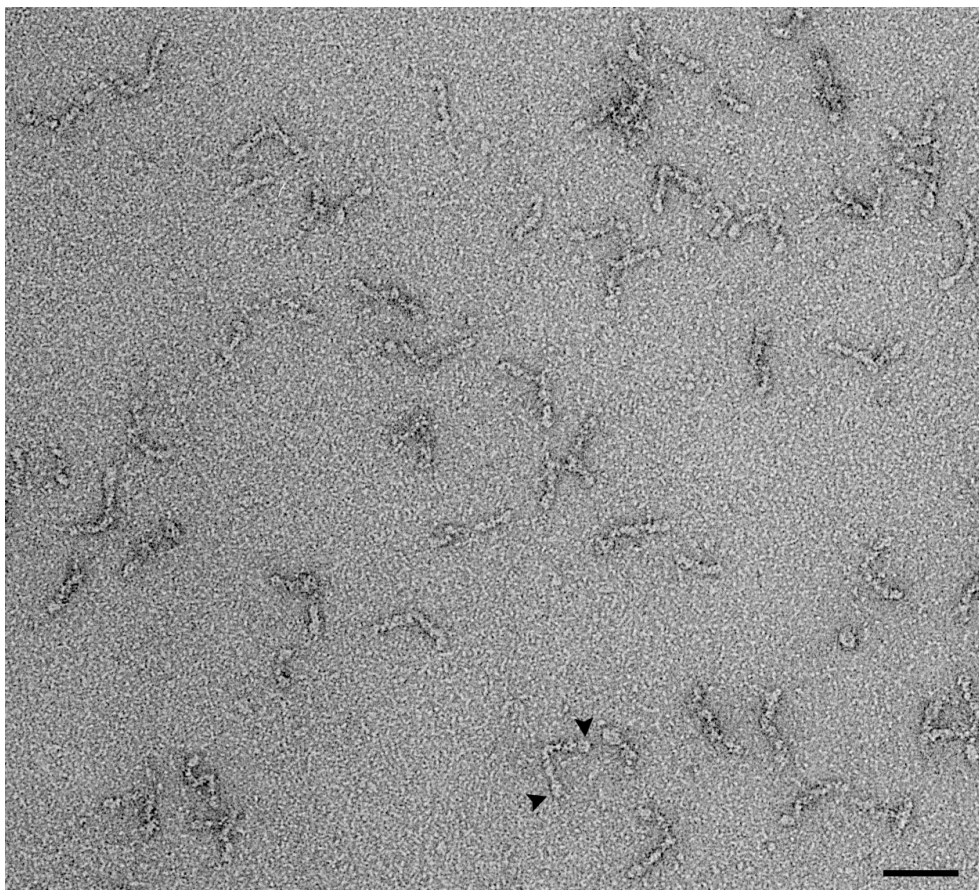
(b)
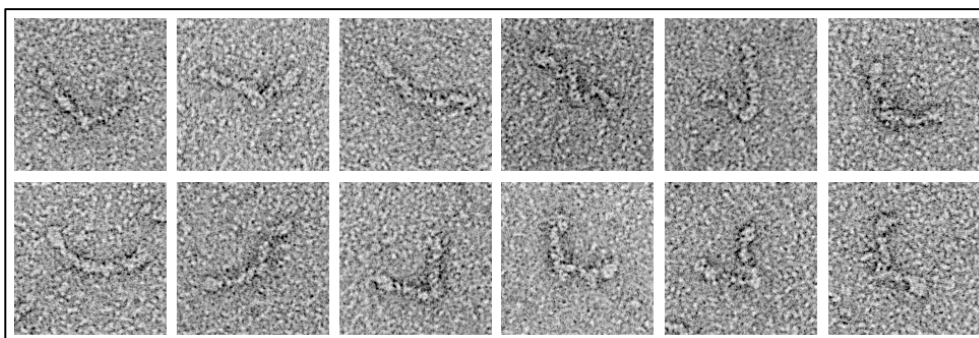
Figure 3.1: Micrograph images of negatively-stained myosin 5a-HMM molecules.

(a) field of negatively-stained myosin 5a-HMM molecules. Each molecule is identified as a V-shape with the two globular motor domains visible at its tips (arrowheads). Field of view shown is ~1/6 of the total digitized micrograph. Scale bar 50 nm. (b) montage of individual molecules from the micrograph in (a). Each image is 78 nm wide. Microscopy performed by Dr Matthew Walker at University of Leeds. Grid conditions: 100 mM KCl, 5 mM MOPS, 3 mM MgCl$_2$, 1 mM EGTA, pH 7.5. 1 % uranyl acetate stain.

The micrograph images were subsequently re-surveyed by the author of this thesis to ensure that any original bias towards picking molecules of a predominant shape (*e.g.* with straighter levers) was avoided. An additional 1,562 head coordinates were added to the dataset through this review. Aside from general selection criteria for particle picking *i.e.*

molecules in optimum stain depth, not crowded or near to micrograph edges, heads were picked only if a globular-shaped motor domain was identifiable at the end of the head and if the lever section was apparently complete (*i.e.* intact and fully stained), preferably joining with, and not obscured by, the second head in the HMM molecule. In the large majority of cases both heads of HMM molecules were picked.

The final dataset comprised 24,720 coordinates of the motor domains of HMM molecules, referencing 103 micrograph image files. The subsequent processing and alignment procedures applied to the images are described in the following section.

3.3   Alignment of myosin 5a-HMM head images

Image pre-processing and reference-free alignment described in this section was performed by the author of this thesis using *SPIDER* (v19.05) and *WEB* software (Wadsworth Center, New York State Department of Health, Albany NY, USA) (Frank et al., 1996). Processing tasks were performed using *SPIDER* commands, denoted hereafter in courier font in parentheses *e.g.* (WI). Sequences of commands were executed via *SPIDER* scripts ('procedures'), most of which were adaptations of scripts previously developed in the EM group by Dr Stanley Burgess (University of Leeds, UK).

3.3.1   Image pre-processing

Individual molecule images of 160×160 pixels, each centred on a picked motor domain coordinate, were excised from the larger micrograph images (WI command). Excessively bright or dark pixels (*e.g.* caused by film scratches or dust), at > 5 S.D. from the mean intensity, were 'cleansed' by setting to the 5 S.D. threshold value (TH command). A ramp subtraction was made to correct for gradual intensity variations across the image (*e.g.* from stain depth changes) (RA command), and the pixel intensities in each image were normalised to a mean of 0 and S.D. of 1 (AR command). The 24,720 molecule images generated were saved in an image stack.

3.3.2   Reference-free image alignment

To help focus image alignment on just the HMM head with its motor centred, images were first cropped to 120×120 pixels (WI command), ramp subtracted and normalised again as above (RA and AR commands). A soft-edged circular mask with 55 pixel radius and 4 pixel falloff, sufficient to encompass the motor and lever of the centred head, was applied to the images (MA command).

The alignment strategy adopted involved rounds of reference-free rotational alignment (AP RA command, using outer radius 55 pixels, inner radius 11 pixels) followed by reference-

free translational alignment (`AP SA`, with maximum shift of 16 pixels). The constraints on both alignments were carefully chosen after consideration of the best alignment strategy, see 3.3.2.1 below. After each type of alignment, the entire image stack was aligned *en bloc* with a model image, to maintain orientation and keep the alignment centred between iterations. The resulting aligned image stack was used as input to the next round of alignment. The process was iterated 15 times in total, with rotational and translational average and variance images produced for each iteration (`AS R` command). The final alignment parameters obtained for each image (rotation angle, *x*-shift, *y*-shift) were then applied to the original unmasked 160×160 images via a single-step interpolation (calculated by `SA P` and executed by `RT SQ` commands), thus avoiding additive interpolation errors from multiple transformations. The final aligned images were cropped to 100×110 pixels around the aligned head of interest (`WI`), oriented with the motor domain top centre and the lever extending downwards along the centre vertical line. Ramp subtraction and normalisation were performed again (`RA` and `AR`), and all pixel intensities were shifted by a constant (`AR`) to ensure that the minimum value across the stack was > 0, a requirement of the subsequent image classification algorithm. Finally, the aligned images were averaged (`AS R` command) to produce 'global average' and 'global variance' images.

The final aligned image stack was then screened for badly aligned images. This is normally done via image classification and removal of images belonging to classes with poorly resolved averages. However, due to the quality of staining of individual molecules and the clearly identifiable head shape, it was possible in this case to examine the raw stack images by eye. 1,470 images (~6% of the total) were discarded through manual screening. These were images where the motor domain was significantly off-centre, or where the lever projected in the wrong direction (*e.g.* 90-180° to the vertical) (but not through sharp lever bending). This generated a final working image stack of 23,250 well-aligned head images. New global average and variance images were created (`AS R` command).

3.3.2.1   Head alignment strategies
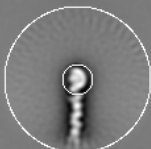
Crucial to the success of image alignment are the constraints applied to the process. The rotational alignment requires an inner and outer pixel radius that specify an annulus in the image. The rotational alignment is calculated only by considering the pixels within the annulus. The inner radius has a minimum of 5 pixels, since, below this, rotation artefacts in the centre of the image become too significant. The translational alignment can be

constrained by a maximum permitted shift. If the shift needed to align an image exceeds the maximum, the image is not translated.

For this thesis, which focuses on the lever, the best strategy for alignment by structural features of this domain was sought. Several rotational alignment strategies were tested: (i) alignment by the beginning section of the lever, (ii) alignment by the full length of the lever, and (iii) alignment on the outer features of the motor domain. The strategies tested and their advantages and disadvantages are detailed in Table 3.1.

Table 3.1: Three rotational alignment strategies for the myosin 5a head.
Pictures in the left column show the inner and outer rings constraining the alignment region, superposed on a global average of the head from a pilot dataset.

(i)	**Alignment by first part of lever:** Alignment constrained to the beginning section of the lever, ~CaMs 1 & 2, defined by the region between the concentric rings at 11 and 22 pixel radii. The motor domain is excluded. Strategy aligns the start of levers well and the position at which the lever emerges relative to motor does not matter. Drawback: alignment annulus is narrow, so probability of alignment success per image is reduced.
(ii)	**Alignment on the full length of the lever:** Alignment on annulus formed by radii at 11 and 55 pixels. Motor excluded. Strategy brings levers into same general orientation. Levers of similar shape and with similar features align and reinforce in average. Emergence position of lever from motor domain does not influence alignment. Drawbacks: unpredictable result of aligning curved levers with straighter ones. Start and end points of levers may not align well across images.
(iii)	**Alignment on outer features of the motor domain**, Alignment on annulus formed by 5 and 11 pixel radii. Lever domain excluded. Alignment is on motor only but allows lever to radiate with range of directions. Drawbacks: very small alignment annulus reduces alignment success. Wide range of lever angle projections blur-out in image averages.

After testing, strategy (ii) was chosen as the most effective method. By this strategy – alignment by the full lever length – levers are brought in to the same *general* alignment (oriented vertically in the images as presented) by maximising the area of the molecule considered. The motor, which may vary independently in position relative to the lever is purposely excluded from consideration. Consistent structural features of the individual CaMs in the lever reinforce through alignment.

Translational alignments were constrained to a 16 pixel maximum shift, approximately the width of the motor domain. In the raw images, the motor can have a similar globular appearance to one or two CaMs (see Figure 3.1(b)). To limit mistaken alignment of CaM-1 with the motor or vice versa, the smallest shift possible, still allowing for the inaccuracies in hand-marking the motor domain centre, was imposed. The 16 pixel constraint was

chosen after inspection of a random sample of 200 raw images with 30×30 pixel crosses superposed on their centres, Figure 3.2. Provided the motor domain centre fell within or on the circle (radius 15 pixels) defined by the cross-shaped targets, the head would be translationally aligned. For the 200 random images checked, all motor domains fell within the 30×30 target area.
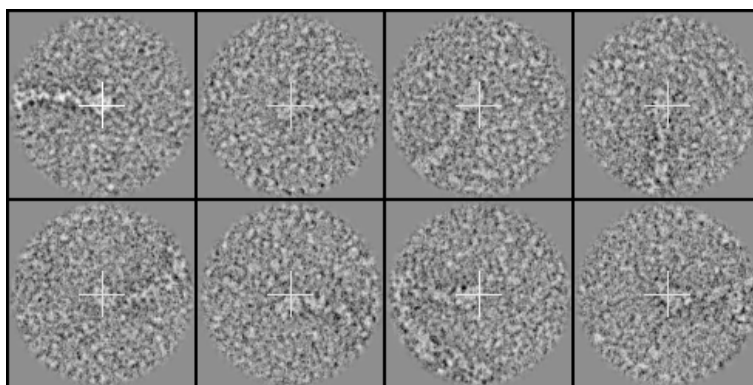
Figure 3.2: Shift alignment check on raw head images

Example images of marked heads from micrograph images with 30×30 pixel target crosses (white) superposed on their centres. For a molecule to be subject to translational alignment, the motor centre must fall on, or within the circle defined by the target cross. Note; in one image (third from left bottom row), the second head of the HMM molecule is visible.

3.3.3   Results of alignment and averaging

The global average and variance images produced from the lever-aligned head images are shown in Figure 3.3(a) and (c). For comparison, a view of a crystal structure of the myosin 5a (apo) motor and first light chain (PDB accession: 1OE9) (Coureux et al., 2003) is shown alongside in Figure 3.3(b). The global average produced from the aligned image stack shows a very clearly delineated head structure. The level of detail resolved, especially in the lever domain, is high.

By comparison with the crystal structure, the motor within the head is clearly identifiable as the ear-shaped domain at the top of the images, with a curved right-hand-side and a straight left-hand-side bearing two protrusions. From the average, the motor measures ~11 nm along its long axis and is ~7.8 nm across at its widest. The specific apo motor appearance seen can also been likened to a face seen in profile. The face 'looks' towards the left, with the two protrusions on the motor's left side being a pointed chin (identifiable from the crystal structure as the SH3 domain) and the nose (part of the lower 50k). The face's eyes would be level with the actin-binding cleft, which is only partially resolved in the average – identifiable as a slightly darker patch at the motor's edge in Figure 3.3(a). The term face-in-profile is adopted hereafter, as a convenient label for this characteristic motor view.
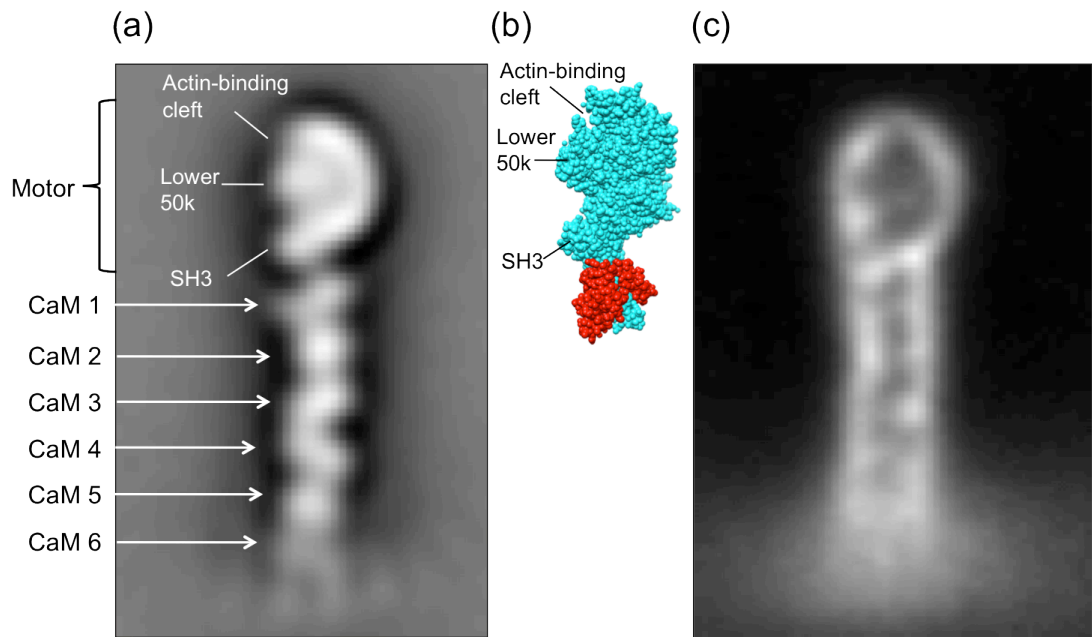
Figure 3.3: Global average and variance images of myosin 5a-HMM head.

(a) global average of 23,250 lever-aligned head images. Identifiable structural features are labelled. (b) view of crystal structure of the myosin 5a motor (heavy chain in cyan) and ELC (red) (PDB: 1OE9) (Coureux et al., 2003). (c) variance image of 23,250 heads. Whiter pixels indicate more variance across the images. Images computed by *SPIDER* (`AS R` command).

The variance image in Figure 3.3(c) indicates that there is significant variety in the appearance and staining pattern around the actin-binding cleft and protrusions on the left side of the motor, and also at the motor's bottom right corner at the interface with CaM-1. Such variance can arise simply from variation in staining around the molecule, but it could also be an indication that the motor may not always be oriented with its face looking left in all images in the dataset. Variety in motor appearance within the data is investigated further in 3.5.1.

The six-CaM substructure of the lever is also clearly resolved in the global average. Each CaM is visible as a circular or elliptically-shaped blob. Each blob merges with the next, indicating that the CaMs are closely packed and that each may form contacts with its neighbours (or CaM-1 with the motor). The major axes of the elliptically-shaped CaMs (numbers 1, 3, 4 and 6) make different angles with the vertical (the assumed line of the lever helix). CaMs 3 and 4 together make a characteristic '<' shape, with their left sides apparently in contact and right sides far apart. The distinct CaM appearances and differently oriented angles indicates that, on average, the CaMs are seen at different azimuthal orientations about the lever axis.

Similar to the motor region, the variance image indicates that there is also significant variability in lever appearance across the images. Variance in the stain pattern immediately around the CaMs is high, and suggests potential variability in CaM orientations. This

77

possibility is explored in detail in 3.5.2. A fan shape of increasing variance also extends outwards from both sides of the lever, beginning around CaM-3 and broadening across the width of the image by CaM-6. This is likely due to the variable lever shapes in the images. The fan shape is actually slightly asymmetric, beginning slightly lower at CaM-4 on the right side. This may indicate a bias in lever curvature towards the left of vertical, a point investigated further in chapter 4. The variance envelope seen is likely due to significant lever bending and the range of straight, curved and kinked conformations present in the raw images (see Figure 3.1). The variance image shows that at the lever's proximal end (nearest the motor), CaMs 1-3 apparently form a straighter, more tightly-coupled unit, whereas there is increasing positional variance progressing through CaMs 4-6 towards the lever's distal end. The CaM-6 region exhibits the most variance and is therefore the most indistinct CaM in the global average. A complicating factor that must be remembered here is that, in some images, the second non-aligned head of the HMM molecule may also be present (or partially present) and so it too will add to the average and variance representations (see example in Figure 3.2). The symmetrical distribution of high variance either side of CaM-6 suggests that second heads are found equally right or left of the centred head, and radiate most commonly at ~90° to the aligned lever (vertical axis).

Inspection of the global average and variance images alone allow only general comments on the flexibility of the head to be made. However, computational subdivision of the images into homogenous groups – image classification, offers a powerful means to examine conformational variety and molecule flexibility in more detail. The technique is ideally suited to sorting large image sets, in which molecules are found in different orientations and with continuously flexible domains (Burgess, Walker, Thirumurugan, et al., 2004). In later sections of this chapter and the next, *SPIDER* is used to classify the lever-aligned head images (using its K-means clustering algorithm (`CL KM` command)), focusing on different regions of interest in the head. Motor domain flexibility and CaM rotational freedom about the lever axis are studied in this chapter, while chapter 4 analyses the conformations resulting from lever bending.

Before embarking on an examination of conformational variability by image classification, the global average in Figure 3.3(a) is first compared to atomic models of the myosin 5a head; to test how well the newly-processed HMM data matches existing models and to create a new working model if required.

3.4 Atomic models of the myosin 5a head

This section first describes a comparison of the average image of the myosin 5a-HMM head obtained by negative stain EM with existing atomic models. One of these models, constructed by Liu et al. (2006), to fit a 3D cryo-EM density map of the folded full-length molecule, is then studied in more detail, before being adapted to make a closer match to the HMM head appearance.

3.4.1 Comparison of HMM head average to existing models

The global average of the myosin 5a head obtained in 3.3.3 was compared to three existing atomic models. Figure 3.4 depicts the EM average alongside views of the three models considered.
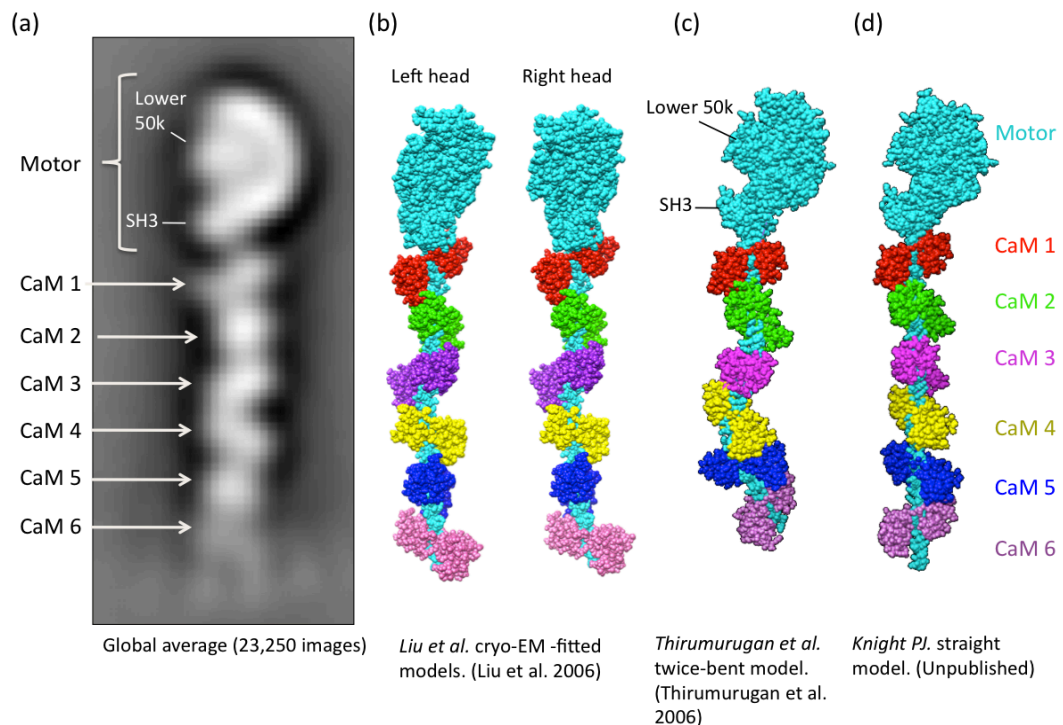
Figure 3.4: Comparison of the HMM head average with different myosin 5a head models.

(a) shows the global average of 23,250 lever-aligned images of myosin 5a-HMM heads (in the absence of ATP) obtained by negative stain EM. Note the six delineated CaM structures along the lever, with varying shapes that indicate different azimuthal orientations. The motor has the face-in-profile appearance with the lower 50k (nose) and SH3 (chin) subdomains identified. (b)-(d) show different atomic models of the myosin 5a head. Each is positioned so that CaMs 1-2 are orientated to give the same outline profile as seen in the average (*i.e.* CaM 1 with both lobes visible, CaM with only one lobe seen). **(b)** Liu et al. (2006) model ('left' and 'right' heads – see text) fit to cryoelectron tomography density **(c)** Thirumurugan et al. (2006) model with twice-bent lever for negative stain EM study **(d)** Knight model with straight lever (unpublished). The orientations of all six CaMs in the Liu model (left and right heads) create a good match to the profile of the lever in the global average, while the Thirumurugan and Knight models cannot be orientated so that CaMs 1-2 and 3-6 simultaneously match the global average profile.

These models, compared in Figure 3.4, were:

(i) *Liu et al. full length myosin 5a model in inhibited state – 'left' and 'right' heads* (PDB: 2DFS) (Liu et al., 2006).

To build this model, the authors constructed a model of a single head by joining the X-ray structure of the myosin 5a (ADP) motor domain plus essential light chain (PDB: 1WJI) (Coureux et al., 2004), to repeats of the X-ray structure of the calmodulin-like myosin light chain (Mlc1p) binding to IQ motifs 2-3 of yeast myosin 5 homologue Myo2p (PDB: 1N2D) (Terrak et al., 2005). A copy of the single-head model was then docked into each of the two head regions of the authors' 2.4 nm resolution 3D density map of myosin 5a, obtained by cryoelectron tomography of 2D crystals of inhibited molecules (see Figure 3.5(a)-(c)). Each docked head was then independently fitted to the local region of the density map using normal modes flexible fitting (Tama et al., 2004).

The independent fitting resulted in two differing head models, labelled in Figure 3.4 as Liu et al. 'left head' and 'right head'. The left head is the head which in the 3D density map lies closer to the folded back tail, see Figure 3.6(b). In negative stain EM images this appears as the left head of the two when motors are orientated at the top of an image (see Figure 3.6(a)), hence its name.

(ii) *Thirumurugan et al., model of myosin 5a in inhibited state with 'twice bent' lever* (Thirumurugan et al., 2006).

This model was constructed by the authors to compare to their negative stain EM images of individual myosin 5a molecules in inhibited state. Following a suggestion in (Houdusse et al., 1996), the lever helix was formed from alternating sections of straight helix for 23 residue-spaced IQ motifs (brush border myosin 1 sequence) and bent sections of helix for 25 residue spacings (scallop myosin 2 lever, PDB: 1WDC) (Houdusse and Cohen, 1996). CaMs were added by superposing the heavy chain backbone atoms of the IQ motif of a CaM-heavy chain model (PBD: 1AJI) (Houdusse et al., 1996) onto the lever helix. A motor domain was added by superposing the first IQ motif of the myosin 5a (ADP.BeF$_x$) motor structure (with first IQ motif and ELC bound) (PDB: 1W7J) onto the first IQ of the lever model.

(iii) *Knight PJ. (University of Leeds), model of myosin 5a head with straight lever* (Unpublished).

This model, constructed as alternative comparison to (ii), was constructed with a entirely straight lever helix (brush border myosin 1 sequence), generated by *SWISS-MODEL* software to have 3.6 residues per turn of α-helix. CaMs were superposed onto the IQ motifs and a motor added as in model (ii).

Each model in Figure 3.4 has been orientated so that the first and second CaMs create a profile that matches the shapes in the global average, namely CaM-1 with both lobes visible (C-lobe closest to the motor), and CaM-2 with only one lobe visible (C-lobe uppermost, N-lobe hidden behind). It is clear from the profile comparison that the CaM shapes for the entire lever domain, are well matched along the whole of the lever for the Liu et al. model (both left and right heads), while neither models in (c) or (d) can be orientated so that all their CaM profiles simultaneously match the EM average. From this, one can conclude that the myosin 5a-HMM lever is not entirely straight, nor exactly 3.6 residues per α-helix turn as in model (iii). Moreover any bends that are present in the lever are less pronounced than the myosin 2 lever (where the IQ spacing is 26 residues) as used in model (ii). The comparison suggests that the Liu model's lever section (left or right head) is also a good model for the HMM lever. From this, one can conclude that the lever structure does not significantly alter when the molecule changes from an active unfolded state (the HMM molecule) to a folded inhibited state (the Liu et al. model). Moreover, the processes of staining, adsorption and drying on the carbon film of the EM grid do not seem to have adversely distorted the HMM lever's appearance.

Though the Liu et al. model's lever section is a good match to the HMM average, the profile of the model's motor region (left or right head) is not a good match. The motor profiles for models (ii) and (iii) depicted in Figure 3.4(c) and (d) are, however, significantly better; exhibiting the characteristic face-in-profile appearance, with an identifiable protrusion of the SH3 (chin) and lower 50k (nose) subdomains (separated by a cleft) at the lower- and mid-left side of the motor. The motor's right side is also curved similarly to the HMM average in the models depicted in (c) and (d), but not in (b). The only discrepancy with the motor appearance in (c) and (d) is that the model motors are tilted to the right of vertical (by ~30°) compared to the straighter HMM average.

Thus, none of the existing models compared here entirely match the new HMM head average. The Liu et al. model matches the lever but not the motor, while the Thirumurugan et al. and Knight models match motor and CaMs 1-2, but not CaMs 3-6 in the lever. In the

following section, the Liu et al. model is examined more closely with a view to reconciling the motor domain mismatch with the HMM average.

3.4.2   Analysis of the Liu et al. model

As mentioned above, the Liu et al. model was constructed by the authors by docking and independently fitting two copies of a single-head model into their 3D density map, obtained by cryoelectron tomography of 2D crystals of inhibited myosin 5a molecules formed on a lipid monolayer. The inhibited molecules formed 2D crystals with a hexagonal unit cell and a repeating six-petal 'coneflower' motif, comprising six V-shape molecules arranged in a ring, with their motors pointing inward, Figure 3.5(a) (Liu et al., 2006).

The head termed 'left head' here is the one lying closer to the folded-back tail in each petal of the coneflower, while the 'right head' is the one slightly further away (see Figure 3.5(a)). Confusingly, these labelling terms derive from the head positions as seen in negative stain EM studies of individual molecules (Thirumurugan et al., 2006), which are seen from the opposite side to how they appear in the coneflower motif when viewed from the solution side (*i.e.* the molecules in the 2D crystal form on the lipid monolayer the other way around to how they adsorb to the carbon surface in negative stain EM).

Both left and right heads were compared to the EM HMM average in Figure 3.4. It is apparent from Figure 3.4(b) and Figure 3.5(c) that, despite the independent fitting of each single-head into the density map, the appearance and orientations of the CaMs in the lever remain very similar. Each lever fits well in its region of the 3D density map. However, the motor orientations differ between left and right heads, and, as described above, neither is a good match to the HMM face-in-profile average. These observations indicate that the main movement of the single-head model in the normal modes fitting step has been to flex and rotate the motors about the motor-lever junction. The lever helix and orientations of the bound CaMs remain largely unchanged, see Figure 3.5(b)-(c). This is well illustrated by Figure 3.5(d), where the right head model has been aligned by the backbone atoms of its lever helix onto the left head. Here, the lever helices and CaMs superpose well, while the difference in motor position is evident.
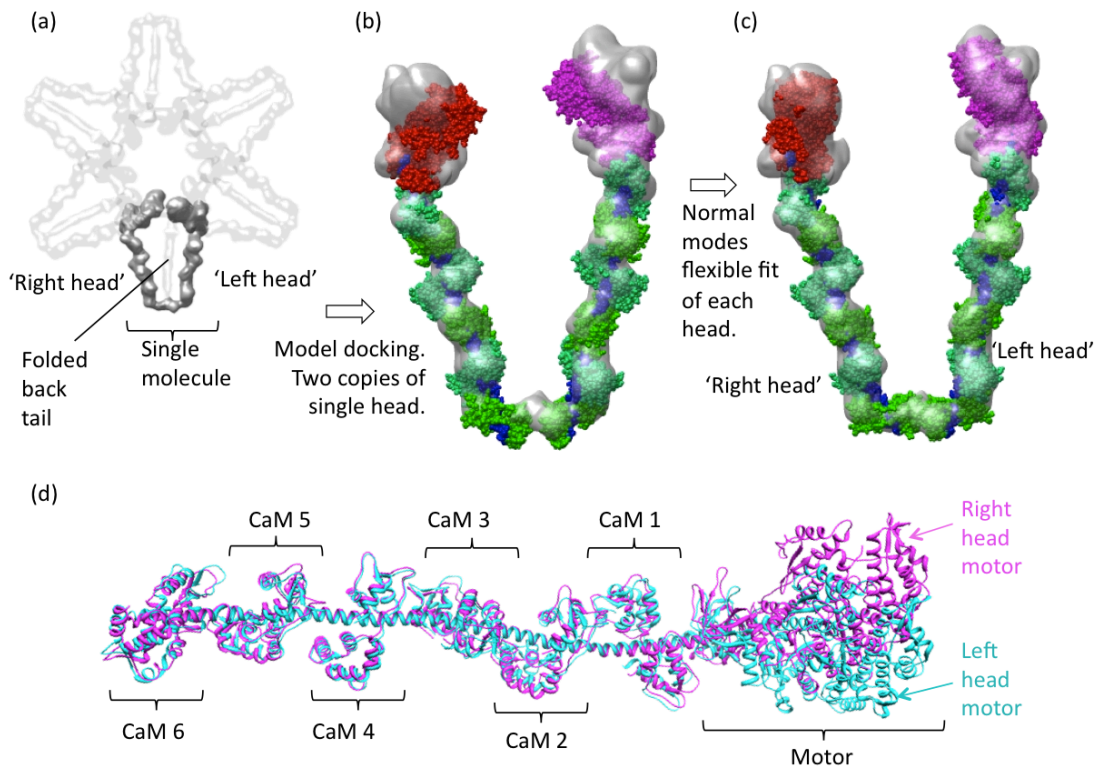
Figure 3.5: Liu et al. model of myosin 5a in inhibited state.

(a) the six-petal coneflower-shaped density map that forms the unit cell of 2D crystals of folded myosin 5a molecules, obtained by Liu et al. (2006) using cryoelectron tomography. Each coneflower contains six overlapping V-shaped molecules with motors inwards. One is highlighted (darker grey). Note; the coneflower motif is shown as presented in Liu et al. (2006) and is seen from the opposite side to how it appears in negative stain EM studies of individual folded molecules from which the labels 'left' and 'right' heads come. (b) the initial position of the single-head atomic model, two copies of which were docked by Liu et al. into the head regions in one petal of the coneflower map. Note that after initial docking, the motor domains (red and magenta) do not fit within the density map. (c) shows the single-head atomic models in the density map after normal modes flexible fitting by Liu et al. The most significant change from (b) is that the motor domains have moved into the map, by rotation about the motor-lever junctions. The lever domain heavy chain (blue) and CaMs (green) do not move much. (d) ribbon depiction of the 'right head' (magenta) of the Liu et al. model aligned by the lever helix backbone to the 'left head' (cyan) after normal modes flexible fitting. Images in (a)-(c) are adapted from supplementary information Liu et al. (2006).

What is the reason for the mismatch in the appearance of the Liu et al. model motor domains and the HMM face-in-profile average in Figure 3.4? Three possibilities for the discrepancy were considered:

1. *Nucleotide state of the motor domain*.

   The Liu et al. model used a motor with ADP bound (PDB: 1W7I), whereas the HMM data was collected in nucleotide-free conditions. An apo motor structure combined with the Liu et al. lever might therefore be a more appropriate comparison to the HMM average.

2. *Conformational difference between HMM and inhibited full length molecules.*

   The motors may adopt different orientations to the HMM average in the inhibited full length molecule, because they contact the globular tail domains (GTDs) of the folded back tail to form a compact stable structure. In the opened HMM molecule the truncated tail is not folded back and lacks GTDs.

3. *Erroneous positioning of the motor domains in the Liu et al. model.*

   The possibility that the motor positions found by the original normal modes flexible fitting were incorrect was also considered. This possibility was entertained in the light of a reinterpretation of the molecule arrangement in the coneflower motif made by Sellers et al. (2008), on the basis of negative stain EM images of individual folded molecules by Thirumurugan et al. (2006). Sellers et al. reinterpreted the coneflower motif as being formed by six molecules in a tiled 'rosette' arrangement, with the motors of each molecule contacting their own GTDs, rather than contacting the GTDs of adjacent molecules (so called 'domain swapping'), as originally suggested by Liu et al., see Figure 3.6. In the Sellers et al. reinterpretation, the density region associated with the motor domain has a notably closer resemblance to the face-in-profile motor of both individual folded molecules and the new HMM average in this chapter. In particular the protruding lower 50k 'nose' and SH3 'chin' on the left side of the motor are evident, see Figure 3.6(c). This is a strong indication that the reinterpretation has merit.

In the next section, the above possibilities for the motor appearance discrepancy are considered in the construction of a new atomic model to match the HMM average. In particular, the qualitative reinterpretation put forward by Sellers et al. is tested to see whether it makes sense quantitatively in the cryo-EM density map of Liu et al. and in comparison with the HMM appearance.
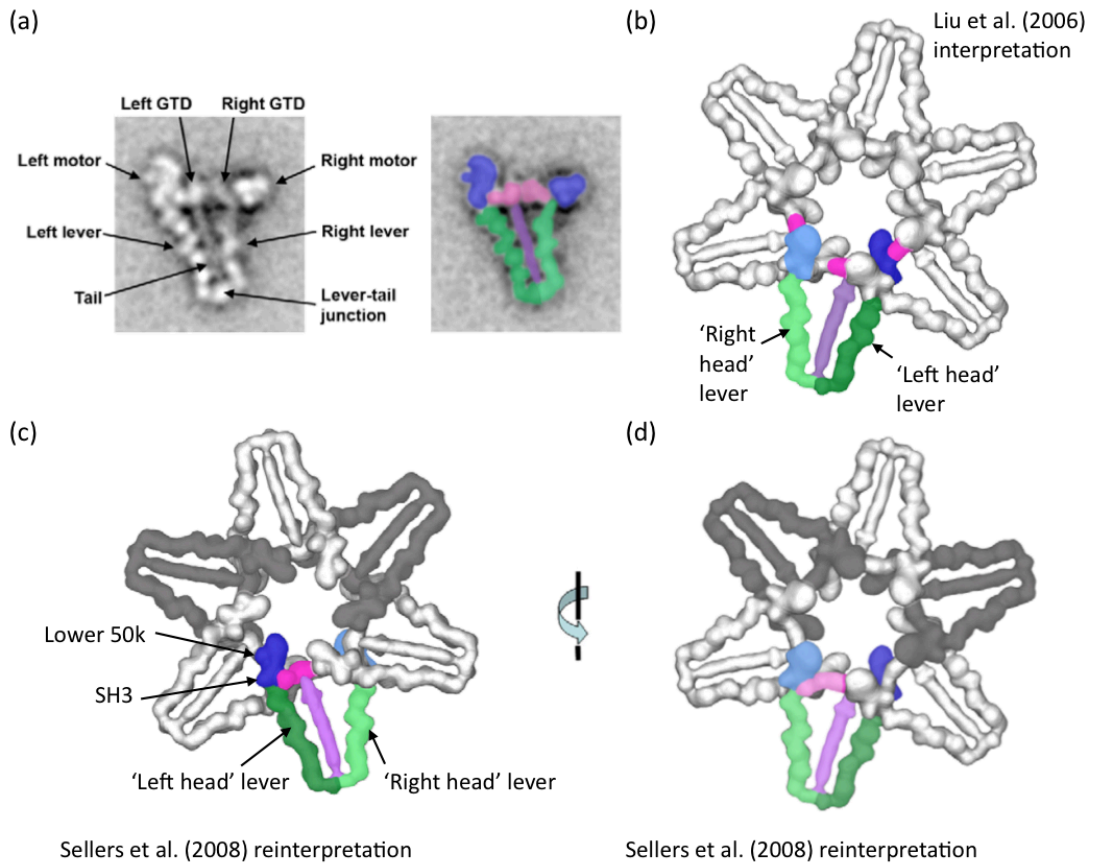
Figure 3.6: Reinterpretation of Liu et al. cryo-EM density map by Sellers et al.

(a) an annotated classification average from Thirumurugan et al. (2006) of single myosin 5a folded molecules left, with molecule domains colour-coded right: blue=motor, green=lever, purple=tail, pink=globular tail domain (GTD). (b) shows the Liu et al. coneflower density map with the authors' interpretation of the various domain regions within the map. Note that the motor domains are interacting with the GTDs in adjacent molecules in the coneflower (so-called 'domain swapping'). (c) is the unit cell flipped over from the view in (b), showing the Sellers et al. reinterpretation of the location of the motor domains and GTDs (Sellers et al., 2008). In this interpretation the motors are seen interacting with the GTDs of the same molecule, consistent with the images in (a). (d) shows the reinterpretation in (c) flipped back over for comparison with the view in (b). Images in (a)-(d) are adapted from Sellers et al. (2008).

3.4.3 Construction of new atomic models

The models in this section were constructed using *UCSF Chimera* (v1.4) molecular modelling software (Resource for Biocomputing, Visualization and Informatics, University of California, San Francisco, USA).

Owing to its good fit in the 3D density map and the close resemblance in the lever region to the HMM average, the Liu et al. left head model was chosen as a starting point for a new model. The left head of the two was chosen for comparative reasons, as it is the head also seen with a face-in-profile-like motor in the negative stain images of individual folded molecules, Figure 3.6(a). Moreover, as established above, there is no substantive difference between left and right head *lever* regions in the two-headed Liu et al. model.

For brevity, the Liu et al. left head model will henceforth be referred to simply as the Liu model.

3.4.3.1   Apo head model by lever-helix alignment

To begin modelling, the X-ray structure of myosin 5a apo motor plus first IQ motif and bound ELC (PDB: 1OE9) (Coureux et al., 2003), was superposed onto the Liu model, by aligning the backbone atoms of the lever helix of the two structures (with resulting RMSD of 1.3 Å for 156 atom pairs). The result, depicted in Figure 3.7 and Figure 3.8, has the lever helices and the ELC and CaM of the two models well aligned, while a difference in the position and profile of motors is evident. Since, as shown later in Figure 3.11, myosin 5a-apo, -ADP and -ADP.BeF$_x$ motor structures are actually very similar in overall structure, the difference with the Liu motor seen here must be due largely to the flexible fitting movement applied to the ADP motor by Liu et al.
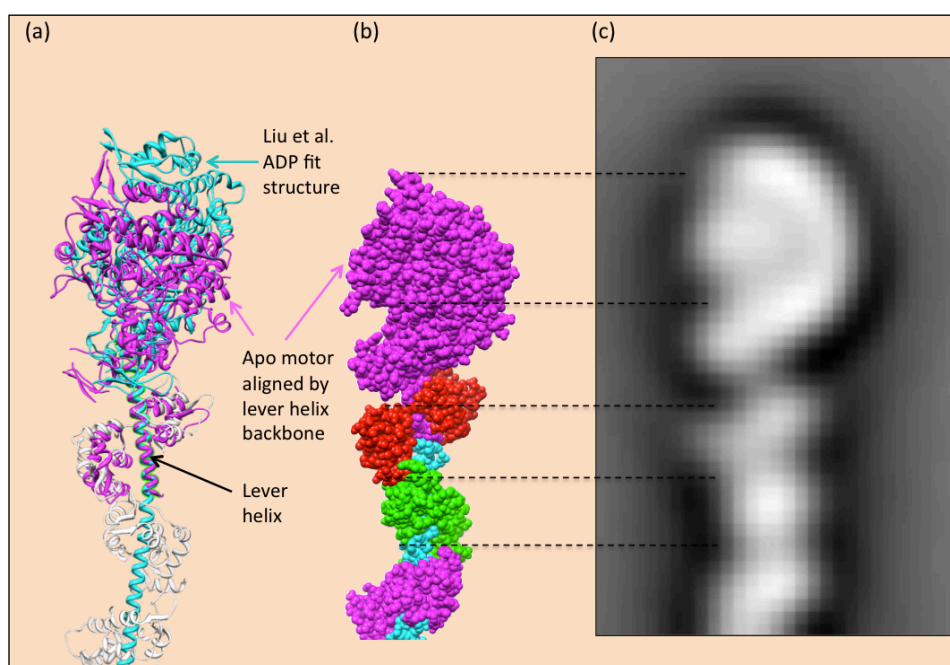
Figure 3.7: Apo motor aligned by lever helix to the Liu et al. 'left head' model.

(a) ribbon depiction of the apo motor plus ELC (PDB: 1OE9) (magenta) aligned to the lever helix backbone of the Liu left head model (heavy chain cyan, CaMs white). (b) shows a spacefill depiction of the lever-aligned apo motor (magenta) and first three CaMs of the Liu model lever. (c) is a section of the EM average of the HMM head for comparison to (b). The dashed lines between (b) and (c) are to guide the eye between structural features in the model and the average. Note that the long axis length of the apo motor is shorter in the model than the average, and the cleft between lower 50k and SH3 domains is lower down.

The profile of the lever helix-aligned apo motor has a closer, though still not exact, resemblance to the HMM EM average than does the Liu flexibly-fit motor. In particular, the cleft between the left side SH3 domain and lower 50k protrusion is evident in the lever-aligned apo motor, as is a more curved right hand side of the motor. However, the

cleft is not in line in the apo model with its position in the HMM average (it is lower in the model), and the top of the model motor is also lower than in the average (see dashed guide lines between Figure 3.7(b)-(c)). The apo modelled-motor appears shorter and squashed in the long axis direction compared to the EM average.

A 90° rotation of the new apo model about the lever axis is informative, Figure 3.8(a). The motor's long axis is seen to make a significant angle (> 45°) with the lever helix axis. This means that in the face-in-profile view, the motor actually projects upwards out of the plane of the field of view, thus foreshortening the length of the motor's 'face'. Figure 3.8(b) shows that, when viewed in the cryo-EM density map, the angled apo motor also projects significantly out of the density envelope.
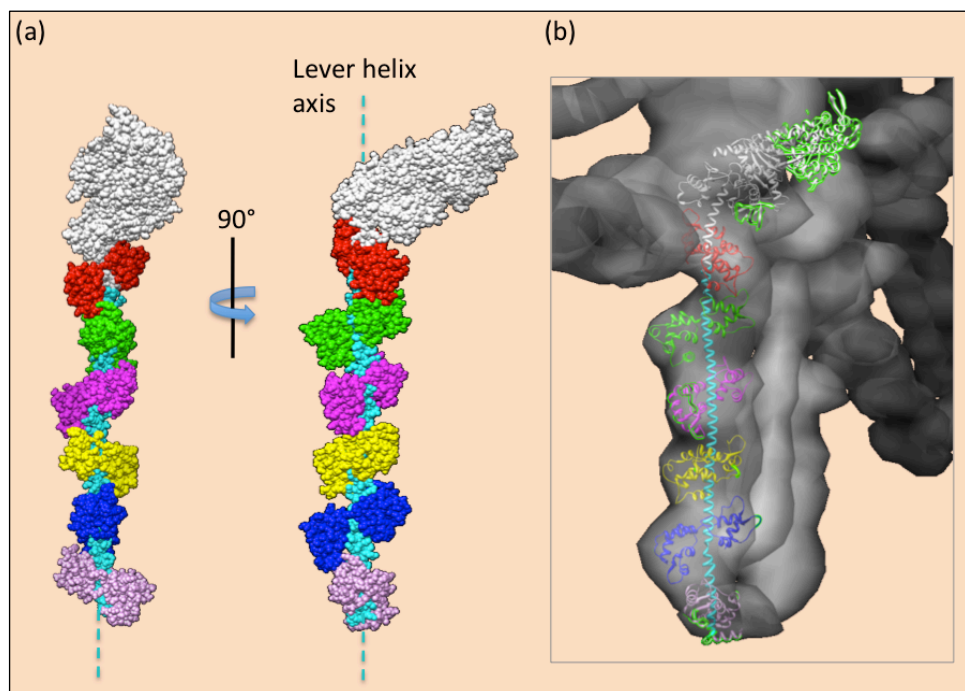
Figure 3.8: Face-in-profile and side views of lever-aligned apo model.

(a) two spacefill views of the head model with apo motor (white) aligned to the lever helix of the Liu model (cyan). On the left is the familiar the face-in-profile view of the head and on the right is a view rotated 90° about the lever helix axis, showing how the motor long axis makes an angle with the helix axis. (b) ribbon view of the 90°-rotated view in (a) within the cryo-EM density map (grey) of Liu et al. A significant part of the top of the motor (green outline highlight) projects out of the density map.

When adsorbed to the carbon substrate of the EM grid, and dried in the presence of stain, it is unlikely that the HMM head retains an upward-projecting motor structure like the apo aligned model. In negative stain EM, surface tension forces of the evaporating stain film are expected to push down on the molecules as the grid is dried (see 1.9.1). Consequent flattening and distortion of molecules is a known phenomenon (Boisset et al., 1990; Cheng et al., 2006; Knight and Trinick, 1984; Radermacher et al., 1994). Figure 3.9(a)-(b) illustrates that if the apo aligned motor is tilted backwards (pivoting about the motor-lever

junction) to make a flatter model, the resulting face-in-profile view would have a very close resemblance to the EM average. This is evidence that such flattening of the head on the EM grid indeed occurs. It is noticeable however, that the tilted motor domain of the flattened model is now tilted beyond the density for the motor region in the Liu et al. cryo-EM density map, Figure 3.9(c). Thus, in solution, the true position of the motor domain may, in fact, lie somewhere between the upward-projecting aligned apo motor and the flattened molecule that matches the EM average.
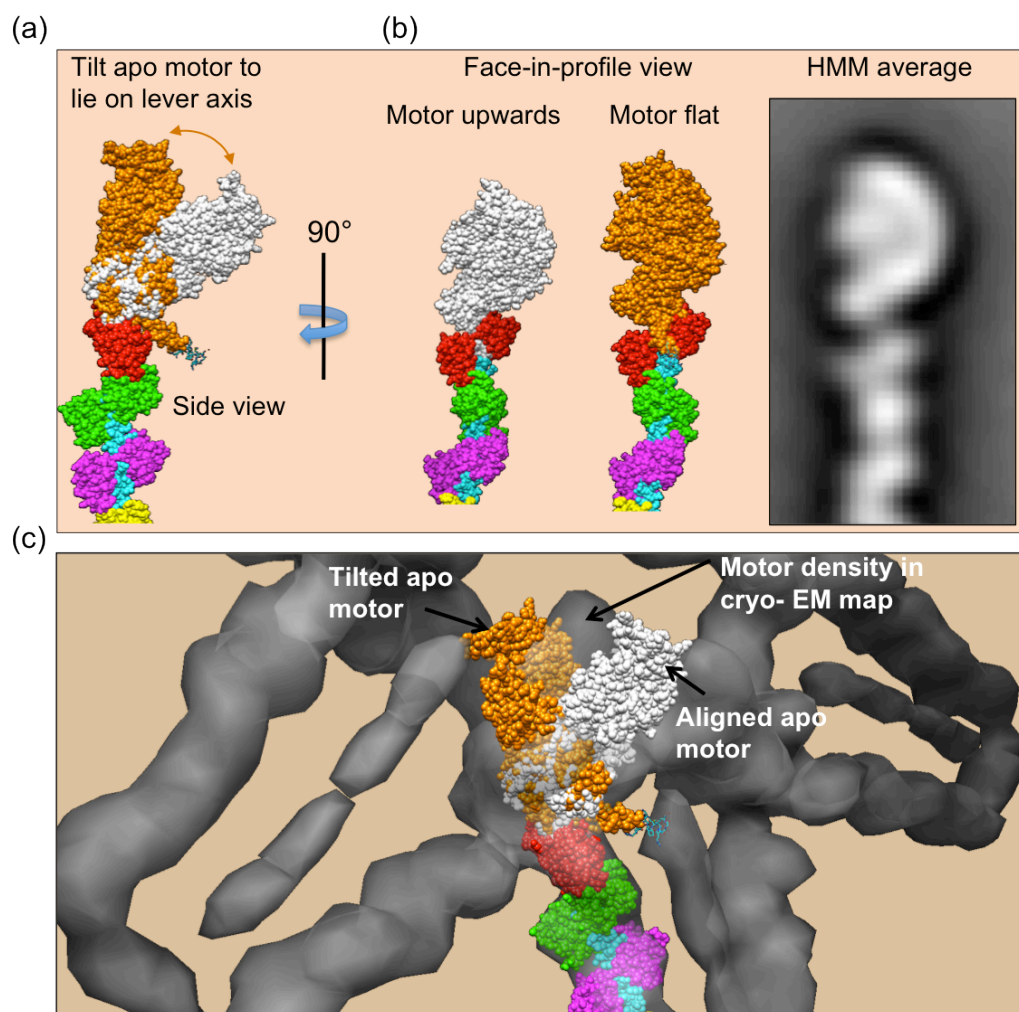
Figure 3.9: The effect of tilting the motor domain to make a flattened head.

(a) side view of the top part of the apo motor head model (motor in white), and the same motor structure in orange, tilted backwards (pivoting about the motor-lever junction) to lie 'flat' *i.e.* with its long axis in line with the lever axis. (b) resulting face-in-profile views (a 90° rotation) for the structures in (a), with the HMM average on the right. (c) the two structures in (a) located in the Liu et al. cryo-EM density map (grey).

3.4.3.2   Apo head model with independently fit motor domain

For comparison, and as a quantitative test of the Sellers et al. coneflower motif reinterpretation, an alternative head model, constrained by the Liu et al. density map, but following the Sellers et al. reinterpretation of the motor location was also constructed.

The apo motor structure (PDB: 1OE9) minus its ELC, was first positioned approximately at the Sellers et al. reinterpreted motor location in the Liu et al. density map, Figure 3.10(a), and then fitted more accurately using the *UCSF Chimera* 'fit in map' function (which uses steepest-ascent local optimization) to find the optimum local fit to the density. An optimum local fit was found for the motor - fitting 6,185 atoms of the motor, with 716 atoms left outside the confines of the map when contoured to contour-level 0.9, Figure 3.10(b).
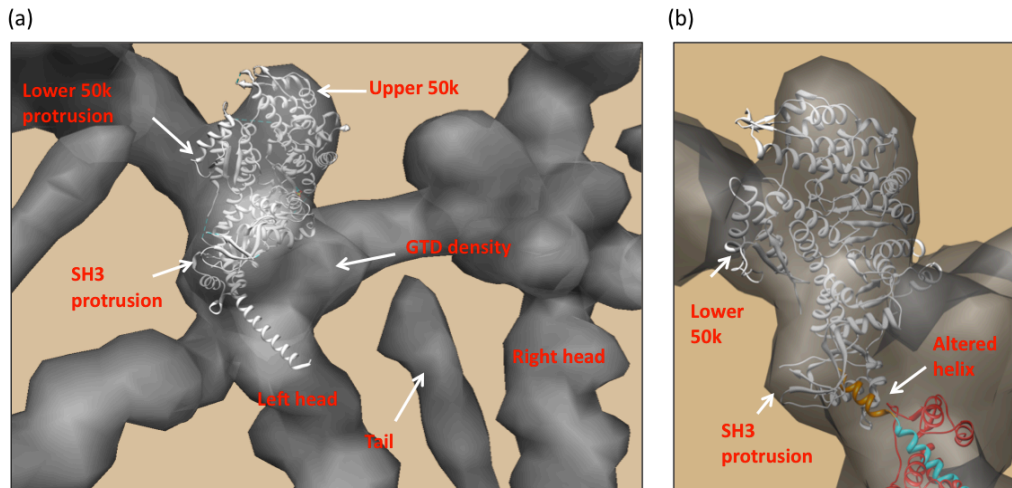
Figure 3.10: Local fitting of an apo motor structure into the Liu et al. cryo-EM density map.

(a) the apo motor domain structure (minus ELC) positioned, by-hand, in the Liu et al. cryo-EM density map (grey), in the region reinterpreted in Sellers et al. (2008) to most likely accommodate the motor domain. The motor is positioned so that the SH3 and lower 50k protrusions fit well into protruding lobes of density. **(b)** a close up of the motor domain in the density map following local optimisation fitting using *UCSF Chimera*. The motor has moved down and rotated a little from (a). To marry the locally fit apo motor to the Liu et al. lever helix (cyan), a section of the Liu lever helix (residues 755-767, orange) was moved, see main text.

To test whether different nucleotide state motors fitted better in the map, the same local fitting procedure was repeated using ADP.BeF$_x$ (PDB: 1WJ7) and ADP (PDB: 1W7I) motor structures. Results are compared in Figure 3.11. Though each motor was fitted independently (from different initial starting positions), each found a similar local minimum position in the density map. Figure 3.11 also shows that the different motor structures superpose closely and that there are no gross conformational differences between them (smaller differences are limited mostly, as would be expected, to helices and loops around the nucleotide binding site). The independently fit motors make a more satisfying fit to the density map in the reinterpreted location than the original Liu et al. interpretation. In particular the SH3 'chin' and lower 50k 'nose' regions occupying protruding lobes of the density map well. From this, it is concluded that the Sellers et al. reinterpretation of the motor location is preferable to the original Liu et al. assignment.

However, the new fits are still not perfect, with some surface loops of the upper 50k region (left side) and helices near the nucleotide binding region (right side) protruding beyond the confines of the map (at contour-level 0.9) (see also Figure 3.14).
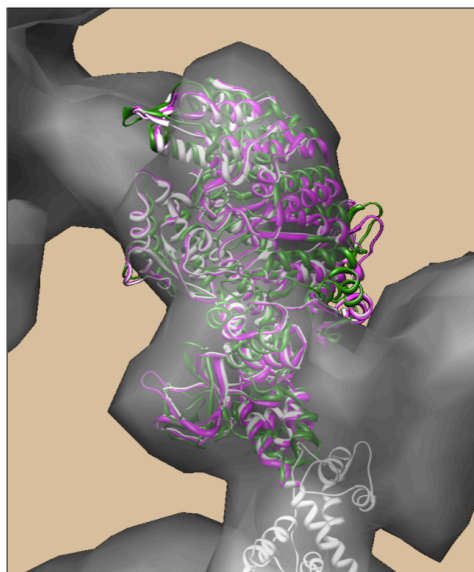
Figure 3.11: Different nucleotide-bound motor structures fit to the cryo-EM density map.

Ribbon depictions of different motor structures independently fit to the Sellers et al. reinterpreted location for the motor domain in the Liu et al. cryo-EM map (grey). Apo=white, ADP.BeF$_x$=green, ADP=magenta. Each motor was moved to the approximate location in the map and fit using *UCSF Chimera*. Although the SH3 and lower 50k regions (lower and mid left side) fit well in the density, loops and helices still extend from the map (at contour-level 0.9) at the top left (upper 50k) and mid right (above nucleotide-binding and GTD-binding regions).

Viewing the locally-fit apo motor and the lever section of the original Liu model together, Figure 3.12(b), one can see that the locally-fit motor also has a close resemblance to the EM average face-in-profile appearance. There is still some motor foreshortening (follow dashed guide lines between Figure 3.12(b)-(c)) due to the domain tilting up out of the plane of the page, but the effect is reduced compared to the lever-aligned apo model in Figure 3.7(b). As discussed above, the foreshortening effect is probably due to flattening of the molecule during to the grid drying process.

While the independently fit motor is a now a good match to the EM average, Figure 3.12(a) shows an obvious difference in the path of the Liu lever helix and that of the locally fit motor. However, the helices of the two structures do come close to intersecting at a point between the motor domain and the first light chain (residue 767 of the Liu helix and residue 771 of the motor structure). This identifies a region at which a bend in one of the lever helices could be introduced to marry the two structures to make a complete head model. Accordingly, the structures were fused by redirecting four turns of the Liu model lever helix (residues 755-767) to meet the helix of the locally-fit apo motor at residue 755, Figure 3.13.
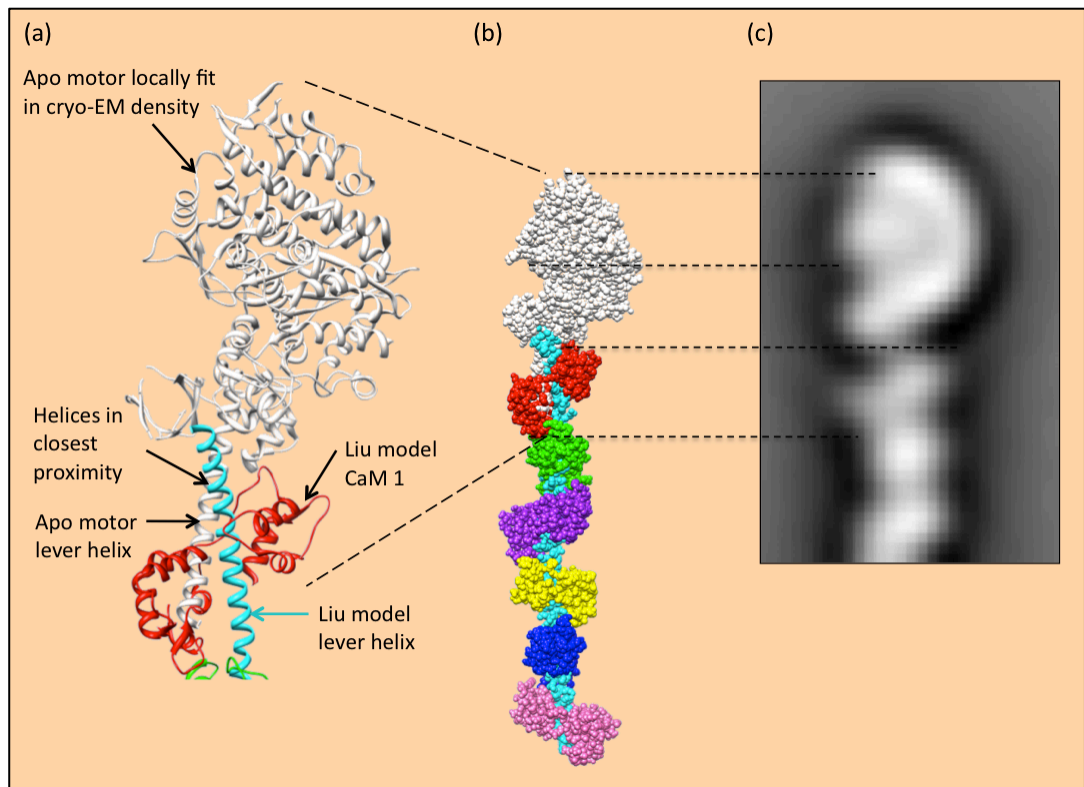
Figure 3.12: Views of independently-fit apo motor and Liu lever model.

(a) ribbon depiction of the lever helix (cyan) and first CaM (red) of the Liu model and the apo motor structure (white) fit to the cryo-EM map of Liu et al. according to the Sellers et al. reinterpretation. Note that the paths of the Liu lever helix and lever helix of the motor are not coincident, but they do come close to intersection at ~ residue 767 of Liu helix. (b) spacefill view of the complete structures in (a). (c) section of the EM HMM average for comparison with (b). Dashed lines between (b) and (c) are guide lines for the eye to compare structural features.
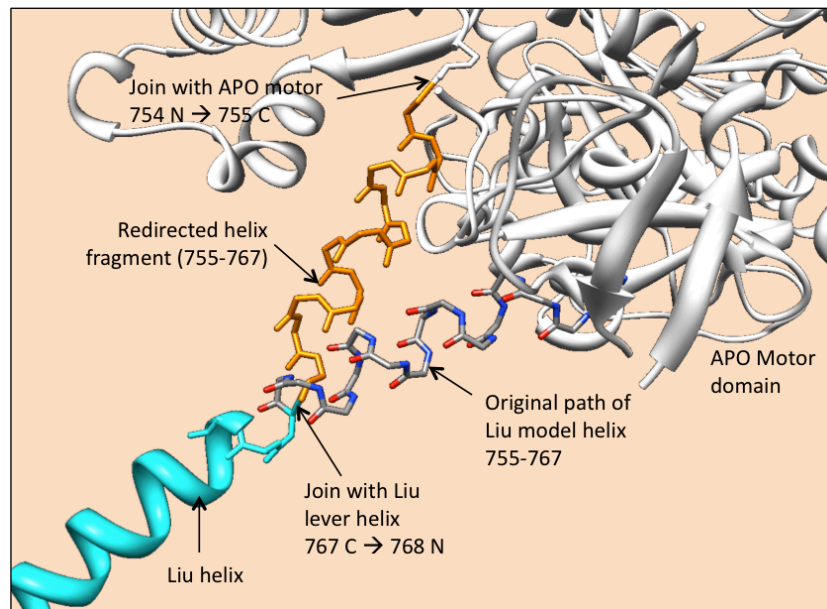
Figure 3.13: Bending of the Liu lever helix to meet the fitted apo motor.

The Liu lever helix is shown bottom left (cyan) and the locally-fit apo motor (white) is show upper right. To marry the two structures, four turns of the Liu helix (residues 755-767) were moved from their original path to a new position (shown in orange) to meet the start of the apo motor's lever helix.

The final atomic model of the head produced is shown in Figure 3.14. It is formed from the locally-fit apo motor, implementing the preferred Sellers et al. coneflower reinterpretation, and joins, via a lever helix bend between motor and IQ motif 1, to the Liu model lever. The whole structure fits well within the 3D cryo-EM density map of Liu et al. and makes a close match to the new EM average of the HMM head. The slight lengthening of the motor in the EM average can be accounted for by head-flattening during adsorption and drying on the EM grid substrate.
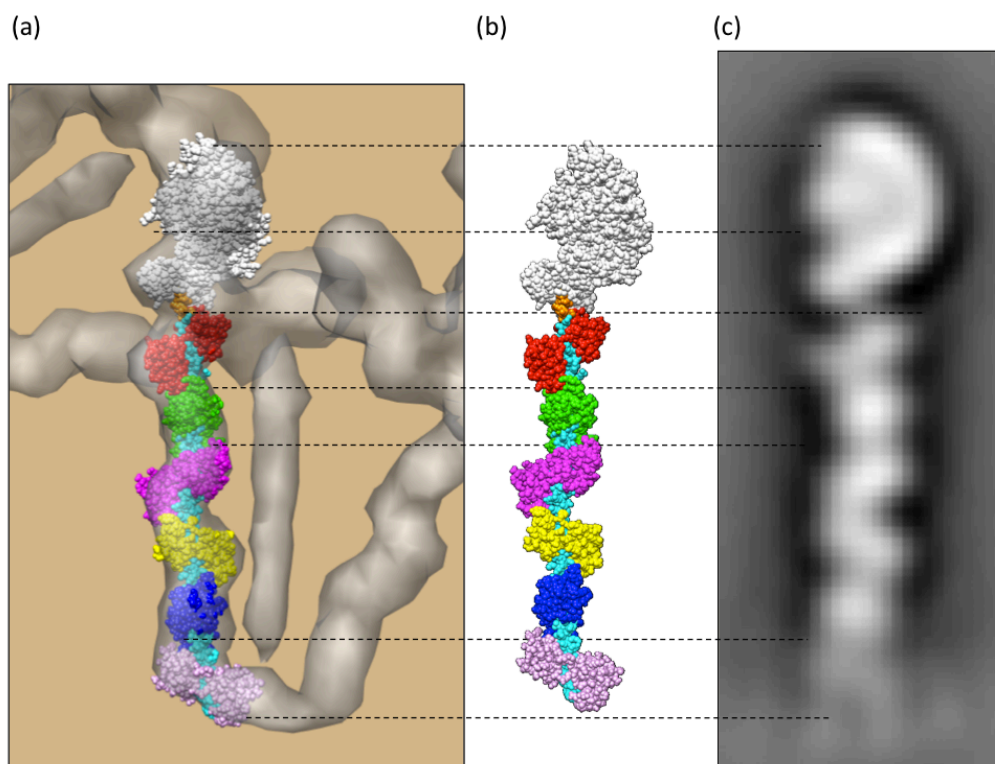
Figure 3.14: View of the final head model in the Liu et al. cryo-EM density map.
(a) the complete apo motor-Liu lever head model in the left head location of the Liu et al. cryo-EM density map, orientated in a view that matches the HMM average shown in (c). (b) the head model without the 3D EM map. Dashed lines are to guide the eye between corresponding structural features of the model and average.

3.4.4 Summary and conclusions on new head models

In summary, the new (apo) HMM average from negative stain EM was compared to existing myosin 5a head models. The varying shapes of differently oriented CaMs in the EM average were inconsistent with two of the models, but were well matched by the profile of the lever section of the Liu model. From this, it can be concluded that the Liu model is a good atomic model for the *lever* in an HMM head and that, on average, the orientations of the CaMs and overall appearance of the lever are not significantly altered when the molecule changes from an open active state to a folded inhibited state. Moreover, there is no evidence for major lever distortion caused by staining and drying on the carbon surface of the EM grid.

In contrast, the profile of the motor in the Liu model was inconsistent with the face-in-profile appearance of the average HMM motor. Therefore, two new model types with adapted motor domain orientations were constructed, to try to better match the HMM appearance. One used an apo motor that was simply aligned by its lever helix to the Liu model. The other model type used independent motor structures (separate apo, ADP-, ADP.BeF$_x$-bound versions were made) which were locally fit to a region of the 3D density map suggested by Sellers et al. (2008). The separate motors all fitted well into the same new optimum position in the density map (there being no gross conformational differences between the structures). The apo helix-aligned model showed a closer match to the face-in-profile HMM average than the original Liu model; but its long axis was significantly tilted with respect to the lever axis and it protruded from the 3D density map, giving a foreshortened appearance. The locally-fit motor structures fitted (by design!) within the confines of the map, but also gave the best match to the HMM face-in-profile average. From this, it is concluded that the Sellers et al. reinterpretation of the motor's location in the 3D map is correct, and that this position is also very close to how the motor is seen in the HMM molecule stained on an EM grid.

To marry the locally-fit motor domains to the Liu lever section, a bend needed to be introduced into the lever helix between the motor and first light chain (at residues 755-767 in the heavy chain). From this, it can be concluded that in both the HMM head as seen by negative stain EM, and in the folded inhibited state of the molecule, there is bending of the lever helix in this region. The presence of such a pliant region of helix between motor and CaM-1 is consistent with observations of motor-lever flexing and rotational movement explored by image classification later in this chapter (3.5.1.1-3.5.1.2). It is also consistent with a pliant region at the junction between motor and lever found in other myosins, such as scallop myosin 1 (Houdusse et al., 2000).

As the model of the head using a locally-fit (apo) motor joined with the Liu lever matches the HMM average and fits well in the 3D EM map of Liu et al., it represents the best atomic model of the myosin 5a head in solution. It is noted, however, that tilting the motor domain further to make a flattened molecule, forms an even closer resemblance to the HMM face-in-profile average. From this, it is concluded that, during the specimen dehydration process, this flattening of the head on the EM grid indeed occurs.
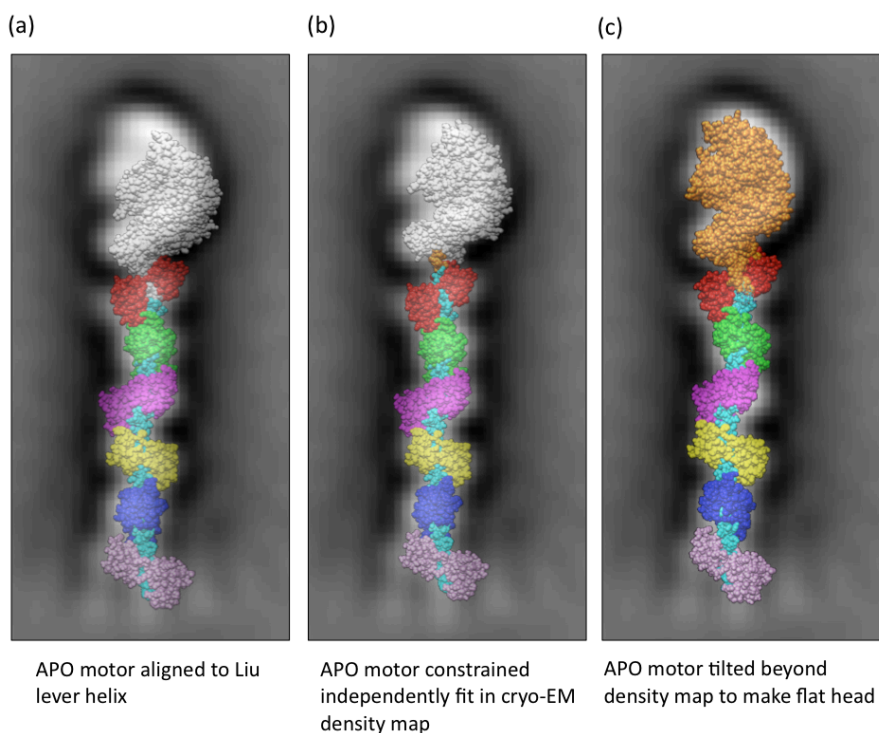
Figure 3.15: Summary of new atomic models of the apo myosin 5a head.

Spacefill views of three atomic models of the HMM head are shown superimposed on the EM HMM average. **(a)** is the model created by aligning the apo motor (white) to the lever helix of the Liu model. **(b)** is the model made by independently fitting an apo motor (white) to the cryo-EM density map and joining it by a helix bend (orange) to the Liu lever model. **(c)** a view of the head created if the upward-projecting motors in (a) or (b) are tilted back to make a flattened HMM head (with motor long axis co-planar with the lever axis).

The validity of the new 'locally fit motor–Liu lever' model of the head can be tested by comparing it to the negative stain EM averages of folded full length molecules obtained by Thirumurugan et al. (2006). Figure 3.16 shows a comparison of the model (with ADP.BeF$_x$ motor) next to a classification average of inhibited full length molecules. It can be seen that, if the head model is rotated 19° about the lever axis from the typical face-in-profile view for HMM, there is an excellent match between the appearance of both the left head's motor domain and the individual CaM shapes of the lever (which, due to the head rotation, make a slightly different pattern to that of the HMM or Liu structure views). This similarity is a validation of the new model, and reinforces the earlier conclusion that there is no gross distortion or conformational change in the motor and CaM average appearances when the molecule changes between unfolded active and folded inhibited states. Individual full length folded molecules do evidently adopt this slightly different rotational orientation (at least of the left head) when adsorbed to the carbon-coated EM grid, compared to open HMM molecules on the same surface, or in the 2D coneflower motif supported on a lipid monolayer.
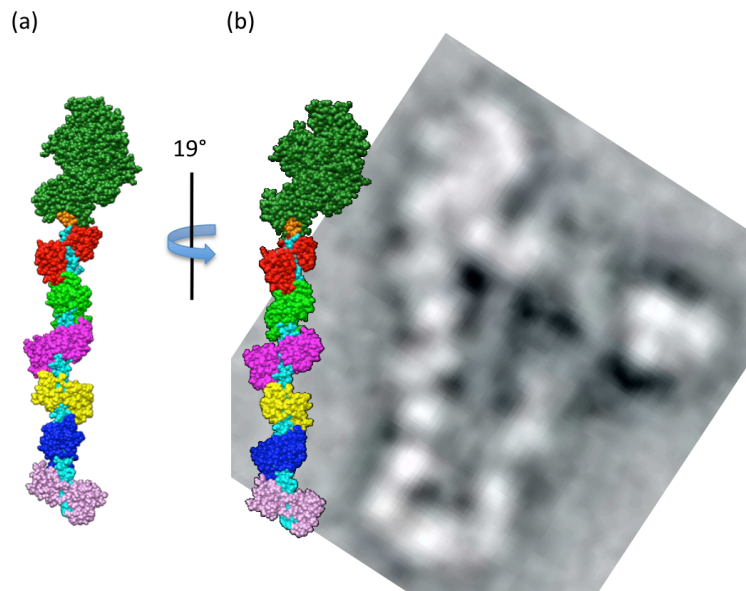
Figure 3.16: Comparison of HMM head model to EM average of folded inhibited state molecules.

(a) shows an atomic model of the HMM head with ADP.BeF$_x$-bound motor (green) (fit to the Liu density map) joined by helix bend (orange) to the Liu model lever. (b) shows the same model rotated 19° about the lever axis to a position where the appearance of the motor domain and the profiles of the CaMs are very similar to the left head of an EM class average of folded full length molecules from Thirumurugan et al. (2006). The angle made by the long axis of the motor profile with the lever (*i.e.* tilted to the right of vertical) and the cleft between upper and lower 50k regions ('forehead' and 'nose') top left of the motor match the average in particular. Note that the lever in the EM average appears to bow inwards towards the tail, a sign of lever bending flexibility.

Finally, one important observation from the comparison in Figure 3.16 is that the overall shape of the lever, *i.e.* the line followed by the lever helix, is clearly more curved in the folded molecule average (bowing inwards towards the folded-back tail) than in the straighter HMM head model, which is an average over many more individual conformations. In fact, a range of lever curvatures were observed in the different averages obtained in the folded molecule study of Thirumurugan et al. (2006). This is evidence for significant bending flexibility in the lever domain in the inhibited-state folded molecule. A rigorous study of lever bending in the free HMM molecules is the subject of chapter 4.

3.5 Investigation of flexibility in the myosin 5a head by image classification

Armed with the new atomic model of the myosin 5a head, this section returns to the stack of individual HMM head images aligned by their lever domains (see 3.3.3). An investigation of conformational variability within the images is undertaken using image classification.

To begin, various classification methods were first compared (work not included in this thesis), using a sample of the myosin 5a-HMM dataset. Methods tested included K-means clustering (implemented in *SPIDER*) (Frank, 2006), hierarchical ascendant classification (implemented in *IMAGIC*) (van Heel, 1984), and a maximum likelihood-based method (Scheres et al., 2007). Of these, *SPIDER*'s K-means clustering algorithm was found to be the most effective, adding to prior experience within the EM group at University of Leeds that suggests this method is particularly suited to classification of molecules, such as myosin, which exhibit continuous flexibility (Burgess et al., 1997).

Given the structural detail seen in the motor and the resolution of differently-oriented CaMs in the HMM average (Figure 3.3), it was decided to first investigate the extent of variability of appearance of these domains individually. This was, in particular, to search for evidence for any degree of torsional flexibility between motor and lever, and within the lever domain itself. Does the HMM head lie in preferred orientations on the EM grid? Are CaMs locked in particular orientations, or do they rotate about the lever helix axis? If so, to what extent can each rotate? Are particular CaM orientations correlated along the lever and is there correlation with the motor domain orientation?

Some published data exists on the myosin 5a (apo) motor adopting different orientations on carbon-coated EM grids exists (Burgess et al., 2002), and motor-lever angular flexibility has been studied as part of a temperature study of myosin 5a-S1 (Thirumurugan et al., 2003). However, classification of individual CaM appearances has, as yet, not been undertaken.

3.5.1 Motor domain classification

To begin the variability analysis, a classification focusing on the motor domain was performed. The 23,250 HMM head images, aligned by their lever domain (3.3.2), were used as the source for classifications. A mask encompassing the motor domain and immediate surrounding area (excluding CaM-1) was drawn, Figure 3.17(a), using the

global variance image (Figure 3.3(b)) as a guide to determine an appropriate width. K-means image classification was then performed using *SPIDER* (v19.05) (`CL KM` command), using the motor-only mask to define the region of interest. In exploratory classifications (data not shown), images were split into an increasing number of classes, from 46-500. Division of the data into 230 classes (~100 images/class) was judged the best comprise between encompassing variability in the images and reinforcement of structural detail through class averaging. Figure 3.17(b) shows a montage of a selection of averages from the classification, representative of the range of motor appearances.

(a)

(b)   (c)
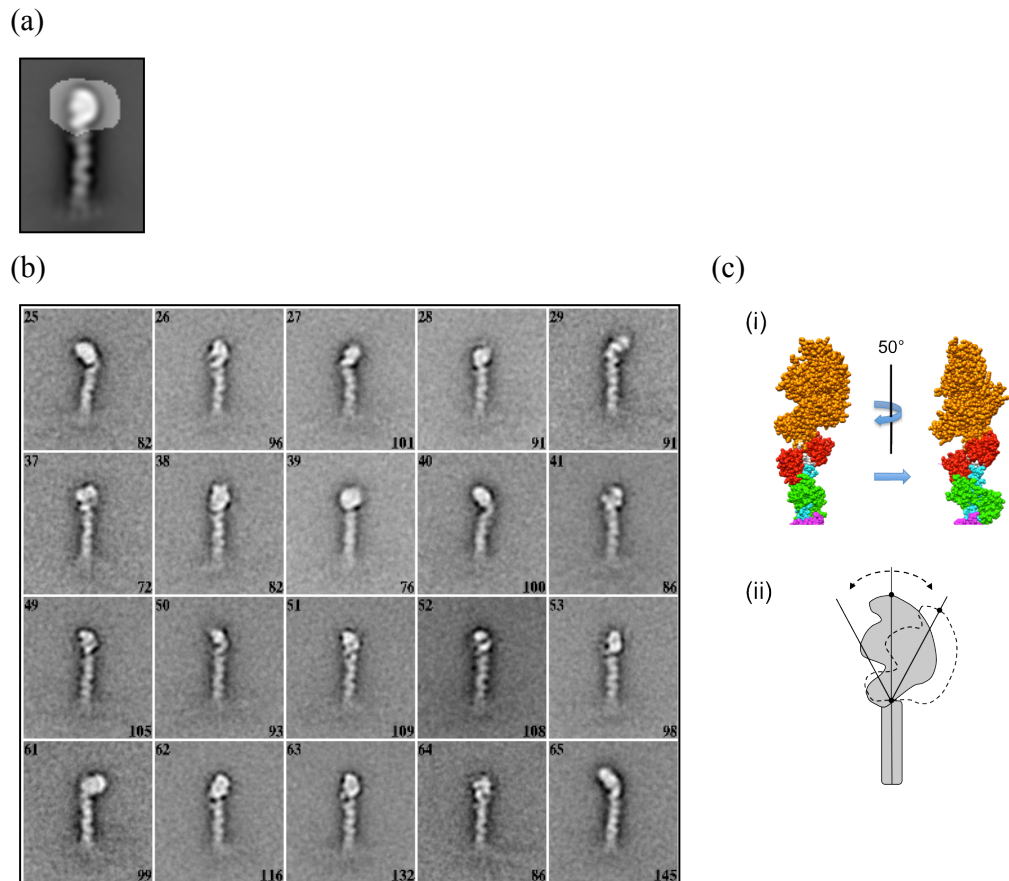
Figure 3.17: Motor domain classification of HMM head images

(a) global average with motor-only classification mask superposed (white transparent region). (b) montage selection from 230 class averages from a classification of 23,250 HMM head images using the mask in (a). Class numbers are top left, number of images contributing to each average, bottom right. (c) two types of variability identified in motor appearance. (i) motor long axis rotation between two principal views, (ii) motor tilting about the motor lever junction.

The 230 class averages showed a lever with a fairly consistent appearance; being straight or near-straight with CaM substructure just resolved, though naturally not as well delimited as in the global average. However, as is evident from Figure 3.17(b), there was clear variation in the motor's appearance. On close inspection of the 230 classes, two main features were identified.

(i) Two specific appearances of the motor tend to reoccur (Figure 3.18), suggesting there are two principal orientations of the motor when adsorbed to the carbon film of the EM grid. These are related by a ~50° rotation about the motor's long axis.

(ii) In either orientation, the motor's long axis makes a range of angles with the lever axis, apparently tilting about the motor-lever junction, moving in the plane of the carbon-coated EM grid (the plane of the field of view).

Both observations are indications of different types of movement at the motor-lever junction, Figure 3.17(c), consistent with previous findings from the University of Leeds EM group (Burgess et al., 2002; Thirumurugan et al., 2003). Each type of movement is analysed in more detail below.

3.5.1.1 Principal motor orientations

The two main views of the motor identified in the class averages are consistent with those in earlier negative stain studies of apo myosin 5a-HMM by Burgess et al. (2002). Each view can be well matched to a specific orientation of the motor in the apo atomic model, Figure 3.18, confirming that the motor tends to adsorb to the carbon-coated EM grid in one of two preferred orientations (though a proportion of 'unidentified' potentially intermediate views are also present in the averages). The two views are related by a rotation of ~50° about the head's long axis. In one view, hereafter termed 'View-1' or 'face-in-profile', the motor has the face-in-profile appearance most similar to the global average, while in the other, hereafter termed 'View-2' or 'face-away', the motor has turned its face ~50° over its right shoulder (*i.e.* into the page) and is facing away.

The chief features identifying View-1 motors (Figure 3.18) are an overall half-circle shape, the right side being curved, with up to three protrusions as one moves up the straighter left side. By comparison to the atomic model, these are identifiable as the SH3 subdomain (lower-left) (the face's 'chin'), the lower 50k region (mid-left) (the 'nose'), and a usually less-pronounced upper 50k protrusion (upper-left) ('forehead') above the actin-binding cleft (sometimes resolved). Additionally, in View-1 motors, dark pools of stain often highlight a notch in the lower-right side of the motors' base (indicating a gap between motor and CaM-1) and also in the large cleft between SH3 and lower 50k protrusions. In View-2, the defining features are a straighter left-hand-side without or with much reduced protrusions (the cleft between SH3 and lower 50k being filled), while the mid-right side now exhibits a small protrusion, formed by part of the upper 50k region. Another characteristic is that the base of the motor is broader at the interface with the

CaM-1, lacking the right-hand-side curvature and characteristic bottom-right notch of View-1.
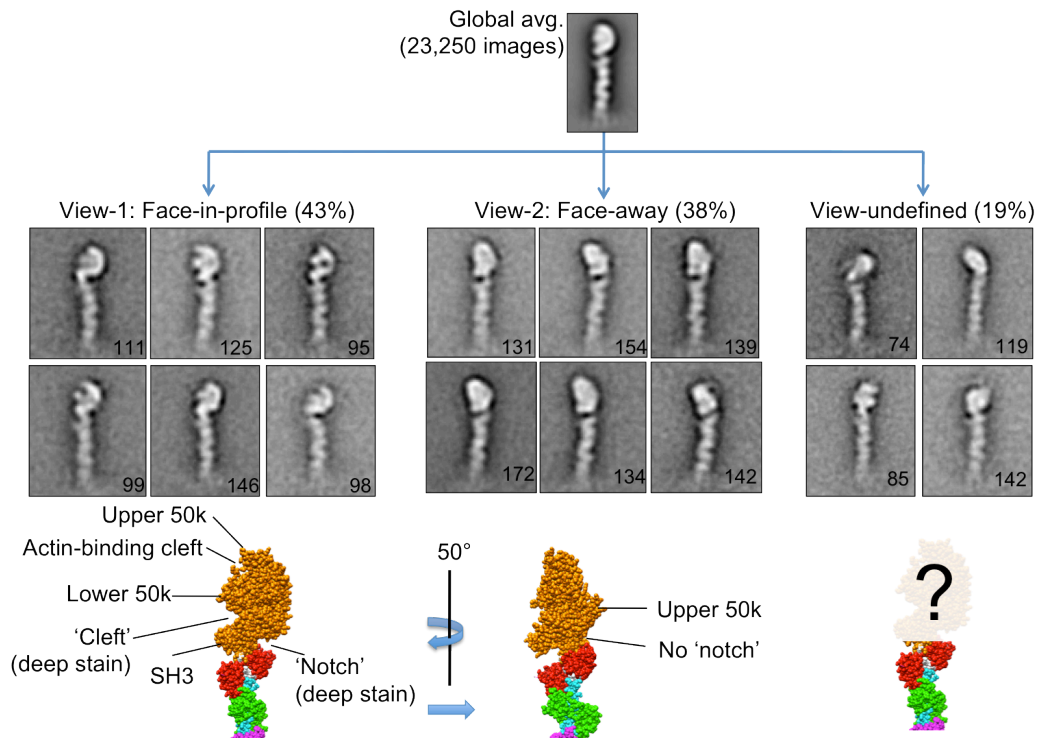
Figure 3.18: Class averages of different motor orientations

Typical example class averages for each view are shown. Numbers indicate the number of images contributing to each average. The atomic models positioned below each orientation are with the artificially flattened apo motor (see 3.4.3.1) to best match the negative stain HMM appearance.

