Development and optimisation of a new single molecule assay to study kinesin-1 activation dynamics

Evelyn Rose Smith

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield

Faculty of Medicine, Dentistry and Health

Department of Neuroscience

April, 2023

Abstract

Kinesin-1 is an essential antereograde motor protein travelling along microtubule tracks distributing cargo throughout the cell. It mainly exists as a heterotetramer composed of two heavy chains and two light chains. The heavy chain consists of a N-terminal motor domain, that hydrolyses ATP to walk along the microtubules, and a C-terminal tail domain that binds the light chains and cargo. It is known that kinesin-1 is autoinhibited to prevent unnecessary movement and ATP hydrolysis, but little is known about how kinesin-1 is then activated. Autoinhibition occurs through direct interaction between the motor domains and inhibitory motifs in the tail domain. This results in two known conformations of kinesin-1; a folded inactive state and a elongated active state, both of which are thought to be stable in physiological conditions. Although significant advances have occurred in our understanding of kinesin-1 movement, mainly through single molecule experiments, knowledge around kinesin-1 autoinhibition and activation is still lacking. This research shows the development of a new single molecule fluorescence resonance energy transfer (smFRET) assay to study conformational changes in kinesin-1. This assay allows kinesin-1 activation to be directly studied, with full rotational freedom at a single-molecule level, for the first time both in the presence and absence of microtubules. This works shows that the current binary on/off model of kinesin-1 autoinhibition is too simplistic and the folded inactive conformation might not be stable. Using this assay the inherent flexibility in kinesin-1 has been uncovered giving insight into the potential complex regulation of kinesin-1.

Declaration

I, Evelyn Smith, confirm that the Thesis is my own work, I am aware of the University's Guidance on the Use of Unfair Means. The work presented and described in this thesis was undertaken at the University of Sheffield between October 2018 and December 2022 under the supervision of Dr. Alison Twelvetrees and Dr. Timothy Craggs. This work has not been previously presented for an award at this, or any other, university.

Acknowledgements

The biggest thank you has to go to the Twelvetrees Lab. Ali Twelvetrees, you have been the best, most amazingly supportive supervisor; especially through a global pandemic! The fact that you offered me a PhD position after the worst application ever and with limited biology knowledge (which has only slightly improved) is astounding. Thank for all your help and support - I can not thank you enough. Annie Savage and Becca Mighell, you helped start off all the kinesin-1 FRET work and were the best masters students ever. Ashleigh Davey and Emma Turner, without you I would have gone even crazier in the lab then I have already. Your continual help, support and distractions have made the lab a happy place. This was especially true in those long difficult COVID shift patterns. Talking of COVID, I have to thank my COVID bubble. Olivia Hill thanks for being the best bubble buddy and for always being there with a glass of Prosecco to either celebrate or commiserate the science. This project would not have been possible without the immense help and support of my second supervisor, Tim Craggs, and his lab. Ben, Elliot, Alice and Mahmoud thanks for all the help with the box but also your friendship along the way. Also Alice your help organising SMaSH (Single Molecules at Sheffield) has been incredible.

Now the non-science people. Amber, Charlotte, Jody, Megan and Meghan, your friendship over the years has meant the world to me. Thanks for all the amazing, hilarious times and your continual support and guidance. You have all been there for me even when I am being rubbish and not replying to messages! Lastly, to

my family. The fairy Godparents, Tim and Ian, for the constant Noodle Doodle. Thank you Grandma and Grandad for always being there for me. To Mum, Dad, Patrick, Michael and Selena - you have supported, advised and actually listened to me talking about science! Even though you are all completely crazy, I couldn't have done it without you. Your support and kindness means the world.

Contents

1 Introduction **17**

1.1 Microtubule based transport 17

1.2 The kinesin superfamily 19

1.3 Kinesin-1 20

1.3.1 The role of kinesin-1 20

1.3.2 The structure of kinesin-1 20

1.4 Single molecule experiments to study kinesin-1 movement 24

1.5 Regulation of kinesin-1 30

1.5.1 Adaptors 30

1.5.2 Autoinhibition and activation 31

1.6 A single molecule approach to directly study kinesin-1 conformation . 35

1.7 Fluorescence resonance energy transfer 36

1.7.1 FRET can be used to study conformational changes in molecules 37

1.7.2 FRET standards have been used to ensure reproduciblity and accuracy of smFRET efficiency measurements 38

1.7.3 A single molecule fluorescence resonance energy transfer (smFRET) custom built confocal instrument 40

1.8 Thesis aims 44

2 Methods **45**

2.1 Molecular Biology 45

2.1.1 Constructs 45

2.1.2 Bacterial culture and cloning 48
2.2 Cell Culture and transfection 56
2.2.1 Culture of HEK293 cells 56
2.2.2 Transfection of HEK293 cells 56
2.2.3 Ligand incubation and washing of cells 57
2.2.4 Collecting transfected and treated HEK293 cells with trypsin . 58
2.2.5 Collecting transfected HEK293 cells without trypsin 58
2.2.6 Cell lysis of HEK293 cells with detergent 58
2.2.7 Cell lysis of HEK293 cells without detergent - homogenisation 59
2.2.8 Sample preparation for SDS-PAGE 60
2.3 smFRET 60
2.3.1 Initial smFRET assay protocol 62
2.3.2 Photo bleaching BSA 62
2.3.3 Optimised protocol for kinesin-1 conformation smFRET assay 63
2.4 Flow cytometry 65
2.4.1 Determining washout of fluorescent ligands 65
2.5 Biochemistry 65
2.5.1 Co-immunoprecipitation 65
2.5.2 SDS page, native-PAGE and western blotting 66
2.6 Imaging 69
2.6.1 Immunofluorescence staining of HEK293 cells 69
2.7 *In vitro* reconstitution 70
2.7.1 Polymerising microtubules 70
2.8 Protein purification 70
2.9 Computional modelling 71
2.9.1 AlphaFold2 71

3 Developing the use of self labelling enzymes in smFRET experiments 75

3.1 Introduction 75
3.1.1 Self-labelling enzymes and organic dyes 75
3.1.2 Using self-labelling enzymes in smFRET assays 76
3.2 Results 79
3.2.1 Validating the use of the self-labelling enzymes and cell lysate samples 79
3.2.2 Optimisation of the smFRET assay 82
3.2.3 Post optimisation smFRET data 98
3.3 Discussion 99

4 Studying kinesin-1 conformation by smFRET 107

4.1 Introduction 107
4.1.1 Two known conformations of the kinesin-1 heavy chain 107
4.1.2 Altering the conformation of kinesin-1 109
4.2 Results 112
4.2.1 smFRET analysis of CLIPf-mKIF5A-SNAPf-HisTag-FLAG . . 112
4.2.2 Heavy chain heterodimer formation 114
4.2.3 Low stoichiometry of the CLIPf-mKIF5A-SNAPf fusion protein 117
4.2.4 Addition of salt to the smFRET assay buffer 122
4.2.5 Deletion mutants of the mKIF5A fusion protein 126
4.2.6 AlphaFold2 predicts a new structure for the KIF5A stalk . . . 128
4.3 Discussion 133

5 Increasing the complexity of the kinesin-1 smFRET assay 138

5.1 Introduction 138
5.1.1 Axonal transport 138
5.1.2 Microtubule tracks 139
5.1.3 The conformation of kinesin-1 attached to cargo 140
5.2 Results 141

5.2.1 Studying heavy chain kinesin-1 conformation when attached to endogenous cargo . . . 141
5.2.2 Studying heavy chain kinesin-1 conformation in the presence of microtubules . . . 151
5.3 Discussion . . . 155

6 Investigating the potential of using smFRET to validate AlphaFold2 predictions 159

6.1 Introduction . . . 159
6.1.1 Computational structural predictions . . . 159
6.1.2 Using smFRET results to help validate structural predictions . 160
6.2 Results . . . 163
6.2.1 inflexible protein FRET standards . . . 163
6.2.2 Investigating the central hinge region in the heavy chains of kinesin-1 . . . 175
6.3 Discussion . . . 181

7 Discussion 188

7.1 Summary . . . 188
7.2 smFRET assay using self labelling enzymes and cell lysate based samples . . . 191
7.3 Heavy chain kinesin-1 activation is more complex then previously thought . . . 192
7.4 Components of axonal transport can affect kinesin-1 autoinhibition . 194
7.5 Using smFRET to validate structure predictions . . . 194
7.6 Future work . . . 195

List of Figures

1.1 Structure of kinesin-1 23

1.2 Biophysical aspects of kinesin-1 movement 29

1.3 Kinesin-1 autoinhibition 34

1.4 Map of smFRET data points 42

2.1 Custom built photo bleaching set up 64

3.1 Absorption and emission spectrum of FRET fluorophores 78

3.2 Raw photon trace of an example smFRET experiment 81

3.3 Initial smFRET results for the CLIPf-SNAPf fusion protein 83

3.4 Example flow cytometry data for optimising washout steps 85

3.5 Optimisation of the CLIP and SNAP washout conditions 87

3.6 Optimisation of the CLIP and SNAP ligand incubation time 89

3.7 Optimisation of the CLIP and SNAP ligand concentration 91

3.8 Cross reactivity of SNAP and CLIP fluorescent ligands 93

3.9 Photon rates of isolated assay components 95

3.10 Background photon rate after optimisation 100

3.11 Optimised CLIPf-SNAPf fusion protein smFRET results 101

3.12 Biological repeats of the CLIPf-SNAPf fusion protein smFRET results 102

4.1 Using smFRET to study kinesin-1 autoinhibition 110

4.2 smFRET results for CLIPf-mKIF5A-SNAPf fusion protein 115

4.3 Biological repeats of the smFRET results for the CLIPf-mKIF5A-SNAPf fusion protein 116

4.4 Co-immunoprecipitation of kinesin-1 heavy chain constructs to check for heterodimers . . . 118
4.5 Stoichiometry of CLIPf-SNAPf and CLIPf-mKIF5A-SNAPf fusion proteins . . . 120
4.6 Possible stoichiometries for monomers and dimers due to differences in labelling efficiency . . . 121
4.7 Possible stoichiometries for co-transfection of CLIPf-mKIF5A and mKIF5A-SNAPf-HisTag-FLAG . . . 123
4.8 FRET efficiency results for single tagged mKIF5A constructs compared to the double tagged mKIF5A constructs . . . 124
4.9 FRET stoichiometry results for single tagged mKIF5A constructs compared to the double tagged mKIF5A constructs . . . 125
4.10 FRET efficiency results for addition of NaCl to the smFRET assay buffer . . . 127
4.11 FRET efficiency results for mutants of the CLIPf-mKIF5A-SNAPf fusion protein . . . 129
4.12 Alphafold2 prediction of mKIF5A (401-1027) . . . 132

5.1 smFRET results for CLIPf-hKIF5B-SNAPf fusion protein . . . 143
5.2 In gel fluorescence for the CLIPf-mKIF5A-SNAPf and the CLIPf-hKIF5B-SNAPf fusion proteins . . . 144
5.3 Confocal microscopy images of the LAMP1-CLIPf-SNAPf fusion protein 147
5.4 Overlap of LysoTracker with the LAMP1-CLIPf-SNAPf fusion protein signal . . . 148
5.5 Cell lysis by homogenisation . . . 150
5.6 Photon trace for the LAMP1-CLIPf-SNAPf and CLIPf-SNAPf fusion proteins . . . 152
5.7 smFRET results for CLIPf-mKIF5A-SNAPf fusion protein in the presence of microtubules . . . 154

6.1 TPR constructs for use as new FRET standards 165
6.2 Purification of bacterial expressed CLIPf-SNAPf fusion protein by StrepTag affinity 167
6.3 Purification of bacterial expressed CLIPf-SNAPf fusion protein by size exclusion chromatography 168
6.4 Purification of bacterial expressed CLIPf-TPR1-SNAPf fusion protein by StrepTag affinity 169
6.5 Purification of bacterial expressed CLIPf-TPR1-SNAPf fusion protein by size exclusion chromatography 170
6.6 Purification of bacterial expressed CLIPf-TPR3-SNAPf fusion protein by StrepTag affinity 171
6.7 Purification of bacterial expressed CLIPf-TPR3-SNAPf fusion protein by size exclusion chromatography 172
6.8 FRET efficiency results for engineered TPR proteins expressed and purified from *E. coli* 174
6.9 FRET efficiency results for engineered TPR proteins expressed from mammalian cells 176
6.10 mKIF5B central hinge constructs 178
6.11 Native PAGE and western blot of central hinge constructs of mKIF5B 180
6.12 FRET stoichiometry of CLIPf-mKIF5B(516-621)-SNAPf and CLIPf-mKIF5B(565-589)-SNAPf fusion proteins 182
6.13 FRET efficiency of CLIPf-mKIF5B(516-621)-SNAPf and CLIPf-mKIF5B(565-589)-SNAPf fusion proteins 183