The reason the motor mainly adsorbs to the carbon in one of the two orientations remains unknown, but may be due to favourable motor-substrate contacts in these orientations. It is also notable that the motor apparently never appears flipped over (180° rotated about the long axis), such that the 'face' looks to the right. Analysis of the hydrophobic, polar and charge properties of the motor surface do not reveal an obvious polarity in the two orientations. Of course, the motor's axial rotation may not be independent of the lever and it may in fact be the surface interaction with the lever which determines the overall orientation of the head.

The class averages of the motor domain classification were segregated into View-1, View-2 or 'View-undefined', on the basis of the characteristics outlined above. Of the 23,250 images in the dataset, the split between the principal views was quite even: 43% (101 classes, 10,119 images) identified as View-1, 38% (82 classes, 8,734 images) as View-2, and 19% (47 classes, 4,397 images) as View-undefined. All individual images contributing to each view were then averaged (*SPIDER*: AS R) to produce Figure 3.19. Interestingly,

the View-undefined average has a motor domain (and CaM pattern) that closely resembles View-1, indicating that the majority of its (poorly-aligned?) contributing images are actually probably close to the View-1 appearance, making this the overall predominant grid orientation of the two outlined.
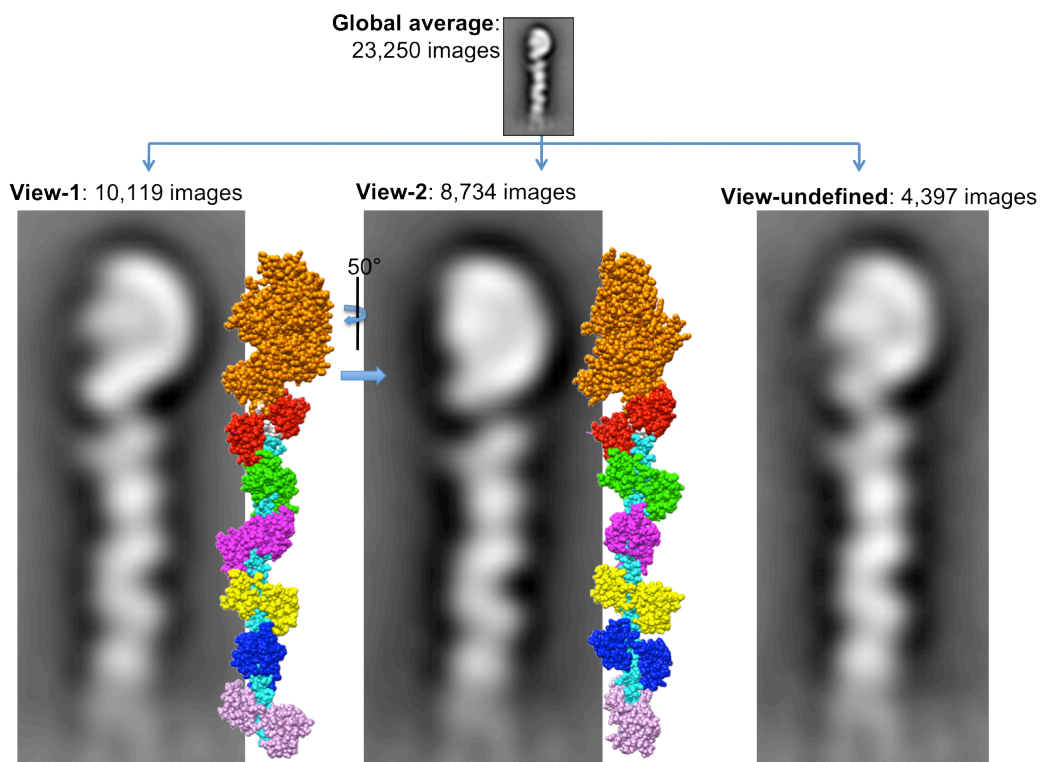
Figure 3.19: Averages of individual heads images categorised by motor view.

> The best corresponding orientations of the head atomic model (with flattened apo motor) are shown next to View-1 and View-2 head averages. Note the correspondence between motors and CaM shapes for View-1, but the clear mismatch between CaM shapes for View-2. In View-2 and View-undefined the CaM shapes remain the same as those for View-1.

Inspection of the atomic model orientations in Figure 3.19 shows that the axial rotation of a fixed head would transform the profiles of CaMs 2, 3, 5 and 6, from an appearance where both lobes are visible (slanted along a diagonal to the lever axis), to one where only one lobe is visible, or *vice versa*. Most notable would be the change in CaM-2 from single- to two-lobe views and in CaMs 3-4 where the characteristic '<' formation would be changed. However, the CaM shapes in the view averages do not change in this way. In fact, each average shows a consistent set of CaM shapes, which follow the orientation pattern for View-1. Hence, the motor appears to rotate between principal views *independently* of the lever, which, on average, appears to adopt a fixed axial orientation on the grid. This picture is also borne out by closer examination of the CaM shapes in the individual class averages (*e.g.* in Figure 3.18). Whilst there is evidence in some class averages of variation in individual CaM profiles along the lever, the variations are not

obviously correlated with different motor domain appearances. Rotation of individual CaM-heavy chain sections within the lever is investigated in detail in 3.5.2.2.

These observations indicate that the motor domain can rotate at least 50° about its long axis, independently of the lever domain. The motor-lever junction, the length of HC α-helix between motor and CaM-1, can apparently act (at least partially) as a rotary joint, the helix presumably deforming or melting locally to allow the movement. Such movement in the head may be important for myosin 5a function, for example in allowing the molecule to find particular F-actin subunits at different azimuths (*e.g.* when crossing over filaments or on occasions when there is a degree of helical disorder in the filament).

Identification of such a pliant region is consistent with the location of the deformation required to reconcile the positions of the motor and lever sections in building the atomic model (see Figure 3.13), though in that case, the movement of the motor was more a rotation about the motor-lever junction rather than rotation about the lever axis. Such tilting about the motor-lever junction is investigated in the following section.

3.5.1.2 Motor tilting about the motor-lever junction

Perhaps more obvious than the motor's axial rotation in Figure 3.17(b) and Figure 3.18, is the different angles the motor's long axis makes with the lever axis. In both View-1 and View-2 orientations, the motor domain appears to tilt about the motor-lever junction, in the plane of the field-of-view. Motors appear in a range of positions, the long axis tilting both left and right of the straight lever position (*i.e.* with motor long axis in line with the lever axis - vertical orientation in the images). Since no nucleotide was present in the sample, the relative motor rotation is likely driven by thermal energy (fluctuations from Brownian motion), rather than resulting from any nucleotide-dependent pre- to post-powerstroke conformational change (which would also involve a more marked change in converter position relative to the SH3 subdomain (Thirumurugan et al., 2006)). In this way, the motor-lever junction appears also to act as a hinge; again, the α-helix presumably bending or melting locally to allow the movement. Hinge-like movement at the motor-lever junction is consistent with images of two-headed myosin 5a-HMM bound to F-actin where (strained) lead levers sometimes appear straight rather than bowed (Burgess et al., 2002).

To quantitate the motor tilting, the distribution of angles formed between motor long axis and lever were measured in the 230-class classification. *ImageJ* software (v1.43) (NIH, USA) was used to measure the angle between two straight line segments connecting three points positioned to define the motor long axis (the line between the middle of the motor-lever junction and the furthest point on motor circumference) and the image vertical, see Figure 3.20(a). Note; for consistency in marking, the image vertical was used as the

implied lever axis rather than attempting to interpret the line of the lever in each class average in turn. This is because variability in the appearance of the start of the lever makes interpretation of its precise angle of projection difficult and each lever average is, in any case, formed from a distribution of differently curved shapes, see chapter 4. For each measurement, the angular deviation about the vertical position, the 'tilt angle', was calculated (*i.e.* 180° − motor-lever angle, Figure 3.20(a)). Angles with motors tilting left of vertical were defined as negative and those right of vertical positive. The tilt angle frequency distributions obtained for both motor orientations are plotted in Figure 3.20(b).
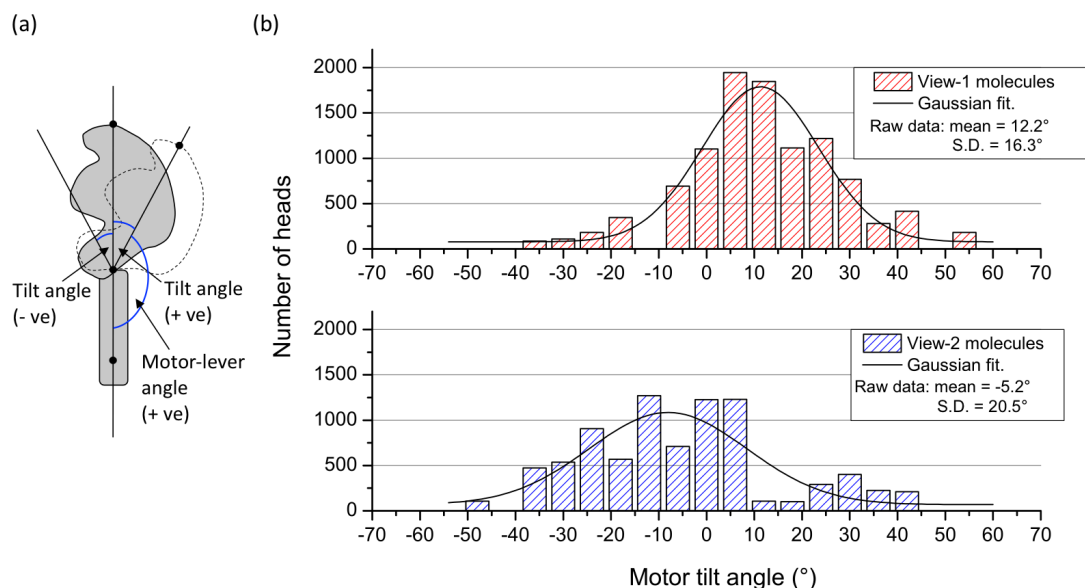
Figure 3.20: Motor-lever tilt angle distributions for View-1 and View-2 orientations.

(a) cartoon of molecule head showing motor tilt angles measured. Angles left of vertical axis are negative and right of vertical positive. **(b)** histograms of View-1 and View-2 tilt angles are shown in the upper and lower plots respectively. Frequencies are calculated from measurements of class average images using a 6° bin size, and have been converted to a number of heads by multiplication by the number of images contributing to each average. Gaussian fits to the (pre-binned) data are shown on the histograms. Note; horizontal and vertical axes are at the same scale to allow comparison of the mean position and magnitude of the frequencies between the views.

The tilt angle distributions exhibit a number of features. Both show broad distributions; View-1 ranging from −37° to 56° and View-2 from −45° to 45°. The mean angle (accounting for class size weighting) is different in each view, 12.2 ± 16.3° for View-1 and −5.2 ± 20.5° (weighted means and S.D.s). Thus, View-1 motor's are more usually tilted to the right of vertical while View-2 is closer to a straight position. This is consistent with the atomic model and the averages presented in Figure 3.19. View-1's distribution is roughly symmetrical and Gaussian-shaped, suggesting thermal fluctuations about an equilibrium angle. View-2's distribution is not Gaussian. It does steadily rise from −45° to a broad peak (~ −14° to 8°), but for greater positive angles there is a sudden drop in frequencies. This profile is more suggestive of thermal fluctuation about the mean, but with right-of-

vertical tilting blocked. One can speculate that this might be due to a steric clash between motor and CaM-1 in this orientation, *e.g.* the N-lobe acting like a 'back stop' to the right, which would be consistent with the broader 'neck' appearance of the motor at the CaM-1 interface seen in View-2 (*i.e.* no lower-right 'notch' as in View-1), see Figure 3.18. However, an initial inspection of the individual class averages does not confirm this idea, so the reason for the 'half-Gaussian' distribution shape is currently unclear.

An effective 'spring constant' for the rotation about the motor-lever joint can be approximated from the distributions by applying the equipartition of energy theorem (Reif, 1965), which equates quadratic terms in the average energy function to thermal energy (specifically, $k_b T/2$ per degree of freedom – k_b Boltzmann constant, T absolute temperature). If the elastic potential energy due to motor tilting U_{tilt} is modelled by a simple harmonic form (a good approximation for the Gaussian-shaped angle distributions),

$$U_{tilt}(\theta) = \frac{k(\theta - \theta_0)^2}{2},$$
(Eq. 3.1)

where θ is the tilt angle, θ_0 the mean angle and k a torsion spring constant, equipartition theorem yields an expression for the average energy

$$\overline{U}_{tilt} = \frac{k}{2}\left\langle (\theta - \theta_0)^2 \right\rangle = \frac{k_b T}{2},$$
(Eq. 3.2)

where angle brackets denote an average over all conformations. $\langle(\theta - \theta_0)^2\rangle$ can be equated with the variance σ^2 from the tilt angle distribution data. Solving for k gives

$$k = \frac{k_b T}{\left\langle (\theta - \theta_0)^2 \right\rangle} = \frac{k_b T}{\sigma^2} \approx \frac{4.1 \text{ pNnm}}{\sigma^2},$$
(Eq. 3.3)

for $T = 293$ K (an estimated value for the EM data).

Applying Eq. 3.3 to the tilt angle data yields a torsion spring constant (or torsion elastic modulus) for the motor-lever joint of 51 pN·nm·rad^{-2} in View-1 and 32 pN·nm·rad^{-2} in View-2. These values are of similar order to (but marginally stiffer than) measurements made in single-molecule fluorescence microscopy experiments of surface-tethered myosin 5a heads labelled with beads (Shiroguchi et al., 2011). In that study motor-lever torsion spring constants of 23 pN·nm·rad^{-2} and 20 pN·nm·rad^{-2} were found for the pre- and post-powerstroke head conformations respectively. The alternative values obtained from the EM image averages here are the first motor-lever torsion spring constants inferred from images of myosin molecules.

Knowledge of motor-lever joint-stiffness is important, as it quantitates the possibility of bending at the motor-lever junction in a myosin 5a molecule with both heads bound to F-

actin. Also, it is an additional non-trivial source of compliance in the head (in addition to cantilever-like lever bending) that is less frequently considered in actomyosin cross-bridge models. Inherent motor-lever joint flexibility may also explain the rather broad range of motor-lever angle values obtained for different myosin 5a states by different measurement techniques (Forkey et al., 2003; Lewis et al., 2012; Sun and Goldman, 2011; Walker et al., 2000).

The extent of lever rotation permitted purely by motor-lever junction flexibility can be gauged by example. For instance, assuming a rigid lever and the View-1 constraints, the energy required to rotate the lever through a large angle *e.g.* 45° (0.7854 rad) relative to a fixed (F-actin-bound) motor, is

$$
\begin{aligned}
U_{\text{tilt}} &= \frac{k(\theta - \theta_0)^2}{2} = \frac{51 \text{ pN} \cdot \text{nm} \cdot \text{rad}^{-2} \times (0.7854 \text{ rad})^2}{2} \\
&= 15.6 \text{ pN} \cdot \text{nm} \\
&\approx 3.8 \cdot k_{\text{b}} T,
\end{aligned}
\tag{Eq. 3.4}
$$

That is, almost four times typical thermal energy for a substantial lever rotation. Hence such large rotations are likely to be rare events. The typical angular deviation expected from thermal energy alone (4.1 pNnm) can be calculated as ~23°.

Figure 3.21(a) depicts a simple model of an actomyosin 5 cross-bridge, such as that typically assumed in a three-bead optical trap assay. If the motor is assumed to maintain a fixed orientation when attached to actin, then when the cross-bridge is strained (*e.g.* by transverse movement of a suspended actin filament), the cross-bridge can deform either through lever rotation about the motor-lever junction (Figure 3.21(b)), or through cantilever-type lever bending, (Figure 3.21(c)), which assumes that the proximal end of the lever (that nearest the motor) cannot rotate. In reality, if both mechanisms are a source of compliance in the head, then the two will combine (e.g. in series) to give an overall cross-bridge compliance.
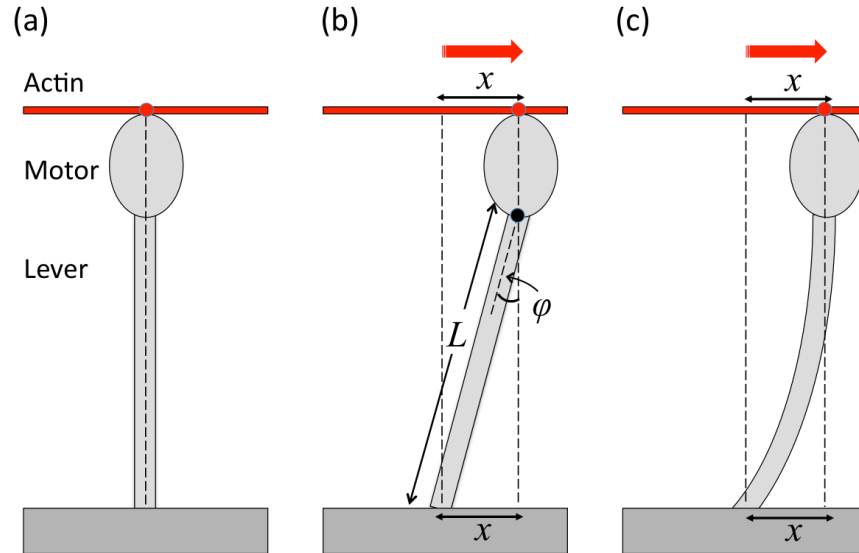
Figure 3.21: Cross-bridge deformation model

(a) simple model of a myosin 5a cross-bridge in a single-molecule trap assay. The motor (oval) is bound to a fixed point on a suspended actin filament (red), and is assumed to maintain a fixed geometry during any strain event. The distal end of the lever is assumed to be bound to a fixed point on a substrate (darker grey). **(b)** a transverse force is applied by movement of the actin filament (red arrow). One way the cross-bridge can accommodate this is to flex at the motor-lever junction (black circle), the lever rotating through an angle φ, resulting in tip displacement x. **(c)** an alternative source of cross-bridge compliance could be cantilever-type bending of the lever, also producing a tip displacement x. In myosin 5a, both types of compliance appear to be present in the head.

For comparison, the motor-lever junction stiffness for myosin 5a can also be expressed as an equivalent cantilever stiffness $k_{cantilever}$ (the ratio of the force applied to cantilever tip's displacement), the measure normally quoted for cross-bridge stiffness assays. To make this comparison, the energy for lever rotation through any angle φ due to motor-lever junction compliance (calculated as in Eq 4.3), can be equated with the energy required to bend a cantilever and displace its tip through a distance x ($U_{bend} = \frac{1}{2} k_{cantilever} x^2$). Both mechanisms result in the same tip displacement, which can be calculated from simple trigonometry, $L\sin(\varphi)$, where L is the lever length external to the motor (Figure 3.21(b)).

$$U_{bend} = \frac{k_{cantilever}x^2}{2} = U_{tilt} = \frac{k_{tilt}\varphi^2}{2}$$

For $x = L\sin\varphi$:

$$\Rightarrow k_{cantilever}(L\sin\varphi)^2 = k_{tilt}\varphi^2$$

$$\Rightarrow k_{cantilever} = \frac{k_{tilt}\varphi^2}{L^2(\sin\varphi)^2} \approx \frac{k_{tilt}}{L^2}, \qquad \text{(Eq. 3.5)}$$

where the approximation $\sin(\varphi) \approx \varphi$ (for small angles) has been applied in the last step. Thus, using $L = 20.6$ nm (see 4.5.2.1), the View-1 torsion spring constant of 51 pN·nm·rad^{-2} is roughly equivalent to a cantilever stiffness of 0.12 pN·nm^{-1}. This is actually less stiff than the value of 0.2 ± 0.01 pNnm^{-1} measured for the actomyosin 5a-S1

cross-bridge by optical trapping by Veigel et al. (2002). Hence, the source of compliance that optical trap and other assays measure could actually be dominated by the motor-lever flexing movement rather than cantilever-type lever bending. However, this must also be considered in the context of the low bending stiffness measurements for the myosin 5a lever obtained in 4.5.3.1 of this thesis.

3.5.2   Individual CaM classifications

In this section, the appearances of the individual CaMs forming the visible lever structure were analysed more closely, using image classification. This work was prompted by the observation that in some class averages, individual CaMs sometimes varied from the average appearance. Were these alternative appearances due to conformational changes or distortions, or did they match to different CaM-IQ views? Were such variations indications of twisting of individual CaM-IQ 'subunits' in the lever, *i.e.* is there torsional flexibility in the lever? If so, is the degree of twisting constant along the lever's length?

For historical reasons concerning availability of data, a subset of only 12,921 of the lever-aligned HMM head images (56% of the 23,250-image total) were used for this CaM appearance analysis. This was, however, still a sufficiently large dataset to yield good results from classification and averaging.

To begin, a general classification of the images by appearance of the entire lever region was made, applying a roughly rectangular classification mask that excluded the motor, Figure 3.22. The images were divided into 120 classes, ~100 images/class, a number judged to be an optimum compromise between image variability and improved signal-to-noise averaging.

Classification averages showed (as expected from preliminary work) considerable variation in the overall shape of the lever, an indication of the significant bending flexibility of the domain. Some levers exhibited significant curves to the left or right of the vertical centre line (the axis of lever alignment for the global average), although in the majority of classes the lever appeared near or approximately straight. A rigorous analysis of lever bending and implications for flexural rigidity of the domain is the subject of chapter 4.

To pursue the CaM appearance analysis it was clearly necessary to first remove the positional variability arising from lever curvature as a variable for subsequent CaM classifications. Accordingly, only classes exhibiting straight levers (where the lever backbone was judged, by-eye, to closely follow the vertical centre line) were selected from

the 120 averages. The individual head images belonging to this subset of classes were combined to create a subset of 6,654 'straight-lever-only' images (~50% of the starting total) for further classification.
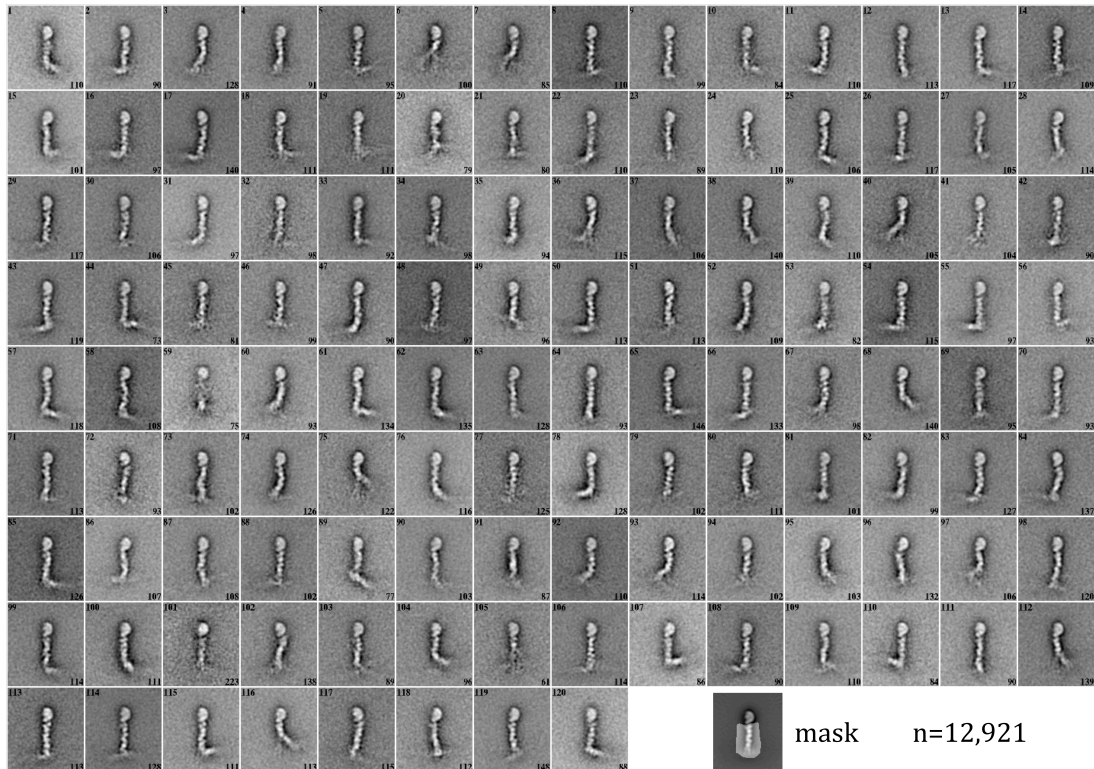
Figure 3.22: 120-class survey classification of HMM heads by lever appearance.

Montage of 120 class averages are shown. The mask used for the classification is shown bottom right superposed on the global average of the 12,921-image dataset. In each average, the number top left identifies the class and the number bottom right is the number of images contributing to the class.

The straight-lever-only images were then classified (*SPIDER* (v17.05), `CL KM` command) according to the appearance of each CaM region in turn. Using the global average as a guide, small rectangular masks (~33 × 9 pixels), covering the regions where each CaM was expected, were created. For each CaM mask, two classifications were performed; one dividing the images into 8 classes (~830 images/class) – to provide an overview of the main appearances within the data, and another dividing the images into 50 classes (~130 images/class) – to capture greater image variety within the data, while preserving signal-to-noise improvement from averaging.

Following general observations on the CaM-by-CaM analysis below, the classification results for each CaM are shown in Figure 3.24–Figure 3.29. A brief description of each CaM's characteristic features is located after each figure and a summary is then presented in Figure 3.30.

3.5.2.1   General observations

The CaM-by-CaM classification results reveal a number of interesting features. Structural details of each CaM are more clearly resolved in the individual CaM class averages than in whole-lever classification. As expected from preliminary work there is evidence of variation in the orientation of the individual CaMs. This variability has not previously been described.

Though close to the resolution limit of the negative stain method (~2 nm), staining is sufficiently good that, in certain averages, the two lobes of the dumbbell-shaped CaM are resolved as two spots, *e.g.* Figure 3.23(a). The spots can be horizontally level or staggered either side of the presumed line of the (vertically-running) lever helix, so that the dumbbell shapes make a varying angle (roughly between 60-90°) with the vertical lever axis. The dumbbell appearances can be well matched to specific views of a CaM-IQ crystal structure rotated about the lever helix axis, Figure 3.23(c). If the dumbbell is tilted, the 'higher' of the two dots (*i.e.* first encountered traversing from motor to lever tip) can be identified as the CaM C-lobe, as this always binds the IQ motif upstream of the N-lobe.

If the two-lobe view is rotated around the lever axis by ± 90° (or ±270°) the two lobes are positioned one behind the other so that only one lobe is seen. Such one-lobe views were also present in the CaM-by-CaM classification averages, visible as a single bright spot positioned directly on the presumed line of the lever helix. A brighter spot with a darker outline (deeper stain) is seen more often for the one-lobe views than the two-lobe ones (*e.g.* see classes 2-6 of the CaM-2 8-class montage in Figure 3.25), and is an indication of the increased height the two stacked lobes make when on the carbon substrate of the EM grid. A similar bold-outlining effect due to increased stain depth was noted in an earlier negative stain study, when describing the end-on view of one of the motor domains of smooth muscle myosin in its shut-down state (Burgess et al., 2007).

When rotated to a position between the one and two-lobe orientations, the CaM view broadens from a circular to an elliptical profile (with the major axis at ~45° to the lever helix), where one lobe partially eclipses the other. Occasional examples of these elliptical views are also seen in the individual CaM class averages, *e.g.* Figure 3.23(b) and classes 6 and 18 of the CaM-2 50-class in Figure 3.25.
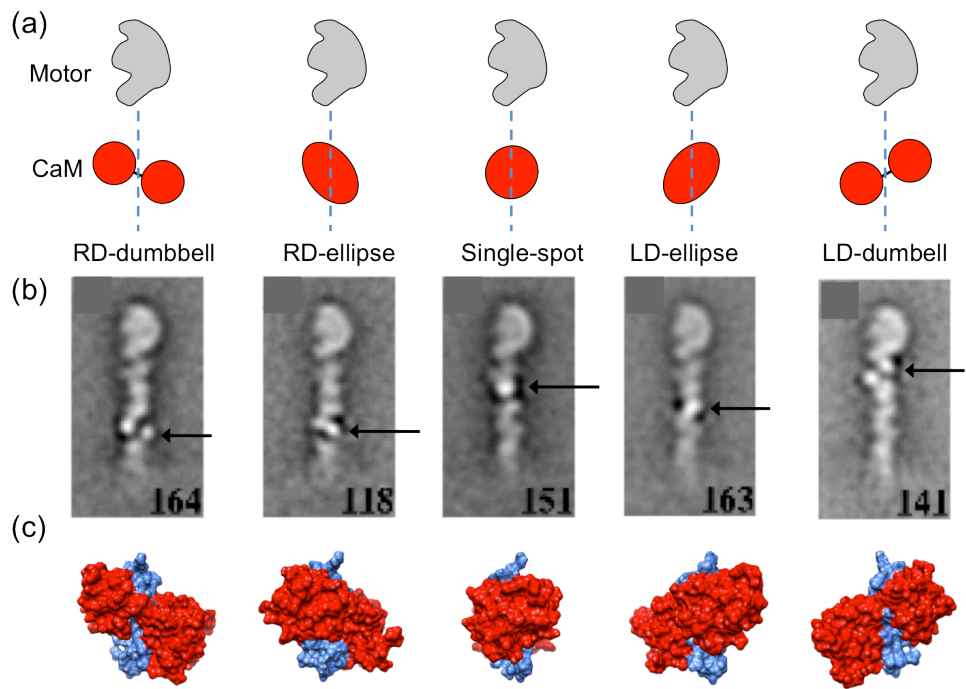
Figure 3.23: Principal CaM views encountered in class averages.

(a) cartoons and labels of principal CaM appearances. The dumbbell and ellipse orientations with the lever axis (dashed blue line) are indicated by their name prefixes: right side lower down = Right down (RD), or left side lower down = Left down (LD). (b) selected class averages of HMM heads where one of the CaMs (highlighted by an arrow) displays the CaM appearance as in the cartoon above. Numbers bottom right indicate the number of images forming the average. (c) surface views of CaM-IQ structure (PDB: 2IX7), oriented to produce the principal appearances to the cartoon above. CaM is red, lever helix is blue. Starting from the left most image the various views are related by consecutive rotations about the lever axis (clockwise rotation if viewed from C-terminal end). RD-dummbell to single-spot is an ~ 90° rotation.

The predominant views seen at each CaM position, and the extent of variation is described below for each CaM in turn. The labels for the five principal profiles possible, as given in Figure 3.23 (Right down (RD)-dumbbell, RD-ellipse, single-spot, Left down (LD)-ellipse, LD-dumbbell), are used.

3.5.2.2   CaM-by-CaM classification results: CaM 1
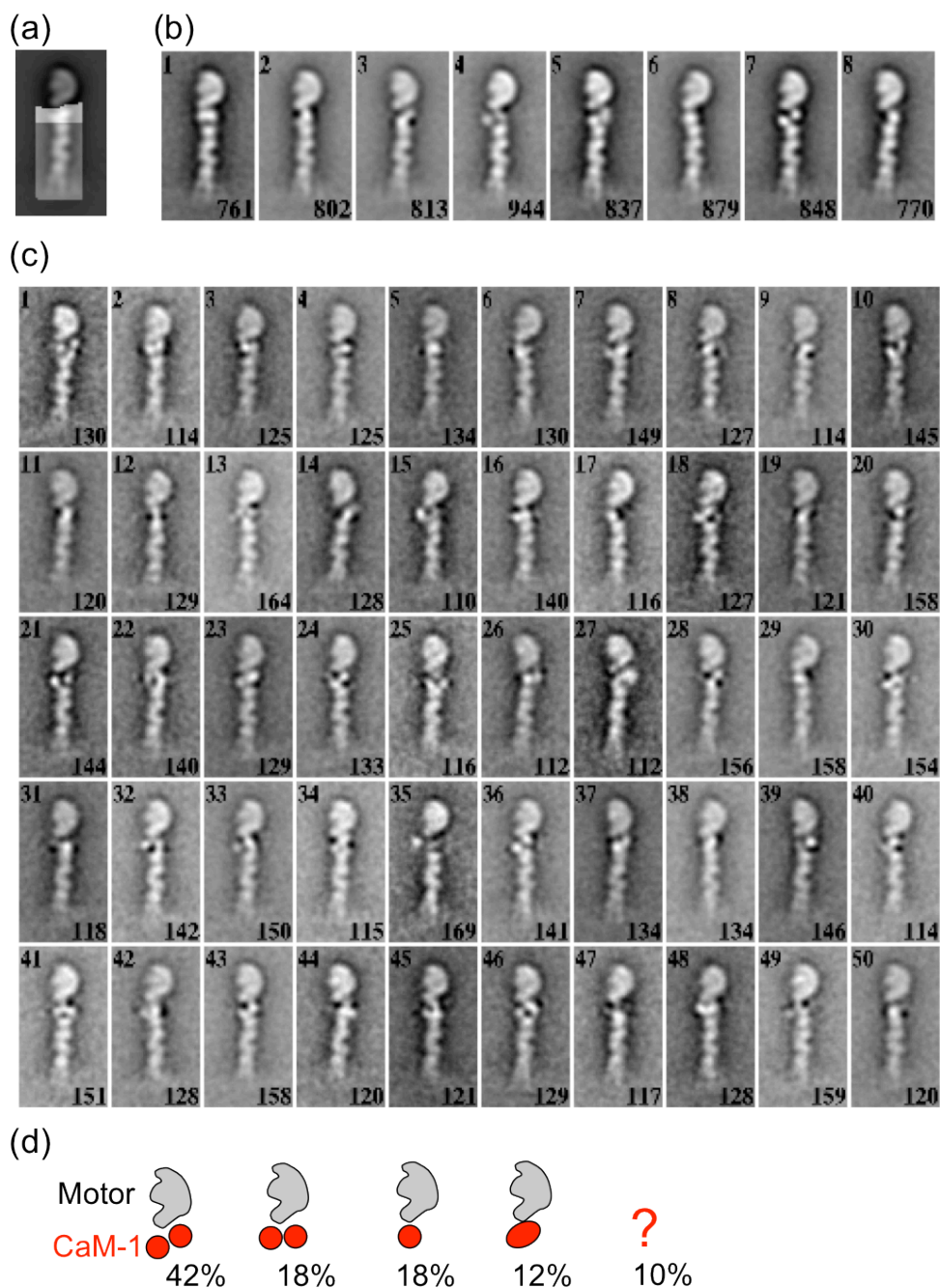
Figure 3.24: CaM-1 classifications

(a) the CaM-1 classification mask used is depicted (small bright white rectangle) superposed on the global average. The remainder of the lever, to be explored in other classifications, is highlighted by the larger less bright white rectangle **(b)** montage of averages from 8-class overview classification using the CaM-1 mask. **(c)** montage for 50-class detailed classification **(d)** cartoons of the principal views seen in the class averages with approximate percentages of individual images exhibiting each appearance below. '?' indicates undefined images, where the CaM appearance was too ill-defined or ambiguous in the average to assign to a view.

CaM-1 appears predominantly with a two-lobe LD-dumbbell profile, *i.e.* with the right lobe higher up than the left (*e.g.* class 7 in 8-class overview, classes 1, 21, 36 in 50-class

montage), identifying the right lobe as the upstream-binding C-lobe. However, there is some variation in the relative positioning of the 'lower' N-lobe which, in some classes, is positioned slightly further up the lever (*e.g.* class 4 in overview, classes 2, 7 and 15 in 50-class montage), and in others appears high enough to be in line with the C-lobe (class 1 in overview, classes 3, 13, 25 in detailed montage). This apparent N-lobe mobility supports the observation from crystal structure analysis, see 1.6.1, that the N-lobe is usually freer from (makes fewer interactions with) the lever helix than the tightly-bound C-lobe.

As well as the predominant two-lobe view, classification reveals some uncommon but clearly resolved examples of CaM-1 with a single-spot profile, corresponding to a one-lobe view (*e.g.* classes 6 and 8 in overview, 11, 17, 20 in detailed montage). This suggests, surprisingly, that a significant (~90°) rotation of CaM-1 is possible. But is this an independent rotation of CaM-1, or is it correlated with a similar rotation of the other CaMs and/or the motor domain? Close inspection of CaM-1 single-spot averages (*e.g.* classes 5, 11, 17, 20, 29 and 50 from 50-class montage), shows CaM-2 adopting a consistently one-lobe appearance, as it most frequently does for CaM-1 two-lobe examples. The motor domain retains the 'face-in-profile' appearance across almost all averages. This suggests that CaM-1 can axially rotate independently of the surrounding parts of the head.

3.5.2.3 CaM-2

(a)   (b)

(c)

(d)
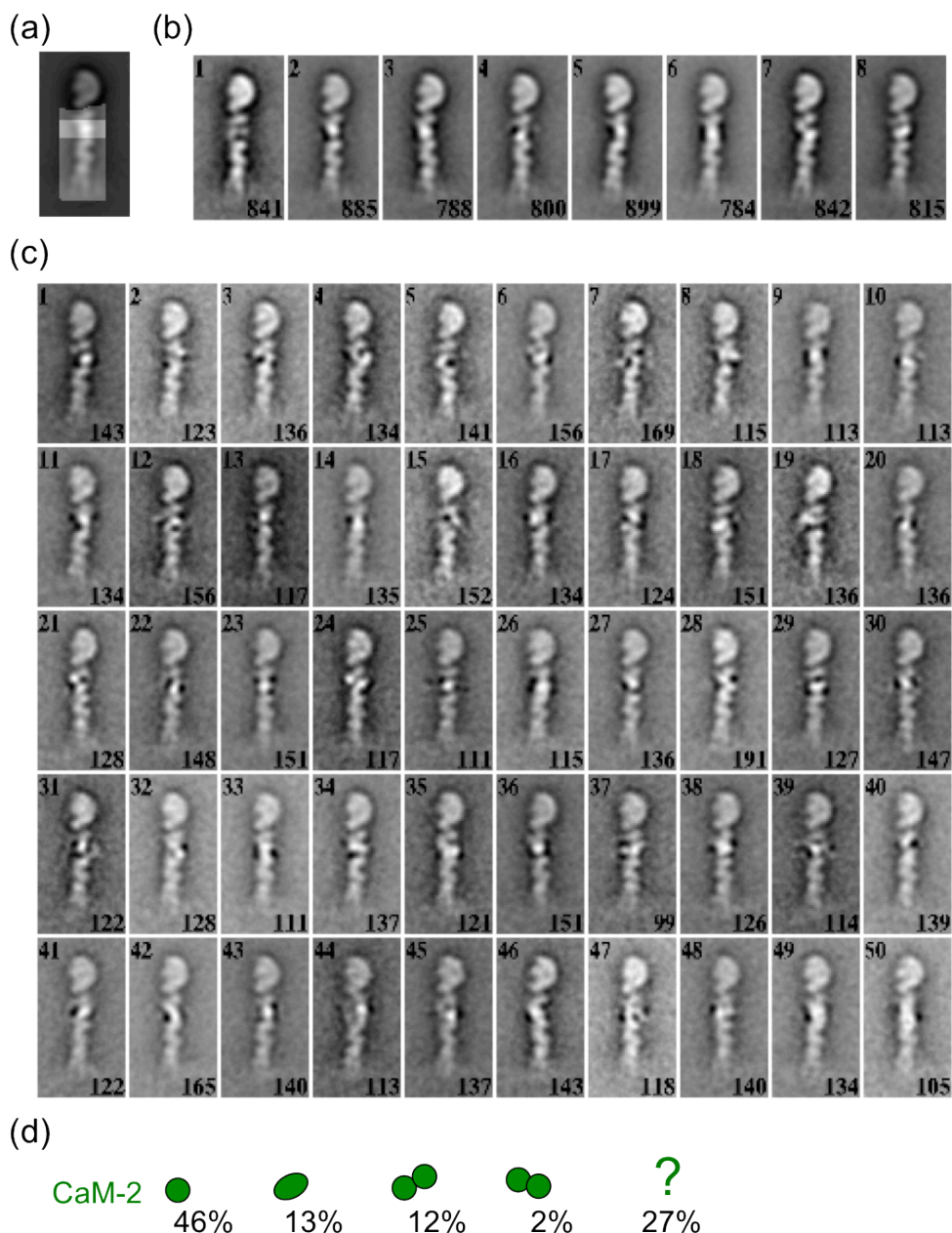
Figure 3.25: CaM-2 classifications

   Parts (a)-(d) depict CaM-2 features as in Figure 3.24

CaM-2 appears most predominantly with a single-spot (one-lobe) profile, *e.g.* classes 2-7 of the 8-class overview and classes 1, 9, 23, 36 of the 50-class montage. The darkly-outlined bright spot tends to contact the left side N-lobe of the two lobes on CaM-1. This is consistent with the arrangement in the CaM 1-2 crystal structure (Houdusse et al., 2006), identifying the view of CaM-2 here to be C-lobe uppermost, in front of its N-lobe.

Less frequently, CaM-2 also appears to have rotated slightly from the one-lobe view, broadening to form an LD-ellipse profile (*e.g.* class 8 in 8-class montage, classes 6, 18, 24 in 50-class montage). In other cases, the rotation is enough to resolve both CaM lobes (*e.g.* class 1 in 8-class montage, classes 12, 35, 37 and 48 in 50-class montage) in LD-dumbbell

112

profile. Which way has CaM-2 rotated to form the LD-ellipse or LD–dumbbell views? A clockwise rotation about the lever axis (viewed from C-terminus) would involve the smaller axial rotation and so is more likely. In the LD-ellipse profiles there appears to be slightly darker staining to the right side of the CaM (classes 6, 18 and 24 in the 50-class montage), indicating that this side projects upwards more from the grid surface. Though a very subtle effect, this likely indicates that the C-lobe is still uppermost but has shifted to the right, confirming a right-handed CaM rotation about the lever helix axis.

3.5.2.4 CaM-3
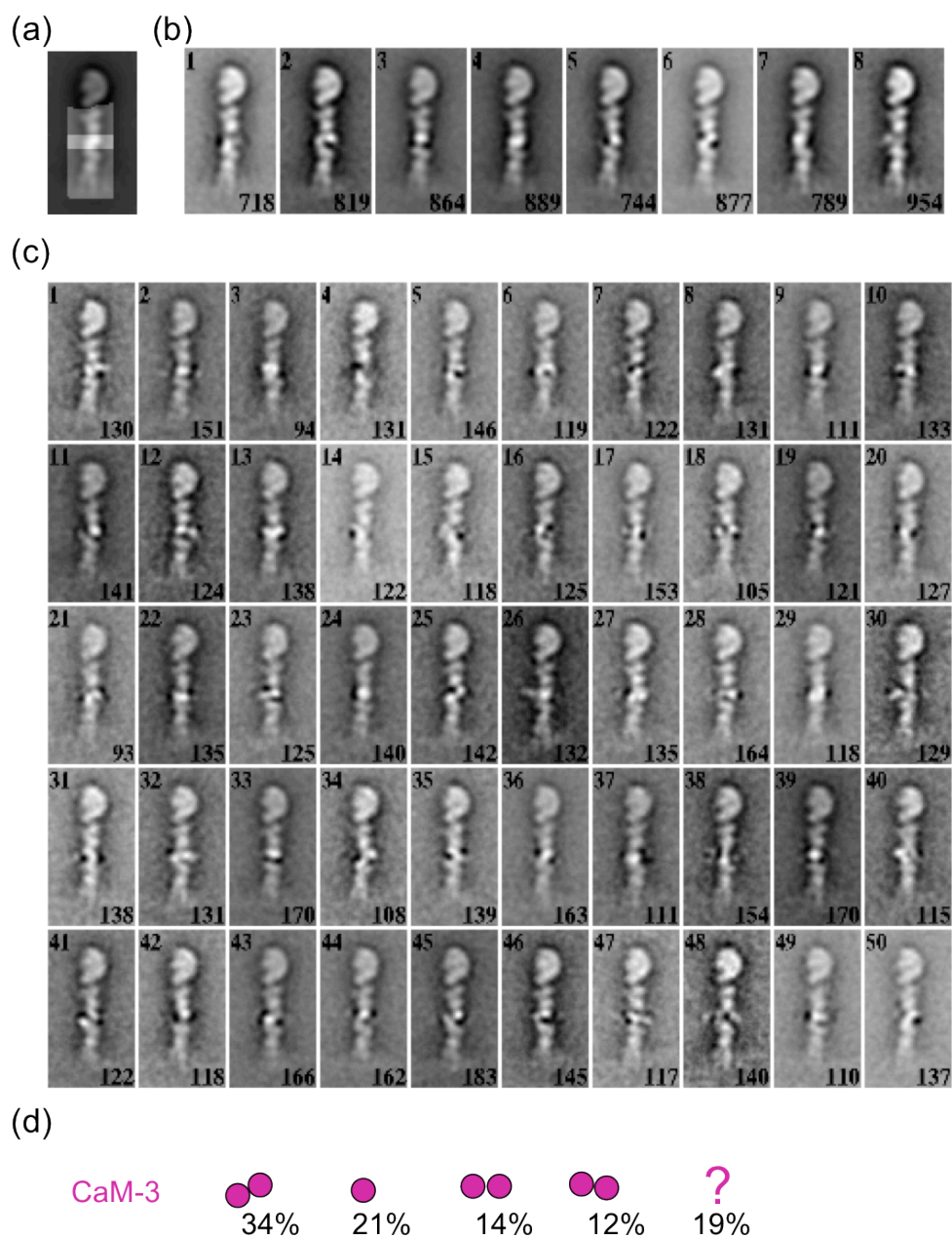
Figure 3.26: CaM-3 classifications

(a)-(d) layout as Figure 3.24.

The CaM-3 classifications show that it is more varied in appearance than CaMs 1 and 2, with a more equal split between the predominant two-lobe LD-dumbbell profile (34 %) and a single-spot alternative view (21%). It is seen most often as a LD-dumbbell (similar to the CaM-1 view) usually with both lobes resolved (*e.g.* classes 1, 8, and 25 in 50-class montage) or forming and LD-ellipse (*e.g.* classes 36, 50). Similar to CaM-1, the lower left-side lobe (likely to be the N-lobe), appears more mobile, sometimes appearing higher up the lever and even in line with the adjacent right lobe (*e.g.* classes 6, 13, 18, 27 and 28 in 50-class montage).

114

As well as the LD-dumbbell/ellipse profile, in some less common instances, the profile is reversed to a RD-dumbbell/ellipse shape (*i.e.* with the left lobe higher than the right *e.g.* class 8 in 8-class montage, classes 5, 10, 12, 35 and 47 in 50-class montage). In these cases, it is not possible to infer with certainty if CaM-3 has rotated a full 180° to reverse the two-lobe slant, or whether freer N- or C-lobes have moved up and down the lever relative to each other.

In a few cases (*e.g.* classes 26, 31), the N-lobe is positioned quite far from the axis of the lever helix as marked by the positioning of the other CaMs. These could be instances of the N-lobe detaching from the lever helix altogether, forming a so-called 'extended' light-chain conformation, as predicted to occur at CaM-3 by Terrak et al. (2005; 2003), based on their crystal structure observations that a bulky residue at position 7 in the IQ motif (in this case Methionine; [822]**IQ**KYW**RM**YVV**R**) can sterically hinder the interaction between the N-lobe and lever helix.

Here, one note of caution should be raised. Recalling that the protein sample under study was originally purified following a GFP-CaM exchange (subsequently deemed too low) (see 3.2), a possible alternative interpretation of the out-of-line spot is that it could be the GFP region of exchanged GFP-CaMs, rather than a detached N-lobe. However, this possibility is deemed unlikely, given the very low estimated exchange of the sample (< 0.9 GFP-CaM per head) and because no similar off-axis spots are seen for CaMs 1-2. Preferential exchange at IQ-motif 3 would be at odds with findings from other groups who have studied myosin 5a CaM exchange, where evidence points to preferential exchange at IQs 1-3 (Lewis et al., 2012; Sakamoto et al., 2005; Trybus et al., 2007). Moreover, if the off-axis spot were GFP, then the partner in-line spot would be a one-lobe CaM view. The relatively weak intensity of the accompanying in-line spots do not suggest this to be the case (see examples in next paragraph).

As well as the two-lobe views, there are also frequent clear examples of a single-spot CaM-3 profiles (*e.g.* classes 3-5 and 7 in 8-class montage, classes 2, 9, 19, 24, 33 and 39 in 50-class montage). In these cases, the one-lobe view seems uncorrelated with differences in appearance of the surrounding CaMs (or motor domain) and so implies independent CaM-3 rotation.

3.5.2.5   CaM-4
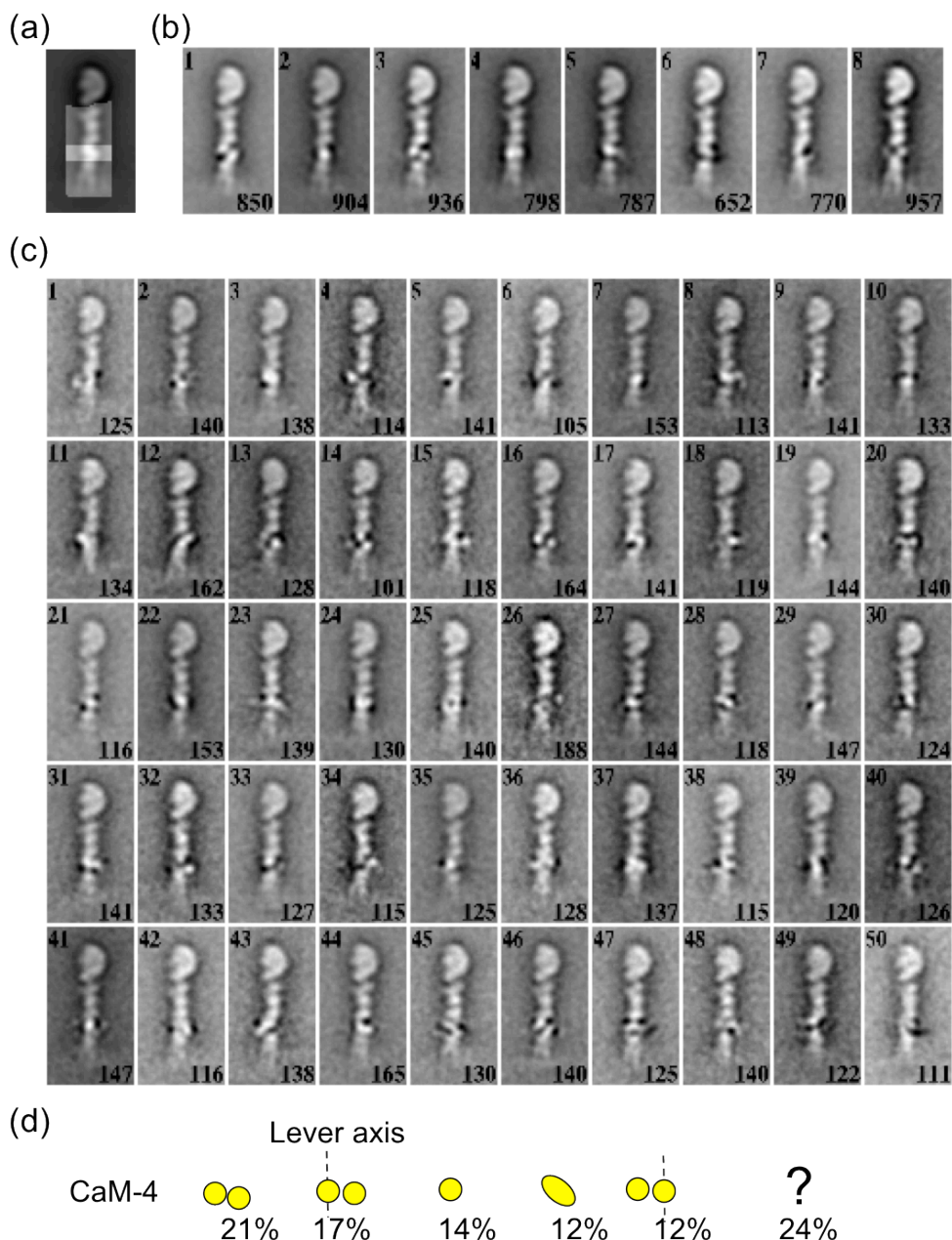
Figure 3.27: CaM-4 classifications
   (a)-(d) layout as Figure 3.24.

CaM-4 also exhibits considerable variety in appearance. The predominant defined view (though only 21%) is a RD-dumbell or –ellipse, with C-lobe left and uppermost. In this view, the two lobes are often clearly resolved, (*e.g.* see classes 1 and 3 of 8-class montage, classes 16, 17, 19 and 28 in 50-class montage). The left (C-) lobe often contacts the left (N-) lobe of CaM-3 such that together CaMs 3 and 4 form a '<' cleft shape, with a dark pool of stain between the two CaMs. This often makes a strong identifying feature of the classification averages, and is also a feature of the global average.

As for the other CaM positions, there is some variation in the angle made between the dumbbell/ellipse and the vertical, varying from a steep angle (~45° to lever, *e.g.* class 40) to the two lobes lining up (class 8 in overview montage, class 20 in 50-class montage). Also, a subset of images (14%) classify to produce a single-spot appearances (class 2 in overview, classes 7, 10, 27 and 35 in 50-class montage), again suggesting that the light chain can rotate significantly (by up to 90°).

A striking feature of the CaM-4 classes is the variation in lateral positioning of the lobes in the two-lobe views. Similar to the extended-lobe classes for CaM-3, there are CaM-4 classes (~17% of images) where the right-side lobe is positioned considerably off-axis to the right (*e.g.* classes 15, 18, 31-34), as well as classes (12% of images) where the left-side lobe is off-axis to the left (*e.g.* classes 1, 4, and 11 in the 50-montage). The IQ motif at CaM-4 (^{845}IQSYLRGYLTR) is canonical, and its sequence doesn't suggest an obvious reason as to why CaM-4 would be prone to lobe detachment. Again, there is a small possibility that the off-axis lobes are GFP labels, but the same counter arguments apply as for CaM-3 and make this unlikely.

117

3.5.2.6 CaM-5

(a)   (b)

(c)

(d)
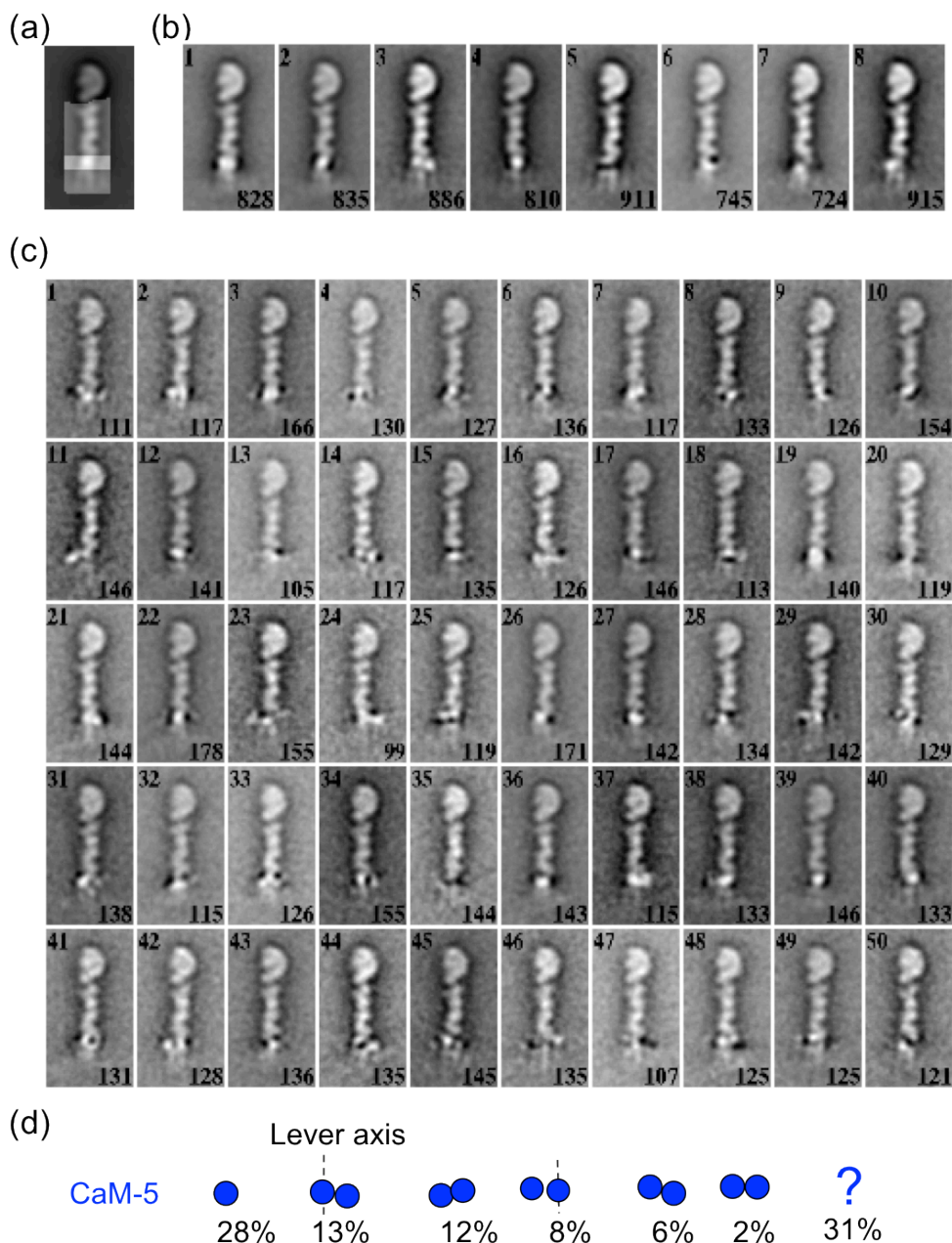
Figure 3.28: CaM-5 classifications

(a)-(d) layout as Figure 3.24.

The single-spot one-lobe appearance is the most frequent (~28% of images) of the well-defined profiles for CaM-5 and is the appearance in the global average. However, alternative two-lobe classes (with varying angles and off-axis lobe positions) together account for 41% of the images. Most of the two-lobe views (29% of images) exhibit a RD-type profile, the same orientation as for CaM-4 (*e.g.* class 3 of 8-class montage), but there are some examples (12% of images) exhibiting the reverse LD profile (*e.g.* classes 7, 25 and 42 in 50-class montage). Again, the variability in class appearances for CaM-5 suggest significant torsional flexibility in the CaM-IQ subunits of the lever, as well as mobility in

precise attachment position of the individual lobes. The CaM-5 classes are particularly hard to interpret and so this CaM has the highest fraction (31%) of undefined images.

3.5.2.7 CaM-6

(a)   (b)

(c)

(d)

Figure 3.29: CaM-6 classifications

   (a)-(d) layout as Figure 3.24.

The CaM-6 classes are the most diverse in appearance, spreading quite evenly among a range of appearances. The global average is a two-lobe view, RD-dumbbell or ellipse, but it is the least well-defined of the CaMs in the average (lighter stain outline, lower edge fades into background noise). The overall two-lobe appearance is accounted for in the classifications by a range of two-lobed classes with lobes either horizontally arranged or

with RD-slant (39% of images), although the (marginally) most predominant individual class profile (16% of images) is the single spot.

Perhaps of most note for CaM-6 is the tendency for either the two-lobe appearances to have one lobe elongated and tapering away to the left (*e.g.* classes 4, 20, 34 and 44 of the 50-montage) or the right (classes 2, 8, 9 and 27 of 50-montage) of the lever axis, or, equivalently in the one-lobe views, to have the single spot positioned off-axis to the left (classes 6, 14 and 16 of 50-montage) or the right (class 50). A possible reason for the frequency of these left or right elongations or deviations may be the close proximity of CaM-6 to the head-tail junction and to CaM-6 of the second unaligned head, which may be a complicating factor in how the images are classified in this region.

3.5.2.8 Summary and discussion of CaM appearances

Figure 3.30 summarises the observations from the CaM-by-CaM classifications. It demonstrates that for each CaM there is a predominant appearance, but also a number of alternatives. The percentage of images falling into the predominant appearance generally decreases from CaM-1 (42%) to CaM-6 (16%), whilst the number of alternative appearances, and the spread of images across the alternatives, concomitantly increases. This suggests that CaM rotation and lobe positioning becomes freer as one moves away from the motor domain, though, as mentioned above, the greater variability exhibited at CaM-6 may also be due to 'interference' in image classification averages, caused by the proximity and variable position of CaM-6 of the unaligned head in the individual molecule images.

In general, the predominant appearance for each CaM classification matches, as one would expect, the appearance in the global average, and so features of the classes can be matched to parts of the atomic model of the head. CaM-6 is a slight exception, being predominantly one-lobe in appearance when individually classified, whilst appearing as a two-lobe DR-dumbbell in the global average. However, the three next largest CaM-6 classes are all two-lobed DR-dumbbells, with slightly differing orientation angles. These are closer to the global average appearance and, if combined, would form a majority two-lobe appearance.
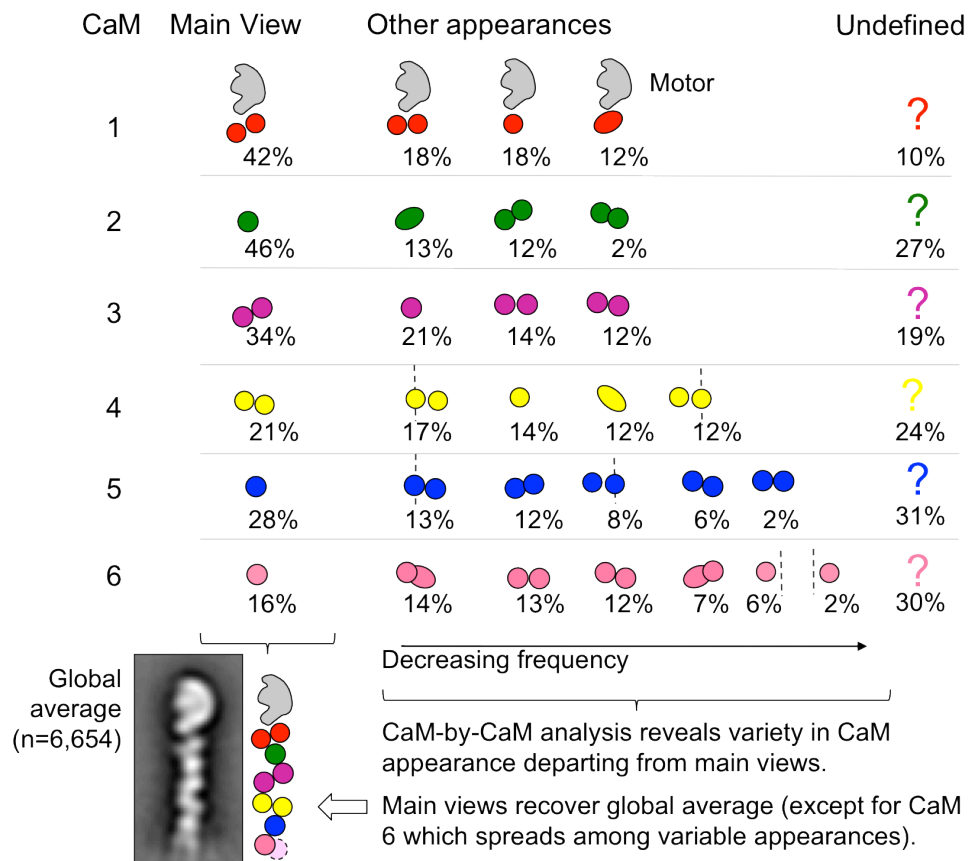
Figure 3.30: Summary of CaM-by-CaM classification analysis.

Cartoons of the CaM appearances found by classifications shown in Figure 3.24-Figure 3.29. Depictions of the appearances of CaMs 1-6 are presented on separate rows of the figure. The predominant appearance is set on the left of each row, with the less frequent appearances shown to the right, arranged by decreasing frequency of occurrence. For each CaM position, the percentage of images falling into the different views are given below each depiction. The percentage of images where the class appearance could not be identified as a particular view is given on the right (under a question mark symbol). The dotted lines shown for some appearances represent the axis of the lever helix, and are shown to indicate an appearance where one of the lobes appeared to be significantly off-axis. Note, the figure should not be read vertically *i.e.* there is no implied correlation between CaM appearances in the same columns.

What can be concluded about the lever from the CaM classifications? The general picture that emerges is one in which, at least under negative stain conditions, there is significant movement of the CaMs about preferred conformations; either entire CaM-IQ subunits within the lever undergoing significant axial rotations or individual lobes shifting up and down or detaching from the lever helix. The degree of CaM-IQ axial rotation implied, often up to 90°, was surprising.

An important question, only partly addressed by the initial work here, is the extent to which the CaM-IQ subunits rotate independently of surrounding parts of the lever and/or the motor. Inspection of the classification averages here indicates that CaM-IQ subunits seem generally uncoupled with their neighbours, with adjacent parts of the molecule

121

tending to appear as in the global average, whatever the appearance of the CaM-IQ subunit studied. However, examination of correlations between movements of larger sections of the lever has not been rigorously pursued within the time restraints of this project, and would be a natural extension of this work. Motor-CaM-1 and CaM-pair classifications were performed, but interpretation of the results is as yet inconclusive. Another potentially fruitful analytical method, applicable in future work, might be to label the orientations of the six CaMs along the lever in each individual head image using the classification assignments. Molecule label sequences could then be sorted and analysed at the single-molecule level for any frequently occurring patterns indicating particular correlations.

The type of CaM movements implied here have not previously been described. The only precedents to suggest CaM mobility in myosin 5a have been a cryo-EM structural study of the motor plus first two IQ motifs bound to F-actin (Trybus et al., 2007), which found differently oriented CaMs in the presence or absence of calcium; and in the spread of angle data specifying 3D CaM orientations in polarised TIRF microscopy experiments (Lewis et al., 2012). Some degree of movement of the CaMs is perhaps not surprising, given the turbulent thermal environment at the length scale of the HMM molecule (*cf.* the thermally-induced fluctuating twist in the F-actin structure (Egelman, 1997; Egelman et al., 1982)). However, the degree of CaM twisting (of up to 90°) seen in the EM images here is much greater, and suggests either a considerable twisting compliance in the lever structure not detected in previous experiments, or alternatively, that substantial torques are applied to sections of the lever as part of the negative staining process and subsequent EM. One could attempt to quantitate the twisting stiffness of the CaM-IQ sections by application of the equipartition theorem to the azimuthal angle distributions corresponding to the CaM views in Figure 3.30. However, given the resolution of the CaMs in image averages, one cannot easily determine if there is a continuous distribution of CaM angles, or whether there are only a few conformations at specific rotations (i.e. conformations separated by significant free energy barriers).

How might the axial rotations of CaM-IQ subunits be possible, and to what extent are the rotations seen in the EM images relevant to any torsional fluctuations in the molecule *in vivo*?

If the rotations reflect real and reversible twisting of CaM-IQ sections of the lever, then the most plausible structural means to allow this would be compliance in the sections of lever helix between the CaM-binding regions. To accommodate such large CaM-IQ twists, these bare α-helical sections would need to unwind or significantly deform (as at the pliant region between motor and CaM-1, which allows for independent motor flexing and

twisting). In this scenario, CaM-IQ units would rotate as a whole rather than the CaMs slipping around an intact lever helix (which would require continual breaking and forming of non-covalent interactions). If such twisting is relevant *in vivo*, how might torsional freedom in lever sections be important to myosin 5a function? One could speculate that this may be a mechanism that imparts important torsional flexibility in the head, which allows the molecule to efficiently find differently oriented F-actin subunits when moving processively or crossing from one filament to another.

Alternatively, and perhaps more probably, it may be that the varying CaM orientations seen in the EM images are actually artefacts and evidence of the surface tension forces that likely come to bear on the molecules during drying on the EM grid (see 1.9.1). Considering a simple model for this process, one could imagine that a substrate-adsorbed myosin 5 molecule breaks the surface of an evaporating film of liquid and that a meniscus develops around the perimeter of the projecting region (the meniscus profile being determined by the adhesive forces between liquid and macromolecule and the surface tension at the liquid-air interface), Figure 3.31. As the meniscus lowers, the surface tension force acting on the sides of the macromolecule (equivalent to the force pulling on a semi-submerged Wilhelmy plate) will have components parallel and perpendicular to (towards) the substrate.
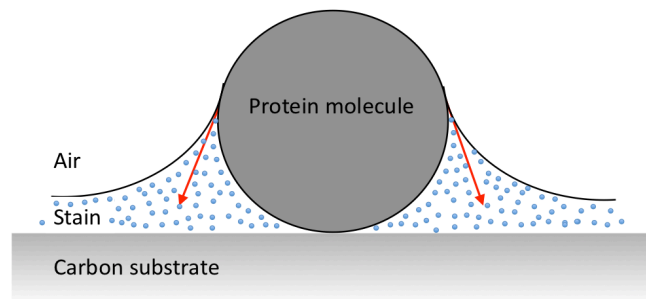
Figure 3.31: Model of surface tension forces acting on a surface-adsorbed protein molecule.

The cross section of a protein molecule (dark grey sphere) adsorbed to a carbon surface (light grey) is shown. As a covering layer of stain solution (blue dots) evaporates, the protein breaks the surface of the liquid and a meniscus may form around the projecting part of the molecule. Forces on the protein due to surface tension around the boundary of the meniscus would be generated (red arrows), their magnitude determined by the surface tension of the liquid-air interface and the contact angle of the meniscus with the protein (*i.e.* the 'wettability' of the protein by the stain solution).

The force components of the surface tension act to spread and pull the molecule down onto the substrate. Moreover, if there is a difference between the surface tension on one side of the macromolecule and the other (e.g. because of differences in wettability, or if the mass of the molecule is asymmetrically distributed, like the two lobes of a CaM) then a torque will be generated about the central axis. Associated collapse of CaM lobes onto the

substrate, or twisting of CaMs under torques produced by the surface forces could therefore account for the variety of appearances found at the different CaM positions.

Taking a simple case where a meniscus is formed around a single CaM-IQ section of the lever, an estimate of the force pulling on the structure is given by $F = \gamma L$, where γ is the surface tension for the liquid-air interface and L is the perimeter length of the meniscus. Taking γ as 0.072 Nm^{-1} for water-air (*i.e.* ignoring complications of stain and other buffer components) and taking $L \sim 5$ nm around the CaM-IQ structure, yields $F = 360$ pN. This rough estimate shows that the surface tension pulling on an object the size of the CaM-IQ complex is sizeable (significantly greater than thermal forces) and thus associated torques produced about the lever axis could readily explain the variable CaM lobe orientations found in the EM images.

3.6   Discussion and Conclusions

In this chapter, a detailed structural analysis of the myosin 5a head was made, using image alignment and classification methods applied to images from a previous negative stain EM study. Owing to the exceptional quality of staining and sample purity, the global average of images aligned by the lever, revealed a high level of detail, resolving features of both the motor and lever's substructure. On average, the motor lies on the carbon-coated EM grid in an orientation that resembles a face in profile, with the SH3, lower and upper 50k regions making characteristic protrusions on the left side. The six-CaM substructure of the lever is clearly resolved as six closely-spaced globule-like spots (round or elliptical), each lying at a different orientations and possibly forming contacts with its neighbours.

Comparison of the EM average to existing atomic models of the head allowed a new model to be constructed. The new adapted model whose profile matched the HMM head and fitted well in the 3D EM density map of Liu et al. (2006), showed that the average orientation of the CaMs and gross structure of lever and motor are unchanged when the molecule changes from an active open state to a folded inhibited one. The new best-fitting model implemented a reinterpretation of the orientation of the motor in the Liu et al. 3D density map suggested by Sellers et al. (2008), confirming this as the better fit. It also required the introduction of a bend in the lever α-helix at the motor lever junction (res. 755-767) in order to marry the motor and lever structures, suggesting this as a pliant point in the head, a fact later confirmed by classification analysis of motor-lever rotation and tilting.

Image classification revealed a number of flexibilities within the head. Firstly, the motor was observed to be able to rotate by ~50° about its long axis, between two preferred

orientations (face-in-profile and face-away). The motor rotation appears independent of the lever which, on average, adopts one preferred axial orientation on the EM grid. The motor also exhibits substantial tilting about the motor-lever junction. The distribution of angles found in either view suggests thermally driven tilting fluctuations about an equilibrium position, the motor-lever junction acting something like a torsion spring, with the motor able to tilt like the bar in a torsion balance. By application of the equipartition theorem to the tilt angle distribution, an effective stiffness for the motor-lever joint of 32-51 $pN{\cdot}nm{\cdot}rad^{-2}$ was estimated. This is consistent with values from recent single-molecule fluorescence microscopy assays (Shiroguchi et al., 2011), and indicates that inherent motor-lever hinge-like flexibility is a significant source of compliance in the head, and one that should not be overlooked in actomyosin crossbridge measurements.

Classification of the appearances of each CaM revealed a surprising degree of variability. Although each CaM had a predominant appearance (a two-lobed dumbbell shape, an ellipse, or a single-lobed spot), variations found at each position implied that different azimuthal orientations of CaMs about the lever axis were also possible. The extent of axial rotation between differing views (up to 90°) was surprising, with some CaMs appearing as a two-lobed dumbbell in one average, and a single-lobed spot in another. CaM rotations appeared uncorrelated with movements of neighbouring CaMs (or with the motor), though more classification work is required to confirm this. These observations suggest either that there is significant torsional flexibility at the junctions between CaM-IQ subunits and between CaM-1 and the motor, or is evidence of distorting forces and torques on the lever, perhaps arising from surface tension as the molecules were dried on the EM grid. Mobility in CaM-lobe positioning along the lever helix and examples of lobe detachment were also observed. Given the large (90°) differences between implied azimuthal orientations of CaMs at a given position, the latter explanation of distorting forces due to drying seems more probable. Because the CaM appearance analysis is at the limit of resolution achievable by negative stain EM, it was not possible to determine if given CaMs occupy a continuum of states between two orientations, or whether the orientations represent discrete conformations separated by energy barrier(s). For this reason, the equipartition theorem was not applied to the classification data to calculate torsional rigidity parameters for specific CaM-IQ subunits.

Finally, lever bending (*i.e.* curving of the path lever helix) was also observed, in passing, during the classification work in this chapter. Clearly, bending is another significant type of flexibility in the lever, and is one that can result in large-scale changes in the appearance of the heads. A full analysis of bending and flexural rigidity in the myosin 5a lever forms the subject of the next chapter.

Chapter 4: Lever Bending Analysis

4.1   Introduction and aims

In this chapter, a study of the flexural rigidity (bending stiffness) of the myosin 5a lever domain is made, from analysis of images of negatively-stained HMM molecules. The focus of the study is on testing a mechanical description of the lever as a thin isotropic rod that is continuously flexible, a model known in polymer physics as the worm-like chain (WLC) (Doi and Edwards, 1988).

Lever bending is perhaps the most obvious manifestation of large-scale flexibility in the myosin 5a head, apparent from the variety of differently-curved lever shapes seen in EM images of free and F-actin-bound molecules (Burgess et al., 2002; Oke et al., 2010; Walker et al., 2000) (this thesis, chapters 3 and 4). The main aims of this chapter are:

1. To establish if the myosin 5a lever exhibits isotropic flexural rigidity, that is, an equal bending stiffness in all directions, and to ascertain if this stiffness is uniform along the lever's length.
2. To determine the flexural rigidity (or appropriate stiffness parameters) of the lever domain from EM images of free HMM molecules.

As has been discussed in 1.8.2.4, the bending stiffness of the lever is of central importance to the function of many myosins and, in particular, of myosin 5a. For myosin 5a to make processive strides along F-actin, its lever must be sufficiently rigid to transmit force generated in the motor and move a detached head and cargo forward through its powerstroke. At the same time, there is evidence that the lever must be flexible enough to bend and adopt functionally-important strained conformations, storing elastic energy that may contribute to the powerstroke when the lever relaxes. Myosin 5a is a good system for the study of lever bending and stiffness, not only because lever rigidity is important to its biological function, but also because the lever's long length makes it particularly amenable to EM studies. Being formed from a six CaM-IQ subunits typical of other myosin levers, gaining an understanding of myosin 5a's lever bending properties and how these relate to the domain's substructure might also provide insight into the mechanical and functional roles played by other myosin levers.

The WLC polymer model is particularly suited to describing stiffer semi-flexible polymers (*e.g.* DNA) that exhibit smoothly curving shapes, as opposed to the freely-jointed chain

model that is flexible only at joints linking discrete segments. The WLC or semi-flexible rod, is the model used for the lever in Vilfan's elastic lever-arm model of myosin 5 (where the flexural rigidity is assumed isotropic and uniform along the lever) (Vilfan, 2005), and in Lan and Sun's alternative model (where two flexural rigidity parameters are used depending on the direction of bending) (Lan and Sun, 2006; 2005). The semi-flexible isotropic rod (WLC) is also the model assumed in most optical trap and muscle fibre measurements when measuring cantilever stiffness of actomyosin crossbridges (see 4.2.3.1). The isotropic rod model is not necessarily the most appropriate model of the myosin 5a lever, especially given its complicated structure of differently orientated CaMs bound to an α-helix. In their myosin 5a modelling study, Craig and Linke (2009) used an alternative lever model of three straight segments connected by flexible hinges. One of the objectives of this chapter is to test the appropriateness of the isotropic rod model for the myosin 5a lever, by examining lever conformations in EM images of myosin 5a molecules. Do the lever shapes seen fit the model of isotropic and continuous flexiblity, or is a more complex model involving points of high compliance required?

The dataset used for the analysis undertaken in this chapter was the same dataset as used in chapter 3 to examine head structure and torsional flexibility in the lever. The dataset comprised 23,250 images of myosin 5a HMM heads (in nucleotide-free state), aligned by their lever domain. For full details of the sample, the EM collection conditions and details concerning alignment by the lever domain see 3.2.

To begin this chapter, the isotropic rod model and the theoretical framework used to describe bending deformations and stiffness is set out. Ideas from this section will be used later in the chapter to analyse lever curvature and flexural rigidity. This is followed by a description of a numerical method devised to analyse a set of filamentous shapes. The work involved the development of a computer program, designed to be a general-purpose tool for analysis of a set of curves, each defined by a series of coordinate values. In this case, the curves were derived from myosin 5a lever shapes, but curves representing any filamentous objects could be used. The program was developed to process the set of curves provided and calculate parameters related to an implicit bending rigidity, such as local tangent angles, radii of curvature and Fourier mode amplitudes.

The analysis method developed is then applied to the aligned and classified image averages of negatively-stained myosin 5a heads. To analyse lever bending, a set of lever curves were created by tracing out the lever helix path, as implied by the position of the lever-bound CaMs in class average images. Through analysis of the radius of curvature at positions along the lever curves, flexibility with lever length could be mapped for the

entire dataset, thus addressing the question of whether the domain is isotropically and uniformly flexible. In addition, by decomposing the lever shapes into a series of Fourier modes, an estimate of the lever's flexural rigidity could be obtained from the variance in the mode amplitudes.

To study the bending characteristics of thin rod-like objects such as cytoskeletal filaments or protein domains like the myosin lever, a means of describing the rod shape and its curvature is required. The mathematical description of the isotropic rod adopted here (the WLC model), and related theory relevant to the analysis later in this chapter is described below. The reader may wish to skip this section if familiar with the theory. A summary of the theoretical concepts used in later analysis are given in Table 4.1.

4.2   Theory: bending rod model for semi-flexible proteins

A thin rod, with length L much greater than its diameter is considered. The rod is constrained to lie in a plane but is free to undergo 2D fluctuations in shape in response to external forces, a key simplifying assumption made for the analysis in the remainder of this chapter (discussed in 4.6.2). The rod's shape can be specified by a set of coordinates, $\{x(s), y(s)\}$, defining a plane curve, where the parameter s is the arc length (distance along the curve) to point (x, y), Figure 4.1(a). Information about the shape of the curve at any arc length can be described by the tangential angle $\theta(s)$, the angle formed between the line tangent to the curve at arc length s and the x-axis. Formally, the tangential angle is defined by the infinitesimal plane curve section (Figure 4.1(b)) by the relation

$$Rd\theta = ds,$$

(Eq. 4.1)

where R is the radius of curvature, the radius of the circle that best fits the curve at the point considered (the 'osculating circle'), Figure 4.1(a). ds and $d\theta$ are the infinitesimal arc length and angle subtended at the curve point, Figure 4.1(b). The tangential angle is also related to the Cartesian coordinates by:

$$\frac{dx}{ds} = \cos\theta, \; \frac{dy}{ds} = \sin\theta.$$

(Eq. 4.2)

Of interest for the studies here is the quantity $1/R$, which specifies the local curvature ρ at arc length s, which, by Eq. 4.1, is the derivative of the tangential angle with respect to arc length,

$$\rho \equiv \frac{1}{R} = \frac{d\theta}{ds},$$

(Eq. 4.3)

a relation also apparent from the diagram in Figure 4.1(b).

(a)

(b)
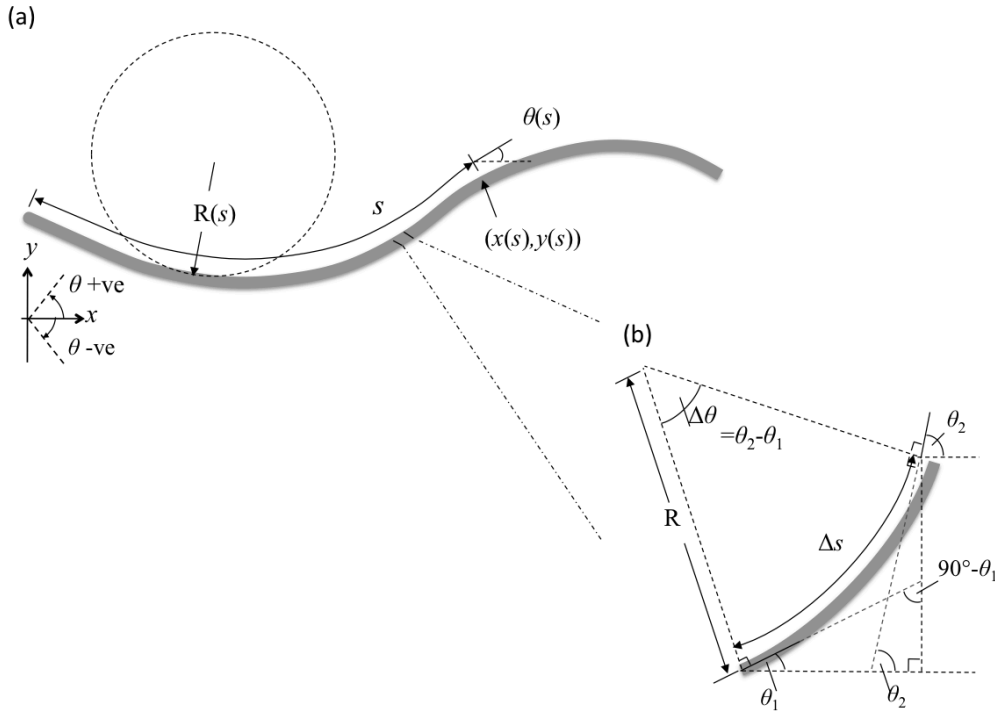
Figure 4.1: Geometry for a bending rod.

(a) a bent rod (grey) described by the set of coordinates $\{x(s),y(s)\}$, where s is the arc length to a particular point. The shape at position s is described by the tangential angle $\theta(s)$, the angle formed between the tangent to the curve at $(x(s),y(s))$ and the x-axis (*NB*: angles formed by anticlockwise rotation from the x-axis are positive, clockwise are negative). The osculating circle (dashed) is the circle that best fits the curve at position $(x(s),y(s))$. Its radius R is the radius of curvature. **(b)** a small segment of the rod with arc length Δs and radius of curvature R. Over the segment length Δs, the tangential angle changes by $\Delta\theta = \theta_2 - \theta_1$. As the segment length decreases, $\Delta s \rightarrow ds$, $\Delta\theta \rightarrow d\theta$ and $1/R = d\theta/ds$. Figure adapted from Howard (2001).

4.2.1   The beam equation and flexural rigidity

The rod will bend in response to applied forces, *e.g.* from an external load or, as is the case for the myosin 5a lever analysis in this chapter, from thermal forces (*i.e.* through collisions with molecules undergoing Brownian motion). How the rod bends depends on the forces applied and the rigidity of the system. In general, the curvature of any bend induced in a rod will be proportional to the applied bending moment M (the force times the distance from the rod's midplane), a relationship known as the beam equation:

$$M(s) = \kappa\rho = \kappa\left(\frac{1}{R}\right) . \qquad \text{(Eq. 4.4)}$$

The constant of proportionality, κ, between bending moment and curvature, is called the flexural rigidity (known also as bending stiffness or bending modulus). It has units of Nm^2 and is a property determined by both the rod material and its internal structure. If the rod is non-isotropic or does not have a circular cross-section then the flexural rigidity will also depend on the direction of bending. However, if the rod *is* isotropic and formed by a

homogenous and elastic material, the flexural rigidity can be shown to be the product of *E*, the Young's modulus of the material, and *I*, the second moment of inertia of the rod cross-section (the area integral of the square of the distance from the middle of the section), see Howard (2001). Using Eq. 4.3, the beam equation can be alternatively expressed in terms of the tangential angle

$$M(s) = \kappa \frac{d\theta}{ds} \quad \left(\overset{isotropic}{=} EI \frac{d\theta}{ds} \right).$$ (Eq. 4.5)

4.2.2 Bending energy

The beam equation is analogous to Hooke's law for extension of a spring where *M* is analogous to the force, κ the spring constant and the curvature the extension. As the rod is bent in response to the external bending moment, energy is stored in the material due to extension or compression of the bonds between atoms on the outer and inner sides of the curved shape. Bending moments due to the tensile and compressive forces in the material develop and balance the external moment. Continuing the Hookean-spring analogy, the energy stored in the bent rod, per unit length, dU/ds, can be formulated as:

$$\frac{dU}{ds} = \frac{1}{2}\kappa \left(\frac{d\theta}{ds} - \frac{d\theta_0}{ds} \right)^2,$$ (Eq. 4.6)

where $\theta_0(s)$ is the rod's relaxed shape in the absence of forces. Eq. 4.6 is analogous to the relation $U = \frac{1}{2} k(x-x_0)^2$ for the stored energy in a spring, with spring constant *k* and extension $x-x_0$ from the natural length. If the rod is isotropic and homogeneous and the flexural rigidity is thus constant, the total bending energy of the rod is the integral of Eq. 4.6., *i.e.* the curvature deviation from the relaxed shape, integrated over the rod's length

$$U = \frac{1}{2}\kappa \int_0^L \left(\frac{d\theta}{ds} - \frac{d\theta_0}{ds} \right)^2 ds,$$ (Eq. 4.7)

where κ can be taken outside the integral only if it is assumed constant.

4.2.3 The small angle beam equation

The beam equation can be solved for different forms of applied force and for different constraints, which dictate the relevant boundary conditions to impose on Eq. 4.6. Firstly, if the tangential angle is always small (*i.e.* bending is slight) the equation itself can be helpfully rewritten in terms of the Cartesian coordinates. If the coordinate system is aligned so that the *x*-axis is parallel with the rod's relaxed axis (for simplicity assumed straight) and θ is always small, then $ds \approx dx$. Therefore $dy/ds \approx dy/dx \approx \theta$ by Eq. 4.2.
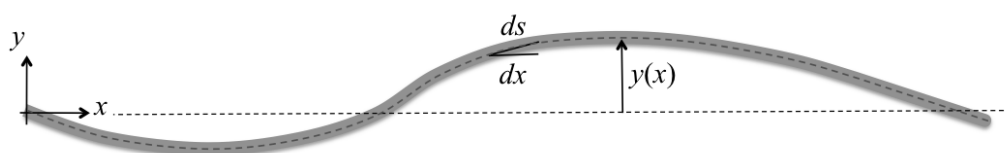
Figure 4.2: The small angle approximation for a bending rod.

If the coordinate system is arranged so that the *x*-axis is parallel to the rod's relaxed shape (here assumed straight for simplicity) and bending is slight, θ is always small and the approximation $ds \approx dx$ can be applied. $y(x)$ is the transverse displacement of the rod from the *x*-axis at position *x*.

This leads to a small angle approximation for the curvature, $\rho = d\theta/ds \approx d^2y/dx^2$, which can be used to write a small angle form of the beam equation:

$$M(x) = \kappa \frac{d^2 y(x)}{dx^2} .$$
(Eq. 4.8)

Solutions to this equation, $y(x)$, under different forms of applied moment, specify the transverse displacement of the rod at any position *x* along the rod's undeformed axis. The solutions for two common situations relevant to the work in this chapter are described below.

4.2.3.1   Rod with one clamped end

One situation is where the rod is clamped at one end and a transverse force *F* applied at the free end, otherwise known as a cantilevered beam, Figure 4.3.
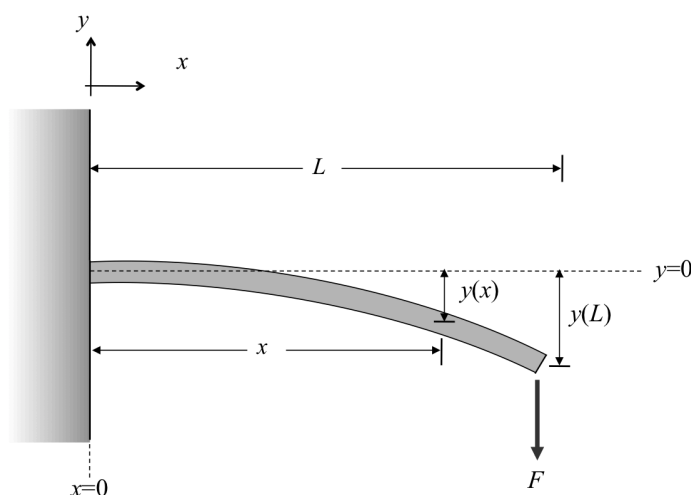
Figure 4.3: Geometry of the cantilevered beam.

A beam is fixed at the left end but free to be deflected at the right end under an applied force *F*. This system is the one typically used to model actomyosin cross-bridges, where the fixed end is often the tail end of the myosin head (*e.g.* 'anchored' in the sarcomere thick filaments or adhered to a surface in a trap assay) and the free end is the actin-binding motor domain to which force is applied (*e.g.* via movement of a bound actin filament). Figure adapted from Howard (2001).

This situation is relevant to the work in this thesis as it is the system considered when modelling the actomyosin crossbridge, such as in muscle structure or in single-molecule crossbridges formed, for example, in '3-bead' optical trap assays. Here, the myosin head – the beam, is considered anchored at one end (often the tail end) and the transverse force is applied at the other (often at the actin-binding motor end). The 'cantilever stiffness' determines how far the free-end tip will be displaced from its relaxed position.

For the cantilever beam arrangement, the bending moment term in Eq. 4.8 is $F \cdot (L-x)$, and the boundary conditions of the system are $y(0) = dy(0)/dx = 0$ (zero deflection and curvature at the clamped end). Eq. 4.8 then has the solution

$$y(x) = \frac{F}{2\kappa}\left(Lx^2 - \frac{x^3}{3}\right)$$

(Eq. 4.9)

for the shape of the beam under applied force. At the free end of the beam ($x = L$) the deflection is $y(L) = FL^3/3\kappa$ and so the deflection at the tip is proportional to the force. This defines an effective 'spring constant' or stiffness for the cantilever $k_{\text{cantilever}}$, relating the tip displacement to applied force

$$k_{\text{cantilever}} = \frac{F}{y(L)} = \frac{3\kappa}{L^3}.$$

(Eq. 4.10)

The cantilever stiffness thus depends on the flexural rigidity κ and is proportional to $1/L^3$. $k_{\text{cantilever}}$ is the per-head stiffness value often quoted for myosin heads in crossbridge arrangements with F-actin. For example, Veigel et al. (2002) measured $k_{\text{cantilever}} = 0.2 \pm 0.01$ pN/nm for the actomyosin 5a-S1 complex in an optical trapping assay, while Lewalle et al. (2008) measured 1.79 ± 0.06 pN/nm in a alternative trap assay of the actomyosin 2-S1 crossbridge. The latter measurement was in agreement with a value of ~1.7 pN/nm deduced from experiments on whole muscle fibres (Linari et al., 2007). For a full list of measured cantilever stiffnesses for different myosin 2 fragments see Table S1 (supplementary information) in Kaya and Higuchi (2010).

4.2.3.2   Fluctuating rod with free ends

Another important case to consider, the one employed in the lever fluctuation analysis in this chapter, is one in which the rod is free at both ends and immersed in a viscous fluid at thermal equilibrium. The mechanics and resulting motions of such a rod are well-studied in physics (Landau and Lifshitz, 1970).

As the rod is buffeted by collisions with fluid molecules it can deform by bending. The bending motion is perpendicular to the rod's long axis (in the small angle approximation) and so at each point along the rod there is a drag force (related to fluid viscosity and the

velocity of motion) opposing the motion. The drag forces create bending moments on the rod which can be equated with the bending moment, $M(x)$, in Eq. 4.8, see Howard (2001). This leads to the hydrodynamic (small angle) beam equation for the bending rod's dynamics

$$\frac{\partial^4 y}{\partial x^4} = -\frac{c_\perp}{\kappa}\frac{\partial y}{\partial t},$$

(Eq. 4.11)

(see Appendix 6.2 of Howard (2001) for derivaition), where $c_\perp$ is the perpendicular drag coefficient per unit length and t is time. Solutions to this fourth-order differential equation, with appropriate initial and boundary conditions, describe the time evolution of the rod's shape, $y(x, t)$. For an unconstrained rod, the relevant boundary conditions (obtained by minimisation of the energy functional) are:

$$\left.\frac{\partial^2 y(x,t)}{\partial x^2}\right|_{x=0,L} = 0, \qquad \left.\frac{\partial^3 y(x,t)}{\partial x^3}\right|_{x=0,L} = 0.$$

(Eq. 4.12a, b)

The physical interpretation of the boundary conditions are that there are no net torques (*i.e.* shear forces) or forces at the rod ends. The solution to Eq. 4.11 under these conditions is a set 'hydrodynamic modes' of the form $y(x,t) = \exp(-t/\tau_n)W_n(x)$, with a time-dependent part, $\exp(-t/\tau_n)$, characterised by a decay constant τ_n (set by the ratio of κ to the viscosity of the environment), and a spatial part $W_n(x)$. The full expressions for the modes, as calculated previously in Gittes et al. (1993) are:

$$y_n(x,t) = e^{-t/\tau_n}\left[\sinh\alpha_n\cos\frac{2\alpha_n}{L}(x-\frac{L}{2}) - \sin\alpha_n\cosh\frac{2\alpha_n}{L}(x-\frac{L}{2})\right]\ (n\ \text{odd}),\ \text{(Eq. 4.13a)}$$

$$y_n(x,t) = e^{-t/\tau_n}\left[\cosh\alpha_n\sin\frac{2\alpha_n}{L}(x-\frac{L}{2}) + \cos\alpha_n\sinh\frac{2\alpha_n}{L}(x-\frac{L}{2})\right]\ (n\ \text{even})\ \text{(Eq. 4.13b)}$$

where α_n is the nth root of the equation

$$\tan\alpha_n = (-1)^n\tanh\alpha_n,$$

(Eq. 4.14)

and the time constant τ_n is

$$\tau_n = \frac{c_\perp}{\kappa}\left(\frac{L}{2\alpha_n}\right)^4.$$

(Eq. 4.15)

α_n is closely approximated by the solutions

$$\alpha_n \approx \left(n+\frac{1}{2}\right)\frac{\pi}{2},\ (n=1,2,3\ldots).$$

(Eq. 4.16)

134

Each individual mode, or a linear sum of different modes, solves the hydrodynamic beam equation under the free-ends boundary conditions. The physical interpretation of the modes is that at any time t, the shape of the rod ($y(x,t)$ or $\theta(s,t)$) can be written as a linear sum of the modes, in the manner of a Fourier series representation of a time-varying signal. The exponential factors, $\exp(-t/\tau_n)$, contain a decay time constant, τ_n, which is in turn dependent on the mode number n, by $\tau_n \propto (1/\pi(n+\frac{1}{2}))^4$ in the α_n approximation. Intuitively (noting the inverse fourth power dependence of τ_n on mode number), the high frequency modes (shapes with more bends and thus greater bending energy) decay faster than the lower modes as the rod relaxes. In an ensemble of rod shapes one would therefore expect the lower mode shapes to dominate. Plots of the first four hydrodynamic mode solutions (using the α_n approximation) for an unconstrained fluctuating rod are shown in Figure 4.4.
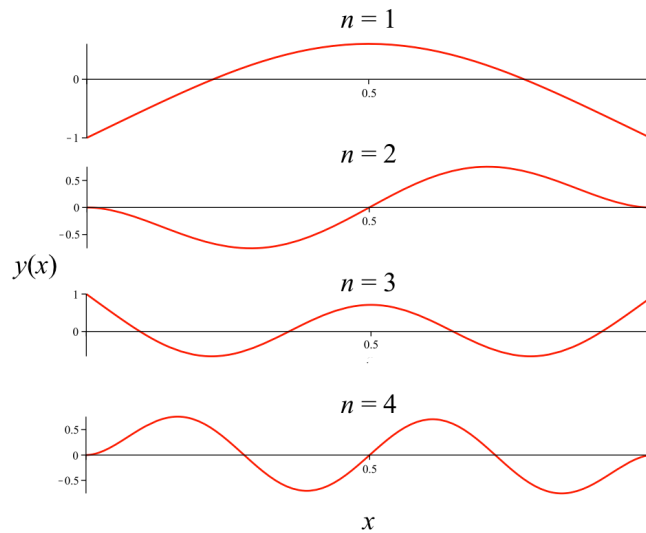
Figure 4.4 Plots of the first four hydrodynamic modes for a rod with free ends.
  The modes are computed from Eq. 4.13 using the α_n approximation. L was set to 1 and the mode amplitudes were normalised for ease of view.

Alternative mode solutions to the hydrodynamic beam equation can also be derived for a variety of different end conditions, such as for a rod with one clamped end (the cantilever case), for one or two hinged ends, or for ends with torques applied, see Appendix B of Wiggins (1998) for a review of general solutions to all cases. In the remainder of this chapter, only the free ends situation will be considered.

4.2.4 Time-independent representation of rod shape by Fourier series

While the hydrodynamic modes define the *dynamics* of a fluctuating unconstrained rod, it is also possible to describe any snapshot rod shape as a linear sum of simpler Fourier "modes" (sine or cosine functions). This is useful if one wants to analyse an ensemble of

shapes independently of time, as will be the case when looking at EM micrographs of negatively-stained molecules later in this chapter.

Any unconstrained rod shape $\theta(s)$ can be expressed as a sum of cosine modes

$$\theta(s) = \sum_{n=0}^{\infty} \theta_n(s) = \sqrt{\frac{2}{L}} \sum_{n=0}^{\infty} a_n \cos\left(\frac{n\pi}{L}s\right),$$

(Eq. 4.17)

where the mode amplitudes coefficients, a_n, are given by

$$a_n = \sqrt{\frac{2}{L}} \int_0^L \theta(s)\cos\left(\frac{n\pi}{L}s\right)ds, \ n \geq 1,$$

(Eq. 4.18)

(Gittes et al., 1993). Note that the choice of only cosine modes in Eq. 4.17 is appropriate for the unconstrained rod, since they meet the zero-curvature boundary conditions for free-ends. For other end constraints a sine or mixed sine-cosine series would be required.

4.2.4.1 Relation between mode amplitudes and flexural rigidity

Decomposition of the rod shape into cosine modes is useful as it reveals an important relation between the amplitude coefficients and the flexural rigidity. By differentiating Eq. 4.17, inserting into Eq. 4.7 and integrating, an expression for the bending energy in terms of the mode series is obtained

$$U = \frac{\kappa}{2} \sum_{n=1}^{\infty} \left(\frac{n\pi}{L}\right)^2 (a_n - a_n^0)^2.$$

(Eq. 4.19)

Here, a_n^0 is the nth mode amplitude coefficient of the rod's relaxed shape $\theta_0(s)$ (the shape in the absence of any applied or thermal forces). For $n \geq 1$, the rod bending energy is thus a quadratic sum of the amplitude coefficients. Note that the 0th order mode does not contribute to the bending energy sum (it adds only a constant to $\theta(s)$ in Eq. 4.17 and merely determines the average orientation of the rod in space). Note also that this relation only holds if the rod is isotropically flexible (*i.e.* κ constant), the key assumption made in formulating the energy expression Eq. 4.7.

The equipartition theorem states that, in equilibrium, each independent quadratic term in an energy expression contributes on average $k_bT/2$ (k_b Boltzmann constant, T absolute temperature) to the total (Reif, 1965). Since the squared amplitude coefficient term in Eq. 4.19 is such a term, this implies

$$\frac{\kappa}{2}\left(\frac{n\pi}{L}\right)^2 \left\langle (a_n - a_n^0)^2 \right\rangle = \frac{k_bT}{2}$$

$$\Rightarrow \ \sigma_n^2 \equiv \left\langle (a_n - a_n^0)^2 \right\rangle = \frac{k_bT}{\kappa}\left(\frac{L}{n\pi}\right)^2 \ \text{for } n \geq 1,$$

(Eq. 4.20)

where the angle brackets denote an average and the quantity $<(a_n\text{-}a_n^0)^2>$ is recognised as the *variance* (*i.e.* the S.D. squared) of the amplitude coefficients for mode *n*. A relationship between the amplitude coefficients and the flexural rigidity is thus established. The relation in Eq. 4.20 can serve as a test of the unconstrained isotropic rod model: the mode amplitude coefficients for a set of shapes would be expected to vary as $\sigma_n^2 \propto 1/n^2$ if the system is consistent with an unconstrained isotropic rod. Deviation from this relationship would imply a different underlying mechanical system. If the isotropic rod model does hold, Eq. 4.20 can provide a means of estimating the flexural rigidity directly from the variance of the amplitude coefficients of different modes, since rearranging Eq. 4.20 yields

$$\kappa = \frac{k_b T}{\sigma_n^{\,2}} \left(\frac{L}{n\pi} \right)^2 . \tag{4.21}$$

4.2.4.2   Assumption of ergodicity

When invoking the equiparition theorem to derive Eq. 4.20, an average over sets of mode amplitudes must be calculated. In previous bending mode studies, a *time average* has usually been taken to calculate σ_n^2 *e.g.* mode amplitudes are computed and averaged for a series of video frames of a single fluctuating filament (Brangwynne et al., 2007; Fakhri et al., 2009; Gittes et al., 1993; Janson and Dogterom, 2004; Käs et al., 1996; Ott et al., 1993). However, in this work, the average must instead be taken over a large collection of independent particles (micrograph images of stained molecules) which are assumed to be a static snapshot of the molecules undergoing 2D thermal fluctuations. Averaging over independent copies of an object forms an *ensemble average*. The *ergodic hypothesis* in thermodynamics states that over long enough time scales the two types of average are equivalent. That is, for long time scales, *ensemble average = time average,* and averaging across many molecules in different conformations on an EM grid is equivalent to averaging over the fluctuating shapes made by a single molecule in a time series. Ergodicity is thus another assumption underlying the analysis work in this chapter.

4.2.5   Relation between persistence length and flexural rigidity

Persistence length L_p is directly related to the flexural rigidity, and is an alternative and perhaps more intuitive measure for polymer stiffness. L_p is the characteristic length over which correlations in the tangential angle are nearly lost (*i.e.* bending due to thermal forces becomes appreciable). It is defined formally by the relation

$$\left\langle \cos[\Delta\theta_{3D}(s)] \right\rangle = \exp\left(-\frac{s}{L_p} \right), \tag{4.22}$$

where $\Delta\theta_{3D}$ is the angle between two tangent vectors to the polymer separated by arc length s. L_p can be visualised as the arc length at which thermal fluctuations cause a polymer to bend away from a straight conformation by roughly 68° on average (when $\langle\cos[\Delta\theta]\rangle = e^{-1} \Rightarrow \Delta\theta \approx 68.4°$). Hence, a stiff polymer has a larger persistence length than does a floppy one. If the polymer is considered confined to a plane (*i.e.* fluctuating in two-dimensions rather than three), the cosine of the angle takes twice as long to de-correlate as there is one less degree of directional freedom. In this case the relation is

$$\langle\cos[\theta(s) - \theta(0)]\rangle = \exp\left(-\frac{s}{2L_p}\right). \tag{Eq. 4.23}$$

Following the calculation in Appendix 6.3 of Howard (2001), by expanding the expression for a small change in the cosine function, and applying the equipartition theorem, L_p is shown to be related to flexural rigidity by

$$L_p = \frac{\kappa}{k_b T}. \tag{Eq.4.24}$$

This accords with the intuitive meaning of L_p, as a ratio of the flexural rigidity to the typical thermal energy causing bending in the system. Using Eq. 4.21 in Eq. 4.24, an expression for L_p in terms of the Fourier amplitude coefficients for a set of bent rod shapes can be written

$$L_p = \frac{1}{\sigma_n^2}\left(\frac{L}{n\pi}\right)^2. \tag{Eq 4.25}$$

This expression will be used later to estimate L_p from mode amplitude variances (and a knowledge of L).

4.2.6   Summary of mathematical framework

In the above section, a mathematical framework for considering the shapes made by an fluctuating isotropic rod has been developed. In particular, attention has been paid to a system in which the rod has free-ends but is constrained to lie in a plane, and where bending fluctuations are driven by the thermal energy of a surrounding viscous fluid. This is the system and the assumptions made later in this chapter when examining the shapes of myosin 5a molecules negatively-stained on the carbon substrate of an EM grid.

The mathematical framework developed started from a description of the rod's shape using the concept of the tangential angle at a given arc length. The local curvature (or equivalently the inverse radius of curvature) was then defined as the derivative of the tangential angle with respect to arc length, and is a useful quantity that will be used as a

measure of molecule curvature later. The beam equation was introduced and relates the bend induced in the rod to the moment applied. The proportionality constant between bending and applied moment is the flexural rigidity, and the concept of a bending energy follows from this.

For small angle bending, the dynamics of a thermally bending rod are governed by the hydrodynamic beam equation, solutions to which are a set of modes that give the time evolution of the rod's shape. The form of the modes depends on the specific boundary conditions of the system, *e.g.* free or clamped ends. At any time, the rod's shape can be considered to be formed from a linear sum of these modes. For a time-independent snapshot of bending rod shapes, a simpler cosine series expression of each shape is also possible. In this case, for a thermally bending isotropic rod (*i.e* with constant flexural rigidity), the variance of different mode amplitudes is expected to vary as $1/n^2$ (for mode number n). For such a system, the persistence length can be estimated directly from the variance of the mode amplitudes.

Table 4.1: Theoretical concepts used for analysis of myosin 5a lever domain EM data

Theoretical concept	How used in analysis in this chapter
Local curvature: $\rho = 1/R = d\theta/ds \; (\cong \; d^2y/dx^2)$	To quantify curvature along the length of the myo5a lever domain in negative stain EM images. To address whether the lever exhibits isotropic, uniform bending.
Variance in cosine mode amplitude coefficient values: $\sigma_n{}^2 \equiv \left\langle (a_n - a_n^0)^2 \right\rangle$	(i) To validate whether lever shapes are characteristic of an isotropic bending rod via the expected relation $\sigma_n{}^2 \propto 1/n^2$ (ii) If applicable, to estimate the persistence length of the myo5a lever domain via the relation: $L_{\mathrm{p}} = \dfrac{1}{\sigma_n{}^2} \left(\dfrac{L}{n\pi} \right)^2$

In the following section the ideas from the theoretical framework are developed into a practical method for capturing and analysing images of fluctuating rod-like shapes from microscopy images of semiflexible rod-like molecules.

4.3   Bending rod shape analysis method

In this section, the theoretical concepts discussed in 4.2 are developed into an analytical method for approximating the curvature profile and estimating the implied flexural rigidity of bending rod-shaped objects from digital images. The method draws on ideas from previous studies of cytoskeletal filaments (Brangwynne et al., 2007; Gittes et al., 1993; Isambert et al., 1995; Käs et al., 1996; McCullough et al., 2008).

To implement the analysis, a computer program was written in the Perl programming language (Wall et al., 2000). The programme is designed to read sets of 2D curve coordinates as input data. In this case, they would be pixel coordinates from EM image averages, but they could originate from any source. From the coordinates, a set of parameters is computed, using simple geometry and numerical approximations. The parameters are stored for each curve analysed and a set of such curves can be processed before analysis of the ensemble as a whole.

4.3.1   Numerical approximation of curve shape

The starting point for the shape analysis method is that the 2D curve of a rod-like molecule (protein domain or filament) has been captured in a set of N coordinates $\{(x_m, y_m)\}$ that specify the path of the curve. The means of capturing the curve coordinates is a separate pre-processing step, and might be accomplished by some automatic shape-recognition procedure, *e.g.* Brangwynne el al. (2007), or, as was the case here, simply by marking the coordinate points by hand using appropriate image software (4.4.2).

By joining the coordinates with straight segments, the curve's path is approximated by a rectilinear one, and the density of points in the curve will determine the accuracy of the approximation.
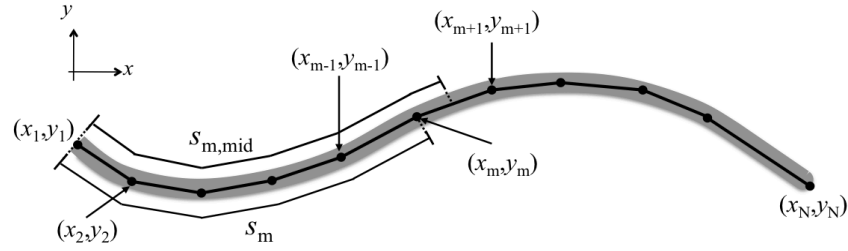
Figure 4.5: A smooth curve (*e.g.* a molecule shape or filamentous object) marked by a series of coordinates.

By joining the coordinates with straight line segments the rectilinear curve approximates the true path of the smooth curve. Arc lengths from the 1st point to the mth point (s_m) and from the first point to the middle of the mth section ($s_{m,mid}$), can be approximated from the sum of straight line segment lengths. The coordinate system used is inherited from the image source, and is defined such that the x-axis is parallel to the relaxed shape (straight) rod axis.

Figure 4.5 shows that s_m, the arc length from the 1st point to the mth point, or $s_{m,mid}$, the arc length from the 1st point to the midpoint of the mth section, can be approximated from sums of the straight line segment lengths:

$$s_m = \sum_{i=1}^{m-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \qquad m = 2, 3,\ldots, N\text{-}1, \qquad \text{(Eq. 4.26)}$$

where s_1 is defined as 0 and

$$s_{m,mid} = s_m + \frac{1}{2}\sqrt{(x_{m+1} - x_m)^2 + (y_{m+1} - y_m)^2} \quad m = 1, 2\ldots, N\text{-}1. \quad \text{(Eq. 4.27)}$$

Next, for each point except the last, a tangential angle can be approximated from θ_m, the angle the mth segment forms with the x-axis (*NB*: angles formed by anticlockwise rotation from the x-axis are defined positive, clockwise from the x-axis as negative)

$$\theta_m(s_{m,mid}) = \tan^{-1}\left(\frac{y_{m+1} - y_m}{x_{m+1} - x_m}\right) \qquad m=1, 2, \ldots, N\text{-}1. \qquad \text{(Eq. 4.28)}$$

Figure 4.6 illustrates how the segment angles are only an approximation for the true tangential angles of the smooth curve sections between pairs of points (the true tangential angle will always vary between any pair of points). The closer the coordinate points are spaced along the curve, the more accurate the segment angle approximates the tangential angle at the segment midpoint. The tangential angles approximated by θ_m are most logically assigned to the position of the midpoint of the mth section (at arc length $s_{m,mid}$).
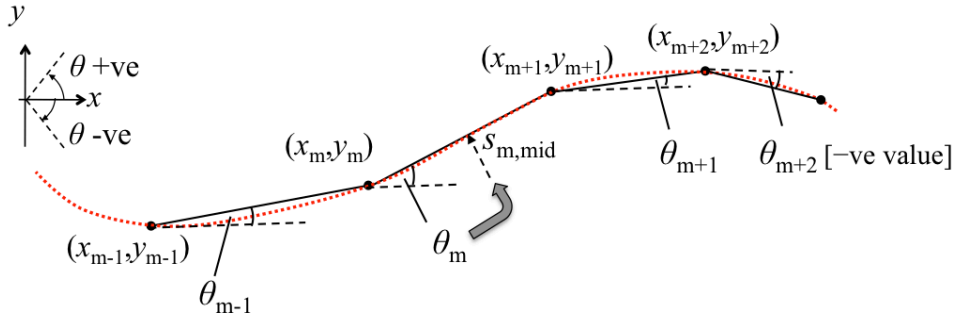
Figure 4.6: Tangential angles of a section of smooth curve approximated by straight segment angles θ_m.

The red dashed line is the path of the curve approximated by the straight line segments (black). The tangential angles at the segment midpoints (at arc length $s_{m,mid}$) are approximated by the segment angles θ_m. Note that in the coordinate system, angles rotating anticlockwise from the x-axis are defined as positive.

Using the set of segment angles, a set of local curvature values ρ_m can be calculated for each arc length s_m along the curve, using the relation

$$\rho_m(s_m) = \frac{1}{R(s_m)} = \frac{d\theta_m}{ds_m} \approx \frac{\Delta\theta_m}{\Delta s_m} \, , \quad m = 2, 3\ldots, N\text{-}1, \qquad \text{(Eq. 4.29)}$$

where, as illustrated in Figure 4.7, $\Delta\theta_m$ is defined as the change in tangential angle in passing through point m, and Δs_m is the accompanying change in arc length, calculated as the sum of the two half segment lengths adjoining point m:

$$\Delta\theta_m = \theta_m - \theta_{m-1}, \quad m = 2, 3,\ldots, N\text{-}1, \qquad \text{(Eq. 4.30)}$$

$$\Delta s_m = \frac{1}{2}\left[\sqrt{(x_{m+1} - x_m)^2 + (y_{m+1} - y_m)^2} + \sqrt{(x_m - x_{m-1})^2 + (y_m - y_{m-1})^2}\right] \qquad \text{(Eq. 4.31)}$$

$$m = 2, 3\ldots, N\text{-}1.$$
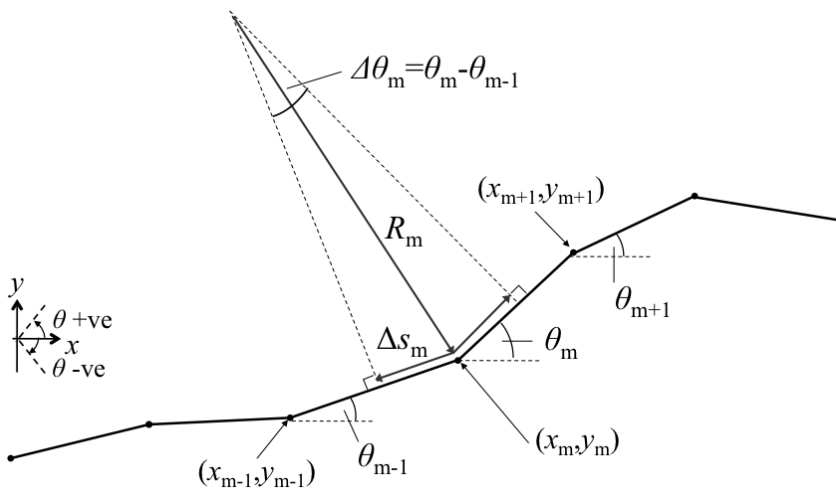
Figure 4.7: Section of rectilinear curve defining parameters for the calculation of local curvature ρ_m ascribed to position s_m.

The curvature at point (x_m, y_m) (at arc length s_m) is approximated from $\Delta\theta_m/\Delta s_m$.

$\Delta\theta_m$ and Δs_m are defined for all curve points except the first and the last (*i.e.* for $m = 2$, $3\ldots$, N-1), since three points (one either side of point m) are required for their calculation. Note that since ρ_m is formed from the difference of two consecutive angles, and angles can be negative or positive in relation to the *x*-axis, R_m and ρ_m may also be positive or negative depending on which direction the line is curving – positive curvature when R_m is drawn above the curve pointing downwards and negative curvature when R_m points upwards from below the curve. Note also that ρ_m values are most logically ascribed to arc lengths s_m (to the *start* of segment *m*) unlike the θ_m tangential angle approximations which are better ascribed to arc length $s_{m,mid}$ (to the *midpoint* of segment *m*).

4.3.2 Numerical approximation of cosine modes

In addition to the above values, for each curve, a set of cosine mode amplitudes corresponding to a Fourier series decomposition of the curve's shape $\theta(s)$ were also derived. This was done to allow the calculation of the mode variances across the curve ensemble as a whole, and so offer a test of the validity of the isotropic rod model. The integral expression for the cosine mode coefficients, a_n, given in Eq. 4.18, was approximated in the analysis program by the finite sum

$$a_n \approx \sqrt{\frac{2}{L}} \sum_{m=1}^{N-1} \theta_m \Delta s_{m,seg} \cos\left(\frac{n\pi}{L} s_{m,mid}\right) \quad n = 1, 2, \ldots, N, \qquad \text{(Eq. 4.32)}$$

where *n* is the mode number, *L* the curve length (the sum of all segment lengths) and $\Delta s_{m,seg}$ the length of the mth segment calculated by

$$\Delta s_{m,seg} = \sqrt{(x_{m+1} - x_m)^2 + (y_{m+1} - y_m)^2} \quad , \quad m = 1, 2, \ldots, N\text{-}1. \qquad \text{(Eq. 4.33)}$$

The number of modes in the summation in Eq. 4.32 is cut off at a physical limit (rather than running to infinity). In practice, a limit of ~*N* modes is sufficient to approximate $\theta(s)$ well. Indeed, modes above a certain frequency threshold are non-physical oscillations (with sub-pixel wavelengths) and the higher modes approaching this threshold represent high frequency oscillations (experimental noise) arising from the imprecision in locating the curve's position, rather than representing the curve's essential shape.

4.3.3 Summary

The parameters calculated and stored for each curve processed by the analysis programme are summarised in Table 4.2.

Table 4.2: Data calculated and stored for each curve processed
The coordinate values (x_m, y_m) originate from an external source (*e.g.* traced EM images), while all other parameters are calculated using the expressions defined above.

Curve: ID								
x_m	y_m	$s_{m,mid}$	s_m	$\Delta s_{m,seg}$	$\theta_m(s_{m,mid})$	$\Delta\theta_m$	Δs_m	$1/R_m(s_m)$
x_1	y_1	$s_{1,mid}$	0	$\Delta s_{1,seg}$	$\theta_1(s_{1,mid})$	–	–	–
x_2	y_2	$s_{2,mid}$	s_2	$\Delta s_{2,seg}$	$\theta_2(s_{2,mid})$	$\Delta\theta_2$	Δs_2	$1/R_2(s_2)$
…	…	…	…	…	…	…	…	…
x_{N-1}	y_{N-1}	$s_{N-1,mid}$	s_{N-1}	$\Delta s_{N-1,seg}$	$\theta_{N-1}(s_{N-1,mid})$	$\Delta\theta_{N-1}$	Δs_{N-1}	$1/R_{N-1}(s_{N-1})$
x_N	y_N	–	–	–	–	–	–	–

Cosine mode coefficients of $\theta(s)$: $a_1, a_2, a_3, ..., a_N$

Curve length: L

Curve end-to-end distance: $[(x_N-x_1)^2+(y_N-y_1)^2]^{1/2}$

Curve class-weight: optional number (if using image averages)

– = value undefined

4.4 Application of analysis method to images of myosin 5a

In this section, the method adopted for the capture of myosin 5a lever shapes from negative stain EM images is described.

As described previously, the myosin 5a lever is formed from a single α-helix that projects out of the motor domain, onto to which six CaMs bind at the IQ motifs in the heavy chain. The global average of the head (see chapter 3) and the example class average in Figure 4.8 demonstrate that the lever structure appears 'lumpy', due to stain accumulation around the lobes of the individual closely-packed CaMs, which adopt different orientations along the lever helix. The α-helix itself is largely enveloped by the bound CaMs and individual sections of it are not seen directly by staining. It is also clear from both individual molecule images, and from class averages of the lever later in this chapter, that the lever can bend significantly, and exhibit a variety of curved shapes. The bending is *implied* by the positions of the stained CaMs leading away from the motor domain, but it is the trajectory of the unseen lever helix that is assumed to have deformed. Its curve defines the bending in the lever domain and the coordinates of its implied path are those sought for analysis here.
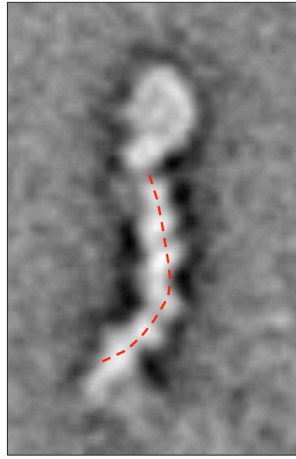
Figure 4.8: Example of lever classification average.

A class average of the HMM head formed from 60 images. The red dotted line marks the possible path of the underlying lever helix as implied by the six irregularly-shaped CaMs. The path is subject to interpretation, especially in the placement of the end point on the six CaM.

4.4.1 Consideration of automatic shape recognition procedure

In previous studies of images of longer filaments (primarily F-actin or microtubules) various automatic filament localisation techniques have been successfully employed, capturing the coordinates of filament shapes by automated image filtering and scanning algorithms *e.g* (Brangwynne et al., 2007; Fakhri et al., 2009; Janson and Dogterom, 2004; Käs et al., 1996; Ott et al., 1993). These techniques are possible if the filamentous objects under study are relatively uniform in structure (approximating a cylindrical profile at the magnification studied). The advantage of such automated methods is that many images can be quickly processed and that bias and errors associated with human interpretation of shapes are avoided.

Despite the more complex, inhomogeneous sub-structure of the myosin 5a lever, consideration was given to employing similar automated techniques in the analysis of the EM stain images here. In particular, a method of raster scanning, searching for the peak intensities in successive strips of the lever region, was tested. Here, the idea was to automatically assign the helix path coordinates from the peak intensity positions (located by *SPIDER*'s 'peak search' command) of successive strips of the lever region. However, this method was ultimately rejected as unreliable, the coordinates produced tending to oscillate too much about the implied path of the underlying lever helix. This effect arose due to the asymmetric shapes of some of the CaM appearances, and because in many orientations, *e.g.* when seen in the two-lobe dumbbell-like profile, the CaM's most intense lobe regions do not always coincide with the path of the underlying lever helix.

Consideration was also given to automatically assigning CaM shapes in an image to particular spatial orientations, by cross correlation comparison with a set of rotated CaM-

IQ crystal structure projections. The line of the lever helix might then be inferred from the matched structure. However, such an approach also proved unworkable, due to the high variability in CaM appearance and the variable dark stain outline around the CaM lobes, making matches to particular crystal structure orientations too imprecise.

Ultimately, given the lumpy and variable appearance of the CaMs, it was decided that the best means to trace the path of the lever helix was simply by marking the images manually. In this instance, the ability of human interpretation to make sense of the complex structure presented in an image, drawing on a knowledge of possible CaM appearances, stain outlines, and the feasible paths the lever helix could describe, was deemed superior to any automated recognition procedure.

4.4.2   Manual image marking method

A set of image averages from a classification of the lever appearance of aligned HMM head images were used for marking-out the variety of curves exhibited by the lever domain.

Firstly, to enable more precise hand-marking and to avoid artefacts arising from the constraints of the low resolution pixel grid (see 4.4.3), the source images were first interpolated, by bicubic Fourier-based interpolation (*SPIDER* command 'IP FS'), to ×8 original size (from 100×110 to 800×880 pixels). Files were then converted to 8-bit TIFF format and opened in *ImageJ* (v1.43r) graphics software for marking. The following processing steps were then performed, with the help of a specially written macro to semi-automate the process.

First, the path of the lever was marked out, progressing from motor-lever junction to the start of the tail, by clicking with a mouse along the implied lever helix path, using the 'segmented line' tool. Care was taken to be as consistent as possible across the images marked – always positioning the first mark at a similar point at the centre of the motor lever junction, and, where possible, ending with a mark on the distal edge of the 6th CaM, continuing the implied line of the lever helix towards the tail. The number of coordinates marked per image was variable, but averaged ~13 along the lever – that is ~2-3 marks per CaM traversed. Images were only marked if all 6 CaMs were clearly stained and their position could be unambiguously determined, otherwise the image was rejected. The coordinates marked along the lever (at ×8-size) (joined by straight line segments for display) were then exported and saved.

In addition to the hand-marked coordinates, a smoothed version of the same lever curve was then generated, by applying the *ImageJ* 'Fit Spline' command to fit a cubic spline through the active hand-marked points. Coordinates for this smooth spline (by default

restricted to integer pixel values) were also exported. In this case a much greater number of coordinates, with much smaller separation, are automatically generated – typically ~150 points per curve, *i.e.* at a resolution ~10 times greater than the hand-marked points. This extra step of spline-fitting and exporting the high resolution coordinates, was performed not only to create a smoothed version of the original rectilinear curve, but also to act as a high resolution reference curve, the coordinates of which could be re-sampled at a later stage, at a specified lower resolution (*i.e.* wider spacing), which could be made consistent across the curve set. The ability to control point spacing (density) proved to be important to the outcome of the curvature analysis, see 4.4.3.

Following the smooth spline-fit and extraction of the high resolution coordinates, an image of each class average, overlayed with the hand-marked coordinates and the fitted spline, was saved for reference. At this stage, trial analyses were performed to determine an appropriate coordinate re-sampling frequency, see 4.4.3. Once determined, spline coordinates were re-sampled at the selected frequency (*i.e.* every n^{th} coordinate retained) using a Perl script. The re-sampled coordinates were then passed to the main processing programme. At the outset of curve processing, two further transformations were applied: (i) *x* and *y* coordinates were switched so that the longitudinal axis of the lever was coincident with the horizontal *x*-axis rather than the image vertical (a requirement of the analysis) and (ii) the coordinate values were divided by the factor previously applied to enlarge the original images (*i.e.* ×1/8), scaling the curve coordinates back to the size of the original images.

4.4.3 Determination of appropriate point density

Following trial analyses of different lever curves, the importance of the effect of the number and spacing of points in the curve was noticed. Figure 4.9 illustrates how, if the high-resolution spline coordinates are sampled at different frequencies (*e.g.* taking every 20th coordinate or taking every 5th coordinate along the spline), the derived tangential angle and local curvature functions differ markedly. Competing effects, not generally discussed by previous studies, are at play here.

If one undersamples the curve coordinates, *e.g.* such that density of points in the curve is less than the original hand-marked coordinates, then the curve is artificially smoothed. That is, fluctuations on a length scale shorter than the point-spacing will be missed (*e.g.* see 1/20 sampling frequency plots in Figure 4.9(b)). Undersampling smoothes the tangential angle function (*i.e.* *changes* between angles are reduced), in turn reducing the magnitude of the derived curvature values. Also, any arc lengths associated with the curve will tend to be underestimated by fewer straight line sections.

Conversely, as one samples at an increasingly high frequency, approaching the point density of the high-resolution spline, an artificial roughness in the curve emerges. This is an artefact of the discrete pixel size in the source image. The exported spline coordinates are fixed to integer pixel values of the underlying pixel grid, and so as the points in the curve are sampled closer and closer together, the straight lines joining points are constrained to adopt only certain angles (since the endpoints of the segment lines must start and finish on integer coordinates). In the extreme case, the shortest segment can extend from a given pixel position to only one of its eight surrounding neighbours, defining only eight possible angles.

By interpolating the source image to a larger size prior to coordinate-marking, thereby increasing the number of pixels in the image, the impact of this roughening artefact can be confined to a length scale shorter than the curve fluctuations of interest. However, at a certain point-spacing threshold, the effect will still become significant, even for an interpolated image. The roughness artefact is well illustrated in Figure 4.9. Figure 4.9(c) shows a magnified section of a high-resolution spline curve in (b), and illustrates how, despite the curve appearing smooth at lower magnification, at high magnification its points are actually related by abrupt angle changes dictated by the constraints of the integer pixel coordinates. In Figure 4.9(b), as the sampling frequency increases above that of the original hand-marked coordinates, short length scale 'wobbling' between coordinates is magnified as fluctuations in the θ_m function, and magnified further still in the $1/R_m$ function (which is calculated from *changes* in θ_m) – see plots for 1/5 sampling and for the high-resolution spline coordinates themselves. As the density of points in the curve is increased above a certain sampling threshold, the magnitude of $1/R_m$ values increases dramatically (note the scale changes in the $1/R_m$ plots), and the function oscillates erratically. The $1/R_m$ oscillations are clearly non-physical and are not related to the original shape of the marked lever curve.

Because of the competing effects of over- and undersampling, for the curvature analysis to be usefully applied, an appropriate regime must be found where the number of points in the curves analysed yield θ_m and $1/R_m$ functions that capture only the essential shape of the object marked-out by hand, and one where the $1/R_m$ values obtained are not highly sensitive to changes in the number of points in the curve. In the case of the myosin 5a lever shapes considered here, trial sampling of a number of curves was performed and a spline sampling frequency of 1-in-12 (*i.e.* using every 12th spline coordinate), was deemed to be the most appropriate. The 1/12 re-sampled points were typically spaced slightly wider than the original hand-marked curves (*e.g.* the example in Figure 4.9(b) has one point less in the 1/12 curve than in the marked coordinate case).

Why trouble to fit splines and re-sample points if the hand marked coordinates are close to the optimum point-spacing regime? The main argument for the process is that unintended variations in point density across different images and along a single curve are 'evened-out', making comparison of curvature profiles across a dataset fairer. For example, in one instance a curve may have been sparsely marked by-hand, while in another case a similar-shaped curve (or curve section) may have marked with more closely-spaced points that fluctuate around the curve backbone more. Quite different $1/R_m$ functions would result from the different degrees of marking of similarly shaped images. Provided a regime is found where the point-sampling still captures the shortest length scale fluctuations in all the images, re-sampling coordinates from high resolution reference curves therefore provides a means of counteracting variability arising from interpretative hand-marking. Re-sampling curve points to an appropriate spacing is the one-dimensional analogue of the practice of filtering two-dimensional images to remove non-physical high frequency components (random noise) above the maximum resolution of the observation system used.

(a)

(b)

(c)
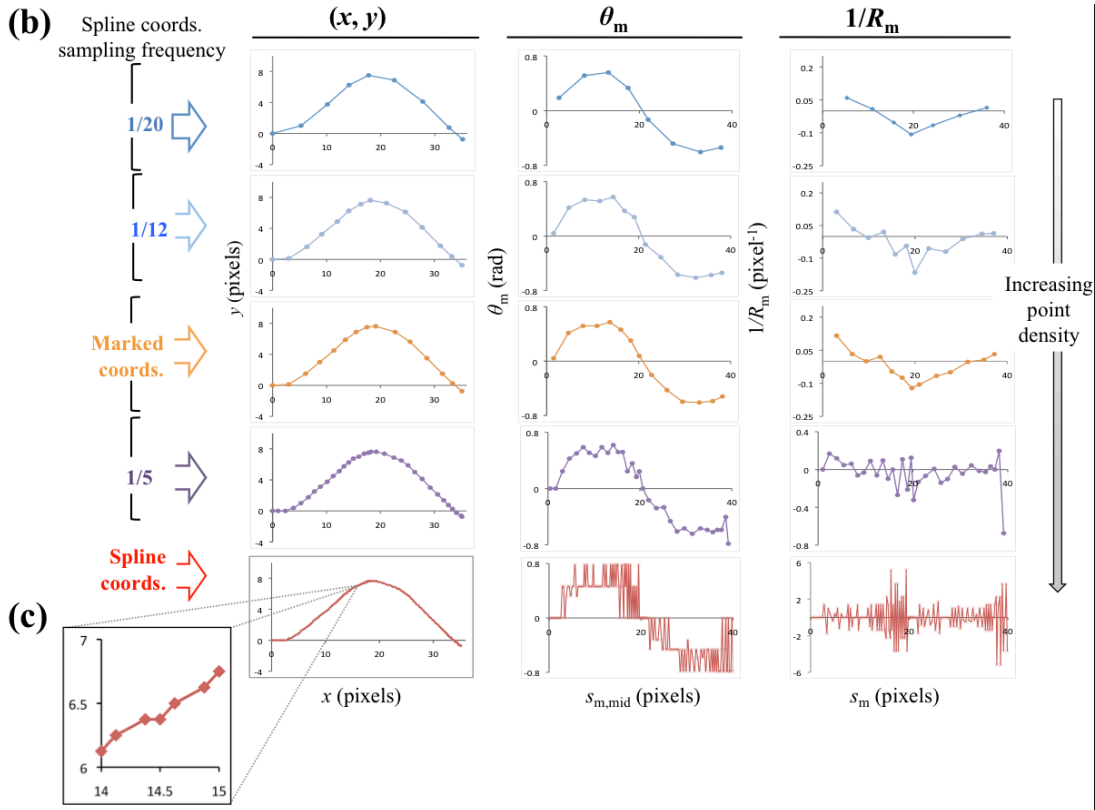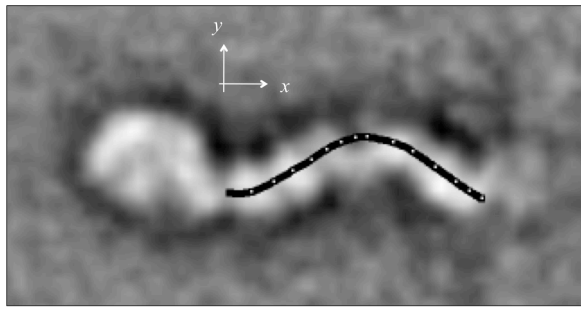
Figure 4.9: Artefacts of under and oversampling spline coordinates.

(a) example of a 62-image class average of the myo5a-HMM head with the curve of its lever marked by-hand (white dots) and its fitted spline overlayed (black). (b) plots of the curve coordinates (x,y), tangential angles (θ_m) and local curvature values ($1/R_m$) derived from the fitted spline coordinates when sampled at different frequencies along the spline (*e.g.* 1/20 = every 20th spline coordinate sampled) The original hand-marked coordinates and derived parameters are shown in the third row of graphs (orange points). As the points are sampled more closely, artificial roughness artefacts emerge as the points in the curve are constrained to integer pixels values in the original image. (c) shows an enlarged section of spline curve with short lengthscale roughness due to the integer pixel constraints of the original image. The short lengthscale roughness has a dramatic effect on the θ_m and $1/R_m$ functions. Note the increase in the y-axis scale of the $1/R_m$ plots as the point density is increased.

To implement the coordinate re-sampling discussed above, an additional Perl script was written. The script operates by reading-in the coordinate values of the high-resolution

spline curves and writing-out only every n^{th} coordinate, where n can be specified by the user. (*NB*: The first and last points of the spline are always retained). These re-sampled coordinate values were then passed to the main curve analysis script for further processing.

4.5 Results

4.5.1 Lever shape classifications

Before beginning lever shape classification, the 23,250 HMM head images, aligned by their lever domain, were split according to their principal motor domain orientations as established by the work in chapter 3 (3.5.1.1). That is, with motor orientated in View-1 ('face-in-profile') (10,119 images), in View-2 ('face-away') (8,734 images) and in View-3 (undefined) (4,397 images). Figure 4.10 is a reminder of the principal motor orientations found by negative staining.
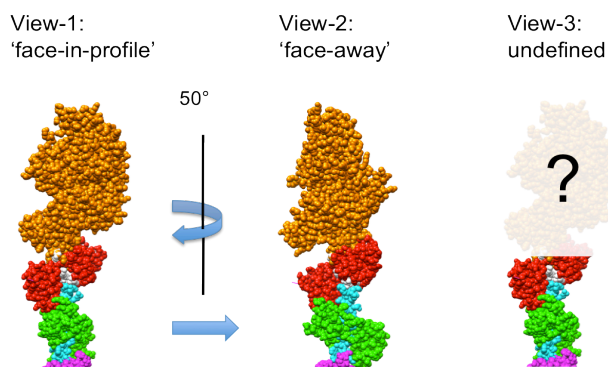
Figure 4.10: Principal motor domain orientations in negative stain EM

Although a conclusion from chapter 3 is that the lever tends to lie on the carbon of the EM grid in only one general orientation about its long axis (the motor rotating independently between its different views), it was never-the-less decided to maintain the separation of the molecules by motor domain orientation for lever shape classification and bending analysis. This was done in case there proved to be a correlated difference in the range or type of lever shapes seen for the different motor-orientated molecules. If no difference was discernable between views, results could later be aggregated.

The most populous View-1 images were classified first, using *SPIDER* to perform a K-means classification of the lever region appearance. An oval-shaped mask was drawn (using the variance image as a guide) (Figure 4.11) to define the region of interest and encompass the full range of expected lever shapes. Based on trial classifications, the 10,119 View-1 images were divided into 200 classes, such that there were ~50 images per class – the optimal balance judged between structure reinforcement from averaging and representation of variety of appearance within the images. View-1 classification results are shown in Figure 4.11. Molecules with motors in View-2 and View-3 were then similarly

classified, using the same mask and, in each case, dividing the images into sufficient classes to achieve ~50 images/class. Results are shown in Figure 4.12 and Figure 4.13.

In the classification averages, the main parts of the head are usually well defined. The CaMs marking the contour of the lever helix are frequently discernable as six irregularly-shaped spots extending from the larger half circle-shaped motor domain at the top of the heads as shown. Detail of the individual CaM lobes are only sometimes resolved. In most cases, averaging is sufficient that an entire lever is seen, with a few exceptions where the lever appears truncated or blurs towards its C-terminus end (caused by heterogeneity in the images averaged together in that class). An obvious feature of some averages is the presence of the C-terminus half of the lever belonging to the second head of the HMM molecule. This is identifiable as it joins the first head at the sixth CaM, frequently at an abrupt angle. Though usually less defined than the lever of the principal aligned head, the appearance of the second head indicates that there are groups of molecules with similar head-head angles. Though the classification mask was drawn to focus the classification on the appearance of the aligned lever, in some cases the second head will still enter the mask region and thus influence the classification. Some images may have been grouped together because they have a second head at a common angle and which lies close to the principal aligned head, rather than on the basis of the appearance of the principal head's lever.

For each motor orientation, a wide range of lever conformations are seen, clearly demonstrating a significant bending flexibility in the lever. In general, many of the levers appear straight, or almost straight, but there are also frequent examples of smooth curves to the left and right of the vertical axis, see Figure 4.14 for examples. Less frequently, a lever may curve to one side then the other, approximating an S-shape (or its long axis mirror), though the amplitude of fluctuations in these cases tends to be small. The range of curved shapes are reminiscent of modes of oscillation. The degree of curvature appears to vary smoothly across the averages, from straight to a pronounced curve, indicative of a continuum of possible shapes rather than a discrete set of lever conformations. The evenly populated classes are also consistent with a continuingly varying shape. Pronounced curvature examples appear less frequently than the straighter ones. Overall, the shapes form a distribution which suggests the lever is fluctuating about a straight (or near straight) minimum energy conformation. These fluctuations are probably thermally-induced, that is, driven by collision with solvent molecules undergoing Brownian motion.
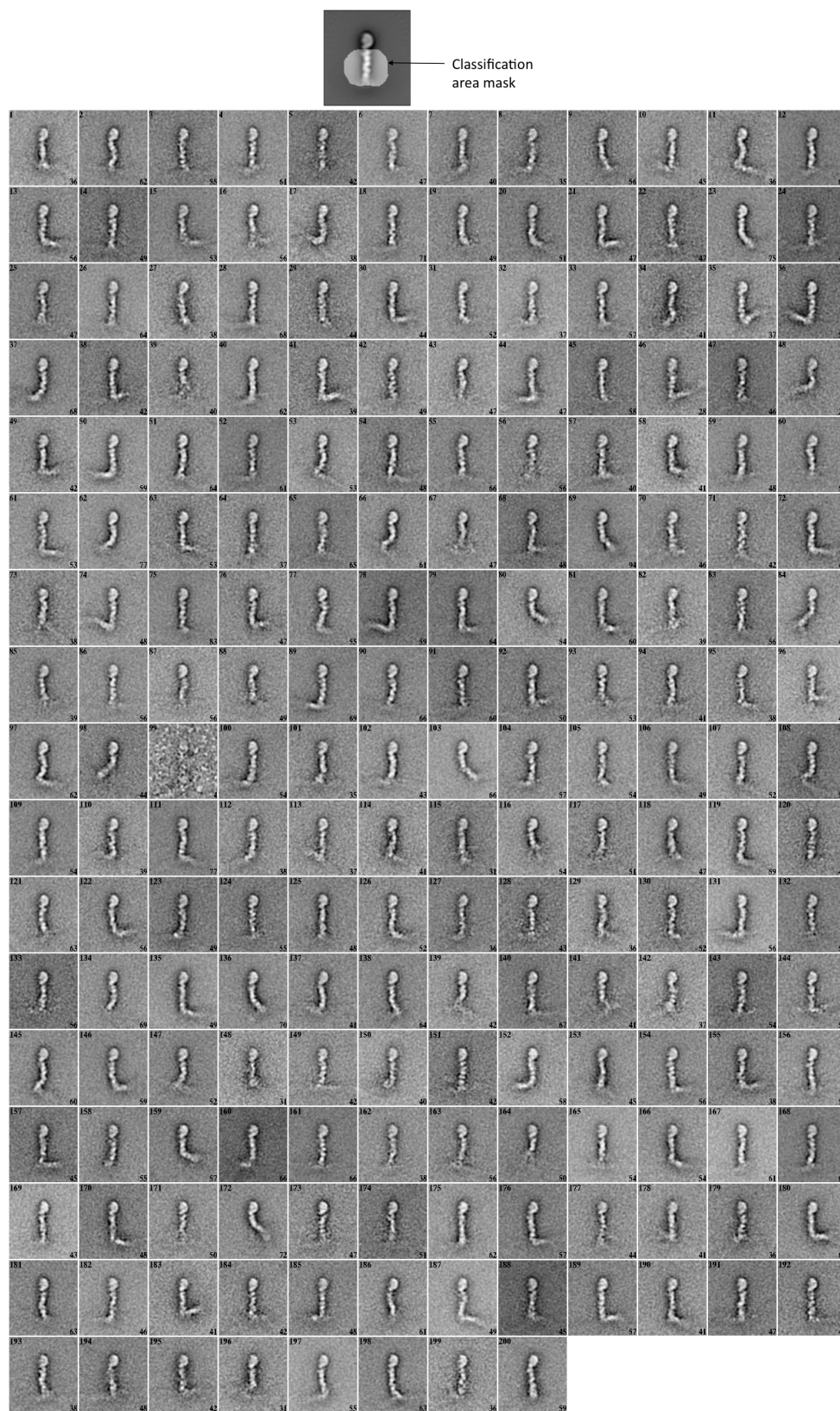
Figure 4.11: Lever region classification of View-1 heads (motor 'face-in-profile').

The global average image (all molecules) is shown above with the classification mask used to focus on the lever region superimposed. The 200 class averages produced by K-means classification by SPIDER are shown below. Numbers top left indicate class number, and numbers lower right are number of images contributing to the class.
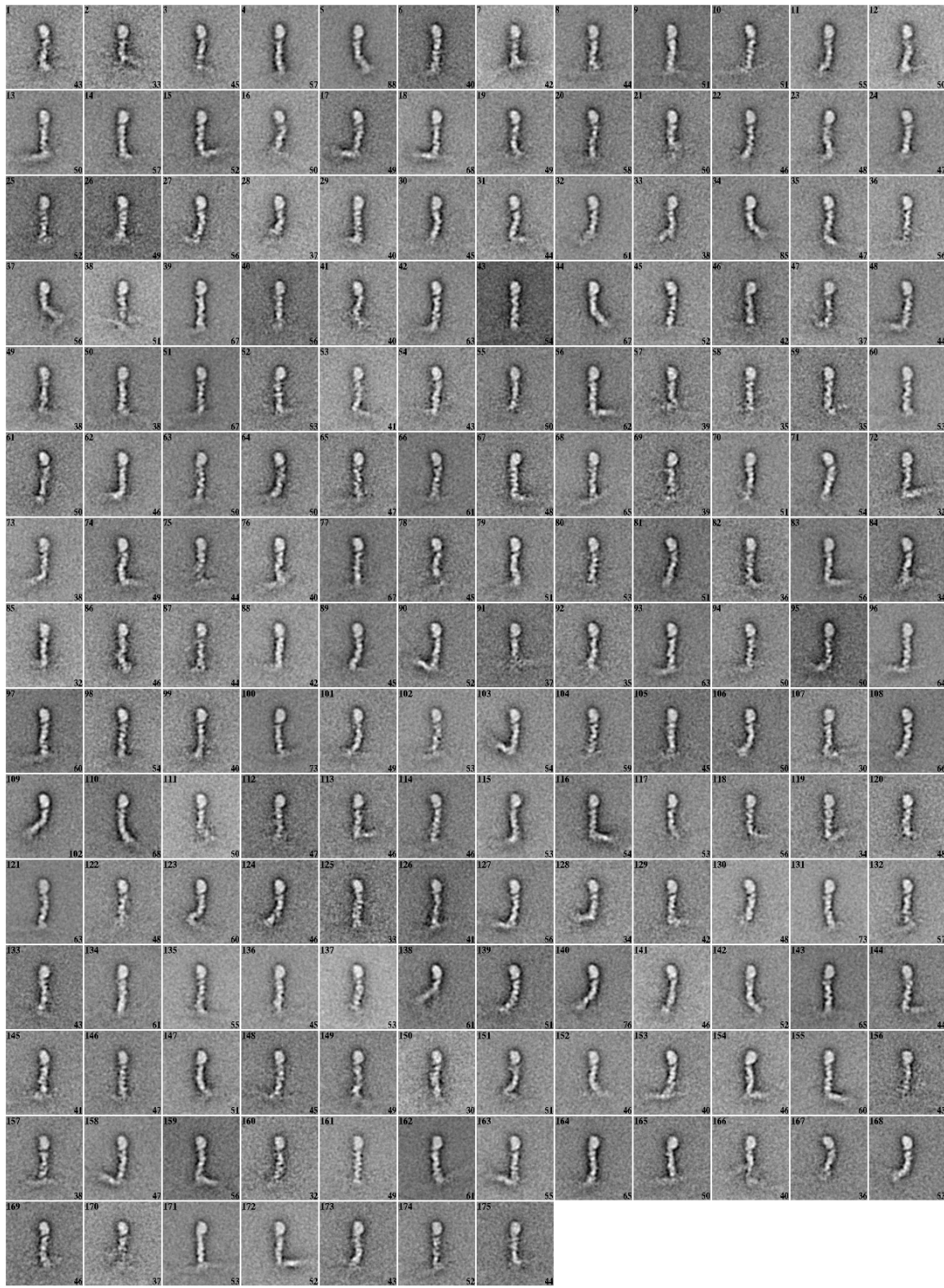
Figure 4.12: Lever region classification of View-2 heads (motor 'face-away')
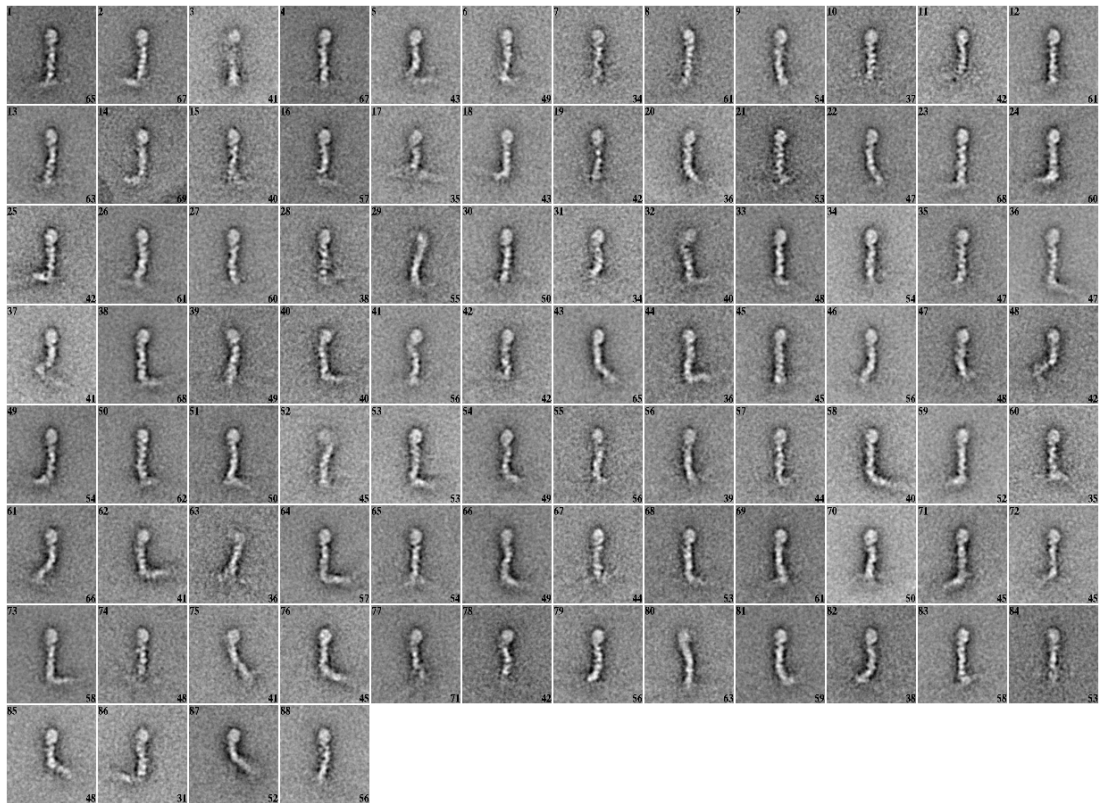
Figure 4.13: Lever region classification of View-undefined heads.

The generally smoothly curving shapes of most of the levers are contrasted with occasional averages where levers exhibit a sharp kink approximately half way along their length, see Figure 4.14 for examples. Smooth curvature is consistent with a lever that is isotropically and uniformly flexible, while the rare kinked examples suggest there may be particular pliant points in the lever's structure prone to deform under certain circumstances.

To ascertain the spread of different lever shapes and to see if there were significant differences in the types of shapes for different motor orientations, the averages in each montage were categorised by eye, according to whether they were: straight, curving left of vertical, right of vertical, S-shaped or shape undefined (*e.g.* because a lever section is missing or blurred in the average). *NB*: the rare kinked averages were placed in either the curve left or curve right classes. The number of separate molecule images contributing to each average were summed to enable the fraction of molecules in each shape category to be calculated. The respective distributions obtained for each motor orientation are shown in Figure 4.15.

Curve left

Straight

Curve right

S-shape

Kinked
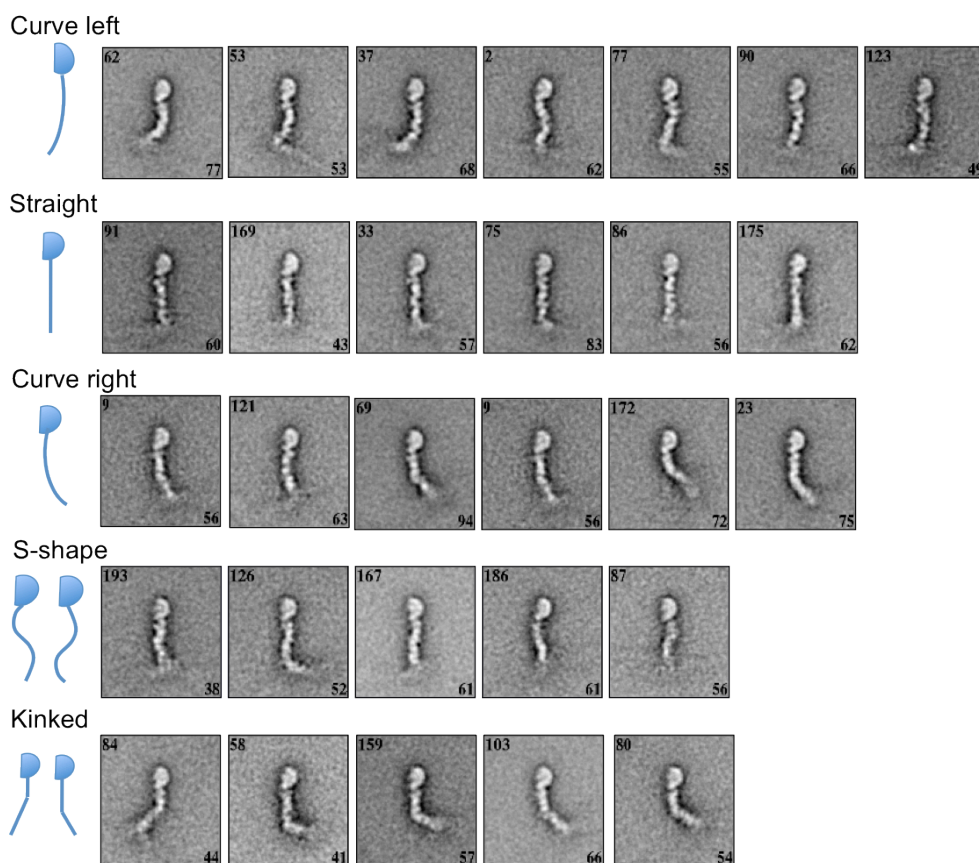
Figure 4.14: Example average lever shapes from classification of images with motor in view-1.

Selected examples of class averages exhibiting the principal types of lever shape found are shown. Cartoons of each lever shape category are depicted in blue left. All averages shown here are drawn from the classification of images which have a motor domain categorised as view-1 (face-in-profile). Class numbers are indicated in the top left of images, number of images forming the average bottom right.
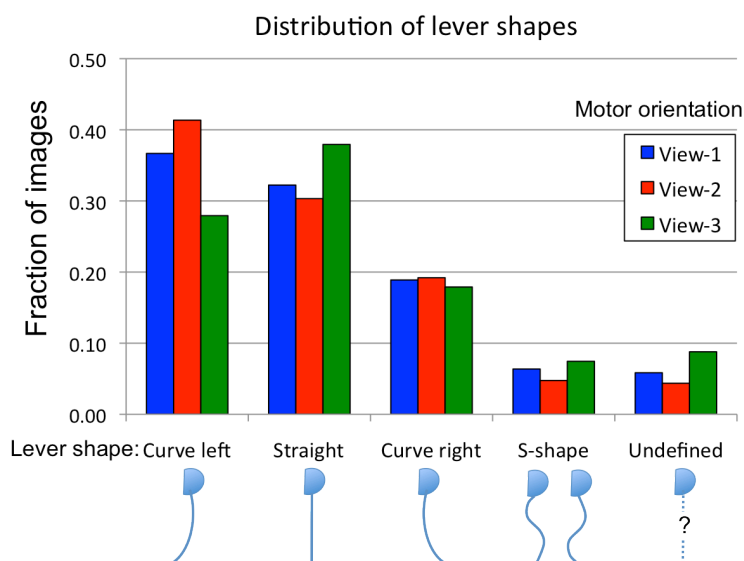
Figure 4.15: Relative frequencies of different lever shapes in classification averages.

The fraction of images falling into classes of different lever shapes are shown. Blue bars show the distribution of lever shapes for molecules with motors in view-1 (face-in-profile) (10,119 images), red bars are for view-2 (face-away) (8,734 images) and green bars are view-3 (undefined) (4,397 images).

The grouping of the class averages in the separate classifications into different lever shapes, Figure 4.15, shows a broadly similar distribution for each motor orientation. The only difference is that View-3 images (motor view undefined) have a majority of straight levers, whereas in motor view-1 and -2 classifications, images with levers curving left form the majority. The cause of this discrepancy is unclear, but its significance is reduced given that there are less than half the number of images in the view-3 distribution compared to view-1 and -2. The general similarity in shape distributions is consistent with the finding in chapter 3, that the general orientation of the lever is independent of the motor's orientation. Combining images for all motor views yields the overall shape distribution in Table 4.3.

Table 4.3: Average distribution of lever shapes

Curve left	Straight	Curve right	S-shape	Undefined
35 %	33 %	19 %	6 %	6%

A significant observation is that the shape distributions are not symmetric about a straight shape. The majority of images have levers exhibiting some degree of left curvature, straight being the second most populous shape. Levers with right curves are significantly less frequent (almost half) than those with right curves. If the lever shapes form a distribution around a minimum energy shape, then the class averages imply that the lever's 'relaxed' shape is curved slightly to the left when viewed in negative stain EM. This is consistent with the global average of all images, Figure 4.16, which presents the same information. Recognition of the relaxed lever shape as slightly curved has important consequences for calculation of the lever's persistence length in 4.5.3.
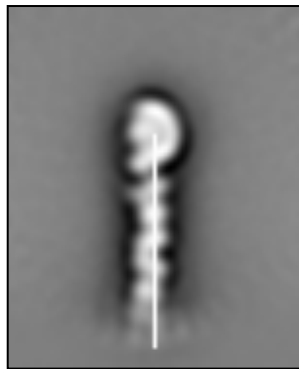
Figure 4.16: Global average of all images aligned by the lever domain

The image is an average of 23,250 free HMM head images aligned by their levers (see chapter 3). The white line is the vertical axis, added to indicate that, on average, the lever deviates from a straight shape. The vertical can be made to run through the centre of CaMs 2-4, but CaMs 1, 5 and 6 all fall to the left of the axis.

4.5.2 Mapping curvature in the myosin 5a lever

To address the question of whether the lever exhibits isotropic and uniform bending with length, lever shapes from the classifications were captured by hand-marking the coordinates of the implied lever helix, following the method explained in 4.4.2. Due to time constraints, the most populous motor View-1 class averages were chosen as a representative subset of data for this work.

Of the 200 class averages produced for View-1 images (Figure 4.11), 183 were able to be marked, the remaining 17 omitted due to the lever's contour being too ambiguous or ill-defined. Figure 4.17 shows representative examples of the lever curves captured. It is acknowledged that marking the path of the lever is an interpretative process. Marked curves were reviewed several times. Most difficulty was encountered in marking the direction of the lever's path at the C-terminal end through the sixth CaM. This is due to the sixth CaM generally being the most indistinct, and because the direction of the lever can depend on an interpretation of where the second head is, if it is visible in a given average.

The lever curve coordinates obtained by marking and spline-fitting were then processed using the curve analysing program described in 4.3. As explained in 4.4.3, the analysis depends on choosing a coordinate sampling density that adequately captures all the physical fluctuations in the lever's shape, but isn't oversampled such that non-physical roughness dominates the analysis. Through a number of trials, a coordinate sampling density of 1/12 (every 12th spline coordinate used) was found to be most appropriate.

Figure 4.18 illustrates the curve, tangential angle and curvature functions derived for example lever shapes.

(a)

(b)
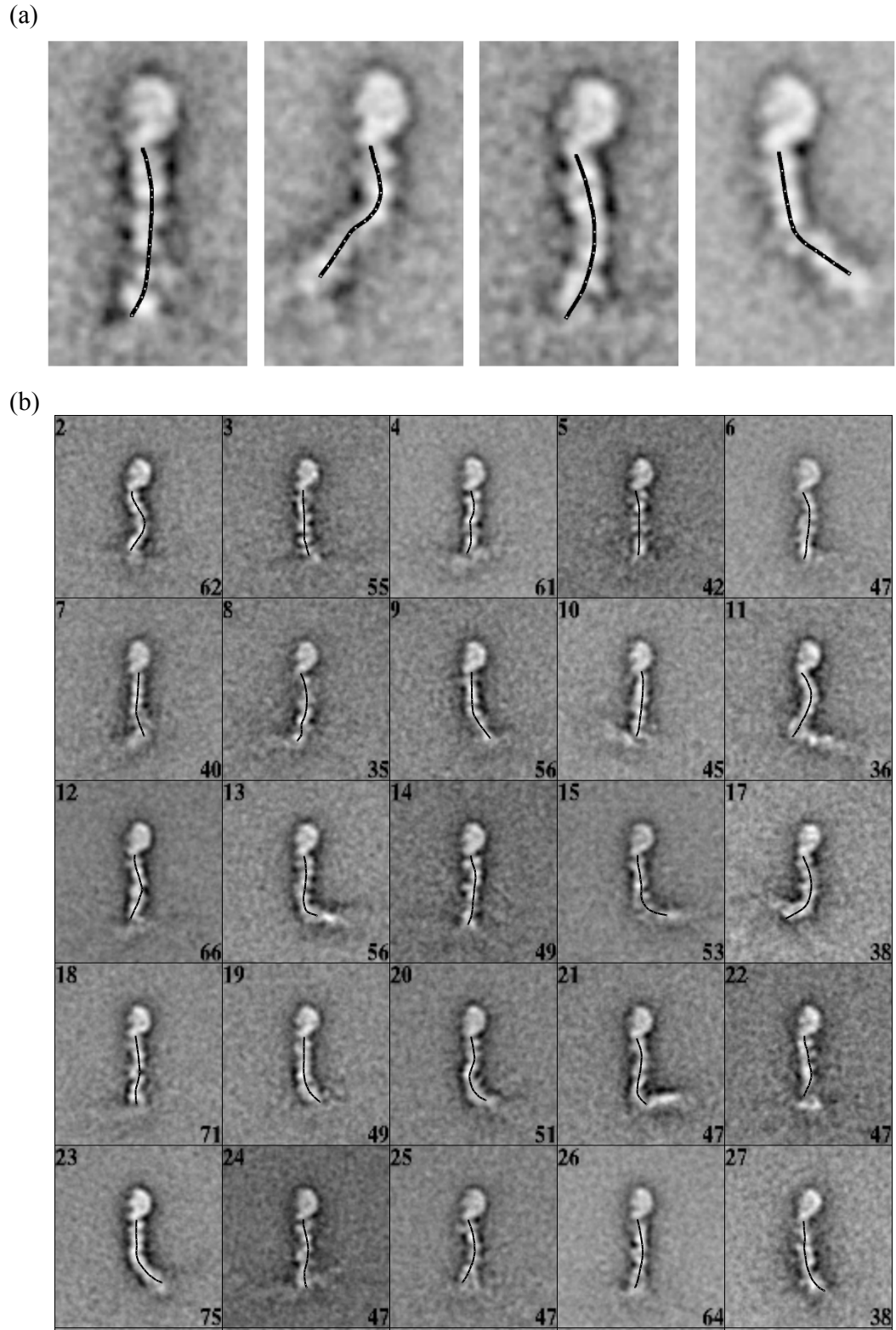
Figure 4.17: Example lever shapes marked by hand with fitted spline curves

(a) enlarged examples of View-1 class averages with differently curved lever shapes. Original hand-marked points are shown (white dots) over the spline fit through these points (black line). **(b)** montage of further View-1 class average examples as in (a). Only a representative selection of the 183 averages analysed are shown. Top left numbers identify the class, bottom right numbers are the number of images contributing to the class.
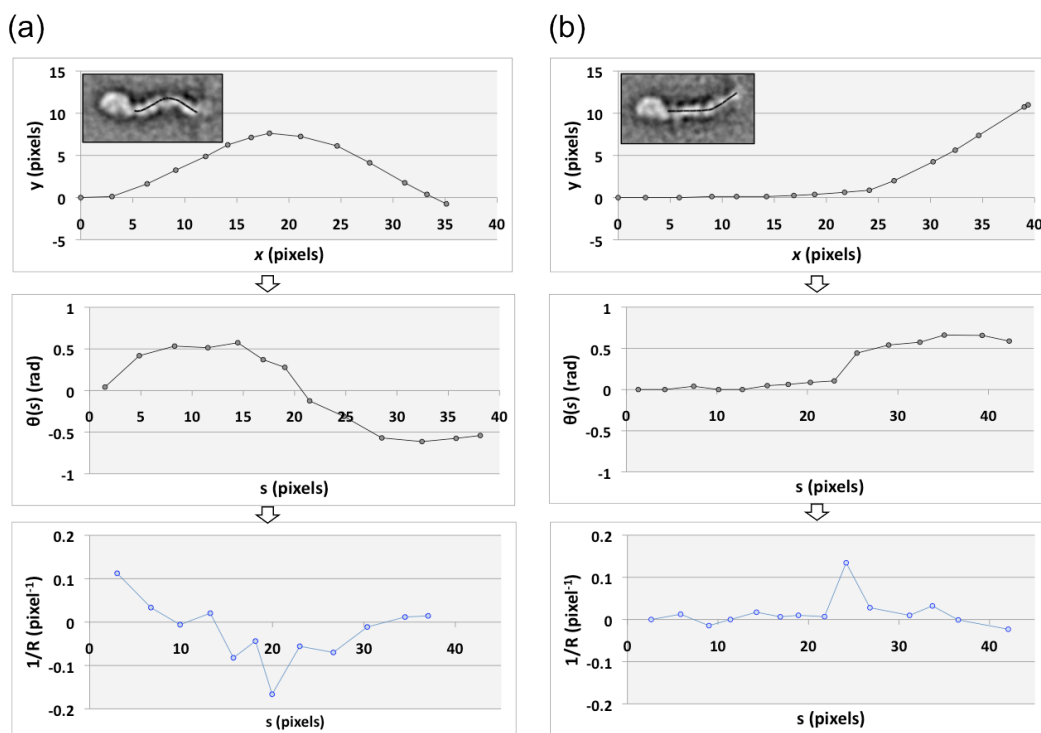
Figure 4.18: Example lever shapes approximations and derived $\theta(s)$ and 1/R(s) functions. Examples of two lever shapes are shown in separate columns (a) and (b).

4.5.2.1 Lever length value

A result of the capture of the lever shape data is that a length measurement for the lever region that extends beyond the motor domain can be obtained. This is useful as it can be compared with that predicted by the atomic model of Liu et al. (2006). It also provides a check on the validity of the class average images and indicates whether there are artefacts from staining, adsorption and drying (*e.g.* molecule stretching or shrinking). Figure 4.19 shows the distribution of lengths obtained from the class average data. The distribution is clustered about a peak value, but with a longer tail on the short side of the mean. This is likely due to blurring in the averages near the head-tail junction. The data yield a mean ± S.D of 20.6 ± 1.4 nm, for the length of the lever beyond the motor domain. This is lower than the value of 22.8 nm measured by Oke (2004), but lies within 1 nm of the 21.4 nm value predicted by measurement of the Liu et al. atomic model. The value measured here equates to a length of 3.43 nm per CaM-IQ.
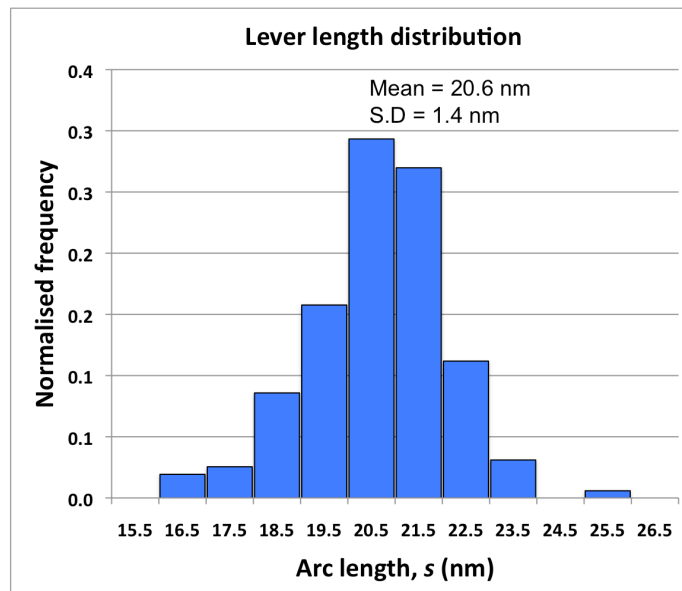
Figure 4.19: Histogram of lever lengths

Bin size 1 nm, *n* =183 image measurements. Arc length labels are the mid-point values of each bin. The data has been weighted according to class sizes and the frequencies are normalised (summing to 1).

4.5.2.2 Lever curvature profile

Each curve in the dataset contributes a set of curvature values 1/R at arc lengths along the lever. To determine if there are particular pliant positions (regions exhibiting high curvature) or anisotropies of bending in the lever, curvature values across the entire set of curves can be aggregated in one scatter plot. Figure 4.20 presents this information for the 183 curves in the motor View-1 dataset. Recalling that high curvature (small *R*) means more pronounced bending, the features of Figure 4.20(a) can be described.

The highest density of points is found clustered on or near the zero curvature (1/*R*=0) axis along the lever's entire length. As one moves further from this axis, towards $\pm$ 0.2 nm^{-1}, the density of points encountered at any arc length tends to decrease rapidly, points beyond $\sim \pm$ 0.2 nm^{-1} being high-curvature outliers. This pattern confirms the general observation made of the lever shape classifications (Figure 4.11-Figure 4.13) – that the lever adopts a near-straight equilibrium shape with fluctuations around this. The higher the curvature fluctuation, the lower the probability of occurrence. This is a picture consistent with thermally-driven bending of an elastic lever, where highly bent conformations are high energy states and straight conformations are near minimum.

(a)
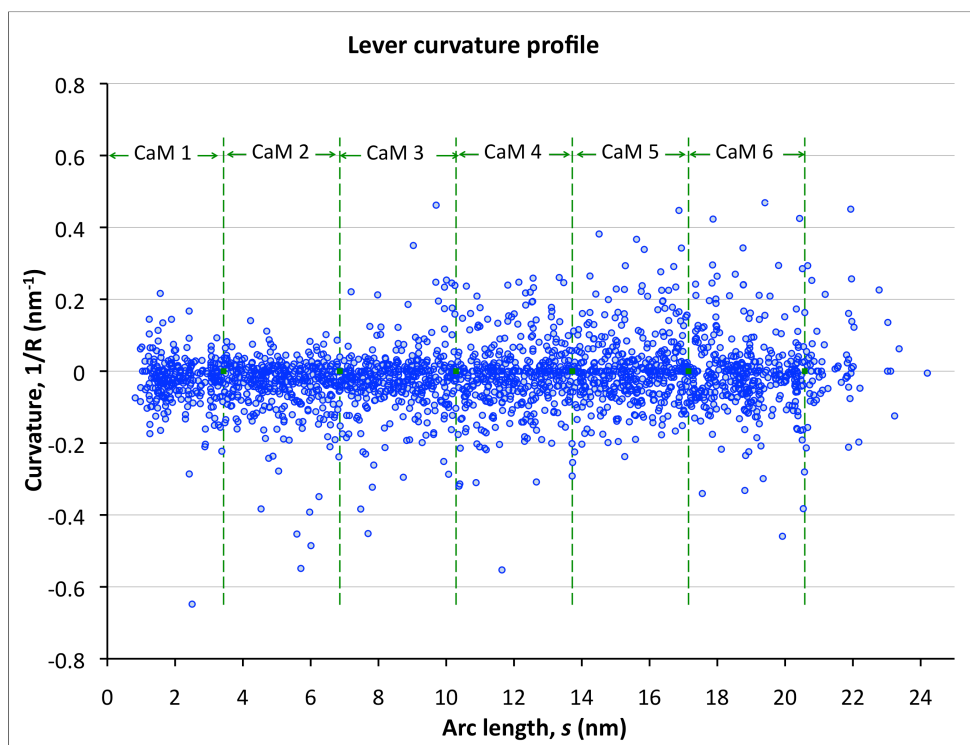
(b)
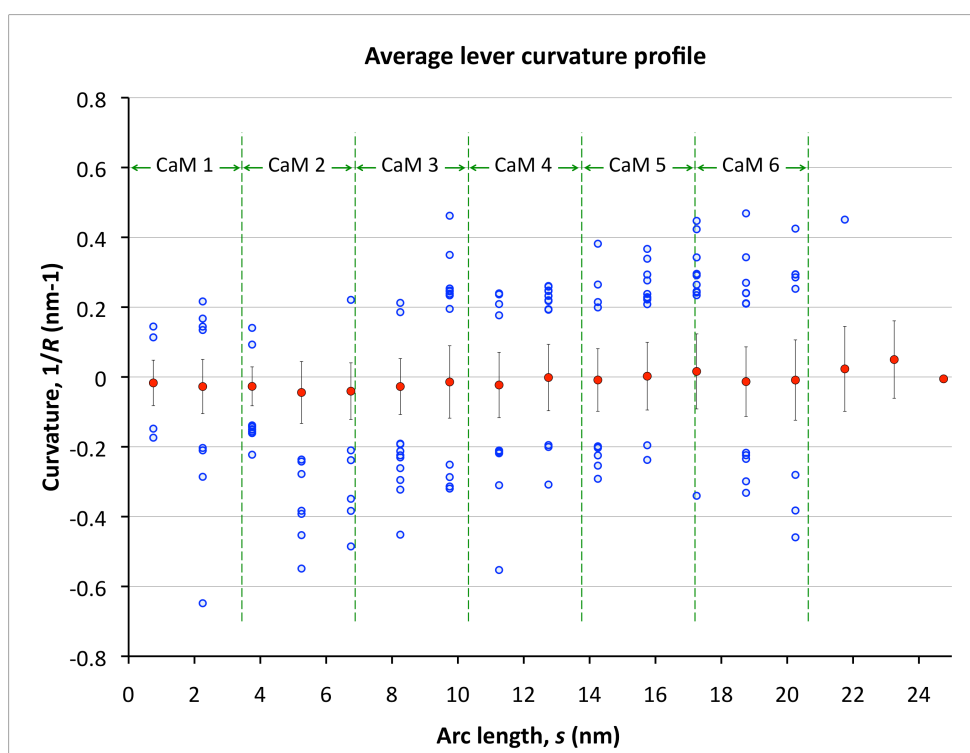
Figure 4.20: Lever domain curvature plots for 183 View-1 class averages.