List of Tables

2.1 Bacterial culture and cloning media 49

2.2 General PCR protocol 50

2.3 Primers used in PCR 51

2.4 Optimised ligand labelling conditions 58

2.5 Lysis buffers 59

2.6 Homogenisation buffers and solutions 61

2.7 smFRET assay buffers 63

2.8 Laser, filter and channel specifications for ligands 65

2.9 IP buffers and solutions 67

2.10 SDS-PAGE solutions 72

2.11 Native PAGE solutions 73

2.12 Transfer and western blotting solutions 73

2.13 Commercial primary antibodies 74

2.14 Buffers and solutions for polymerising microtubules 74

2.15 Protein purification solutions 74

3.1 Comparison of labelling conditions 104

List of Equations

1.1 FRET stoichiometry 41
1.2 FRET efficiency 41
3.1 Alpha FRET correction factor 96
3.2 Delta FRET correction factor 97
6.1 The distance between the two fluorophores 161
6.2 Föster radius (R_0) 161

Abbreviations

AA acceptor-only

ALEX alternating-laser excitation

ALS amyotrophic lateral sclerosisanoside

APD avalanche photodiodes

APS ammonium persulfate

ATP adenosine triphosphate

BSA bovine serum albumin

DA donor-acceptor

DD donor-only

DMEM Dulbecco's Modified Eagle Medium

DTT dithiothreitol

E FRET efficiency

EDTA ethylenediaminetetraacetic acid

EGTA ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid

FBS fetal bovine serum

FRET fluorescence resonance energy transfer

HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid

LAMP1 lysosomal-associated membrane protein 1

LB Luria-Bertani

PAE predicted aligned error

PBS phosphate-buffered saline

PIPES piperazine-N,N-bis(2-ethanesulfonic acid)

pLDDT predicted local-distance difference test

PLL poly-L-lysine

PMSF phenylmethylsulfonyl fluoride

S stoichiometry

SDS sodium dodecyl sulfate

SDS-PAGE sodium dodecyl-sulfate polyacrylamide gel electrophoresis

SiR silicon rhodamine

smFRET single molecule fluorescence resonance energy transfer

TAME tosyl-L-Arginine Methyl Ester

TB Terrific Broth

TBST tris-buffered saline with 0.1% tween

TEMED N,N,N',N'-tetramethyl ethylenediamine

TIRF total internal reflection fluorescence

TPR tetratricopeptide repeat

Chapter 1

Introduction

1.1 Microtubule based transport

Microtubule based transport is crucial for cell survival and function. Cellular cargo (for example proteins, vesicles and even organelles) are distributed throughout the cell to their necessary sites of function by microtubule based transport. This transport of cellular cargo becomes extremely important in large cells like neurons. Neurons are highly specialised cells that have to last your entire lifetime. In humans they can be around one metre in length with the cell body at one end and the synapse at the other. The axon is the part of the cell which separates the cell body from the synapse. Most material is synthesised in the cell body and then needs to be transported along the axon to its site of use. This is done by a specific type of microtubule based transport known as axonal transport where motor proteins, like cytoplasmic dynein and kinesin-1, deliver cargo along microtubules tracks within the axon.

Microtubules have a hollow cylindrical structure formed of thirteen protofilaments. These are long rod-like polymers made of α- and β-tubulin dimers (Downing and Nogales 1998). These tubulin dimers polymerise in a end-to-end fashion resulting in a overall polarised microtubule structure with a more stable minus-end and

a fast growing plus-end. Crucial for microtubule function is the dynamic instability of this structure resulting in the rapid switching between growth and depolarisation ("catastrophe") of the microtubule structure. The dynamics of the structure is controlled by the binding and hydrolysis of GTP by β-tubulin in the tubulin dimer (Alushin et al. 2014). Growth occurs at the plus-end of the microtubules where the α-tubulin of an incoming tubulin dimer interacts with the exposed β-tubulin at the end of the microtubule. This results in GTP hydrolysis at this β-tubulin and a buried GDP molecule. This GDP is unable to be exchanged for a molecule of GTP preventing another hydrolysis step and release the newly added tubulin dimer (Nogales et al. 1999). This creates a "GTP-cap" stabilising the microtubule and allowing growth. Rapid depolymerisaion occurs once this cap is lost resulting in catastrophe events. This is due to GTP exchange with the previously buried GDP molecules (Mitchison and Kirschner 1984).

Microtubules can be stabilised by microtubule associated proteins known collectively as MAPs. They usually bind to more than one tubulin dimer crosslinking these subunits and therefore stabilising the structure. An example of this type of MAP is the protein Tau which stabilises mictubules in the axon (E. Mandelkow and E. M. Mandelkow 1995). Another way of regulating microtubules is through post-translational modifications; PMTs. This gives rise to microtubule diversity allowing a wide array of functions for microtubules. Examples of PMTs are: acetylation, polyglycylation, polyglutamylation, tyrosination/detyrosination, phosphorylation and palmitoylation. These modifications could play a part in the control of axonal transport (Westermann and Weber 2003).

Polarisation of microtubules helps organise the direction of cellular transport as different motor proteins walk in different directions along microtubules. In axons all microtubules are positioned with the the minus-end directed towards the cell body giving uniform polarity along the axon (Baas et al. 1988). This leads to tight control of cargo movement along the axon. For example, cytoplasmic dynein is a minus-end

directed motor and therefore travels from the plus-end of the microtubule to the minus-end whereas kinesin-1 is a plus-end directed motor. This thesis focuses on kinesin-1.

1.2 The kinesin superfamily

The kinesin superfamily is composed of motor proteins with a conserved, globular, catalytic motor domain which can bind microtubules and hydrolyse ATP. The rest of the protein, usually the stalk and tail regions, are highly variable across the superfamily (Hirokawa and Noda 2008).

Kinesins can be grouped within the superfamily by the position of the motor domain. M-kinesins have the motor domain in the middle of the protein sequence whereas N- and C-kinesins have the motor at the N- and C-terminus respectively. Direction of travel along the microtubule varies for different members of the kinesin superfamily. This is necessary as kinesins have a large variety of roles within the cell; mainly involving intracellular transport and mitosis. Broadly N-kinesins are plus-end directed motors whereas C-kinesins walk towards the minus-end of microtubules. M-kinesins do not carry cargo along microtubule tracks but instead depolymerise the microtubule removing tubulin dimers from the ends of the filament (Hirokawa and Noda 2008).

In 2004 a standard nomenclature for the kinesin superfamily was set and defined 14 recognised kinesin families; kinesin-1 to kinesin-14. These families were based on the full length protein sequence alignment along with kinesin motor core alignment, and phylogenetic analyses. Any ungrouped kinesins are placed in a mixed "orphan kinesins" family (Lawrence et al. 2004). Of the 14 families, 12 are N-kinesins with only kinesin-13 being a M-kinesin and kinesin-14 a C-kinesin. This thesis focuses on the kinesin-1 family.

1.3 Kinesin-1

1.3.1 The role of kinesin-1

Kinesin-1 is responsible for transporting a vast array of different cellular cargo, (for example lipid bound organelles, proteins and parts of the cytoskeleton) to cargo specific destinations by walking along tracks of microtubules (Maday et al. 2014). It usually exits as a hetereotetramer made up of two heavy chains and two light chains (Kuznetsov, E. A. Vaisberg, et al. 1988). Both ends of the heavy chain, the motor and the tail domains, can bind to microtubules but only the N-terminal motor domain can hydrolyse ATP (Navone et al. 1992). This ATPase activity is stimulated by microtubules and results in kinesin-1 walking along the microtubule tracks (Hackney 1988).

Kinesin-1 was the first kinesin of the superfamily to be discovered. This happened in 1985 by two separate research groups. The new protein was found in both chick brain (Brady 1985) and squid axoplasm (Ronald D Vale et al. 1985), with both papers describing a novel ATPase and which binds to microtubules. Vale *et al.* named the protein kinesin from the Greek word kinein which means "to move" (Ronald D Vale et al. 1985 and Brady 1985). Throughout the rest of this decade a lot of research went into understanding the structure of this novel protein.

1.3.2 The structure of kinesin-1

Kinesin-1 usually exits as a heterotetramer made up of two heavy chains and two light chains (Kuznetsov, E. A. Vaisberg, et al. 1988). In humans there are three genes which encode for the heavy chain (KIF5A, KIF5B and KIF5C) and four for the light chain (KLC1, KLC2, KLC3 and KLC4). KIF5A and KIF5C are neuro-specific whereas KIF5B is ubiquitously expressed (Kanai et al. 2000). Soon after it's discovery, a fragment of the heavy chain purified from bovine brain of just 45 kDa was shown to be capable of binding microtubules and hydrolysing ATP, but didn't

have the ability to transport cargo (Kuznetsov, E. A. Vaisberg, et al. 1988). Electron microscopy showed a globular motor domain with the light chains and cargo binding to the other end of the heavy chain; around 80 nm apart (Hirokawa, Pfister, et al. 1989). These studies, and others, allowed the whole kinesin-1 molecule to be split into domains based on function. The overall structure of kinesin-1 is summarised in Figure 1.1. The kinesin-1 heavy chain molecule can be split into four main domains. These are: the motor (KIF5B 1-322), the neck linker (KIF5B 323-335), the stalk (KIF5B 336-914) and the tail domain (KIF5B 915-963).

The motor domain is the previously described globular region at the N-terminus of the heavy chain. It can bind to both microtubules and ATP (Friedman and R D Vale 1999). It then hydrolyses the ATP, generating force and a conformational change in the neck region (Tomishige et al. 2006). The mechanism behind kinesin-1 movement and how it was elucidated through single molecule experiments, will be discussed in detail later in this chapter.

The motor domain is the only part of the kinesin-1 heavy chain with detailed structural information. Over the years, many crystal structures have been solved for the motor domain of different species; human (Kull et al. 1996, Sindelar et al. 2002, Cao, W. Wang, et al. 2014, and Cao, Cantos-Fernandes, et al. 2017), rat (Sack et al. 1997) and *Drosophila* (Kaan et al. 2011). These showed high conservation of this domain across these species. At the moment all the human kinesin crystal structures are just from the KIF5B isoform but mutants and cysteine-light structures have been solved (Cao, W. Wang, et al. 2014 and Cao, Cantos-Fernandes, et al. 2017). The largest fragment for the human construct was KIF5B(1-349) (Cao, W. Wang, et al. 2014); the rat kinesin-1 fragment was slightly larger at 354 residues (Sack et al. 1997). Cryo-EM has also been used to study kinesin-1; for example Shang et al. 2014 and Weijman et al. 2022. High resolution structures were obtained for the monomeric human kinesin motor domain (KIF5B 1-349) attached to microtubules. The structure before and after ATP binding was captured revealing conformational

changes upon ATP binding (Shang et al. 2014). Resolution of the whole kinesin-1 heavy chain or molecule structure has been limited to 2D class average analysis of negative-stain EM (Weijman et al. 2022). This is probably due to the low electron density of the stalk and tail domain compared to the globular motor domains.

The stalk domain is where the two heavy chains dimerise together. This results in a coiled-coil α-helical structure believed to be rigid and rod-like (Amos 1987, Hirokawa, Pfister, et al. 1989, Yang et al. 1989 and Thormählen et al. 1998). The first evidence of a folded conformation of kinesin-1 came from electron microscopy in 1987. Amos *et al.* purified kinesin-1 from pig brain and observed a possible flexible joint near the middle of the rod structure (Amos 1987). This was confirmed in later electron microscopy studies (Hirokawa, Pfister, et al. 1989) and also by study of the DNA sequence of *Drosophila*, where a breakdown in the hydrophobic periodicity was seen as a potential hinge region (Yang et al. 1989). The hinge region was thought to be involved in autoinhibition of ATPase activity of the kinesin-1 molecule. This shown in 1999 by Friedman and Vale where the region was deleted (K963(Δ505-610)) and resulted in a increase in ATPase activity (Friedman and R D Vale 1999).

The tail domain was originally described as a fan-shaped C-terminus of the kinesin-1 heavy chain (Hirokawa, Pfister, et al. 1989) and is responsible for cargo and light chain binding. This cargo binding can be light chain independent (Skoufias et al. 1994). The tail domain has also been found to bind to microtubules but, unlike in the motor domain, this interaction is not influenced by ATP (Navone et al. 1992). Microtubule binding to both ends to the kinesin-1 heavy chain can result in cross links between separate microtubule polymers (Hirokawa, Pfister, et al. 1989). Both of the light chains interact with the tail domain to result in the heterotetrameric structure of the kinesin-1 molecule (Kuznetsov, E. A. Vaisberg, et al. 1988).

The light chains are smaller then the heavy chains and thought to be responsible for cargo binding. This occurs primarily through a six tetratricopeptide repeat (TPR) motif. These light chains can themselves be autoinhibited but it is not

known how this relates to the overall autoinhibition of kinesin-1 (Yip et al. 2016).