Plots are derived from fitted spline curve coordinate data from 183 class averages of View-1 molecules. Spline coordinate sampling was at every 12th point. Green dashed lines mark the expected positions of the CaM junctions, every 3.43 nm. (a) curvature values from all data plotted with arc length along the lever. (b) binned curvature plot of the data in (a), with the arc length bin size at 1.5 nm. The plot shows the average curvature for each bin (red points) with ± 1 S.D error bars. Points shown in blue are any outlying points that lie > 2 S.D from the mean.

4.5.2.3 Anisotropies in pronounced curvature

Whilst the general spread of points about a minimum suggests thermally-induced bending, there are also features of the curvature plot that suggest possible anisotropies in the response of the lever.

In the CaM-2 region, high $1/R$ values (≥ 0.2 nm^{-1}) are always negative. Recalling that negative curvature results from the sign of the θ_m angles, this indicates that when significant bending in the CaM-2 region occurs, it is in the direction that would curve left of vertical as marked in the class averages, right of vertical bending being somehow inhibited. Tracing some of the high-curvature CaM-2 points to their source images confirms this, Figure 4.21. Inspection of these averages demonstrates that the high curvature values appearing in the CaM-2 region derive from shapes where the pronounced curve is manifested at the CaM-2–CaM-3 junction.
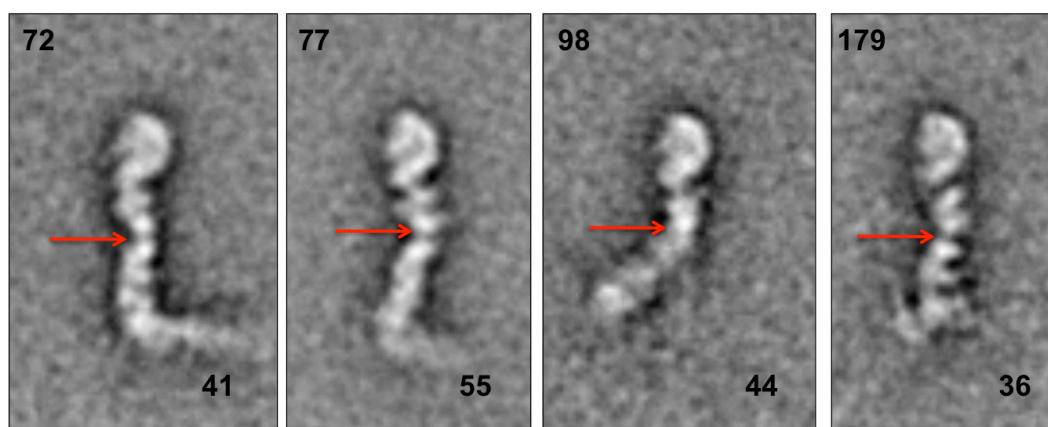
Figure 4.21: Examples of pronounced anisotropic bending at the CaM-2–CaM-3 junction.
Red arrows indicate the CaM-2–CaM-3 junction where there is pronounced bending left. Class numbers are indicated top left and the number of images contributing to each class are bottom right in the images.

CaM-2–CaM-3 is a 25-residue IQ-motif spacing. Aligning the atomic model structure from chapter 3 to best match the CaM profiles of CaMs 1-3 in the example averages of Figure 4.21, suggests that a possible reason for the anisotropy in bending may be simple steric clash of CaM-2 and -3. Figure 4.22 shows how there is more space to allow significant left side curvature before CaM-2 and -3 lobes come close together, while only slight bending to the right side of the lever could result in a clash between CaMs 2 and 3 (in particular Lys 94 and Asp 95 of CaM-3 C-lobe might clash with Lys 148 of CaM 2, or indeed the heavy chain helix).
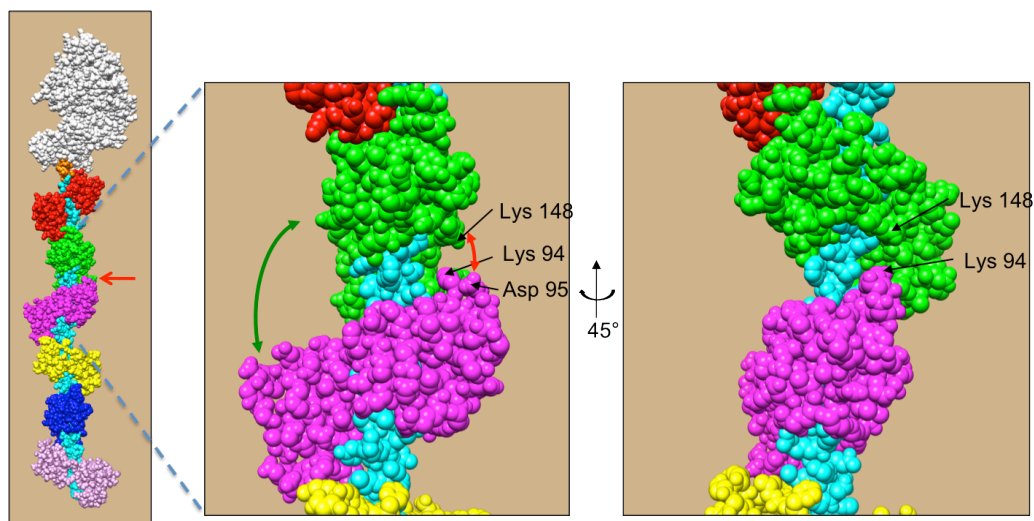
Figure 4.22: Atomic model at the CaM-2–CaM-3 junction.

Left panel is the atomic model of the myosin 5a HMM head developed in chapter 3 (apo motor with bent lever helix at the pliant motor-lever junction (orange)). The middle and right panels show enlargements of the CaM-2 (green) CaM-3 (magenta) region. The heavy chain helix is cyan. The middle panel is at the same orientation as the full head depiction which has CaMs 1-3 aligned as in the averages in Figure 4.21. The right panel is rotated 45° about the lever axis as shown, to look down upon the C-lobe of CaM-3. The red arrow (left, mid panels) indicates how, due to side chains in the C-lobe of CaM-3 being in close proximity to CaM-2 and the heavy chain, bending to this side of vertical may be restricted. On the other side of the lever there is more space to allow lever flexing (green arrow).

A similar anisotropy in high curvature direction can be seen in the CaM-5 region of Figure 4.20(a). Here, pronounced curvature values (>0.2 nm^{-1}) are all positive in value. Tracing some of the high curvature values to specific class averages confirms that pronounced bending in this region tends to be towards the right of vertical, and is manifested at the CaM-5–CaM-6 junction.
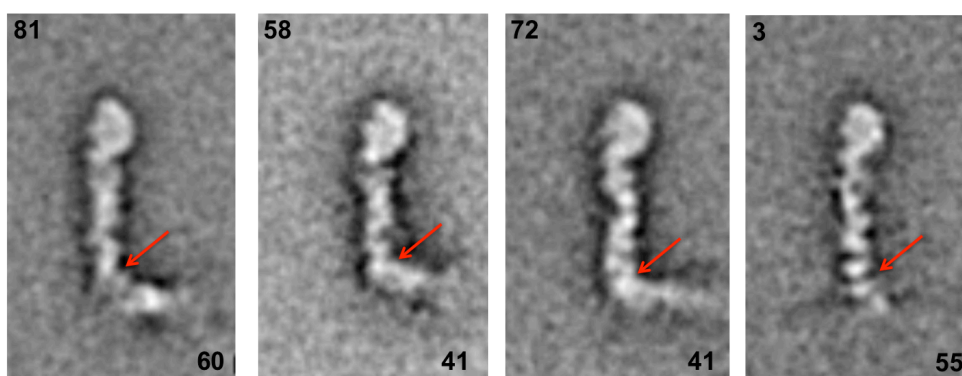
Figure 4.23: Pronounced bending at the CaM-5–CaM-6 junction

Example class averages that contribute high curvature values in the CaM 5-6 region are shown. Red arrows mark the CaM-5–CaM-6 junction where there is pronounced curvature to the right of vertical.

Unlike CaMs 2-3, CaMs 5-6 bind at a 23-residue IQ spacing. However, inspection of the atomic model in this region, Figure 4.24, shows that, structurally, the two CaMs are orientated similarly to in the CaM 2-3 region, but mirrored about the lever axis. Thus there

164

appears more room for bending on the right side of the lever and less on the left, the reverse of the CaM 2-3 junction.
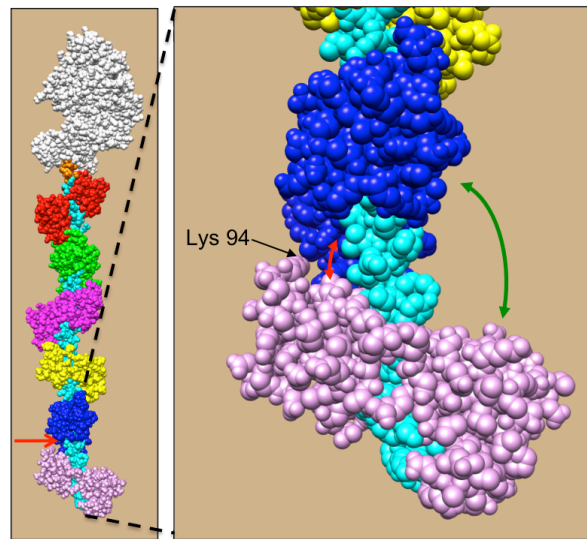
Figure 4.24: Atomic model of the CaM-5–CaM-6 junction.

Right panel shows an enlargement of the CaM-5 (blue) –CaM-6 (pink) region. The red arrows (left and right panels) highlight the crowded residues of the adjacent CaMs (and heavy chain) on the left side of the lever, while the green arrow (right panel) shows how there is more space for the CaMs to move towards each other on the right side of the lever.

A final general observation of the curvature plot of Figure 4.20 is that curvature values become increasingly more spread in the C-terminal half of the lever. That is, higher magnitude curvatures become more frequent and generally increase in size from around the CaM 3-4 junction onwards. What structural feature could explain the higher amplitude bending in the C-terminal half of the lever? One possibility may relate to the observation of increasing rotational freedom in CaMs 4-6, described in 3.5.2.8. It is possible that if the CaMs are freer to rotate around the lever axis in the C-terminal half, steric clashing of adjacent CaMs is more often avoided. Of course, the causal link in this conjecture is not proved. It could also be argued that increased tendency for bending in the C-terminal region of the lever causes CaMs to reposition more frequently.

4.5.3   Bending mode analysis of lever shapes

In this section, a bending mode analysis of the lever classification shapes in Figure 4.11 is undertaken. An implicit assumption in the analysis is that the flexible lever region can be treated as an isotropic elastic rod that responds uniformly with length to bending deformations. From the evidence of anisotropic curvature with arc length presented in 4.5.2.2, it could be argued that this assumption is, to some extent, already disproved. However, application of the mode analysis is still worthwhile if one argues that the anisotropies of curvature are outlying rare events (*e.g.* > 2 S.D), as Figure 4.20(b)

suggests, and that, on average, the lever appears to fluctuate evenly about its near-straight low energy conformation. Indeed, application of the bending mode analysis can also be applied as a further test of the assumption that the lever responds as an uniform elastic rod. Results deviating from the expectations of the model would be confirmation that the model is inaccurate. The bending mode analysis is also useful as (if proved valid) it can provide an estimate of the bending stiffness of the lever, which may be compared to existing experimental values.

The lever coordinate data of the 183 class averages (motor view-1) already captured for 4.5.2 were used for the mode analysis. As for the curvature analysis, the coordinate data used was re-sampled from the high resolution spline fits through the hand-marked coordinates, at a 1/12 point sampling frequency. Each re-sampled curve in the dataset was then decomposed into a series of n cosine modes (see 4.2.4) with the amplitude coefficient of each mode a_n computed by Eq. 4.32.

The relationship between the mode amplitudes and the flexural rigidity of the lever was obtained by studying the distributions of each mode number across all curves in the dataset. This information for the 183 lever shape averages is illustrated in Figure 4.25.

The plot of amplitudes for modes 1-8 in Figure 4.25(a) shows that the amplitudes of each mode fluctuate roughly symmetrically around a mean (indicated by the red line in each panel), with the spread of points decreasing rapidly with n to a minimum threshold. The width of each cloud of points (*i.e* the standard deviation σ) is related to κ, with the expected relationship that $\sigma^2 \propto 1/n^2$ for an isotropic rod (Eq. 4.20). The reduction in the width of the clouds with mode seen in Figure 4.25(a) therefore indicates that the average lever shapes are likely produced by thermal fluctuations about an equilibrium. If the shapes were purely random (or dominated by measurement error), the amplitude variances would be constant, or rise across the first ~10 modes (see discussion below). Modes 1 and 2 exhibit non-zero mean amplitudes (mean $a_1 = 0.487$ nm$^{1/2}$, mean $a_2 = 0.168$ nm$^{1/2}$) (red lines in Figure 4.25(a)), which is due to the intrinsic curvature of the lever's relaxed shape, consistent with the observations of the classified lever shapes and the global average described in 4.5.1.

(a)

(b)
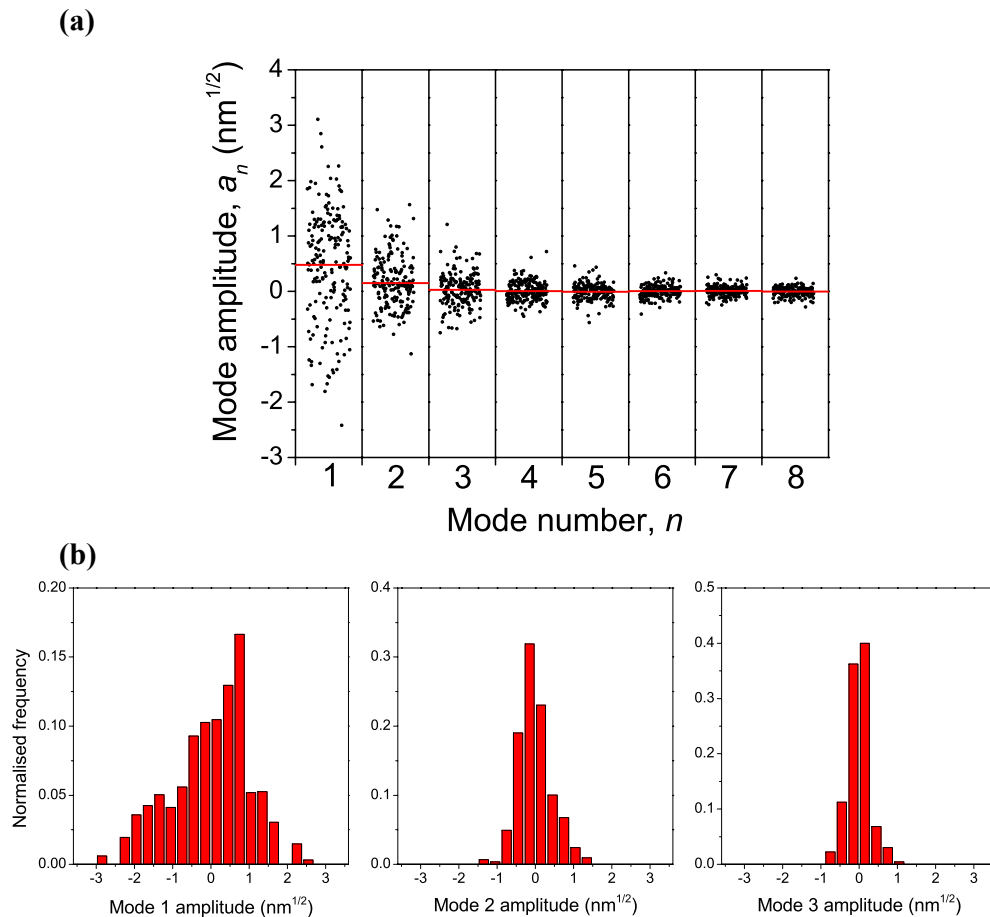
Figure 4.25: Bending mode amplitudes of the myosin 5a lever domain.

Amplitude values plotted are for a cosine series expansion of 183 view-1 lever shapes, where spline curve fits to the lever coordinates were sampled at every 12th coordinate. (a) the amplitudes for each mode number plotted in separate panels (points spread along the horizontal axis in each panel for ease of viewing). The mean amplitude for each mode is shown by a red line. Non-zero mean amplitudes seen for modes 1 (0.487 nm$^{1/2}$) and 2 (0.168 nm$^{1/2}$) reflect an intrinsic bend in the relaxed lever shape. The width of each cloud of points is related to the lever's persistence length. (b) histogram plots of the amplitude distributions (bin size 0.3 nm$^{1/2}$) of modes 1-3 after subtraction of their mean amplitudes. Amplitude values have been weighted according class sizes and the frequencies are normalised (summing to 1). Note the change in the frequency scale as the distributions get narrower and taller with increasing mode number. Gaussian-shaped distributions (centred about 0) would be expected for the isotropic rod model.

To calculate the absolute variance values for each mode one must first make corrections for the rod's relaxed shape and for any fluctuations in the lever curves caused by sources of experimental noise *e.g.* imprecision in manually marking the lever's position with a mouse pointer. The latter will always add to the measured variance as a systematic error (Gittes et al., 1993).

Subtraction of the rod's relaxed shape was accomplished by computing the mean amplitude of a mode, a_n^0, and subtracting this from each amplitude in that mode. Figure 4.25(b) shows the amplitude distributions obtained for modes 1-3, after this correction is

167

made. The normalised histograms show that the amplitude fluctuations have approximately Gaussian profiles (as expected), although mode 1 is slightly skewed towards positive (possibly reflecting non-uniform bending in the lever). An additional correction made here relates to the fact that each lever curve is itself an average of individual molecule images, and so each class average carries a different weight, according to the number of images contributing to that class. The data forming the histograms in Figure 4.25(b) was weighted by class size accordingly.

Subtraction of noise contributions that add to the observed variances has been modelled elegantly by Gittes et al. (1993). By a geometrical argument they show that for the cosine expansion, the additional noise fluctuation contribution to the variance is given by

$$(\sigma_n^2)_{\text{noise}} = \frac{4}{L}\varepsilon^2\left[1 + (N-2)\sin^2\left(\frac{n\pi}{2(N-1)}\right)\right],$$ (Eq. 4.34)

where ε^2 is the mean-squared error in locating the position of the curve, N is the number of position points sampled along the curve, n the mode number and L the curve length. In other words, the total measured variance is the sum of thermally-induced variance and a noise variance:

$$(\sigma_n^2)_{\text{measured}} = (\sigma_n^2)_{\text{thermal}} + (\sigma_n^2)_{\text{noise}} = \left(\frac{L}{n\pi}\right)^2\frac{1}{L_p} + \frac{4}{L}\varepsilon^2\left[1 + (N-2)\sin^2\left(\frac{n\pi}{2(N-1)}\right)\right],$$

(Eq. 4.35)

where Eq. 4.21 has been used to formulate an expression for $(\sigma_n^2)_{\text{thermal}}$, the variance component due purely to thermal bending. The noise contribution alone, given by Eq. 4.34, is expected to rise monotonically with n over the first ~10 modes, (see Figure 4.26). By plotting $(\sigma_n^2)_{\text{measured}}$ values obtained for each mode, an estimate of the error parameter ε can be obtained from a fit to the higher order modes (where the error function dominates). In turn, the noise term can then be subtracted from each mode's measured variance, such that each mode can provide an unbiased estimate of the lever's persistence length:

$$L_p = \frac{\kappa}{k_bT} = \left(\frac{L}{n\pi}\right)^2\frac{1}{\left[(\sigma_n^2)_{\text{measured}} - (\sigma_n^2)_{\text{noise}}\right]}.$$ (4.36)

Figure 4.26 shows plots (one with linear axes, one with log-log axes) of the variance for each mode (uncorrected for noise), measured across the 183 lever shape averages in the dataset (*NB*: the variances *are* corrected for relaxed shape contributions and the data has been weighted according to class size).

(a)

(b)
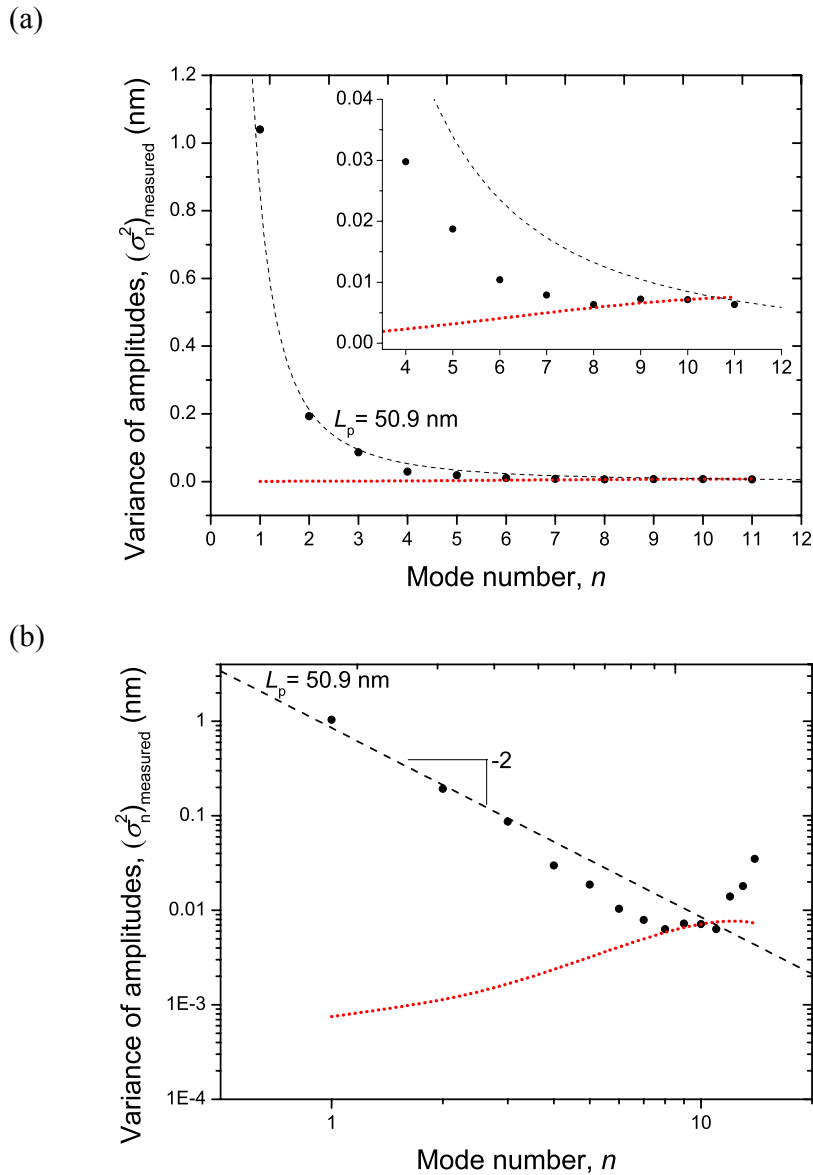
Figure 4.26: Amplitude variances against mode for 183 average lever shapes.

(a) linear scale axes, (b) Log-Log scale axes. The points plotted are the measured variance values derived from the spline fit curves of the lever shapes (without correction for noise errors). The red dotted line is a plot of the noise function with RMS error parameter $\sqrt{\varepsilon^2} = 0.0572$ nm as an example. The dashed line is the theoretical curve expected for thermal fluctuations alone, obeying the relationship $(\sigma_n^2)_{\text{thermal}} \propto n^{-2}$ with $L_p = 50.9$ nm. The inset graph in (a) is an enlargement of the data in the modes 4-11 range. The measured variance values in this region are consistently less than the sum expected from thermal fluctuations and noise terms.

The plots confirm how the amplitude variances decreases rapidly with mode number, resembling a power law relationship between variance and n. The noise function Eq. 4.34 (rising dotted red line) is plotted for an RMS error $\sqrt{\varepsilon^2} = 0.0572$ nm (equivalent to 0.11 pixels in the image), to show an example of its form. In fact, the variance values in at least the first 7 modes seem unaffected by sources of noise following this form, the noise contribution is negligible until the higher modes ($n>7$). An upper limit of $\varepsilon = 0.0572$ nm was therefore used to calculate noise-subtracted variance values. For modes 1-3, the

measured variances follow the n^{-2} relationship expected for thermal fluctuation theory well (dashed line), while the modes 4-7 variances are underrepresented (falling below the value expected for n^{-2}). This is most apparent from the Log-Log scale plot in Figure 4.26(b), where the expected n^{-2} relationship forms a straight line with gradient equal to the exponent, -2.

4.5.3.1 Bending stiffness results

Using Eqs. 4.21, 4.10 and 4.36, independent calculations of κ, $k_{cantilever}$ and L_p were made using the variances of modes 1, 2 and 3 in turn (*i.e.* using only the modes that follow the $(\sigma_n)^2 \propto n^{-2}$ expectation in Figure 4.26(b)). The small correction for experimental noise in the variance (given by Eq. 4.34) was calculated and incorporated in each case. The stiffness values obtained for each mode and the weighted means for all three modes (individual values weighted according to $1/(S.E)^2$) are given in Table 4.4.

Table 4.4: Flexural rigidity, cantilever stiffness and persistence length values derived from mode analysis of myosin 5a lever region

Mode, n	Amplitude variance (nm)	κ^a (pN nm^2)	$k_{cantilever}^b$ (pN nm^{-1})	L_p (nm)
1	1.04	168 ± 125	0.037 ± 0.028	42 ± 31
2	0.192	227 ± 142	0.050 ± 0.031	56 ± 35
3	0.0851	228 ± 130	0.050 ± 0.029	56 ± 32
Weighted averages:		**206 ± 75^c**	**0.045 ± 0.017^c**	**50.9 ± 18.6^c**

Uncertainties for mode parameters are derived from standard error formulas following the method in (Gittes et al., 1993).
[a] Values calculated assuming 293.15 K final grid temperature (estimated).
[b] Stiffness calculated assuming mechanical lever length of 23.9 nm.
[c] Weighted mean ± standard error.

The stiffness parameters obtained in Table 4.4 are consistent across the first three modes, although the relative uncertainties (calculated by standard error-propagation formula, following the method in Gittes et al. (1993)) are high (>50%). Taking the weighted average of each stiffness parameter reduces the relative uncertainty, but only to 37%.

The weighted mean persistence length of 50.9 ± 18.6 nm obtained for the myosin 5a lever is a surprising result. It is notably lower than the 225 nm value implied by the 0.2 ± 0.01 pNnm^{-1} cross-bridge stiffness measured by Veigel et al. (2005; 2002) in the optical trap. In other words, the free HMM heads as analysed here, appear considerably more flexible, at least under negative stain EM conditions, than the previous studies imply actin-bound myosin 5a heads are. Possible reasons for this discrepancy, and the assumptions and experimental factors that may have lead to an underestimated L_p value by the mode variance analysis are discussed further in 4.6 below.

In the following section, the lever contour data from the EM image averages is first analysed by an alternative method, as a means of cross checking the mode variance stiffness result.

4.5.3.2 L_p estimate by tangential angle correlation method.

To validate the L_p value obtained by mode variance analysis, the lever shape data was also analysed by 'tangential angle decorrelation method'. This approach, initially used by Frontali et al. (1979) to measure DNA flexibility, uses Eq. 4.23, which relates the cosine of the tangential angle at a given arc length to L_p. Taking the natural logarithm of Eq 4.23 yields the relation:

$$\ln\langle\cos[\theta(s) - \theta(0)]\rangle = -\frac{s}{2L_p}. \qquad (4.37)$$

Thus, a plot of the logarithm of cosine of tangential angles at arc lengths along the curve, averaged over all conformations, should yield a straight line with gradient $-1/(2L_p)$. A requirement for the use of this statistically-based method is that the measured angle distributions be Gaussian (Frontali et al., 1979). Figure 4.27 shows a plot of $\ln\langle\cos[\theta(s) - \theta(0)]\rangle$ derived from the lever coordinate data as used in the mode analysis. An estimate of ~29-62 nm for L_p is obtained from linear regression of the points, the precise value varying depending on what range of arc length data is used for the fit.

Linear fits over s ranges:		
s data range	Slope	L_p
0-8 nm	-0.0081 nm^{-1}	62 nm
8-20.5 nm	-0.0173 nm^{-1}	29 nm
19.5-24 nm	-0.0538 nm^{-1}	9 nm
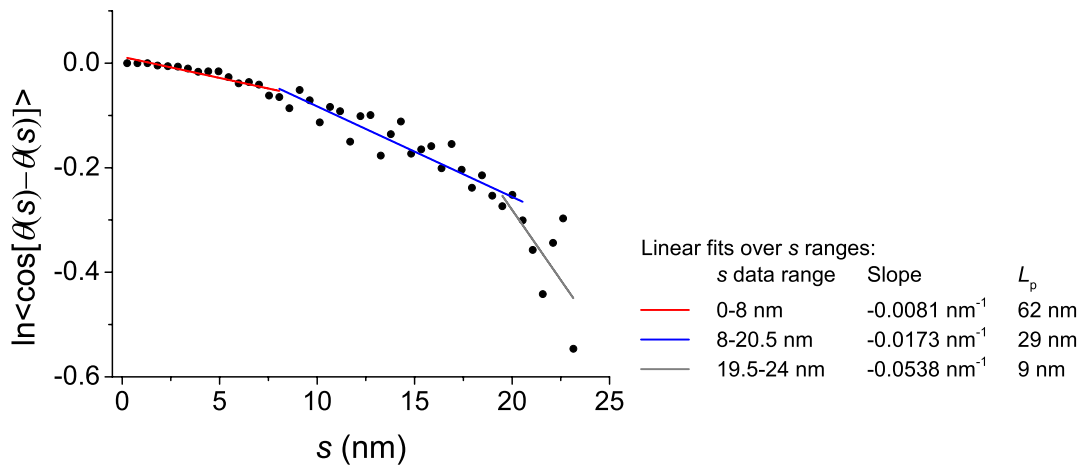
Figure 4.27: Tangential angle correlation analysis of lever curve data

ln<cos[$\theta(s)$ - $\theta(s)$]> values for data from the 183 class averages of lever curves are plotted. A linear trend would be expected if the angular data for $\theta(s)$ was Gaussian. Values appear to deviate from a single linear trend. Three linear fits are shown for different data ranges in s. From the gradients of the lines, three different L_p are estimated.

The plot in Figure 4.27 deviates from the linear relationship expected for a Gaussian distribution in $\theta(s)$, an indication that the statistical polymer chain model does not fit the lever shapes particularly well. The negative gradient of the logarithm data tends to

increase with s (implying greater flexibility/shorter L_p with increasing s) so that linear fits to different ranges of data points yield different L_p estimates (Figure 4.27). The longest L_p estimate of 62 nm (considering only the $s = 0$-8 nm data range) is close to the 51 nm estimate from mode analysis, but the majority of the data in the middle range of the plot (s = 8-20.5 nm) falls below the mode estimate. However, an important caveat to this method is that the decorrelation calculation as implemented here does not account for any curvature in the relaxed lever shape, which the classification averages indicate myosin 5a possesses (4.5.1), and which *was* accounted for in the mode analysis. As recent EM and molecular dynamics studies of tropomyosin have shown (Li, Holmes, et al., 2010; Li, Lehman, et al., 2010), neglecting intrinsic curvature of a protein can lead to significant underestimation of the true L_p value. Hence, the 29-62 nm L_p estimate by tangential angle correlation is likely an underestimate.

4.6 Discussion

To mean persistence length of 50.9 ± 18.6 nm obtained for the myosin 5a lever can be compared to that of other biological macromolecules. It is longer (*i.e.* stiffer) than that of titin ($L_p \sim 13.5$ nm) (Tskhovrebova and Trinick, 2001) or of a typical stable single α-helix ($L_p \sim 15$ nm) (Sivaramakrishnan et al., 2009), but interestingly of similar length (stiffness) to that of DNA ($L_p \sim 50$ nm) (Frontali et al., 1979; Hagerman, 1988). Could the levers of a myosin 5a molecule be replaced by equivalent lengths of DNA and the molecule still be mechanically competent and able to move processively on F-actin?

The lever stiffness results given in Table 4.4 can also be compared to the actomyosin 5a cross-bridge stiffness measured by optical trap by Veigel et al. (2005; 2002). At the time of writing, the trap data from these studies is the only other directly measured stiffness for myosin 5a in the literature. A number of other stiffness values do exist, but are either estimates to fit models (Craig and Linke, 2009; Lan and Sun, 2006; 2005; Skau et al., 2006), or have been inferred indirectly from step size data and an assumption of a particular mechanical model of the molecule (Moore et al., 2004; Vilfan, 2005).

From the mode analysis above, the mean cantilever stiffness of 0.045 ± 0.017 pN nm^{-1} (Table 4.4) was calculated from the mean flexural rigidity ($k_{cantilever} = 3\kappa/L^3$), using a lever length L of 23.9 nm (the preferred mechanical lever length used in this thesis, see Table 1.2). This stiffness can be directly compared to the value 0.20 ± 0.01 pN nm^{-1} for the actomyosin 5a-S1 cross-bridge stiffness, as measured by Veigel et al. (2005; 2002). Clearly, the value obtained by the mode analysis of EM images is less than (~25% of) that measured by Veigel et al. in the optical trap studies. It is also lower than any of the

previous modelling study estimates (which range from $k_{cantilever} = 0.09$ pN nm^{-1} ($L_p = 120$ nm) (Lan and Sun, 2005; Skau et al., 2006) to $k_{cantilever} = 0.33$ pN nm^{-1} ($L_p = 370$ nm) (Vilfan, 2005)). Thus, subject to the analytical assumptions made and potential uncertainties or artefacts arising from experimental methods used in this study (discussed in 4.6.2 below), the new stiffness result obtained here indicates that the free myosin 5a head is significantly more flexible than the trapping studies have thus far indicated for the actin-bound complex.

How does the EM image-acquired myosin 5a stiffness compare to measurements of the myosin 2 cross-bridge stiffness? Because of its importance in the context of muscle contraction, more measurements of the myosin 2 cross-bridge stiffness have been made. The values obtained vary widely depending on the technique and experimental conditions used, with single-molecule optical trap experiments ranging from 0.17 to 1.79 pN nm^{-1} (Lewalle et al., 2008; Molloy et al., 1995) and single muscle fibre experiments ranging from 1.7 all the way up to 5 pN nm^{-1} (Linari et al., 2007; Piazzesi et al., 2002). For a summary of cross-bridge stiffness measurements of myosin 2, see supporting information Table S1 in Kaya and Higuchi (2010). Since myosin 2 has only a 2-IQ lever that is ~1/3 the length of myosin 5a's, its cantilever stiffness would therefore be expected to be ~$3^3 = 27$ times greater than that of myosin 5a (since $k_{cantilever} \propto 1/L^3$). Extrapolating from the myosin 5a lever stiffness obtained from the EM image data, one would therefore estimate that the myosin 2 cantilever stiffness would be roughly 27×0.045 pN nm$^{-1} = 1.22$ pN nm^{-1}. This compares quite favourably with the value of 1.79 pN nm^{-1} obtained in the most recent optical trap study by Lewalle et al. (2008), which was also in close agreement with the single fibre study of Linari et al. (2007). In contrast, if the Veigel et al. myosin 5a cross-bridge stiffness is extrapolated to a myosin 2 stiffness, a value of 27×0.2 pN nm$^{-1} = 5.4$ pN nm^{-1} is obtained, which is higher than any of the current empirical values for myosin 2. Thus, the myosin 5a stiffness result obtained in this study is more consistent with the myosin 2 data than the Veigel et al. optical trap value, which is actually surprisingly high.

4.6.1   Evaluation of lever bending stiffness value obtained

Is the comparatively low myosin 5a lever stiffness obtained by mode the variance analysis justified?

In addressing this question, one should first ask if the $L_p = 51\pm19$ nm result is consistent with the lever shapes marked in the image averages in Figure 4.17? Recalling that L_p is the arc length over which correlations in the tangential angle are nearly lost (the arc length at which the difference in start and end tangential angles average 68.4°) (see 4.2.5), and that,

173

when confined to two dimensions, the decorrelation occurs over twice the distance; $L_p \sim 50$ nm implies that in EM images, the tangential direction made by the start of the lever should be decorrelated with the tangential angle at a distance of $\sim 2L_p \approx 100$ nm. This is approximately four times the length of the lever itself. Inspection of the image examples in Figure 4.17 suggests that this is a reasonable assertion for the set of lever curves captured. For example, the tangential direction at the end of the lever is often at $\sim \pm 45°$ to the starting angle, and this change occurs over only $\sim 1/4$ the distance expected for 2D decorrelation.

Another way to roughly judge the accuracy of the L_p value is to compare the lever shapes with other filamentous objects which, like the myosin 5a lever, have a persistence length similar to (or ~ 2 times) their typical length. One such example is actin, which forms filaments of order μm's and which has $L_p \sim 10\text{-}20$ μm (Brangwynne et al., 2007; Gittes et al., 1993; Isambert et al., 1995; McCullough et al., 2008). Video stills of actin filaments undergoing 2D thermal fluctuations from these studies indeed exhibit shapes with a roughly similar degree of curvature to those seen in the myosin 5a lever averages – a verification that the L_p value obtained is, at least, of the right order.

In summary, approximate, 'by-eye', estimates of L_p and comparison with the value calculated by tangential angle correlation, confirm that the surprisingly low L_p value obtained by mode analysis is of the right order for the set of average lever shapes analysed. The decorrelation calculation yielded a similar maximum L_p, though the result is actually less reliable than the mode analysis owing to the lever shapes deviating from expectations for a statistical polymer chain model and the calculation not taking account of the lever's curved relaxed shape.

4.6.2 Review of modelling assumptions made.

It is possible that the low lever stiffness value obtained by mode analysis results from inaccurate assumptions made in the underlying mechanical model of the lever, in the treatment of the system as a equilibrated rod undergoing 2D thermal fluctuations and in the measurement methods used. Key assumptions made in this study are recapped and appraised below:

    1. The lever behaves as a thin isotropically elastic rod.

        As discussed above, the primary mechanical assumption that the lever is continuously and isotropically flexible is contradicted to some extent by rare examples of sharp anisotropic bending at certain points along the lever seen in 4.5.2.3. The slight deviation of first mode amplitudes from a symmetrical Gaussian distribution in Figure 4.25(b) may also betray an element of anisotropic bending. However, in support of the isotropic model, it could be

argued that the sharply bent examples are exceptional cases (perhaps caused by unusual surface adsorption processes) and that otherwise, the lever bending appears smooth and uniform. The finding that the variances of the first three modes follow the expected n^{-2} relationship supports the isotropic model.

2. *The lever can be treated as 'free' at both ends.*

   This assumption was implicitly made in adopting a pure cosine expansion of the lever curves. Clearly, the assumption is not exact in the case of myosin 5a lever, since the N-terminal end embeds into the globular motor domain, and the C-terminal end is connected to the second head and the tail section at the head-tail junction. The degree to which these attached structures influence the fluctuations of the aligned lever is uncertain. Viscous forces on the globular motor domain may act to dampen fluctuations at the N-terminal end, while fluctuation in the second head and the tail at the C-terminal end may couple to the fluctuations in the first head in a complex manner. From analysis in 3.5.1.1-3.5.1.2, it was established that the motor-lever junction is flexible, allowing independent flexing and axial rotations of the motor in different directions. At the C-terminal end of the lever, the secondary structure connecting the lever helix to the head-tail junction is unknown (see 1.8.2.1.2.1). Lacking contradictory empirical evidence, the assumption made in this thesis is that the connecting polypeptide is unstructured, and thus fluctuations in the second head or tail sections are likely to be only weakly coupled (if at all) to the lever of the primary head.

   To further test the free-ends assumption, alternate mode analyses satisfying different end constraints could be performed for comparison. In addition, mode analysis of free myosin 5a-S1 molecules could be made to remove the potential complicating factor of the second head and tail sections at the C-terminal lever end.

3. *The captured lever shapes represent snapshots of thermally-induced fluctuations in two dimensions.*

   This assumption concerns the process of attachment of the molecules to the carbon film substrate of the EM grid, a process which, for negative stain EM, is not well understood. The assumption underlying the mode analysis in this chapter is that the molecules are able to re-equilibrate in 2D during attachment to the substrate, and that interaction with the adsorbing substrate does not alter the local bending rigidity of the molecules (Frontali et al., 1979). Such a process would generally require a weak interaction strength between molecule

and substrate, in the range of thermal energy (Mücke et al., 2009). The contrasting scenario is that the molecule-substrate interaction is much greater than thermal energy, and that the molecule binds irreversibly to the surface, beginning at a few contact points from where adsorption is quickly propagated. In this process, the molecule is 'trapped' in a frustrated conformation before 2D equilibration can occur. In 'ideal' trapping, the un-equilibrated conformation on the surface can be modelled as a projection of the molecule's 3D conformation onto the plane of the substrate (Frontali et al., 1979; Mücke et al., 2009; Rivetti and Guthold, 1996). Such 'trapping' and 3D-projection has the effect of condensing the molecule shapes and gives rise to a shorter apparent L_p (Mücke et al., 2009), so this process might be an explanation for the low lever L_p value obtained here.

Which scenario best describes the myosin 5a head-carbon film interaction? Evidence in this thesis that heads adopt preferred orientations on the EM grid substrate, 3.5.1.1, suggesting that the surface interaction is sufficiently weak that heads have time to rotate about their long axis before irreversible adsorption occurs. This would suggest that 2D equilibration is also possible. The 2D equilibration model has also been verified in a number of other EM studies; of DNA (stain/rotary shadowing on cytochrome-coated substrate) (Frontali et al., 1979), of titin (rotary shadowing on mica substrate) (Tskhovrebova and Trinick, 2001) and of vimentin intermediate filaments (negative staining on carbon substrate) (Mücke et al., 2004). However, in each case, the filament lengths being measured were significantly longer than the persistence lengths obtained. In the case of the myosin 5a lever the reverse is true, the expected L_p being at least twice the lever's length.

The phenomenon of non-thermal surface forces on the grid should also be noted in relation to this assumption. Evidence of molecule flattening in 3.4.3.1 of this thesis demonstrates that additional forces, *e.g.* surface tension resulting from drying-down, can affect molecule conformations. It is possible that such forces are also able to induce non-equilibrium bending conformations, such as those seen with sharp kinks. The effect of non-thermal forces would generally reduce the apparent L_p, and so may be another contributing factor in the low L_p value obtained for the lever.

In future work, the effect of surface forces on the L_p estimate could be studied by comparison with a similar length similarly-stained control sample of known

L_p (*e.g.* titin or tropomyosin fragments). Imaging on alternative substrates, *e.g.* rotary shadowing on mica or AFM studies on glass might also be informative.

4. *The method of hand-marking and spline-fitting captures the essential lever curves accurately*.

   The inherent uncertainty and interpretative nature of manually tracing the lever curves from the position of the CaMs has been discussed at length in 4.4.2. Given the inhomogeneous appearance of the lever it was decided that manual tracing was superior to any automated curve assignment method. However, this process may still be one of the biggest sources of uncertainties in the process of calculating L_p.

   To mitigate the effects of uneven coordinate spacing and short length-scale roughness, the spline-fitting and coordinate re-sampling process was devised. Nevertheless, this process still requires an interpretative judgement as to best sampling frequency that satisfactorily captures the lever's shape but avoids artefacts.

5. *Image averages can be treated as weighted examples of individual molecule conformations*.

   In principle, the analytical methods in this chapter apply only to conformations of *individual* molecules (*e.g.* a requirement of the ensemble average assumption 4.2.4.2). However, the low signal-to-noise ratio in the individual images necessitates classification and *averaging* before the lever shapes can be interpreted and manually traced. The assumption is thus made that the K-means classification groups molecules with similar lever shapes together, which may then be treated as a single (weighted) object. Clearly, grouping and averaging will hide a degree of conformational variability in the individual molecules. In general, one would expect the averaging to increase the apparent stiffness (and L_p) of the object (since the average shape is expected to be straighter), and so it is unlikely that the image averaging assumption accounts for the low lever stiffness value obtained in this chapter.

   In future work, one could attempt to quantify the effect of averaging on apparent stiffness, by successively expanding an average into increasing numbers of sub-averages and making flexibility measurements from the shapes represented. If a trend between decreasing class size and increasing flexibility is observed it might be possible to extrapolate this effect to single-molecule level.

Having cross-checked the L_p value and considered the potential inaccuracies in modelling and methods, the possibility that the lever *is* more flexible than implied by optical trapping measurements should also be considered! Indeed, the optical trap measurements are themselves subject to potential sources of inaccuracy. In particular (and not often discussed in the literature), the mechanism of attachment of the C-terminal end of the myosin head in the trap experiments is typically by non-specific binding to a nitrocellulose substrate (Lewalle et al., 2008; Veigel et al., 2005; 2002). It is reasonable to assume that the myosin heads could in fact bind the surface at different points along the lever domain (rather than all molecules binding the substrate exactly at the C-terminus of the lever) and thus a distribution of effective head lengths are sampled by the trap measurement. This would have the effect of shortening the effective lever length measured, potentially by as much as one half, which would in turn significantly increase the calculated value of $k_{cantilever}$ ($\propto L^{-3}$), by a factor of $2^3 = 8$. If this argument is applied to the Veigel et al. myosin 5a stiffness result (0.2 pN nm^{-1}), the implied cantilever stiffness, as if measured at the tip of the full length lever, would actually reduce to $0.2 \times 1/8 = 0.025$ pN nm^{-1} – a value similar to the stiffness obtained from the EM image averages here.

A final point in evaluating the myosin 5a L_p value obtained, is to note that the EM measurement performed here was on free myosin 5a HMM molecules, as opposed to the optical trap measurements which were made on the actomyosin 5a-S1 complex. It is conceivable that in binding to F-actin, the flexural rigidity of the lever is significantly increased over the free state, *e.g.* through induced structural rearrangements and associated interactions between light-chains and motor domain (*cf.* phosphorylation and calcium-induced changes in regulatory mechanisms of different myosins). In future work, a way to test this possibility would be to perform a bending mode analysis of the lever shapes of myosin 5a-S1 molecules bound to F-actin.

4.7 Conclusions and future work

The work in this chapter is the first attempt to quantitate bending flexibility in the myosin 5a lever and relate this directly to its structure. To achieve this, a large data set of well-stained HMM molecules was utilised. Image processing and classification techniques were successfully used to align and group the images to the improve signal-to-noise ratio. This revealed a continuum of lever conformations, demonstrating the flexibility of this region. New computer tools and procedures were developed to analyse the flexibility implied by the molecule images.

Development of the tools to trace the lever shapes and to analyse curvature and bending stiffness are in themselves useful outcomes from this chapter. The scripts were written to be as generic as possible, so that they might be easily re-used to analyse flexibility in other rod-shaped molecules, protein domains or filaments. The observation of curve roughening or smoothing artefacts dependent on point density was a useful insight (not widely reported in previous studies), and lead to the development of the extra spline-fitting and coordinate re-sampling procedures.

A number of new results emerged from the analysis of the myosin 5a lever domain:

The length of the lever extending beyond the motor domain was measured from image averages as 20.6 ± 1.6 nm (mean ± S.D). This is close to the value from atomic model measurement in chapter 1 (1.8.2.1.2.1), and confirms that staining and grid-adsorption have not excessively stretched or compressed the lever.

A distribution of straight and curving lever shapes was observed in the averages. The degree of lever curvature varied smoothly across evenly populated classes, indicative of a continuum rather than a discrete set of conformations. Left-curving, straight, right-curving and S-shape levers were seen with decreasing frequency, along with rare sharply kinked cases. The distributions of shapes appeared invariant to the orientation of the motor, confirming the observation of chapter 3 that the lever generally adopts a preferred axial orientation on the carbon of the EM grid, and that a pliant motor-lever junction probably allows the motor to rotate independently between its preferred orientations. The shape distribution is not symmetric about a straight lever conformation. Rather, the lever is more often seen curving left of the vertical in the aligned images. This bias reflects the fact that the relaxed lever shape is intrinsically curved, an observation confirmed by inspection of the global average and by non-zero mean mode amplitudes. In general, the distribution of shapes seen is consistent with a picture of a flexible lever undergoing thermal fluctuations about a gently curved equilibrium conformation.

The curvature profile produced by numerical analysis of the lever shapes provides a map of flexibility along the lever's length. Most curvature values are low and cluster around the zero-curvature axis along the lever's entire length. At a given position, the number of points found as one moves to higher curvatures decreases rapidly, with ρ = 0.2-0.4 nm^{-1} (R = 5.0-2.5 nm) representing unusually high curvature (> 2 S.D. from the mean). These features are consistent with a model of the lever as a continuously flexible rod undergoing thermal fluctuations, with more bent conformations being high energy (low probability) states, and near straight conformations being low energy (high probability) states. This evidence supports the elastic rod model adopted by Vilfan (2005) and Lan and Sun (2005).

179

The distribution of curvature values with position along the lever is generally symmetric in the low bending regime ($\rho < |\pm0.2|$ nm^{-1}), supporting the idea that, for smaller forces, lever bending is isotropic. However, in the high bending regime ($\rho \geq |\pm0.2|$ nm^{-1}) there were two notable exceptions. Anisotropic bending is seen at the CaM2-3 junction, where pronounced bending is always negative (to the left of vertical in the images as presented), and at the CaM5-6 junction where sharp bending is always positive (to the right of vertical). In both cases, a plausible structural basis for the anisotropy is the steric clash of closely-spaced lobes of adjacent CaMs (or of CaM and heavy chain), inhibiting bending on one side of the lever. On the opposite side of the lever there is more space for the CaM lobes to come together and permit high bending. The CaM 2 and 3 IQ binding sites are separated by 25 amino acids, while CaMs 5 and 6 are spaced by 23, so there is no obvious relationship between the IQ spacing and flexibility. Pronounced bending is not limited to these CaM junctions, indeed high bending values are seen across most of the lever's length (with the exception of the CaM 1-2 region which appears stiffer), but the two locations identified are the only ones where high bending is apparently inhibited in one direction. It is possible that similar high bending anisotropies exist in the directions in and out of the plane of the grid surface, but these directions are not probed by the EM experiment, as the lever lies generally in one orientation.

The picture of the lever that emerges from this initial curvature mapping is one in which the lever responds to forces below a certain threshold as an elastic and isotropic rod, making small bends uniformly along its length, in all directions. Above the threshold, further bending at particular regions becomes inhibited in certain directions by steric clashes of the light chains at certain CaM junctions. The inhibited directions are dictated by the orientations of the bound CaMs, which are in turn dictated by the IQ motif spacing in the heavy chain. The evidence here does not support the three-segment lever model with flexible hinges at the 25-residue IQ-spacings proposed by Terrak et al. (2005) and modelled by Craig and Linke (2009). Rather, a hybrid model is suggested, where bending is isotropic at low forces (as in the Vilfan model), but is blocked at certain regions in certain directions at a particular limit. Two such inhibited points were identified here. Further study of the 3D model of the head could identify if there are other potential regions of directionally limited bending, *e.g.* where neighbouring CaMs have the same relative orientations as those at the CaM 2-3 and 5-6 junctions.

These curvature results ideally need confirming by repeat analysis of separate data *e.g.* of S1 or more HMM molecules. An obvious starting point would be to analyse the remaining images from the HMM data used in this chapter *i.e.* the molecules with differently-oriented motors. Are the anisotropies repeated?

Despite the indications of anisotropies in lever flexibility, the bending mode analysis was never-the-less pursued as an additional test of the isotropic model, and to see if it was possible to estimate an overall L_p value. In fact, the amplitude variances of the first three modes of the lever shapes followed the isotropic rod relationship (n^{-2}) quite well, while higher modes were underrepresented. Interestingly, some published mode variances for actin filaments and microtubules follow a very similar trend to that seen here (Brangwynne et al., 2007). If the mode variance results are interpreted at face value, then a surprisingly short persistence length of 51±19 nm is calculated for the lever. This is similar to that of DNA, and would suggest the lever is considerably more flexible than optical trap measurements or literature estimates. Could a myosin 5a molecule still function if the levers were replaced by DNA strands of equivalent length? By-eye, the L_p value obtained is roughly consistent with the EM image averages, and cross-checking the value with another method produces a result of similar order. There are many possibilities that modelling and analytical assumptions produce an artificially low L_p and these have been discussed at length above. Perhaps of most importance is the interpretive nature of marking out the lever curves, and the simplifying assumption that the molecules are equilibrated in two-dimensions before grid adsorption. Repeat measurements on separate data and on a control molecule would naturally be useful here.

Chapter 5: Study of Myosin 5a Lever Mutants

5.1   Introduction and aims

This chapter describes the design, production and initial characterisation of a set of myosin 5a mutants with altered lever domains. The mutants were designed to probe the effect of altering the native IQ motif spacing on lever domain rigidity and the ability of the molecules to move processively along F-actin.

Each lever helix in the two-headed myosin 5a molecule bears six successive light chain-binding IQ motifs. The 11-residue 'cores' of these motifs (with consensus sequence IQxxxRGxxxR) are spaced at intervals of exactly 23-25-23-25-23 amino acids (a/a) along the lever helix. This alternating pattern is unique to myosin 5 and the plant homologue myosin 11, where the pattern is conserved in all myosin 5a, 5b and myosin 11 sequences. The conservation in plants and animals suggests a strong evolutionary advantage for the feature, perhaps providing the molecule with particular structural and/or mechanical characteristics important for function, *e.g.* allowing the molecule to efficiently make straight steps along an actin filament.

In the work presented here, a series of heavy meromyosin (HMM)-like myosin 5a constructs, with altered IQ-core spacings were constructed. The aim was to change the native 23-25 alternating pattern and create lever mutants with 23-23-23-23-23 (All-23) and 25-25-25-25-25 (All-25) spacing patterns. These constructs were conceived to address the following questions:

1. Is the 23-25 alternating IQ motif pattern necessary for proper function of myosin 5a, in particular its ability to move processively along F-actin?
2. How does myosin 5a lever rigidity derive from its structure?
3. Does the 23-spaced structure confer greater lever rigidity than the 25-spaced structure?

Question 3. refers to an idea forwarded in crystal structure studies, that the 23 a/a IQ-spacings allow neighbouring LC-LC interactions that may stiffen the lever, while the 25 IQ-spacings are such that neighbouring LCs do not interact and thus may be more flexible regions (Terrak et al., 2005). Certainly, aside from differing LC-LC interaction possibilities, an inspection of the orientations of neighbouring LCs in the 23 and 25-spaced structures (see Figure 1.18 in chapter 1 of this thesis), suggests that the directional bending properties of these different lever 'subunits' will vary according to the constraints of differing steric clashes for the two types of structure. Assuming an alternating rigid-soft

183

lever as suggested by Terrak et al., an All-23 myosin 5a lever (also comparable to the myosin 7a 5-IQ lever structure) might therefore be expected to be more rigid than native, while an All-25 lever might be more flexible. This is a testable hypothesis for the study in this chapter. If lever mechanics are modulated by specific IQ spacings, one might expect to see effects not only in a lever mutant's appearance, but also in its ATPase and motile properties in the presence of F-actin. In the study described here, initial data characterising these structural and functional properties is presented for each lever mutant.

The new All-23 and All-25 HMM constructs described in this chapter, build on some earlier work of Sakamoto et al. (2003), who made a lever mutant ('HMM-2Ala-6IQ') with one IQ-spacing changed from 23 to 25 a/a, by insertion of two alanine residues between IQ-cores 3 and 4. Subsequent EM studies of the '2Ala' mutant showed a molecule in which the mutated lever was unbroken and still able to bind CaMs along its length, and which appeared to bind to F-actin with a similar head-head spacing as wild-type (Oke et al., 2010). However, measurement of various mechanical and biochemical properties of the construct suggested that the lever had, to some extent, been mechanically weakened by the double a/a insertion halfway along its length. In particular, the HMM-2Ala-6IQ's working stroke, as measured by optical trapping, was found to be 10.9 ± 0.6 nm, less than ~1/2 the wild-type control's size and equivalent to a lever with only three CaM-IQ repeats, Figure 5.1 (Sakamoto et al., 2003). Stopped-flow kinetic analysis using the ATP analogue deac-aminoATP also showed that HMM-2Ala-6IQ's ADP dissociation rate from its lead head increased slightly over wild-type (to 0.04 s^{-1} from 0.015 s^{-1}) (Oke et al., 2010). This is also suggestive of increased flexibility (*i.e.* reduced capacity to sustain intramolecular strain) in the mutated lever.
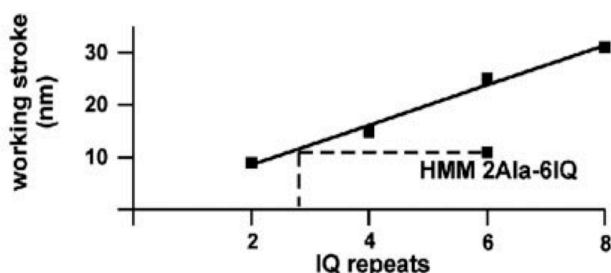
Figure 5.1: Working stroke measurements of HMM-2Ala-6IQ and lever length mutants.

Figure is taken from Sakamoto (2003). The graph shows a plot of the measured working strokes of a series of lever length mutants with 2, 4, 6 and 8 IQ motifs, along with the HMM-2Ala-6IQ construct. The working strokes were measured using an optical trap experiment. The stroke size measured for HMM-2Ala-6IQ (10.9 ± 0.6 nm) is less than half that of the 6-IQ control, and is equivalent to a theoretical construct with 3 CaM-IQ repeats.

Whilst there have been numerous previous studies varying the number of IQ motifs forming the myosin 5a lever (*i.e.* altering lever length) (Lewis et al., 2012; Moore et al.,

2004; Oke et al., 2010; Purcell et al., 2002; Sakamoto et al., 2005; 2003), apart from the Sakamoto et al. HMM-2Ala-6IQ precedent, very few have attempted to alter the actual substructure of the lever (one exception being Baboolal et al. (2009)). There has also been no systematic study of the importance of the IQ motif spacing in myosin 5a, or indeed in other myosin types.

After careful consideration of how best to perform the lever mutations while preserving CaM binding, All-23 and All-25 HMM-like constructs, plus a wild-type control, were designed, cloned and expressed. Construction and expression of the clones was carried out on visits to Dr James Sellers' Laboratory of Molecular Physiology, National Heart Lung and Blood Institute, National Institutes of Health (LMP-NIH) (Bethesda, USA), and with the subsequent help of Dr Attila Nagy of the group.

The mutant design strategy, cloning, expression and purification methods are described below, followed by initial characterisations results. Directions for future work are discussed in the conclusion.

5.2 Design of myosin 5a-HMM lever mutants

The principal aim in constructing the myosin 5a-HMM lever mutants was to alter the native IQ motif spacing to All-23 or All-25, but to try to maintain full CaM binding to the altered lever helix sequence. To preserve the most important CaM-lever helix interactions at the IQ binding sites, mutations were targeted at the 12 or 14 a/a sequences that 'bridge' between the 11-residue consensus 'cores' (IQxxxRGxxxR), at the heart of the binding sequences. The six IQ-cores plus bridge sequences (together forming an IQ motif) in the mouse myosin 5a sequence (the sequence used in this work) are:

	IQ-core	Bridge sequence	Amino acids to next IQ-core
IQ1:	^{773}**IQ**KTI**RG**WLL**R**	KRYLCMQRAAIT^{795}	23
IQ2:	^{796}**VQ**RYV**RG**YQA**R**	CYAKFLRRTKAATT^{820}	25
IQ3:	^{821}**IQ**KYW**RM**YVV**R**	RRYKIRRAATIV^{843}	23
IQ4:	^{844}**IQ**SYL**RG**YLT**R**	NRYRKILREYKAVI^{868}	25
IQ5:	^{869}**IQ**KRV**RG**WLA**R**	THYKRTMKAIVY^{891}	23
IQ6:	^{892}**LQ**CCF**RR**MMA**K**	RELKKLKIEARSVE...	–

Figure 5.2: The six successive IQ-cores and bridge sequences for mouse myosin 5a.

The IQ-core consensus residues are highlighted in bold. The sequence number of the first and last residues in each sequence is given in superscript. The spacing pattern counts inclusively from the first consensus residue of one IQ-core up to and including the residue prior to the next. NCBI GenBank sequence accession: NM_010864.

After careful consideration, it was decided to use two different strategies for constructing the mutants, in the hope that if one strategy failed to produce a CaM-binding molecule, the alternative might still be fruitful. One strategy used simple double a/a insertion or deletion similar to the HMM-2Ala-6IQ mutant of Sakamoto et al., while the other approach used a strategy of bridge sequence splicing.

5.2.1 'Insert' and 'Delete' style lever mutants

In the first mutation strategy, 23 to 25 or 25 to 23 IQ spacing changes were made by either double residue deletion in a 14-residue IQ-core bridge sequence (converts a 25 a/a spacing to 23) or by double residue insertion in a 12-residue bridge (converts 23 a/a spacing to 25). These constructs were given the names 'All-25-Insert' or 'All-23-Delete', the postfix reflecting the mutation method used. The rationale for the location of the insertions or deletions and the choice of which a/a to insert is detailed below.

5.2.1.1 Locating double residue insertion/deletion site in lever helix

To locate the best position to make the double insertions or deletions with least disruption to potential light chain (LC)-heavy chain (HC) interactions, crystal structures of CaM bound to IQs 1-2 of the mouse myosin 5a lever helix (a 23-spaced structure) (PDB: 2IX7) (Houdusse et al., 2006), and the Mlc1p LC bound to IQs 2-3 of lever helix of myo2p, the myosin 5a homologue in yeast (a 25-spaced structure) (PDB: 1N2D) (Terrak et al., 2005), were examined in detail using *UCSF Chimera* molecular modelling software (v1.4) (Resource for Biocomputing, Visualization, and Informatics, University of California, San Francisco, USA).

The two structures were searched for any potential non-covalent interactions between HC helix and LC atoms, searching residue-by-residue along the HC for any atoms with van der Waals radii coming within an overlap distance of ≥ -0.4 Å of LC atoms (where overlap distance is defined as the sum of the two atomic radii minus the distance between atomic centres). *NB*: hydrogen atoms were explicitly added to the structures before analysing. The number of potential interactions per HC residue were mapped, to determine the typical extent and location of interactions along the lever helix around IQ-cores, including potential hydrogen bonds between HC and LC. Interaction maps for the two crystal structures are shown in Figure 5.3.
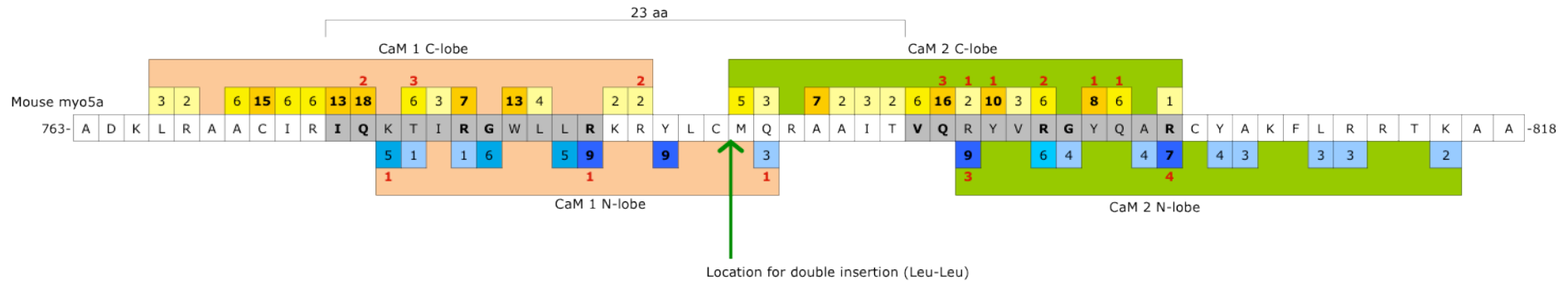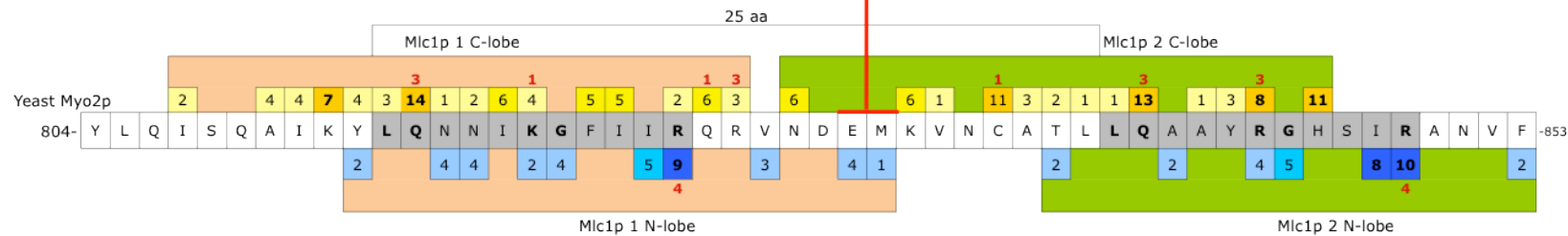
Figure 5.3: Light chain-heavy chain interaction maps for 23 and 25 amino acid-spaced IQ motifs for different myosin 5a structures.

The numbers of potential interactions (atoms with van der Waals radii overlap ≥ −0.4 Å) between heavy chain (HC) and light chain (LC) atoms are shown above and below the HC sequences. Red numbers indicate the number of potential hydrogen bonds within the interaction totals. LC C-lobe interactions appear in yellow above the helix sequence, while LC N-lobe interactions are in blue below the sequence (darker shades indicate more interactions). First LC-binding region with tan background, the second with green background. The maps show that the LC C-lobe typically has more potential interactions with the HC and binding can extend typically 7 residues upstream of the IQ motif core. The N-lobes bind less tightly and can extend downstream of the core. LC binding can thus dovetail across IQ-bridge sequences. The position chosen for double insertion/deletion causing minimal potential disturbance to LC binding (in particular C-lobe recognition) is shown.

The maps show that the LC C-lobe typically has more interaction possibilities with the HC and these typically extend 7 residues upstream of an IQ-core (up to 11 residues upstream in one case for the yeast structure). The N-lobes bind less tightly (fewer interaction possibilities), but interaction possibilities can still extend downstream of the IQ-core. Due to the extent of upstream and downstream interaction possibilities, the binding and/or stabilizing HC interactions of adjacent LCs probably dovetail across the bridge sequences linking successive IQ-cores. Figure 5.4 illustrates this for the 23-residue structure for IQ-cores 1-2 in the mouse crystal structure (PDB 2IX7). The network of potential LC-HC interactions in the IQ-bridge region is suprising.
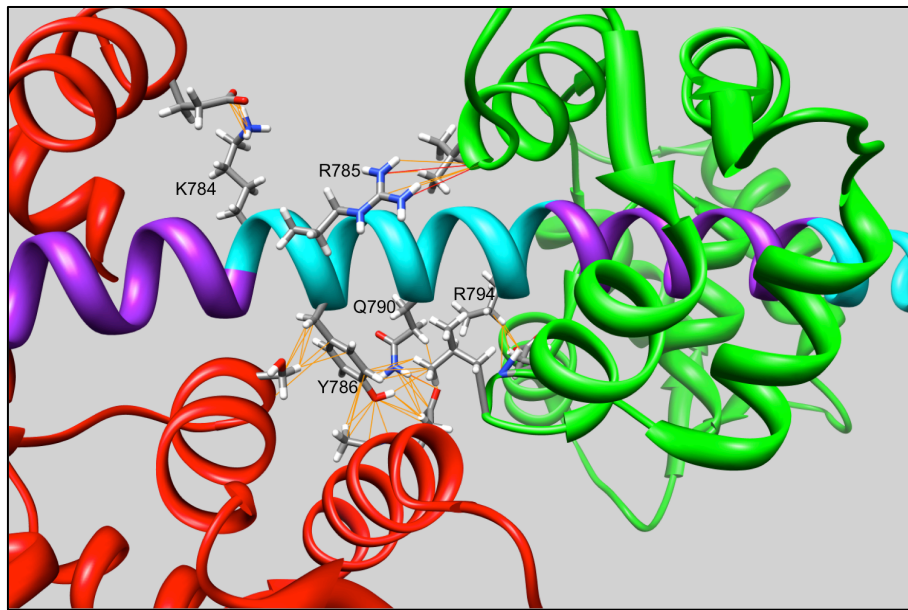
Figure 5.4: Light chain interactions with heavy chain IQ-bridge residues.

Region of the crystal structure of apo-CaMs binding to IQ motifs 1-2 of mouse myosin 5a is shown. HC helix is cyan with IQ-core residues purple. CaM-1 is red, CaM-2 green. Potential CaM-HC interactions (atoms with van der Waals radii overlap ≥ -0.4 Å) in the IQ-bridge section are shown in orange, hydrogen bonds in red. Side chains for any residues that participate in the interactions are shown in stick representation. Residue labels are for the HC helix residues involved in the potential interactions.

Any insertions or deletions made in the bridge sequence will therefore likely disrupt at least some interactions. Since the C-lobe grips the lever helix most tightly (makes more interactions), it was decided to prioritise protection of the key binding interactions for this lobe of the LC. To preserve the 7-residue sequences upstream of the IQ-cores, while also presenting minimal disruption to N-lobe interactions downstream of IQ-cores, the optimal location for a double residue insertion or deletion was determined from the interaction map to be 6 and 7 residues downstream of the last consensus R in the IQ-core. That is:

$$|\textbf{IQ}\text{xxx}\textbf{RG}\text{xxx}\textbf{R}|\text{xxxxx}\Delta\Delta\text{xxxxx}[\text{xx}]|\textbf{IQ}\text{xxx}\textbf{RG}\text{xxx}\textbf{R}|,$$

where $\mathbf{\Delta\Delta}$ denotes the position for the double insertions or deletions ([xx] are the extra residues that occur in the 25-spaced motifs).

5.2.1.2 Choice of amino acids for lever helix insertion

Sequence alignment of myosin 5a lever domains for different organisms was performed to determine what amino acids to choose for the double insertions into the lever helix (to convert 23 to 25 IQ motif spacings). Figure 5.5 shows alignments of the 23 and 25 a/a-spaced IQ motifs for selected myosin 5a sequences.
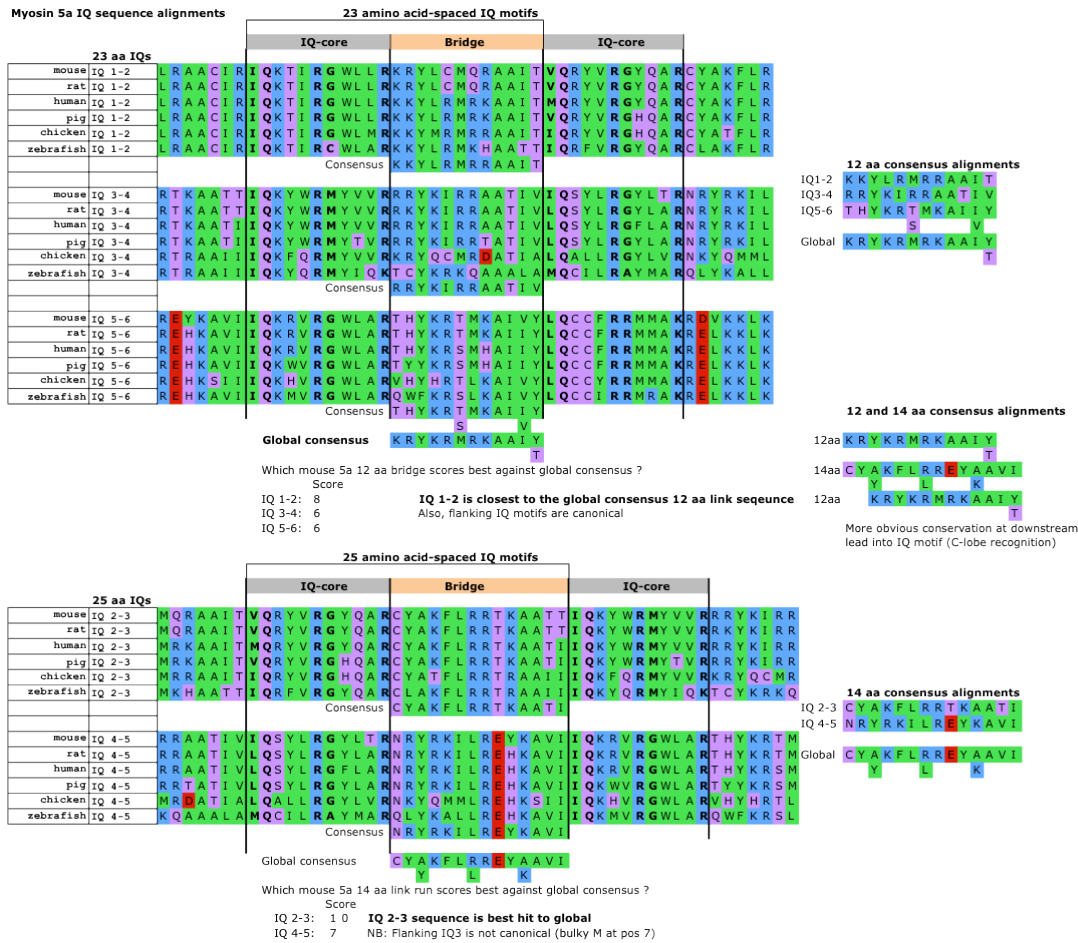
Figure 5.5: Selected myosin 5a sequence alignments of lever domain IQ motif sections.
Green = hydrophobic side-chain, blue = positively charged, red = negatively charged, pink = polar uncharged.

From this, leucine-leucine (LL) or leucine-arginine (LR) were suggested as possible insertions, as these residues occur most frequently in the equivalent insert positions in the two native 25 a/a-spaced sequences. Heptad net analyses of the lengthened 12 a/a bridge sequences with either LL or LR insertions were made, to check that the insertions did not result in excess clustering of neighbouring charged residues, which could be destabilizing to proper α-helix formation. An LL insertion was determined as preferable to LR, due to the arginine side chain being more positively charged and potentially destabilizing to the lever helix. Reservations concerning the introduction of leucine pairs was the unknown

190

behaviour of the hydrophobic and rather bulky side chains in potentially solvent-exposed sections of α-helix, and also that, with the exception of IQ 4-5 in zebrafish sequence, LL does not naturally occur at the equivalent insert positions in any of the other myosin 5a sequences checked in Figure 5.5. Never the less, the LL insertion strategy was pursued, with the knowledge that the parallel bridge splicing approach would be pursued as an alternative.

5.2.1.3   Sequences for All-23-Delete and All-25-Insert mutants

To make the 'All-23-Delete' mutant, residues LR (812-813) and IL (860-861) were deleted from the 14-a/a bridge sequences of the two 25-spaced IQ motifs.
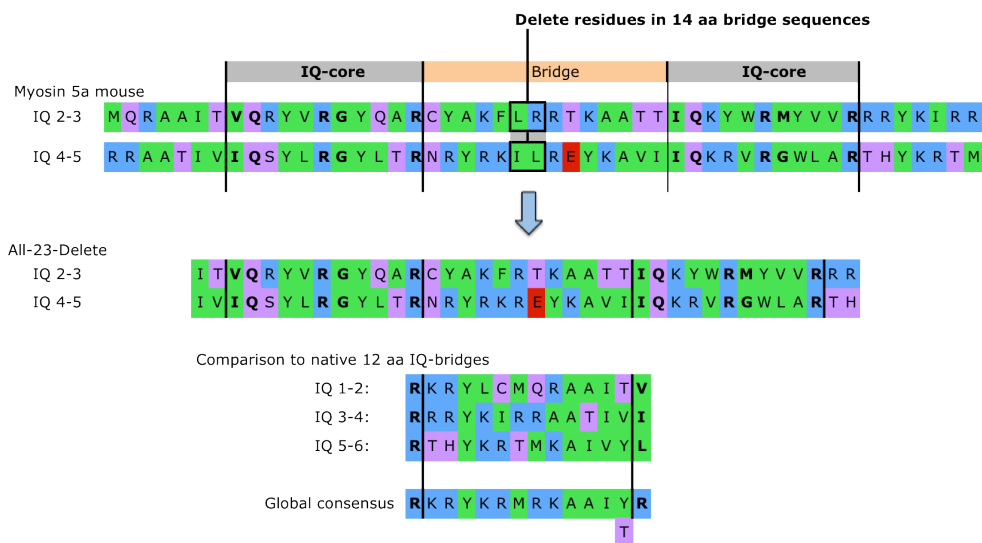
Figure 5.6: Mutation scheme to make All-23-Delete lever mutant.

To make the 'All-25-Insert' mutant, LL was inserted between the residues CM (788-789), IR (836-837) and RT (884-885) in the three 12 a/a bridge sequences of the 23-spaced IQ motifs.
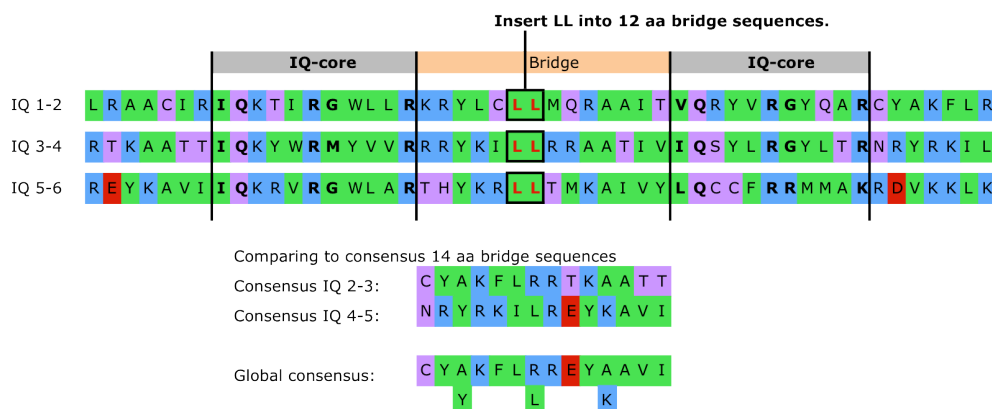
Figure 5.7: Mutation scheme to make All-25-Insert lever mutant.

5.2.2   'Splice' style lever mutants

When analysing the potential LC-HC interactions (Figure 5.3) it became apparent that residues extending upstream and downstream of each IQ core might also be important in

binding or at least stabilizing bound CaMs through non-covalent interactions. The network of possible CaM-helix interactions extending into the IQ-bridge sequences was surprising, see Figure 5.4. With this in mind, an IQ-bridge 'splicing' strategy was devised as an alternative to the 'insert/delete' style mutants, in case these failed to express, fold correctly or bind CaMs.

In the splicing strategy, entire native 12- or 14-residue IQ-bridge sequences were spliced into other positions in the lever, replacing the existing IQ-bridge sequences as appropriate. The logic here was that the native bridge sequences support CaM binding/stabilization (at least in their native positions) and so might support CaM binding when spliced into other IQ-bridge sections.

5.2.2.1 All-23-Bridge construct.

By sequence alignment, see Figure 5.5, the 12 a/a IQ 1-2 bridge sequence (^{784}KRYLCMQRAAIT^{795}) was chosen as most consensus-like and used as the bridge sequence to make the All-23 spliced lever, by replacing bridges between IQ-cores 2-3 and 4-5.
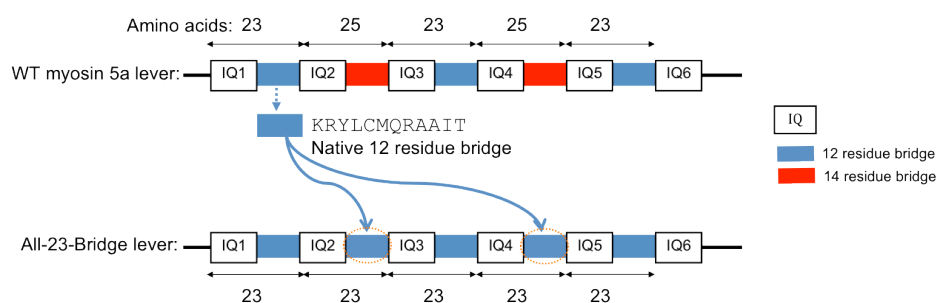
Figure 5.8: Design scheme for the lever of the All-23-Bridge construct.

That is, the nucleotide sequence

^{2390}AAGAGATACCTGTGTATGCAGAGGGCAGCCATCACA^{2425} (^{784}KRYLCMQRAAIT^{795})

was used to replace the two sequences

^{2459}TGCTATGCTAAGTTTCTGCGCAGAACCAAGGCAGCAACCACC^{2500} (^{807}CYAKFLRRTKAATT^{820})

^{2603}AATAGGTATCGCAAGATACTCCGTGAATACAAAGCAGTCATC^{2644} (^{855}NRYRKILREYKAVI^{868}).

The construct created was named All-23-Bridge.

5.2.2.2 All-25-Bridge construct.

Similarly, the 14 a/a IQ 2-3 bridge sequence (^{807}CYAKFLRRTKAAT^{820}) was chosen as the most consensus-like bridge sequence to make the All-25 spliced lever, replacing bridge sequences between IQ-cores 1-2, 3-4, and 5-6.
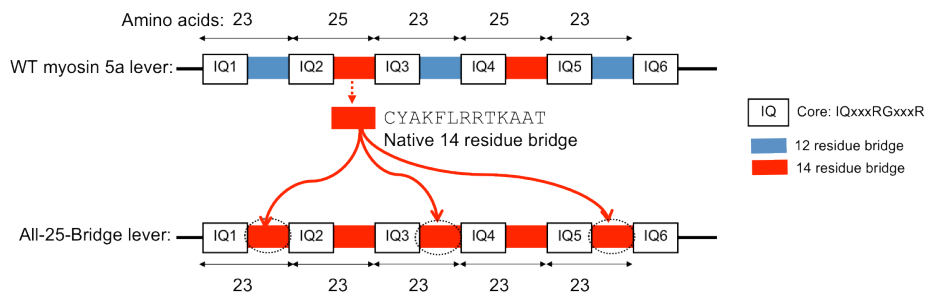
Figure 5.9: Design scheme for All-25-Bridge

Specifically, the nucleotide sequence:

^{2459}TGCTATGCTAAGTTTCTGCGCAGAACCAAGGCAGCAACCACC^{2500} (^{807}CYAKFLRRTKAATT^{820})

was used to replace the three sequences

^{2390}AAGAGATACCTGTGTATGCAGAGGGCAGCCATCACA^{2425} (^{784}KRYLCMQRAAIT^{795})

^{2534}AGGAGGTACAAGATTAGACGAGCTGCCACGATTGTT^{2569} (^{832}RRYKIRRAATIV^{843})

^{2678}ACACATTATAAGAGGACCATGAAAGCCATCGTCTAC^{2713} (^{880}THYKRTMKAIVY^{891})

This construct was named All-25-Bridge.

5.3 Methods and Materials

The following sections detail the methods used for cloning, expression and purification of the lever mutants and wild-type HMM clones, plus information on the various characterisation techniques used.

5.3.1 Construction of wild-type myosin 5a HMM clone

A wild-type myosin 5a HMM clone (HMM-WT) was created at the outset of the project, to provide a DNA template for subsequent construction of the lever mutants and to act as a wild-type control in characterisation work. The cDNA sequence used for the HMM-WT clone follows amino acids 1-1090 of the mouse myosin 5a heavy chain sequence (NCBI GenBank accession: NM_010864). The sequence length was chosen to be consistent with myosin 5a HMM-like clones used in earlier studies of myosin 5a (Wang et al., 2000). A 'FLAG tag' epitope (DYKDDDDK) was added to the C-terminus to facilitate purification.

The HMM-WT clone was constructed from a pre-existing myosin 5a construct (a myosin 5a-HMM chimera with a Caldesmon domain added at its C-terminus) kindly supplied by Dr Attila Nagy (LMP-NIH). Source cDNA used to derive HMM-WT was in the cloning site of a pFastBac™1 donor plasmid (Invitrogen – Life Technologies). DNA for HMM-WT was generated from this template by polymerase chain reaction (PCR), using a Platinum Pfx DNA polymerase PCR kit (Invitrogen – Life Technologies) and oligonucleotide primers (Eurofins MWG Operon, Huntsville AL, USA) designed for the 5′ and 3′ ends of the HMM-WT sequence, Table 5.1. To facilitate later insertion into a

pFastBac™1 plasmid, the primers were designed to introduce flanking restriction site sequences for the restriction endonucleases *BssH*II and *Spe*I at the 5′ and 3′ ends of the HMM-WT sequence.

Table 5.1: Primers and PCR cycle used to generated HMM-WT DNA sequence.

Forward primer: *MVA5*	5′-ATAATAGCGCGCATGGCCGCGTCCGAGCTCTACACC-3′
Reverse primer: *MVA3*	5′-TATTATGACTAGTGACTCCTCCAGGCGACTG-3′
PCR reaction: (Pfx Platinum DNA polymerase PCR kit, Invitrogen)	Reaction mix (100 μL volume): 10 μL 10x Pfx amplification buffer (Invitrogen), 3 μL dNTP mix (10 mM), 2 μL MgSO$_4$ (50 mM), 3μL primer *MVA5*, 3μL primer *MVA3*, 0.5 μL template DNA, 77.7μL H$_2$O, 0.8 μL Pfx (added last). Heating cycle: 1. 95 °C, 3 min 2. 95 °C, 1 min 3. 52 °C, 1 min 4. 68 °C, 3 min 5. Repeat 2.-4. ×31 times 6. 68 °C, 5 min 7. 4°C hold indefinite

The HMM-WT PCR product was checked for its size by agarose gel electrophoresis, running a 1% agarose gel (supplemented with 0.5 μg/mL ethidium bromide) in TAE buffer (40 mM Tris base, 40 mM acetic acid, 1 mM EDTA, pH 8.0). The DNA band identified at correct size for HMM-WT (~3.3 kbp) was excised from the gel under UV light illumination and the DNA purified using a QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany).

Purified DNA was then inserted into a pFastBac™1 donor plasmid between *BssH*II and *Spe*I restriction sites by restriction digest and ligation reactions. The target vector had also already been modified (Dr Attila Nagy, LMP-NIH) to contain the FLAG tag nucleotide sequence (GACTACAAGGACGACGATGATAAG) followed by a stop codon (TGA), between the *Not*I and *Xho*I restriction sites, downstream of *Spe*I, see Figure 5.10. After cloning, the complete amino acid expression sequence expected for the HMM-WT construct was therefore:

[Res. 1-1090 mouse myosin 5a]-**[SLVAAA]**-**[DYKDDDDK]**

From *Spe*I and *Nco*I restriction site seqeunces in vector cloning site.   FLAG tag.
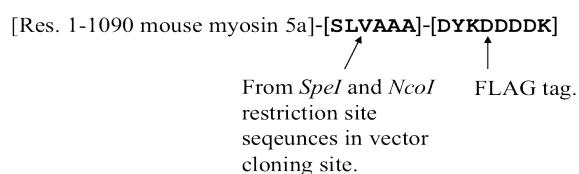
The sequence SLVAAA between the myosin 5a fragment and FLAG tag result from the *Spe*I and *Not*I restriction site sequences still in the vector multiple cloning site.