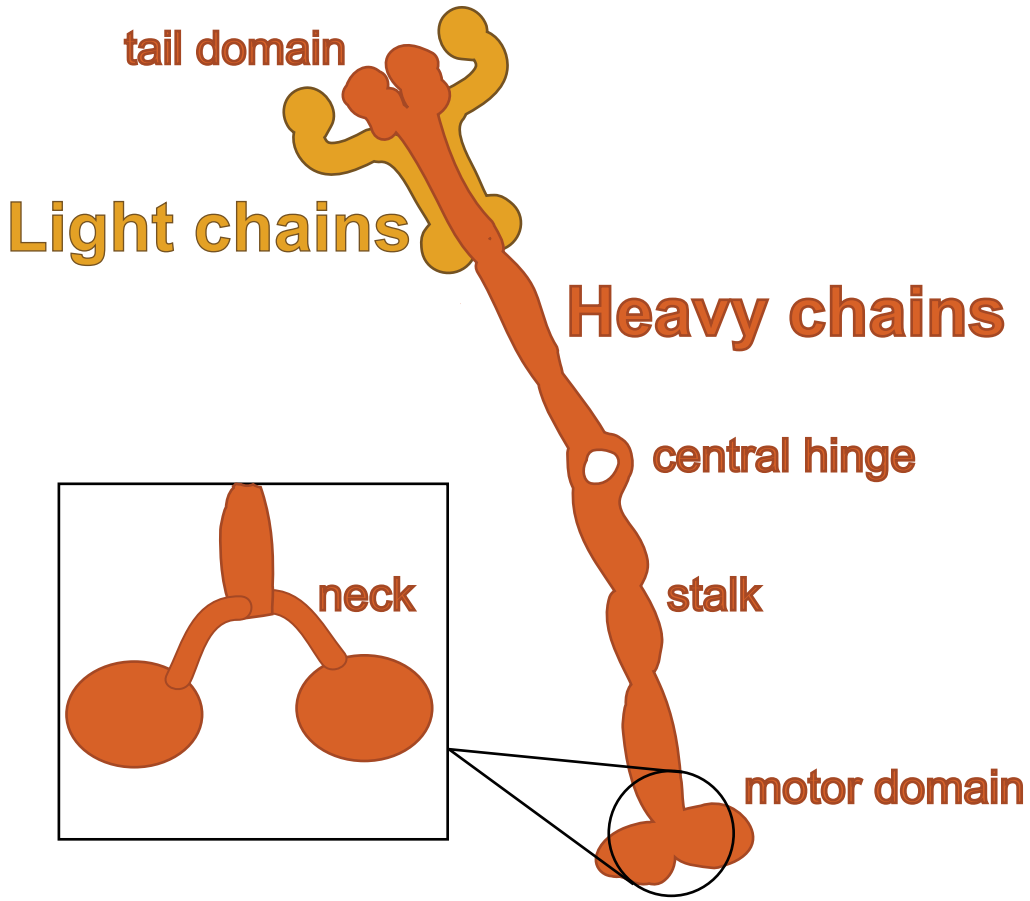

Figure 1.1: **Structure of kinesin-1**

Cartoon representation detailing the structure of kinesin-1. The heavy chains (KIF5A, B or C) are shown in dark orange whereas the light chains (1, 2, 3 or 4) are shown in light orange. All four domains of the heavy chain (motor, neck, stalk and tail) are shown as well as the central hinge region. Adapted from Twelvetrees 2020.

1.4 Single molecule experiments to study kinesin-1 movement

The detailed mechanism by which kinesin-1 travels along the microtubule has been elucidated through single molecule experiments. This type of experiment allows the movement and force generation of a single kinesin-1 molecule to be measured. Without single molecule experiments the intricacies of movement would not have been found due to ensemble averaging and lower resolution experiments. The first single molecule experiment conducted on kinesin-1 started in 1989 only four years after its discovery (Howard et al. 1989). Through extensive work the highly detailed mechanism for kinesin-1 movement has unfolded to such an extent that it is now used to test new biophysical setups (Wolff et al. 2022). It is now known that kinesin-1 movement: generates a stall force of around 5-6 pN, is highly processive, hydrolyses one ATP molecule per 8 nm step occurs in a hand-over-hand asymmetric fashion and through conformational change in the neck domain. All these points will be discussed in detail in this section.

In 1989 Howard *et al.* took the first single molecule approach to study kinesin-1. In their *in vitro* assay the tail end of the purified bovine brain kinesin was adsorbed onto a glass slide. Microtubules were then added and their sliding observed. Using progressively lower concentrations of kinesin-1 it was found that only one kinesin-1 molecule is needed to move a microtubule, independent of its length. The speed of microtubule sliding was dependant on the ATP concentration and not the number of kinesin-1 molecules (Howard et al. 1989).

The stall force of a single kinesin-1 molecule is around 5-6 pN (Svoboda and Block 1994). This was measured by optical trapping where single kinesin-1 molecules, purified from the optic lobe of squid, were attached to a silica bead. The complex was placed on a microtubule using optical tweezers and the bead position, within a stationary optical trap, measured to subnanometer accuracy. In this assay, the

displacement of the bead is proportional to the force generated. This technique allows individual kinesin-1 molecules to be investigated until the motor disengages, which was usually after a couple of minutes (Svoboda and Block 1994).

Although the high processivity of kinesin-1 was suggested in 1989 it was not confirmed until later (Howard et al. 1989). In 1989 single molecules of kinesin-1 were studied but the output of the experiment was microtubule sliding not kinesin-1 movement. The tail end of the kinesin-1 molecule was adsorbed onto a glass slide preventing any movement away from the microtubule (Howard et al. 1989). So the kinesin-1 molecule could have stopped moving the microtubule multiple times and then just rebound and restarted. Different single molecule experiments, using optical trapping and total internal reflection fluorescence (TIRF) microscopy (Svoboda, Schmidt, et al. 1993 and R D Vale, Funatsu, et al. 1996), have reported high processivity of around 150 steps per run. This only occurred with molecules containing two motor domains, with an average distance travelled of 600 nm. Molecules containing a single motor did not show processivity (R D Vale, Funatsu, et al. 1996).

The optical trap experiments showing processivity also allowed the step size of kinesin-1 to be determined as 8 nm; this distance corresponds to that between α-/β-tubulin dimers (Svoboda, Schmidt, et al. 1993). A few years later it was then confirmed, again using optical trapping, that only one ATP molecule is required for each step (Schnitzer and Block 1997). Optical trap experiments do not look at the movement of each head domain but instead, the overall movement of the bead attached to the protein. High resolution of these steps has been challenging. This is due to the labelling of the motor domain usually being much larger then the domain itself. For example, the use of gold beads which are 40 nm in diameter (Isojima et al. 2016) or silica beads with a diameter of 500 nm (Schnitzer and Block 1997). This has led to discrepancies in results. More recently, use of an interferometic MINIFLUX microscope, which has precision down to 1.7 nm within less than 1 ms, has allowed the study of site-specific cysteine labelled kinesin-1 motor

movement along a microtubule at different ATP concentrations. The small size of the fluorophore allowed 8 nm steps, and 4 nm substeps to be resolved with a median precision of 0.63 nm (Wolff et al. 2022).

Multiple studies have dissected the ATP hydrolysis cycle of kinesin-1. ATP hydrolysis is stimulated by the presence of microtubules due to the accelerated release of ADP. In the absence of microtubules, ATP binding and hydrolysis along with phosphate release are very rapid. The rate limiting step is the release of ADP but this is increased when microtubules are added (Hackney 1988). This allows the molecule to "step" along the microtubule track as in the weakly bound ADP state it can detach and rebind further along the track. Precise detail over aspects of this cycle are still being uncovered with new microscopy techniques giving better resolution and precision (Wolff et al. 2022).

There were two clear models of how kinesin-1 could move along the microtubule. These had to allow for the previously found parameters; one ATP molecule per 8 nm step, highly processive and with ADP release being the rate limiting step. The two models were the inchworm and the hand-over-hand. In the inchworm model only one motor domain is catalytically active. This results in there being a constant lead and a trailing motor domain throughout movement without the switching of roles. This results in no rotation of the coiled coil stalk domain (Hua et al. 2002). In contrast, the hand-over-hand model results in constant switching of the leading and trailing roles between the two motor domains. This model was suggested by the high processivity of the kinesin-1 molecule (R D Vale, Funatsu, et al. 1996). The hand-over-hand model can be either symmetrical or asymmetrical depending on the rotation of the stalk domain. A symmetric hand-over-hand model would mean the whole kinesin-1 molecule rotated with each step. In an asymmetric model there are different conformations at each step for the two motor domains (Asbury et al. 2003 and Yildiz et al. 2004).

It was determined, through optical force-clamp experiments, that kinesin-1 walks

in an asymmetric hand-over-hand fashion (Asbury et al. 2003). A truncated kinesin-1 molecule, only containing a short part of the stalk domain and the two motor domains (*Drosophila melanogaster* kinesin heavy chain 1-401), was seen to "limp" along the microtubule by alternating fast and slow periods between steps. This rules out symmetric models including the inchworm model (Asbury et al. 2003). The hand-over-hand model was further proved through fluorescent labelling of a single motor domain and seeing how its position changes with each step. If the inchworm model was true the dye would only move 8 nm for every step but, in the hand-over-hand model the dye would move 16 nm to the next available β-tubulin. Single molecule nanometer resolution was key to be able to differentiate between these two models. By tracking the position of the dye it was determined that kinesin-1 walked in an asymmetric hand-over-hand fashion (Yildiz et al. 2004).

The role of the neck linker was shown through by a number of experimental techniques, such as cryo-EM, electron paramagnetic resonance (EPR) and fluorescence resonance energy transfer (FRET). The reversible docking of the neck linker, from a loose undocked state on microtubule binding and ADP release to a docked rigid state on ATP binding, is essential for processive movement (S. Rice et al. 1999). The neck linker of kinesin-1 is defined as a 12-15 residue peptide which joins the motor to the stalk domain. Due to its size it was very hard to visualise any conformational change upon kinesin-1 movement. Also many studies used ensemble averaging or high ATP concentrations. Single molecule measurements were needed to investigate the conformation of the neck linker as two populations, depending on whether the motor is in the leading or trailing position, were likely. These problems were eventually overcome through the use of smFRET and dual labelling of only one of the neck linkers per molecule. Two different conformations for the neck linker were observed depending on whether it is attached to the leading or trailing head (Tomishige et al. 2006)

The biophysical aspects of kinesin-1 movement are summarise in Figure 1.2.

The detail now known around the mechanism of kinesin-1 movement would not be possible without single molecule techniques allowing nanometer precision.

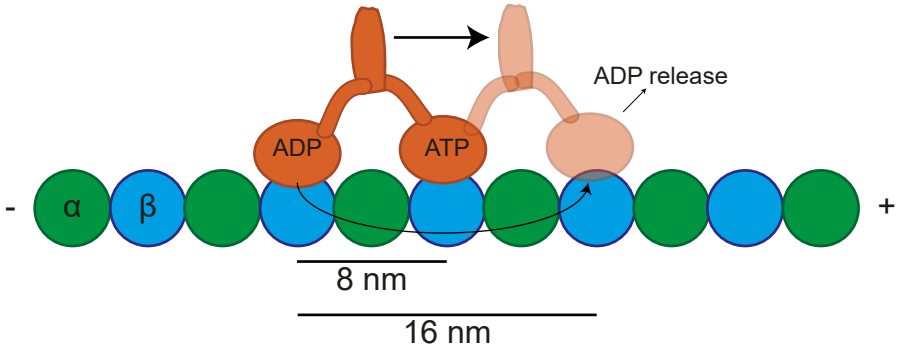

Figure 1.2: **Biophysical aspects of kinesin-1 movement**

The mechanism of kinesin-1 movement has been well characterised through the use of single molecule experiments. It: is highly processive, occurs through conformational change in the neck domain, occurs in an asymmetric hand-over-hand fashion, hydrolyses one ATP molecule per 8 nm step and generates a stall force of around 5-6 pN.

1.5 Regulation of kinesin-1

1.5.1 Adaptors

There are a group of proteins known as adaptors that are necessary for the transport of cargo. These can form bridges between the cargo and kinesin-1 or can alter the activation of the motor protein (Adio et al. 2006).

An example of a bridging adaptor of kinesin-1 is the protein miro. Miro associates with the outer membrane of the mitochondria and the kinesin-1 heavy chain forming a large complex which can transport mitochondria throughout the cell (Macaskill et al. 2009). Not only does this miro complex facilitate the transport of mitochondria but it also allows the regulation of its distribution favouring high energy demand places. This occurs through the interaction of miro to the complex which is highly calcium dependant. At high concentration of calcium causes detachment from the kinesin-1 heavy chain and release of the mitochondria (Macaskill et al. 2009).

A bridging adaptor crucial in the axonal transport of syntaxin-1 is syntabulin (Su et al. 2004). Syntaxin-1 plays a crucial role in the release of neurotransmitters and is needed at the presynaptic membrane. To reach this destination it is transported along the axon by kinesin-1. This distribution was found to be dependant on syntabulin as when it was disabled, through the use of siRNA, syntaxin-1 wasn't transported and instead clustered near the nucleus (Su et al. 2004).