194

To insert the HMM-WT sequence into the target vector, the FLAG-modified vector and insert PCR product were first digested with *Spe*I and then *BssH*II restriction enzymes (New England BioLabs Inc. (NEB), Ipswich MA, USA), following supplier reaction conditions. After each digest, insert and linearised vector reaction products were checked by electrophoresis on a 1% agarose gel, before recovery and purification using QIAquick gel extraction kit (Qiagen). The linearised vector was further treated with shrimp alkaline phosphatase (NEB) to prevent re-ligation. Ligation of HMM-WT DNA sequence into the target vector was carried out using T4 DNA ligase (Fermentas – Thermo Fisher Scientific Inc., Glen Burnie MA, USA). Insert DNA was added in 3:1 excess over vector. The reaction (2 μL enzyme, 2 μL buffer (Fermentas – Thermo Fischer Scientific), 4 μL vector, 12 μL insert) was run for 1 hr at 20 °C.
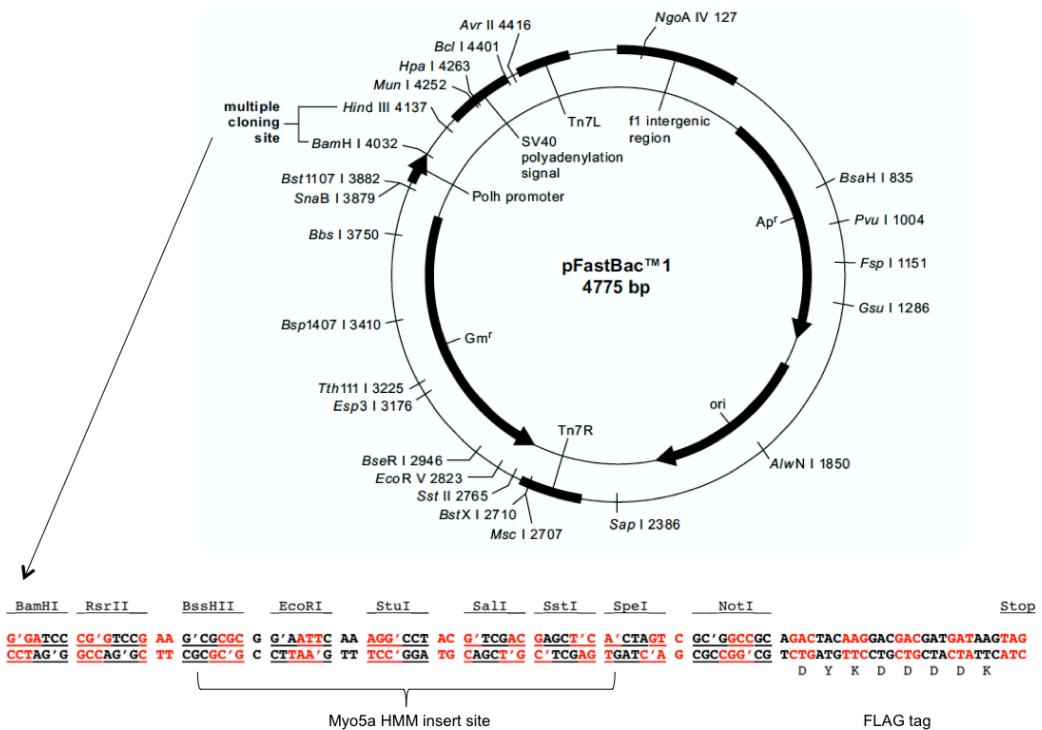
Figure 5.10: Map of Invitrogen pFastBac™1 vector with enlargement showing the multiple cloning site sequence.

The vector was digested with *BssH*II and *Spe*I restriction enzymes to form the HMM-WT insertion site. A FLAG tag and stop codon had already been engineered in the vector downstream of *Not*I (Dr Attila Nagy, LMP-NIH). The complete expressed a/a sequence expected was: [res. 1-1090 mouse myosin 5a]-[SLVAAA]-[DYKDDDDK].

The HMM-WT-pFastBac™1 ligation reaction was transformed into Max Efficiency DH5α Competent *E. coli* cells (Invitrogen) using a heat shock protocol: 6 μL of the ligation reaction was added to a thawed aliquot of cells, incubated on ice for 45 mins, heated-shocked for 60 s at 42 °C in a water bath, cooled on ice for 2 mins, added to 800 μL of Super Optimal broth with Catabolite repression medium (S.O.C) (Invitrogen) in a 15 mL Falcon tube, before shaking incubation for 1 hr (270 RPM, 37 °C). 100 and 400 μL

volumes of this culture were spread on Luria Bertani (LB)-agar-ampicillin (100 μg/mL) plates (KD medical, Columbia, MD, USA) and incubated overnight at 37 °C. Colonies were picked from the plates and cultures grown by adding them to 10 mL of LB with 100 μg/mL ampicillin and shaking overnight in an incubator (270 RPM, 37 °C). Plasmid DNA was then extracted and purified from the cultures using a QIAprep Spin Miniprep Kit (Qiagen), eluting the purified DNA with H_2O.

The final HMM-WT DNA insert in the modified pFastBac™1 plasmid was verified by DNA sequencing over the entire cloned DNA sequence (ACGT Inc., Wheeling, IL, USA) and by a separate PCR reaction to confirm that a HMM-WT-sized sequence (~3.3 kbp) could be re-amplified from the plasmid using the 5′ and 3′ primers *MVA5* and *MVA3* (Table 5.1) and seen by agarose gel electrophoresis.

5.3.2   Construction of lever mutants

The myosin 5a lever mutants, termed All-23 and All-25, were constructed using the HMM-WT clone DNA as a template. Two mutagenesis strategies were devised, one using a series of PCR reactions to alter the native IQ-bridge sequences along the lever one-by-one, working stepwise from the N-terminal to the C-terminal end, and another single-reaction strategy, involving ligation of an entire mutated lever-tail sequence and motor domain into a target plasmid in one reaction. For the stepwise strategy, each mutagenesis product formed the template DNA for the next reaction, and a set of intermediate lever mutants were thus formed as the final All-23 or All-25 target constructs were approached. For the double ligation strategy, so-called In-Fusion® reactions (Clonetech Laboratories Inc. (Clonetech), Mountain View, CA, USA) were performed using a proprietary enzyme that allows separate DNA fragments with overlapping ends to anneal.

Manufacture of both types of All-23 and All-25 mutants, that is 'insert/delete' or 'splice' type mutants (sections 5.2.1-5.2.2), were attempted by both the stepwise and In-Fusion strategies. This resulted in an outcome of four possible mutant sequences: All-23-Delete, All-23-Bridge, All-25-Insert and All-25-Bridge. Each was approached by two mutation strategies, a total of eight reaction sequences (four stepwise and four In-Fusion) to perform. Details of the stepwise and In-Fusion ligation protocols are given in the following sections.

5.3.2.1   Stepwise mutagenesis protocol

In the stepwise approach, DNA for each lever mutant was created by introducing the mutations required in each IQ-bridging sequence one IQ motif at a time, working in an N- to C-terminus direction along the lever. A separate PCR reaction was carried out for each

23 to 25 a/a IQ-bridge change, or *vice versa*. A Quikchange II XL Site-Directed Mutagenesis Kit (Stratagene) (which uses the high fidelity DNA polymerase *PfuTurbo*) and a pair of specifically-designed oligonucleotide primers (synthesized by Eurofins MWG Operon) were used in the separate PCRs, each reaction following a common protocol, Table 5.2. All lyophilised primer stocks were dissolved in H_2O to ~1250 μg/mL before use. The forward sequences of the DNA primers pairs used to mutate specific IQ-bridge sequences are listed in Table 5.3. Each pair of complimentary primers was designed to anneal to the template DNA around the desired mutation site, their sequences deviating from the template sequence only to encode the mutant sequence required. After an initial round of annealing and polymerisation the PCR should begin amplifying the mutated DNA sequence required.

Following each reaction, the PCR product was treated with 1 μL of *Dpn*I enzyme (Quikchange II kit) at 37 °C for 1 hr, to digest methylated parental DNA. Treated reactions were then ligated into a linearised pFastBac™1 (containing the FLAG-stop sequence) and transformed into DH5α cells (Invitrogen) by adding 7 μL of ligation reaction to an aliquot of freshly-thawed cells and following the heat shock protocol as described for HMM-WT-creation in 5.3.1. 10 mL cultures were grown from bacterial colonies picked from agar-ampicillin plates and the plasmid DNA extracted and purified by mini-prep, also as described in 5.3.1.

Table 5.2: PCR reaction components and heating cycle used in all stepwise mutation reactions.

PCR reaction:	Reaction mix (51 μL volume):
Quikchange II XL Site-Directed Mutagenesis Kit (Stratagene) with *PfuTurbo* DNA polymerase.	1 μL of a 100× dilution of template DNA, 1 μL of a 10× dilution of primer 1, 1 μL of a 10× dilution of primer 2, 1 μL of dNTP mixture (10 mM), 3 μL Quik solution. (Stratagene), 5 μL 10x running buffer (Strategene), 38 μL H_2O, 1 μL *PfuTurbo* DNA polymerase (added last)
	Heating cycle:
	1. 95 °C, 1 min
	2. 95 °C, 50 s
	3. 60 °C, 50 s
	4. 68 °C, 8 min
	5. Repeat steps 2.-4. ×17
	6. 68 °C, 7 min
	7. 4 °C, hold indefinitely

Table 5.3: Forward DNA primer sequences used in PCR reactions (Table 5.2) to create lever mutant constructs.

NB: Only the forward primer sequences used to produce the mutations in IQ-core bridging sequences are listed. For each primer, a reverse complement sequence was also obtained and used in the PCR.

Delete mutant: All-23-Delete	
IQ2 – IQ3	5'-GATGCTATGCTAAGTTTAGAACCAAGGCAGCAACCACC-3' Encodes: CYAKF[..]RTKAATT ([..] = LR in native)
IQ4 – IQ5	5'-CAAGAAATAGGTATCGCAAGCGTGAATACAAAGCAGTCATC-3' Encodes: RNRYRK[..]REYKAVI ([..] = IL in native)
Insert mutant: All-25-Insert	
IQ1 – IQ2	5'-GGAAGAGATACCTGTGTCTGCTGATGCAGAGGGCAGCCATCAC-3' Encodes: KRYLC**LL**MQRAAI (**LL**= insertion)
IQ3 – IQ4	5'-CGCAGGAGGTACAAGATTCTGCTGAGACGAGCTGCCACGATTG-3' Encodes: RRRYKI**LL**RRAATI (**LL** = insertion)
IQ5 – IQ6	5'-CCGTACACATTATAAGAGGCTGCTGACCATGAAAGCCATCGTCTAC-3' Encodes: RTHYKR**LL**TMKAIVY (**LL** = insertion)
Splice mutant: All-23-Bridge	
IQ2 – IQ3	5'-GGGGCTATCAGGCTCGAAAGAGATACCTGTGTATGCAGAGGGCAGCCATC ACAATTCAAAAGTACTGGCG-3' Encodes: GYQAR**KRYLCMQRAAIT**IQKYW (Bold = 12 a/a splice-in)
IQ4 – IQ5	5'-GAGGCTACTTGACAAGAAAGAGATACCTGTGTATGCAGAGGGCAGCCATC ACAATTCAGAAACGTGTCCG-3' Encodes: GYLTR**KRYLCMQRAAIT**IQKRV (Bold = 12 a/a splice-in)
Splice mutant: All-25-Bridge	
IQ1 – IQ2	5'-CGTGGGTGGCTTCTAAGGTGCTATGCTAAGTTTCTGCGCAGAACCAAGGC AGCAACCACCGTGCAGCGATACGTGCGG-3' Encodes: RGWLLR**CYAKFLRRTKAATT**VQRYVR (Bold = 14 a/a splice-in)
IQ3 – IQ4	5'-CGCATGTATGTGGTCCGCTGCTATGCTAAGTTTCTGCGCAGAACCAAGGC AGCAACCACCATTCAGTCTTACTTGAGAG-3' Encodes: RMYVVR**CYAKFLRRTKAATT**IQSYLR (Bold = 14 a/a splice-in)
IQ5 – IQ6	5'-CGTGGCTGGCTGGCCCGTTGCTATGCTAAGTTTCTGCGCAGAACCAAGGC AGCAACCACCCTTCAGTGCTGCTTCCGG-3' Encodes RGWLAR**CYAKFLRRTKAATT**LQCCFR (Bold = 14 a/a splice-in)

For each new lever construct produced, DNA sequencing was carried out over the entire lever section (ACGT Inc., Wheeling, IL, USA). Sequence-verified construct DNA was stored at −20 °C for future use.

5.3.2.2   In-Fusion mutagenesis protocol

The In-Fusion strategy was devised as a potentially faster, more direct, method to create the complete All-23 and All-25 lever mutants in single-step reactions.

Four ~950-bp dsDNA sequences encoding the complete lever-plus-tail sequences for the All-23 and All-25 constructs (*i.e.* a sequence for each of the 'insert', 'delete' and two 'splice'-type mutants) were synthesized as 'minigenes' by Integrated DNA Technologies (IDT) Ltd (Coralville, IA, United States). On delivery, the minigene sequences were PCR-extracted and amplified from their plasmids, and purified as linear dsDNA from agarose gel for later use in the In-Fusion ligation reactions. Using another set of specially designed primers, four motor domain sequences (each running to IQ 1 in the lever), one to match each lever-tail minigene, were PCR-amplified using HMM-WT DNA as a template. *NB*: four unique motor sequences (each slightly varied in codon sequence at its C-terminus end) were required in order to dovetail with each minigene, because, for synthesis and cloning reasons, the codon sequence of each overlap in the lever-tail minigene (at IQ-core 1) had been varied by IDT Ltd. A pFastBac™1 plasmid (pre-modified with FLAG-stop) was cut at the *Stu*I restriction site using *Stu*I restriction enzyme (Roche) in supplier-recommended conditions (Buffer B, Roche) to prepare for the ligation reaction.

Using the appropriate motor fragment, lever-tail minigene and linearised vector (prepared as above), multiple insert cloning reactions were set-up using an In-Fusion Dry-Down PCR Cloning kit (24) (Clonetech). In each reaction, the aim was to join the motor fragment with its lever-tail minigene and to ligate the whole motor-lever-tail construct into the target vector, Figure 5.11. These multiple insert ligations are accomplished by using the In-Fusion enzyme which requires a 15 bp overlap at the fragment ends. The enzyme works by creating single-stranded regions at the ends of the double-stranded vector and any insert fragments, which then fuse in a PCR reaction due to the 15 bp homology at their ends.

4 μL lever-tail DNA, 4 μL motor DNA and 2 μL linearised pFastBac1-*Stu*I were added to a In-fusion Dry-Down PCR bead, vortexed and then incubated for 15 min at 37 °C, then 15 min at 50 °C. The reaction was transformed into DH5α cells, plated on LB-agar-ampicillin and colonies grown overnight using the protocols as previously described in 5.3.1. Colonies were then PCR-screened to see if an HMM-sized fragment could be amplified from the plasmids and seen on an agarose gel. Mini-cultures were grown from positive colonies and plasmid DNA isolated also as described in 5.3.1. To verify the In-Fusion cloning reaction, DNA sequencing of the recombinant plasmid was carried out for

each mutant (ACGT Inc., Wheeling, IL, USA), sequencing over a section encompassing the motor-lever boundary through to the end of HMM tail and FLAG tag.
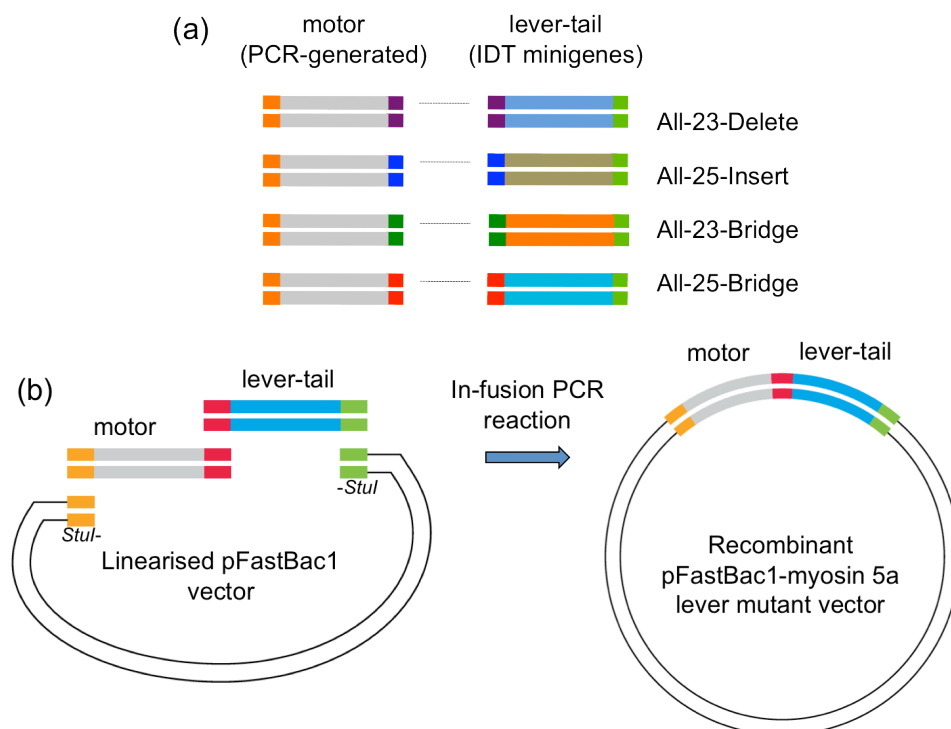
Figure 5.11 Reaction scheme for In-fusion multiple insert cloning reaction.

(a) dsDNA of four motor fragments (running to IQ-core 1) were PCR-generated to dovetail with four lever-tail minigenes (IDT Ltd.), one corresponding to each lever mutant. The overlap section between motor and lever-tail was 15 bp and varied in exact codons used for each (for IDT synthesis/cloning reasons). (b) schematic diagram of the In-Fusion multiple insert cloning reaction. Based on the unique 15 bp overlaps at the ends of the fragments, the motor, lever-tail and linearized target vector are recombined using the Clonetech In-Fusion enzyme.

5.3.3 Protein expression and purification

HMM-WT and all lever mutant constructs were expressed and purified using the same protocols. Proteins were expressed using an Invitrogen *Bac-to-Bac*® Baculovirus Expression System, using *Spodoptera frugiperda* (Sf9) insect cells. Following expression, cells were lysed and the myosin 5a constructs were isolated by binding to anti-FLAG antibody-coated resin, and purified by liquid chromatography.

The Baculovirus/Sf9 cell expression process is summarised in Figure 5.12 with more detailed methods for each step given in 5.3.3.1-5.3.3.5 thereafter. In outline, a recombinant pFastBac™1 donor plasmid containing a gene of interest is transformed to DH10Bac™ *E. coli* cells, where the cloned gene is transposed into a plasmid, known as a bacmid, which also carries the genomic DNA for the baculovirus. Cells containing recombinant bacmid DNA are cultured against antibiotics to increase the amount of bacmid DNA, which is then isolated. The recombinant bacmid is then transfected into Sf9 cells which, once infected,

begin to generate the recombinant baculovirus and foreign gene of interest (here a myosin 5a construct). The virus multiplies and infects more cells, which express more of the foreign gene and produce more copies of the virus. Recombinant virus is harvested and amplified to a higher 'titer' (concentration) through rounds of infection of fresh Sf9 cells, until the concentration is high enough to produce sufficient quantities of the protein of interest. The process is a balance between generating sufficient quantities of virus to generate the protein of interest, whilst not allowing the infection to mature to the point that too many cells are killed before protein is harvested.
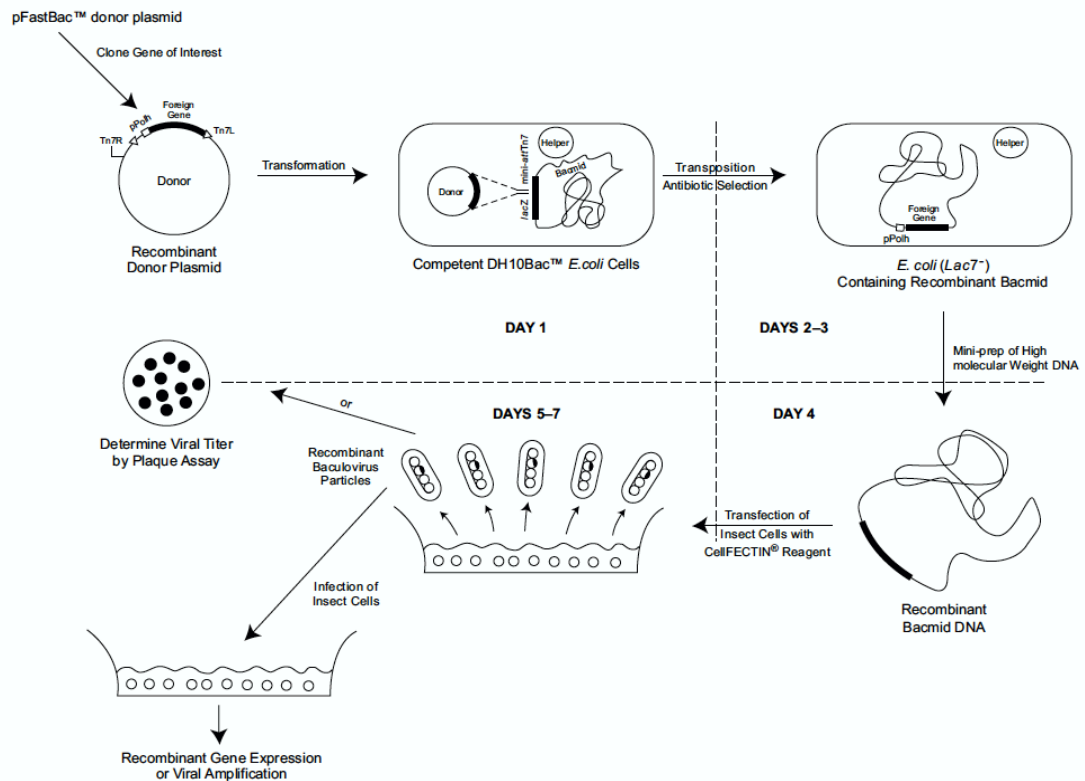
Figure 5.12: Scheme for Invitrogen Bac-to-Bac baculovirus and insect cell expression system.

These steps were followed (by AN and/or DR see Table 5.5) to produce amplified baculovirus and recombinant protein overexpression of HMM-WT and each lever mutant. This figure is adapted from the Invitrogen Bac-to-Bac® expression system manual.

Invitrogen's standard *Bac-to-Bac*® protocols were followed to produce each recombinant protein. The main steps, including recombinant gene transposition, protein expression and purification are detailed below.

5.3.3.1   Donor plasmid transformation into DH10Bac™ cells

Recombinant pFastBac™1 donor plasmids containing myosin 5a constructs were transformed into DH10Bac™ *E. coli* cells (Invitrogen) (where the cloned insert is transposed into baculovirus genomic DNA) by heat shock: 5 µL of plasmid DNA was

gently mixed with a thawed aliquot of DH10Bac™ cells and incubated on ice for 30 mins. The mixture was shocked at 42 °C for 45 s in a water bath, re-chilled on ice for 2 mins, transferred to 900 μL S.O.C medium (Invitrogen) in a 15 mL round bottom tube and left in a shaking incubator for 4 hrs (37 °C, 270 RPM). 100 and 400 μL volumes of culture were spread on pre-prepared LB-agar plates containing antibiotics (50 μg/mL kanamycin, 7 μg/mL gentamicin, 10 μg/mL tetracycline) plus 100 μg/mL Bluo-gal and 40 μg/mL isopropyl β-D-1-thiogalactopyranoside (IPTG). Plates were incubated for ~48 hr at 37 °C.

5.3.3.2 Isolation of recombinant bacmid DNA

DH10Bac™ colonies were screened by blue-white selection from the plates (through Bluo-gal use in the plate mixture, white-coloured colonies should contain the desired recombinant bacmid DNA). Four white colonies and a single blue were first streaked to fresh plates and incubated overnight to confirm their colour. Positively-identified white colonies were picked and added to 5 mL of LB containing kanamycin, gentamicin and tetracycline (concentrations as section 5.3.3.1) and cultured overnight (37 °C, 270 RPM shaking). Bacmid DNA was isolated from cultures using steps 1-3 of the microcentrifuge protocol from a QIAprep Spin Miniprep Kit (Qiagen), before precipitation of the DNA by adding 800 μL isopropanol to each 600 μL of centrifuged cell lysate and storing at −20 °C for 5 mins. This mixture was centrifuged on the bench (15 min, $13,800 \times g$), the supernatant discarded and 1 mL of 70% ethanol (at −20 °C) added to the pellet left behind. This was benchtop-centrifuged (5 min, 13,800 g), the supernatant carefully removed by wicking with a tissue and the DNA pellet remaining dried in air. Finally, 40 μL H_2O was added to re-dissolve each pellet and the isolated bacmid DNA samples were stored at 4 °C before use (freeze-thaw cycles can damage long bacmid DNA).

5.3.3.3 Verification of transposition to bacmid.

To confirm the transposition of the myosin 5a constructs to the bacmid, PCR analysis was used. Here, pCU/M13 primers are directed at sequences flanking the gene insertion site in the bacmid. If transposition is successful, a PCR product at 2.3 kbp plus clone insertion size (3.3 kbp) (*i.e*. 5.6 kbp total) should be seen by agarose gel electrophoresis. The PCR screening reaction used is detailed in Table 5.4.

Table 5.4: PCR screening reaction to verify transposition of genes to bacmid

PCR reaction:	Reaction mix:
	1 μL of 10-fold dilutions of pCU/M13 primer stocks and 1 μL of bacmid DNA were added to a PuReTaq Ready-To-Go dry PCR bead (GE Healthcare Life Sciences) pre-dissolved with 22 μL H$_2$0.
	Heating cycle: 1. 93 $^\circ$C, 3 min 2. 93 $^\circ$C, 45 s 3. 55 $^\circ$C, 45 s 4. 72 $^\circ$C, 7 min 5. repeat 2.- 4., ×35 times 6. 72 $^\circ$C, 10 min 7. 4°C, hold indefinitely

Reactions samples were analysed by agarose gel electrophoresis to look for amplification of the bacmid insertion product.

5.3.3.4 Transfection of Sf9 cells with recombinant bacmid DNA

Successfully-verified recombinant bacmid DNA was transfected into Sf9 insect cells following the *Bac-to-Bac*® protocol. Typically, 5 μL of bacmid DNA and ~2 x 10^6 cells in 2 mL of Sf-900 II serum free medium (Sf-900 II SFM) (Invitrogen), containing Cellfectin® reagent (Invitrogen) and antibiotics (penicillin at 50 units/mL and streptomycin at 50 μg/mL final concentrations) were added to a 35 mm well of a 6-well plate to begin transfection. The plate was covered and incubated in a humid environment (a sealed plastic bag containing moist tissue paper) for 5 hrs at 27 °C, before removal of the transfection mixture from the cell monolayer formed in the well and refreshment with 2 mL of Sf-900 II SFM containing antibiotics (penicillin and streptomycin as above). The first baculovirus harvest (known as P1) was taken from the supernatant after 72 hrs of further incubation at 27 °C.

For all cell transfections, virus amplification and final infection stages, a fresh supply of Sf9 cells was maintained by culturing them in Sf-900 II SFM in 100 mL volumes in 250 mL conical flasks. These were incubated at 27 °C and shaken at 150 RPM. Cell density was maintained at 1-4 × 10^6 cells/mL by splitting cells to fresh medium when required. Cell density and condition was monitored by staining 10 μL cell-culture samples with 10 μL Trypan blue and examining under a light microscope in a haemocytometer (0.1mm depth, 0.0025 mm^2 grid squares) (Paul Marienfeld GmbH & Co. KG, Lauda-Königshofen, Germany).

5.3.3.5 Virus amplification and protein expression

Recombinant baculoviruses for the myosin 5a constructs were amplified from P1 stage through 3 (or 4) more cycles to P4 (or P5) stage, according to the *Bac-to-Bac*® protocol. At each step, virus was harvested from infected cell culture supernatant (clarified by 5 min, $500 \times g$ centrifugation) and used to re-infect fresh Sf9 cells in Sf-900 II SFM:

- *To obtain P2 stock*: 1 mL of P1 virus stock was added to 25 mL of fresh Sf9 cells (at 1×10^6 cells/mL) in Sf-900 II SFM and incubated at 27 °C with 150 RPM shaking for 72 hrs.
- *To obtain P3 stock*: 10 mL of P2 virus was added to 100 mL Sf9 cells (1×10^6 cells/mL), incubated for 72 hrs (27 °C, 150 RPM).
- *To obtain P4 stock*: 10 mL of P3 virus added to 100 mL Sf9 cells (1×10^6 cells/mL), incubated for 72 hrs (27 °C, 150 RPM) or until infected cells began bursting.

For final large-scale protein expression, 50 mL of P4 (or P5) virus was added to 1 L of Sf9 cells ($1.5\text{-}2.0 \times 10^6$ cells/mL) in Sf-900 II SFM. 25 mL of a recombinant baculovirus for *Xenopus* calmodulin (CaM) (which has an identical sequence to mouse) (produced by Dr Attila Nagy, LMP-NIH) was also added to the final infection, so that the myosin 5a construct and CaM were co-expressed in the Sf9 cells. Cell cultures were incubated for 3-4 days (until cells are enlarged but at least 90% are still alive) at 27 °C, 150 RPM. Final cell cultures were centrifuged (Sorval RC50 centrifuge, Sorvall SH-3000 rotor, $1000 \times g$, 5 min, 4 °C) and cell pellets retained. These were quickly frozen in liquid nitrogen and stored at −80 °C.

5.3.3.6 Protein purification by FLAG tag

NB: In the following purification procedure, cells pellets, lysate and proteins samples were kept on ice throughout working, and flasks and buffers were pre-chilled at 4° C.

Frozen cell pellets were each re-suspended in 50 mL of $1.5 \times$ extraction buffer, the $1 \times$ formulation being 10 mM MOPS, 200 mM NaCl, 10 mM MgCl$_2$, 1 mM EGTA, 3 mM NaN$_3$, 0.1 mM DTT, 0.1 mM PMSF, 1 mg/L leupeptin, 2 mM ATP and a protease inhibitor cocktail, pH 7.3. Cells were mechanically lysed using a glass homogenizer and 4 mins of pulse sonication. Lysate sample was checked under a microscope (checking for any remaining intact cells to determine if lysis was sufficient) and 25 mL volumes transferred to 50 mL centrifuge tubes, before 15 mins centrifugation at 4 °C (Sorvall RC50, SS-34 rotor, $47,800 \times g$). Supernatant was retrieved and 3 mL of Anti-FLAG M2 resin (A2220, Sigma-Aldrich) (washed according to supplier instructions) was added to each supernatant, which was then gently rocked for 2 hrs (or alternatively overnight) at 4

°C. Resin mixture was centrifuged (Beckman Coulter Allegra X-22R centrifuge, SX4250 rotor, 2 min, 113 × g 1 kRPM, 4 °C) and pellets re-suspended with 15 mL re-suspension buffer (10 mM MOPS, 0.5 M NaCl, 0.1 mM EGTA, 3 mM NaN$_3$, 0.1 mM DTT, 0.1 mM PMSF, 1 mg/L leupeptin, 1 mM ATP, 5 mM MgCl$_2$, pH 7.2) through 5 cycles. The final resin re-suspension was loaded onto a column (0.8x4 cm Poly-Prep Chromatography Column, Bio-Rad Laboratories Ltd, Hemel Hempstead, UK) at 4 °C and washed with 1 column volume re-suspension buffer followed by 3 column volumes of HMM buffer (10 mM MOPS, 0.1 mM EGTA, 3 mM NaN$_3$, 0.1 mM DTT, 0.1 mM PMSF, 1 mg/L leupeptin, pH 7.2). Elution of protein was initiated by washing with 3 column volumes of elution buffer (0.3 mg/mL FLAG peptide (American Peptide Company, Sunnyvale CA, USA) dissolved in 80% HMM buffer, 20% re-suspension buffer minus its MgCl$_2$ and ATP). Ten ~1.0 mL fractions of the column elution were collected under gravity flow (at ~1-2 mins elution time per fraction).

Fractions were characterised using a Bradford assay: 10 µL fraction samples were added to 50 µL Bradford reagent (Bio-Rad Protein Assay, Bio-Rad Laboratories Ltd) drops in a 96-well plate. Fractions identified as containing protein (usually numbers 1-4) were pooled and either dialysed directly (see paragraph below) or, in later preparations, loaded onto a Q-sepharose column for an additional ion-exchange purification step before dialysis. When used, the Q-sepharose column (containing ~1 mL of fast flow Q-sepharose resin (GE Health Care) was equilibrated in a 4 °C environment with 10 mL of HMM buffer supplemented with 0.1 M NaCl. Protein-containing fractions were loaded onto the column which was then washed with another 10 mL of 0.1 M NaCl HMM buffer. Protein was eluted using a step increase in salt (rather than a gradient) by adding ~8 mL of the re-suspension buffer (contains 0.5 M NaCl) and collecting 0.5-1.0 mL fractions under gravity flow. Collected fractions were analysed again by Bradford assay.

Protein-containing fractions (either from AntiFLAG-column directly or from the Q-sepharose ion exchange column if used) were pooled and then dialysed (dialysis cassettes, MWCO 10 kDa, Thermo Fischer Scientific) at 4 °C against a 2 L volume of HMM buffer supplemented with 500 mM KCl (*NB*: 100 mM KCl was used in some earlier preparations) for ~16 hrs, refreshing the buffer after ~12 hrs.

To characterise the purification process and final preparation, samples from stages throughout the cell lysis, FLAG/Q-sepharose purification and final dialysis were taken for examination by SDS-PAGE. 1 mm thick 4-20 % polyacrylamide Tris-Glycine gradient gels were used to examine the samples unless otherwise stated. Methods are given in 5.3.5.1.

The HMM (*i.e.* two-headed) molecule concentration was determined from the final purified samples by UV-visible light absorption measurement as detailed in 2.6.1.

Finally, fresh protein samples were kept on ice or, if not intended for use within 2-3 days, were drop-frozen by dropping 25 µL droplets into a basin of liquid nitrogen and storing the frozen droplets in a liquid nitrogen Dewar.

5.3.4   Division of labour for cloning and protein preparation.

All cloning, protein expression and purification was carried out either by the author of this thesis (DR) or by Dr Attila Nagy (AN) at LMP-NIH. AN contributed about 50% of the work and was responsible for the baculovirus and protein expression stages for all constructs. Purifications of repeated preparations of a construct were shared. On some later occasions, purifications were carried out by DR at University of Leeds, using frozen cell pellets shipped on dry-ice from LMP-NIH. Table 5.5 summarises who contributed which elements in the work.

Table 5.5: Summary of labour division in preparation of protein constructs.
DR = Derek Revill, AN = Dr Attila Nagy (LMP-NIH)

Construct	Cloning method	Cloning to pFastBac-1	Bacmid production	Baculovirus & expression	Purification
HMM-WT	PCR	DR	DR	AN	AN & DR
All-23-Delete	*In-Fusion*:	DR	DR/AN	AN	AN
	Step-wise:	DR/AN	-	-	-
All-25-Insert	*In-Fusion*:	-	-	-	-
	Step-wise:	DR/AN	AN	AN	AN & DR
All-23-Bridge	*In-Fusion*:	AN	AN	AN	AN & DR
	Step-wise:	DR	DR/AN	-	-
All-25-Bridge	*In-Fusion*:	AN	AN	-	-
	Step-wise:	DR/AN	DR/AN	-	-

DR/AN = begun by DR, completed by AN.
AN & DR = Process carried out on two or more separate occasions either by AN or DR.

As can be seen from Table 5.5, some of the lever constructs remained unfinished. This was usually because one cloning method (In-fusion or step-wise) had produced the desired clone, so cloning by the other technique was halted. In one case, the All-25-Insert construct, IDT Ltd were unable to produce the full minigene sequence required, so this construct was pursued only by step-wise mutagenesis. In the case of the All-25-Bridge construct, cloning attempts by both In-fusion and step-wise methods both failed, so this construct was never produced (though an intermediate construct from the step-wise method, with IQs 1-2 changed to a 25 a/a spacing, was expressed and purified). Protein was successfully expressed and purified at least once for HMM-WT, All-23-Delete,

All-23-Bridge and All-25-Insert, plus the intermediate construct with one IQ change for All-25-Bridge.

5.3.5   Characterisation techniques used

The principal techniques and assays used to characterise the lever mutants were SDS-PAGE and Western Blotting, negative stain EM, NADH-coupled ATPase assay and a dual-colour TIRF microscopy assay.

5.3.5.1   SDS-PAGE

1 mm thick Tris-Glycine gels with of a 4-20% polyacrylamide gradient (Novex, Life Technologies) were used for SDS-PAGE of myosin 5a-HMM constructs. The use of the gradient allows simultaneous resolution of both the myosin 5a-HMM heavy chain and the lower MW CaM bands. The MW marker used was typically See Blue plus 2 (Invitrogen – Life Technologies). To denature protein samples prior to electrophoresis, 40 μL protein samples were mixed with 10 μL of 5× SDS denaturing buffer (5× composition: 0.25% w/w Bromophenol blue, 0.1 M DTT, 50% v/v glycerol, 10% w/w SDS) and boiled at 100 °C for 3-5 mins, before loading 25 μL into the gel wells for electrophoresis. Electrode buffer was 25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3 and gels were run at ~150 V for 90 mins. Gels were stained for ~2 hours in 0.1% Coomassie brilliant blue R-250 solution (0.1% coomassie blue, 40% methanol, 10% acetic acid) and de-stained using a 40% methanol, 10% acetic acid solution. Gels were scanned using a LI-COR Odyssey Infrared Imager (LI-COR Biosciences – Biotechnology, Lincoln, NE, USA).

5.3.5.2   Western Blotting

Standard immunoblotting methods were used to probe for the presence of the FLAG tag epitope, using a monoclonal Anti-FLAG antibody raised in mouse (Sigma-Aldrich, F3165). To perform the assay, protein samples were first separated by SDS-PAGE, before transfer to Amersham Hybond-ECL nitrocellulose membranes (GE Healthcare) by electroblotting. Nitrocellulose membranes were then blocked for 1 hr by incubation (in a falcon tube on a rolling mixer) in a 5 mL solution of 3% non-fat dried milk (Applichem), solubilised in Tris-buffered saline (TBS). 5 μL of the primary Anti-FLAG antibody was then added and allowed to incubate on rollers for another 1 hr. The blocking solution was discarded and three 15 min rolling washes, in 10 mL TBS supplemented with 0.05% Tween detergent (TBS+T), were conducted. Wash buffer was then exchanged for 5 mL blocking solution (3 % non-fat milk in 5 mL TBS as above) and the membrane incubated on rollers for 1 hr. 0.5 μL of the secondary horse-radish peroxidase (HRP)-linked antibody (Anti-mouse HRP raised in goat, Invitrogen) was added and rolling incubation for a

further 1 hr continued. Two further 15 min washes in TBS+T and one final 15 min wash in TBS were performed before storing the membrane (in TBS) at 4 °C for future developing.

Immunoblotted membranes were warmed up (on rollers) for 1 hr to room temperature, and then treated with SuperSignal West Pico Chemiluminescent Substrate (Pierce, Thermo Fisher Scientific) for 5 mins. The membranes were then assessed by chemiluminescence detection on Kodak X-Ray film at various exposure times (though a rapid 'touch exposure' provided more than sufficient). Films were developed and fixed using a Xograph Compact X2 machine.

5.3.5.3   Negative stain EM:

Grid preparation and staining methods used were as described in 2.8.2-2.8.5. The author of this thesis performed all EM at University of Leeds, using a JEOL 1200EX microscope (equipped with tungsten filament) operating at 80 kV, and recording images at 40,000× nominal magnification. (Paramyosin filament images were used as a calibration standard). Micrograph images were recorded on Kodak S0-183 film (Eastman Kodak, Rochester, USA) using a 1 s exposure time. Films were developed in-house. Dr Neil Billington (NB) (LMP-NIH) carried out all EM at LMP-NIH, using a JEOL 1200EX II microscope operating at 80 kV, recording images at 40,000× nominal magnification via a AMT XR-60 CCD camera. (Catalase crystals were used as a calibration standard).

To prepare actomyosin 5a-HMM grids, a procedure for mixing components was followed. First, equal volumes of suitably diluted myoin 5a-HMM and ATP were mixed (typically yielding a ~20 μL total volume). A volume of the myosin 5a-HMM-ATP mix was then added to a suitably dilute equal volume of F-actin and mixed. 5 μL of the actomyosin 5a-HMM-ATP mix was then transferred to an EM grid, and the sample droplet stained as quickly as possible thereafter (~10-20 s from sample mixing to stain application).

5.3.5.4   NADH-coupled ATPase assays

Actin-activated MgATPase assays were performed using the NADH coupled assay methods described in 2.7. The final composition of the assay buffer was: 10 mM MOPS (pH 7.0), 2 mM $MgCl_2$, 0.15 mM EGTA, 40 U/mL *l*-lactate dehydrogenase, 200 U/mL pyruvate kinase, 200 μM NADH and 1 mM phospho(enol)pyruvate. Final KCl concentrations for different samples are stated in results. 1 mM ATP concentration was used and the assay run at 25°C unless stated otherwise.

5.3.5.5   Dual-colour TIRF assay

The set-up and methods used for this assay are described in full in 6.6.1-6.6.2. Any specific sample conditions are referred to in the results section.

5.4 Results

5.4.1 Mutagenesis and cloning

5.4.1.1 Construction of wild-type HMM clone

Construction of the wild HMM-WT clone by PCR was successful. Images of DNA electrophoresis in Figure 5.13 confirm the steps in the cloning process. These are: Figure 5.13(a), successful PCR derivation of the HMM-WT-sized sequence from a myosin 5a-HMM-Caldesmon source plasmid, Figure 5.13(b) successful *Spe*I restriction digests of target plasmid and HMM-WT insert, Figure 5.13(c) shrimp alkaline phosphatase-treated DNA ahead of ligation reaction and Figure 5.13(d)-(e), successful HMM-WT ligation into target plasmid, transformation to *E. coli* and amplification/purification of the new clone.
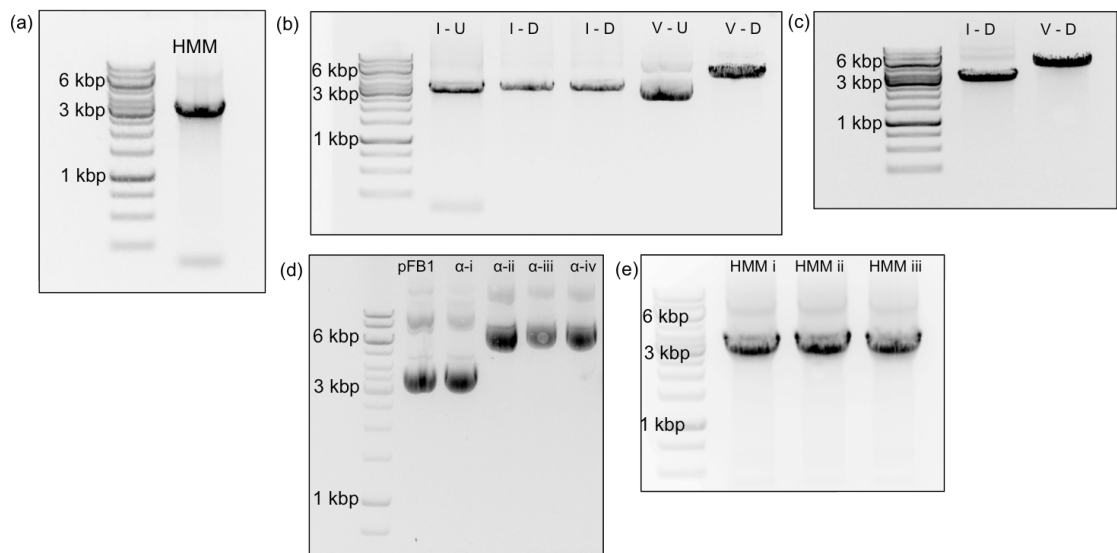
Figure 5.13: DNA electrophoresis gels confirming construction of HMM-WT clone.

NB: Original greyscale of gel images is inverted for ease of view, such that DNA bands appear dark on a pale background. kbp = kilo base pairs (1,000 bp).

(a) 'HMM' lane is linear DNA for HMM-WT, PCR-derived from a myosin 5a-HMM-Caldesmon chimera template. HMM-WT runs at the expected size of ~3.3 kbp. **(b)** HMM insert DNA (I) and target pFastBac™1 vector (V) following *Spe*I digest. I-U is undigested insert, I-D are digested inserts, V-U is undigested pFastBac™1 (circular DNA running as doublet) and V-D is linear digested insert running as single band. **(c)** HMM Insert (I-D) and target vector (V-D) following shrimp alkaline phosphatase treatment **(d)** Purified recombinant plasmid DNA, HMM-WT inserted in pFastBac™1, following mini-prep purifications from different cell cultures. pFB1 is empty pFastBac™1 plasmid, α-i-iv are plasmid purifications from DH5α cell cultures. α-i is negative for the insert (running at same size as pFB1 control) while α-ii-iv are positive for insert. α-ii-iv were selected for further use. All samples are circular DNA running with doublet bands. **(e)** linear HMM-WT DNA (at expected 3.3 kbp size) PCR-amplified from the DH5α plasmids α-ii-iv in (d), α-iii was selected as the final HMM-WT clone and provided the template for construction of all lever mutants.

DNA sequencing results confirmed that the HMM-WT clone chosen for use correctly followed the entire oligonucleotide sequence expected.

HMM-WT DNA was successfully transposed into the bacmid (the plasmid containing genomic baculovirus DNA) as indicated by the ~5.6 kb PCR product (formed from 2.3 kbp of bacmid DNA plus 3.3 kbp HMM-WT insert) seen in lanes 2 and 3 of the gel in Figure 5.14.
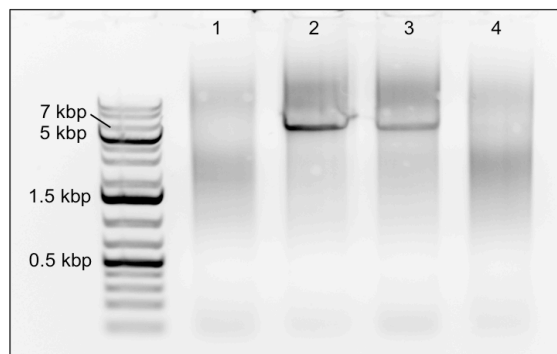
Figure 5.14: DNA electrophoresis gel of a PCR screen for the presence of HMM-WT insertion in bacmids.

Lanes 2 and 3 show DNA bands running between 5-7 kbp which are indications of positive transposition of the HMM-WT insert. A PCR product of ~5.6 kbp (formed from 2.3 kbp bacmid DNA plus 3.3 kbp HMM-WT insert) is expected for positive transposition.

5.4.1.2 Construction of lever mutants

As explained above, two alternative forms of mutation strategy (double a/a insertion or deletion and native IQ-bridge splicing) were conceived for both All-23 and All-25 lever constructs, making a total of four unique sequences to produce: All-23-Delete, All-25-Insert, All-23-Bridge and All-25-Bridge. Since two cloning methods (stepwise mutations or In-Fusion reaction) were attempted in parallel for each construct, a total of eight reaction pathways were followed. Moreover, since the stepwise method involved sequential reactions, intermediate lever constructs (*e.g.* with only one 23 to 25 a/a change) were also naturally generated. The outcome of cloning (and expression) attempts for the many constructs is summarised in Table 5.6 below. (*NB*: Table 5.5 summarises who carried out the work for each step).

Table 5.6: Summary of lever construct cloning results.

Construct	Cloning method	Cloning to pFastBac-1	Bacmid production	Baculovirus & expression	Purification
All-23-Delete	*In-Fusion*:	✓	✓	✓	✓
	Step-wise:	✓	–	–	–
All-25-Insert	*In-Fusion*:	–	–	–	–
	Step-wise:	✓	✓	✓	✓
All-23-Bridge	*In-Fusion*:	✓	✓	✓	✓
	Step-wise:	✓	✓	–	–
All-25-Bridge	*In-Fusion*:	–	–	–	–
	Step-wise:	–	–	–	–

As Table 5.6 indicates, recombinant pFastBac™1 DNA for all the lever constructs (and related intermediates), with the exception of All-25-Bridge, was successfully produced, thanks to collaborative work by AN (LMP-NIH). Despite repeated attempts, the All-25-Bridge mutant could not be produced by either the multiple-insert In-Fusion reaction or by stepwise mutagenesis. Since the DNA sequence for All-25-Bridge (as produced by step-wise mutagenesis) would contain four close-proximity repeats of the same 14 a/a IQ-bridge nucleotide sequence (one native and three spliced-in bridges), this may account for the why the construct proved too difficult to clone. In DNA sequencing results for several reaction attempts (by step-wise mutagenesis) it was found that the lever sequence often skipped from IQ 1 or IQ 2 straight to IQ 6. This may be due to mis-annealing of primers across looped sections of similar-sequence template DNA during PCR.

It should also be mentioned that, despite repeated attempts, the ~950 bp lever-tail 'minigene' sequence for the All-25-Insert construct, could not be successfully synthesized by IDT Ltd. (attempts always resulted in base pair deletions or mutations), so an In-Fusion reaction attempt was never possible for this construct.

Also noticeable from Table 5.6, is that the multi-insert In-Fusion reactions, using the pre-synthesized lever-tail minigenes, were ultimately only successful in producing one of the final constructs, All-23-Delete. Mostly it was found that the transformed In-Fusion reactions produced few bacterial colonies on LB-agar-ampicillin plates, and recombinant plasmid DNA recovered from these was usually negative for the required insert sequence. Step-wise PCR mutation of the IQ bridge sequences ultimately proved the faster method of clone production.

Of the four possible All-23 or All-25 constructs sequences, three were successfully cloned: All-23-Delete (by In-Fusion method), All-23-Bridge (by step-wise method), All-25-Insert (by step-wise method). Specific cloning results for these mutants are detailed in sections

5.4.1.2.1 and 5.4.1.2.2 below. For each construct, DNA sequencing results confirmed the full sequences of the clones in the pFastBac™1 plasmids.

5.4.1.2.1 All-23-Delete by In-Fusion reaction

Figure 5.15 shows DNA electrophoresis gels confirming the cloning stages for All-23-Delete, the only lever mutant produced by In-Fusion reaction. These steps are: Figure 5.15(a), PCR-amplified motor and lever fragments (from HMM-WT-pFastBac™1 and IDTdna plasmids), (b) PCR-screen of DH5α colonies (using primers for motor start and lever end) to amplify joined motor plus lever-tail fragments. (c), purified recombinant All-23-Delete-pFastBac™1 DNA and (d), PCR check on positive All-23-Delete plasmid.
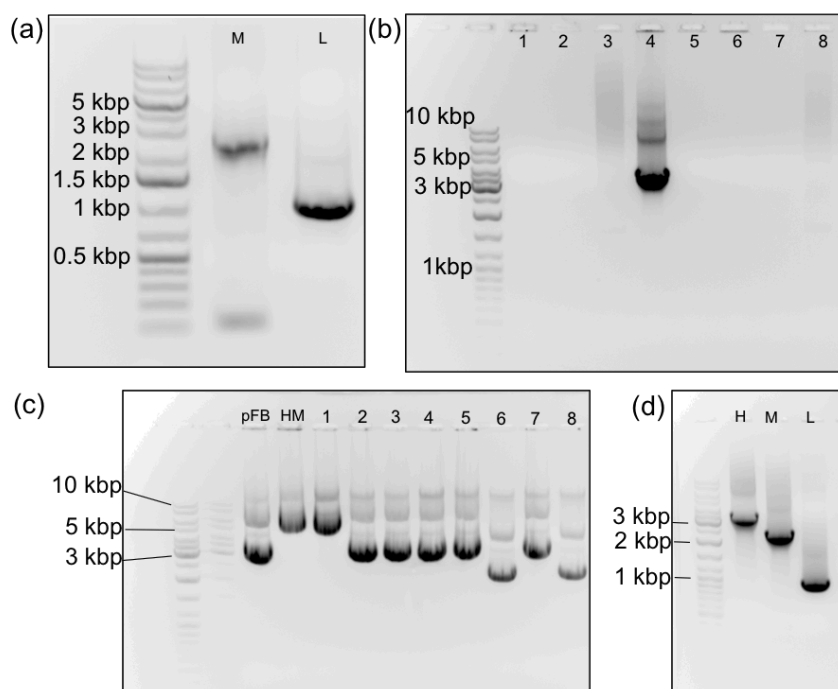
Figure 5.15: DNA gel images confirming steps in the production of All-23-Delete.

(a) linear DNA fragments prepared for In-Fusion reaction. 'M' is a motor domain sequence derived using PCR from the HMM-WT-pFastBac™1 plasmid, at expected size ~2.3 kbp. 'L' is 942 bp lever-tail sequence (PCR-amplified from IDT delivery plasmid) to be joined to the motor sequence in a multi-insert In-Fusion reaction. **(b)** In-Fusion reaction result. PCR-screen to amplify DNA from 8 bacterial colonies (lanes 1-8) containing recombinant All-23-Delete-pFastBac™1 plasmids. Lane 4 is positive, showing an intense band at ~3.2 kbp corresponding to the, now joined, motor plus lever-tail fragments from (a). **(c)** lanes 1-8 are purified recombinant All-23-Delete-pFastBac™1 plasmids (circular DNA running as two bands) from DH5α bacterial colonies. Lane 'pFB' is empty pFastBac™1 and 'HM' is HMM-WT-pFastBact™1 as control comparisons. Lane 1 is derived from the same positive colony as screened by PCR in (b) and runs, as expected, like the HMM-WT-pFastBac™1 control. All other lanes are negative for the All-23-Delete insert (mostly running like empty pFastBac™1). **(d)** final PCR-check on positive All-23-Delete plasmid DNA from (c) lane 1. 'H' is HMM, 'M' is motor domain and 'L' is lever domain linear DNA, each PCR-amplified from the same All-23-Delete plasmid source. The fragments appear at the expected 3.2, 2.3, 0.9 kbp sizes respectively.

Figure 5.16 shows the successful transposition of the All-23-Delete construct to the bacmid, work done by Dr Attila Nagy (LMP-NIH).
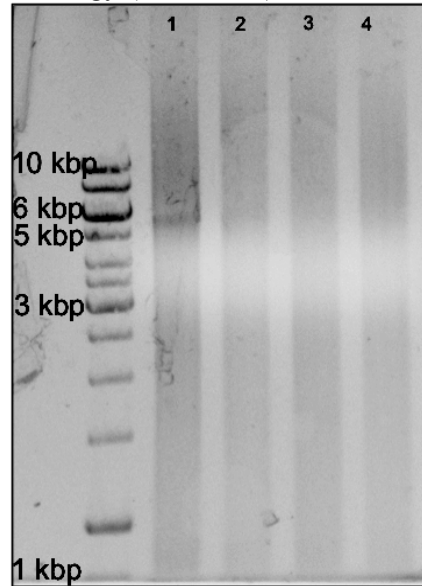
Figure 5.16: DNA gel of PCR-screen for All-23-Delete transposition to bacmid.

Work performed-by and image courtesy-of Dr Attila Nagy (LMP-NIH). Lanes 1-4 are PCR-screens of purified DNA from four DH10Bac cell cultures, to search for the presence of an All-23-Delete-sized insert in the bacmid. A faint band is visible in lane 1 (between 5 and 6 kbp) corresponding to the expected size of 5.6 kbp (2.3 kbp bacmid DNA plus 3.3 kbp for All-23-Delete construct). This is a positive indication that All-23-Delete DNA was successfully transposed for this sample.

5.4.1.2.2   All-23-Bridge and All-25-Insert by stepwise mutagenesis

Similar DNA gel images to those in Figure 5.15 and Figure 5.16 for All-23-Delete confirm the final cloning steps in the construction of All-23-Bridge and All-25-Insert by step-wise mutagenesis method, see Figure 5.17. The final rounds of PCR mutation and transposition to bacmid were carried out by Dr Attila Nagy (LMP-NIH).

In summary, the cloning work lead to successful construction of three lever mutants; two forms of an All-23 construct, All-23-Delete and All-23-Bridge, and one form of All-25 construct, All-25-Insert. These constructs, possessing the most perturbed IQ spacing patterns (rather than the intermediate constructs) were the main focus of expression, purification and characterisation work.

(a)

(b)
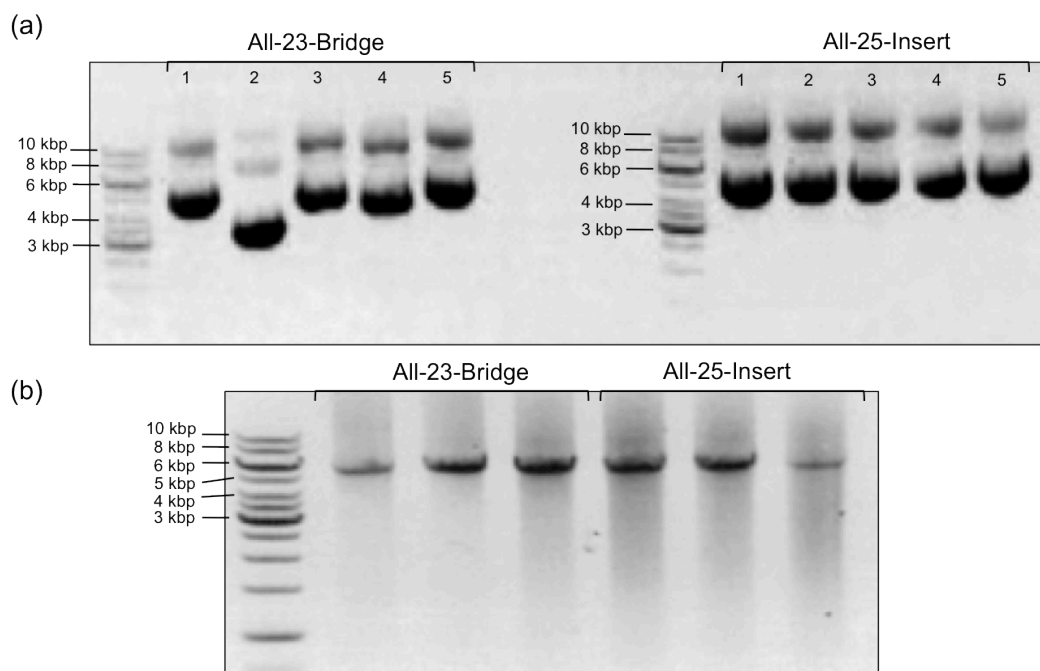
Figure 5.17: DNA gels confirming cloning and transposition to bacmid of All-23-Bridge and All-25-Insert constructs.

Work and images courtesy of Dr Attila Nagy (LMP-NIH)

(a) Purified recombinant pFastBac™1 plasmids (circular DNA running as two bands) from DH5α bacterial cultures, containing All-23-Bridge, left five lanes, and All-25-Insert, right five lanes. All samples were verified positive by DNA-sequencing, except for lane 2 All-23-Bridge DNA, which on the gel runs at a different rate to the other samples (because it lacks or contains only a partial insert). (b) PCR control screen of bacmid samples for All-23-Bridge and All-25-Insert. Bands at the expected size of ~5.6 kbp are seen for all samples. Samples corresponding the left-most lanes for each construct were used for transfection of Sf9 cells to begin baculovirus amplification.

5.4.2 Protein expression and purification

Multiple expression and purification attempts of HMM-WT and the lever mutant constructs were made. All baculovirus amplification and protein expression in Sf9 cells was carried out by Dr Attila Nagy (AN) at LMP-NIH. Separate purifications of expressed proteins were shared between AN and Derek Revill (DR). DR carried out all characterisation work with some EM images being acquired by Dr Neil Billington (NB) at LMP-NIH (instances of this are credited where appropriate).

In general, the purification procedure was successful in isolating myosin 5a-HMM molecules. However, SDS-PAGE analysis showed a common feature of most of the purifications was the presence of specific degradation fragments or contaminants as well as the expected bands for the myosin 5a-HMM heavy chain (~129 kDa) and CaM (~17 kDa), Figure 5.18.
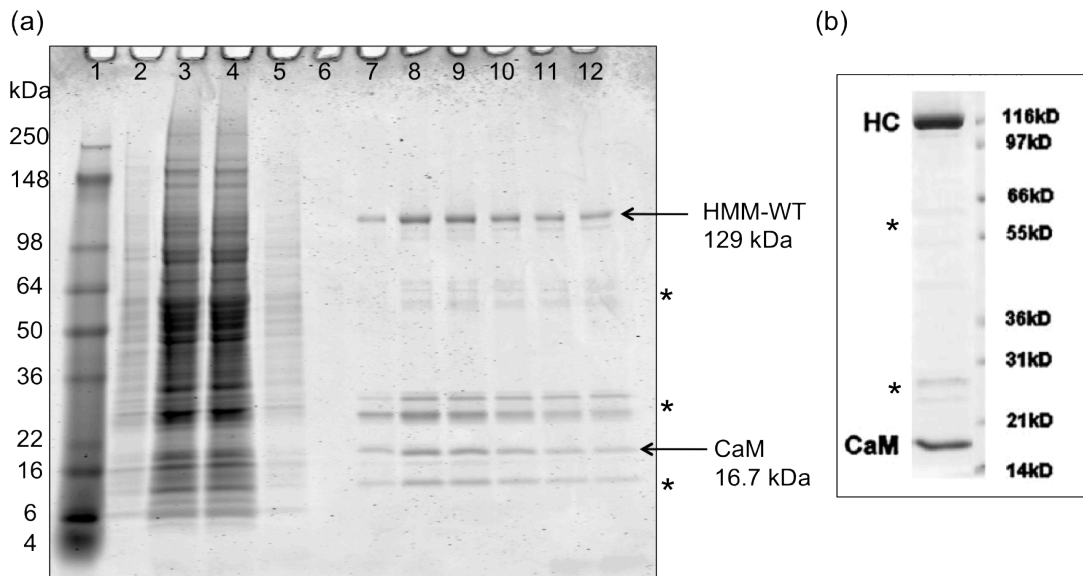
Figure 5.18: SDS-PAGE of HMM-WT construct.

(a) SDS-PAGE of HMM-WT purification by FLAG tag. Lanes 7-12 are samples from eluted fractions (prior to final dialysis) following FLAG peptide addition. Bands corresponding to the HMM heavy chain (HC) and CaM are evident. So too are additional bands, one at ~10 kDa and two between 22-32 kDa (marked with *) representing contaminants or proteolytically-produced fragments. Lanes 2-6 are various samples taken during the purification process. Lane 1: MW marker (See Blue plus 2, Invitrogen), 2: dilute sample of crude Sf9 cell lysate, 3: lysate supernatant following centrifugation, 4-5: supernatant from lysate-anti-FLAG resin wash centrifugations, 6: first wash of anti-FLAG resin column. **(b)** SDS-PAGE of an expressed and purified myosin 5a-HMM construct (identical to HMM-WT) from Wang et al. (2000). Note the pattern of similar-sized additional bands (marked by *) to HC and CaM bands as in (a).

The pattern of additional bands was consistent across preparations and the same for HMM-WT as the different lever constructs, suggesting a common rather than random degradation process or source of contamination in the purifications. In particular, three usually faint bands appeared on the SDS gels between 55-65 kDa and two (or possibly an unresolved series) of usually more intense bands were seen migrating between 22-32 kDa. On some of the gels, a band at ~10 kDa was also apparent. The same or very similar additional band patterns (albeit relatively less intense) were seen in SDS-PAGE results from the earliest recombinant myosin 5a-HMM purifications made by Wang et al. (2000), see Figure 5.18(b), and Moore et al. (2001).

The cause of the 55-65 kDa and 10 kDa bands remains unclear, but a western blot using the Anti-FLAG antibody, carried out on All-23-Bridge and All-23-Delete samples, indicates that the 22-32 kDa bands include the FLAG tag, and so must be myosin fragments that encompass the C-terminus tail of the molecule where the FLAG epitope is, see Figure 5.19. (*NB*: the FLAG peptide itself is much smaller, at ~1 kDa, and, if present, would not be seen on either the SDS gels or western blot.)
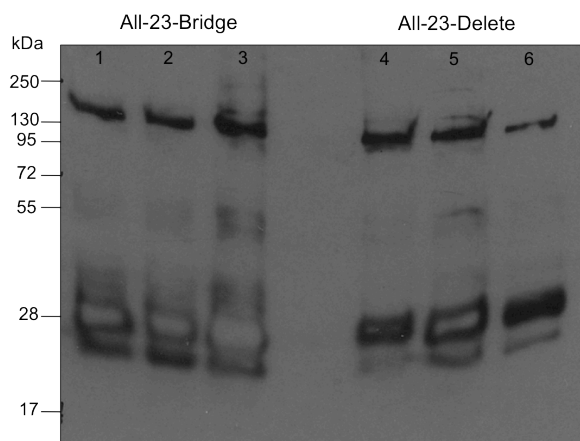
Figure 5.19: Anti-FLAG western blot of purified All-23-Bridge and All-23-Delete constructs.

A 4-20% SDS-polyacrylamide gel was used to create the blot. Lanes 1-3 is purified All-23-Bridge construct loaded in increasing amounts (lane 1: ~ 10 μg protein, lane 2: ~ 20 μg, lane 3: ~ 30 μg). Lanes 4-6 are purified All-23-Delete construct loaded in increasing amounts (lane 4: ~ 10 μg protein, lane 2: ~ 20 μg, lane 3: ~ 30 μg). The blot probes for the presence of the FLAG epitope using an anti-FLAG antibody. Note that the 'letter box' bands with pale centres (owing to substrate depletion) at ~28 kDa indicate very high intensity, *i.e.* a high concentration of FLAG. These low MW FLAG-tagged bands are more intense than the HMM heavy chain bands at ~129 kDa. They are probably proteolytically-produced lever-tail fragments, see main text.

An explanation for the presence of these 22-32 kDa FLAG-tagged tail fragments across preparations is proteolysis of the HMM molecules, such that one or both of the heads are cleaved from the coiled-coil tail. Wang et al. (2000) showed that during Sf9 cell expression, one (or presumably both) of the heads in a double-headed myosin 5a molecule could be readily cleaved close to the head-tail junction by proteases, leaving either one head and its tail still coupled with the tail fragment from the cleaved head, a species known as single-headed HMM (shHMM), or, if both heads were cleaved, just the coiled-coil tail fragment (Wang et al., 2000). On FLAG-tag affinity purification, any untagged cleaved heads are removed, while the FLAG-tagged shHMM species and any FLAG-tagged tail fragments (with MW of ~25 kDa per HC) will remain.

From the molecular mass of the 22-32 kDa gel bands in Figure 5.18 and Figure 5.19, it can be predicted that the head-tail proteolysis is actually likely to be occurring within the lever domain, perhaps at one or more of the hydrophobic IQ motifs if they have become unprotected by CaMs after expression. Based on the amino acid lever-tail sequence for the HMM construct, the following fragment masses are predicted for cleavage at the different IQ motifs.

IQ1…FLAG:	40.3 kDa	IQ4…FLAG:	31.5 kDa
IQ2…FLAG:	37.5 kDa	IQ5…FLAG:	28.3 kDa
IQ3…FLAG:	35.0 kDa	IQ6…FLAG:	25.5 kDa

The pair (or unresolved series) of FLAG-tagged fragments seen migrating between 22-32 kDa on the electrophoresis gels therefore indicates that the majority of proteolysis is occurring around IQs 4-6, in the C-terminus half of the lever.

An SDS gel run for a full length myosin 5a purification (protein cloned and expressed by Dr James Sellers group (LMP-NIH), purified by Dr Sellers at University of Leeds) further supports this lever proteolysis theory. For the SDS gel of the full length molecule, Figure 5.20, there are no bands between 22-32 kDa, but instead there are three bands migrating between 95 and 130 kDa. These would correspond to the same proteolytically-produced fragments as in the HMM case, but with the expected greater mass due to the longer full length tail (expected size of lever IQs to C-terminus for full length molecule is ~106-120 kDa).
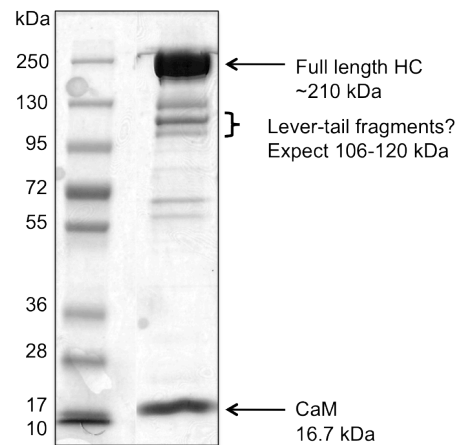
Figure 5.20: SDS-PAGE of a purification of full length myosin 5a.
Lane 1: MW standard, lane 2: purified protein sample.

In the case of the HMM SDS-gels in particular, the relative intensity of the lever-tail bands and the very intense 'letter box' reaction in this region on the western blot, indicate that many of the molecules have actually lost both heads. The intensity of the coomassie-stained lever-tail bands and intense western blot reaction is too high to be just the partner coiled-coils of shHMMs. If this were so, the band intensity would be expected to be roughly 1/5th that of the intact heavy chain (based on the ratio of masses) but gel densitometry measurements (section 5.4.2.1 below) showed that the lever-tail bands are frequently as, or indeed more intense than the HC band. On this basis, one could estimate that for every intact HMM molecule there could be five or more coiled-coil tail fragments (or for every shHMM molecule there could be two or more coiled-coil tails). Negatively stained EM images of purified samples showed that in reality there was usually a mixture of both single and double-headed species in the preparations, see Figure 5.21. The coiled-coil lever-tail fragments are unlikely to be visible by EM (the coiled-coil tail is rarely seen in negative stain EM images of myosin 5a-HMM), but, from the SDS-PAGE, it can be

assumed that they are present in the purified samples and likely to be at least twice as numerous as the HMM/shHMM molecules themselves.
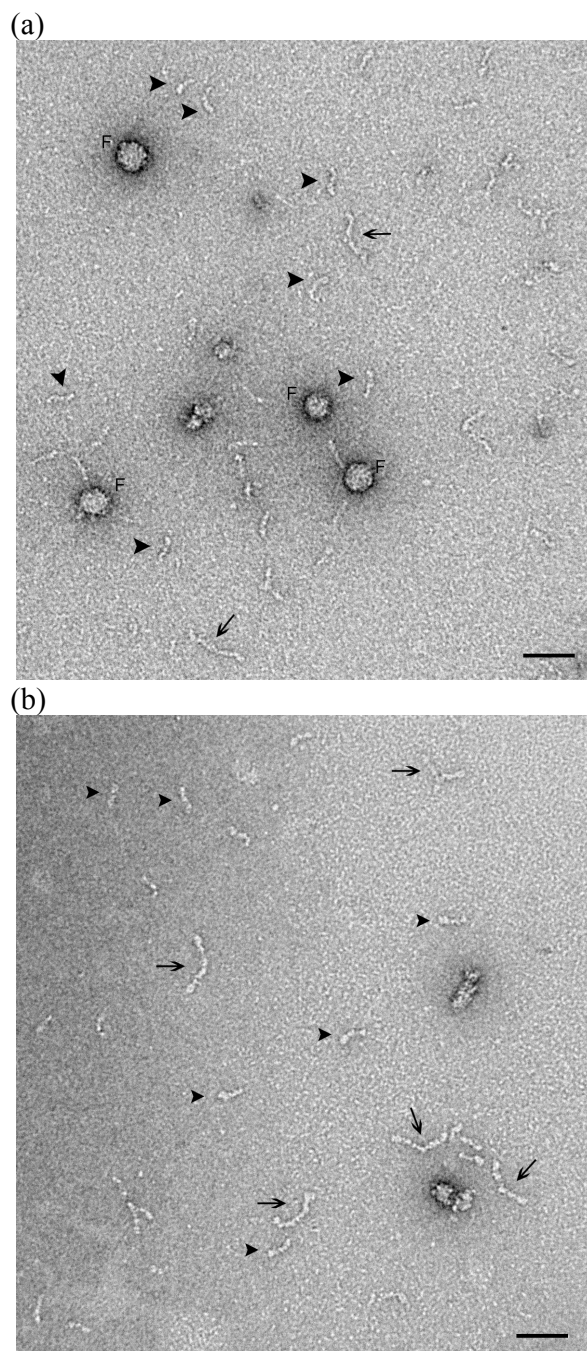
Figure 5.21: Negative stain EM images of purified HMM-WT and All-23-Bridge proteins. Images acquired by NB (LMP-NIH). Scale bars are 50 nm. **(a)** field of free HMM-WT molecules following purification. Two-headed molecule examples are highlighted by arrows, but these are outnumbered by examples of single head fragments highlighted by arrowheads. The single heads appear to have a motor domain and ~ 3 or 4 CaMs bound. *NB*: the ~20 nm diameter circular objects, marked by 'F's, are thought to be contaminants originating from the batch of FLAG peptide used in purifications, see main text. **(b)** field of free All-23-Bridge molecules following purification. Two-headed molecule examples are highlighted with arrows but, similarly to the HMM-WT example, these are outnumbered by single head fragments marked by arrowheads. Grid conditions prior to staining with 1% uranyl acetate: 125 mM KCl in (a), 50 mM KCl in (b), 10 mM MOPS, 0.1 mM EGTA, 2 mM MgCl$_2$, pH 7.0.

Recognition of contamination of both HMM-WT and lever mutant preparations by lever-tail fragments and shHMM is very important, as the lever-tail fragments will contribute to (and thus tend to overestimate) HMM molecule concentration measurements by absorption spectroscopy. Note also that the short lever-tail fragments may also still bind one or more CaMs at intact IQ motifs, thus adding even more to the overestimated HMM concentration. Preparations containing mixed populations of single and double-headed molecules also pose difficulties for interpretation and analysis of actin-activated ATPase or TIRF motility assays.

The EM micrograph example in Figure 5.21(a) also reveals the presence of another type of contamination frequently encountered in the preparations made for this work, namely the appearance of unidentified circular objects with a variable diameter of ~20 nm, (see objects marked 'F' in Figure 5.21(a)). It is thought (investigative work by Dr Neil Billington (LMP-NIH)) that these circular objects originate from the source of FLAG peptide used to elute protein in the anti-FLAG affinity purification step, perhaps being aggregates of peptide or more likely (given their large size) being a contaminant of the original peptide synthesis process. Again, recognition of the presence of these relatively abundant 'FLAG contaminants' is important as, if they contain FLAG peptide or protein material, they might also add to 280 nm absorption. They are also an unwelcome distraction in EM negative stain imaging.

Attempts were made to improve the purification process to eliminate the lever-tail fragments, shHMM, and FLAG contaminants from final preparations. In earlier purifications, dialysis was the only method employed to clean-up samples following the FLAG affinity chromatography step. In later preparations, the additional Q-sepharose ion exchange column was employed, using a step increase in salt (from 0.1 to 0.5 M NaCl) to elute protein. This was apparently sufficient to separate the circular FLAG contaminants from myosin molecules, but it was still not sufficient to separate the tail fragments and shHMM species from the two-headed HMM molecules. Although not attempted due to time limitations, it is noted that other laboratories have used an ion exchange column with a salt *gradient* (0.1-0.5 M KCl) rather than step, to fractionate protein, and this may be necessary to achieve proper separation between tail fragments, single and double-headed molecules (Forgacs et al., 2008).

To limit proteolysis at the lever and formation of shHMM and tail fragments in the first place, one could consider increasing the amount of CaM baculovirus used to co-infect the Sf9 cells (a 2:1 myosin to CaM virus ratio was used in preparations here). Increased CaM concentration in the Sf9 cells might better protect the IQ motifs from proteases. However,

due to time limitations, no variation in CaM co-infection levels was made across preparations.

5.4.2.1 SDS-PAGE densitometry measurements

Because of the continued presence of the lever-tail fragments and other contaminants in the samples, even following ion-exchange chromatography and dialysis, densitometry analysis of the SDS-PAGE images was undertaken. This was done chiefly to achieve a better estimate of the actual myosin head concentrations within the samples as, for the reasons discussed above, the OD absorption measurements were almost certainly an overestimate. Having a more precise head concentration was a necessity for calculating accurate ATPase rates (section 5.4.3.3).

Band intensity measurements were carried out on scanned images of SDS gels of final protein preparations using the standard gel macro tools in *ImageJ* software (v1.43) (NIH, USA). Images were first converted to 8-bit greyscale and the background intensity was subtracted using the ImageJ 'rolling ball' algorithm. To find the myosin head concentration as a fraction of total protein in the sample, a lane was selected, the band intensity profile plotted and the integrated area under the intensity peaks corresponding to gel bands calculated, see Figure 5.22. From the peak areas, the fraction of myosin heads was estimated from the combined area of the HC plus CaM peaks, as a fraction of the total area from all peaks.

As Table 5.7 shows, the gel densitometry values indicate that the heavy chain and CaM bands taken together typically accounted for only ~40-50 % of the total protein in the samples measured, with the lever-tail fragment contributing the next largest fraction at ~35% of the total.
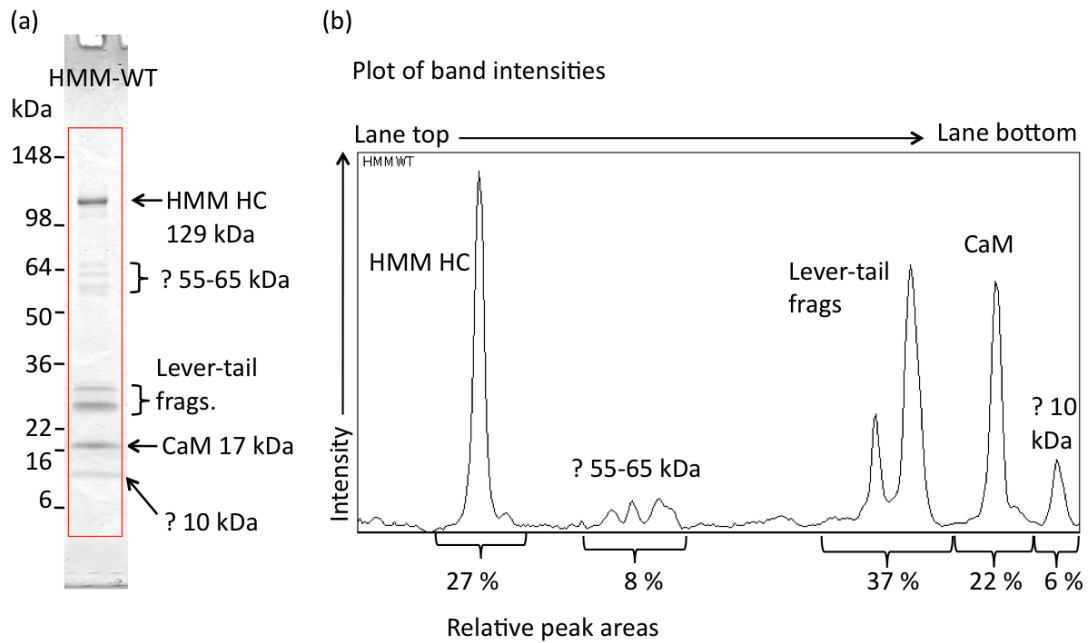
Figure 5.22: Illustration of SDS-PAGE densitometry analysis for HMM-WT sample.

Example of a gel densitometry on HMM-WT sample (March 2011 purification). **(a)** a single lane is selected (red box) from an image of an HMM-WT purified sample run on a 4-20% SDS-polyacrylamide gel. The HMM heavy chain (HC) band is identified along with CaM and the various contaminating fragments such as the lever-tail bands. **(b)** using *ImageJ* software the image intensity profile of the bands from top to bottom (reading left to right on intensity plot horizontal axis) is shown. The percentages for the integrated peak areas calculated for this example are shown.

Table 5.7: SDS-PAGE band intensity percentages for separate purifications of myosin 5a HMM constructs.

NB: Band intensity % values for the March 2011 purifications are averages over two gel analyses, Sept. 2011 values are from single gel measurements.

| | March 2011 purifications | | | Sept. 2011 purifications | |
| | Band intensities (%) | | | Band intensities (%) | |
SDS gel band(s)	HMM-WT	All-23-Bridge	All-25-Insert	HMM-WT	All23-Bridge
Heavy chain	26	26	19	27	29
CaM	17	18	0*	22	21
Lever-tail fragments	35	28	40	37	33
Other (10, 55-65 kDa)	22	28	41	14	17
Heavy chain + CaM	43	44	19	49	50

* The striking result that no CaM band is present for the All-25-Insert construct is discussed further in 5.4.4.1.

The results for the All-25-Insert mutant are exceptional among the samples analysed. Most strikingly there was no trace of a CaM-sized band on the gel (nor on similar gels for other purifications) and the percentage of lever-tail and other degradation fragments is higher (at the expense of intact heavy chain) than in the other HMM-WT and All-23-Bridge samples. The absence of CaM for this mutant is discussed in more detail in 5.4.4.1.

For the HMM-WT and All-23-Delete purifications, adjusted head concentrations were derived from the densitometry data accordingly. Previously obtained protein concentrations (from OD measurements) were simply multiplied by the heavy chain-plus-CaM fractions obtained from densitometry.

Table 5.8: Protein stock concentrations as measured by spectrophotmetry and revised by SDS-PAGE gel densitometry measurements.

		Stock conc. HMM molecules (from OD) (µM)	Myo5a plus CaM fraction (from densitometry)	Revised stock conc. HMM molecules (µM)	Myo5a head concentration (for ATPase) (µM)
March 2011 purifications	HMM-WT	1.6	0.43	0.69	1.4
	All-23-Bridge	2.0	0.44	0.88	1.8
Sept. 2011 purifications	HMM-WT	3.0	0.49	1.47	2.9
	All-23-Bridge	3.4	0.50	1.70	3.4

SDS gel densitometry was used to make one further correction (in addition to the % heavy chain-plus-CaM adjustment) to the working concentration of one of the All-23-Bridge samples (March 2011 purification). An SDS gel of theoretically equal protein mass loadings of this mutant and a corresponding HMM-WT sample showed the All-23-Bridge heavy chain band was in fact significantly less intense (~½ intensity) than the wild-type. The same concentration mismatch was also observed in an EM comparison of the same samples (Neil Billington (LMP-NIH)). The reason for the discrepancy between OD-measured protein concentration and apparent concentration in SDS-PAGE and EM was not clear, though one possible reason might have been the presence of more un-removed FLAG peptide in the mutant sample. Accordingly, for the actin-activated ATPase rate calculations in 5.4.3.3, it was decided to scale down the All-23-Bridge head concentration relative to HMM-WT, by the ratio of the heavy chain band intensities measured from the comparative SDS gel (All-23-Bridge Intensity/HMM-WT intensity = 0.57). For purposes of the ATPase assay, 5.4.3.3, the working head concentration for the March 2011 All-23-Bridge construct was altered from 3.4 µM (already adjusted for heavy chain-plus-CaM content) to 3.4 µM × 0.57 = 1.9 µM.

5.4.2.2 All-23-Delete mutant

The three fully-mutated lever constructs produced (All-23-Delete, All-23-Bridge, All-25-Insert) were all expressed and purified at least once. However, results for the first preparation of All-23-Delete as compared to All-23-Bridge were disappointing. Although All-23-Delete heavy chain and CaM bands were present on SDS-PAGE, actin-activated ATPase was low compared to All-23-Bridge, and it proved very difficult to find any examples of intact two-headed molecules either free, or bound to F-actin by EM imaging.

For this reason, it was decided to concentrate on the more promising All-23-Bridge and All-25-Insert mutants. Results of initial characterisations of these constructs are described in the following sections.

5.4.3 Characterisation of All-23-Bridge lever mutant

5.4.3.1 Expression and purification

The All-23-Bridge mutant was expressed and purified four times, including one version of the construct (cloned by Dr Attila Nagy (LMP-NIH)) with an additional Green Fluorescent Protein (GFP) sequence added to the C-terminus after the HMM tail (This version of the construct was made for potential use in fluorescence microscopy experiments).

SDS-PAGE analysis of samples from various preparations indicate that the All-23-Bridge mutant expresses and purifies as successfully as HMM-WT, see Figure 5.23 and Figure 5.29(c). Bands at comparable intensities to HMM-WT purifications are evident at ~129 kDa, the expected size for the HMM heavy chain, and also at ~17 kDa for CaM. CaM LCs co-purifiy with the myosin if they are bound to the heavy chain after expression, so presence of a CaM-sized band on the gels is a first indication that the mutations to the All-23-Bridge lever have not disrupted all CaM binding.
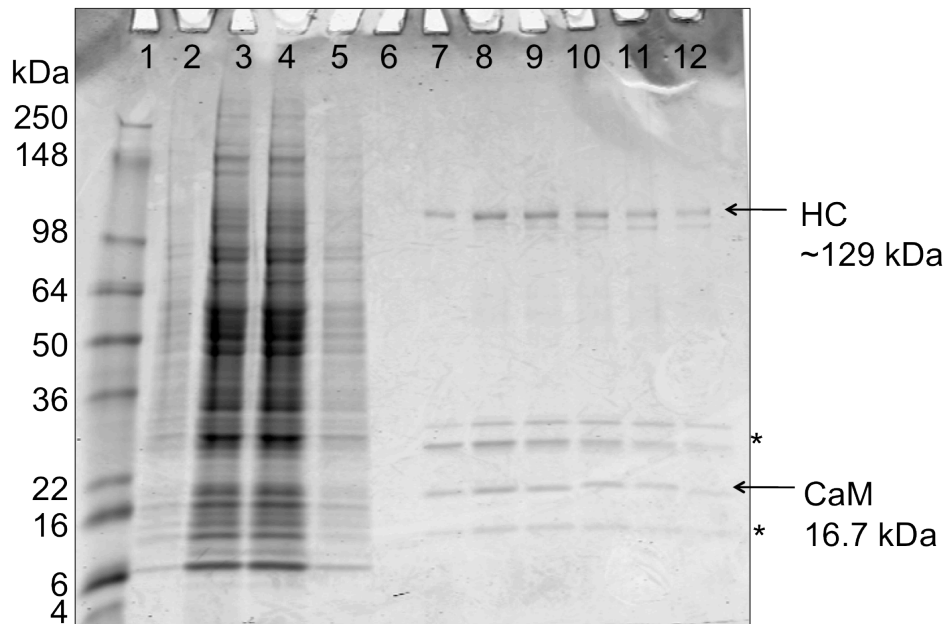
Figure 5.23: SDS-PAGE of an All-23-Bridge purification.