Direct activation of kinesin-1 by adaptor binding and relieving the autoinhibition of the heavy chains has been shown (Twelvetrees, Lesept, et al. 2019 and F. Sun et al. 2011). For example, the protein Sunday Driver (also known as syd and JIP3) can bind directly to both the heavy chains and light chains of kinesin-1. Activation was shown by an increase in heavy chain velocity and run length in *in vitro* motility assays. This was determined to be through interaction with the heavy chains only as mutants that were unable to bind to the light chains resulted in a similar activation of transport compared to the wildtype Sunday Driver protein (F. Sun et al. 2011). An

increase in the velocity of kinesin-1 heavy chain movement was also seen on addition of GRIP1 and HAP1a. Although the rate in which the heavy chain landed on the microtubule was increased upon addition of either one of these adaptor proteins, both were needed to result in an increase of these landing events resulting in motility and for the increase in overall speed. It is thought that GRIP1 and HAP1a employ a co-operative activation mechanism possibly through stabilisation of the central hinge of the kinesin-1 heavy chain (Twelvetrees, Lesept, et al. 2019).

1.5.2 Autoinhibition and activation

A detailed mechanism of kinesin-1 autoinhibition, the prevention of ATPase activity, and activation, the unlocking of kinesin-1 autoinhibition so it can bind to cargo/microtubules and hydrolyse ATP, is still unknown. This lack of information, especially compared to the vast information known now about kinesin-1 movement, is probably due to the lack of single molecule experiments that directly study kinesin-1 conformation. The current knowledge around kinesin-1 autoinhibition will be summarised in this section.

Kinesin-1 autoinhibition is vital within cells to prevent unnecessary use of energy. Without some control mechanism kinesin-1 would be constantly walking along the microtubule tracks even when no cargo was attached. This would result in a lack of available kinesin-1 molecules when cargo needed to be transported. Constant walking would also require constant ATP hydrolysis, wasting energy moving the kinesin-1 down the microtubule track unnecessarily.

The presence of a hinge region in the structure of kinesin-1 indicated the possibility of an autoinhibition mechanism before any biochemical confirmation (Amos 1987 and Hirokawa, Pfister, et al. 1989). This region was first observed in electron microscopy studies of kinesin-1. Purified kinesin-1 could be seen adopting multiple different conformations. Molecules consisted of a globular domain at one end and a fan shaped domain at the other. These two domains were mainly joined by

a straight rod-like structure. In some molecules, this rod-like structure was bent. This was the first indication of a central hinge region which could be involved in an autoinhibitory mechanism (Amos 1987 and Hirokawa, Pfister, et al. 1989).

The next indication of an autoinhibitory mechanism was the difference in ATPase activity between different kinesin-1 mutants. The microtubule-stimulated ATPase activity of the heavy chain homodimer is lower when the light chains are also present (Hackney, Levitt, and Wagner 1991) but, the activity is greatest when just motor domains (a 45 kDA fragment) are present (Kuznetsov, Y. A. Vaisberg, et al. 1989). These two findings suggest autoinhibition of just the heavy chains and with the light chains adding to this inhibitory effect.

Evidence of a conformational change was seen by differences in sedimentation coefficients for kinesin-1, purified from bovine brain, and the full length *Drosophila* heavy chain (Hackney, Levitt, and Suhan 1992 and Stock et al. 1999). For kinesin-1, two conformations were observed; one at 9.4 S and one at 6.5 S. This shift in sedimentation coefficient could be induced through addition of salt. Low ionic conditions ($\sim$0.2 M NaCl) resulted in the lower coefficient (6.5 S, indicating the folded conformation) and higher ionic conditions ($\sim$1 M NaCl) the higher coefficient; 9.4 S corresponding to the extended conformation. This conformational change occurred in the presence of the light chains and in their absence, suggesting it is the heavy chains undergoing the change. The extended conformation is thought to be active, with the motor and tail domain free to bind to microtubules and cargo respectively, and the global conformational change resulting in an inactive folded conformation (Hackney, Levitt, and Suhan 1992 and Stock et al. 1999).

Global and local conformational changes were also seen in fluorescence resonance energy transfer (FRET) measurements in COS cells. Fluorescent proteins, monmeric Citrine (mCit) and monomeric ECFP (mECFP), were fused to the N- and C-termini of the kinesin-1 heavy chain (rat KIF5B) respectively. Transfected cells were transiently permeabilised before addition of high ionic strength buffer.

An overall shift in conformation, from folded to unfolded was observed through a decrease in FRET efficiency measured across the whole cell. As well as this global conformational change a local conformational change was also seen in the motor domains. When the kinesin-1 light chains were also present the two motor domains were pushed apart in a closed inactive conformation (Cai et al. 2007).

The QIAK motif in the tail of the kinesin-1 heavy chain has been shown to interact with the motor domain, in the folded inactive conformation (Kaan et al. 2011). This motif is highly conserved and in *Drosophila* was shown to be essential for autoinhibition of the motor's ATP activity (Hackney and Stock 2000). A "double lockdown" mechanism of autoinhibition was later suggested by comparison of the crystal structure of the motor domain dimer with and without one bound tail domain (Kaan et al. 2011). The tail domain consisted of a small peptide (937-952) containing the QIAK region. The crystal structure allowed theories such as tail domain inference with either the microtubule binding site or the nucleotide binding site, to be ruled out. Instead cross-linking occurs at two positions limiting the movement of the motor domains and potentially preventing the undocking of the neck linker and ADP release needed for kinesin-1 movement (Kaan et al. 2011).

Currently kinesin-1 activation is modelled by a simple binary on/off model (see Figure 1.3), but little detail of the mechanism involved is known. In order to activate kinesin-1, and unlock autoinhibition, many interactions need to be overcome resulting in local and global conformational changes within the molecule. At the moment, the minimum requirements for this process and the mechanism behind it is unknown.

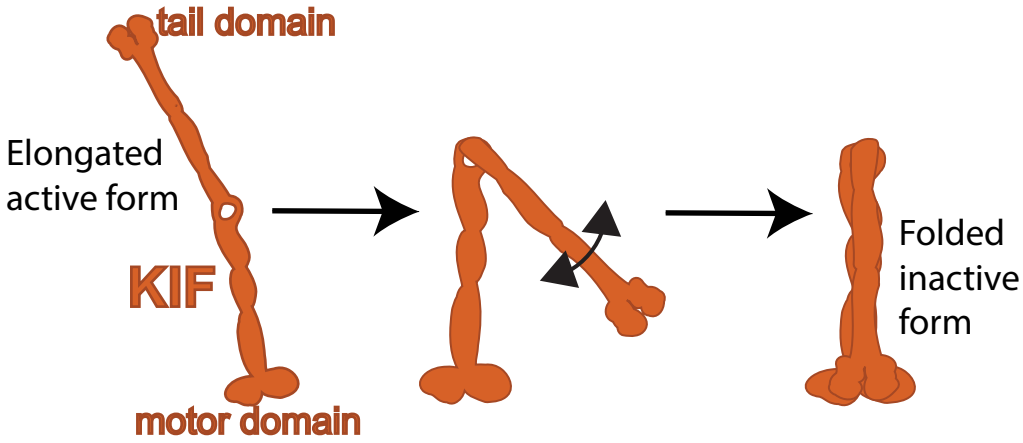

Figure 1.3: **Kinesin-1 autoinhibition**

Kinesin-1 is known to be autoinhibited resulting in two known conformations; an elongated active form and a folded inactive form. There is a direct interaction between the motor domains an the inhibitory motif in one of the tail domains.

1.6 A single molecule approach to directly study kinesin-1 conformation

The activation of kinesin-1 has been studied since autoinhibition was first proposed in 1992 by Hackney *et al.*, but there is still no single molecule experiment where the direct readout is kinesin-1 conformation (Hackney, Levitt, and Suhan 1992). Instead conformation and activation are inferred from the other experimental results. For example, *in vitro* reconstitution total internal reflection fluorescence (TIRF) walking assays are a common technique to study kinesin-1. Flow chamber are created on clean coverslips and coated with α-tubulin antibodies. The rest of the surface is blocked to prevent protein aggregation on the surface. Fluorescent microtubules are then flowed through the chamber and bind to the antibodies. Fluorescently tagged kinesin-1, either purified samples or overexpressed cell lysate samples, are then be added to the chamber. Single molecules of kinesin-1 walking along the microtubule track can then be tracked and monitored. Differences in velocity, run length and processivity of different constructs are compared. Constructs with a decreased velocity are characterised as less active. However, when kinesin-1 is walking along the microtubule, activation has already occurred so changes in activation due to specific conditions cannot be observed. By studying the change in the conformation of kinesin-1 instead, a more detailed and accurate view of kinesin-1 activation could be determined. Also, currently all single molecule approaches to study kinesin-1 require the presence of microtubules. It is not known if activation occurs in response to microtubule or cargo binding, so having microtubules always present prevents this question being resolved.

Whole cell fluorescence resonance energy transfer (FRET) has been used to directly study kinesin-1 conformation. These experiments looked at the conformational changes upon activation averaged over the entire kinesin-1 population within the cell. For the first time a local conformation change in the motor domains was

suggested along with the global conformation change; folded to unfolded. This showed the potential benefits of a having kinesin-1 conformation as the direct experimental readout (Cai et al. 2007). An *in vitro* single molecule approach would allow the affect of the individual components of transport to be observed, along with a more detailed look at the individual characteristics of the kinesin-1 populations. Single molecule experiments allow a heterogeneous mixture of protein conformations to be resolved into multiple smaller populations which is not possible with ensemble measurements.

1.7 Fluorescence resonance energy transfer

Fluorescence resonance energy transfer (FRET) is the highly distance dependent energy transfer from one donor fluorophore to one acceptor fluorophore (Y. Sun et al. 2011). This results in acceptor emission following donor excitation. The non-radiative energy transfer, from the excited state of the donor fluorophore to the ground state of the acceptor fluorophore, can only occur when the two fluorophores are within 3-10 nm due to the efficiency of the energy transfer process (Ingargiola et al. 2016). This efficiency of energy transfer between the two fluorophores can the quantified and used to calculate absolute distances (Hellenkamp et al. 2018). This means that not only does FRET allow resolution of distances below the diffraction limit of light (around 200 nm), but also below the resolution of super resolution microscopy (around 20 nm). This energy transfer is also dependant on the spectral overlap between the emission of the donor and the absorption of the acceptor and the orientation of the two fluorophores (Y. Sun et al. 2011).

The addition of alternating-laser excitation (ALEX) to confocal FRET has increased the accuracy of FRET efficiency measurements and therefore calculated distances (Lee et al. 2005). This is done by applying fluorophore specific correction factors which take into account background, detection efficiency and cross-talk be-

tween the two fluorophores. When ALEX is used, the lasers that excite the donor and acceptor fluorophore are switched on and off in an alternating fashion allowing the probing of both fluorophores instead of just the donor fluorophore. This results in a new dimension, stoichiometry, which allows the sorting of the molecules into donor-only, acceptor-only and double labelled populations. ALEX also removes instrumental factors meaning that measurements taken on different instruments can be compared more easily (Lee et al. 2005). This was shown very effectively in single molecule FRET measurements by Hellenkamp *et al.* where FRET efficiency measurements from 20 labs were compared and shown to be highly precise and accurate with standard deviations between ± 0.02 and ± 0.05 (Hellenkamp et al. 2018).

1.7.1 FRET can be used to study conformational changes in molecules

Changes in FRET efficiency can be used to investigate conformational changes in biomolecules like proteins, where these changes are crucial for their function. This can be done by determining the FRET efficiency of the molecule in different conditions, as either ensemble or single molecule measurements.

FRET has previously been used to study conformational changes that occur in MCAK; a member of the kinesin superfamily also known as kinesin-13 (Ems-McClung et al. 2013). MCAK conformation is controlled by phosphorylation of the molecule. Unlike kinesin-1, the active state of MCAK is the closed conformation and switching to an open conformation prevents microtubule binding and MCAK's microtubule depolymerisation activity. Ems-McClung *et al.* added mCitrine and mCerulean to either end of full-length MCAK and to MCAK deletion mutants. The fluorescently labeled full length MCAK construct resulted in a higher FRET efficiency indicating that in solution MCAK exists in a closed conformation. This FRET efficiency could then be reduced, indication a conformational change to the open conformation, on addition of Aurora B, a known molecule that phosphorylates

MCAK and inhibits its activity. They also combined FRET with fluorescence lifetime imaging (FLIM) to discover the conformation of MCAK on different parts of the microtubule lattice. A reduction in the fluorescent lifetime indicated FRET was occurring. Using this method, they showed that MCAK bound to the microtubule ends are in a more closed conformation then the MCAK bound to the microtubule lattice. This is consistent with MCAK being active at the microtubule ends resulting in depolymerisation (Ems-McClung et al. 2013).

Single molecule fluorescence resonance energy transfer (smFRET) experiments provide insight on the conformational dynamics of the molecule that ensemble measurements can't, due to the inhomogeneous nature of the dynamics within the structure. It also allows transient intermediates to be uncovered (He et al. 2015 and Hellenkamp et al. 2018).

smFRET has been used to study conformational changes in enzymes and understand how regulators change the conformation distribution, both between different states and the lifetime of these states. An example of this was done in 2015 by He *et al.*. They investigated nitric oxide synthase catalysis and how this was changed by the known regulator, calmodulin. Using smFRET, they found that calmodulin reduced the mean FRET efficiency. The distribution of FRET efficiencies was also tighter when it was present. They also looked at the time spent in these different conformational states by suspending single molecules in 1 % agarose and seeing how the FRET signal changed over time. From this they found that calmodulin allows faster fluctuations between conformational states. Both these effects had the potential to increase the rate of the nitric oxide synthase enzymes (He et al. 2015).

1.7.2 FRET standards have been used to ensure reproducibility and accuracy of smFRET efficiency measurements

In 2018 twenty labs from around the world took part in a comparative blind study, using double labelled DNA duplexes, in order to establish a generalised protocol and

data analysis pipeline for single molecule FRET (Hellenkamp et al. 2018). Up until this point FRET efficiencies had been mainly used to study relative changes between different conformations of biomolecules. A lack of the reproduciblity, and known inaccuracy of the data, stopped the wide spread use of distance calculations from FRET efficiency measurements being applied to help determine structural information. There was a wide variety in data acquisition and analysis methods, usually using different custom-built set ups, making comparing results and validating distance measurements extremely hard (Hellenkamp et al. 2018). This is an extremely useful potential application of smFRET, as other structural determination methods lack temporal resolution meaning information about the heterogeneity of the population is lost (Agam et al. 2022). As well as the generalised protocol and data analysis, a set of standard DNA samples were made to compare and determine the precision and accuracy of the FRET efficiencies from different custom set ups. These benchmark samples were made to cover the whole FRET efficiency range (one for low, mid and high FRET efficiencies respectively). Across the labs in the study, a high consistency was seen with a standard deviation of less then 0.05 for the FRET efficiency measurements (Hellenkamp et al. 2018).

Recently a pre-print manuscript has shown similar reliability and accuracy in protein FRET efficiency measurements (Agam et al. 2022). This study was similar to that done by Hellenkamp *et al.* in 2018, as 19 labs from around the world took part in a blind study in order to determine the uncertainty of the FRET efficiency measurements. Instead of a set of standard double labelled samples where the position of the dyes changed for each sample, resulting in a ruler across the FRET efficiency range, two protein systems which undergo known conformational changes were used. Again a very low uncertainty between different lab measurements was observed with a standard deviation of 0.06. The two proteins in this study were purified and then labelled by site specific labelling of accessible cysteine-light residues (Agam et al. 2022). This labelling technique is widely used to study proteins

but requires the production of cysteine variants of the protein of interest. Although these are not hard to produce by site-directed mutagenesis it does mean the wildtype protein sequence, and potentially the structure, is not being directly studied. As it requires protein purification it also prevents the study of the proteins in their native environment.

1.7.3 A single molecule fluorescence resonance energy transfer (smFRET) custom built confocal instrument

The smFRET instrument, used for the studies in this thesis, is a custom built confocal smFRET platform (Ambrose et al. 2020). It has two lasers; an 80 mW 515 nm laser to excite the donor fluorophore and a 100 mW 613 nm laser to excite the acceptor fluorophore. These are used in an ALEX laser duty cycle where the donor laser is on for 45 µs, then a 5 µs break where all lasers are off before the acceptor laser is switched on for 45 µs. Then all lasers are switched for a further 5 µs. This cycle is repeated over the duration of the experiment. Any fluorescence is then detected by avalanche photodiodes (APD). The instrument contains two APDs; again one for the donor fluorophore and one for the acceptor fluorophore. Separate filters in front of the APD control the wavelengths of emitted light detected. The filter in front of the the donor APD only lets through light between 535 to 607 nm whereas the filter in front of the second APD, for the acceptor, removes wavelength other than that between 660 to 700 nm. Any photon detected generates a timestamp and records the corresponding detector. The laser that resulted in the emission is then inferred from the laser duty cycle. This allows the photons to be split into three distinct categories; donor-only (DD, photons as a result of the donor laser being on and donor emission being observed), acceptor-only (AA, where acceptor laser excitation lead to detection of acceptor emission) and donor-acceptor (DA, where excitation from donor laser resulted in acceptor emission). This last group occur due to FRET taking place so are the ones used for further analysis.

After photon detection and categorisation, photons are grouped into bursts. A burst is defined as many photons in a small amount of time. For each burst a stoichiometry value (S) and a FRET efficiency value (E) can be calculated from the number of photons in each of the three channels; donor-only (DD), acceptor-only (AA) and donor-acceptor (DA) (see Equation 1.1 and 1.2). Stoichiometry (S) is the ratio of donor to acceptor labelling. So if S = 0.5 the labelling is equal between the two fluorophores. S >0.5 shows more donor labelling and S <0.5 for greater acceptor labelling. FRET efficiency (E) indicates the efficiency of the energy transfer from the donor fluorophore to the acceptor. This is indirectly proportional to the distance between the fluorophores. For example if E is between 0.8-1.0 this is a high FRET signal and the two fluorophores are extremely close together (around 3 - 4 nm). As E decreases the distance between the fluorophore increases until E = 0 and there is more than 10 nm between the fluorophores meaning FRET can not occur. This is all summarised in the map of smFRET data points showing the different possible photon populations (see Figure 1.4). Samples are diluted to around fM concentration to achieve single molecule resolution; around one burst per second within the confocal volume. These molecules then diffuse across the confocal volume, around a fl, throughout the experiment.

$$S = \frac{DD + DA}{DD + DA + AA} \tag{1.1}$$

$$E = \frac{DA}{DD + DA} \tag{1.2}$$

smFRET offers distinct advantages over other single molecules assays currently used to study kinesin-1 activation. Firstly, molecules are freely diffusing and not immobilised onto a surface. This allows conformational freedom of all molecules. Also smFRET requires very small amounts of sample meaning that labelled protein

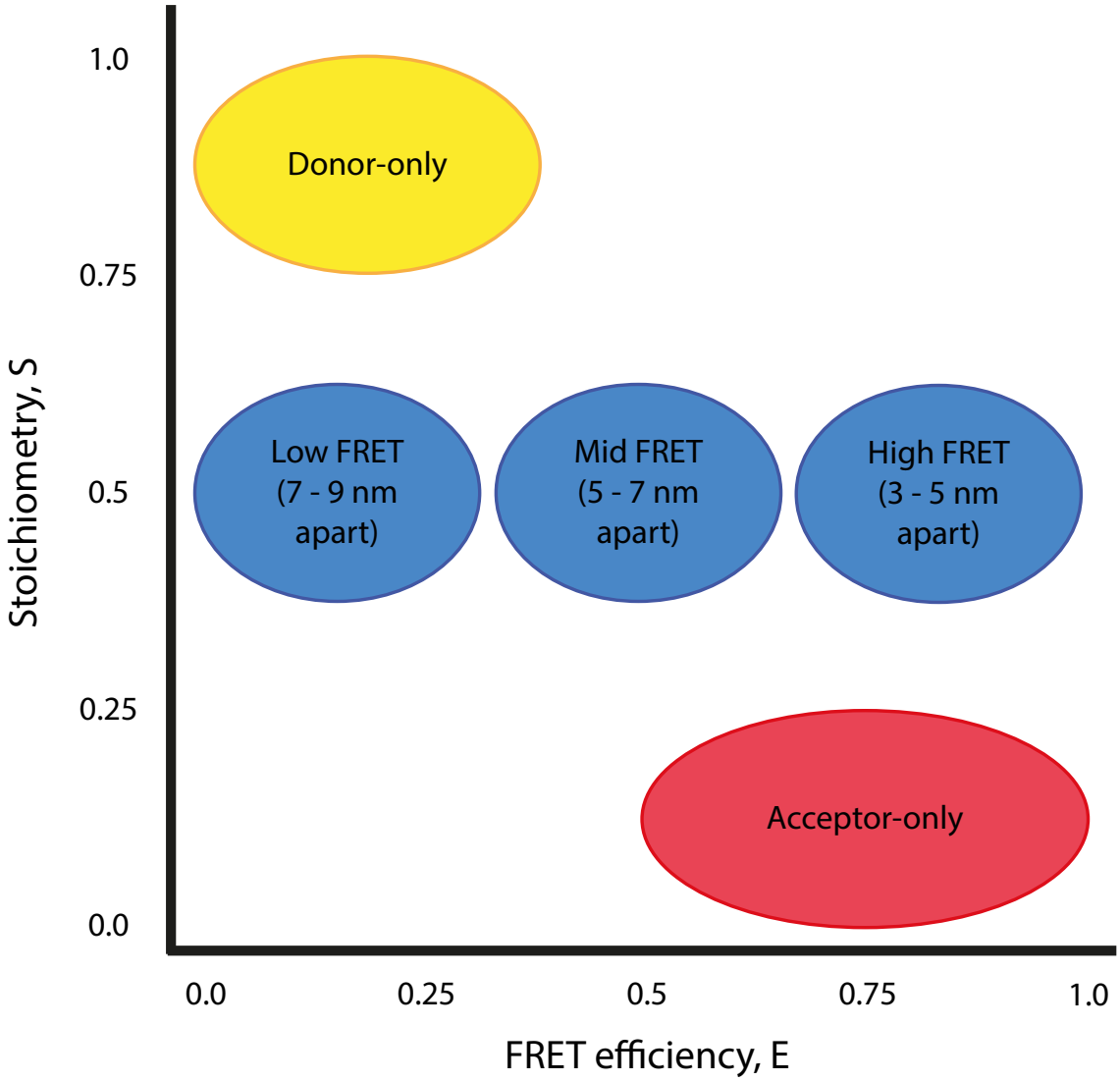

Figure 1.4: **Map of smFRET data points plotting FRET efficiency, E, against stoichiometry, S**

Each burst of photons has a value for E, how efficient the energy transfer between the two fluorophores was, and S, the ratio of photons detected between the two fluorophores. A high FRET efficiency indicates the energy transfer was very efficient and therefore the two fluorophores are very close together. As the FRET efficiency decreases the distance between the two fluorophores increases until they are more then 10 nm apart and E = 0. If S = 0.5 then there is equal labelling between the two fluorophores in that burst. A higher S indicates more donor labelling whereas a lower S shows more acceptor labelling is present.

samples might not need to be purified. Instead cell lysate samples might be able to be used meaning that labelled mammalian protein can be made by mammalian cells and kept in a more physiological environment. Additionally different components of axonal transport can be added in separately, meaning that kinesin-1 conformation can be studied in the absence of microtubules.

1.8 Thesis aims

Single molecule experiments have allowed the intricacies of kinesin-1 movement to be uncovered. This same level of detail into kinesin-1 autoinhibition and the underlying mechanism behind it is still needed. To help gain this information a single molecule assay, where the direct readout is the conformational state of kinesin-1, would be beneficial. smFRET seems a likely technique on which to base this assay. Therefore the aims of this thesis are as follows:

- Validate the use of an smFRET assay based on the chosen labelling strategy.
- Optimise this smFRET assay to establish robust and highly reproducible results.
- Study kinesin-1 conformation using mutations and conditions known to affect the activation or conformation of kinesin-1.
- Investigate the individual effect of other components of transport upon kinesin-1 conformation.
- Investigate the potential of using this smFRET assay to gain structural information about kinesin-1.

Chapter 2

Methods

2.1 Molecular Biology

2.1.1 Constructs

Primers for PCR or deletion mutagenesis are all listed in Table 2.3. The whole insert for each new construct was sequence verified.

2.1.1.1 CLIPf-SNAPf fusion protein constructs

CLIPf-SNAPf-HisTag-FLAG

The CLIPf-SNAPf-HisTag-FLAG construct was designed by Dr. Twelvetrees and synthesised by GeneArt (Invitrogen by Thermo Fisher Scientific). After the CMV and T7 promoter there is a CLIPf-tag and then a SNAPf-tag. These two tags are separated by a small linker of 12 amino acids (GSAGSAAGSGEF, Waldo et al. 1999). Between the CLIPf-tag and the linker is a BamHI restriction site. An EcoRI restriction site exists between the linker and the SNAPf-tag to enable easy incorporation of inserts between the CLIPf-tag and the SNAPf-tag. After the SNAPf-tag there is a polyhistidine tag followed by a FLAG tag. These were included for protein purification and immuno assays. The resultant construct is CLIPf-SNAPf-HisTag-FLAG.

CLIPf-HisTag-FLAG and SNAP-HisTag-FLAG

Single tag versions of the CLIPf-SNAPf-HisTag-FLAG, where only the CLIPf-tag or the SNAPf-tag is present, were made by deletion mutagenesis (see Section 2.1.2.2) of the construct. This resulted in two constructs: the CLIPf-HisTag-FLAG construct and the SNAPf-HisTag-FLAG construct.

2.1.1.2 Kinesin-1 heavy chain constructs

Mouse KIF5A (mKIF5A) constructs

The CLIPf-mKIF5A-SNAPf-HisTag-FLAG construct was made by Dominic Bingham. Full length mouse KIF5A(1-1027) was amplified (see Section 2.1.2.2) from the pRK5-Myc-KIF5A construct (Twelvetrees, Yuen, et al. 2010) to include complementary overlaps with the vector. After restriction digest of the CLIPf-SNAPf-HisTag-FLAG vector (see Section 2.1.2.4), the amplified mKIF5A fragment was inserted by isothermal assembly (see Section 2.1.2.6). The resultant construct is CLIPf-mKIF5A-SNAPf-HisTag-FLAG.

Single tag versions of the CLIPf-mKIF5A-SNAPf-HisTag-FLAG construct were made by deletion mutagenesis (see Section 2.1.2.2) by Annabelle Savage. This resulted in two constructs: the CLIPf-mKIF5A construct and the mKIF5A-SNAPf-HisTag-FLAG construct.