> Lanes 7-12 are samples from eluted fractions following anti-FLAG liquid chromatography purification (prior to final dialysis). Bands corresponding to the myosin 5a-HMM heavy chain (HC) at ~129 kDa and CaM at 16.7 kDa are evident. So too are additional bands, one at ~10 kDa and two between 22-32 kDa (marked with *) representing contaminants or proteolytically-produced fragments as in HMM-WT. Lanes 2-6 are various samples taken during the purification process. Lane 1: MW marker (See Blue plus 2, Invitrogen), 2: dilute sample of crude Sf9 cell lysate, 3: lysate supernatant following centrifugation, 4-5: supernatant from lysate-anti-FLAG resin wash centrifugations, 6: first wash of anti-FLAG resin column.

The gel in Figure 5.23 also shows a similar pattern of degradation/contamination to HMM-WT, with an unknown band at ~10 kDa, and two bands between 22-32 kDa representing proteolytically-produced lever-tail fragments. Again this is an indication that single heads would be expected to be found in samples studied by EM.

5.4.3.2 Electron microscopy

EM images of negatively-stained All-23-Bridge molecules, both free and in the presence of F-actin and low (~1 μM) ATP were captured. EM was performed on fresh and previously drop-frozen samples, and no noticeable difference was observed. Note that in the following images, the ~20 nm diameter round objects are thought to be the contaminants (originating from the FLAG peptide source) and are not related to the (acto)myosin sample (see 5.4.2 and Figure 5.21).

5.4.3.2.1   Appearance of Free Molecules
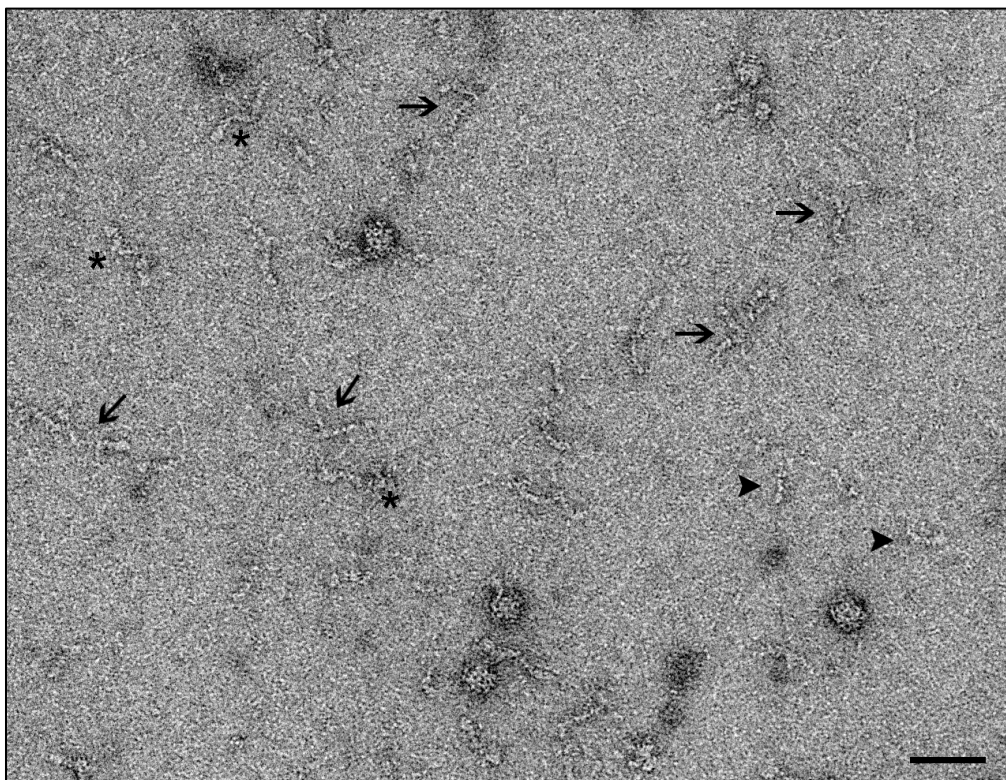
Figure 5.24: Field of free, negatively-stained All-23-Bridge molecules without ATP.

Examples of V-shaped two-headed molecules with an appearance similar to HMM-WT are highlighted with arrows. Two CaM-bound levers join at the head-tail junction. Bulbous tips on the ends of the levers are the motor domains. Molecules marked with asterisk exhibit levers with sharp kinks (a feature also common to wild-type). Arrowheads mark examples of single head fragments. Scale bar 50 nm. Grid conditions prior to staining: 75 mM KCl, 10 mM MOPS, 2 mM EGTA, 1 mM MgCl$_2$, 2 mM K-phosphate, pH 7.0. 1% uranyl acetate stain. See also Figure 5.21(b).

The field of free molecules seen in Figure 5.24 show mainly two-headed molecules with an overall V-shape (see arrowed examples), an appearance very similar to that observed

for HMM-WT. See also Figure 5.21 (b). Motor domains appear as bulbous tips on the end of the two levers that join together at the head-tail junction. No coiled-coil tail is visible. Head-head angles vary from ~20-180°. Levers are visible with a lumpy appearance because they have CaMs bound, apparently along their entire length, and they vary from being quite straight (*e.g.* arrowed) to having sharp kinks in them (*e.g.* asterisks). These features have all been observed for wild-type HMM molecules previously (Burgess et al., 2002) (3.2 in this thesis). As well as two-headed molecules, single heads with varying lengths of CaM-bound lever are seen in Figure 5.24 and Figure 5.21(b) (arrowheads). This is consistent with the expectations of a shHMM species from SDS-PAGE analysis.

5.4.3.2.2   Appearance with F-actin

Images of All-23-Bridge molecules (at ~ 0.5 μM concentration) in the presence of F-actin and ATP (at ~ 0.5-1.0 μM) were obtained using the methods in 5.3.5.3. These conditions have previously produced examples of HMM-WT molecules bound by one or both motor domains to actin filaments, most frequently at a motor-motor spacing of 36 nm (13 actin subunits) (with less frequent examples also at 11 and 15 subunit spacings) (Oke et al., 2010; Walker et al., 2000). Example images of the All-23-Bridge construct in similar conditions are shown in Figure 5.25.

As for the free molecules, the acto-myosin 5a images are generally consistent with observations of wild-type HMM. Numerous examples of molecules forming a V-shape, with both motors bound to a single filament were seen, Figure 5.25(c). Instances of molecules crosslinking two filaments (Figure 5.25(b)) or binding by only one motor (Figure 5.25(a),(d)) were also recorded. In the latter case, the molecules generally adopt an opened-out conformation (*i.e.* with head-head angle ~180°) with the unattached motor some distance from the filament. By inspection, motor domains for doubly-bound molecules appear, like HMM-WT, to bind at around a ~36 nm spacing (scale bar in (c)-(d)), the 13-subunit pseudorepeat of the F-actin; though there is clearly some variation and precise distance measurements for a large number of molecules remain to be made.

(a)

(b)

(c)

(d)
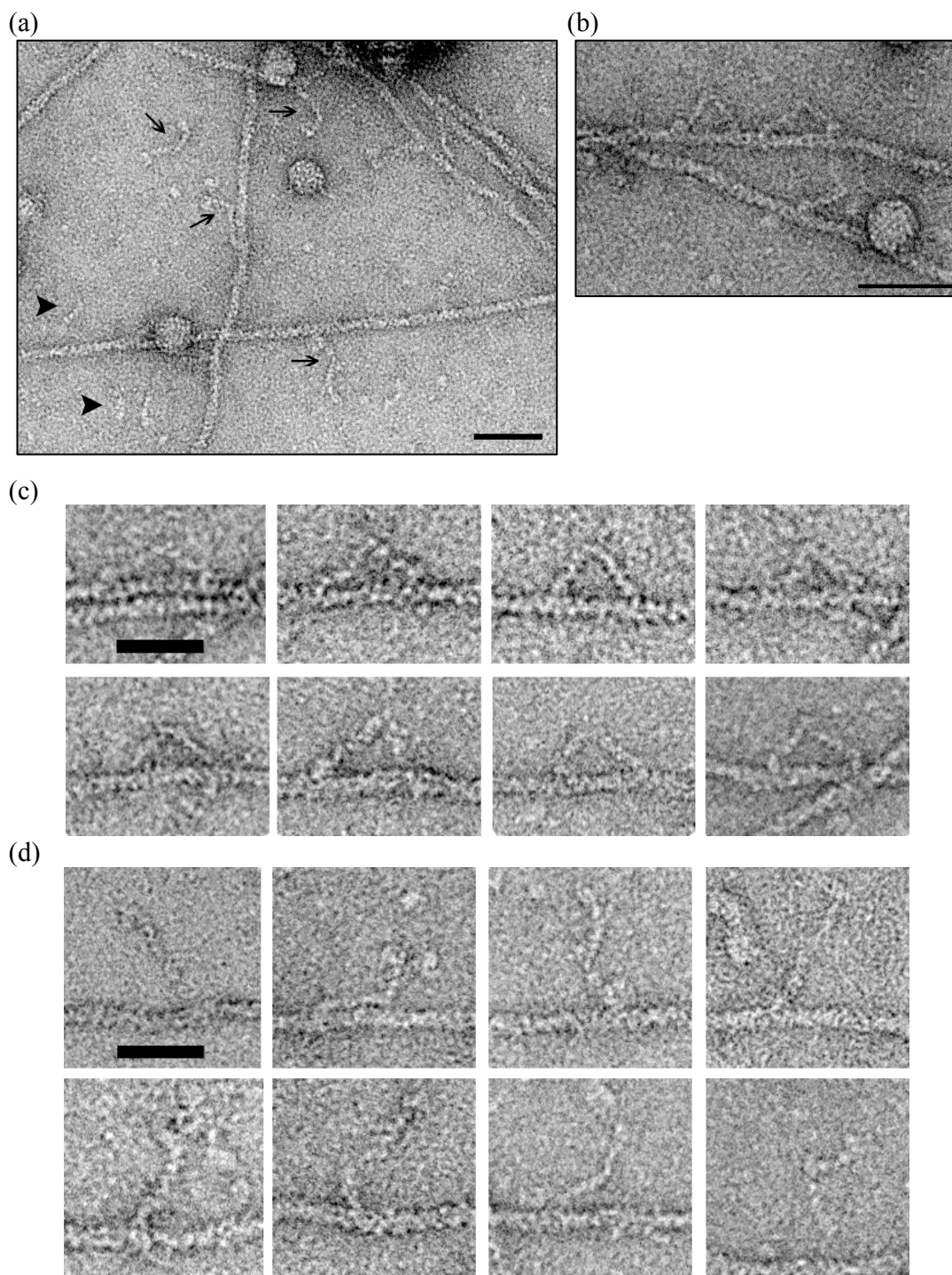
Figure 5.25: Negatively-stained EM images of All-23-Bridge molecules in the presence of F-actin and ATP.

In the images shown, HMM molecules are at ~0.5 μM, actin at ~1 μM and ATP at ~0.5 μM. Grid conditions at point of staining: 75 mM KCl, 10 mM MOPS, 2 mM EGTA, 1 mM MgCl$_2$, 2 mM K-phosphate, pH 7.0. 1% uranyl acetate stain. Scale bar in (a) and (b) is 50 nm, and is 36 nm in (c) and (d). **(a)** field of molecules (two-headed arrowed or single-headed with arrowhead) next-to or bound to F-actin. In this example, only molecules bound by one head are seen. **(b)** trio of HMM molecules with both heads bound. Note the apparently different motor-motor spacing of the adjacent molecules above the actin filament. The third molecule is cross-linking two filaments. **(c)** gallery of molecules bound to F-actin by both motor domains or in **(d),** by one motor only.

Levers in the actin-bound All-23-Bridge molecules appear to bind CaMs along their entire length and, on first inspection, no unusual structural features are discernable from the individual molecule images. One very tentative observation from surveying the doubly-bound molecules (n ~100) is the reoccurrence of a slight inward buckle (towards the actin filament) in one or both of the levers, about 1/3rd of the way along its length *e.g.* Figure 5.25(c) 2nd and 4th images from the left (left levers). However, a larger dataset, and multiple samples must be surveyed before assigning significance to this. For the singly-attached images, it is noticeable that the levers appear largely straight and that the unattached head is positioned far from the actin filament. Again, the data set is too small to speculate if this is by chance, or whether it is due to an intrinsic property of the mutated molecules (*e.g.* a more rigid lever domain) or the manner in which the unattached heads and actin filament tend to adsorb to the carbon substrate.

5.4.3.3 Actin-activated MgATPase

To measure the relative catalytic activity of the All-23-Bridge mutant, and thus gain a potential insight into its processivity on F-actin, the steady-state actin-activated MgATPase rates for separate preparations of both All-23-Bridge and HMM-WT were compared, in two NADH-coupled assays (see 2.7 and 5.3.5.4). In the first assay (March 2011), ATP was at 1 mM final concentration, KCl ~10 mM and the assay temperature ~23 °C (laboratory temperature), though the latter two variables were not well controlled. In the second assay (September 2011), ATP was at 1 mM, KCl at 50 mM and the temperature controlled to be 25 °C. The calculated MgATPase rates V, at varying actin concentrations, are plotted in the graphs in Figure 5.26. *NB*: the rates stated are per myosin head (*i.e.* per active site, two per HMM molecule), and the head concentrations used in the calculations were those obtained following the SDS-PAGE densitometry corrections to OD measurements (Table 5.8), with an additional correction made to the All-23-Bridge concentration (March 2011 purification) following SDS-PAGE comparison of heavy chain band intensities (5.4.2.1).

From fits of the Michaelis-Menton equation ($V = V_{max}\cdot$[actin]/(K_{ATPase} + [actin])) to the points corresponding to each protein, values for V_{max}, the maximal activity and K_{ATPase}, the actin concentration at half maximum rate were obtained. These are given along with the measured myosin-only basal rates (no actin present) in Table 5.9. For comparison, equivalent published parameters for myosin 5a HMM-WT (obtained by NADH-coupled assays in similar conditions) are given in Table 5.10.

(a)

(b)
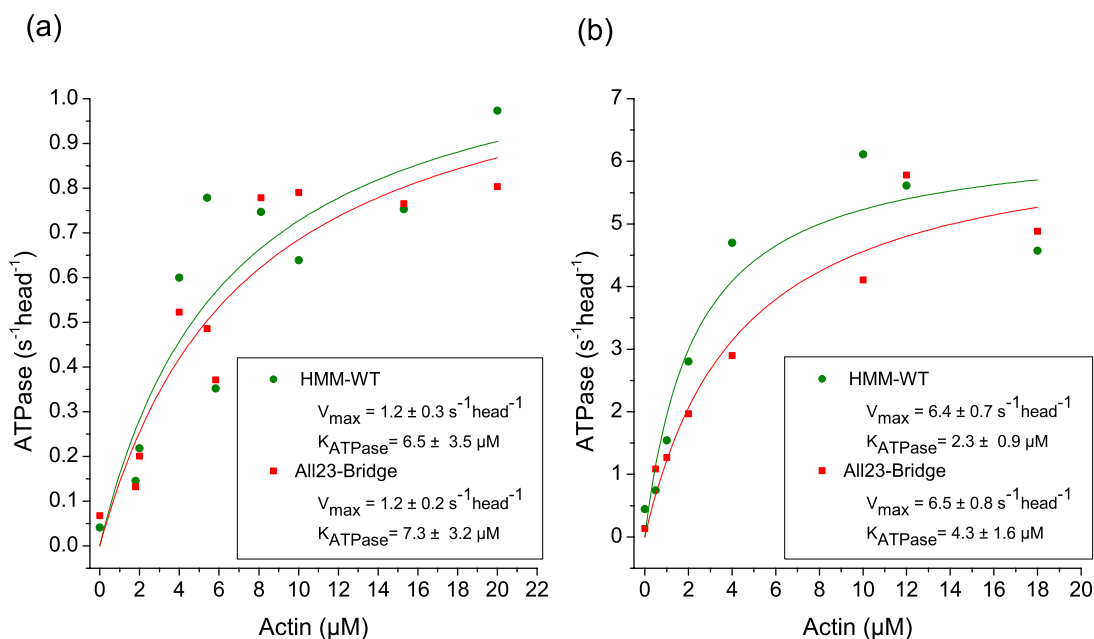
Figure 5.26: Steady-state actin-activated MgATPase rates for HMM-WT and All-23-Bridge constructs.

ATPase rates for March 2011 protein preparations shown in (a), September 2011 rates in (b). HMM-WT rates (green points) and All-23-Bridge rates (red points) were recorded in the both assays. Coloured lines are fits of the Michaelis-Menton equation ($V = V_{max}$[actin]$/(K_{ATPase}$ + [actin])) to the corresponding points for each construct. Values for V_{max} and K_{ATPase} (actin concentration at half-maximal rate) obtained from the fits are given in the graph legends (± SE). *NB*: active head concentrations used to calculate rates were those obtained following gel densitometry adjustments, see Table 5.8. In (a): KCl was ~10 mM final concentration, ATP 1 mM and temperature ~23 °C. In (b): KCl was 50 mM, ATP 1 mM, and temperature 25 °C. Note the different vertical axis scales and the lower rates obtained in (a). The protein purifications made for (a) were not subject to ion-exchange chromatography before dialysis, whereas those in (b) were. This may be one reason rates are lower in (a), perhaps owing to contamination, *e.g.* by FLAG peptide, contributing to the OD measurements and so inflating the apparent concentration of active molecules.

Table 5.9: Steady-state actin-activaed MgATPase parameters obtained from myosin 5a HMM-WT and All23-Bridge NADH-coupled assays.

Assay	Construct	Basal rate (s^{-1}head^{-1}) (±SD)	V_{max} (s^{-1}head^{-1}) (±SE)	K_{ATPase} (µM) (±SE)
March 2011: KCl ~10 mM Temp ~23 °C	HMM-WT	0.04 ± 0.01	1.2 ± 0.3	6.5 ± 3.5
	All23-Bridge	0.07 ± 0.02	1.2 ± 0.2	7.3 ± 3.2
Sept 2011: KCl = 50 mM Temp = 25 °C	HMM-WT	0.45	6.4 ± 0.7	2.3 ± 0.9
	All23-Bridge	0.14	6.5 ± 0.8	4.3 ± 1.6

Table 5.10: Published steady-state actin-activated MgATPase parameters for myosin 5a HMM and S1 fragments.

Molecule type	Basal rate ($s^{-1}head^{-1}$)	Vmax ($s^{-1}head^{-1}$)	K_{ATPase} (μM)	[KCl] (mM)	Reference
HMM	< 0.04	3.8 ± 0.6	0.15 ± 0.06	80	(Wang et al., 2000)
	0.021	8.2	0.55	25	(Forgacs et al., 2009)
	< 0.07	9.7 ± 1.0	0.32 ± 0.09	50	(Sakamoto et al., 2003)
S1	< 0.04	4.5 ± 0.4	4.4 ± 1.0	80	(Wang et al., 2000)
	0.027	12.4	2.6	25	(Forgacs et al., 2009)
	< 0.07	15.3 ± 4.3	3.6 ± 1.0	50	(Sakamoto et al., 2003)

All parameters listed were obtained by NADH-coupled assay at 25 °C

A number of observations can be made from the results of the two assays and comparison to literature values. Firstly, it is clear that the activated rates and V_{max} value obtained from the March 2011 assay are significantly lower (~5-fold) than those obtained in September 2011 assay, and are also lower than any published results for HMM-WT in similar conditions, see Table 5.10. However, in the September 2011 assay, the V_{max} values (HMM-WT = 6.4 ± 0.7 s^{-1}, All-23-Bridge = 6.5 ± 0.8 s^{-1}) do fall within the range of values previously reported for myosin5a HMM-WT (3.8 ± 0.6 s^{-1} – 9.7 ± 1.0 s^{-1}). Two reasons may account for the unexpectedly lower March 2011 rates. Firstly, the KCl salt concentration was not well controlled and allowed to drop to ~10 mM (or lower in some cases) in the March 2011 assay. Such low ionic strength may have caused aggregation of the myosin and/or F-actin, which would reduce the effective concentration of active protein and so the calculated rates. Secondly, and perhaps more significantly, the lower March rates may have been due to a difference in the purification of the proteins used. For the March assay, no ion-exchange chromatography step was performed following the Anti-FLAG column elution and before final dialysis, whereas this extra step was carried out for the September assay. It is possible that dialysis alone is not sufficient to remove all of the FLAG peptide used to elute protein from the Anti-FLAG column. If significant amounts of the peptide remain in the preparation, this would certainly add to OD absorbance measurements (the peptide sequence, DYKDDDDK, contains a tyrosine) and so overestimate the myosin concentration and underestimate the ATPase rates.

Despite the rate differences between the two assays, for each assay taken separately, results show no striking difference (at least within error of the experiment) in the rates or fitted curve profiles between the All-23-Bridge mutant and the corresponding HMM-WT control. Similar V_{max} values were obtained for both mutant and control in each case. One could note that the K_{ATPase} values are a little lower for All-23-Bridge than HMM-WT in

both assays, with the mutant requiring more actin substrate before its ATPase rises to V_{max}. However, there is significant relative uncertainty (~40-50 %) in the curve fits for the K_{ATPase} parameter (due to the small number and spread of data points), so this slight difference should not be over interpreted as an indicator of reduced ATPase or processivity in the mutant. Repetition of the assay with further All-23-Birdge and HMM-WT preparations would help confirm if the difference in the mutant is genuine.

For both assays, the K_{ATPase} values obtained, ranging from 2.3-7.3 µM actin, are in all cases much higher compared to published values for HMM-WT molecules (0.15-0.55 µM actin). From the published values, Table 5.10, it can be seen that high K_{ATPase} values are more typical of an S1, *i.e.* single-headed, preparation. Since both HMM-WT and mutant samples exhibited high K_{ATPase} values this suggests that the protein preparations assayed may have contained significant numbers of single rather than two-headed molecules. The presence of single heads in the preparations is consistent with the results from SDS-PAGE (Figure 5.18 & Figure 5.23), where gel bands suggested the presence of shHMM fragments, and also from EM, where single heads were seen in micrograph images directly (Figure 5.21).

Finally one can observe, as expected for myosin 5a, significant activation of the ATPase activity in the presence of actin, the basal rates increasing 17-30 fold (March 2011 assay) and 14-46 fold (September 2011 assay), although this activation is not as dramatic as reported in the literature (typically $\geq$ 400-fold increase). Curiously, the basal rates recorded in the September assay were ~10-fold higher than expected, though the rates were the result of only one assay measurement, as opposed to the March assay which are an average over 5 separate measurements.

5.4.3.4 Dual colour TIRF assay

Total internal reflection fluorescence (TIRF) microscopy assays, using two colour illumination and detection, were conducted on the All-23-Bridge construct. The main aim of the assay was to ascertain whether the altered lever mutant could still make steps along actin filaments, and, if so, what its motile characteristics were, *e.g.* run length, speed and step size.

The full set-up of the TIRF assay, plus a description of the development a new fluorophore-labelled CaM mutant for myosin 5a labelling, can be found in chapter 6. In outline, All-23-Bridge and HMM-WT molecules are first fluorescently labelled by exchanging one or more endogenous CaMs with CaMs site-specifically labelled with AlexaFluor 488® (AF488-CaM) (see 6.5.2). To conduct the assay, a flow-cell (a small chamber formed between a glass microscope slide and coverslip), containing rhodamine-

labelled actin filaments lightly tethered to one surface, is assembled. A solution containing ATP and AF488-CaM-labelled myosin 5a molecules is then introduced, allowing the myosin to bind/unbind and potentially move along the actin filaments at the flow-cell surface. The flow-cell is illuminated by two laser sources at a glancing angle to create TIRF excitation (to ~100 nm depth) at the flow-cell-sample interface. One laser source emits at a 488 nm wavelength, for specific excitation of the AF488-CaM-labelled myosin, and the other emits at 568 nm, for excitation of the rhodamine-phalloidin-labelled F-actin. Fluorescence emission from the sample viewed through a microscope objective is split into two readout channels, one filtered for actin emission, the other for myosin emission, allowing a split-image video of the surface-tethered actin and myosin from a single field of view (FoV) to be recorded. Spots of light from single myosin molecules can be seen in the upper half of the image (the myosin-FoV), and any static actin filaments in the same region are seen in lower part of the image (the actin-FoV), *e.g.* Figure 5.27. By tracking the myosin spots, one can determine if they 'land' and move along the actin filaments and what, if any, motile characteristics they have.

When examined in the TIRF assay, AF488-CaM-labelled All-23-Bridge protein was clearly visible. At high concentration (~10 nM, well above the level for single molecule resolution) a clouded appearance with darker high-intensity regions caused by fluorescence from many labelled molecules was seen in the myosin-FoV, see Figure 5.27. Although at this concentration, the myosin is saturating the surface, one can observe that the protein material is sticking both non-specifically to the flow cell surface, in regions where no actin is visible (*e.g.* white arrow Figure 5.27), but also that it most often localises strongly to the actin filament positions (*e.g.* black arrow Figure 5.27). The higher density of fluorescent myosin localised to the F-actin reveals the underlying pattern of filaments, simultaneously visible in the actin-only FoV below.
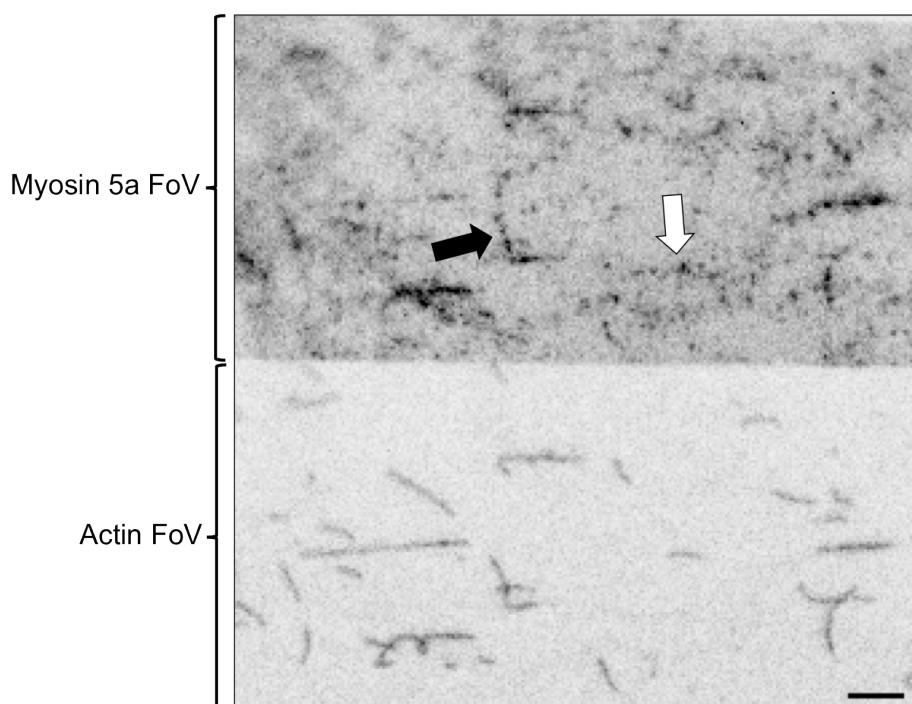
Figure 5.27: Single video frame from TIRF assay of All-23-Bridge construct.

Upper part of the image is the myosin field of view (FoV) and is the fluorescence signal from AF488-CaM-labelled All-23-Bridge molecules excited by a TIRF illumination field. Darker spots are greater intensity. The lower part of the image is the same FoV in the sample but shows only the fluorescence signal from rhodamine-labelled surface-tethered F-actin, darker spots are greater intensity. Scale bar 5 μm. Frame exposure 100 ms. The concentration of All-23-Bridge molecules is high, such that the myosin FoV is saturated with many spots in a given frame. Preferential binding to the F-actin is revealed by the density of dark spots (*e.g.* black arrow) which trace out the actin filament shapes corresponding to the actin FoV. There is also evidence of non-specific surface binding in regions where no actin is present (*e.g.* white arrow).

When the myosin concentration was decreased (to ~0.1 nM), individual spots in the myosin-FoV were then resolved. Many of these spots, thought to be individual fluorescently-labelled molecules, could be observed rapidly appearing and disappearing (typically within 100-200 ms, *i.e.* within 1-2 100 ms exposure frames) across the entire FoV, *i.e.* not confined to actin filament locations. It is thought that these are myosin molecules in solution that fleetingly enter and exit the TIRF excitation field, or momentarily land on the surface before photobleaching or diffusing away. As well as the fleeting spots, videos also showed myosin-FoV spots that persisted over longer timescales, of order 0.5-2.0 s. These longer-lived spots appeared, by eye, to be always static in position, whether localised to an actin filament or not, Figure 5.28.

(a)

(b)

(i)

(ii)
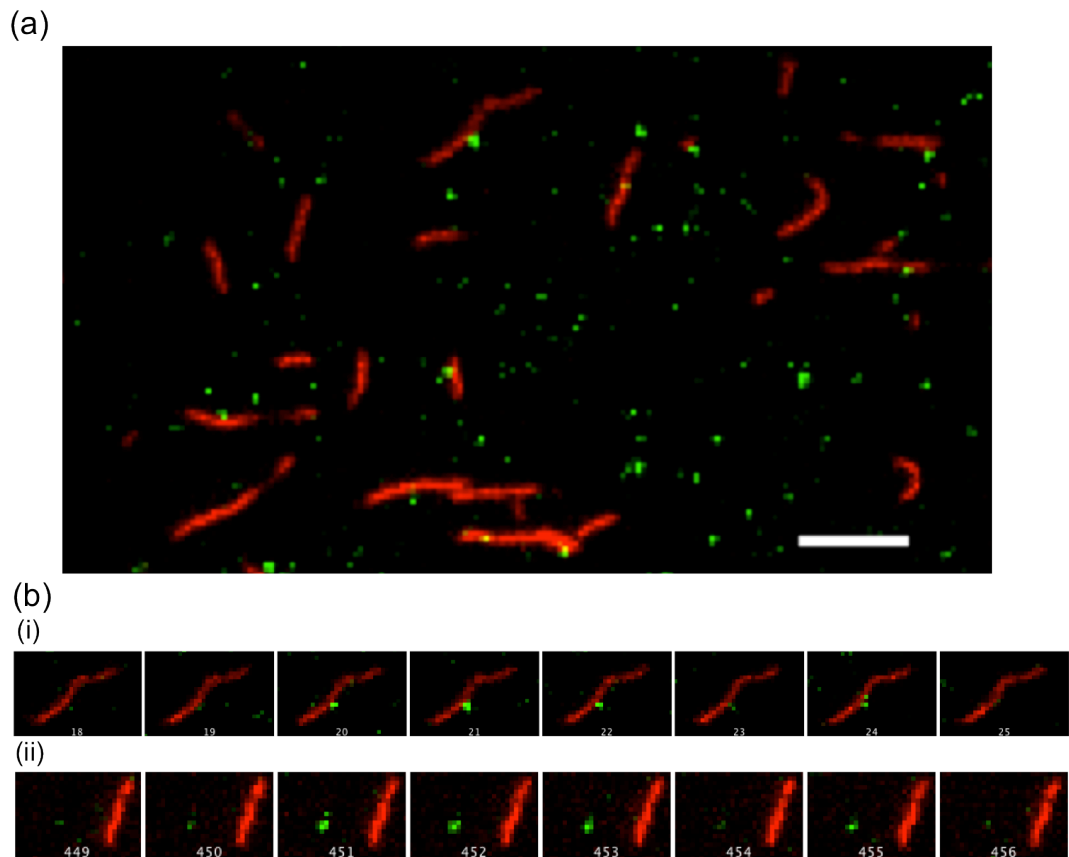
Figure 5.28: Two-colour TIRF assay on All-23-Bridge construct.

Images from the TIRF two-colour detection assay. Green spots are fluorescence from AF488-CaM-labelled All-23-Bridge molecules, red is fluorescence from rhodamine-labelled F-actin. Scale bar 5 μm. (a) single frame (100 ms exposure) showing fluorescence from single All-23-Bridge molecules which rapidly appear and disappear throughout the video sequence. Some spots coincide with actin filament positions while others appear and disappear at non filament locations. Some spots persist for longer, appearing in ~5 or more consecutive frames (≥ 0.5 s). (b) montage sequences showing consecutive frames (100 ms/frame) of two example regions from the FoV in (a). (i) a myosin spot localising to an actin filament but remaining static in position over 5 frames, (ii) a myosin spot localising to the flowcell surface away from an actin filament, also remaining static over ~5 frames.

A number of conclusions can be drawn from these initial TIRF assay results:

1. *CaMs still bind/unbind the All-23-Bridge lever sequence.*
   This straightforward observation supports the same observation made for the negative stain EM, and is evidenced here by the very fact that bright spots can be seen in the myosin-FoV at all. The All-23-Bridge molecules fluoresce only because at least some AF488-labelled CaMs have exchanged onto the altered lever helix. (The possibility remains that the fluorescent spots are simply unbound AF488-CaMs but the intensity of the spots is similar to another HMM-WT assay in which spots moved unidirectionally along F-actin (see 6.6.3), defining them in that assay as labelled HMM-WT molecules).

2. *All-23-Bridge binds F-actin.*
   This is evidenced most clearly by the higher concentration TIRF assay where myosin localised most to the actin filament positions. This observation is again consistent with EM where All-23-Bridge molecules were imaged binding to F-actin by one or two heads, Figure 5.25.

3. *The All-23-Birdge March 2011 purification is non-motile, but a wider conclusion cannot be inferred.*
   This conclusion is perhaps the most frustrating. Although there was localisation to actin filaments, the spots were static, and no obvious examples of spots moving on F-actin were observed across numerous videos of different fields of view. While this leads to the conclusion that this *particular* preparation is non-motile, the wider conclusion that the All-23-Bridge mutant is generally unable to move on F-actin cannot be drawn, since exactly the same non-motile results were seen for the corresponding assay on the March 2011 purification of HMM-WT, *i.e.* the positive control expected to show movement on F-actin was lacking.

Without a good positive control one cannot draw many conclusions on the mutant preparation. The lack of movement in both mutant and control March 2011 purifications in fact suggests a more fundamental problem with these particular protein samples. Since the SDS-PAGE results and EM indicate the presence of a large proportion of single heads and proteolytically-produced lever-tail fragments, this may account for the static myosin spots seen in TIRF. Single heads can bind/unbind from F-actin (and undergo a working stroke if hydrolysing ATP), but without coupling to a second head, unidirectional motion through processive stepping is not possible.

In summary, although the TIRF assay looks promising as a means for characterisation of the All-23-Bridge and other lever mutants (*i.e.* AF488-CaMs exchange onto its lever and are visible in TIRF), the initial results reported here are inconclusive regarding any motile properties or lack thereof. It is clear that All-23-Bridge can be successfully labelled with fluorescently-labelled CaM and that it is able to bind F-actin. However, repeat experiments must be carried out on new purifications of both the lever mutant and an HMM-WT control. Moreover these new preparations should be purified and checked as carefully as possible (*e.g.* including use of salt gradient ion-exchange chromatography) to reduce the complicating factors of the presence of significant amounts of single heads or degradation products. Ideally, the mutant and HMM-WT purifications would be produced at the same time and the HMM-WT control, when assayed in TIRF, should exhibit clear unidirectional movement before an analysis of the mutant is made in the same conditions.

5.4.3.5 Conclusions for All-23-Bridge mutant

The initial results presented for the All-23-Bridge mutant suggest that the alterations to the IQ core spacing to All-23, via native IQ-bridge sequence splicing, do not prevent CaMs binding to the lever. Indeed, on first inspection, the mutant molecules resemble HMM-WT in appearance, both free from and bound to F-actin. Two-headed V-shaped molecules form, CaMs appear to bind levers along their entire length and the molecule can bind by one or both heads to F-actin, with an ~36 nm head-head spacing. Whether the All-23-Bridge construct can step and has similar processivity to HMM-WT is still to be determined, as TIRF motility assays were as yet inconclusive and lack a positive control. No significant difference in catalytic activity was detected between the mutant or wild-type sample, but it is likely that the presence of shHMM fragments in both mutant and control samples was a significant complicating factor that should be addressed (*e.g.* by use of salt gradient ion-exchange fractionation) before future experiments are attempted.

5.4.4 Characterisation of All-25-Insert lever mutant

The All-25-Insert mutant was expressed and purified only twice, including one version of the construct, 'All-25-Insert-GFP', with a C-terminus GFP extension (clone created and expressed by Dr Attila Nagy (LMP-NIH) for potential use in fluorescence microscopy). For this mutant, EM characterisation was performed by Dr Neil Billington (at LMP-NIH), but all images were processed and analysed by DR.

5.4.4.1 Expression and purification

A striking observation, immediately apparent on purification of the expressed All-25-Insert mutant, was a lack of CaM co-purifying with the myosin HC. The SDS-polyacrylamide gel in Figure 5.29(a) has no CaM band in the FLAG peptide elution fractions (lanes 7-12), although bands for the intact HC plus the familiar proteolytic fragments are present. The lack of CaM is most apparent in Figure 5.29(c), where the All-25-Insert sample is compared alongside HMM-WT and All-23-Bridge purifications, which do have a CaM-sized band.

To discount the possibility that loss of CaM was a result of simple human error during protein expression (*e.g.* omission of CaM baculovirus during Sf9 cell transfection), new purifications were made from frozen cell pellets of adapted HMM-WT and All-25-Insert constructs – each engineered with an additional C-terminus GFP (added for potential use in fluorescence microscopy experiments). These constructs, 'HMM-WT-GFP' and 'All-25-Insert-GFP', were cloned from the non-GFP constructs and expressed by Dr Attila Nagy (LMP-NIH). When purified (by DR), the same result was noted, with no CaM band

apparent in SDS-PAGE for the All-25-Insert-GFP construct, but present for HMM-WT-GFP, Figure 5.29(b).
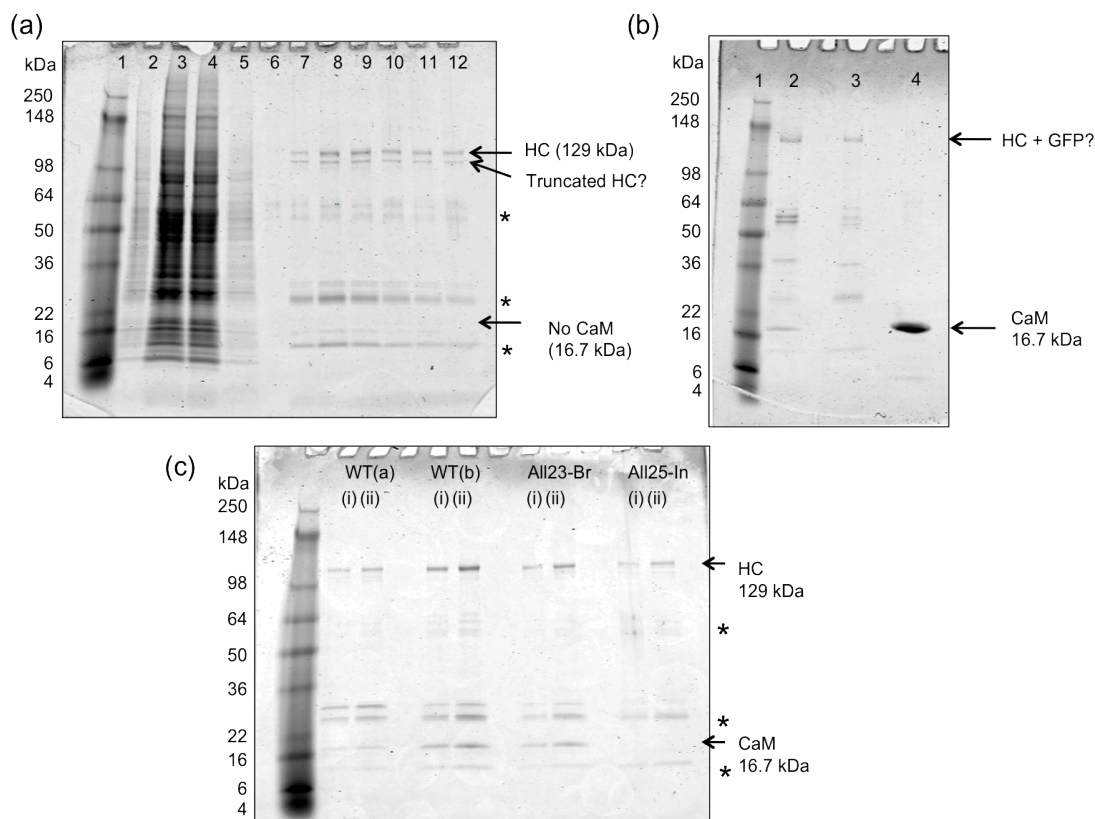
Figure 5.29: SDS-PAGE analysis of All-25-Insert(-GFP) constructs.

(a) SDS-PAGE of FLAG tag purification of All-25-Insert. Flag peptide elution fractions in lanes 7-12, stages from the cell lysis and purification in lanes 2-6 (arranged as for earlier gels in Figure 5.18 and Figure 5.23). Note, no CaM band is present in any of the eluted fractions. **(b)** comparison of purified samples of HMM-WT-GFP (lane 2), All-25-Insert-GFP (lane 3) and CaM control (lane 4). Again, the All-25-Insert construct has no CaM. In both cases, the HC band runs at a greater MW than the construct without GFP, but not at the expected size of 157 kDa. **(c)** comparative SDS-PAGE of two HMM-WT preparations (WT(a) and WT(b)), an All-23-Bridge preparation (All23-Br), and an All-25-Insert sample (All25-In). Two protein loadings were made in each case (i) ~3.7 μg and (ii) ~7.4 μg total mass of protein. No CaM band is seen for All-25-Insert samples. * marks reoccurring degradation/contaminant bands across all preparations.

HMM-WT-GFP and All-25-Insert-GFP bands are seen migrating just below the 148 kDa marker, and are presumably the respective HMM-GFP HCs. Although these bands are shifted, as expected, to a slightly higher mass compared to the non-GFP constructs (with HCs at ~129 kDa), they are nevertheless migrating at a smaller mass than the expected size of ~157 kDa (129 kDa HMM + 28 kDa GFP). This same anomaly was noted also for the HC band of an All-23-Bridge-GFP mutant (data not shown). The presence of the All-25-Insert(-GFP) HCs on the gels of each type of construct argues against the possibility that additional HC proteolysis, or a significant sequence mutation has occurred *e.g.* causing a frame shift and change to the HC sequence (and stop position), that could account for the

236

loss of CaM in the purified proteins. The possibility of sequence mutation was also checked and ruled-out by re-sequencing the DNA sequences (GATC Biotech AG) of the inserts in the source pFastBac™1 plasmids for the All-25-Insert(-GFP) constructs. The SDS-polyacrylamide gels for All-25-Insert(-GFP) also indicate a similar degree of lever-tail proteolysis to all the other constructs, with similar lever-tail(-GFP)-sized bands migrating between ~25-40 kDa (note, GFP constructs exhibit a heavier species fragment migrating at ~40 kDa, presumably due the additional GFP sequence after the tail).

In the baculovirus/Sf9 cell expression process, separate myosin and CaM baculoviruses are allowed to co-infect Sf9 cells, leading to overexpression of both myosin and CaM. On expression (and with the involvement of chaperones) the myosin HCs are thought to fold, bind CaM and dimerise *via* their coiled-coil sequences, such that the CaMs are then retained throughout the anti-FLAG and ion-exchange chromatography steps of purifiation. Clearly in the case of All-25-Insert(-GFP), this process is disrupted due to the LL insertions in the lever helix. CaMs either do not bind, or bind only weakly to the modified lever sequence, and are lost after purification.

5.4.4.2 Electron microscopy

The appearance of the All-25-Bridge-GFP construct was examined by negative stain EM, both in the absence and presence of F-actin and ATP. Dr Neil Billington (LMP-NIH) performed the EM, while DR conducted all image analysis and processing.

5.4.4.2.1 Appearance of free molecules

Figure 5.30 shows example images and results of single-particle image processing of free All-25-Insert-GFP molecules. Figure 5.30(a) and (b) show a typical field of 'particles' and a montage of individual particles, picked from a series of micrographs and computationally aligned using *SPIDER* (v17.05) image processing software (Frank et al., 1996). (For information on rotational and translational alignment procedures and *SPIDER*'s K-means classification protocol, see 3.3.2 and 1.10). It is noticeable that the free particles do not resemble features of either single or double-headed HMM-WT. As might be expected from the SDS-PAGE analysis, there are no obvious CaM-bound lever domains. The montage of particles are globular and heterogeneous in appearance, with an approximately 20-25 nm diameter. Figure 5.30(c) shows the global average and variance images of the aligned particles with a montage below showing results of image classification and averaging on the small dataset of images. Image classification computationally subdivides the images into groups with similar features and the images in each group can then be averaged. For classification, *SPIDER*'s K-means classification

algorithm was used (see 2.9) and the image region of interest was specified by a circular mask (radius 22 pixels (~11.5 nm)) at the centre of the image.

(a)

(b)

(c)

Average Variance (n=196)
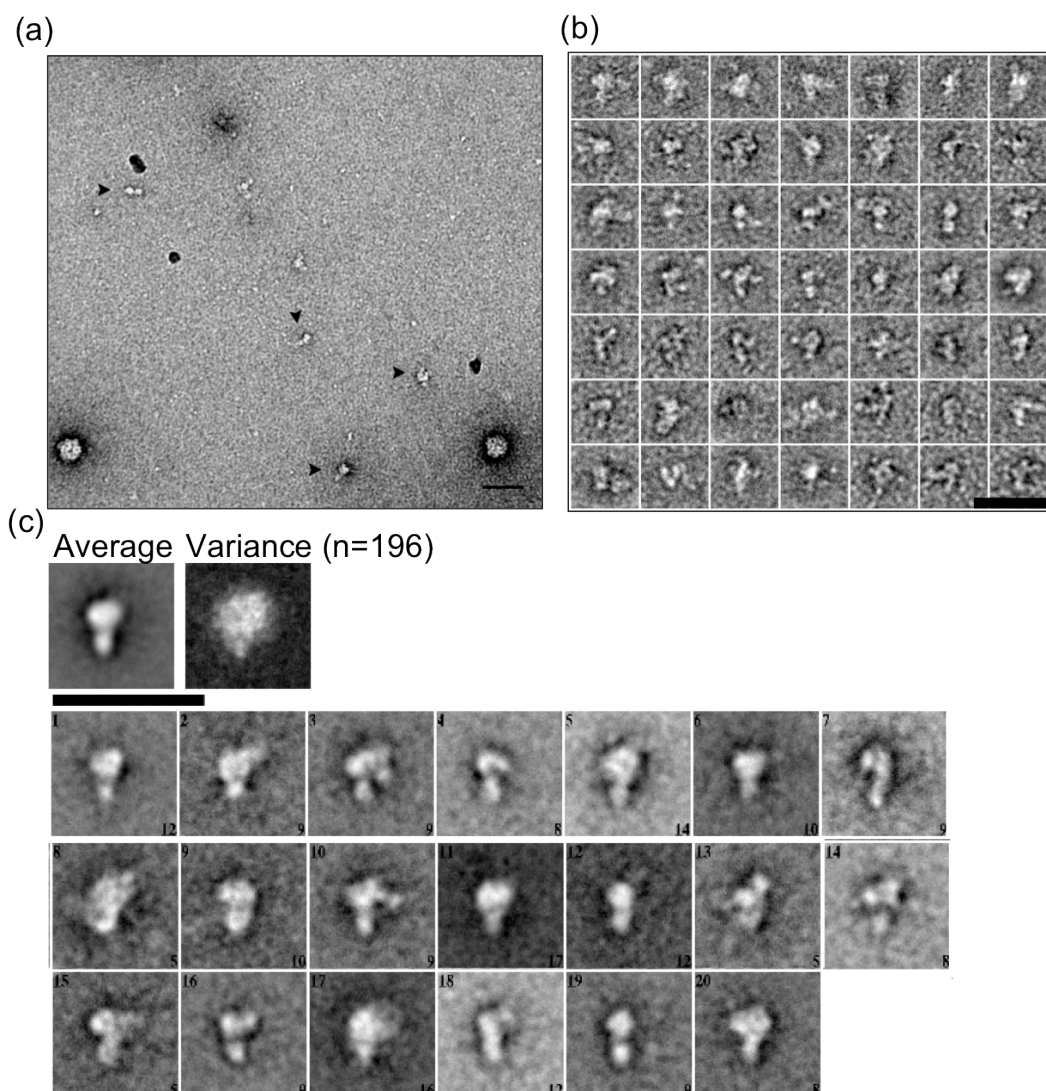
Figure 5.30: Negative stain EM images and single-particle classification of All-25-HMM-GFP.

(a) micrograph example showing sparse scattering of heterogeneous particles, examples highlighted by arrowheads. Some of the particles, *e.g.* example in centre, show features similar to a motor domain appearance. NB: the ~20 nm diameter round objects bottom left and right are contaminants from the anti-FLAG purification. **(b)** montage of selected aligned particles showing variability in appearance. **(c)** global average and variance images for 196 aligned particles, with a montage of image averages after segregation of the images into 20 classes by *SPIDER* K-means classification. The variance image (higher intensity = greater variance) gives an indication of the high degree of variability among particles, which is also reflected in the range of structures in the montage of class averages. In the montage, the averages are numbered 1-20 top left and the number of images combining to make each average are bottom right. Classes 3 and 10 are similar with a globular centre and two projecting arms, suggestive of two similar molecules stuck together. Sample conditions prior to application of 1 % uranyl acetate stain: 125 mM KCl, 10 mM MOPS, 0.1 mM EGTA, 2 mM $MgCl_2$ concentration. Scale bars in all images are 50 nm.

In Figure 5.30(c), the global variance image shows a high degree of variability across the images, and the classification averages show a range of structures. There is perhaps some repetition of a 'P' shape (classes 4, 6, 16, 18) and its long axis mirror (classes 7, 9, 15, 20). Although this may faintly echo the shape of a single myosin (nucleotide-bound) motor domain with a truncated (*e.g.* single CaM) lever, the resemblance is not strong, and inspection of the individual particles forming these classes are still highly variable. Two class averages, numbers 3 and 10, share an appearance of two arms projecting from a globular centre. This shape raises the possibility that these classes are two molecules stuck together, and that the others are single molecules. Overall, the high variability and lack of resemblance to myosin features suggest that the All-25-Insert(-GFP) construct is misfolded, presumably as a result of the LL insertions in its lever helix and the associated inability to bind CaM. Incubation of free molecules with high concentrations of CaM did not noticeably alter the particle appearance in EM.

5.4.4.2.2   Appearance bound to F-actin

Negative stain images of All-25-Insert-GFP in the presence of F-actin (at ~0.1 μM) and ATP (at ~1 μM) are shown in Figure 5.31. The micrograph section in Figure 5.31(a) shows that despite their misfolded appearance, the molecules do appear to bind to F-actin and moreover exhibit a degree of polarity in the angle they make with the filaments. Virtually no free molecules were seen in the background of F-actin micrographs. (*NB:* As before, ~20 nm circular objects, some sticking to the F-actin, are contaminants from the anti-FLAG purification). Figure 5.31(b) shows a selection of *SPIDER*-aligned F-actin-bound molecules, with the global average and variance images right. These indicate a globular but variable structure bound to one side of the filament. Segregation of particles into varying numbers of classes indicated some consistency in features for groups of molecules. In particular, particles varied in the direction they 'point' along the F-actin, and in the region where they contact with the filament. Classifying on only the thin F-actin contact region, Figure 5.31(c), illustrates these features. Classes 1 and 4 are symmetrical and show a similar object that leaves the actin filament at ~45° and seems to have one right-angled contact point with an actin subunit. Beyond this, a more globular region projects away from the filament. In form, these classes are similar to a single myosin 5a head bound to F-actin, although the trailing globular region is clearly very unlike a HMM-WT six-CaM lever. The F-actin attachment point is suggestive of a single myosin 5a motor domain binding to F-actin (for comparison see wild-type motor average images in Figure 5.31(d) (i) and (ii)). Classes 3 an 5 are also symmetrical and similarly globular in appearance, but here the contact with the actin filament is more diffuse and may be at more than one actin subunit. For comparison, averaged negative stain images of

nonmuscle myosin 2a & 2b HMM molecules, with two heads binding to adjacent subunits are shown in Figure 5.31(e) (i) and (ii). Examination of the All-25-Insert-GFP classes does not determine if the molecules are two-headed, single-headed or a mixture. It does establish that although clearly very unlike HMM-WT in appearance, an F-actin-binding region of the head is still present and capable of attaching to F-actin and orientating the remainder of a malformed lever-tail section along the filament.
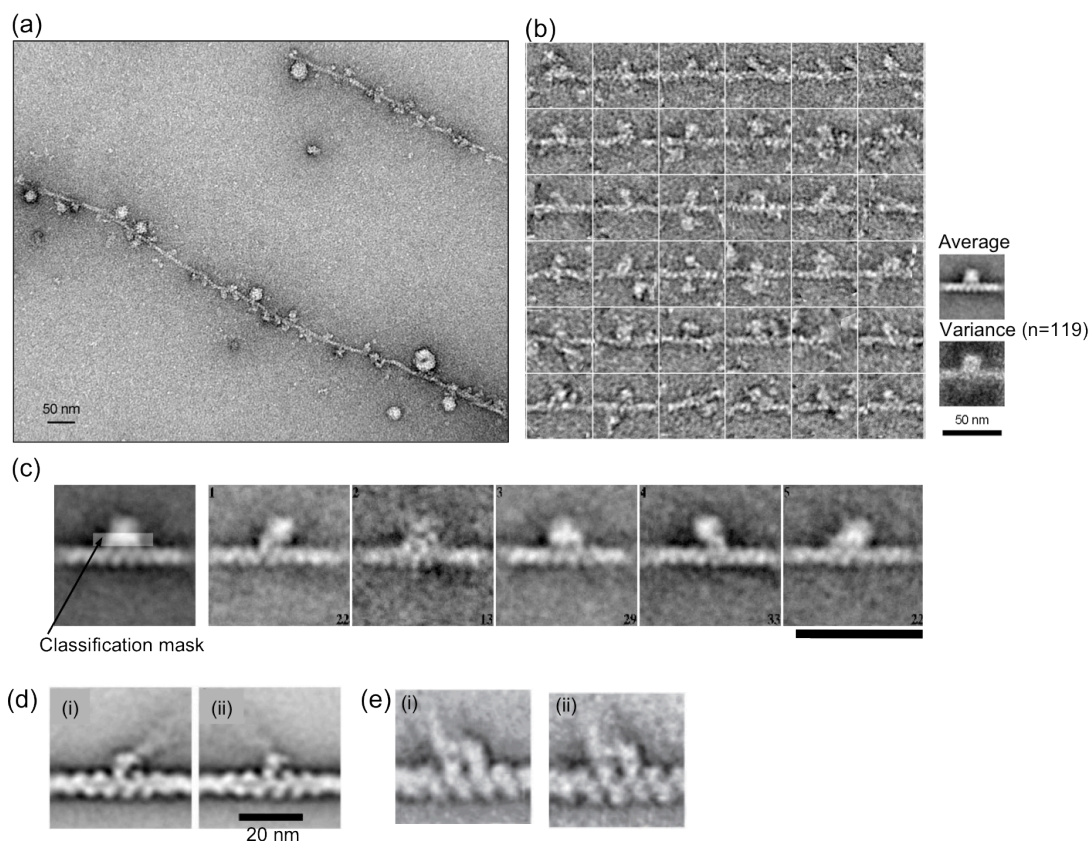
Figure 5.31: Negative stain EM images and single-particle classification of All-25-HMM-GFP in the presence of F-actin and ATP.

Samples conditions prior to staining with 1 % uranyl acetate: 125 mM KCl, 10 mM MOPS, 0.1 mM EGTA, 2 mM $MgCl_2$, pH 7.0, **(a)** Example micrograph section showing molecules binding to two actin filaments. *NB*: as in earlier EM images, the larger ~20 nm diameter round objects are contaminants from the anti-FLAG purification. The molecules attached to the lower filament have a general polarity in the angle they make with the filament (forming arrowheads with tails to the right). **(b)** montage of selected aligned molecules bound to F-actin. The actin filament runs horizontally. Global average and variance images are shown right. The global average shows a compact particle bound to one side of the filament. **(c)** classification averages after image segregation into 5 classes. The left most image shows the global average with the classification mask superposed (transparent white rectangle). The mask focuses the classification on the appearance of the molecules around the actin-binding interface. Class averages are numbered 1-5 top left, with the number of images forming each average bottom right. For comparison, **(d)** shows previously-obtained image averages of myosin 5a HMM-WT trail (i) and lead (ii) heads bound to F-actin (Walker et al., 2000), while **(e)** shows averages of non-muscle myosin 2a (i) and 2b (ii) HMM molecules with the two heads binding to adjacent actin subunits (Kovács et al., 2007).

5.4.4.3 Actin-activated MgATPase

Using the NADH-coupled assay, the actin-activated MgATPase of the All-25-HMM-GFP sample was tested and compared to a HMM-WT-GFP control, see Figure 5.32. Buffer conditions were as for the HMM-WT/All-23-Bridge assay (section 5.4.3.3) though KCl concentration was ~75 mM for All-25-HMM-GFP but only ~45 mM KCl for wild-type control. Temperature was ~23°C but not accurately controlled.

Within the 0-10 μM actin range tested, All-25-Insert-GFP molecules showed no actin-activation above a low basal rate (< 0.03 s^{-1}head^{-1}), unlike the wild-type control, which did show an actin-activation and low K_{ATPase} of 0.42 μM, characteristic of myosin 5a-HMM (see Table 5.10). The HMM-WT-GFP rates rise to an apparent V_{max} of 0.33 s^{-1}head^{-1}. Although this rate is very low, *e.g.* compared to HMM-WT (Table 5.9), this is probably a reflection of an overestimated molecule concentration as, in this assay, the HMM concentrations were not adjusted by gel densitometry analysis, and the preparation had not been subject to the ion-exchange chromatography step to remove peptide and other contaminants. Thus, the basal rate of All-25-Insert-GFP could really be as high as 0.5 s^{-1}head^{-1}. Never-the-less, the significant result here is the lack of ATPase activation by F-actin for the lever mutant compared to WT. The result suggests that the loss of CaM and/or misfolding of All-25-Insert(-GFP) severely impairs the catalytic function of the acto-myosin complex. This might be due to a reduced actin-binding affinity (though the EM analysis shows at least some actin binding capacity) or, most likely, is due to the fact that the molecule has a malformed structure.
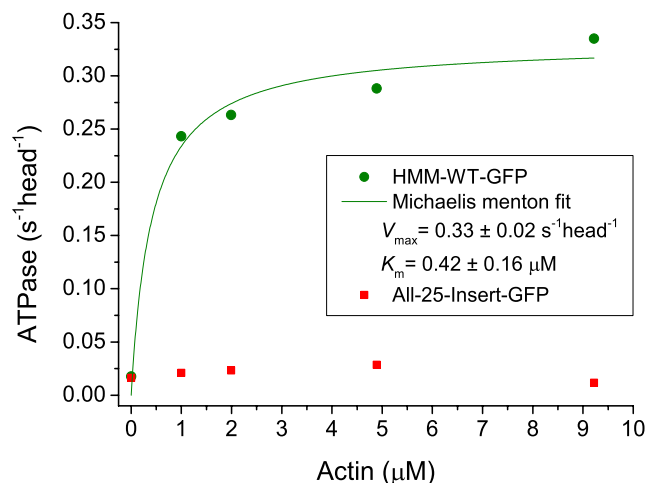
Figure 5.32: Actin-activated ATPase data for All-25-HMM-GFP and WT-HMM-GFP.

For the 0-10 μM actin, the All-25-Bridge-GFP construct stays at a low basal rate (< 0.3 s^{-1}head^{-1}) while HMM-WT-GFP shows a marked actin-activation (and low K_{ATPase}) consistent with All-23-Bridge/HMM-WT assays (Figure 5.26). *NB*: KCl concentration was ~75 mM for the All-25-HMM-GFP sample and ~45 mM for wild-type control.

5.4.4.4   Conclusions of All-25-Insert lever mutant

Taken together, the SDS-PAGE, EM and actin-activated ATPase results for All-25-Insert(-GFP) show that the three LL insertions in the bridge sequences between IQ-cores 1 & 2, 3 & 4 and 5 & 6, disrupt all CaM binding (even to IQ motifs that retain their native sequence), resulting in a misfolded molecule that is unable to bind MgATP and/or release product efficiently when bound to F-actin. This result is surprising. Sequence alignments for myosin 5a (see Figure 5.5 and Figure 1.17 in chapter 1) indicate that leucine is the most frequently occurring a/a in the equivalent insertion positions in the native 25-spaced IQ motifs. Heptad net representations of the mutated sequences did not suggest a destabilising clustering of like charges in the portions of altered helix. However, with the exception of IQs 4-5 in zebrafish, LL is not a naturally occurring sequence in any of the other myosin 5a sequences checked. Also, it is noted that in two out of the three LL insert cases (IQs 1-2 and 3-4), the LL insertions result in a run of three hydrophobic side chains in the middle of the α-helical bridge structure. Such a pattern does not occur naturally in any of the native IQ motifs checked, and may be fatal for proper formation of the lever helix.

CaMs might fail to bind the All-25-Insert helix if important stabilizing CaM-CaM interactions are disrupted *e.g.* by altered CaM azimuths, or simply if the lever helix fails to form correctly. Likely as a result of this and/or because of exposure of stretches of hydrophobic helix residues, the molecule becomes significantly misfolded in the lever-tail region, giving the free molecules a heterogeneous globular appearance and the F-actin-bound molecules an overall globular shape or globular lever-tail region. The residual ability of the molecules to bind and orientate on F-actin indicates that at least part of the motor domain remains intact, though the lack of significant actin-activated MgATPase indicates that the actin-induced catalytic function and certainly any processivity of the molecule is destroyed. EM images do not definitively determine if two HCs are still able to dimerise to make a two-headed molecule, though one might speculate that if the lever helix is misfolded, the adjacent tail could do also, and this could prevent dimerisation by the coiled-coil sequence. Although the size and form of some of the free and actin-bound averages are suggestive of single myosin 5a heads, the possibility that the attached molecule images were of two tightly packed heads binding to proximate actin subunits (comparable to the myosin 2a/b HMM averages) remains.

From the previous characterisation of the HMM-2Ala-6IQ construct by Oke et al. (2010) and from EM characterisation of an All-25-Bridge intermediate construct cloned (by DR) and expressed (by AN) for this project, it is possible to conclude that alteration of a single

23 to 25 a/a IQ-spacing in itself does not abolish CaM binding to the lever. Figure 5.33 shows EM images of negatively-stained molecules of these two 'extra 25' IQ constructs. In Figure 5.33(a) the molecules have the IQ spacing pattern 25-25-23-25-23, where the first IQ spacing was changed to 25 by splicing-in the native bridge sequence ([807]CYAKFLRRTKAAT[820]) from IQs 2-3. In Figure 5.33(b) images of the HMM-2Ala-6IQ mutant from (Oke, 2004) are shown. In this case the 23 to 25 conversion was made by insertion of an alanine pair into the IQ 2-3 bridge sequence (giving a 23-25-25-25-23 pattern).

(a)

(b)
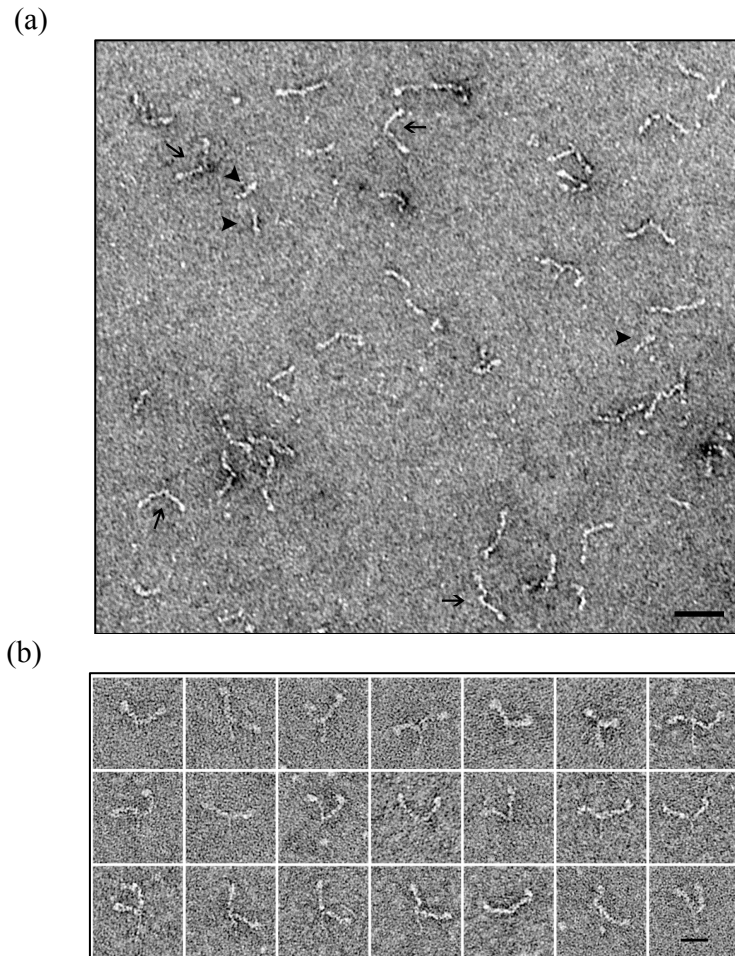
Figure 5.33: Negative stain EM images of lever mutants with one IQ spacing altered to 25 amino acids.

(a) field of molecules of an expressed intermediate of the All-25-Bridge construct. The molecule has a 25 a/a spacing between IQs 1-2, formed by splicing-in the native bridge sequence from IQs 2-3. The overall IQ spacing pattern is 25-25-23-25-23. Arrows indicate examples of two-headed V-shaped molecules with an HMM-WT appearance. CaMs apparently bind along the length of the levers. Arrowheads highlight examples of single head fragments. Scale bar 50 nm. Image acquired by Dr Neil Billington (LMP-NIH). Grid conditions: 100 nM HMM. 25 mM KCl, 10 mM MOPS, 0.1 mM EGTA, 2 mM MgCl$_2$, pH 7.0. 1 % uranyl acetate stain. (b) images of individual 'HMM-2Ala-6IQ' mutant from Oke (2004). This mutant had its third IQ motif altered from 23 to 25 a/a by insertion of two alanines. The molecules apparently still bind CaMs along their entire lengths despite the altered lever helix sequence. The molecules exhibited no major differences with an HMM-WT comparison. Scale bar 25 nm.

In both mutant cases molecules with an V-shaped appearance very similar to wild-type were seen. CaMs can apparently bind along the lever lengths and no increased kinking of the lever at the mutation position was observed over the wild-type controls.These observations suggest that the cause of the failure of the All-25-Insert construct was to do with the particular choice of two leucines as additions in the lever helix rather to with changing 23 to 25 IQ-spacings. It is frustrating therefore that the alternative sequence All-25-Bridge construct (which uses spliced-in native sequences) was not amenable to cloning (or synthesis) within the timescale of this project. The original research question about the structural and functional effects of changing IQ spacings from 23 to 25 a/a remains unanswered.

5.5   Discussion and future prospects for lever mutants

The aim of work in this chapter was to create and characterise a set of myosin 5a HMM lever mutants, where the native IQ motif spacing was altered from 23-25-23-25-23 to All-23 or All-25. Clones were successfully produced for 3 out of 4 of the target sequences, where two alternative mutated sequences were pursued for both the All-23 and All-25 cases. Of the target sequences, only the All-25 construct formed using repeats of the native IQ 2-3 bridge sequence resisted cloning or synthesis by step-wise or In-Fusion reaction.

Protein purifications made generally suffered from various sources of contamination, including unidentified circular objects from the FLAG peptide batch, possibly unremoved FLAG peptide itself, additional fragments or unknown proteins at ~10 and 55-65 kDa, and, most significantly, lever-tail fragments and shHMM molecules caused by proteolysis at the C-terminus half of the lever domain. These contaminations are a significant issue, as they make estimation of the concentration of intact HMM molecules difficult. SDS gel densitometry can at least provide a rough head concentration estimate *e.g.* for ATPase rate calculations, but the sample heterogeneity is still far from ideal, since the mix of single and double-headed molecules makes interpretation of the ATPase, TIRF or other assays difficult.

Some improvements were made to the purification procedure. Firstly, storage of purified protein in 0.5 M KCl resulted in less aggregation. Secondly, addition of ion-exchange chromatography with salt step (0.1-0.5 M NaCl) in the purification protocol, which helped remove the FLAG contaminants, but was apparently still not sufficient to separate the 10, 55-65 kDa fragments/proteins or single heads from the two-headed HMM molecules. In future work, use of a KCl salt gradient (0.1-0.5 M) (rather than step) in the ion-exchange chromatography would be an important further improvement to try. Future work should

also address the high levels of proteolysis seen in these purifications. An expression study varying the levels of CaM virus used to co-infect the Sf9 cells could be made as a higher level of CaM present during expression might help protect the hydrophobic IQ motifs of the lever helix from proteases. Supplementary CaM could also easily be added to the Sf9 cell lysate during purification, as has been be done in previous studies involving expressing of recombinant myosin 5a-HMM (Moore et al., 2001).

The preliminary characterisation of the two lever mutants All-23-Bridge and All-25-Insert, showed contrasting results. The All-23-Bridge mutant at first sight appears to be very similar to wild-type, both in its appearance and its ATPase properties, while the All-25-Insert construct showed a catastrophic loss of CaM binding and a misfolded enzymatically impaired molecule.

These results form the basis for future investigations that could seek to answer some of the key questions that remain unanswered from the initial work in this chapter. Perhaps of most interest from a biological perspective is the outstanding question of whether the All-23-Bridge can still step processively along F-actin, and, if so, if it has similar or altered characteristics to wild-type. The TIRF assay was conceived as the most informative experiment to address this question. However, though the fluorophore-labelled CaM exchange was successful, and single All-23-Bridge molecules were visualized in TIRF, the lack of a motile positive control, plus doubt over the sample quality and number of active two-headed molecules present, leaves the motility question currently unanswered. Repetition of the TIRF assay with better quality samples would be a promising and priority future experiment.

Although the results for All-25-Insert are more clear cut, the mutant failing to bind CaM and misfolding in lever-tail region, the broader question of whether it is the All-25 spacing itself, or just the specific mutations chosen that were responsible, remains to be further investigated. The evidence from the HMM-2Ala-6IQ mutant and from the All-25 IQ bridge intermediate created for this chapter, both show that change of a single 23 to a 25 is possible without being catastrophic for binding of CaMs or, in the HMM-2Ala-6IQ case, the continued (albeit impaired) functioning of the molecule. To continue the work begun in this chapter, it would perhaps be most interesting to next pursue expression of the second All-25-Bridge intermediate (for which a clone has already been constructed) with two out of the three of the 23 to 25 changes made via the splicing strategy. Characterisation of this mutant, to see if all CaMs are bound to the lever, any ATPase and motile properties, would help answer the question of whether the specific choice of 'LL' insertions in the All-25-Insert mutant were the cause of lost CaM binding, or whether there

is some cooperative binding effect for CaMs along the lever, requiring that certain IQ-spacings be 23 a/a. Comparing a 'Bridge'-type All-25 intermediate to the 'Bridge'-style All-23-Bridge construct would also be a more systematic comparison.

Finally, a natural future direction for the lever mutant study would be to capture more negative stain EM data of images of attached and unattached mutant molecules, and apply to them the same single-particle image processing, classification and analysis techniques as applied to HMM-WT in chapters 2 and 3. Does the mutant's appearance and mechanical properties of the head differ from that of HMM-WT? Do the CaMs in the lever of the All-23-Bridge construct adopt a different pattern of orientations to that of wild-type? Does an All-23 lever exhibit the same CaM-IQ twisting flexibility to wild-type and are the lever's flexural rigidity properties the same?

Chapter 6: Calmodulin Labelling Study

6.1   Introduction and aims

This chapter presents results from a standalone project related to myosin 5a lever mutant studies in chapter 5. The project involved the production and use of two calmodulin (CaM) point mutants, created to allow site-specific attachment of a fluorophore or other functional label. The labelled CaM molecules were designed to be able to bind to the lever of myosin 5a (or indeed to other myosin levers) and serve as a probe in different microscopy experiments.

In particular, this chapter covers experiments performed to exchange single fluorophore-labelled CaMs onto myosin 5a-HMM molecules for tracking in total internal reflection fluorescence (TIRF) microscopy experiments (Axelrod, 1989; Funatsu et al., 1995). In such experiments, individual fluorescently-labelled myosin molecules can be seen as spots of light under TIRF illumination conditions (Sakamoto et al., 2000; Tokunaga et al., 1997). Myosin localisation with F-actin and any motile characteristics (*e.g.* velocity, run length, and step size) can potentially be measured from video images of the labelled molecules (Churchman et al., 2005; Sakamoto et al., 2008; 2005; 2003; 2000; Yildiz et al., 2003).

A future use envisaged for the CaM mutants is for gold particle labelling and use in EM, as an aid to potentially understanding torsion in the myosin 5a lever. For example, a gold nanoparticle could be site-specifically attached to a CaM mutant and then exchanged onto myosin 5a-HMM, as done in Dunn and Spudich (2007), where light scattering from the gold particle was used to track myosin 5a head movement. If the tagged CaMs were imaged by negative stain EM for example, the gold particles might appear as high contrast markers of the CaMs along the lever domain. From their position relative to the myosin molecule this could indicate if particular CaMs preferentially exchange and also yield information on any specific rotational orientations the CaMs might adopt on a carbon substrate in EM stain conditions.

6.2   Design of CaM cysteine mutants

It was decided to create a CaM cysteine mutant for site-specific labelling, as this was preferable to using other non-specific labelling methods (such as targeting accessible lysine residues with N-hydroxysuccinimide ester-derivatized labels) where there is less

247

control over the number and positioning of the label. Maleimides carrying an attached fluorophore or other label will readily react with thiol groups such as that found on cysteine. The native *Xenopus* CaM sequence used in the work here (UniProt accession: P62155) has no cysteine, so the creation of a single cysteine residue as the reactive target offers the advantage that a unique labelling site can be created, and one has control over the maximum number of labels that can attach to each CaM. Moreover, the labelling site can be positioned so as to reduce the possibility of any attached label interfering with potentially important protein structure and sites of interaction with the myosin heavy chain (HC) or other CaMs.

A suitable residue was sought for mutation to cysteine. The criteria for selection of the location were:

1. *To be surface-accessible and solvent-exposed* (in both Ca^{2+}-bound and apo CaM conformations). This is important to maximize the likelihood of reaction with the maleimide-dye and to reduce the chance of disrupting the protein structure or causing the CaM to misfold.

2. *To be located away from the binding site for the myosin heavy chain.* For obvious reasons it was important to preserve the binding to the myosin lever domain IQ motifs.

3. *To be located away from any potential CaM-CaM interaction area.* This was particularly important for the lever flexibility analysis. If CaM-CaM interactions between neighbouring CaMs are important for local lever stiffness which may be important for function *in vivo*, then introduction of the maleimide-dye should not disrupt these interactions.

4. *To be located away from the motor domain.* The attached maleimide-dye on CaM 1 (closest to the motor) should not interfere with the myosin motor domain (*e.g.* contact the converter) in its different positions in the myosin conformational cycle.

5. *To be located so as to preserve the CaM Ca^{2+}-binding capacity.* This property of CaM and its associated 'closed' to 'open' conformational change is an important switch that facilitates binding or unbinding from myosin or other target proteins. A principal method by which CaM can be purified utilises this Ca^{2+}-binding conformational change, and it is also exploited as a method for swapping exogenous (*e.g.* labelled) CaMs into myosin molecules, a process known as CaM 'exchange'. For these reasons, the preservation of the Ca^{2+} binding capacity of all four binding sites was important.

Crystal structures of CaMs 1 and 2 bound to a fragment of the myosin 5a lever helix (Houdusse et al., 2006) (PDB code: 2IX7) and the myosin 5a motor-plus-first-light-chain in different nucleotide states (Coureux et al., 2004; 2003) (PDB codes: 1OE9, 1W7J, 1W8J, 1W7I) were inspected to find a suitable mutation site. In addition, literature was searched for any precedents for CaM cysteine mutants that might be appropriate for reuse here. After consideration, two alternative mutation sites were selected. One, Q135C[†], following the precedent of Churchman et al. (2005) (later reused by Dunn and Spudich (2007)), and a second, E139C[†], a new point-mutant created specifically for this study, see Figure 6.1.

6.2.1 CaM mutation Q135C

This mutation follows an existing precedent first created by Churchman et al. (2005) for CaM-labelling with fluorescent Cy3 and Cy5-maleimide dyes, for myosin 5 tracking by single-molecule high-resolution colocalization (Churchman et al., 2005). In their work, the single cysteine was introduced into sea urchin vertebrate-like CaM via the mutation Q143C. In the *Xenopus* CaM sequence (using the numbering system adopted for this chapter, see footnote †) this is homologous to Q135C. Dunn and Spudich used the same CaM mutant for biotin-malemide reaction to provide an attachment point for streptavidin-coated gold nanoparticles for myosin 5a head tracking by light scattering (Dunn and Spudich, 2007).

Figure 6.1(a) and (b) show the site of Glutamine 135 highlighted on a CaM-heavy chain crystal structure. It can be seen that the residue is located in a surface loop in the C-lobe of the CaM. The loop is part of the fourth Ca^{2+}-binding helix-turn-helix or 'EF hand' motif. The residue is in 'position 7' in the Ca^{2+}-binding loop (see 1.6) but the side-chain plays no role in co-ordinating the Ca^{2+} cation (rather the carboxy group of the polypeptide backbone does). The site would be accessible, is unlikely to create a clash with the motor domain, is distant from the myosin binding site and is reasonably far away from any CaM-CaM interaction area. Reservations about choosing this site are that it is possibly a little buried and because of the positioning in the loop, the side-chain with a bulky dye attached via a flexible carbon chain might project towards the CaM-CaM interaction zone. The

---

[†] The numbering used for residues in the *Xenopus* CaM sequence (UniProt: P62155) in this thesis is based on CaM's *final* polypeptide sequence following post-translational cleavage of the initiator methionine *i.e.* the second residue, alanine, in the full (uncleaved) sequence is labelled here as residue 1. Hence, the CaM sequence referred to here runs [1]ADQLTEE…MMTAK[148].

advantage of adopting this mutation is that the earlier studies have already demonstrated that the mutation is not disruptive to CaM-folding, that the Ca^{2+}-induced conformation change is maintained and that the motile properties of myosin 5a with the mutant bound are unchanged.
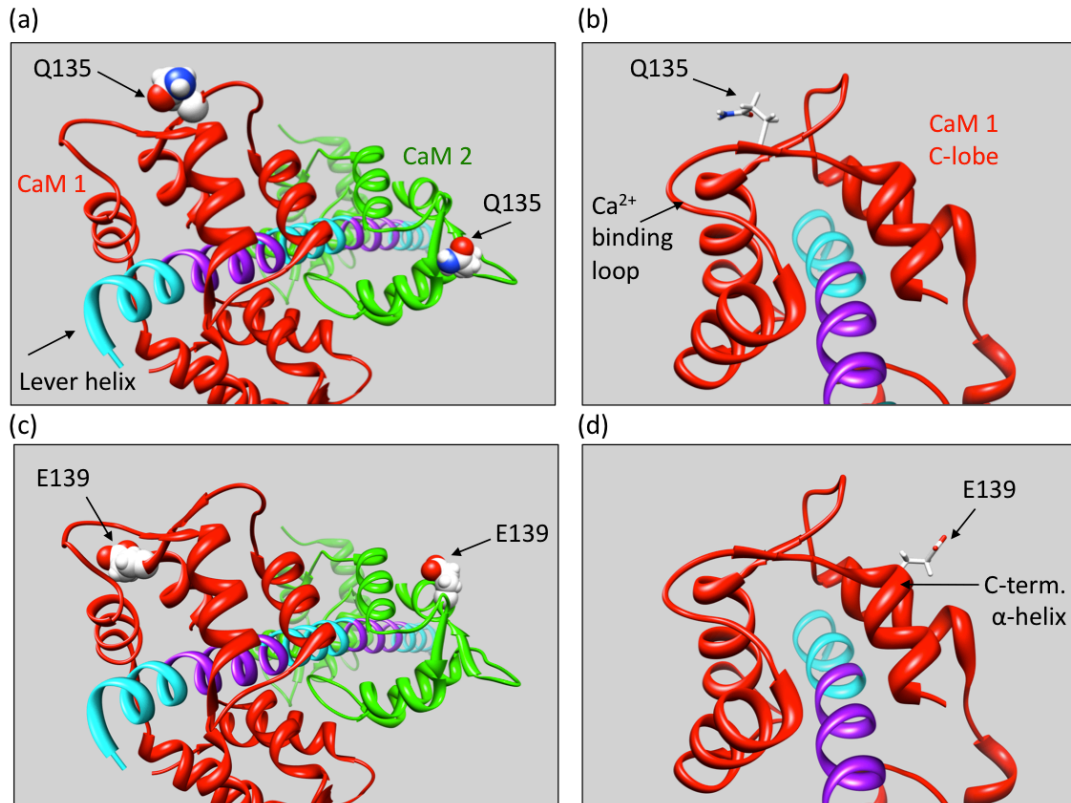
Figure 6.1: Views of the CaM mutation sites selected.

Views of the crystal structure of apo CaMs 1-2 bound to the first two IQ motifs of the myosin 5a heavy chain (PBD: 2IX7) (Houdusse et al., 2006). CaM 1 is in red, CaM 2 green, the lever helix in cyan with the two IQ motifs highlighted purple. (a) and (c) are a view looking along the lever helix (in N to C terminus orientation) with CaM 1 in the foreground and CaM 2 in the distance. Residues Q135 and E139, located in the C-lobe region of the CaMs (the region of CaM 1 positioned above the lever helix in (a) and (c)), are shown in spacefill. (b) and (d) are a view looking down on the N-terminus section of lever helix at the first IQ motif, showing just the C-lobe region of CaM 1, with Q135 and E139 shown in stick representation. Q135 is located in the final Ca^{2+}-binding loop of the last EF-hand (but does not participate in Ca^{2+} coordination) while E139 is at the start of the last helix of the CaM. Although only separated by three residues, the two side chains are evidently orientated differently. (c) shows how E139 projects into the solvent in a direction away from the neighbouring CaMs, while (a) shows that Q135 of CaM 2 is projecting towards CaM 1.

6.2.2 CaM mutation E139C

Glutamate 139 is shown highlighted in the CaM-heavy chain structure in Figure 6.1(c) and (d). The residue is located at the start of the final helix in the CaM C-lobe as part of the final EF hand motif. The residue is at position 11 in the Ca^{2+}-binding loop but has no cation coordination role. Because of its position at the start of the final helix, although being only 4 residues downstream of Q135, the side chain projects in a different direction.

250

It projects into the solvent and in a direction more likely to orientate an attached dye away from the CaM-CaM interaction zone. It also avoids clash with the motor domain and is distant from the myosin IQ binding site. As such, it meets all the criteria for the labelling site. One reservation for this mutation is that it is located at the start of an α-helix, and change of amino acid here could be disruptive to correct formation of the helix and correct structure formation of the CaM, especially replacing a negatively charged Glu with neutral Cys or a large label. However, a heptad-net analysis of the amino acids in the mutated helix sequence was made (Baboolal et al., 2009), and it was assessed that introduction of a Cys would be unlikely to be destabilising.

It was judged that of the two mutation sites identified, the new E139 site was preferable in its projection away from the CaM-CaM interaction zone. However, since there was a precedent to demonstrate successful folding, Ca^{2+}-binding and myosin 5a exchange for Q135C it was decided that the best policy was to pursue the production of both CaM mutants. Q135C would serve as a back-up to E139C if the latter proved to be problematic to overexpress and purify, or to be non-functional in any way.

6.3   Production of CaM cysteine mutants

6.3.1   Overview

In outline, the two CaM cysteine mutants CaM-E139C and CaM-Q135C were produced using rounds of polymerase chain reaction (PCR) using a source cDNA sequence for native CaM (*Xenopus laevis*, UniProt accession: P62155) and DNA primers designed to introduce the required mutations. Following cDNA mutation and amplification by PCR, the mutant sequences were each cloned into pCR™2.1-TOPO® vectors (Invitrogen – Life Technologies, Grand Island, USA) and, after sequence verification, the DNA inserts were transferred to a pET-28a(+) vector (Novagen® – Merck KGaA, Darmstadt, Germany) for use in protein overexpression in *E. coli* cells. CaM overexpression was induced by addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to 1 L bacterial cultures. The CaM proteins were purified from cell lysate using an established liquid chromatography protocol (Dedman and Kaetzel, 1983; Gopolakrishna and Anderson, 1982), first by binding (via hydrophobic interaction) to phenyl-sepharose resin in their Ca^{2+}-bound conformation and then releasing in the apo (no Ca^{2+} bound) conformation when a EDTA-containing buffer was applied.

More detailed descriptions of the methods and materials used for the mutagenesis, expression and purification steps are described below with corresponding results sections thereafter.

251

6.3.2 Mutagenesis

Polymerase chain reaction (PCR) was used to generate cDNA sequences for the two separate CaM mutants, CaM-Q135C and CaM-E139C. Source cDNA for the native CaM sequence for use in the reactions was a kind gift from Dr Melanie Colegrave (formerly University of Leeds, UK). The native cDNA had already been manipulated to contain flanking restriction enzyme sequences for *NcoI* and *SalI* at the 5′ and 3′ ends, and was contained in the cloning site of a pFastBac™ vector (CaM-FB) (Invitrogen - Life Technologies) which had already been linearised by digestion with *EcoRI* enzyme. PCR primers were synthesized by Integrated DNA Technologies (Coralville, USA) or were a gift from Dr Colegrave (University of Leeds).

PCR and ligation reactions were carried out using a Eppendorf Mastercycler® Gradient PCR machine (Eppendorf AG, Hamburg, Germany) using reagents from a GoTaq® qPCR kit (Promega, Madison, USA). PCR products were cloned into a pCR™2.1-TOPO® vector (Invitrogen – Life Technologies) and transformed into NEB 5-*α* high efficiency chemically competent *E.coli* cells (New England Biolabs (NEB), Ipswich, USA) for amplification. Cells were cultured in Lennox L Broth Base (LB) mixture (Invitrogen – Life Technologies), mixed at 10 g LB powder per 500 mL of milliQ H_2O and pre-sterilised by autoclave. Bacterial colonies were grown on LB-agar plates, the plate mixture formed from 2 g LB, 1.5 g Agar (Oxoid Ltd., Basingstoke, UK), 100 mL milliQ H_2O, pre-sterilised by autoclave, before addition of appropriate antibiotics (on cooling) and pouring into petri dishes to set.

DNA products from PCR and restriction digest reactions were examined by agarose gel electrophoresis. 1.5 % agarose gels were made by mixing and heating 0.75 g agarose in 50 mL TAE buffer (40 mM Tris base, 40 mM Acetic Acid, 1 mM EDTA), adding 2.5 μL of ethidium bromide as the mixture cooled. Gels were run submerged in TAE buffer in a horizontal electrophoresis gel tank, typically at 100 V input voltage.

6.3.2.1 CaM-Q135C PCR protocol

To make the DNA sequence for CaM-Q135C, suitable primers (Table 6.1) and a three-reaction protocol was devised, see Figure 6.2. In reaction 1 the reverse primer *CaM135C_R* binds the template cDNA at base pair (bp) 397-417 in the native sequence, but deviates to introduce the required CAA→TGC base mutation (mutates glutamine to cysteine) at bp 406-8 (red in Figure 6.3). The product of reaction 1 is the first 419 bp of the CaM-Q135C sequence. In reaction 2, the forward primer *CaM135C_F* also binds at bp 397-417 and similarly incorporates the mutation of the native sequence. This reaction produces the final 60 bp of CaM-Q135C sequence. The PCR reaction mixtures for both

reactions 1 and 2 were run on a 1.5% agarose gel and sample DNA for both products recovered by excision (using a 20 μL pipette tip) of small fragments of the gel DNA bands under UV light illumination. The recovered DNA was mixed with 5 μL of milliQ-H$_2$O. Reaction 3 brings both PCR DNA products together. The sequences overlap and so complementary strands will anneal at the mutation site. The result of the final PCR is an amplification of the complete 458 bp CaM-Q135C sequence. The primer sequences, PCR mixtures and reaction conditions used are detailed in Table 6.1 and Table 6.2.
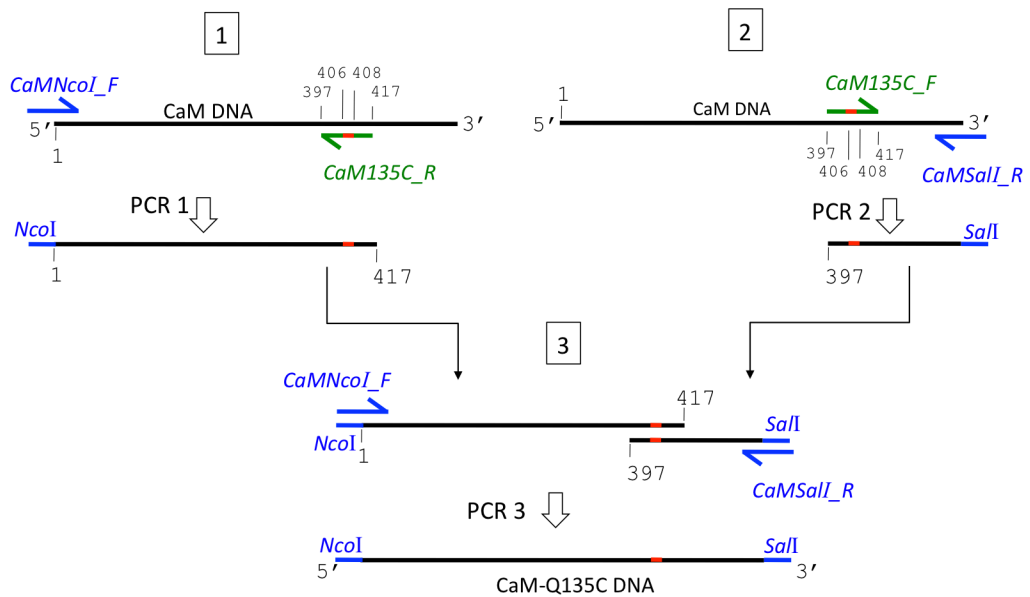
Figure 6.2: Three-reaction PCR scheme to produce CaM-Q135C cDNA

Template cDNA for the CaM sequence (black line) was used in reactions 1 and 2. In reaction 1, the forward primer *CaMNcoI_F* (blue) binds the 5' end of the template DNA and carries a flanking restriction site sequence for *NcoI*, while the reverse primer *CaM135C_R* (green) binds at the desired mutation site and introduces the base mutation for Q135C (shown in red). In reaction 2, the forward primer *CaM135C_F* (green) binds at the mutation site and introduces the same mutation (red), while the reverse primer *CaMSalI_R* (blue) adds the flanking restriction site sequence for *SalI* to the 3' end of the template DNA. In reaction 3, the complementary PCR products of reactions 1 and 2 are brought together and anneal, producing the final CaM-Q135C cDNA sequence with flanking restriction site sequences for *NcoI* and *SalI*.

Table 6.1: Primers sequences for CaM-Q135C construction

Primer	Sequence
CaMNcoI_F	5'-CCATGGCTGACCAACTGACAGAA-3'
CaM135C_F	5'-GGTGACGGCTGCGTAAACTAC-3'
CaM135C_R	5'-GTAGTTTACGCAGCCGTCACC-3'
CaMSalI_R	5'-GTCGACTCACTTTGCTGTCATCATT-3'

Table 6.2: PCR mixtures and conditions for CaM-Q135C reactions

Template DNA and primers used	*Components common to all reactions*
Reaction 1: 1 µL 10 µM linearised CaM-FB plasmid 1 µL 10 µM *CaM_NcoIF*, 1 µL 10 µM *CaM135C_R*. **Reaction 2**: 1 µL linearised CaM-FB plasmid (1/100 diln. stock), 1 µL 10 µM *CaM135C_F*, 1 µL 10 µM *CaMSalI_R*. **Reaction 3**: 5 µL reaction 1 PCR product, 5 µL reaction 2 PCR product, 1 µL 10 µM *CaM_NcoIF* 1 µL 10 µM *CaMSalI_R*	10 µL 5× GoTaq green reaction buffer, 1 µL 10 mM dNTPs, 35.75 µL MilliQ H_2O, 0.25 (5U/µL) GoTaq polymerase (added last).

Common heating cycle:
   1. 95° C for 2 mins
   2. 95° C for 30 s
   3. 56° C for 30 s
   4. 72° C for 30 s
   5. Repeat 2.-4. four times.
   6. 95° C for 30 s
   7. 56° C for 30 s
   8. 72° C for 30 s
   9. Repeat 6.-8. nineteen times
   10. 4° C hold indefinitely	

6.3.2.2 CaM-E139C PCR protocol

CaM-E139C DNA sequence was produced by a simpler single step PCR using two primers (Table 6.3), see Figure 6.3. The single step reaction was possible (as opposed to the three reactions required for CaM-Q135C) as the E139C mutation site is closer to the 3' end of the template DNA sequence compared with Q135C, such that a single long primer can be designed to bind at the 3' DNA end and run upstream to beyond the mutation site. The reverse primer, *CaM139C_R*, was designed for this purpose. It is 43 bp long and extends from the 3' template DNA end to 4 bp upstream of the mutation site. The primer also carries the additional sequence for a flanking *SalI* restriction site (blue line in Figure 6.3) and complements the template DNA sequence, except for a 3 bp deviation at bp 418-420 to introduce the desired mutation GAA→TCC (mutates glutamic acid to cysteine) (red in Figure 6.3) for E139C. The forward primer *CaMNcoI_F* (as used in the Q135C reaction 1) was the second primer in the PCR, and is used to anneal at the 5' DNA end and introduce a flanking *NcoI* restriction site. The entire mutated sequence including the flanking restriction site sequences is generated and amplified by the single PCR reaction. Primer sequences and reaction conditions are detailed in Table 6.3 and Table 6.4.
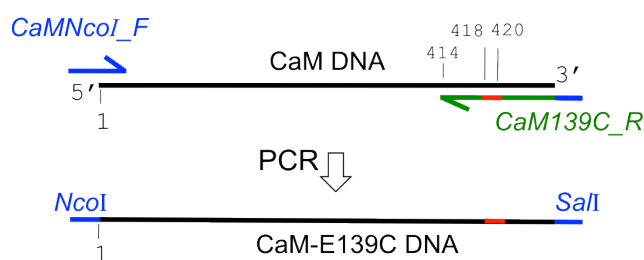
Figure 6.3: Single-reaction PCR scheme to produce CaM-E135C cDNA

Table 6.3: Primers sequences for CaM-E139C construction

Primer	Sequence
CaMNcoI_F	5'-CCATGGCTGACCAACTGACAGAA-3'
CaM139C_R	5'-GTCGACTCACTTTGCTGTCATCATTTGTACAAACTCGCAGTAG-3'

Table 6.4: PCR mixtures and conditions for CaM-E139C reaction

Reaction mixture	Heating cycle:
1 µL 10 µM linearised CaM-FB	1. 95° C for 2 mins
1 µL 10 µM *CaM_NcoIF*,	2. 95° C for 30 s
1 µL 10 µM *CaM139C_R*.	3. 56° C for 30 s
10 µL 5× GoTaq green reaction	4. 72° C for 30 s
buffer,	5. Repeat 2.-4. four times.
1 µL 10 mM dNTPs,	6. 95° C for 30 s
35.75 µL MilliQ H_2O,	7. 63° C for 30 s
0.25 (5U/µL) GoTaq polymerase	8. 72° C for 30 s
(added last).	9. Repeat 6.-8. nineteen times
	10. 4° C hold indefinitely

6.3.2.3 Cloning and isolation of plasmid DNA

PCR DNA for both CaM-Q135C and CaM-E139C were cloned into a pCR™2.1-TOPO® vector using a TOPO® TA cloning kit (Invitrogen – Life Technologies). The pCR™2.1-TOPO® plasmid is supplied pre-linearised with Topoisomerase I enzyme covalently bound to overhanging 3′ deoxythymidine (T) residues. The vector-bound Topoisomerase I will bind to any *Taq*-amplified PCR products which have 3′ deoxyadenosine (A) overhangs and thus ligate the PCR product into the plasmid. In each reaction, 1 µL of PCR product DNA was mixed with 0.5 μL of kit salt solution, 1 μL of milliQ H_2O and 0.5 μL of pCR™2.1-TOPO® vector and incubated at room temperature for 5 mins before storing at −20 °C.

Recombinant CaM-pCR™2.1-TOPO® plasmid DNA was then transformed into 5-*α* chemically competent *E. coli* cells (NEB) using a heat shock protocol: 1 µL of plasmid DNA was mixed with ~25 μL of thawed cells in a 15 mL falcon tube and incubated on ice for between 5-30 mins. The cells were shocked by submersion in a water bath at 42 °C for

45 s before returning to ice for 2 mins. 250 μL of LB (or S.O.C) mixture was added to each cell aliquot and the tubes transferred to a shaking incubator running at 220 RPM at 37 °C for 1 hour. 100 μL of the cell cultures were then spread onto LB-Agar plates containing kanamycin antibiotic (Bioline Reagents Ltd.) at 50 μg/mL final concentration and X-gal compound at 50 μg/mL, the latter added to allow blue/white colony screening. The plates were inverted and incubated overnight at 37 °C.

Positive (*i.e.* white-coloured) colonies were picked from the plates using a sterile pipette tip and used to seed 5 mL LB-kanamycin (0.05 mg/mL) volumes which were cultured overnight in a shaking incubator (220 RPM, 37 °C). Recombinant plasmid DNA was recovered and purified from the cultures by cell lysis and DNA adsorption to silica gel membrane using a QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany). Each recovered plasmid sample was then checked for the presence of a ~460 bp CaM-sized insert by restriction enzyme digest by *EcoRI* (5 μL of plasmid in 12 μL of milliQ-H$_2$0, 1 μL *EcoRI-HF* and 2 μL reaction buffer 4 (NEB), at 37 °C for 1 hrs), the digest result viewed by agarose gel electrophoresis. Samples containing the correct-sized insert were then checked by DNA sequencing (GATC Biotech AG, Konstanz, Germany).

Positively-sequenced insert DNA was transferred to the cloning site of a pET-28a(+) vector (Novagen® – Merck KGaA) to be used for protein overexpression in *E. coli* cells. CaM-Q135C and CaM-E139C insert DNA were first liberated from their pCR™-2.1-TOPO® vectors by restriction enzyme digest with *SalI* and *NcoI* (15 μL insert DNA, 26 μL milliQ-H$_2$O, 2 μL *NcoI* (Roche), 2 μL *SalI,* 5 μL buffer 3 (NEB), at 37 °C for 1 hrs) and the pET-28a(+) plasmid opened with the same restriction enzymes (5 μL plasmid DNA, 11.5 μL milliQ-H$_2$O, 0.75 μL *NcoI* (Roche), 0.75 μL *SalI* and 2 μL reaction buffer 3 (NEB), at 37 °C for 1 hrs). Digest results were verified by gel electrophoresis and the DNA for both vector and inserts were recovered by excising the bands from a gel and purifying with a GenElute Gel Extraction Kit (NA111) (Sigma-Aldrich).

Purified insert DNA and pET-28a(+) linearized vector were checked by gel electrophoresis before ligation using T4 DNA Ligase (NEB) (2 μL vector, 6.5 μL insert, 1 μL T4 DNA ligase and 0.5 μL 10x T4 DNA Ligase buffer (NEB), incubated on melting ice overnight). Finally, using the methods and materials as described above the re-ligated pET-28a(+)-CaM recombinant plasmids were transformed by heat shock into 5-α *E.coli* cells (NEB), the cells cultured in 250 μL media volumes, 100 μL of which were plated onto LB-Agar-kanamycin-Xgal and incubated overnight at 37 °C. Positive (white) colonies were picked and used to seed 5 mL LB-kanamycin cultures before recovery and purification of the plasmid DNA by QIAprep Spin Miniprep Kit. As a final verification check, samples of the

purified pET-28a(+)-CaM plasmids for both mutants were re-digested by *SalI* and *NcoI* restriction enzymes (as above) to verify insert presence and size by gel electrophoresis.

6.3.3 Overexpression

To begin overexpression of the CaM mutants the pET-28a(+)-CaM recombinant plasmids were each transformed into BL21-Gold (DE3) competent cells (Stratagene – Agilent Technologies, Santa Clara, USA) by the heat shock as described in 6.3.2.3. Cells were cultured in 250 μL SOC (1 hr, 37 °C, 220 RPM) before plating onto LB-Agar-kanamycin (50 μg/mL) for overnight incubation. Sample colonies were picked from the plates and used to seed 5 mL cultures to incubate overnight (37 °C, 220 RPM). The 5 mL cultures were then used to seed 50 mL LB-kanamycin overnight cultures which were in turn used to seed 0.5 L LB-kanamycin cultures for large scale CaM production (two 0.5 L cultures per CaM mutant). The 0.5 L cultures were incubated at 37 °C with 220 RPM shaking until their optical density at 600 nm (measured by spectrophotometer) reached a value of ~1.0 and bacterial growth was in an exponential phase. At this time point, 95 min since incubation, IPTG was added to each shaking culture to 1 mM final concentration, to induce overexpression of the CaM mutants. Shaking incubation with IPTG was continued for a further 4 hrs 30 mins before harvesting the cells by centrifugation (3500 g, 30 mins at 4 °C). Cell pellets were retained and immediately frozen at −80 °C for future use.

To proceed with purification, frozen cell pellets (two per mutant, each from a 0.5 L culture) were thawed on ice. Each was suspended in 10 mL of lysis buffer (2.4 M sucrose, 40 mM Tris-HCl, 10 mM EDTA, 1 mM PMSF, pH 8.0) by stirring and incubated on ice for 30 mins with frequent shaking. 60 mL of enzyme buffer (50 mM MOPS, 100 mM KCl, 1 mM EDTA, 1 mM DTT, pH 7.5) was added to each resuspended pellet and incubated on ice for 10 mins while stirring. Lysozyme (Sigma-Aldrich) was added to 100 μg/mL and stirring incubation continued for another 60 mins. 5 U of DNAse enzyme (amplification grade, Invitrogen – Life Technologies) and $MgCl_2$ to 10 mM was added to each (now pooled) mutant lysate, and stirring continued at room temperature for 30 mins. Lysate was then centrifuged (20 mins, 20,000 RPM, 4 °C, J-20 rotor, Beckman Avanti J-25 centrifuge). Supernatants were retained, PMSF added to 1 mM, and stored at −80 °C for further purification.

6.3.4 Purification

Fast protein liquid chromatography (FPLC) was used to separate the CaM mutant proteins from their clarified lysate contents. Following a well-established procedure (Dedman and Kaetzel, 1983; Gopolakrishna and Anderson, 1982) two ~10 mL columns packed with phenyl-sepharose resin (phenyl-sepharose 6 high flow low substitution, GE Healthcare

Life Sciences) were used. The first column is used as a first-pass filtration of the lysate: CaM in apo conformation (no Ca^{2+} bound) does not bind and appears in the flow-through. The second column is used to trap the CaM through hydrophobic interaction with the phenyl-sepharose. CaM is loaded in a Ca^{2+}-bound conformation and binds prior to the introduction of EDTA to re-chelate the Ca^{2+} causing the CaM to unbind. The two phenyl-sepharose columns were connected in turn to an ÄKTAprime FPLC system (GE Healthcare Life Sciences, Pittsburgh, USA) measuring ion concentration (conductivity) and protein concentration (280 nm absorbance) of the out flow. Further detail of this purification procedure is given below.

Clarified cell lysate was first thawed from storage and filtered through a 0.2 μm syringe filter. The filtered lysate was applied to phenyl-sepharose column 1, pre-equilibrated in buffer C (50 mM Tris-HCl, 1 mM EDTA, pH 7.4) and then washed with 100 mL of buffer C. $CaCl_2$ was added to 5 mM to the collected initial flow-through containing the CaM, allowing the CaM to bind Ca^{2+} ions and adopt its closed conformation. This mixture was then applied to phenyl-sepharose column 2, pre-equilibrated in calcium-containing buffer A (50 mM Tris-HCl, 1 mM $CaCl_2$, pH 7.5). Column 2 was washed with 50 mL buffer A and then with 50 mL high salt buffer B (50 mM Tris-HCl, 1 mM $CaCl_2$, 0.5 M NaCl, pH 7.5) to remove remaining non-CaM contaminants, before a final 50 mL wash in buffer A again. 50 mL buffer C was then applied to the column, releasing the CaM from the phenyl-sepharose resin. 5 mL elute fractions were collected until the 280 nm absorbance peak(s) of any protein elution had been collected.

Note, it was found that owing to the small column volumes used that there was significant (up to 70%) CaM-breakthrough during loading of column 2 as the phenyl-sepharose resin became oversaturated with bound CaM. It was found that the first flow-through and buffer A wash elutions for column 2 could in fact be run back over column 2 to harvest more CaM by the same Ca^{2+}-dependent binding and release protocol.

All FPLC flow-through, washes and elution fractions were kept for analysis by SDS-PAGE. Tris-HCl-based 15 % acrylamide gels (with 4 % stacking gel section) were polymerised specifically for CaM analysis. Gels samples were denatured by boiling in a modified Laemmli buffer (32 mM Tris-HCl, 1% SDS, 5% glycerol, 2.5% beta-mercaptoethanol, 3M urea and bromphenol blue) (Laemmli, 1970), and gels were run in a electrophoresis gel tank submerged in Tris-Glycine-SDS running buffer (0.025 M Tris, 0.192 M Glycine, 0.1% SDS), typically at 150 V for ~1 hrs. Since CaM has a molecular weight of ~17 kDa, a protocol for fixing and staining small peptides in the protein gel was adopted (Schägger, 2006). Gels were first placed for 40 mins in fixing solution (10%

acetic acid, 50% methanol, 100 mM ammonium acetate) then transferred to staining solution (0.025% Coomassie Brilliant Blue R-250, 10% acetic acid) for 80 mins, before placing in destain solution (10% acetic acid) for 60 min prior to analysis and scanning.

CaM concentrations in elution fractions identified (by SDS-PAGE) were quantified by UV-visible light absorbance, using a molar extinction coefficient at 276 nm of $3006 \text{ M}^{-1}\text{cm}^{-1}$ as detailed in 2.6.3. The reducing agent tris-(2-carboxyethyl)phosphine (TCEP) (Sigma-Aldrich) was then added to 1 mM to each CaM fraction to reduce any CaM-CaM disulfide bonds between the cysteine thiol groups on different CaMs. The reduced CaM fractions were then divided into 0.5 mL aliquots and frozen at −80 °C for storage.

6.3.5 Results

6.3.5.1 Mutagenesis

DNA sequences for both mutant CaMs were successfully produced by PCR. Figure 6.4 shows DNA gels of the PCR products produced by the three CaM-Q135C reactions and the single CaM-E139C reaction. The DNA fragments produced were of the expected size in each case.
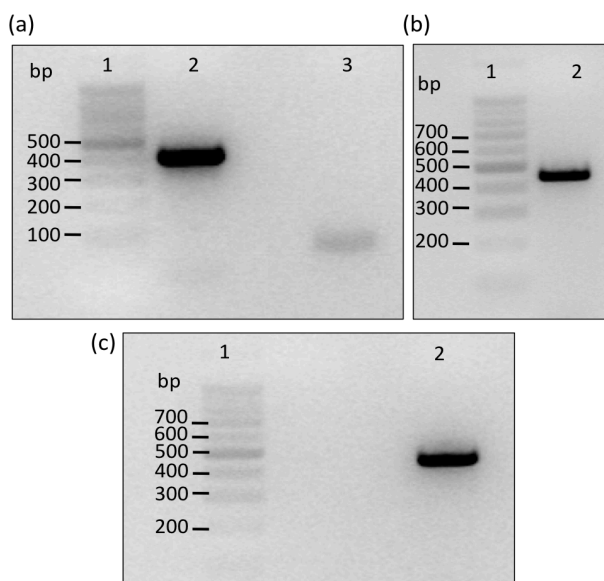
Figure 6.4: DNA gels of PCR reactions for CaM-Q135C and CaM-E139C

(a) DNA products for CaM-Q135C PCR reactions 1 and 2 (see Figure 6.2). Lane 1: MW marker, 2: PCR–1 expected to produce the first 419 bp of CaM-Q135C sequence, 3: PCR 2 expected to produce the last 60 bp of sequence (band is less intense, *i.e.* less concentrated DNA than lane 2). (b) PCR–3 expected to combine products of PCR–1 and –2 to produce the final ~460 bp DNA sequence for CaM–Q135C. MW marker in lane 1, PCR-3 result in lane 2. (c) PCR product (lane 2) for single step CaM-E139C reaction (see Figure 6.3). Band expected at ~460 bp is seen. MW marker–lane 1.

The mutant DNA sequences produced by PCR were successfully cloned into pCR™2.1-TOPO® vectors, and then successfully transferred to pET-28a(+) vectors. Figure 6.5 shows

DNA gels of verification restriction enzyme digests of both recombinant plasmid types, to check for the presence of a CaM-sized sequence in the cloning site. Both gels show that, for both mutant types, a CaM-sized fragment (~460 bp), was liberated from the plasmids by restriction digest (with the exception of one negative CaM-Q135C sample for the pCR™2.1-TOPO® digest – lane 3 of Figure 6.5(a)).
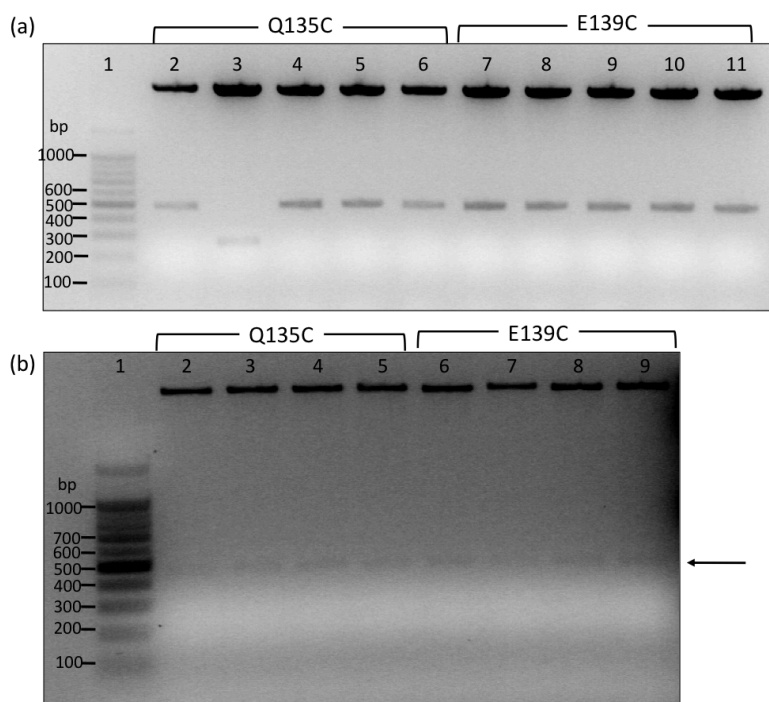
Figure 6.5: DNA gels of verification restriction enzyme digests of recombinant CaM-pCR™2.1-TOPO® and CaM-pET-28a(+) plasmids.

(a) Digest of pCR™2.1-TOPO®-CaM plasmids using *EcoRI* restriction enzyme. The digest is expected to liberate a ~460 bp CaM-sized insert if positive. Lane 1: MW marker, 2-6: CaM-Q135C samples, 7-11: CaM-E139C samples. All samples are positive for a ~460 bp band except for lane 3, which has a band at too small a size (~250 bp). The more intense high MW bands (>>1 kbp) are the linearised plasmid DNA minus insert. (b) Double restriction digest of pET-28a(+)-CaM plasmids using *NcoI* and *SalI* to liberate a CaM-sized insert. Lane 1: MW marker, 2-4: CaM-Q135C samples, 5-8: CaM-E139C samples. All lanes are positive for the insert, with a faint band visible at ~460 bp (arrow right) plus the high MW band for the linearised plasmid minus insert.

The DNA sequences for both CaM-Q135C and CaM-E139C from the pCR™2.1-TOPO® restriction digest were verified by nucleotide sequencing (GATC Biotech) before transfer to the pET-28a(+) vector. For each mutant, several clones were rejected due to sequence mutations. Clones with correct sequence were selected for each mutant for transfer to pET-28a(+) and subsequent expression.

6.3.5.2 Overexpression and purification results

To determine the optimum expression period following IPTG addition to the bacterial cultures and to verify whether the cell lysis procedure would yield sufficient CaM in the soluble lysate fraction, a pilot expression and lysis study was carried out on a 50 mL scale. The pilot expression and lysis was performed following the procedures described in 6.3.3, scaling-down reagent volumes accordingly. Figure 6.6 shows a time-course SDS-PAGE analysis of the soluble and insoluble lysate fractions for each mutant following IPTG induction.
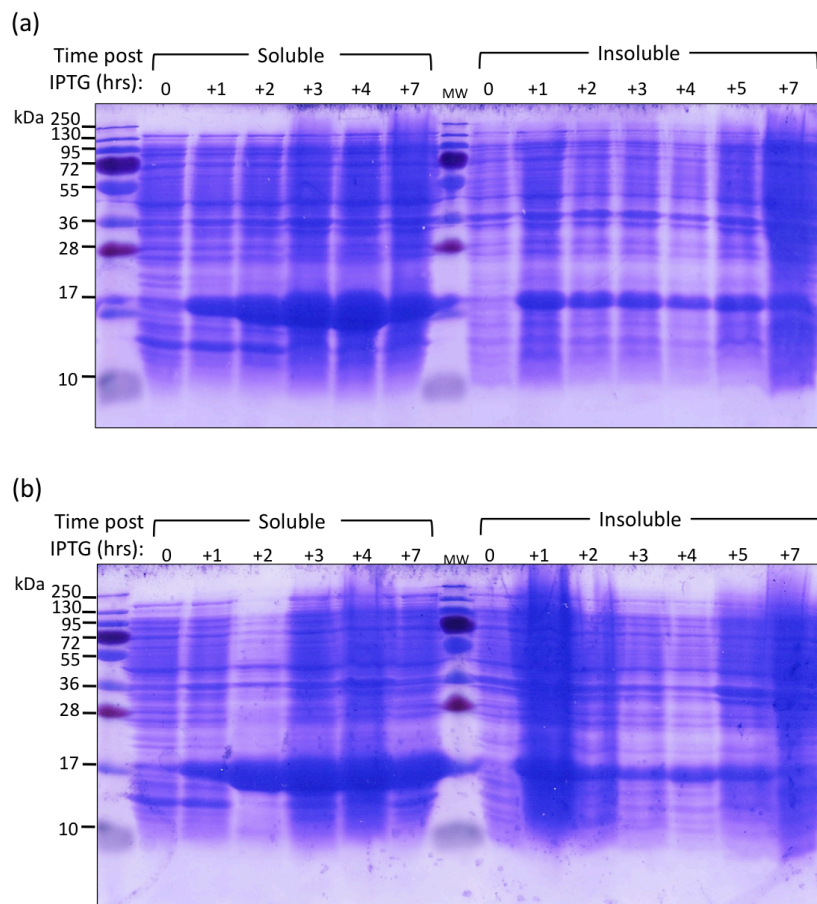
Figure 6.6: SDS-PAGE time course of 50 mL scale pilot expression of CaM mutants.

(a) CaM-Q135C samples, (b) CaM-E139C samples. Both gels are arranged with samples of the soluble fraction from cell lysate at increasing times after addition of IPTG in the first 6 lanes on the left and with the insoluble lysate fraction at the same time points in the last 6 lanes on the right. The numbers above the lanes indicate the time in hours after addition of IPTG to bacterial cultures. CaM is expected to run at ~17 kDa and is evident as an intense (wide) band growing in intensity with time in the soluble fraction. NB: to make a fair comparison between the soluble and insoluble fractions, insoluble material was always re-suspended in a volume so as to equal the volume of the corresponding soluble fraction before a gel sample was taken. Equal volumes of soluble/insoluble samples were then denatured and loaded to the gels.

The SDS-PAGE results for the pilot study indicate similar results for each mutant. In each case, a CaM-sized band (~17 kDa) is seen increasing in intensity with time in the soluble fraction. A CaM-sized band is also apparent in the insoluble fraction (at times after 1 hrs) but it is less intense than the soluble fraction bands and does not increase with time. Because sample volumes loaded to the gel were equalised (see caption Figure 6.6) the soluble fraction thus contains the majority of expressed CaM after lysis and clarification. The soluble bands peak in intensity between 4 and 7 hrs post induction (+7 hrs band is fainter than +4 hrs band). From this, an optimum (and practical) expression time of 4-5 hrs was determined for the large scale expression.

Purification of both mutant proteins from cell lysate via hydrophobic-interaction FPLC procedure was successful. An example FPLC trace and corresponding SDS-PAGE analysis of the purification steps for CaM-E139C is shown in Figure 6.7, while Figure 6.8 shows a corresponding SDS-PAGE analysis or the same purification procedure for CaM-Q135C. Together, Figure 6.7 and Figure 6.8 confirm the successful steps in the phenyl-sepharose Ca^{2+}-dependent binding procedure for both mutants. Firstly, apo CaM is seen passing through the first filtration column, see intense CaM-sized bands in flow-through lanes on SDS-PAGE gels. Secondly, a significant amount of Ca^{2+}-bound CaM is then held by the phenyl-sepharose of column 2, while low and high salt washes are applied to remove other contaminants. This is evidenced by the narrow single elution peak on the absorbance trace and also by the pure CaM-sized bands in gel lanes for the fractions eluted with EDTA.
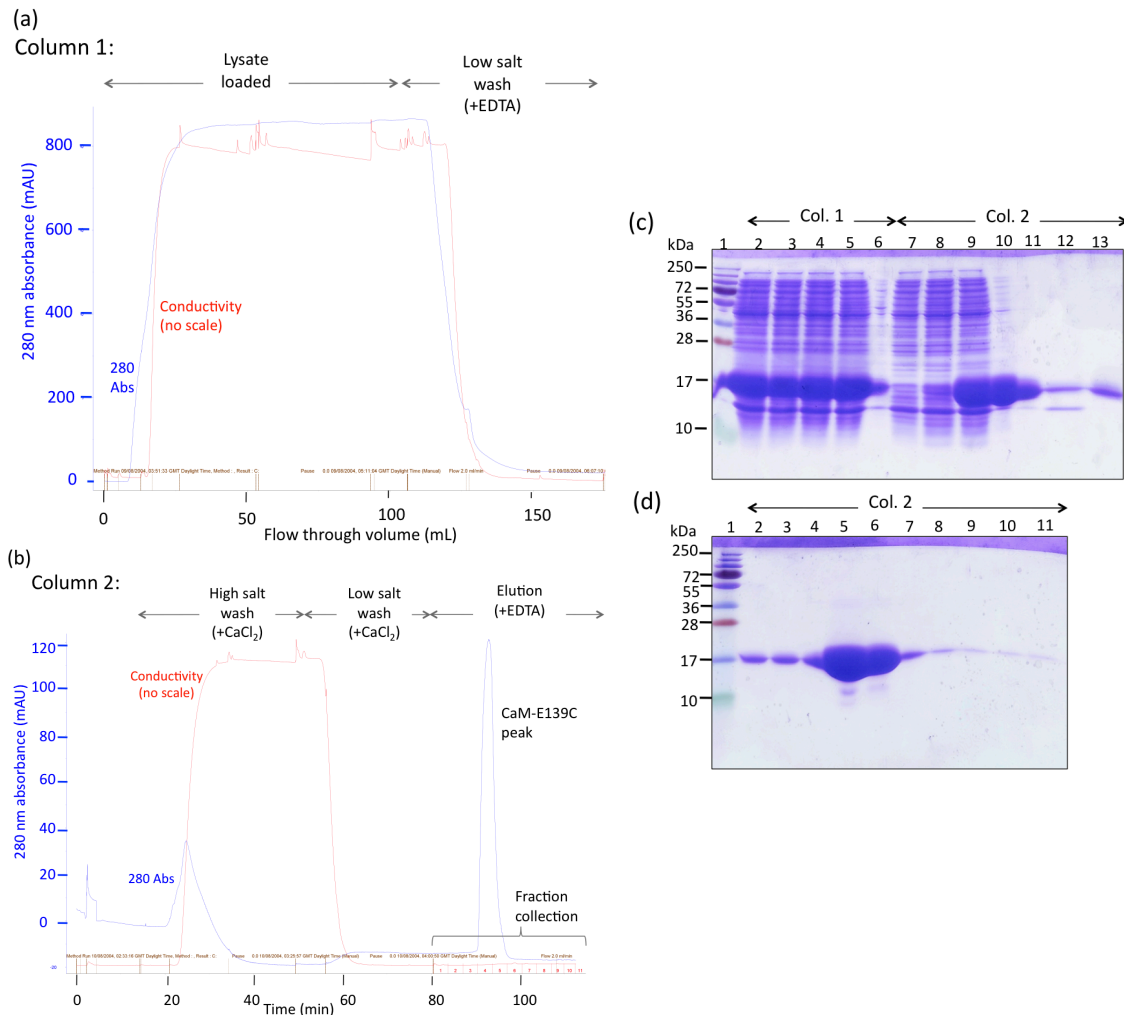
Figure 6.7: FPLC traces and SDS-PAGE for purification of CaM-E139.

(a) and (b) show FPLC measurements of absorbance at 280 nm (blue line) and conductivity (red line) of flow-through from phenyl-sepharose column 1 and column 2 in the CaM purification procedure. A sharp absorbance peak corresponding to the elution of CaM on addition of the EDTA-containing buffer can be seen at the end of the column 2 trace in (b). (c) and (d) are an SDS-PAGE analysis of samples corresponding to the stages of the purification. (c) lane 1: MW standard, 2: sample of clarified lysate applied to col. 1, 3-5: flow-through from lysate loading of col. 1, 6: low salt wash of col. 1, 7-9: flow-through from loading of col. 2, 10-11: low salt wash of col 2., 12: high salt wash of col 2., 13: low salt wash col 2. (d) lane 1: MW standard, 2-11: fractions collected from col 2. washing with EDTA-containing elution buffer.
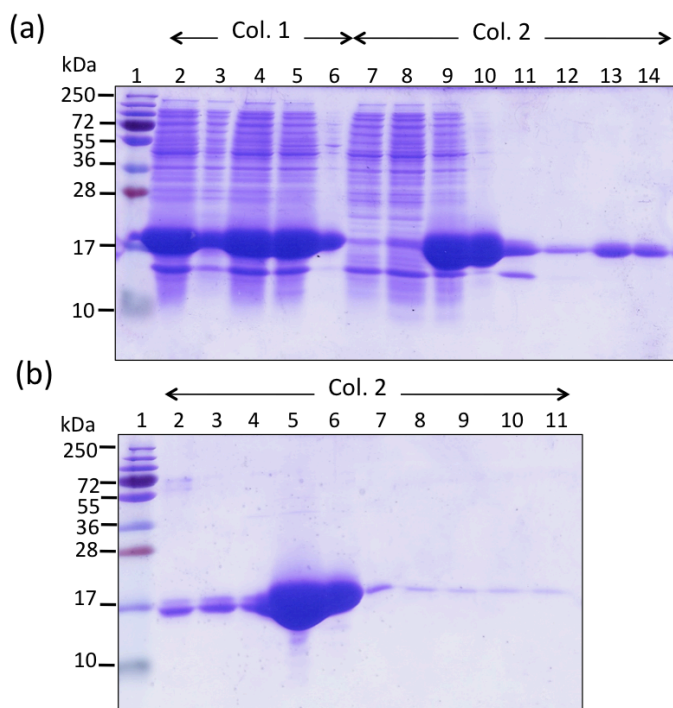
Figure 6.8: SDS-PAGE analysis of purification of CaM-Q135C

Samples from the purification procedure are shown in a similar layout to Figure 6.7(c)-(d). (a) lane 1: MW marker, 2: clarified lysate sample applied to col. 1, 3-5: col. 1 flow-through of lysate loading, 6: low salt wash col. 1, 7-9: lysate loading col. 2, 10: low salt wash col. 2, 11-12: high salt wash col. 2, 13-14: low salt wash col. 2. (b) lane 1: MW marker, 2-11: collected fractions from col. 2 elution.

Lanes 9 and 10 in the SDS gels in Figure 6.7(c) and Figure 6.8(a) also bring attention to the fact that significant amounts of CaM (estimated up to ~70%) breakthrough column 2 on loading, presumably due to oversaturation of the relatively small volume of phenyl-sepharose resin used. For both mutants, intense CaM bands are seen in the flow-through samples of the last lysate loading volume and subsequent low salt wash of column 2 (lanes 9-10 Figure 6.7(c) and Figure 6.8(a)), *prior* to application of the elution wash. This observation led to the reapplication of both the final loading and wash flow-through volumes to column 2, to harvest more of the unbound CaM via the same FPLC procedure.

Each FPLC purification run typically produced two concentrated 5 mL fractions during the final elution, evidenced by the intense CaM-sized bands in the SDS-PAGE analysis of the fractions (lanes 5-6 Figure 6.7(d) and Figure 6.8(b)). UV-visible light absorbance spectra of the fractions confirmed this and allowed concentration estimates to be made. Example spectra of fractions obtained by FPLC for each of the mutants can be seen in Figure 6.9. The concentrated fractions exhibit the distinctive absorbance profiles for CaM, with five absorbance peaks increasing in size at approximately 252, 259, 265, 269 and 276 nm, with a broad shoulder beginning at ~280 nm (Cheung, 1980). Typically, concentrations of 150-250 μM were obtained for the most concentrated fractions and 45-100 μM for less concentrated ones.
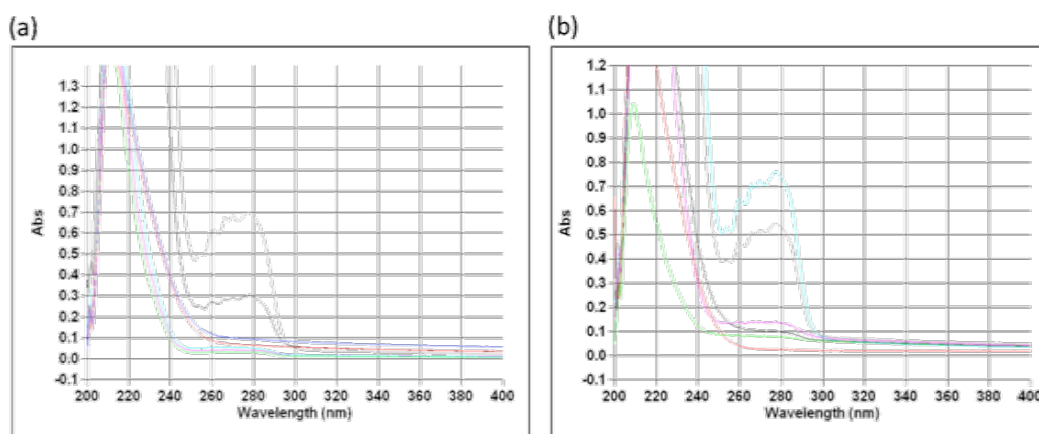
Figure 6.9: Example absorption spectra for purified CaM mutant fractions.

(a) CaM-Q135C fractions, (b) CaM-E139C fractions. Each purification run typically produced two concentrated fractions as seen in both mutant spectra. The concentrated fractions exhibit the distinctive absorbance profiles for CaM, with five absorbance peaks of increasing size at approximately 252, 259, 265, 269 and 276 nm, with a broad shoulder beginning at ~280 nm (Cheung, 1980).

In summary, construction of both CaM-Q135C and CaM-E139C clones was successful. Results from a pilot expression trial showed that the CaM mutants overexpressed well in *E. coli* with an optimum expression period of 4-5 hrs. Moreover, the soluble fraction of cell lysate contained the majority of CaM. Liquid chromatography purification of the CaM mutants by calcium-dependent binding to phenyl-sepharose was also successful, providing pure and concentrated yields of protein for use in future labelling work. The calcium-dependent binding to phenyl-sepharose is also a good indication that the CaM mutants are properly folded. For future purifications, use of a larger column volume of phenyl-sepharose (at least 20 mL) is recommended, to avoid oversaturation and the need for additional purification runs.

6.4 Fluorescent labelling of CaM mutants

Once successfully overexpressed and purified, labelling reactions were initiated to attach single fluorophores to the cysteine residues introduced into each CaM mutant.

6.4.1 Overview

A thiol-reactive label was added in excess to each mutant CaM sample, the protein having been prepared in a reduced state using TCEP to break disulfide bonds and make cysteine sulphydryl groups accessible. Labelling reactions were quenched with an excess of 2-Mercaptoethanol (β-ME) and excess un-reacted label removed by dialysis. Success of labelling was checked with ionization time-of-flight mass spectrometry and quantified by absorbance spectrophotometry.

The following sections cover the selection of an appropriate label, the protocols adopted for conjugation reaction and techniques used to quantify the success and degree of labelling. Methods and materials are discussed first with corresponding results sections thereafter.

6.4.2 Selection of thiol-reactive label

Alexa Fluor® 488 C_5 maleimide (A488M) (Invitrogen – Life Technologies) was selected as the thiol-reactive label to use, Figure 6.10(a). The maleimide group is known to react well at physiological pH with accessible sulphydryl groups (such as in cysteine), attaching through a thioether bond in the conjugation reaction, Figure 6.10(b). Maleimides are similar to iodoacetamides as thiol-reactive reagents but are thought to be more thiol-selective as they do not react with histidine or methionine, an advantage for site-specific cysteine labelling (Life Technologies and Haugland, 2010). The Alexa Fluor® 488 green dye has similar excitation and emission spectra (in the 400-600 nm range) to fluorescein, making it suitable for fluorescence microscopy using standard FITC filters. It offers the advantages that it is a brighter, more photostable and self-quenches less readily that fluorescein (Panchuk-Voloshina et al., 1999), making it preferable for TIRF assays of myosin movement. The $(CH_2)_5$ carbon backbone linking dye and maleimide was chosen to be as short as possible to anchor the fluorophore as closely as possible to the CaM labelling site. A five carbon backbone was the shortest available from the vendor.

(a)

(b)
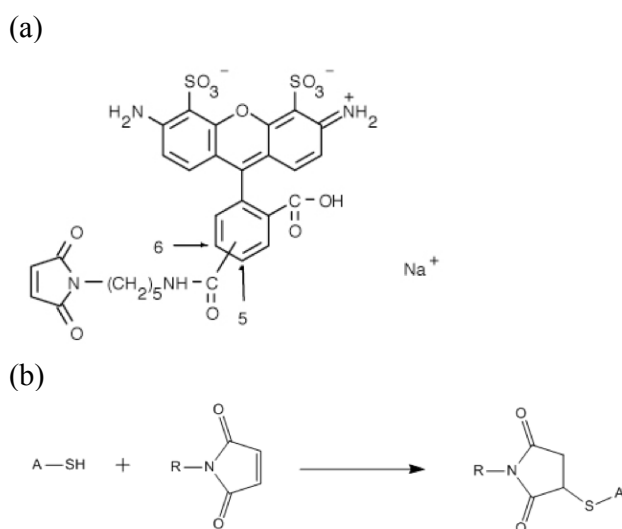
Figure 6.10: Structure of AlexaFluor 488 C_5-maleimide and maleimide reaction scheme.

(a) AlexaFluor 488 C_5-maleimide structure (Invitrogen – Life Technologies). The maleimide reactive group (lower left) is attached to the 4-ring Alexa fluor 488 fluorescent dye via a $(CH_2)_5$ linker. (b) Maleimide-SH group reaction scheme. An accessible protein side chain (A) bearing an SH group (*e.g.* cysteine) will react with the maleimide to form a bond with the sulphur atom, thus attaching itself and its linked fluorescent dye (R) to the protein.

6.4.3 CaM-Alexa Fluor® 488 C$_5$ conjugation reaction

A 6 mM stock solution of A488M was made by dissolving 1 mg of the lyophilized solid in 232 μL of dimethyl sulfoxide (DMSO). The stock was kept covered in foil to protect the dye from light and stored at −20 °C when not in use. A reference absorption spectrum (200-700 nm) of a 3 μM AF488M solution in CaM buffer C (6.3.4) was taken using a spectrophotometer, both to verify the A488M stock concentration from its 495 nm absorption peak, using an molar absorption coefficient at 495 nm of 71,000 M^{-1}cm^{-1} (Life Technologies and Haugland, 2010), and to obtain the relative A488M absorption in the 270-290 nm region. The latter was done to calculate a correction factor for the dye absorption contribution at 276 and 280 nm in subsequent AF488M-labelled CaM and myosin 5a-HMM absorbance spectra.

The conjugation reaction was initiated by adding a small volume of the reactive dye to thawed samples of each CaM mutant, the dye added to make a 10-fold molar excess (typically 0.5 mM) over the CaM concentration (typically 50 μM) in a 200-350 μL reaction volume. The reaction was quickly vortexed, the tube covered in foil to protect from light and either left to incubate for 2 hrs at room temperature or overnight at 4 °C. Importantly, the CaM mutant samples used in the reactions already contained 1 mM TCEP that had been added to reduce any disulfide bonds between cysteines, to make them accessible to the reactive dye. Use of TCEP over other low MW reducing agents such as DTT was advantageous as TCEP is not itself reactive with maleimides and so does not compete with the protein in the reaction and does not need to be removed beforehand. The conjugation reaction was quenched by addition of a 5-fold molar excess of β-ME (typically 2.5 mM) over A488M.

The quenched reaction volumes were transferred to dialysis tubes (Novagen D-Tube™ Midi 3-6 kDa MWCO dialyzer tubes, Novagen – Merck KGaA) and extensively dialysed at 4 °C against CaM buffer C, to remove quenched free dye and excess β-Me. Dialysis beakers were kept covered in foil to protect the dye at all times. A minimum dialysis period of 48 hrs against 500 mL buffer volumes (refreshed ≥ 4 times) was established by monitoring absorption spectra of both reaction and dialysis beaker contents of trial reactions. Indeed, where practicable the dialysis was left for even longer (up to 4.5 days) to maximize the chance of all free dye being removed.

6.4.4 Verification and quantification of labelling

To verify if the A488M label successfully attached to the CaM mutants, samples were analysed by electrospray ionization time-of-flight mass spectrometry (MS) by Dr. James Ault (Mass Spectrometry Facility, University of Leeds). CaM-Q135C and CaM-E139C

samples both before and after labelling, as well as an unlabelled wild type CaM control (a gift from Dr. Melanie Colegrave, University of Leeds) were first prepared for analysis by overnight dialysis in 50 mM ammonium acetate (pH 7.5). Samples were analysed by Z-spray nanoelectrospray ionisation MS using a quadrupole-IMS-orthogonal time-of-flight mass spectrometer (Synapt HDMS, Waters MS Technology Centre, Manchester, UK), using gold/palladium-coated nanospray needles prepared in-house. Molecular weights of any species present in the samples were obtained from analysis of the mass/charge spectra using MassLynx software v4.1 (Waters, Milford Massachusetts, USA).

To quantify the degree of labelling achieved in the conjugation reaction (*i.e.* the fraction of CaM that acquired a label) absorption spectrophotometry was used. A UV-visible light absorption spectrum (200-600 nm) was measured for each sample. From this, the dye concentration was determined from the 495 nm absorption value (A_{495}). The dye absorption contribution at 276 nm was calculated from a 276 nm correction factor (CF_{276}) (obtained from the ratio A_{276}/A_{495} in the free dye reference spectrum) multiplied by A_{495}. The CaM concentration was calculated from the 276 nm absorption (A_{276}) corrected by subtraction of the calculated dye contribution. The final degree of labelling (D.O.L) expressed as the dye-per-CaM ratio was calculated from the formula:

$$\text{D.O.L} = \frac{[\text{AF488M}]}{[\text{CaM}]} = \frac{A_{495}/\varepsilon_{\text{dye},495}\cdot l}{(A_{276} - (CF_{276}\cdot A_{495}))/\varepsilon_{\text{CaM},276}\cdot l}, \qquad \text{(Eq. 6.1)}$$

where $\varepsilon_{\text{dye},495}$ is the 495 nm extinction coefficient for the reactive dye, $\varepsilon_{\text{cam},276}$ is the CaM 276 nm extinction coefficient and l is the absorption path length, in this case 1 cm.

6.4.5 Results

Mass spectrometry (MS) analysis of the dialysed CaM-A488M reactions showed a measure of success in the labelling reactions for both mutants. Deconvoluted mass spectra for both mutant samples showed two species to be present, in each case one with a MW corresponding to the unlabelled mutant and the other with a MW corresponding to the mutant CaM plus the A488M label, Figure 6.11. The measured masses of labelled and unlabelled mutants were accurate to within 0.5 Da of their expected values, Table 6.5.

The precision of the mass spectrometry technique not only demonstrates successful attachment of the A488M label, but also confirms the exact MW's of the unlabelled CaM mutants (see Table 6.5 values and plots in Figure 6.11). The different single amino acid change to cysteine in the two mutants gives each a slightly different protein mass, both differing from that of wild type CaM. All the masses obtained from the mass spectrometry

are accurate to within 1 Da of the expected values and thus offer a further verification that the overexpressed and purified mutant CaMs have the correct amino acid composition.
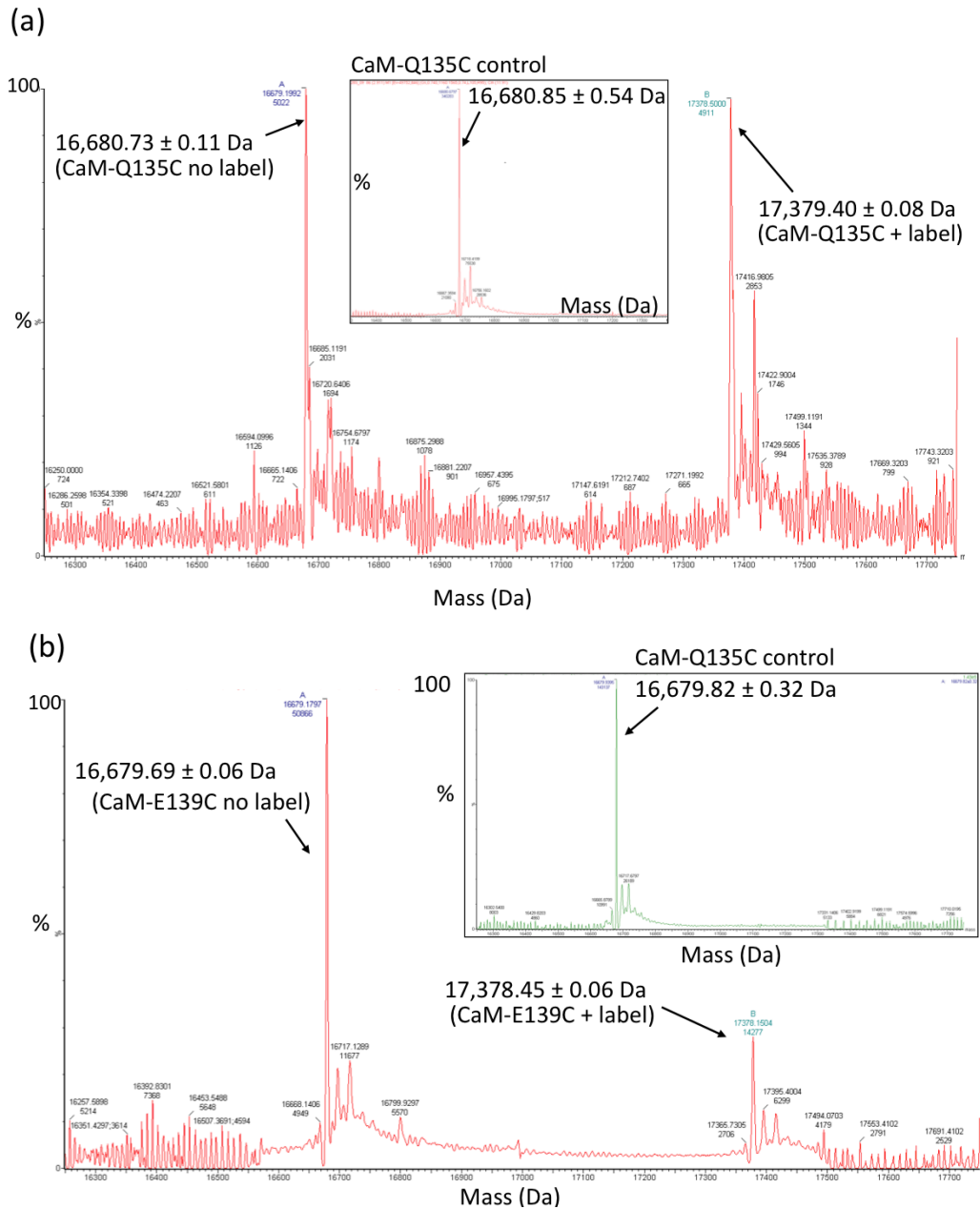
Figure 6.11: Deconvoluted mass spectra for A488M-labelled CaM mutants from ionizing time-of-flight mass spectrometry.

(a) deconvoluted mass spectrum for labelled CaM-Q135C sample. Two peaks corresponding to masses of unlabelled and labelled CaM-Q135C (in 1:1 ratio) are seen. The inset graph is a mass spectrum for an unlabelled CaM-Q135C control sample, in which, as expected, only one mass species is seen. **(b)** deconvoluted mass spectrum for labelled CaM-E139C sample, inset graph is an unlabelled control. Two mass species corresponding to unlabelled and labelled species are seen in the main plot. The ratio of unlabelled:labelled peak sizes is 3:1, indicating that labelled CaM accounts for only ~25% of this E139C sample.

NB: the mutant CaM samples measured by MS were prepared from A488M labelling reactions quenched after 2 hrs 40 min, and dialysed for ~48 hrs to remove unreacted dye.

Table 6.5: Expected and measured MW's of labelled CaM mutants.

	Expected MW[§] (Da)	MW measured by MS (Da)
AF488-C$_5$-M	697.7	Not measured
CaM wild type	16706.3	16705.96 ± 0.29
CaM-Q135C (unlabelled)	16681.4	16680.85 ± 0.54
CaM-E139C (unlabelled)	16680.4	16679.82 ± 0.54
[CaM-Q135C]-[AF488-C$_5$-M]	17379.1	17379.40 ± 0.08
[CaM-E139C]-[AF488-C$_5$-M]	17378.1	17378.45 ± 0.09

[§] CaM MW's are calculated from the amino acid sequence, minus the first post-translationally cleaved methionine.

Also of note in the mass spectra in Figure 6.11 are the relative intensities of the labelled and unlabelled species peaks. In the case of CaM-Q135C the unlabelled and labelled peaks are of similar intensity, while in the CaM-E139C spectrum the unlabelled peak is approximately three times the labelled species. If the relative intensity of these peaks is directly proportional to the species abundance (and this is constant across the spectral range), then these spectra indicate that the A488M conjugation reactions were only partially successful, in the case of E139C-CaM leaving 75 % of sample unlabelled (D.O.L = 0.25) and in the Q135C-CaM case leaving 50 % unlabelled (D.O.L = 0.50). The partial labelling success for these particular samples is confirmed by similar D.O.L values of < 1 obtained from separate absorbance spectra measurements, Table 6.6. It is noted that the absorbance-calculated D.O.L values are ~0.25 higher than those calculated from MS. Since the absorbance measurements were made first (after only 29 hrs of dialysis) while the MS samples underwent a further 19 hrs dialysis (to make an additional buffer change), the higher absorbance-calculated values are likely attributable to the continued presence of unreacted free dye at the time of absorbance measurement, rather than indicating any greater fraction of CaM labelling. Indeed, the later MS analysis did not detect any peak around 698 Da (the expected MW of A488M) to indicate the presence of free dye in either sample. This suggested that dialysis for at least 48 hrs (with regular buffer changes) is required for proper separation of unreacted dye.

Figure 6.11 shows example absorption spectra for each CaM mutant after an overnight labelling reaction plus 29 hrs dialysis. Table 6.6 is a summary of CaM concentrations and D.O.L values obtained from all labelling reactions conducted. In the later reactions, D.O.L values closer to 1.0 (*i.e.* 100 % labelling) were obtained. Values > 1.0 suggest the continued presence of free dye (since there cannot be more than one label per CaM). Some D.O.L variability from preparation to preparation is seen, though this is probably within the sensitivity of the absorbance measurement technique.
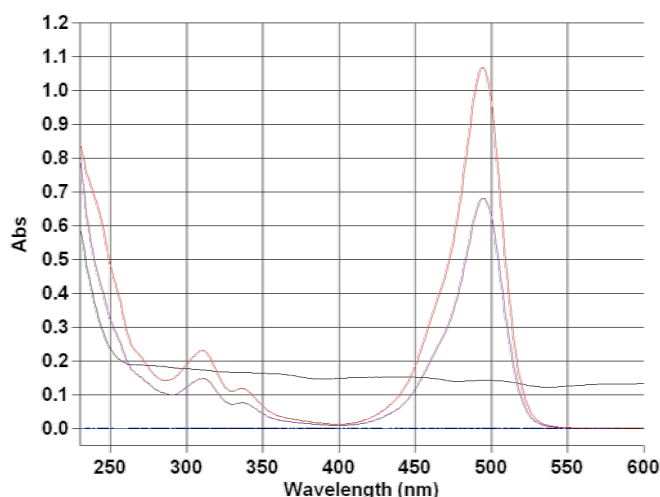
Figure 6.12: Absorbance spectra for CaM mutants after an overnight A488M labelling reaction.

Labelled CaM-Q135C spectrum (red) and CaM-E139C (purple). The black line is the spectrum for the buffer used to blank the spectrophotometer prior to measurements. The samples shown here were labelled by A488M in overnight reactions (quenched after 20 hrs 30 min) and dialysed for 29 hrs thereafter. 1/5 dilutions of the reaction samples were made for measurement of the spectra shown. The large peak at 495 nm is due to the AF488 dye and enables a dye concentration estimate to be made. Absorption in the 270-320 nm region is due to CaM, but there is also a contribution from the AF488 dye (calculated as $A_{495} \times 0.109$, the latter correction factor obtained from a reference A488M spectrum, not shown).

Table 6.6: List of CaM-AF488M labelling reactions and outcomes.

CaM-AF488M reaction (duration)	Dialysis time post reaction	Final concentrations measured (µM)		Degree of CaM labelling (D.O.L) (AF488M/CaM)
		CaM	AF488	
Q135C-AF488M (2 hrs 40 min)	42 hrs	86	45	0.52
E139C-AF488M (2 hrs 40 min)	42 hrs	45	23	0.51
Q135C-AF488M (20 hrs 30 min)	29 hrs	97	75	0.77 (~0.50 by MS)
E139C-AF488M (20 hrs 30 min)	29 hrs	90	48	0.53 (~0.25 by MS)
Q135C-AF488M (Overnight)	5 days	23	26	1.13
E139C-AF488M (Overnight)	5 days	46	59	1.28
Q135C-AF488M (Overnight)	7 days	41	29	0.71
E139C-AF488M (Overnight)	7 days	54	59	1.09

6.5 Exchange of labelled CaM into myosin 5a-HMM molecules

6.5.1 Overview

Endogenous CaMs were replaced by labelled CaM molecules on wild-type myosin 5a-HMM molecules using a protocol devised specifically for this work. This process of swapping in labelled for endogenous CaMs is called CaM exchange. The exchange protocol used was developed from previous methods (Beausang et al., 2007; Churchman et al., 2005; Sakamoto et al., 2000) and utilised calcium-induced unbinding of CaM from myosin 5a (Krementsov, 2004; Lu et al., 2006; Nascimento et al., 1996; Nguyen and Higuchi, 2005; Sellers et al., 2008) to facilitate the labelled CaM exchange. In outline, the protocol, discussed in more detail in the following sections, involved incubation of myosin 5a-HMM molecules with an excess of labelled CaM. $CaCl_2$ was introduced to cause unbinding of endogenous CaMs from myosin 5a-HMM and to allow exchange with labelled CaMs to occur. The exchange was stopped by addition of excess EGTA (reducing the free calcium concentration) and the myosin 5a-HMM molecules (now bearing a mix of labelled and unlabelled CaMs) were separated from free CaM using a low-binding cellulose membrane centrifugation filter. The degree of labelled CaM exchange was quantified using absorption spectroscopy and the enzymatic activity of the CaM-labelled myosin 5a-HMM molecules were compared to a wild-type control in an actin-activated ATPase assay.

6.5.2 Exchange protocol

The exchange reaction was set-up by mixing excess labelled CaM with wild-type myosin 5a-HMM molecules in exchange buffer (250 mM KCl, 20 mM Imidazole-HCl, 2 mM $MgCl_2$, 1 mM EGTA, 5 mM DTT, pH 7.6). Prior to the reaction, labelled CaM-Q135C and CaM-E139C samples (thawed from −80 °C storage) were in CaM buffer C (6.3.4) and myosin 5a-HMM samples (thawed beads from liquid nitrogen storage) were in 0.5 M KCl HMM buffer (5.3.3.6). Labelled CaM was added to the myosin 5a-HMM and incubated for 5 mins at room temperature. CaM:myosin 5a-HMM ratios from 6:1 to 20:1 were tried for different reactions, with myosin 5a-HMM typically diluted to 0.4-1.0 μM (two-headed molecule concentration) in a 150-250 μL reaction volume. Exchange was initiated by addition of $CaCl_2$ to 1 mM final concentration (~10 μL of 25 mM $CaCl_2$ added depending on reaction volume), the reaction vortexed and then incubated at room temperature for 5-10 mins. Exchange was terminated by addition of EGTA to 4 mM (~10 μL of 100 mM EGTA added depending on reaction volume), the reaction vortexed and then incubated for a further 10 mins at room temperature, before placing on ice for ≥ 2 mins.

CaM-exchanged myosin 5a-HMM molecules were separated from free CaM by spin column centrifugation. Reaction volumes were transferred to Amicon Ultra-0.5 mL centrifugation filters with 100 kDa MWCO low-binding cellulose membrane filters (Merk Millipore, Billerica MA, USA). The filter membranes had been pre-washed with exchange buffer prior to adding the exchange reaction solution. Each filter column containing the exchange reaction was centrifuged at 5000 g for 10 mins at 4 °C (Eppendorf centrifuge 5417 R, 11,000 RPM). Filtrate volumes were pipette-measured to calculate the retained volumes remaining in columns and the spin columns were refreshed with corresponding new volumes of the desired experiment buffer (usually motility buffer (6.6.2) for TIRF experiments). The new column solutions were gently agitated by pipette (taking care not to touch the filter) to resuspend any membrane-adhered protein, and the 5000 g / 4 °C centrifuge-and-replenish routine just described was repeated a further 2-3 times (*i.e.* the myosin 5a-HMM-CaM solution was centrifuge-filtered 3-4 times in total). The final retained volumes were recovered from the filters (by inverting the spin columns into fresh tubes and centrifuging at 3000 RPM for 2 min at 4 °C (Eppendorf centrifuge 5417 R)). The recovered volumes (typically ~25 μL) were placed on ice and diluted with final experiment buffer to ~50-100 μL for use in experiments.

The degree of labelled CaM exchange onto myosin 5a-HMM molecules was estimated by absorption measurement. Since there can be only one dye per CaM, the concentration of labelled CaM was determined via measurement of the 495 nm dye absorption (A_{495}). The concentration of myosin 5a-HMM molecules (two-headed molecule concentration) was determined from the 280 nm absorption value (A_{280}). The A_{280} value was first corrected by subtraction of the dye's absorption contribution at 280 nm (calculated from A_{495} multiplied by the ratio A_{280}/A_{495} (CF_{280}) determined previously from a free dye reference spectrum). No additional correction for light scattering was made as absorbance values were found to be very low in the dilute samples measured, Table 6.7. The myosin 5a-HMM concentration was then calculated from the adjusted A_{280} value using an extinction coefficient per HMM head ($\varepsilon_{hmm-head,280}$) of 0.12 $\mu M^{-1}cm^{-1}$ (Forgacs et al., 2009). The final two-headed myosin 5a-HMM concentration was then calculated as ½ the HMM head concentration.

$$[\text{AF488M}] = \frac{A_{495}}{\varepsilon_{dye,495} \cdot l}. \qquad\qquad \text{(Eq. 6.2)}$$

$$[\text{myosin 5a HMM}] = \frac{A_{280} - (CF_{280} \cdot A_{495})}{2 \cdot \varepsilon_{hmm-head,280} \cdot l}, \qquad\qquad \text{(Eq. 6.3)}$$

where l is the absorption path length, here 1 cm. The degree of CaM exchange, effectively the number of labels per myosin 5a-HMM, could be calculated as a simple ratio of the two concentrations.

$$\text{Labelled CaMs per HMM} \; = \frac{[\text{AF488M}]}{[\text{myosin 5a HMM}]}. \qquad \text{(Eq. 6.4)}$$

An assumption made here is that all free labelled CaM has been removed from the sample by centrifugation filtering and that the A_{495} value is due only to labelled CaM bound to myosin 5a-HMM.

6.5.3   Results

Absorbance spectra of CaM-exchanged spin-filtered myosin 5a-HMM samples indicated that both the labelled CaM types could be successfully exchanged into wild-type myosin 5a-HMM molecules, with varying success from sample-to-sample under differing CaM:myosin 5a-HMM reaction ratios. Figure 6.13 is an example of exchanged myosin 5a-HMM absorbance spectra from which the degree of labelled CaM exchange was estimated. Table 6.7 summarises the absorbance spectrophotometry measurements and exchange estimates made for a number of exchange attempts, using both CaM mutants.

It is clear from the example spectra in Figure 6.13 that the measured absorbance values were very low (a result of necessary stock dilutions to fill 100 µL spectrophotometry cuvettes). The AlexaFluor® 488 dye has the strongest absorbance characteristic at 495 nm, but reached only 0.07 across the samples measured, corresponding to a maximum dye concentration of ~1 µM in the cuvette dilutions. 280 nm absorbance values were of similar order (0.00-0.07) equating to only ~0.0-0.3 µM myosin 5a-HMM in the dilutions after dye absorbance correction. Ideally the sample concentration should be 10-fold greater (yielding absorbance values of order 1.0) for more reliable estimate of CaM-exchange. The spectra also suffer from a significant signal fluctuation (S.D ~0.4 Abs.), possibly owing to a spectrophotometer set-up fault (*e.g.* misalignment of the beam with cuvette). The baseline set by the blank buffer will also have a similar fluctuation, and since the sample absorbance readings are obtained by subtraction of the fluctuating baseline from the fluctuating sample signal, the relative uncertainty in the final measurement is increased further (uncertainties add in quadrature). The uncertainties given in Table 6.7 are estimated by the standard error propagation formula, where the error in sample absorbance and baseline were considered and taken as the S.D of the fluctuations sample signal across the 550-600 nm range. Owing to the dilute samples and signal fluctuations, the exchange values calculated in Table 6.7 can only, at best, be viewed as rough estimations.
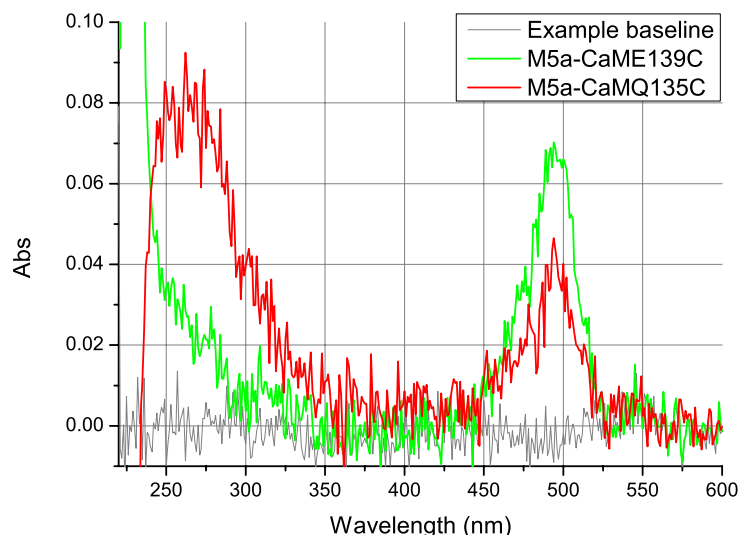
Figure 6.13: Example absorbance spectra for myosin 5a-HMM after exchange reaction with different A488M-labelled CaMs.

Green plot is myosin 5a-HMM exchanged with labelled CaM-E139C, red is myosin 5a-HMM exchanged with labelled CaM-Q135C. Measurements were made on 1/6 dilutions of protein stocks. Grey plot is an example measurement of the blank buffer solution used to set the baseline. Note the absorbance values are very low (vertical axis scale), ~1/10 ideal values for measurement, indicating the retained protein after exchange is dilute. Moreover, each line (including the baseline) is subject to a significant fluctuation (SD ~0.004 Abs), perhaps due to a spectrometer set-up fault, making the relative uncertainty in the low A_{280} measurement high. The lower A_{280} value for Myosin 5a-CaM-E139C increases its relative uncertainty. Presence of the AF488 dye is indicated by the peaks seen at 495 nm.

Table 6.7: Summary of AF488-labelled CaM exchange reactions with myosin 5a-HMM

AF488-CaM type exchanged	Exchange reaction CaM:HMM ratio (µM)	Abs readings			Stock concentration post exchange (µM)		Degree of CaM exchange (label/HMM)
		A_{495}	A_{280}	S.D	HMM	Label	
Q135C	20:1	0.045	0.0723	0.004	1.7 ± 0.1	3.8 ± 0.3	**2.3 ± 0.2**
E139C	20:1	0.0703	0.025	0.003	0.4 ± 0.1	5.9 ± 0.3	**14.8 ± 2.9**
Q135C	10:1	0.026	0.016	0.004	0.6 ± 0.1	3.9 ± 0.3	**6.4 ± 2.4**
E139C	10:1	0.0142	N.M	0.004	N.D	1.4 ± 0.3	**N.D**
E139C	6:1	N.M	N.M	N.M	N.D	N.D	**N.D**
Q135C	10:1	0.007	0.034	0.004	0.1 ± 0.1	0.1 ± 0.3	**0.7 ± 0.4**

S.D: standard deviation of baseline absorbance measurements.
N.M: not measured.
N.D: not determined.

Despite these qualifications on the data, Table 6.7 values do at least suggest that both labelled CaM mutants exchanged to some degree into wild-type myosin 5a-HMM. A 10:1 CaM:myosin 5a-HMM ratio appears to be sufficient for the reaction, as halving the excess CaM concentration from 20:1 did not result in significantly less exchange. In fact, for

275

CaM-Q135C the estimated exchange went up from 2.3 CaMs/HMM to 6.4 CaMs/HMM when decreasing the CaM excess. Interestingly, CaM-E139C seemed to exchange much more readily than CaM-Q135C at the 20:1 exchange ratio, resulting in an apparent average of 14.8 ± 2.9 labelled CaMs per myosin 5a-HMM. Since there can be a maximum of only 12 CaMs exchanged per HMM molecule (6 on each lever) the 14.7 calculated for CaM-E139 must be an overestimate of the true exchange. This may be encompassed by the high relative uncertainty, but may also indicate the continued presence of excess labelled CaM not removed by the centrifugation filtering. *NB*: no labelled CaM-only controls were measured to test if the centrifugation efficiently removed all free CaM from the myosin 5a-HMM solution. In future work, such controls should be performed.

In short, the spectrophotometry measurements suggest that some CaM-exchange occurred for both mutants, but as a means of quantifying the degree of exchange the method was in this case not very reliable, given the dilute and small volumes of protein recovered from the centrifugation filtering. Nevertheless, the indication of some degree of CaM exchange was sufficient to pursue a TIRF motility analysis of the samples. In the TIRF assay, only small amounts of very dilute labelled protein are required for study at a single-molecule level.

6.5.3.1   Actin-activated ATPase comparison of labelled myosin 5a HMM

Before proceeding with the TIRF motility assay, a test of the enzymatic function of myosin 5a-HMM bearing each type of labelled CaM mutant was made, to ascertain if the AF488M-labelled CaMs affected the ATPase of the wild-type two-headed molecule in the presence of F-actin (and so by inference perhaps the ability of the myosin to walk along F-actin). A single actin-activated ATPase assay was performed to compare a wild-type myosin 5a-HMM sample that had not been subject to a CaM exchange with myosin 5a-HMM samples that had received either labelled CaM-E139C or CaM-Q135C through exchange reaction. The actin-activated ATPase assay was the NADH-coupled assay as described in 2.7. In the assay, actin was set at a saturating final concentration, 20 μM, the myosin 5a-HMM final concentrations (of two-headed molecules) were 0.1 μM, ATP 2 mM and KCl 75 mM. 100-200 μL reaction volumes were prepared (depending on concentration and availability of the myosin sample) by mixing the F-actin, ATP and all assay reagents excluding the myosin in an Eppendorf tube. The reactions were initiated by adding the remaining myosin volumes (5-17 μL depending on stocks) and quickly transferring the mixtures to quartz micro-cuvettes in a Cary UV-visible light spectrophotometer recording 340 nm absorbance at regular time points. The assay was performed at 23 °C and buffer-only and actin-only control samples were also included.

Figure 6.14 shows the results of the actin-activated ATPase assay. The rates were calculated from the 340 nm absorbance traces by linear fits through to the decreasing absorbance values (before ATP is depleted and the values plateau), as detailed in 2.7. *NB*: actin-only control gradients were subtracted from the sample gradients before calculating the rates. The calculated rates shown in Figure 6.14 are given per myosin 5a head (*i.e.* active site).

Estimated ATPase rates ($s^{-1}head^{-1}$)

M5:WT-CaM: 3.22

M5:CaM-E139C: 3.23

M5:CaM-Q139C: 0.73

Actin control: 0.012
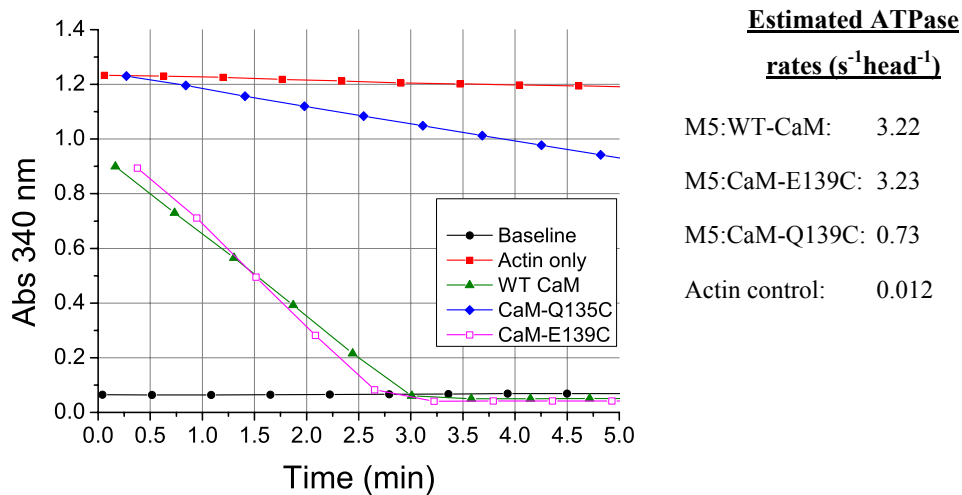
Figure 6.14: Actin-activated ATPase assay for wild-type myosin 5a-HMM exchanged with AF488M-labelled mutant CaMs.

Points plotted are 340 nm absorbance values for each sample with time. Actin concentration was 20 μM, ATP 2 mM, KCl 50 mM, HMM molecules ~0.1 μM, assay temp. ~23°C. Rates (given per head) are estimated from the gradients of linear fits to the descending straight line sections for each sample. The estimated degree of labelled CaM exchanged in the CaM-Q139C sample was 2.3±0.2 labelled CaMs/HMM molecule, 14.8±2.9 labelled CaMs/HMM in the CaM-E139C sample. The slopes of the myosin 5a-wild type CaM and myosin 5a-CaM-E139C points are similar, while the myosin 5a-CaM-Q135C sample has a slope of ~1/2 the size.

The graph and the calculated rates suggest that the wild-type (no CaM exchange) and labelled CaM-E139C-exchanged samples exhibit similar activity, ~3.2 $s^{-1}head^{-1}$, while the labelled CaM-Q135C-exchanged sample is significantly lower at 0.7 $s^{-1}head^{-1}$. The rate for the non-exchanged and CaM-E139C-exchanged samples is slightly lower than maximum rates of 3.8-9.7 $s^{-1}head^{-1}$ (at 25 °C, 25-80 mM KCl) reported for myosin 5a-HMM previously (Forgacs et al., 2009; Sakamoto et al., 2003; Wang et al., 2000) (and 5.4.3.3 of this thesis), but unlike in 5.4.3.3, no adjustments to myosin concentrations were made (*e.g.* by performing SDS gel densitometry analysis) to account for any non-active degradation fragments in the samples. From the rates observed here, one might conclude that labelled CaM-E139C has no affect on the maximum ATPase rate of wild-type myosin 5a-HMM, while labelled CaM-Q135C inhibits the maximum activity somehow (and thus may affect the ability of the wild-type myosin 5a-HMM walking on F-actin). However, caution must be exercised here as the assay was a single measurement at 20 μM actin only. No repeat

measurements were made to verify the CaM-Q135C difference seen (since myosin stocks were exhausted). Moreover, no exploration of the ATPase rate at lower actin concentrations were made, so one cannot determine the general ATPase-actin concentration profiles and if the rates have a similar difference across all actin concentrations. Also, as stated above, protein concentration measurements were from single absorbance measurements and were not tested by other means (*e.g.* SDS gel densitometry or EM imaging) to check for levels of degradation.

The results of this ATPase measurement showed sufficient evidence of enzymatic activity for the labelled CaM-E139C exchanged sample to continue with a TIRF assay. Although the single measurement of the labelled CaM-Q135C-exchanged sample showed lower activity, the rate was still markedly higher (~75 fold) than would typically be the case for myosin 5a-HMM in the absence of F-actin (≤ 0.02 s^{-1}) (Forgacs et al., 2009; Sakamoto et al., 2003), *i.e.* there was still a marked actin-activation. Therefore, it was decided that a TIRF investigation of a labelled CaM-Q135C-exchanged sample would also be of interest, at least to see if any marked inhibition of motility would confirm the ATPase result.

6.6 Dual colour TIRF assay

6.6.1 Overview

TIRF microscopy assays were performed with the aims of establishing if the AlexaFluor® 488-labelled CaM mutants bound to myosin 5a-HMM lever domain would be visible under TIRF illumination and be sufficiently bright and stable to yield high resolution information on any myosin motility on F-actin (*e.g.* step-size, run lengths). A comparison between the two labelled CaM mutants could be made to see if there was any observable difference in myosin 5a-HMM motility when exchanging-in one or the other. Beyond these aims it was hoped to extend the TIRF assays to examine if the labelled CaMs could be exchanged into the 'All-23' and 'All-25' myosin 5a-HMM lever mutant molecules, and to compare their motility characteristics with wild-type (see 5.4.3.4 for these results).

The dual colour TIRF assay is demanding and technically complex. In outline, labelled F-actin and myosin 5a-HMM samples were prepared in flow cells formed from a glass coverslip mounted on a microscope slide. Two different coloured fluorophores were used to enable split channel viewing of the differently labelled protein components. In the myosin 5a assay described the AlexaFluor® 488 fluorophore labelled CaM (a probe for the myosin lever) while actin filaments were labelled with rhodamine-phalloidin (RhPh). Each

flow cell was mounted on an the stage of an IX70 Olympus microscope (Olympus, Center Valley PA, USA) set-up for objective-type TIRF. The sample was illuminated by two lasers emitting at 488 nm (for CaM AlexaFluor® 488 excitation) and 568 nm (for rhodamine F-actin excitation). With the aid of a motorised TIRF launcher module, the laser light was directed to illuminate the flowcell at a glancing angle, such that total internal reflection was achieved at the coverslip-sample interface. This provided an electromagnetic excitation field (exponentially decreasing with distance from the interface) to a depth of ~100 nm into the sample. Fluorescent light emitted from the sample was split and passed through 510 and 605 nm filters (matched to the emission peaks for AlexaFluor® 488 and rhodamine respectively) for split-channel viewing.

The components of the flow cell were introduced in layers designed to tether fluorescently-labelled actin filaments to the inward-facing surface of the coverslip, which prior to forming the cell was coated in nitrocellulose. A layer of biotinylated bovine serum albumin (BSA) was first deposited on the nitrocellulose surface, followed by a covering of streptavidin which binds tightly to the biotin on the BSA. 10 % biotinylated (*i.e.* ~1/10 actin subunits bears biotin) RhPh actin filaments (RhPh-BFA) were introduced into the cell and allowed to partially tether to the surface (via biotin binding to remaining streptavidin binding sites on the streptavidin-coated surface). Actin remaining in solution was washed out of the cell before the surface was blocked by a coating of BSA, to prevent non-specific surface binding when myosin is introduced. Myosin 5a-HMM labelled via AlexaFluor® 488-CaM exchange was then introduced into the cell in an assay mixture containing ATP, excess free CaM (unlabelled wild type) and components of an oxygen scavenging system (to reduce rapid photobleaching). In this way, single labelled myosin molecules could 'land' from solution on a surface-tethered actin-filament within the TIRF excitation field and be monitored for movement as ATP hydrolysis occurred.

The precise methods and materials used in the TIRF assay are described below, followed by a discussion of the results of wild-type myosin 5a-HMM analysis.

6.6.2 Assay protocol

Flow cells for the TIRF assay were prepared directly before conducting the assay. Glass coverslips (22×30×0.16 mm) (Corning cover glass # 1, Corning Incorporated, New York, USA) were first cleaned by submersion in concentrated sulphuric acid for ≥ 12 hrs before rinsing with deionized H_2O and transfer to a 2 L beaker of H_2O which was then boiled for ~30 mins. Coverslips were then dried with a stream of nitrogen before application of 3 μL of 1% nitrocellulose in amyl acetate to one side of the glass (applied using a 20 μL pipette tip as a wand to spread the solution across one face). Nitrocellulose-coated coverslips were

then air-dried for ~5 mins before placing, nitrocellulose side down, onto two parallel strips of double-sided adhesive tape (Scotch 3M, USA) positioned along the long sides of a glass microscope slide. A flow cell channel of ~30 µL was thus formed between the supporting microscope slide, upper coverslip and adhesive tape sides. Solutions were sequentially introduced to the flowcell by pipetting into one end of the channel and allowing the liquid to flow through the channel under the action of gravity or by drawing the liquid through using torn Whatman filter paper as a wick placed at the channel exit.

TIRF assays were carried out in a final assay mixture containing 20 mM MOPS, pH 7.4, 5 mM $MgCl_2$, 0.1 mM EGTA, 40 mM KCl, 40 mM DTT, 0.25 mg/mL BSA, 1 µM wild-type CaM, ~1 µM ATP (actual concentration varied for different experiments), 25 µg/mL glucose oxidase, 45 µg/mL catalase, 2.5 mg/mL glucose, at 25^0 C. A motility buffer (20 mM MOPS, pH 7.4, 5 mM $MgCl_2$, 0.1 mM EGTA) was used to dilute all other assay reagents and for washing the flow cell. To conduct the assay, 30 µL of 1 mg/mL biotinylated-BSA solution was applied to the flow cell, incubated for 4 min, and washed with motility buffer (2×200 µL). 60 µL streptavidin (1 mg/mL) (Sigma-Aldrich) was then applied and incubated for 2 minutes, and washed with motility buffer (2×200 µL). Filaments of ~1 µM 10% biotinylated actin, stabilized with rhodamine-phalloidin (see 2.2.3) were applied to the flow cell and incubated for 2 min. 1 mL of motility buffer was applied to wash off unbound F-actin followed by a solution of BSA (1 mg/mL) (1×100 µL) incubated for 1 mins to block the surface. Myosin 5a-HMM samples labelled by CaM-Q135C or CaM-E139C exchange were diluted to required concentration in 100 µL of final assay mixture and introduced into the flow cell, which was then mounted onto the microscope stage for imaging. The microscope used was an Olympus IX70 equipped for objective type TIRF using an Olympus apo N (×60, numerical aperture 1.49) oil immersion objective lens, Olympus motorized multicolor cell[TIRF] illuminator and a relay lens (PE5, ×5 or PE2.5, ×2.5; Olympus) connected to a two-color observation system (Dual view system, Optical Insights) and iXon EMCCD camera (Andor Technologies, Belfast, Northern Ireland). The excitation light sources were 488 and 568 nm diode lasers (IK Series He-Cd laser, Kimmon, Prairie Technologies, Middleton WI, USA) and the dual view observation filters used were 510 nm (20 nm FWHM) for viewing the AlexaFluor[®] 488 CaM fluorescence, and 605 nm (50 nm FWHM) for viewing rhodamine-labelled F-actin. Video data was captured using MetaMorph software (Molecular Devices, LLC, Sunnyvale, CA, USA), typically 20-30 secs of data capturing at 100-200 ms/frame.

6.6.3 Results

Examples of myosin 5a-HMM molecules localised-to or 'landing' on F-actin were observed for both kinds of labelled CaM mutant by the TIRF microscopy assay. The observation via the 'green' (510 nm-filtered) channel of spots of light localising to the position of the actin filaments (simultaneously observed on the 'red' 610 nm channel) gives a definitive confirmation that both AlexaFluor® 488-labelled CaM-mutants had (to some degree) exchanged onto wild-type myosin 5a-HMM molecules. This confirms what could only be estimated by the weak absorbance spectra measurements. Moreover, the green myosin spots observed were bright enough and sufficiently long-lived before bleaching for tracking measurements to be made. Thus the AlexaFluor® 488 labelled CaMs were suitable for the myosin 5a-HMM TIRF assay. The optimum assay conditions were sought separately in each experiment, but in both cases this was found to be: myosin 5a-HMM diluted to ~1-10 nM and actin at 50 nM. ATP was normally set at ~1 μM, KCl was at 40 mM and the assays were performed at ambient temperature (~23 °C).