A C-terminal truncation mutant of mKIF5A was made by deletion mutagenesis (see Section 2.1.2.2) of the CLIPf-mKIF5A-SNAPf-HisTag-FLAG construct by Rebecca Mighell. This resulted in the CLIPf-mKIF5A(Δ954-1027)-SNAPf-HisTag-FLAG construct, making mKIF5A equivalent in length, based on homology, as mKIF5B and mKIF5C.

A double deletion mutant of mKIF5A (central hinge region and tail, including QIAK, region deleted) was made by sequential deletion mutagenesis (see Section 2.1.2.2) by Rebecca Mighell. First the central hinge region was deleted from the CLIPf-mKIF5A-SNAPf-HisTag-FLAG construct. This resulted in the CLIPf-

mKIF5A(Δ505-610)-SNAPf-HisTag-FLAG construct. Then the tail region, including QIAK domain, was deleted from this construct. This resulted in the CLIPf-mKIF5A(Δ505-610 and Δ915-1027)-SNAPf-HisTag-FLAG construct.

Mouse KIF5B (mKIF5B) constructs

The KIF5B(mouse) hinge fragments were chosen using paircoil2 (discussed in detail in Section 6.2.2). These four fragments were bought from GeneArt with a BamHI restriction site at the N-terminus of the hinge fragment and a EcoRI restriction site at the C-terminus. After restriction digest of the hinge fragments and the CLIPf-SNAPf-HisTag-FLAG vector (see Section 2.1.2.4), the hinge fragments were inserted by ligation (see Section 2.1.2.7). This subcloning was done by Rebecca Mighell. This resulted in the following four constructs: CLIPf-mKIF5B(516-621)-SNAPf-HisTag-FLAG, CLIPf-mKIF5B(529-607)-SNAPf-HisTag-FLAG, CLIPf-mKIF5B(545-593)-SNAPf-HisTag-FLAG and CLIPf-mKIF5B(565-589)-SNAPf-HisTag-FLAG.

Human KIF5B (hKIF5B) constructs

The CLIPf-hKIF5B-SNAPf-HisTag-FLAG construct was made by Annabelle Savage. KIF5B(human) was amplified (see Section 2.1.2.2) from the pXPE vector which was a gift from Ron Vale (Addgene plasmid 15284 ; http://n2t.net/addgene:15284 ; RRID:Addgene15284). It was amplified to include an N-terminal BamHI restriction site and a C-terminal EcoRI restriction site. After restriction digest of the PCR insert and the CLIPf-SNAPf-HisTag-FLAG vector (see Section 2.1.2.4), the amplified hKIF5B fragment was inserted by ligation (see Section 2.1.2.7). The resultant construct is CLIPf-hKIF5B-SNAPf-HisTag-FLAG.

2.1.1.3 LAMP1 construct

Lysosomal-associated membrane protein 1 (LAMP1) was amplified from Lamp1-mScarlet-I, a gift from Dorus Gadella (Addgene plasmid 98827 ; http://n2t.net/addgene:98827 ; RRID:Addgene98827), to include complementary overlaps with the vector. KpnI was used to digest the CLIPf-SNAPf-HisTag-FLAG construct before

inserting the lamp1 PCR fragment by isothermal assembly.The resultant construct is LAMP1-CLIPf-SNAPf-HisTag-FLAG.

2.1.1.4 TPR constructs

TPR constructs codon optimised for bacterial expression

The TPR constructs codon optimised for bacterial expression were designed by Dr. Twelvetrees and myself. The design of the TPR section is described in detail in Section 6.2.1.1. The FLAG-CLIPf-3xTPR-SNAPf-Strep construct was synthesised by GeneArt (Invitrogen by Thermo Fisher Scientific). Deletion mutagenesis (see Section 2.1.2.2) of this construct was performed to produce: FLAG-CLIPf-2xTPR-SNAPf-Strep, FLAG-CLIPf-1xTPR-SNAPf-Strep, FLAG-CLIPf-SNAPf-Strep. This final construct has no linker between the CLIPf-tag and the SNAPf-tag.

TPR constructs codon optimised for mammalian expression

The TPR constructs codon optimised for mammalian expression were designed by Dr. Twelvetrees and myself. The design of the TPR section is described in detail in Section 6.2.1.1. They were synthesised by GeneArt (Invitrogen by Thermo Fisher Scientific). The resultant constructs are: FLAG-CLIPf-3xTPR-SNAPf-Strep, FLAG-CLIPf-2xTPR-SNAPf-Strep, FLAG-CLIPf-1xTPR-SNAPf-Strep, FLAG-CLIPf-SNAPf-Strep. This final construct, with no TPR sections present, contained a small linker between the CLIPf- and SNAPf-tags (GSAGSAAGSGEF, Waldo et al. 1999).

2.1.2 Bacterial culture and cloning

2.1.2.1 Growth media and agar plates

All plasmids were transformed and maintained using *E. coli* DH5-αbacteria apart from the human KIF5B plasmids. These were maintained and transformed using *E. coli* Stbl3 bacteria. All *E. coli* strains were grown in Luria-Bertani (LB) (see Table

2.1) at 37 °C. Ampicillin resistant plasmids were grown in 100 µg/ml Ampicillin whereas Kanamycin resistant plasmids were grown in 50 µg/ml Kanamycin.

E. coli BL21(DE3)pLyS bacteria were used for protein expression. They were grown in Terrific Broth (TB) (see Table 2.1) at 37 °C, in 50 µg/ml Ampicillin and 35 µg/ml Chloramphenicol.

	Component	**Final concentration**
LB media	Tryptone Yeast extract NaCl	1 % w/v 0.5 % w/v 171 mM
LB agar	LB media Agar	- 2 % w/v
SOC media	Tryptone Yeast extract NaCl KCl $MgCl_2$ $MgSO_4$ Glucose	2 % w/v 0.5 % w/v 10 mM 2.5 mM 10 mM 10 mM 20 mM
TB media	Tryptone Yeast extract Glycerol KH_2PO_2 K_2HPO_2	12 g/L 24 g/L 0.4 % 17 mM 72 mM
Lithium borate buffer	Lithium acetate duhydrate Boric acid	10mM 10 mM
2x PBS+DTT solution	PBS DTT	2x 1 mM

Table 2.1: Bacterial culture and cloning media

2.1.2.2 PCR and deletion mutagenesis

PCR was carried out using Q5 High-Fidelity DNA Polymerase (New England BioLabs) following the manufacturer's instructions. 1 ng of template DNA was incubated with 12.5 µl Q5 High-Fidelity 2X Master Mix and 0.5 µM of both the forward and the reverse primer. The final volume of 25 µl was made up with nuclease free water. A PCR programme was carried out using the protocol in Table 2.2 where Ta stands for the annealing temperature of the specific primer pair used (see Table 2.3). If the annealing temperature was greater than or equal to 72 °C then a two step PCR protocol without the annealing temperature step in the cycle was used. The PCR insert was usually then purified (see Section 2.1.2.3) or inserted into the desired vector by isothermal assembly (see Section 2.1.2.6).

Step	**Temperature (°C)**	**Time (seconds)**
Initial denaturation	98	30
30 cycles	98 Ta 72	10 30 30 per kb
Final extension	72	120

Table 2.2: General PCR protocol

Deletion mutagenesis was performed by PCR, as stated above, with a few minor changes to the PCR protocol. Firstly, both denaturation steps were conducted at 95 °C and secondly the final extension step was five minutes. The PCR products were then mixed with 1 µl of 20,000 units/ml Dpn1 restriction enzyme (New England BioLabs) to digest any template vector. This mixture was left to incubate at 37 °C for 15 minutes before transformation (see Section 2.1.2.8).

Construct	**Template**	**Forward primer (5' to 3')**	**Reverse primer (5' to 3')**	Annealing temperature (Ta/°C)
CLIPf-HisTag-FLAG	CLIPf-SNAPf-HisTag-FLAG	GGAAGCGCCGGATCTGCA	GAATTCGCCGCTTCCAGC	68.0
SNAPf-HisTag-FLAG	CLIPf-SNAPf-HisTag-FLAG	GGATCCGCTGGATCTGCT	GGTACCAAGCTTAAGTTTAAACG	64.0
CLIPf-mKIF5A-SNAPf-HisTag-FLAG	pRK5-Myc-KIF5A	GGGAAAACCTGGCCTCGGAG GATCCATGGCGGAGACTAACAAC	TCATTTCACAATCCTTATCG AATTCGCTGGCTGCTGTCTCTTG	62.6
CLIPf-hKIF5B-SNAPf-HisTag-FLAG	pRK5-Myc-KIF5A	CATCATGGATCCATGGCGGACC TGGCC	GTAGTAGAATTCTTACACTTGT TTGCCTCCTCC	66.0
science CLIPf-mKIF5A	CLIPf-mKIF5A-SNAPf-HisTag-FLAG	TCGAGTCTAGAGGGCCCG	GCTGGCTGCTGTCTCTTG	67.0
mKIF5A-SNAPf-HisTag-FLAG	CLIPf-mKIF5A-SNAPf-HisTag-FLAG	ATGGCGGAGACTAACAAC	AAGCTTAAGTTTAAACGCTAG	60.0
CLIPf-mKIF5A(Δ954-1027)-SNAP-SNAPf-HisTag-FLAG	CLIPf-mKIF5A-SNAPf-HisTag-FLAG	GAATTCGATAAGGATTGTGAAATG	GTTCTGGAAGAGGTTGTTG	60.0
CLIPf-mKIF5A(Δ505-610)-SNAP-SNAPf-HisTag-FLAG	CLIPf-mKIF5A-SNAPf-HisTag-FLAG	CTCCAGGTGGAGTGTCAT	CTCCTGGGACTTCTGGTC	65.0
CLIPf-mKIF5A(1-915; Δ505-610)-SNAPf	CLIPf-mKIF5A(Δ505-610)-SNAP-SNAPf-HisTag-FLAG	GAATTCGATAAGGATTGTGAAATG	ATGGCCCCGCTTGCC	60.0
LAMP1-CLIPf-SNAPf-HisTag-FLAG	Lamp1-mScarlet-I	TAGCGTTTAAACTTAAGCTTATGGC GGCCCCCGGCAGC	ATCTCGCAGTCCTTGTCCATGAT AGTCTGGTAGCCTGCGTGACTCC TCTTCCTG	72.0
FLAG-CLIPf-TPR2-SNAPf-Strep	FLAG-CLIPf-TPR3-SNAPf-Strep	TGGAATTAGATCCGAACAATGCA GAAGCCAAACAGAATC	ATTGTTCGGATCTAATTCCAG	60.1
FLAG-CLIPf-TPR1-SNAPf-Strep	FLAG-CLIPf-TPR3-SNAPf-Strep	GGAACTGGACCCGAATAATGCA GAAGCCAAACAGAATC	ATTATTCGGGTCCAGTTC	59.2
FLAG-CLIPf-SNAPf-Strep	FLAG-CLIPf-TPR3-SNAPf-Strep	TGGGCAAACCTGGTTTAGGCGAC AAAGACTGTGAGATGAAAC	GCCTAAACCAGGTTTGCC	61.8

Table 2.3: Primers used in PCR. Sequence in blue anneals to the template DNA. Underlined sequence indicates restriction enzyme sites.

2.1.2.3 PCR product purificiation

DNA fragments produced by PCR, requiring restriction digest before insertion into a vector, were purified using Monarch PCR and DNA Cleanup Kit (New England BioLabs) following the manufacturer's instructions. Briefly, the PCR samples were diluted in DNA Cleanup Binding Buffer at a ratio of binding buffer to sample of 2:1 and mixed by pipetting up and down gently. The diluted samples were loaded into the DNA binding column and centrifuged at 16,000 RCF for 1 minute. The supernatant was discarded and 200 µl of DNA Wash Buffer was added to the column before centrifuging again at 16,000 RCF for 1 minute. This step was repeated for the second wash. The DNA was eluted by adding 37 °C nuclease free water to the column and leaving to incubate at room temperature for one minute. The column was then spun for 1 minute at 16,000 RCF. The final concentration of DNA was determined by the absorbance at 260 nm using a NanoDrop (Labtech International).

2.1.2.4 Restriction digests

High-Fidelty restriction enzymes from New England BioLabs were used along with CutSmart Buffer for all restriction enzyme reactions.

1 µg of DNA to be cut was added to 5 µl of CutSmart buffer and 1 µl of each restriction enzyme. The total reaction volume was made up to 50 µl with nuclease free water. The reactions were incubated at 37 °C for 15-30 minutes before adding 10 µl 6x loading dye to stop the reaction. Samples were run on a 1 % agarose gel, in lithium-borate buffer (see Table 2.1), at 80 volts for 30 minutes and the DNA was then extracted (see Section 2.1.2.5).

2.1.2.5 Gel extraction

DNA fragments were extracted from agarose gel using Monarch DNA Gel Extraction Kit (New England BioLabs) following the manufacturer's instructions and all centrifugation steps were carried out at room temperature. Briefly, the gel fragments

containing the desired DNA were cut out and diluted in four times the amount of Gel Dissolving Buffer (e.g. 100 mg gel fragment was diluted in 400 µl of the buffer). The sample was heated at 50 °C for around 10 minutes until the gel slice was dissolved. The solution was then loaded onto a DNA binding column and centrifuged at 16,000 RCF for 1 minute. The supernatant was discarded and 200 µl of DNA Wash Buffer was added to the column before centrifuging again at 16,000 RCF for 1 minute. This step was repeated for a second wash. To elute the DNA, nuclease free water (heated to 37 °C) was added to the column and left to incubate for 1 minute before a final spin for 1 minute at 16,000 RCF. The concentration of the eluted DNA was determined by the absorbance at 260 nm using a NanoDrop (Labtech International).

2.1.2.6 Isothermal assembly

DNA fragments were assembled following the New England BioLabs HiFi DNA Assembly Protocol. A DNA molar ratio of 1:5 (vector to insert) was used and the amount of each fragment was calculated using the New England BioLabs online NEBioCalculator. Briefly, 10 µl of NEBuilder HiFi DNA Assembly Master Mix was added to the DNA fragments. Nuclease free water was then added to make a total reaction volume of 20 µl. This mixture was incubated at 50 °C for 30 minutes before transformation into chemically competent cells (see Section 2.1.2.8).

2.1.2.7 Ligation

Ligation of two restriction digest DNA fragments was done using the New England BioLabs T4 ligation protocol. The amount of each fragment was again calculated using the New England BioLabs online NEBioCalculator for a final DNA molar ratio of 1:5 (vector to insert). Briefly, 2 µl of T4 DNA Ligase Buffer was added to the DNA fragments. 1 µl of T4 DNA ligase was added along with nuclease free water to make a total reaction volume of 20 µl. This mixture was incubated at room

temperature for 10 minutes before heat inactivation at 65 °C for 10 minutes. 2 µl of the sample was then transformation into chemically competent cells (see Section 2.1.2.8).

2.1.2.8 Transformation of bacteria with plasmid DNA

Plasmid DNA was transformed into chemically competent *E.coli* cells following the New England BioLabs Transformation Protocol. Briefly, 50 µl of *E. coli* cells were thawed on ice for approximately ten minutes. 5-10 ng of DNA was then added and gently mixed without vortexing. This mixture was incubated on ice for 30 minutes before being heat shocked at 42 °C for 30 seconds. 950 µl of room temperature SOC media (see Table 2.1) was added and the cells incubated with shaking at 37 °C for 60 minutes. Cells were then spun down at 2500 RCF for 2 minutes at 4 °C. 900 µl of the supernatant was removed and the pellet resuspended in the remaining 100 µl. This was spread onto room temperature selection plates and incubated at 37 °C overnight.

2.1.2.9 Maxi-preparation of plasmid DNA

PureYields Plasmid Maxiprep System (Promega) was used to extract and purify plasmid DNA from bacterial cultures of 250 ml. Manufacturer's instructions were followed and any solutions mentioned are provided by the manufacturer. All centrifugation steps were done at room temperature. Briefly, bacterial cultures were grown overnight at 37 °C with shaking. The cells were pelleted by centrifugation at 5,000 RCF for 10 minutes. The supernatant was discarded and the pellet resuspended completely in 12 ml of Cell Resuspension Solution. The same amount of Cell Lysis Solution was then added and the solution mixed by inversion (3-5 times) before leaving to incubate for 3 minutes at room temperature. 12 ml of Neutralization Solution was then added to the lysed cells and the solution again mixed by inversion (10-15 times). To pellet the cellular debris the lysate was centrifuged at

14,000 RCF for 20 minutes. The supernatant was placed in a blue PureYield Clearing column which was placed on top of a white PureYield Maxi Binding Column on a vacuum manifold. Maximum vacuum was applied until all the lysate had passed through both the clearing and the binding columns. The clearing column was then removed and 5 ml of Endotoxin Removal Wash was poured into the binding column. Maximum vacuum was applied until all the solution was pulled through the column. 20 ml of Column Wash was then poured into the binding column. Again maximum vacuum was applied until all the wash had been drawn through the column and then left on for a further 5 minutes in order to dry the membrane. To elute, 1 ml of nuclease free water was added to the membrane and left to incubate for 1 minute before application of maximum vacuum for a further minute. The concentration of the eluted DNA was determined by the absorbance at 260 mm using a NanoDrop (Labtech International).

2.1.2.10 Batch Expression *E.coli* BL21 bacteria

A single colony from a fresh streak out plate was added to 50 ml of LB media (see Table 2.1) with 50 µg/ml Ampicillin and 35 µg/ml Chloramphenicol added. The culture was incubated overnight at 37 °C with shaking. This overnight culture was added to 1 L of TB (see Table 2.1) with 50 µg/ml Ampicillin and 35 µg/ml Chloramphenicol added. This day culture was incubated at 37 °C with shaking until the OD_{595} reached 0.8-1. The culture was then placed on ice until it reached room temperature. Isopropyl B-D-1 thiogalactopyranoside (IPTG) was added, to induce protein expression, at a final concentration of 0.2 mM. Cultures were then incubated overnight, at room temperature, with shaking.

The next day the culture was centrifuged at 4000 RCF for 15 minutes at 4 °C. The supernatant was removed and the pellet resuspended in 10 ml of 2x PBS+DTT solution (see Table 2.1) before centrifugation at 4000 RCF for 20 minutes at 4 °C. The supernatant was removed and the pellet flash frozen in liquid nitrogen before

storage at -80 °C until needed.

2.2 Cell Culture and transfection

2.2.1 Culture of HEK293 cells

Human Embryonic Kidney (HEK293, 293 (ECACC 85120602)) cells were kept under 5 % CO_2 and at 37 °C. They were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % fetal bovine serum (FBS) and penicillin-streptomycin (final concentration of 50 U/mL). When 70-80 % confluency was reached (usually every 2-3 days) the cells were passaged in T75 flasks. The used media was removed before washing the cells with 5ml phosphate-buffered saline (PBS). Cells were lifted by adding 2 ml of trypsin and leaving the cells to incubate at 37 °C for around five minutes. Fresh media (8 ml) was added to neutralise the trypsinisation process before the cells were collected in a centrifuge tube. The cell suspension was spun at 400 RCF for four minutes and the supernatant removed. The cell pellet was resuspended in 5 ml of supplemented media and added to a new T75 flask according to the desired sub-culturing ratio (usually 1 ml) along with 12 ml of supplemented media.

2.2.2 Transfection of HEK293 cells

Volumes for each well of a six-well plate:

HEK293 cells were transfected using SignaGen Laboratories LipoD293 DNA In Vitro Transfection Reagent following the manufacturer's instructions. Briefly for one well of a 6-well plate, the day before transfection 200 µl of the resuspended cell pellet (resuspended in 5 ml) was added to each well along with 3 ml supplemented media. 30-60 minutes before transfection, 2 ml of the media was removed (leaving 1 ml on the cells). In a microcentrifuge tube, a volume of DNA equal to 1 µg was diluted in 50 µl of DMEM (not supplemented) and gently mixed. 3 µl LipoD293

reagent was diluted and gently mixed in 50 µl of DMEM (not supplemented) in a separate microcentrifuge tube. This diluted transfection reagent was added dropwise to the diluted DNA solution and gently mixed. This mixture was left to incubate for 10 minutes at room temperature and then added dropwise to the well. The reagent was then left on the cells.

Volumes for a 10 cm dish:

The same protocol was followed when transfecting a 10 cm dish but the volumes were: 1 ml of the resuspended cell pellet (resuspended in 5 ml) with 12 ml supplemented media, remove 7 ml media (leaving 5 ml on the cells), 5 µg of DNA diluted in 250 µl of DMEM (not supplemented) and 15 µl LipoD293 reagent diluted in 250 µl of DMEM (not supplemented).

For the ligand labelling optimisation, differences in transfection efficiencies between wells were controlled by transfecting a 10 cm dish and then replating the cells in a 6-well plate. Briefly, 5 hours after transfection of a 10 cm dish, cells were passaged (see Section 2.2.1) with the resultant cell pellet being resuspended in 6 ml of supplemented media. 1 ml of cell suspension was then placed in each well of a 6-well plate.

2.2.3 Ligand incubation and washing of cells

Two ligands are used in this thesis; CLIP-Cell TMR-Star and SNAP-Cell 647-SiR. Both are maunfactured by New England BioLabs. On arrival, 50 µl of DMSO is added to the vial, containing 30 nmol of ligand, and mixed thoroughly. The resuspended ligand (0.6 mM) is then aliquoted into 1 µl one use samples and kept in the dark at -20 °C until used.

Ligands were added to the cells in a well of 6-well plate according to the specific incubation time and concentration optimised for that dye (see Table 2.4). The cells were then washed twice with 500 µl of supplemented media and left to incubate for the optimised time specific for that dye (see Table 2.4 and Section 3.2.2.1).

Unoptimised labelling conditions are given specifically in the text when used.

Ligand	Wash incubation	Ligand incubation	Ligand final concentration
CLIP-Cell TMR-Star	2 hours	Overnight	0.5 µM
SNAP-Cell 647-SiR	2 hours	2 hours	0.5 µM

Table 2.4: Optimised ligand labelling conditions

2.2.4 Collecting transfected and treated HEK293 cells with trypsin

Volumes for each well of a six-well plate:

Media was removed before washing with 1 ml PBS. 0.5 ml trypsin was then added and the cells left to incubate at 37 °C for around 5 minutes. Cells were then resuspended in 1 ml of cell culture media before transfer to a 1.5 ml microcentrifuge tube. The cell suspension was spun at 500 RCF for 4 minutes at 4 °C. The supernatant was removed and the cell pellet was resuspended in PBS (volume depends whether cells are being prepared for flow cytometry or being lysed for SDS-PAGE or smFRET).

2.2.5 Collecting transfected HEK293 cells without trypsin

2.2.6 Cell lysis of HEK293 cells with detergent

Volumes for each well of a six-well plate:

The cell pellet (end of Section 2.2.4) was resuspended in 1 ml 1x PBS. This cell suspension was spun down at 500 RCF for 4 minutes at 4 °C. The resulting pellet was resuspended in 100 µl of ice cold lysis buffer (see Table 2.5) and left to incubate

on ice for 10 minutes. Samples were then spun down at 17,000 RCF for 10 minutes at 4 °C. The resultant supernatant was used as necessary for onward experiments.

	Component	Final concentration
Lysis buffer	HEPES (pH 7.5)	40 mM
	EDTA (pH 8)	1 mM
	NaCl	120 mM
	Triton X-100	0.05 %
	Aprotinin	1 µg/ml
	Leupeptin	10 µg/ml
	Pepstatin A	1 µg/ml
	PMSF	100 µM
	TAME	10 µg/ml
smFRET lysis buffer	HEPES (pH 7.5)	40 mM
	EDTA (pH 8)	1 mM
	NaCl	120 mM
	Triton X-100	0.05 %
	Aprotinin	1 µg/ml
	Leupeptin	10 µg/ml
	Pepstatin A	1 µg/ml
	PMSF	100 µM
	TAME	10 µg/ml
	BSA	0.1 mg/ml
	Mg.ATP	1 mM

Table 2.5: Lysis buffers

2.2.7 Cell lysis of HEK293 cells without detergent - homogenisation

This protocol is based on Araújo et al. 2015 and Aguado et al. 2016.

Volumes for a 10 cm dish:

Cells were collected without trypsin to prevent breakdown of proteins and loss of protein-protein interactions. The 10 cm dish of cells was placed on ice and washed three times in 5 ml of ice cold PBS. 3 ml of ice cold PBS+ (see Table 2.6) was added to the dish and the cells were deattached from the dish by scraping (always being

immersed in the PBS+). The cell suspension was collected in a 15 ml centrifuge tube and spun down at 4 °C for 5 minutes at 400 RCF. The resultant cell pellet was gentle loosened in 3 ml of ice cold homogenisation buffer (see Table 2.6) but not completely resuspended. This suspension was spun down at 1300 RCF for 10 minutes at 4 °C and the pellet then resusupended in 3x the pellet volume (usually 300 µl) of ice cold homogenisation buffer. During this centrifuge step, the cell homogeniser (Isobiotec Cell Homogenizer), containing a 4 µm clearance ball bearing, and two gas tight syringes (SGE Gas Tight Syringes, Luer Lock) was washed with cold homogenisation buffer to remove all air bubbles. The cell suspension was then passed through the homogeniser ten times before collection in a 1.5 ml microcentrifuge tube. This solution was diluted in homogenisation buffer in a ratio of 1:0.7 (usually 0.5 ml homogenate is diluted in 0.35 ml of homogenisation buffer), before being spun down for 10 minutes at 4 °C at 17968 RCF. The resultant pellet is the nuclear pellet which is taken on for onward experiments. The supernatant is spun again at 17968 RCF for 10 minutes at 4 °C where the resultant supernatant is the postnuclear supernatant which is taken on for onward experiments.

2.2.8 Sample preparation for SDS-PAGE

Cell lysate samples were diluted in 3x sample buffer (See Table 2.10) and boiled at 100 °C for 5 minutes. Samples were stored at -20 °C before analysis by SDS-PAGE, in gel fluorescence and/or western blotting.