In the case of CaM-Q135C exchanged assay, very clear unidirectional movement of the myosin spots along the actin filaments were seen, see Figure 6.15. This is good evidence for processive movement of the myosin 5a-HMM molecules in the manner expected for the wild-type protein with endogenous CaMs. Run lengths before detachment or photobleaching were estimated at ~1-2 μm with an approximate speed of 120-220 μm/s (molecule tracking was not undertaken), data consistent for wild type HMM recorded by other groups (Armstrong et al., 2012; Baker et al., 2004; Nagy et al., 2008; Pierobon et al., 2009; Sakamoto et al., 2003).

In the case of a CaM-E139C exchanged sample, although myosin localisation to the actin filaments was observed, see Figure 6.16, there was no observation of obvious movement of the myosin along the filaments, and certainly not equivalent to the CaM-Q135C-exchanged example discussed above. Rather, at low enough concentration, single myosin spots appeared to fleetingly 'land' on the actin filaments before detaching into solution or photobleaching. Processive movement, if present, was not detectable by eye.
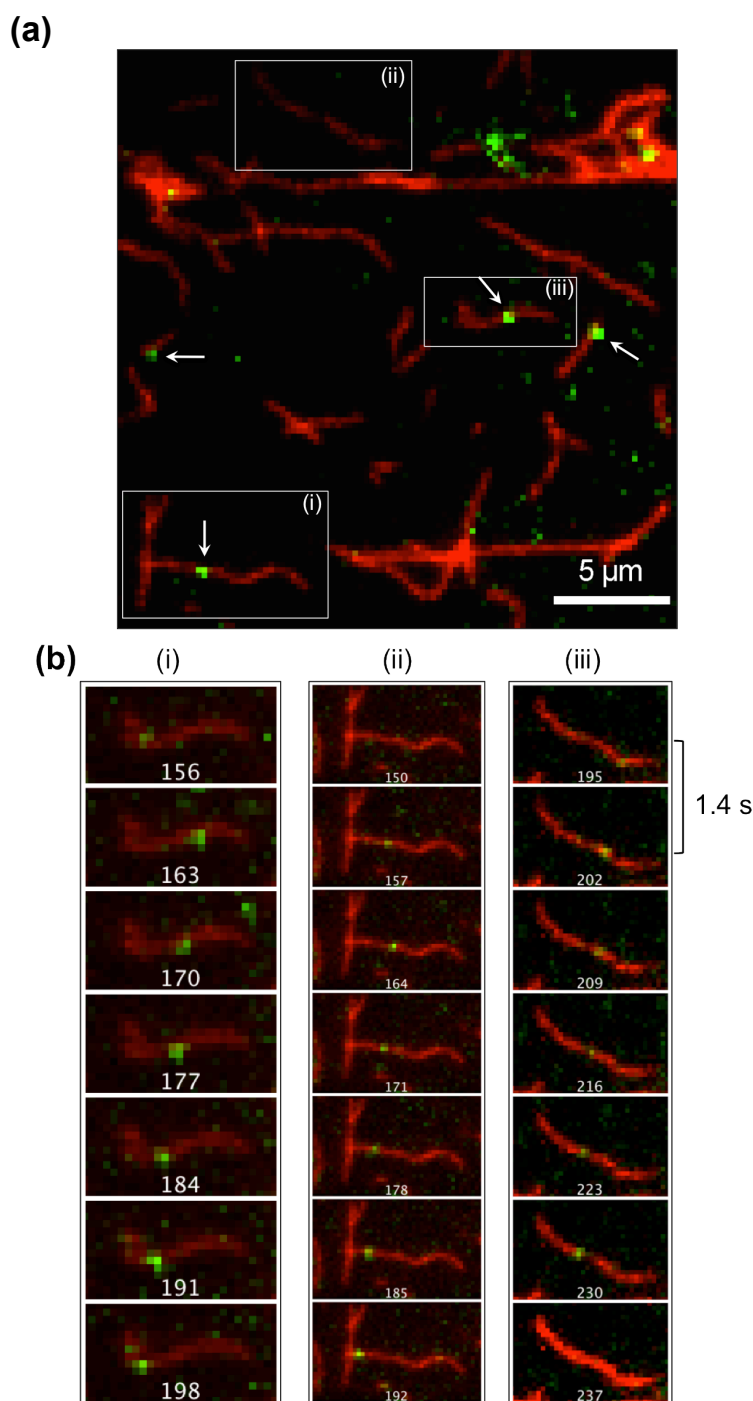
Figure 6.15: TIRF images of myosin 5a-HMM labelled with AF488M-CaM-Q135C.

(a) a field of view for a single frame (0.2 s exposure) in a TIRF assay video sequence. Single myosin 5a-HMM molecules labelled with AF488M-CaM-Q135C are seen as bright green spots overlayed on the rhodamine-phalloidin actin filaments in red. Arrows highlight examples of actin-bound molecules identified as undergoing unidirectional movement. Time sequence montages of the boxed regions (i)-(iii) are shown in (b). Each column of images in **(b)** is a frame sequence showing every seventh frame (*i.e.* 1.4 s time gaps) from part of the assay video (frame numbers label the images). In each sequence a green spot can be seen progressing ~1-2 µm along the actin filament in a 8.4 s time period (*i.e.* with speed ~120-240 nm/s). Assay conditions: 50 nM actin, ~10 nM myosin 5a-HMM, 0.75 µM ATP, 50 mM KCl, temp. ~23°C. Estimated degree of AF488M-labelled CaM-Q135C exchange, 0.7 ± 0.4 per HMM.
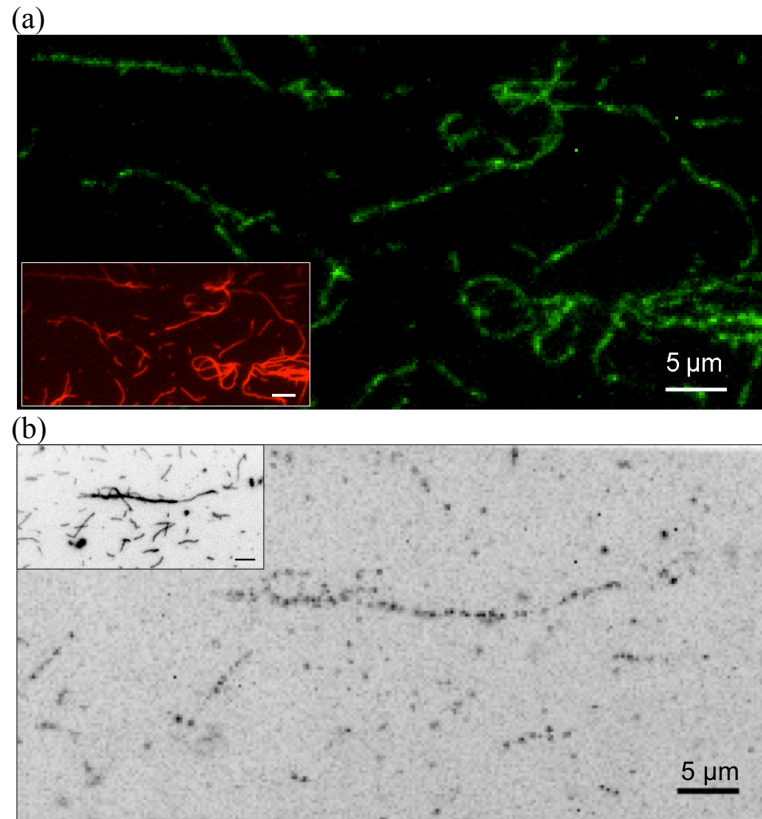
Figure 6.16: TIRF assay images for myosin 5a-HMM labelled with AF488M-CaM-E139C.

(a) field of view (FoV) for a single frame (0.2 s exposure) of a TIRF video, showing fluorescence from myosin 5a-HMM molecules labelled with AF488M-CaM-E139C (green, main image). The smaller inset image shows the rhodamine-phalloidin actin in the same FoV in red. The myosin concentration is too high to resolve individual spots, but the image clearly demonstrates that the myosin 5a-HMM molecules localise to and decorate the actin filaments. (b) main image is a maximum intensity projection through a 300-frame video of another FoV, with the myosin 5a-HMM concentration significantly reduced from that in (a). Darker spots in the main image are the myosin 5a-HMM fluorescence signal. The smaller inset image is the corresponding actin FOV. The myosin 5a-HMM image demonstrates that, although no processive movement was detected by-eye, over the 300-frames, the myosin 5a-HMM molecules preferentially land (and remain long enough to emit significant fluorescence) at the actin positions, thus recovering the pattern of the underlying filaments. Assay conditions: 50 nM actin, 0.75 μM ATP, 50 mM KCl, temp. 23°C.

The observed difference in the motility of the samples exchanged with different labelled CaM types (Q135C/E139C) could not be unambiguously ascribed to the difference in the labelled CaM type exchanged, since the two TIRF assays performed also used different preparations of myosin 5a-HMM heavy chain. In the case of the non-motile E139C-exchanged sample, SDS-PAGE analysis of the particular (un-exchanged) preparation used indicated that there was likely significant proteolysis of the molecules into single headed species, an indication that was later confirmed by EM imaging of the same preparation, Figure 6.17. A myosin 5a-HMM preparation dominated by single heads might be expected

to show the kind of static decoration or fleeting actin landing events observed by TIRF, as the single heads can bind or unbind actin (and undergo a powerstroke) but cannot participate in the processive 'walking' cycle characteristic of the two-headed molecule. The actin-activated ATPase assay (Figure 6.14) also supports the view that the low quality of the myosin 5a-HMM preparation likely accounts for the lack of CaM-E139C myosin motility seen in the TIRF assay. In the ATPase assay, the same myosin 5a-HMM preparation was used for all samples tested (making a fairer test of the effect of exchanging different CaM types) and results showed that CaM-E139C sample matched the ATPase of a wild-type control more closely than the CaM-Q135C sample (*i.e.* the opposite way around to the TIRF results). Moreover, the particular myosin 5a-HMM preparation used in the ATPase was the same one as used in the motile CaM-Q135C TIRF assay, whereas the CaM-E139C TIRF assay was the only experiment that used the separate myosin 5a-HMM preparation shown to contain significant numbers of single heads.
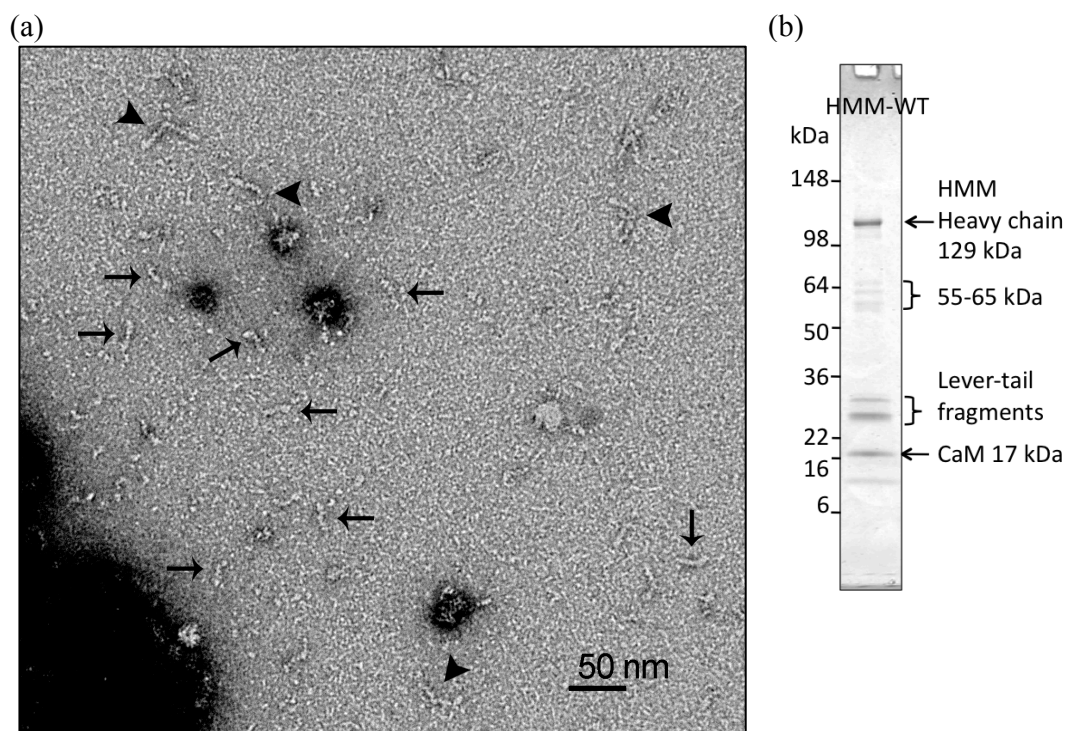
Figure 6.17: EM and SDS-PAGE evidence of mixed single and doubled-headed molecules in myosin 5a-HMM preparation.

(a) electron micrograph image and (b) SDS-PAGE analysis of the wild type myosin 5a-HMM sample used for CaM-E139C exchange and TIRF microscopy. The EM image in **(a)** shows negatively stained myosin 5a-HMM molecules. Classic two-headed species are seen (arrow heads) but are outnumbered by single heads (arrows) in this field of view. EM performed by Dr Neil Billington (LMP-NIH), grid conditions: 0.5 M KCl, 10 mM MOPS, 0.1 mM EGTA, 1 mM DTT, 1% uranyl acetate stain, carbon-coated copper mesh grid. **(b)** shows an SDS-PAGE analysis (4-20% Tris-glycine gel) of the same preparation pictured in (a). The additional bands at ~25-30 kDa are likely proteolytic-produced lever-tail fragments, see 5.4.2, suggesting the single head species seen by EM are single-headed HMMs (one heavy chain bound to the tail section of another cleaved head).

In summary, the initial results from the dual-colour TIRF assays demonstrated that both AlexaFluor 488®-labelled CaM mutants exchanged onto wild type myosin 5a-HMM molecules and that the fluorescence was sufficiently bright and long-lived to resolve and track single myosin 5a-HMM molecules on actin filaments. One sample, exchanged with labelled CaM-Q135C showed clear unidirectional movement along actin filaments, achieving a velocity of ~120-220 μm/s and run lengths of ~1-2 μm, characteristic of wild-type myosin 5a-HMM. The motility confirmed that the AF488M-labelled CaM-Q135C did not obviously disrupt the processive movement of the myosin molecule when exchanged onto the lever domain, making it suitable as a TIRF microscopy probe for myosin 5a-HMM (or other myosin) studies in the future. The CaM-E139C exchanged sample, bound actin but didn't show the same motility, but it is thought likely that this was due to the myosin 5a-HMM sample quality rather than an intrinsic effect of the AF400-M-CaM-E139C. Though there was insufficient time in the present studies, TIRF analysis of more myosin 5a-HMM preparations would be necessary to establish this definitively.

6.7   Conclusions and future work

The aim of the work in this chapter was to produce a CaM mutant that could be site-specifically labelled and exchanged onto myosin 5a-HMM molecules to act as a lever domain probe in microscopy experiments. The CaM mutant was then to be labelled with a fluorophore, exchanged onto myosin 5a-HMM molecules and tested in dual colour actomyosin 5a-HMM TIRF assays, to confirm its appropriateness for the assay. Each of these overall aims were met as judged by initial assays, with an excellent example of myosin 5a-HMM processive movement on F-actin recorded by tracking labelled and exchanged CaMs in TIRF illumination. Some revisions to the exchange protocol would be desirable and more investigative work is needed to determine conclusively if there is any inhibitory effect on myosin 5a-HMM actin-activated ATPase and motility, and if this is more pronounced for one CaM mutant or the other of the two produced.

Using the criteria to preserve CaM calcium-binding and to avoid label interference with CaM-CaM and CaM-HC interactions, two CaM cysteine point mutants were conceived. One, CaM-Q135C, followed a existing precedent already known to overexpress, fold and exchange onto myosin 5a-HMM (Churchman et al., 2005). A second, CaM-E139C was chosen since this location was considered slightly more suitable for avoiding potential label-CaM interactions, of particular importance to studies in this thesis. Both mutants were successfully cloned by PCR, overexpressed in *E. coli* and purified by liquid chromatography using a calcium-dependent interaction with phenyl-sepharose. Since the

hydrophobic phenyl-sepharose interaction depends on a CaM conformational change from closed to open on binding Ca^{2+}, the successful purification via this method is also a direct indication that both mutants are correctly folded and maintain their capacity for calcium-binding and linked conformational change.

Labelling of the mutants with AlexaFluor® 488 malemide was successfully achieved, being most definitively confirmed by mass spectrometry results. However, evidence from mass spectrometry (relative peak intensities) and spectrophotometry absorbance measurements indicated that the labelling reaction was less than 100% efficient. This could perhaps be further optimized, for example investigating longer incubation times and de-oxygenating reaction buffers. In future, a more effective form of separating unreacted dye from the labelled CaM could also be considered. In the present study extensive dialysis (up to 4.5 days) was required to separate the unreacted dye, and moreover the extent of removal remained unknown after the dialysis. A potentially faster, more efficient and quantitative alternative could be to use a desalting column, which separates the labelled protein from the low MW dye by size exclusion chromatography, or alternatively by use of the phenyl-sepharose column (as used in the CaM purification) to perform the free dye separation. Study of UV-visible light absorbance of the elution fractions (at 280 and 495 nm) from such a column would give better information on the separation and quantities of unreacted dye and labelled CaM.

Results showed that exchange of both labelled CaM mutants onto myosin 5a-HMM was possible. This was most clearly demonstrated by observation of fluorescent myosin 5a-HMM molecules in the TIRF assay. For both CaM types, myosin 5a-HMM molecules were visible and seen localising to F-actin. In the case of CaM-Q135C, one experiment showed clear unidirectional processive movement along the actin filaments in the presence of ATP. Such a clear example of myosin 5a-HMM motility when bearing labelled CaM-E139C was not obtained, though it is thought that this was due to the particular myosin preparation containing a significant number of single heads rather than a specific inhibitory affect of the CaM type exchanged. The fact that the opposite observation of the CaM mutants was made in an ATPase assay, namely that exchanged CaM-Q135C myosin 5a-HMM had a reduced activity while CaM-E139C matched a wild-type control, underlines the likelihood that the myosin preparation quality (and fraction of single heads) is the significant factor. Although time did not permit in this project, a TIRF examination of more myosin 5a-HMM preparations and ATPase assays covering a range of actin concentrations would be necessary to unequivocally determine if there were any inhibitory effects on myosin 5a-HMM of one labelled CaM or the other.

A further practical point emerging from the work in this chapter concerns the use of a cellulose-membrane centrifugal filter to separate free CaM from myosin 5a-HMM bearing CaM. Although it is a technique used by other research groups it was found here that centrifugal filtering yielded only very dilute and low volumes of labelled myosin 5a-HMM, presumably due significant protein loss by adsorption to surfaces. Though not an impediment to the TIRF assay (only small dilute quantities are required), subsequent quantification of the degree of exchange by absorbance spectrophotometry was not reliable or practicable. An alternative method for free CaM separation might be to use anti-FLAG resin (as used in the myosin purification) to selectively bind (and later release) the FLAG-tagged C-terminus of the recombinant protein (and its bound CaMs). Alternatively a simple ion exchange column (*e.g.* Q-sepharose) may be sufficient for the separation.

The successful demonstration of the use of the CaM mutants for myo5-HMM labelling and detection in TIRF microscopy assays provides a basis for future work. Indeed, the mutant CaMs are not restricted to use with myosin 5, but could be used more widely as a general probe to label other myosin classes at their light chain binding domain. In terms of the focus of this thesis, application of the CaM labelling to analyse the appearance and characteristics 'All-23' and 'All-25' myosin 5a-HMM lever mutants would be the next step. Indeed, this work has been started and initial observations indicating that the labelled CaMs do exchange into these lever mutants (such that they are visible in TIRF) are presented in other chapter 5. Following the precedent of Dunn and Spudich (2007), one could also now pursue work to attach a small gold nanoparticle to the mutant CaMs rather than a fluorophore. When exchanged onto myosin 5a-HMM (or other myosins) and viewed by EM, the gold particles might be useful markers (of higher contrast to a fluorophore) to reveal CaM orientation information and any torsional characteristics of the lever such as were investigated in chapter 3.

Chapter 7: Summary and Concluding Remarks

7.1 Summary

This thesis has provided new information on structure and flexibility within the myosin 5a head. A useful outcome of the work was also the development of a method for analysing the curvature profile and flexural rigidity of filaments or elongated protein domains from digital images. In the second part of the thesis an investigation into the importance of light chain spacing in the myosin 5a lever was undertaken. Although only initial characterisations of two mutants are reported, results were contrasting and highlight the importance of the lever's construction to the function of the molecule. The considerations made in designing the lever mutants and the outcomes from cloning, expression and purification lay foundations for potential future research, see 7.4. Another useful contribution of this project was the production of two CaM cysteine mutants, which can be used in future myosin studies for site-specific labelling and exchange onto the lever.

7.2 General remarks

A knowledge of the structure and flexibility within the myosin 5a head is important not only for understanding the mechanics and function of myosin 5, but also of myosins in general, which share the motor-plus-lever structure. The lever plays a key mechanical role in myosin function. It acts as a distance multiplying element, transforming small nucleotide-dependent movements within the motor into a nm-sized movement of the head-tail junction. Its mechanical properties are therefore important. To properly function, the domain must be sufficiently rigid to rotate as a semi-rigid body and sustain force, *e.g.* in thick-thin filament sliding in muscle contraction, whilst in other circumstances, *e.g.* during myosin 5a 'walking', it must be flexible enough to bend and accommodate functionally-important strained conformations. Variety in structural components and in the number and spacing of IQ motifs across the myosin family suggests possible divergence of lever structure for different mechanical functions. However, our knowledge of specific mechanical properties of myosin levers and an understanding of how these derive from the domain's substructure is still somewhat lacking. Work in this project has contributed some new information in this area with regard to myosin 5a, though there remains much still to investigate.

Myosin 5a was chosen as a model system for study as its has an especially long lever formed from canonical CaM-IQ subunits, and it is clear that its length and stiffness play an important role in the processive behaviour of the molecule. In addition, it has a unique alternating IQ spacing pattern, a pattern that is conserved in animals and also in myosin 11, the plant counterpart of myosin 5. The reason for the spacing pattern has not been determined, but may have to do with dictating a particular 3D orientation of the light chains, which perhaps creates a favourable lever structure for the molecule's function. Testing of this hypothesis was the aim of the second part of this project.

The use of images from negative stain EM formed an important part of this thesis. Although this technique has the potential for producing staining or drying artefacts, and is limited to a lower resolution than crystallography, here it offered the advantage that entire molecules (including those bound to F-actin) could be readily imaged in a range of conformations (unlike in crystallography where only the conformation(s) favourable to crystallisation are selected). Moreover, through the use of single-particle image processing, the information gained from micrographs was increased and the signal to noise ratio greatly improved by classification and averaging. Owing to the high quality of staining and imaging in the original data, a resolution of about ~2 nm (the working limit of negative stain EM) was achieved. The molecule conformations revealed could be compared to existing crystal structures to judge the technique's reliability, and the variation across conformations could be analysed to understand the types and extent of flexibility. Though perhaps inferior to cryo-EM, which can be used to image hydrated, unstained specimens to higher resolution (in 3D orientations), there is still much that can be learned from negative stain images of myosin molecules.

7.3   Summary of specific findings

The principal findings from the different parts of this thesis are summarised below with possible directions for future work given thereafter.

7.3.1   Structural details of the myosin 5a head

In chapter 3, image alignment and averaging resolved structural features within the myosin 5a motor and lever. Within the (apo) motor, the SH3 domain, parts of the lower and upper 50k, and the actin-binding cleft were identified from comparison with crystal structures The close correspondence found between the motor domain's predominant appearance in the EM averages (here called 'face-in-profile') and an orientation of the motor domain crystal structure provided an indication of the reliability of the staining process.

Within the lever, all six CaMs were clearly delineated, but with different shapes, either elliptical or circular, corresponding to varying azimuths of view. In some cases regions of neighbouring CaMs appeared closely packed, possibly contacting, while at other locations there was more space between the adjacent light chains. This suggested that CaM-CaM stabilising interactions and the constraints of steric clashing could be important factors for lever bending.

7.3.2 New atomic model of the myosin 5a head

Comparison of the EM average to atomic models of the entire head showed that the profile of each head of an (enzymatically active) HMM-like molecule, stained on a carbon substrate, is very similar to one of the heads of a folded full-length (inhibited) molecule, supported on a lipid monolayer, as obtained by Liu et al. (2006). The profile of all six CaMs of the head model of Liu et al. matched the shapes seen in the new EM average (which was not the case for other models compared), validating it as a good working model of the myosin 5a lever section. However, the motor profile matched less well, until it was reoriented in the author's 3D-EM map to a position, as suggested by Sellers et al. (2008), that made a more satisfying fit to the density map and created a better correspondence to the new negative stain HMM average. From this, it was concluded that the Sellers et al. reinterpretation of the 3D-EM map is indeed preferable to the original Liu et al. interpretation. An adapted atomic model closely matching the new HMM average and fitting the Liu et al. 3D-EM map was made accordingly. This represents a new working model of the myosin 5a head that best fits current empirical data. Of importance to the flexibility studies, was the observation that construction of the new model required the introduction of a bend in the lever helix at the motor lever junction (at residues 755-767), highlighting this as a potential pliant region in the head. In addition, it was found that if the motor in the new model was tilted by about 45° about the motor-lever junction, to make a flatter head, this matched the EM head average even more closely. This is evidence of head flattening occurring on the carbon film of the EM grid, probably as a result of the surface adsorption and drying processes.

7.3.3 Flexibility in the myosin 5a head

Classification of aligned head images provided evidence of particular flexibility at the motor-lever junction. Two primary views of the motor, related by a rotation of ~50° about the lever axis were identified. This is consistent with the findings in Burgess et al. (2002), although it is new in this thesis that the axial rotation is apparently independent of the orientation of the lever (which on average tends to adopt only one axial orientation).

Significant flexing at the motor-lever junction was also observed for both motor orientations. The distribution of angles suggested thermally-driven fluctuations about a mean, with a torsional spring constant of ~32-51 pN·nm·rad^{-2} for movement of the motor relative to the lever about the motor-lever junction. This represents a significant source of compliance in the unattached head, roughly equivalent to 0.11 pN·nm^{-1} if it was mistakenly attributed to cantilever-type lever bending in an actomyosin crossbridge. A pliant region at the motor-lever junction is consistent with evidence from scallop smooth and striated muscle myosin 2 where bending at this point is also found (Burgess et al., 2007; Houdusse et al., 2000), and so may be a wider spread feature in other myosin classes. Thus, in future crossbridge models of myosin 5a and other myosins, this type of compliance should be considered.

Independent motor rotation and flexing confirm that the motor-lever junction is a pliant region in the molecule, an idea first suggested in Burgess et al. (2002). Flexibility at the motor-lever junction is one way (the other being lever bending) that a lead head in myosin 5a can deform and adopt a pre-powerstroke conformation when both heads are attached to an actin filament at the preferred 13 subunit spacing.

Independent variation in the CaM orientations at each IQ motif was an unexpected (and not previously reported) observation from this study. Individual lobes of CaMs were clearly resolved in class average images and allowed a rough identification of the azimuthal orientations of CaMs around the lever axis to be made. The observation of varying orientations of CaMs at a given position (apparently independent of the orientation of other parts of the lever) suggests either that there are significant (previously unreported) regions of torsional flexibility between the CaM-binding regions of the lever, or is evidence of distorting forces acting on the molecules during drying on the EM grid. If a genuine feature of the molecule *in vivo*, the purpose of such torsional flexibility is unclear but might, for instance, be a mechanism that helps the molecule to efficiently find F-actin subunits at similar azimuths while walking. Dynamic azimuthal rotation of CaM-IQ subunits might also be a mechanism that allows the lever to adopt more highly bent conformations if required. However, in considering the possibility that the CaM rotations are merely an artefact of the negative staining process, an estimate made of the surface tension forces associated with a liquid film drying around the grid-adsorbed molecules was surprisingly large, of the order of hundreds of pNs. Therefore, the possibility that such forces distort the CaMs in the lever, *e.g.* applying torques to different CaM-IQ sections, currently seems a more likely explanation for the results.

Lever bending was studied by image classification and by a method of numerically approximating the lever's shape and mapping curvature with arc length. The purpose of the work was to test whether the flexural rigidity of the lever was isotropic and uniform, and what rigidity parameter(s), if appropriate, describe the bending stiffness. Within a low force-low bending regime (up to a radius of curvature of magnitude ~5 nm) bending was found to be uniform and consistent with the isotropic rod model (an isotropic worm-like chain) adopted by Vilfan (2005). However, for bending above this threshold, specific anisotropies were observed, with further bending apparently blocked in certain directions at the junctions between CaMs 2 & 3, and between CaMs 5 & 6. Study of the atomic model of the head suggested this blocking may be due to constraints of steric clashes between adjacent CaM lobes. Evidence for greater flexibility at junctions between the 25-residue spaced IQ motifs was not found. There may also be bending limitations in directions perpendicular to those measured, but these directions are not probed by the 2D negative stain images.

Assuming a simple isotropic rod model for the lever (*i.e.* ignoring the high force anisotropies above), a flexural rigidity for the lever was estimated from a decomposition of average lever shapes into cosine modes. Despite the complications of anisotropies in bending and simplifying assumptions made in the analysis (*e.g.* treating the lever as an unconstrained rod that equilibrates in 2D) the first three modes were consistent with the expectations of the WLC model and suggested a flexural rigidity for the lever of 206 ± 75 pNnm^2. This is equivalent to a persistence length of 50.9 ± 18.6 nm or a cantilever stiffness of 0.045 ± 0.017 pNnm^{-1}. The high relative uncertainty (~37%) is due to the relatively small set of lever shapes analysed. This stiffness number is surprisingly low, ~1/5-1/4 the cantilever stiffness measured by Veigel et al. for actomyosin 5a S1 (Veigel et al., 2005; 2002), and roughly equivalent to the bending stiffness of DNA. Could the myosin 5a levers be replaced with lengths of DNA and still function? The low flexural rigidity found was cross-checked by analysis of tangential angle correlations. Reasons why the optical trap measurements of Veigel et al. may be overestimates of the bending stiffness were discussed.

Overall, the data obtained on motor rotation, motor-lever flexing, CaM-IQ twisting and lever bending suggest that the apo, unattached myosin 5a head, at least under negative stain EM conditions, is more flexible than has previously been measured by optical trap experiments.

7.3.4 The importance of IQ motif spacing

In the second half of this thesis, work was begun to test the importance of IQ motif spacing on the properties of the myosin 5a lever domain and on the behaviour of the molecule on F-actin. Although only initial characterisation work was completed, contrasting results for two different IQ lever mutants were found.

In one case, alteration of the motif spacing pattern to 'All-23' (by a strategy of replacing IQ-bridge sequences with stretches of native sequence) produced a molecule that resembled wild-type in appearance and enzymatic activity. Frustratingly, the molecule's motile properties were not determined due to problems of sample degradation and purity, issues which could be addressed in future work.

In the second case, alteration of the IQ spacing to 'All-25', by insertion of pairs of leucine residues (the only cloning method that worked), caused a dramatic change in the molecule. The construct did not bind calmodulin and it had no actin-activated MgATPase. The molecule appeared misfolded but retained residual actin-binding properties. Clearly, the insertion of leucine residues into the lever helix was disruptive in some way. Even though individual leucines do occupy the equivalent insert positions elsewhere in myosin 5a lever sequences, the sequence LL itself actually only ever occurs in one instance known sequences, in zebrafish. In hindsight, perhaps a less bulky/hydrophobic residue (*e.g.* alanine) would have made a better choice for insertion. Unfortunately, the equivalent All-25 construct produced by native sequence splicing could not be cloned.

Production in this project of an intact intermediate IQ construct, with the first IQ spacing changed from 23 to 25 residues, and the existence of the 'HMM-2Ala-6IQ' construct (Oke et al., 2010; Sakamoto et al., 2003), indicate that a change from a 23 to a 25 spacing is not *per se* destructive to the lever.

7.3.5 Production of CaM mutants for labelling studies

Finally, a useful outcome of the latter part of the project was the production of two CaM cysteine mutants that can be used for site-specific labelling. The two mutants produced were Q135C and E139C. Both were well expressed and purified, the purification technique indicating that the mutants retain the ability to undergo the conformational change on binding or release of calcium. Each mutant was successfully labelled with the fluorescent dye AlexFluor488-maleimide, and, after successful exchange onto myosin 5a-HMM molecules, was visible at single molecule concentration in TIRF microscopy. Either CaM mutant could be used in future myosin labelling experiments.

7.4 Future directions

The work documented in this thesis has been informative, and there are many possibilities for continuation of the research.

For the studies of bending flexibilities in the myosin 5a lever, the analytical tools developed here could be usefully applied to more data, to corroborate the results found here and potentially act as a control. This could include the remainder of the HMM dataset used in this project, S1-like molecules (which, lacking the second head, are a simpler system to model) as well as both these molecules types bound to F-actin. Moreover, the technique could also be applied to other filamentous molecules or flexible protein domains *e.g.* caldesmon, tropomyosin, or the coiled-coil stalk of dynein, to map their flexibility and determine bending stiffness values.

With regard to the myosin 5a IQ motif spacing study, there remains more work to do to fully answer some of the research questions of this PhD. One priority would be to try to establish the motile properties of the All-23-Bridge lever mutant compared to wild-type. This would give a clear indication of whether changing the IQ spacing alters the mechanical properties of the lever, and would answer an important biological question about myosin 5a function. Use of the TIRF assay and labelling of the molecules with one of the fluorophore-labelled CaM mutants already looks promising in this regard. The requirement here would be to produce consistently intact and pure (two-headed) protein preparations. For the All-25 construct one could attempt a characterisation of one of the well-expressed intermediate constructs *e.g.* with only one 23 spacing changed to 25.

Finally, an obvious future step, and one that relates the two aspects of this PhD, would be to analyse the appearance of the heads of the All-23-Bridge mutant and the All-25 intermediate, using the same negative stain EM and single-particle image processing techniques as applied to wild-type. Indeed, at the time of writing this thesis a large well-stained dataset of free All-23-Bridge molecules has been obtained. Many questions are therefore close to being answered. Are the average orientations of the CaMs altered as one would expect for changed IQ motif spacings? Does the lever lie in the same orientation on the EM grid as wild-type? Is the same CaM-IQ subunit twisting seen, and is the lever bending profile the same?

With this PhD thesis as a starting point, it is hoped that future researchers can answer some of these open questions and come to a better understanding of the relationship between structure and function in this important class of motor protein.

References

Alberts, B., 1998. The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell 92, 291–294.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002. Molecular Biology of the Cell, 4th ed. Garland Science, New York.

Ali, M.Y., Homma, K., Iwane, A.H., Adachi, K., Itoh, H., Kinosita, K., Yanagida, T., Ikebe, M., 2004. Unconstrained steps of myosin VI appear longest among known molecular motors. Biophysical Journal 86, 3804–3810.

Ali, M.Y., Uemura, S., Adachi, K., Itoh, H., Kinosita, K., Ishiwata, S., 2002. Myosin V is a left-handed spiral motor on the right-handed actin helix. Nat Struct Mol Biol 9, 464–467.

Armstrong, J.M., Krementsova, E., Michalek, A.J., Heaslip, A.T., Nelson, S.R., Trybus, K.M., Warshaw, D.M., 2012. Full-length myosin Va exhibits altered gating during processive movement on actin. Proceedings of the National Academy of Sciences 109, E218–E224.

Axelrod, D., 1989. Total internal reflection fluorescence microscopy. Methods Cell Biol. 30, 245–270.

Baboolal, T.G., Sakamoto, T., Forgacs, E., White, H.D., Jackson, S.M., Takagi, Y., Farrow, R.E., Molloy, J.E., Knight, P.J., Sellers, J.R., Peckham, M., 2009. The SAH domain extends the functional length of the myosin lever. Proc Natl Acad Sci USA 106, 22193–22198.

Baker, J.E., Krementsova, E.B., Kennedy, G.G., Armstrong, A., Trybus, K.M., Warshaw, D.M., 2004. Myosin V processivity: multiple kinetic pathways for head-to-head coordination. Proc Natl Acad Sci USA 101, 5542–5546.

Ballardini, R., Balzani, V., Credi, A., Gandolfi, M.T., Venturi, M., 2001. Artificial molecular-level machines: which energy to make them work? Acc. Chem. Res. 34, 445–455.

Bähler, M., Rhoads, A., 2002. Calmodulin signaling via the IQ motif. FEBS Letters 513, 107–113.

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

Beausang, J.F., Sun, Y., Quinlan, M.E., Forkey, J.N., Goldman, Y.E., 2007. Orientation and Rotational Motions of Single Molecules by Polarized Total Internal Reflection Fluorescence Microscopy, in: Selvin, P.R., Taekjip, H. (Eds.), Single-Molecule Techniques: a Laboratory Manual. Cold Spring Harbor Laboratory Press, New York, pp. 121–148.

Benashski, S.E., Harrison, A., Patel-King, R.S., King, S.M., 1997. Dimerization of the highly conserved light chain shared by dynein and myosin V. J Biol Chem 272, 20929–20935.

Bensimon, A., Simon, A., Chiffaudel, A., Croquette, V., Heslot, F., Bensimon, D., 1994. Alignment and sensitive detection of DNA by a moving interface. Science 265, 2096–2098.

Bensimon, D., Simon, A., Croquette, V., Bensimon, A., 1995. Stretching DNA with a receding meniscus: Experiments and models. Physical Review Letters 74, 4754–4757.

Berg, J.S., Powell, B.C., Cheney, R.E., 2001. A millennial myosin census. Mol. Biol. Cell 12, 780–794.

Bloemink, M.J., Geeves, M.A., 2011. Shaking the myosin family tree: Biochemical kinetics defines four types of myosin motor. Seminars in Cell & Developmental Biology 22, 961–967.

Boisset, N., Taveau, J.C., Lamy, J., Wagenknecht, T., Radermacher, M., Frank, J., 1990. Three-dimensional reconstruction of native Androctonus australis hemocyanin. J Mol Biol

References

216, 743–760.

Brangwynne, C.P., Koenderink, G.H., Barry, E., Dogic, Z., MacKintosh, F.C., Weitz, D.A., 2007. Bending dynamics of fluctuating biopolymers probed by automated high-resolution filament tracking. Biophysical Journal 93, 346–359.

Burgess, S., Walker, M., Wang, F., Sellers, J., White, H., Knight, P., Trinick, J., 2002. The prepower stroke conformation of myosin V. J Cell Biol 159, 983–991.

Burgess, S., Yu, S., Walker, M., Hawkins, R., Chalovich, J., Knight, P., 2007. Structures of Smooth Muscle Myosin and Heavy Meromyosin in the Folded, Shutdown State. J Mol Biol 372, 1165–1178.

Burgess, S.A., Walker, M.L., Sakakibara, H., Knight, P.J., Oiwa, K., 2003. Dynein structure and power stroke. Nature 421, 715–718.

Burgess, S.A., Walker, M.L., Sakakibara, H., Oiwa, K., Knight, P.J., 2004. The structure of dynein-c by negative stain electron microscopy. J Struct Biol 146, 205–216.

Burgess, S.A., Walker, M.L., Thirumurugan, K., Trinick, J., Knight, P.J., 2004. Use of negative stain and single-particle image processing to explore dynamic properties of flexible macromolecules. J Struct Biol 147, 247–258.

Burgess, S.A., Walker, M.L., White, H.D., Trinnick, J., 1997. Flexibility within Myosin Heads Revealed by Negative Stain and Single-Particle Analysis. The Journal of Cell Biology 139, 675–681.

Bustamante, C., Chemla, Y.R., Forde, N.R., Izhaky, D., 2004. Mechanical Processes In Biochemistry. Annual Review of Biochemistry 73, 705–748.

Cameron, L.C., Carvalho, R.N., Araujo, J.R., Santos, A.C., Tauhata, S.B., Larson, R.E., Sorenson, M.M., 1998. Calcium-induced quenching of intrinsic fluorescence in brain myosin V is linked to dissociation of calmodulin light chains. Arch Biochem Biophys 355, 35–42.

Cappello, G., Pierobon, P., Symonds, C., Busoni, L., Gebhardt, J., Rief, M., Prost, J., 2007. Myosin V stepping mechanism. Proc Natl Acad Sci USA 104, 15328–15333.

Cheney, R.E., Mooseker, M.S., 1992. Unconventional myosins. Current Opinion in Cell Biology 4, 27–35.

Cheney, R.E., O'Shea, M.K., Heuser, J.E., Coelho, M.V., Wolenski, J.S., Espreafico, E.M., Forscher, P., Larson, R.E., Mooseker, M.S., 1993. Brain myosin-V is a two-headed unconventional myosin with motor activity. Cell 75, 13–23.

Cheney, R.E., Riley, M.A., Mooseker, M.S., 1993. Phylogenetic analysis of the myosin superfamily. Cell Motil. Cytoskeleton 24, 215–223.

Cheng, Y., Wolf, E., Larvie, M., Zak, O., Aisen, P., Grigorieff, N., Harrison, S.C., Walz, T., 2006. Single particle reconstructions of the transferrin-transferrin receptor complex obtained with different specimen preparation techniques. J Mol Biol 355, 1048–1065.

Cheung, W.Y., 1980. Calmodulin plays a pivotal role in cellular regulation. Science 207, 19–27.

Churchman, L.S., Okten, Z., Rock, R.S., Dawson, J.F., Spudich, J.A., 2005. Single molecule high-resolution colocalization of Cy3 and Cy5 attached to macromolecules measures intramolecular distances through time. Proc Natl Acad Sci USA 102, 1419–1423.

Collins, J.H., 1991. Myosin light chains and troponin C: Structural and evolutionary relationships revealed by amino acid sequence comparisons. Journal of Muscle Research and Cell Motility 12, 3–25.

Coluccio, L., Ridley, A., Frampton, J., 2008. Myosins: A Superfamily of Molecular Motors. Proteins and Cell Regulation 7, 475.

Coluccio, L.M., 2007. Myosin I, in: Coluccio, L.M. (Ed.), Myosins: a Superfamily of Molecular Motors. Springer, Dordrecht, pp. 95–124.

Cooke, R., 1986. The mechanism of muscle contraction. CRC Crit. Rev. Biochem. 21, 53–118.

Coureux, P., Sweeney, H., Houdusse, A., 2004. Three myosin V structures delineate essential features of chemo-mechanical transduction. EMBO Journal 23, 4527–4537.

Coureux, P.-D., Wells, A., Menetrey, J., Yengo, C., Morris, C., Sweeney, H., Houdusse, A., 2003. A structural state of the myosin V motor without bound nucleotide. Nature 425, 419–423.

Craig, E.M., Linke, H., 2009. Mechanochemical model for myosin V. Proceedings of the National Academy of Sciences 106, 18261–18266.

Craig, R., Woodhead, J.L., 2006. Structure and function of myosin filaments. Current Opinion in Structural Biology 16, 204–212.

Dedman, J.R., Kaetzel, M.A., 1983. Calmodulin Purification and Fluorescent Labeling. Meth Enzymol 102, 1–8.

Doi, M., Edwards, S.F., 1988. The Theory of Polymer Dynamics. Clarendon Press, Oxford.

Dominguez, R., Freyzon, Y., Trybus, K.M., Cohen, C., 1998. Crystal structure of a vertebrate smooth muscle myosin motor domain and its complex with the essential light chain: visualization of the pre-power stroke state. Cell 94, 559–571.

Dunn, A.R., Spudich, J.A., 2007. Dynamics of the unbound head during myosin V processive translocation. Nat Struct Mol Biol 14, 246–248.

Dunn, B.D., Sakamoto, T., Hong, M.-S.S., Sellers, J.R., Takizawa, P.A., 2007. Myo4p is a monomeric myosin with motility uniquely adapted to transport mRNA. J Cell Biol 178, 1193–1206.

Egelman, E.H., 1997. New angles on actin dynamics. Structure 5, 1135–1137.

Egelman, E.H., Francis, N., DeRosier, D.J., 1982. F-actin is a helix with a random variable twist. Nature 298, 131–135.

El-Mezgueldi, M., Bagshaw, C.R., 2007. The myosin family: Biochemical and kinetic properties, in: Coluccio, L.M. (Ed.), Myosins: a Superfamily of Molecular Motors. Springer, Dordrecht, pp. 55–93.

Elliott, A., Offer, G., Burridge, K., 1976. Electron Microscopy of Myosin Molecules from Muscle and Non-Muscle Sources. Proceedings of the Royal Society of London. Series B. Biological Sciences 193, 45–53.

Espindola, F.S., 1992. Biochemical and immunological characterization of p190-calmodulin complex from vertebrate brain: a novel calmodulin-binding myosin. The Journal of Cell Biology 118, 359–368.

Espindola, F.S., Suter, D.M., Partata, L.B., Cao, T., Wolenski, J.S., Cheney, R.E., King, S.M., 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 Motil. Cytoskeleton 47, 269–281.

Espreafico, E.M., Cheney, R.E., Matteoli, M., Nascimento, A.A., De Camilli, P.V., Larson, R.E., Mooseker, M.S., 1992. Primary structure and cellular localization of chicken brain myosin-V (p190), an unconventional myosin with calmodulin light chains. J Cell Biol 119, 1541–1557.

Fakhri, N., Tsyboulski, D.A., Cognet, L., Weisman, R.B., Pasquali, M., 2009. Diameter-dependent bending dynamics of single-walled carbon nanotubes in liquids. Proceedings of the National Academy of Sciences 106, 14219–14223.

Finn, B.E., Evenäs, J., Drakenberg, T., Waltho, J.P., Thulin, E., Forsén, S., 1995. Calcium-induced structural changes and domain autonomy in calmodulin. Nat Struct Mol Biol 2, 777–783.

Fischer, S., Windshügel, B., Horak, D., Holmes, K.C., Smith, J.C., 2005. Structural mechanism of the recovery stroke in the myosin molecular motor. Proc Natl Acad Sci

USA 102, 6873–6878.

Fisher, A.J., Smith, C.A., Thoden, J.B., Smith, R., Sutoh, K., Holden, H.M., Rayment, I., 1995. X-ray structures of the myosin motor domain of Dictyostelium discoideum complexed with MgADP.BeFx and MgADP.AlF4-. Biochemistry 34, 8960–8972.

Flicker, P.F., Wallimann, T., Vibert, P., 1983. Electron microscopy of scallop myosin. Location of regulatory light chains. J Mol Biol 169, 723–741.

Forgacs, E., Cartwright, S., Sakamoto, T., Sellers, J.R., Corrie, J.E.T., Webb, M.R., White, H.D., 2008. Kinetics of ADP dissociation from the trail and lead heads of actomyosin V following the power stroke. J Biol Chem 283, 766–773.

Forgacs, E., Sakamoto, T., Cartwright, S., Belknap, B., Kovács, M., Tóth, J., Webb, M.R., Sellers, J.R., White, H.D., 2009. Switch 1 mutation S217A converts myosin V into a low duty ratio motor. J Biol Chem 284, 2138–2149.

Forkey, J.N., Quinlan, M.E., Shaw, M.A., Corrie, J.E.T., Goldman, Y.E., 2003. Three-dimensional structural dynamics of myosin V by single-molecule fluorescence polarization. Nature 422, 399–404.

Foth, B.J., Goedecke, M.C., Soldati, D., 2006. New insights into myosin evolution and classification. Proc Natl Acad Sci USA 103, 3681–3686.

Frank, J., 1990. Classification of macromolecular assemblies studied as 'single particles'. Q. Rev. Biophys. 23, 281–329.

Frank, J., 2006. Three-Dimensional Electron Microscopy of Macromolecular Assemblies. Oxford University Press, Inc., New York.

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., Leith, A., 1996. SPIDER and WEB: Processing and Visualization of Images in 3D Electron Microscopy and Related Fields. J Struct Biol 116, 190–199.

Frontali, C., Dore, E., Ferrauto, A., Gratton, E., Bettini, A., Pozzan, M.R., Valdevit, E., 1979. An absolute method for the determination of the persistence length of native DNA from electron micrographs. Biopolymers 18, 1353–1373.

Fujii, T., Iwane, A.H., Yanagida, T., Namba, K., 2010. Direct visualization of secondary structures of F-actin by electron cryomicroscopy. Nature 467, 724–728.

Funatsu, T., Harada, Y., Tokunaga, M., Saito, K., Yanagida, T., 1995. Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. Nature 374, 555–559.

Geeves, M.A., Holmes, K.C., 1999. Structural mechanism of muscle contraction. Annual Review of Biochemistry 68, 687–728.

Geeves, M.A., Holmes, K.C., 2005. The Molecular Mechanism of Muscle Contraction, in: Squire, J., Parry, D. (Eds.), Advances in Protein Chemistry. Fibrous Proteins: Muscle and Molecular Motors. Elsevier (Academic Press), pp. 161–193.

Gennerich, A., Vale, R.D., 2009. Walking the walk: how kinesin and dynein coordinate their steps. Current Opinion in Cell Biology 21, 59–67.

Gittes, F., Mickey, B., Nettleton, J., Howard, J., 1993. Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape. J Cell Biol 120, 923–934.

Gopolakrishna, R., Anderson, W.B., 1982. Ca^{2+}-induced hydrophobic site on calmodulin: Application for purification of calmodulin by phenyl-Sepharose affinity chromatography. Biochem Biophys Res Commun 104.

Grabarek, Z., 2006. Structural basis for diversity of the EF-hand calcium-binding proteins. J Mol Biol 359, 509–525.

Hagerman, P.J., 1988. Flexibility of DNA. Annu Rev Biophys Biophys Chem 17, 265–286.

Hammer, J.A., Sellers, J.R., 2012. Walking to work: roles for class V myosins as cargo

transporters. Nat Rev Mol Cell Biol 13, 13–26.

Hanson, J., Lowy, J., 1963. The structure of F-actin and of actin filaments isolated from muscle. J Mol Biol 6, 46–IN5.

Hayat, M.A., 2000. Principles and Techniques of Electron Microscopy, 4th ed. Cambridge University Press, Cambridge.

Hirokawa, N., Noda, Y., Tanaka, Y., Niwa, S., 2009. Kinesin superfamily motor proteins and intracellular transport. Nat Rev Mol Cell Biol 10, 682–696.

Hirokawa, N., Takemura, R., 2007. Kinesin Superfamily Proteins, in: Coluccio, L.M. (Ed.), Myosins: a Superfamily of Molecular Motors. Springer, Dordrecht, pp. 79–109.

Hodges, A.R., Krementsova, E.B., Trybus, K.M., 2008. She3p binds to the rod of yeast myosin V and prevents it from dimerizing, forming a single-headed motor complex. J Biol Chem 283, 6906–6914.

Holmes, K.C., 1997. The swinging lever-arm hypothesis of muscle contraction. Curr. Biol. 7, R112–8.

Holmes, K.C., Kabsch, W., 1991. Muscle proteins: actin. Current Opinion in Structural Biology 1, 270–280.

Holmes, K.C., Popp, D., Gebhard, W., Kabsch, W., 1990. Atomic model of the actin filament. Nature 347, 44–49.

Homma, K., Saito, J., Ikebe, R., Ikebe, M., 2000. Ca(2+)-dependent regulation of the motor activity of myosin V. J Biol Chem 275, 34766–34771.

Houdusse, A., Cohen, C., 1995. Target sequence recognition by the calmodulin superfamily: implications from light chain binding to the regulatory domain of scallop myosin. Proc Natl Acad Sci USA 92, 10644–10647.

Houdusse, A., Cohen, C., 1996. Structure of the regulatory domain of scallop myosin at 2 angstrom resolution: Implications for regulation. Structure 4, 21–32.

Houdusse, A., Gaucher, J.F., Krementsova, E., Mui, S., Trybus, K.M., 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 103, 19326–19331.

Houdusse, A., Silver, M., Cohen, C., 1996. A model of Ca2+-free calmodulin binding to unconventional myosins reveals how calmodulin acts as a regulatory switch. Structure 4, 1475–1490.

Houdusse, A., Szent-Györgyi, A., Cohen, C., 2000. Three conformational states of scallop myosin S1. Proc Natl Acad Sci USA 97, 11238–11243.

Houk, T.W., Ue, K., 1974. The measurement of actin concentration in solution: a comparison of methods. Anal. Biochem. 62, 66–74.

Howard, J., 2001. Mechanics of Motor Proteins and the Cytoskeleton. Sinauer Associates, Inc., Sunderland (MA, USA).

Howard, J., Spudich, J.A., 1996. Is the lever arm of myosin a molecular elastic element? Proc. Natl. Acad. Sci. USA 93, 4462–4464.

Höök, P., Vallee, R.B., 2006. The dynein family at a glance. J. Cell. Sci. 119, 4369–4371.

Huxley, H.E., 1957. The double array of filaments in cross-striated muscle. J Biophys Biochem Cytol 3, 631–648.

Huxley, H.E., 1963. Electron Microscope Studies on the Structure of Natural and Synthetic Protein Filaments from Striated Muscle. J Mol Biol 7, 281–308.

Huxley, H.E., 1969. The mechanism of muscular contraction. Science 164, 1356–1365.

Isambert, H., Venier, P., Maggs, A.C., Fattoum, A., Kassab, R., Pantaloni, D., Carlier, M.F., 1995. Flexibility of actin filaments derived from thermal fluctuations. Effect of bound

nucleotide, phalloidin, and muscle regulatory proteins. J Biol Chem 270, 11437–11444.

Janson, M.E., Dogterom, M., 2004. A bending mode analysis for growing microtubules: evidence for a velocity-dependent rigidity. Biophysical Journal 87, 2723–2736.

Johnston, G.C., Prendergast, J.A., Singer, R.A., 1991. The Saccharomyces cerevisiae MYO2 gene encodes an essential myosin for vectorial transport of vesicles. J Cell Biol 113, 539–551.

Kabsch, W., Mannherz, H.G., Suck, D., Pai, E.F., Holmes, K.C., 1990. Atomic structure of the actin:DNase I complex. Nature 347, 37–44.

Kawasaki, H., Nakayama, S., Kretsinger, R.H., 1998. Classification and evolution of EF-hand proteins. Biometals 11, 277–295.

Kaya, M., Higuchi, H., 2010. Nonlinear elasticity and an 8-nm working stroke of single myosin molecules in myofilaments. Science 329, 686–689.

Käs, J., Strey, H., Tang, J.X., Finger, D., Ezzell, R., Sackmann, E., Janmey, P.A., 1996. F-actin, a model polymer for semiflexible chains in dilute, semidilute, and liquid crystalline solutions. Biophysical Journal 70, 609–625.

Kieke, M.C., Titus, M.A., 2003. The Myosin Superfamily: An Overview, in: Schliwa, M. (Ed.), Molecular Motors. Wiley-VCH, Weinheim, pp. 3–44.

Knight, P., Trinick, J., 1984. Structure of the myosin projections on native thick filaments from vertebrate skeletal muscle. J Mol Biol 177, 461–482.

Knight, P.J., Thirumurugan, K., Xu, Y., Wang, F., Kalverda, A.P., Stafford, W.F., Sellers, J.R., Peckham, M., 2005. The predicted coiled-coil domain of myosin 10 forms a novel elongated domain that lengthens the head. J Biol Chem 280, 34702–34708.

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

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

Kollmar, M., Dürrwang, U., Kliche, W., Manstein, D.J., Kull, F.J., 2002. Crystal structure of the motor domain of a class-I myosin. EMBO Journal 21, 2517–2525.

Kon, T., Sutoh, K., Kurisu, G., 2011. X-ray structure of a functional full-length dynein motor domain. Nat Struct Mol Biol 18, 638–642.

Kovács, M., Thirumurugan, K., Knight, P.J., Sellers, J.R., 2007. Load-dependent mechanism of nonmuscle myosin 2. Proc Natl Acad Sci USA 104, 9994–9999.

Köhler, D., Ruff, C., Meyhöfer, E., Bähler, M., 2003. Different degrees of lever arm rotation control myosin step size. J Cell Biol 161, 237–241.

Krementsov, D.N., 2004. Myosin V: regulation by calcium, calmodulin, and the tail domain. The Journal of Cell Biology 164, 877–886.

Krendel, M., Mooseker, M.S., 2005. Myosins: Tails (and Heads) of Functional Diversity 20, 239–251.

Kron, S.J., Spudich, J.A., 1986. Fluorescent actin filaments move on myosin fixed to a glass surface. Proc Natl Acad Sci USA 83, 6272–6276.

Kuboniwa, H., Tjandra, N., Grzesiek, S., Ren, H., Klee, C.B., Bax, A., 1995. Solution structure of calcium-free calmodulin. Nat Struct Mol Biol 2, 768–776.

La Cruz, De, E., 2009. Kinetic and equilibrium analysis of the myosin ATPase. Meth Enzymol.

La Cruz, De, E., Wells, A., Rosenfeld, S., Ostap, E., Sweeney, H., 1999. The kinetic mechanism of myosin V. Proc Natl Acad Sci USA 96, 13726–13731.

La Cruz, De, E.M., Ostap, E.M., 2004. Relating biochemistry and function in the myosin

superfamily. Current Opinion in Cell Biology 16, 61–67.

La Cruz, De, E.M., Sweeney, H.L., Ostap, E.M., 2000. ADP inhibition of myosin V ATPase activity. Biophysical Journal 79, 1524–1529.

La Cruz, De, E.M., Wells, A.L., Sweeney, H.L., Ostap, E.M., 2000. Actin and light chain isoform dependence of myosin V kinetics. Biochemistry 39, 14196–14202.

Laakso, J.M., Lewis, J.H., Shuman, H., Ostap, E.M., 2010. Control of myosin-I force sensing by alternative splicing. Proc Natl Acad Sci USA 107, 698–702.

Laemmli, U.K., 1970. Cleavage of Structural Proteins During Assembly of Head of Bacteriophage-T4. Nature 227, 680–&.

Lagacé, L., Chandra, T., Woo, S.L., Means, A.R., 1983. Identification of multiple species of calmodulin messenger RNA using a full length complementary DNA. J Biol Chem 258, 1684–1688.

Lan, G., Sun, S., 2005. Dynamics of myosin-V processivity. Biophysical Journal 88, 999–1008.

Lan, G., Sun, S., 2006. Flexible Light-Chain and Helical Structure of F-Actin Explain the Movement and Step Size of Myosin-VI. Biophysical Journal 91, 4002–4013.

Landau, L.D., Lifshitz, E.M., 1970. Theory of Elasticity, Course of Theoretical Physics. Pergamon Press, New York.

Larson, R.E., Espindola, F.S., Espreafico, E.M., 1990. Calmodulin-binding proteins and calcium/calmodulin-regulated enzyme activities associated with brain actomyosin. J. Neurochem. 54, 1288–1294.

Larson, R.E., Pitta, D.E., Ferro, J.A., 1988. A novel 190 kDa calmodulin-binding protein associated with brain actomyosin. Brazilian journal of medical and ….

Lehrer, S., 1972. Intrinsic fluorescence of actin - Biochemistry (ACS Publications). Biochemistry.

Lewalle, A., Steffen, W., Stevenson, O., Ouyang, Z., Sleep, J., 2008. Single-Molecule Measurement of the Stiffness of the Rigor Myosin Head. Biophysical Journal 94, 2160–2169.

Lewis, J.H., Beausang, J.F., Sweeney, H.L., Goldman, Y.E., 2012. The azimuthal path of myosin V and its dependence on lever-arm length. J. Gen. Physiol. 139, 101–120.

Li, X.E., Holmes, K.C., Lehman, W., Jung, H., Fischer, S., 2010. The shape and flexibility of tropomyosin coiled coils: implications for actin filament assembly and regulation. J Mol Biol 395, 327–339.

Li, X.E., Lehman, W., Fischer, S., 2010. The relationship between curvature, flexibility and persistence length in the tropomyosin coiled-coil. J Struct Biol 170, 313–318.

Life Technologies, Haugland, R.P., 2010. The Molecular Probes Handbook, 11th ed. Life Technologies.

Linari, M., Caremani, M., Piperio, C., Brandt, P., Lombardi, V., 2007. Stiffness and fraction of Myosin motors responsible for active force in permeabilized muscle fibers from rabbit psoas. Biophysical Journal 92, 2476–2490.

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

Lorenz, M., Popp, D., Holmes, K.C., 1993. Refinement of the F-Actin Model against X-ray Fiber Diffraction Data by the Use of a Directed Mutation Algorithm. J Mol Biol 234, 826–836.

Lowey, S., Slayter, H.S., Weeds, A.G., Baker, H., 1969. Substructure of the myosin molecule. I. Subfragments of myosin by enzymic degradation. J Mol Biol 42, 1–29.

References

Lu, H., Krementsova, E.B., Trybus, K.M., 2006. Regulation of myosin V processivity by calcium at the single molecule level. J Biol Chem 281, 31987–31994.

Lymn, R.W., Taylor, E.W., 1971. Mechanism of adenosine triphosphate hydrolysis by actomyosin. Biochemistry 10, 4617–4624.

McCullough, B.R., Blanchoin, L., Martiel, J.-L., La Cruz, De, E.M., 2008. Cofilin increases the bending flexibility of actin filaments: implications for severing and cell mechanics. J Mol Biol 381, 550–558.

McDonnell, A.V., Jiang, T., Keating, A.E., Berger, B., 2006. Paircoil2: improved prediction of coiled coils from sequence. Bioinformatics 22, 356–358.

McGough, A., Pope, B., Chiu, W., Weeds, A., 1997. Cofilin changes the twist of F-actin: implications for actin filament dynamics and cellular function. J Cell Biol 138, 771–781.

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

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

Mermall, V., Post, P., Mooseker, M., 1998. Unconventional Myosins in Cell Movement, Membrane Traffic, and Signal Transduction. Science 279, 527–533.

Ménétrey, J., Bahloul, A., Wells, A.L., Yengo, C.M., Morris, C.A., Sweeney, H.L., Houdusse, A., 2005. The structure of the myosin VI motor reveals the mechanism of directionality reversal. Nature 435, 779–785.

Molloy, J., Burns, J., Sparrow, J., Tregear, R., Kendrick-Jones, J., White, D., 1995. Single-molecule mechanics of heavy meromyosin and S1 interacting with rabbit or Drosophila actins using optical tweezers. Biophysical Journal 68, 298S–303S; 303S–305S.

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

Moore, J.R., Krementsova, E.B., Trybus, K.M., Warshaw, D.M., 2001. Myosin V exhibits a high duty cycle and large unitary displacement. J Cell Biol 155, 625–635.

Mooseker, M.S., Foth, B.J., 2007. Structural and functional diversity of the myosin family, in: Coluccio, L.M. (Ed.), Myosins: a Superfamiy of Molecular Motors. Springer, Dordrecht, pp. 1–34.

Mukherjea, M., Llinas, P., Kim, H., Travaglia, M., Safer, D., Ménétrey, J., Franzini-Armstrong, C., Selvin, P., Houdusse, A., Sweeney, H., 2009. Myosin VI dimerization triggers an unfolding of a three-helix bundle in order to extend its reach. Mol Cell 35, 305–315.

Mücke, N., Klenin, K., Kirmse, R., Bussiek, M., Herrmann, H., Hafner, M., Langowski, J., 2009. Filamentous biopolymers on surfaces: atomic force microscopy images compared with Brownian dynamics simulation of filament deposition. PLoS ONE 4, e7756.

Mücke, N., Kreplak, L., Kirmse, R., Wedig, T., Herrmann, H., Aebi, U., Langowski, J., 2004. Assessing the flexibility of intermediate filaments by atomic force microscopy. J Mol Biol 335, 1241–1250.

Nagy, A., Piszczek, G., Sellers, J., 2009. Extensibility of the Extended Tail Domain of Processive and Nonprocessive Myosin V Molecules. Biophysical Journal 97, 3123–3131.

Nagy, S., Ricca, B.L., Norstrom, M.F., Courson, D.S., Brawley, C.M., Smithback, P.A., Rock, R.S., 2008. A myosin motor that selects bundled actin for motility. Proceedings of the National Academy of Sciences 105, 9616–9620.

Nascimento, A.A., Cheney, R.E., Tauhata, S.B., Larson, R.E., Mooseker, M.S., 1996. Enzymatic characterization and functional domain mapping of brain myosin-V. J Biol Chem 271, 17561–17569.

Nelson, M.R., Chazin, W.J., 1998. Structures of EF-hand Ca(2+)-binding proteins: diversity in

the organization, packing and response to Ca2+ binding. Biometals 11, 297–318.

Neuhaus, J.M., Wanger, M., Keiser, T., Wegner, A., 1983. Treadmilling of actin. Journal of Muscle Research and Cell Motility 4, 507–527.

Neuwald, A.F., Aravind, L., Spouge, J.L., Koonin, E.V., 1999. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res. 9, 27–43.

Nguyen, H., Higuchi, H., 2005. Motility of myosin V regulated by the dissociation of single calmodulin. Nat Struct Mol Biol 12, 127–132.

O'Brien, E.J., Gillis, J.M., Couch, J., 1975. Symmetry and molecular arrangement in paracrystals of reconstituted muscle thin filaments. J Mol Biol 99, 461–475.

O'Neil, K.T., DeGrado, W.F., 1990. How calmodulin binds its targets: sequence independent recognition of amphiphilic α-helices. Trends in Biochemical Sciences 15, 59–64.

Oda, T., Iwasa, M., Aihara, T., Maéda, Y., Narita, A., 2009. The nature of the globular- to fibrous-actin transition. Nature 457, 441–445.

Odronitz, F., Kollmar, M., 2007. Drawing the tree of eukaryotic life based on the analysis of 2,269 manually annotated myosins from 328 species. Genome Biol. 8, R196.

Ohi, M., Li, Y., Cheng, Y., Walz, T., 2004. Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy. Biol Proced Online 6, 23–34.

Oke, O., 2004. Electron Microscopy of Myosin V Molecules. PhD thesis, University of Leeds.

Oke, O.A., Burgess, S.A., Forgacs, E., Knight, P.J., Sakamoto, T., Sellers, J.R., White, H., Trinick, J., 2010. Influence of lever structure on myosin 5a walking. Proc Natl Acad Sci USA 107, 2509–2514.

Ott, A., Magnasco, M., Simon, A., Libchaber, A., 1993. Measurement of the persistence length of polymerized actin using fluorescence microscopy. Phys. Rev. E 48, R1642–R1645.

Otterbein, L.R., Graceffa, P., Dominguez, R., 2001. The crystal structure of uncomplexed actin in the ADP state. Science 293, 708–711.

Panchuk-Voloshina, N., Haugland, R.P., Bishop-Stewart, J., Bhalgat, M.K., Millard, P.J., Mao, F., Leung, W.Y., Haugland, R.P., 1999. Alexa dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates. J. Histochem. Cytochem. 47, 1179–1188.

Pardee, J., Spudich, J.A., 1982. Purification of muscle actin. Meth Enzymol 85, 164–181.

Parry, D.A.D., Fraser, R.D.B., Squire, J.M., 2008. Fifty years of coiled-coils and alpha-helical bundles: a close relationship between sequence and structure. J Struct Biol 163, 258–269.

Peckham, M., Knight, P., 2009. When a predicted coiled coil is really a single α-helix, in myosins and other proteins. Soft Matter 5, 2493–2503.

Penczek, P., Radermacher, M., Frank, J., 1992. Three-dimensional reconstruction of single particles embedded in ice. Ultramicroscopy 40, 33–53.

Persechini, A., Moncrief, N.D., Kretsinger, R.H., 1989. The EF-hand family of calcium-modulated proteins. Trends in Neurosciences 12, 462–467.

Piazzesi, G., Lucii, L., Lombardi, V., 2002. The size and the speed of the working stroke of muscle myosin and its dependence on the force. J. Physiol. (Lond.) 545, 145–151.

Pierobon, P., Achouri, S., Courty, S., Dunn, A.R., Spudich, J.A., Dahan, M., Cappello, G., 2009. Velocity, processivity, and individual steps of single myosin V molecules in live cells. Biophysical Journal 96, 4268–4275.

Pollard, T.D., Craig, S.W., 1982. Mechanism of actin polymerization. Trends in Biochemical Sciences.

Pollard, T.D., Korn, E.D., 1973. Acanthamoeba myosin. I. Isolation from Acanthamoeba castellanii of an enzyme similar to muscle myosin. J Biol Chem 248, 4682–4690.

References

Purcell, T., Morris, C., Spudich, J., Sweeney, H., 2002. Role of the lever arm in the processive stepping of myosin V. Proc Natl Acad Sci USA 99, 14159–14164.

Purcell, T., Sweeney, H., Spudich, J., 2005. A force-dependent state controls the coordination of processive myosin V. Proc Natl Acad Sci USA 102, 13873–13878.

Radermacher, M., Rao, V., Grassucci, R., Frank, J., Timerman, A.P., Fleischer, S., Wagenknecht, T., 1994. Cryo-electron microscopy and three-dimensional reconstruction of the calcium release channel/ryanodine receptor from skeletal muscle. J Cell Biol 127, 411–423.

Rayment, I., Holden, H.M., Whittaker, M., Yohn, C.B., Lorenz, M., Holmes, K.C., Milligan, R.A., 1993. Structure of the actin-myosin complex and its implications for muscle contraction. Science 261, 58–65.

Rayment, I., Rypniewski, W.R., Schmidt-Bäse, K., Smith, R., Tomchick, D.R., Benning, M.M., Winkelmann, D.A., Wesenberg, G., Holden, H.M., 1993. Three-dimensional structure of myosin subfragment-1: a molecular motor. Science 261, 50–58.

Reck-Peterson, S., Provance, D., Mooseker, M., Mercer, J., 2000. Class V myosins. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1496, 36–51.

Reif, F., 1965. Fundamentals of Statistical and Thermal Physics. McGraw-Hill, New York.

Rhoads, A.R., Friedberg, F., 1997. Sequence motifs for calmodulin recognition. FASEB J. 11, 331–340.

Rice, S., Lin, A.W., Safer, D., Hart, C.L., Naber, N., Carragher, B.O., Cain, S.M., Pechatnikova, E., Wilson-Kubalek, E.M., Whittaker, M., Pate, E., Cooke, R., Taylor, E.W., Milligan, R.A., Vale, R.D., 1999. A structural change in the kinesin motor protein that drives motility. Nature 402, 778–784.

Rief, M., Rock, R., Mehta, A., Mooseker, M., Cheney, R., Spudich, J., 2000. Myosin-V stepping kinetics: A molecular model for processivity. Proc Natl Acad Sci USA 97, 9482–9486.

Rivetti, C., Guthold, M., 1996. Scanning Force Microscopy of DNA Deposited onto Mica: Equilibration *versus* Kinetic Trapping Studied by Statistical Polymer Chain Analysis. J Mol Biol 264, 919–932.

Roberts, A., Numata, N., Walker, M., Kato, Y., Malkova, B., Kon, T., Ohkura, R., Arisaka, F., Knight, P., Sutoh, K., Burgess, S., 2009. AAA+ Ring and Linker Swing Mechanism in the Dynein Motor. Cell 136, 485–495.

Roberts, A.J., 2009. Structural studies on the mechanism of dynein. PhD thesis, University of Leeds.

Rodriguez, O.C., Cheney, R.E., 2002. Human myosin-Vc is a novel class V myosin expressed in epithelial cells. J. Cell. Sci. 115, 991–1004.

Rosenfeld, S., Sweeney, H.L., 2004. A Model of Myosin V Processivity. J Biol Chem 279, 40100–40111.

Ruff, C., Furch, M., Brenner, B., Manstein, D.J., Meyhöfer, E., 2001. Single-molecule tracking of myosins with genetically engineered amplifier domains. Nat Struct Mol Biol 8, 226–229.

Sakamoto, T., Amitani, I., Yokota, E., Ando, T., 2000. Direct observation of processive movement by individual myosin V molecules. Biochem Biophys Res Commun 272, 586–590.

Sakamoto, T., Wang, F., Schmitz, S., Xu, Y., Xu, Q., Molloy, J., Veigel, C., Sellers, J., 2003. Neck length and processivity of myosin V. J Biol Chem 278, 29201–29207.

Sakamoto, T., Webb, M., Forgacs, E., White, H., Sellers, J., 2008. Direct observation of the mechanochemical coupling in myosin Va during processive movement. Nature 455, 128–132.

Sakamoto, T., Yildiz, A., Selvin, P.R., Sellers, J.R., 2005. Step-size is determined by neck length in myosin V. Biochemistry 44, 16203–16210.

Schägger, H., 2006. Tricine-SDS-PAGE. Nat Protoc 1, 16–22.

Scheres, S.H.W., Gao, H., Valle, M., Herman, G.T., Eggermont, P.P.B., Frank, J., Carazo, J.-M., 2007. Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. Nat. Methods 4, 27–29.

Schilstra, M.J., Martin, S.R., 2006. An elastically tethered viscous load imposes a regular gait on the motion of myosin-V. Simulation of the effect of transient force relaxation on a stochastic process. J R Soc Interface 3, 153–165.

Schliwa, M., Woehlke, G., 2003. Molecular motors. Nature 422, 759–765.

Sellers, J., Thirumurugan, K., Sakamoto, T., Hammer, J.A., Knight, P.J., 2008. Calcium and cargoes as regulators of myosin 5a activity. Biochem Biophys Res Commun 369, 176–181.

Sellers, J., Veigel, C., 2006. Walking with myosin V. Current Opinion in Cell Biology 18, 68–73.

Sellers, J.R., 1991. Regulation of cytoplasmic and smooth muscle myosin. Current Opinion in Cell Biology 3, 98–104.

Sellers, J.R., 1999. Myosins. Oxford University Press, USA.

Sellers, J.R., 2000. Myosins: a diverse superfamily. Biochimica et Biophysica Acta.

Sellers, J.R., Weisman, L.S., 2007. Myosin V, in: Coluccio, L.M. (Ed.), Myosins: a Superfamily of Molecular Motors. Springer, Dordrecht, pp. 265–323.

Seperack, P.K., Mercer, J.A., Strobel, M.C., Copeland, N.G., Jenkins, N.A., 1995. Retroviral sequences located within an intron of the dilute gene alter dilute expression in a tissue-specific manner. EMBO Journal 14, 2326–2332.

Shiroguchi, K., Chin, H.F., Hannemann, D.E., Muneyuki, E., La Cruz, De, E.M., Kinosita, K., 2011. Direct Observation of the Myosin Va Recovery Stroke That Contributes to Unidirectional Stepping along Actin. PLoS Biol. 9, e1001031.

Sivaramakrishnan, S., Sung, J., Ali, M., Doniach, S., Flyvbjerg, H., Spudich, J., 2009. Combining Single-Molecule Optical Trapping and Small-Angle X-Ray Scattering Measurements to Compute the Persistence Length of a Protein ER/K α-Helix. Biophysical Journal 97, 2993–2999.

Skau, K., Hoyle, R., Turner, M., 2006. A Kinetic Model Describing the Processivity of Myosin-V. Biophysical Journal 91, 2475–2489.

Smith, C.A., Rayment, I., 1996a. Active site comparisons highlight structural similarities between myosin and other P-loop proteins. Biophysical Journal 70, 1590–1602.

Smith, C.A., Rayment, I., 1996b. X-ray structure of the magnesium(II).ADP.vanadate complex of the Dictyostelium discoideum myosin motor domain to 1.9 A resolution. Biochemistry 35, 5404–5417.

Song, C.F., Sader, K., White, H., Kendrick-Jones, J., Trinick, J., 2010. Nucleotide-dependent shape changes in the reverse direction motor, myosin VI. Biophysical Journal 99, 3336–3344.

Spink, B.J., Sivaramakrishnan, S., Lipfert, J., Doniach, S., Spudich, J.A., 2008. Long single alpha-helical tail domains bridge the gap between structure and function of myosin VI. Nat Struct Mol Biol 15, 591–597.

Spudich, J.A., 1994. How molecular motors work. Nature 372, 515–518.

Steffen, W., Smith, D., Simmons, R., Sleep, J., 2001. Mapping the Actin Filament with Myosin. Proc Natl Acad Sci USA 98, 14949–14954.

Stevens, R.C., Davis, T.N., 1998. Mlc1p is a light chain for the unconventional myosin Myo2p

in Saccharomyces cerevisiae. J Cell Biol 142, 711–722.

Sun, Y., Goldman, Y.E., 2011. Lever-arm mechanics of processive myosins. Biophysical Journal 101, 1–11.

Sun, Y., Schroeder, H.W., Beausang, J.F., Homma, K., Ikebe, M., Goldman, Y.E., 2007. Myosin VI walks "wiggly" on actin with large and variable tilting. Mol Cell 28, 954–964.

Sweeney, H., Houdusse, A., 2010. Myosin VI Rewrites the Rules for Myosin Motors. Cell 141, 573–582.

Syed, S., Snyder, G., Franzini-Armstrong, C., Selvin, P., Goldman, Y., 2006. Adaptability of myosin V studied by simultaneous detection of position and orientation. EMBO Journal 25, 1795–1803.

Szent-Györgyi, A.G., 1953. Meromyosins, the subunits of myosin. Arch Biochem Biophys 42, 305–320.

Takagi, Y., Yang, Y., Fujiwara, I., Jacobs, D., Cheney, R.E., Sellers, J.R., Kovács, M., 2008. Human myosin Vc is a low duty ratio, nonprocessive molecular motor. J Biol Chem 283, 8527–8537.

Tama, F., Miyashita, O., Brooks, C.L., 2004. Normal mode based flexible fitting of high-resolution structure into low-resolution experimental data from cryo-EM. J Struct Biol 147, 315–326.

Taylor, K.A., Schmitz, H., Reedy, M.C., Goldman, Y.E., Franzini-Armstrong, C., Sasaki, H., Tregear, R.T., Poole, K., Lucaveche, C., Edwards, R.J., Chen, L.F., Winkler, H., Reedy, M.K., 1999. Tomographic 3D reconstruction of quick-frozen, Ca2+-activated contracting insect flight muscle. Cell 99, 421–431.

Terrak, M., Rebowski, G., Lu, R., Grabarek, Z., Dominguez, R., 2005. Structure of the light chain-binding domain of myosin V. Proc Natl Acad Sci USA 102, 12718–12723.

Terrak, M., Wu, G., Stafford, W., Lu, R., Dominguez, R., 2003. Two distinct myosin light chain structures are induced by specific variations within the bound IQ motifs - functional implications. EMBO Journal 22, 362–371.

Thirumurugan, K., Burgess, S.A., Trinick, J., Wang, F., White, H.D., Knight, P.J., 2003. Temperature effects on myosin V head conformation. Biophysical Journal 84, 116A–116A.

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

Tokunaga, M., Kitamura, K., Saito, K., Iwane, A.H., Yanagida, T., 1997. Single molecule imaging of fluorophores and enzymatic reactions achieved by objective-type total internal reflection fluorescence microscopy. Biochem Biophys Res Commun 235, 47–53.

Tominaga, M., Kojima, H., Yokota, E., Nakamori, R., Anson, M., Shimmen, T., Oiwa, K., 2012. Calcium-induced Mechanical Change in the Neck Domain Alters the Activity of Plant Myosin XI. J Biol Chem 287, 30711–30718.

Trentham, D.R., Bardsley, R.G., Eccleston, J.F., Weeds, A.G., 1972. Elementary processes of the magnesium ion-dependent adenosine triphosphatase activity of heavy meromyosin. A transient kinetic approach to the study of kinases and adenosine triphosphatases and a colorimetric inorganic phosphate assay in situ. Biochem. J. 126, 635–644.

Trinick, J., Cooper, J., Seymour, J., Egelman, E.H., 1986. Cryo-electron microscopy and three-dimensional reconstruction of actin filaments. J Microsc 141, 349–360.

Trinick, J., Elliott, A., 1979. Electron microscope studies of thick filaments from vertebrate skeletal muscle. J Mol Biol 131, 133–136.

Trinick, J.A., 1981. End-filaments: a new structural element of vertebrate skeletal muscle thick filaments. J Mol Biol 151, 309–314.

Trybus, K., Gushchin, M., Lui, H., Hazelwood, L., Krementsova, E., Volkmann, N., Hanein,

D., 2007. Effect of calcium on calmodulin bound to the IQ motifs of myosin V. J Biol Chem 282, 23316–23325.

Trybus, K.M., Waller, G.S., Chatman, T.A., 1994. Coupling of ATPase activity and motility in smooth muscle myosin is mediated by the regulatory light chain. J Cell Biol 124, 963–969.

Tskhovrebova, L., Trinick, J., 1997. Direct visualization of extensibility in isolated titin molecules. J Mol Biol 265, 100–106.

Tskhovrebova, L., Trinick, J., 2001. Flexibility and extensibility in the titin molecule: analysis of electron microscope data. J Mol Biol 310, 755–771.

Tskhovrebova, L., Walker, M.L., Grossmann, J.G., Khan, G.N., Baron, A., Trinick, J., 2010. Shape and flexibility in the titin 11-domain super-repeat. J Mol Biol 397, 1092–1105.

Ungewickell, E., Branton, D., 1981. Assembly units of clathrin coats. Nature 289, 420–422.

Uyeda, T., Abramson, P., Spudich, J., 1996. The neck region of the myosin motor domain acts as a lever arm to generate movement. Proc Natl Acad Sci USA 93, 4459–4464.

Vale, R.D., Milligan, R.A., 2000. The way things move: looking under the hood of molecular motor proteins. Science 288, 88–95.

van Heel, M., 1984. Multivariate statistical classification of noisy images (randomly oriented biological macromolecules). Ultramicroscopy 13, 165–183.

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

Veigel, C., Wang, F., Bartoo, M., Sellers, J., Molloy, J., 2002. The gated gait of the processive molecular motor, myosin V. Nature Cell Biology 4, 59–65.

Vilfan, A., 2005. Elastic lever-arm model for myosin V. Biophysical Journal 88, 3792–3805.

Volkmann, N., Liu, H., Hazelwood, L., Krementsova, E., Lowey, S., Trybus, K., Hanein, D., 2005. The Structural Basis of Myosin V Processive Movement as Revealed by Electron Cryomicroscopy. Mol Cell 19, 595–605.

Wagner, W., Brenowitz, S.D., Hammer, J.A., 2011. Myosin-Va transports the endoplasmic reticulum into the dendritic spines of Purkinje neurons. Nature Cell Biology 13, 40–48.

Wagner, W., Hammer, J.A., 2003. Myosin V and the endoplasmic reticulum: the connection grows. J Cell Biol 163, 1193–1196.

Walker, M., Knight, P., Trinick, J., 1985. Negative staining of myosin molecules. J Mol Biol 184, 535–542.

Walker, M., Knight, P., Trinick, J., 1991. Properties of the myosin molecule revealed by negative staining. Micron and Microscopica Acta 22, 413–422.

Walker, M., Trinick, J., 1989. Electron microscopy of negatively stained scallop myosin molecules. Effect of regulatory light chain removal on head structure. J Mol Biol 208, 469–475.

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

Wall, L., Christiansen, T., Orwant, J., 2000. Programming Perl, 3rd ed. O'Reiley & Associates, Inc., Sebastopol.

Wallace, R.W., Tallant, E.A., Cheung, W.Y., 1983. Assay of calmodulin by Ca2+-dependent phosphodiesterase. Meth Enzymol 102, 39–47.

Wang, F., Chen, L., Arcucci, O., Harvey, E., Bowers, B., Xu, Y., Hammer, J., Sellers, J., 2000. Effect of ADP and ionic strength on the kinetic and motile properties of recombinant mouse myosin V. J Biol Chem 275, 4329–4335.

Wang, F., Thirumurugan, K., Stafford, W.F., Hammer, J.A., Knight, P.J., Sellers, J.R., 2004. Regulated conformation of myosin V. J Biol Chem 279, 2333–2336.

Wang, Z., Edwards, J.G., Riley, N., Provance, D.W., Karcher, R., Li, X.-D., Davison, I.G., Ikebe, M., Mercer, J.A., Kauer, J.A., Ehlers, M.D., 2008. Myosin Vb mobilizes recycling endosomes and AMPA receptors for postsynaptic plasticity. Cell 135, 535–548.

Warshaw, D.M., Guilford, W.H., Freyzon, Y., Krementsova, E., Palmiter, K.A., Tyska, M.J., Baker, J.E., Trybus, K.M., 2000. The light chain binding domain of expressed smooth muscle heavy meromyosin acts as a mechanical lever. J Biol Chem 275, 37167–37172.

Warshaw, D.M., Kennedy, G.G., Work, S.S., Krementsova, E.B., Beck, S., Trybus, K.M., 2005. Differential labeling of myosin V heads with quantum dots allows direct visualization of hand-over-hand processivity. Biophysical Journal 88, L30–2.

Wells, A.L., Lin, A.W., Chen, L.Q., Safer, D., Cain, S.M., Hasson, T., Carragher, B.O., Milligan, R.A., Sweeney, H.L., 1999. Myosin VI is an actin-based motor that moves backwards. Nature 401, 505–508.

Wiggins, C.H., Riveline, D., Ott, A., Goldstein, R.E., 1998. Trapping and wiggling: elastohydrodynamics of driven microfilaments. Biophysical Journal 74, 1043–1060.

Wolenski, J.S., 1995. Regulation of calmodulin-binding myosins. Trends in Cell Biology 5, 310–316.

Yamamoto, K., 2007. Plant Myosins VIII, XI, and XIII, in: Coluccio, L.M. (Ed.), Myosins: a Superfamily of Molecular Motors. Springer, Dordrecht.

Yang, Y., Baboolal, T.G., Siththanandan, V., Chen, M., Walker, M.L., Knight, P.J., Peckham, M., Sellers, J.R., 2009. A FERM domain autoregulates Drosophila myosin 7a activity. Proc Natl Acad Sci USA 106, 4189–4194.

Yildiz, A., Forkey, J., McKinney, S., Ha, T., Goldman, Y., Selvin, P., 2003. Myosin V Walks Hand-Over-Hand: Single Fluorophore Imaging with 1.5-nm Localization. Science 300, 2061–2065.

Zhang, M., Tanaka, T., Ikura, M., 1995. Calcium-induced conformational transition revealed by the solution structure of apo calmodulin. Nat Struct Mol Biol 2, 758–767.

Zhao, F.-Q., Craig, R., 2003. Capturing time-resolved changes in molecular structure by negative staining. J Struct Biol 141, 43–52.

Zhao, L.P., Koslovsky, J.S., Reinhard, J., Bähler, M., Witt, A.E., Provance, D.W., Mercer, J.A., 1996. Cloning and characterization of myr 6, an unconventional myosin of the dilute/myosin-V family. Proc Natl Acad Sci USA 93, 10826–10831.