2.3 smFRET

The single molecule fluorescence resonance energy transfer (smFRET) confocal microscope used is custom built with separate lasers, filters and detectors for the donor and acceptor channels respectively. A full description of the custom set up including the full specification of parts and instrument layout is given in detail in Ambrose

	Component	Final concentration
PBS+	NaCl	137 mM
	KCl	2.7 mM
	KH_2PO_4	1.5 mM
	Na_2HPO_4	6.5 mM
	$CaCl_2$	1 mM
	$MgCl_2$	1 mM
Homogenisation buffer	Sucrose	250 mM
	HEPES (pH 7.5)	10 mM
	EDTA (pH 8)	1 mM
	Aprotinin	1 µg/ml
	Leupeptin	10 µg/ml
	Pepstatin A	1 µg/ml
	PMSF	100 µM
	TAME	10 µg/ml

Table 2.6: Homogenisation buffers and solutions

et al. 2020. The donor laser, 80 mW but used at 50 % power , excites at 515 nm and any donor emission is filtered through a 571/72 emission filter before being detected by an avalanche photodiode. A 100 mW 638nm laser, used at 15 % power, excites the acceptor fluorophore with any emission filtered (678/41) before being detected by a second APD . 30 µl of diluted cell lysate is added above the lens on a clean coverslip. Cell lysate is diluted in smFRET assay buffer usually around 100,000 times to achieve a single molecule concentration (roughly 1 fM). Data was collected for 15 minutes unless otherwise stated.

The raw data from the smFRET experiments consists of the photon arrival times and the detector can be determined. The laser duty cycle is then applied to determine which laser was on at the time of excitation (Ambrose et al. 2020). Data analysis is done using custom built Jupyter and R notebooks. Examples of these can be seen in the appendix. The Jupyter notebooks were made by the Craggs lab and analyses the raw data using the FRETBursts python module (Ingargiola et al. 2016) The notebook estimates the experimental background based on the time between

detected photons. This is done by plotting the log of the number of photons against the delay time between them. A poisson distribution is fitted and any photons below this line are removed as background photons. The remaining photons are grouped into bursts. The bursts threshold is set to 20 photons so only bursts with 20 photons or more above background were selected.

2.3.1 Initial smFRET assay protocol

This protocol was used in the validation of the smFRET assay (see Section 3.2.1).

HEK293 cells were plated and transfected before overnight incubation with 0.25 µM CLIP-Cell TMR-Star and 0.25 µM SNAP-Cell 647-SiR. The next day cells were washed (wash incubation was 1 hour for both fluorescent ligands), collected and then lysed in 100 µl of lysis buffer (see Table 2.5 and see Sections 2.2.1, 2.2.2, 2.2.4 and 2.2.6). The resultant supernatant was diluted in initial smFRET assay buffer (see Table 2.7) in order to give single-molecule resolution on the smFRET confocal microscope.

Data analysis was done using the original version of the custom built Jupyter notebooks and then plotted using an R notebook (see Appendices - Initial smFRET data analysis). The first Jupyter notebook is use to convert to .txt file produced by the smFRET software to a .HDF5 file for data analysis. This allows the identity of the relevant laser to be matched to the photons depending on the laser duty cycle. The second Jupyter notebook is used for the data analysis; initial elimination of background and the selection of bursts. This notebook does not apply any dye specific correction factors.

2.3.2 Photo bleaching BSA

Lyophilized bovine serum albumin (BSA) powder that was $\geq$ 99 % pure by agarose gel electrophoresis and also essentially globulin free (and fatty acid free for extra purity) was diluted in double distilled water to a give 20 mg/ml solution. The

	Component	Final concentration
Initial smFRET assay buffer	HEPES (pH 7.5)	40 mM
	NaCl	5 mM
	BSA	0.1 mg/ml
	DTT	1 mM
Kinesin smFRET assay buffer	PIPES (pH 6.7)	12 mM
	EGTA (pH 7)	1 mM
	$MgCl_2$	2 mM
	BSA	0.1 mg/ml
	DTT	1 mM
	Mg.ATP	1 mM
TPR smFRET assay buffer	Tris (pH 8.0)	40 mM
	NaCl	5 mM
	BSA	0.1 mg/ml
	DTT	1 mM

Table 2.7: smFRET assay buffers

solution was filtered using a 0.22 µm syringe filter before storing in 1 ml aliquots at -20 °C until needed. The filtered BSA was then photo bleached using a custom built set up (see Figure 2.1) for at least three days at 4 °C. This set up is made of: a thick cardboard box with a magnetically sealed lid and the inside covered in foil, a full spectrum (380-800nm) LED grow lamp, string (or an elastic band) and a clip to hold the BSA. The BSA is contained with in a thin strip of plastic sheet and held between the two bulbs.

2.3.3 Optimised protocol for kinesin-1 conformation smFRET assay

This protocol was used for any smFRET experiments after all optimisation steps were carried out.

HEK293 cells were plated, transfected, incubated with fluorescent ligands, washed

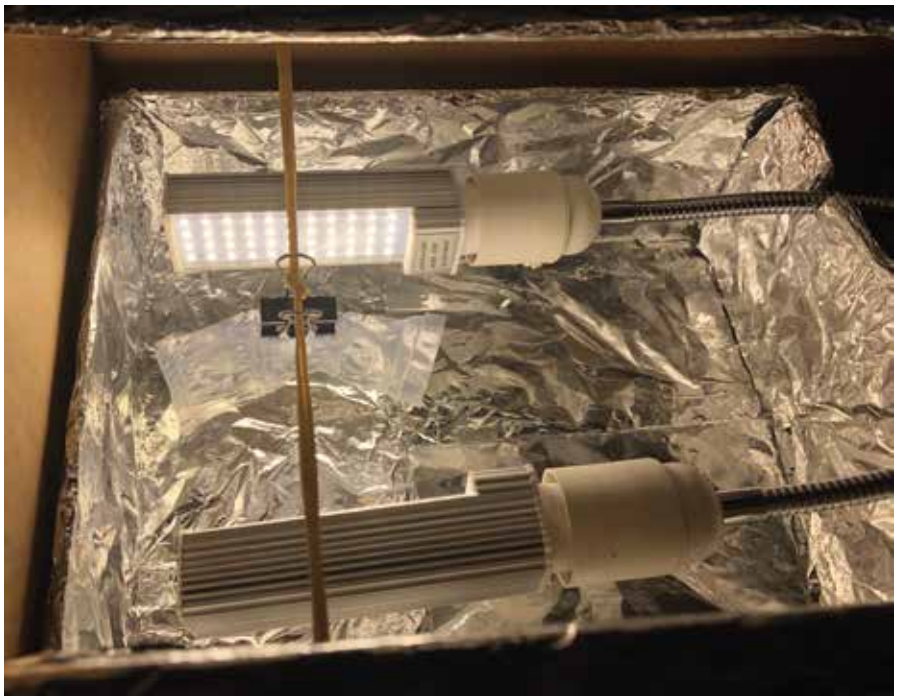

Figure 2.1: Custom built photo bleaching set up for BSA.

(optimised conditions specific to the fluorescent ligand for concentration, ligand incubation and washout incubation was used (see Table 2.4)), collected with trypsin and lysed in 50 µl of smFRET lysis buffer (see Table 2.5) as described previously (see Cell culture and transfection sections). The resultant supernatant was diluted in kinesin smFRET assay buffer (see Table 2.7) in order to give single-molecule resolution on the smFRET confocal microscope.

Data analysis was done using the custom built Jupyter notebooks and then plotted using a R notebook (see Appendices - Optimised smFRET data analysis). The Jupyter notebook is used for the data analysis; initial elimination of background and the selection of bursts. Then the R notebook is used to filter and plot the data. Dye specific correction factors were applied.

2.4 Flow cytometry

2.4.1 Determining washout of fluorescent ligands

Untranfected HEK293 cells were plated, incubated in the fluorescent ligand and washed as described previously (see Sections 2.2.1 and 2.2.3). They were then collected (described in Section 2.2.4) and the resultant pellet resuspended in 0.5 ml PBS. The cell suspension was analysed using a BD FACS Melody flow cytometer using the relevant lasers and filters according to the fluorescent ligand being used (see Table 2.8).

Fluorescent ligand	**Laser**	**FACS filter**	**Licor channel**
CLIP-Cell TMR-Star	561 nm	PE 582/15	600 nm
SNAP-Cell 647-SiR	633 nm	APC 660/10	700 nm

Table 2.8: Laser and filter specifications on the BD FACS Melody flow cytometer and channel specification for Licor Odyssey Fc Imager for each fluorescent ligand.

Data was analysed using a custom built R notebook (see Appendices - Flow cytometry data analysis). Gating of the data did not take place as the cells were not going to be presorted before the smFRET assay so all the fluorescence signal needed to be analysed. Also compensation was not used as only one fluorophore was added to the cells at a time.

2.5 Biochemistry

2.5.1 Co-immunoprecipitation

HEK293 cells were plated in a 6-well dish and transfected as previously stated (see Sections 2.2.1, 2.2.2). A day after transfection the media was removed from the cells and 1.5 ml PBS was added. Cells were deattached by scraping and the cell

suspension transferred to a 1.5 ml microcentrifuge tube before being spun down at 400 RCF for 2 minutes at 4 °C. The pellet was resuspened in 1 ml PBS before centrifugation at 2500 RCF for 5 minutes at 4 °C. The resultant pellet was resuspended in 300 µl of ice cold IP lysis buffer (see Table 2.9). The cell suspension was left to incubate on ice for 10 minutes before being spun down at 17000 RCF for 10 minutes at 4 °C. 40 µl of the supernatant was removed and kept as a sample of the input. The rest (around 250 µl) was placed in a low bind 1.5 ml microcentrifuge tube and incubated with 1 µg of the α-FLAG antibody (Proteintech) with rotation for 1 hour at 4 °C. During this incubation, 25 µl of Dynabeads Protein G (Thermo Scientific) were placed in a low bind 1.5 ml microcentrifuge tube on a magnetic rack and washed with PBS Tween (see Table 2.9). The tubes were placed back on the rack and the supernatant removed. After incubation with the antibody, the cell lysate was added to the beads and left to incubate at room temperature with rotation for 10 minutes. A sample (40 µl) was then taken as the flow through. The bead suspension was washed three times with PBS and put in a fresh 1.5 ml microcentrifuge tube on the last wash. The beads were resuspended in 1x sample buffer (see Table 2.10). This is the sample of the IP.

.

2.5.2 SDS page, native-PAGE and western blotting

2.5.2.1 SDS-PAGE

Polyacrylamide gels formed of a high pH (8.8) 10 % or 15 % resolving gel and a low pH (6.8) 5 % stacking gel (see Table 2.10) were cast in Novex 1.5 mm cassettes. Samples were run for approximately 1.5 hours at 150 Volts on a Hoefer MiniVE Mini Vertical Electrophoresis Unit in running buffer (see Table 2.10).

Polyacrylamide gels not used for either in gel fluorescence or western blotting were incubated for 15 minutes at room temperature in Quick Coomassie Stain (BioServ UK Limited) to visualise any protein present. Several washes with nu-

	Component	Final concentration
IP lysis buffer	HEPES (pH 7.5)	50 mM
	EDTA (pH 8)	1 mM
	$MgCl_2$	1 mM
	NaCl	25 mM
	Triton	0.5 %
	Aprotinin	1 µg/ml
	Leupeptin	10 µg/ml
	Pepstatin	1 µg/ml
	PMSF	100 µM
	TAME	10 µg/ml
PBS Tween	Tween	0.02 %
	PBS	-

Table 2.9: IP buffers and solutions

clease free water were then done over several hours to destain the gel. The next day gels were imaged using a Syngene G:Box with transilluminator.

Gels that were to be used for western blotting were incubated in ponceau solution (see Table 2.10) for five minutes as room temperature and then destained in several washes of nuclease free water.

2.5.2.2 In gel fluorescence

Fluorophore labelled cell lysate samples were run on a SDS-PAGE gel (see Section 2.5.2.1), with a box blocking out any light to prevent bleaching of the sample. The gel was placed in nuclease free water and covered in foil until visualisation. This was done using the Licor Odyssey Fc Imager in the 600/700 nm or 800 nm channel depending on the fluorescent ligand the samples were incubated with (see Table 2.8).

In the Licor Image Studio Lite software bands were selected and the signal, background, area and total signal was recorded for each. This data was exported and analysed in custom written R notebook (see Appendices - In gel fluorescence data analysis). Each biological repeat was normalised to the lowest condition.

2.5.2.3 Native PAGE

Cell lysate samples were stored in non-denaturating sample buffer (see Table 2.11) at -20 °C.

Polyacrylamide native gels formed of a high pH (8.8) 10 % resolving gel and a low pH (6.8) 5 % stacking gel (see Table 2.11) were cast in Novex 1.5 mm cassettes. Samples were run for approximately 1.5 hours at 150 Volts on a Hoefer MiniVE Mini Vertical Electrophoresis Unit in running buffer (see Table 2.10).

Gels were incubated in ponceau solution (see Table 2.10) for five minutes at room temperature and then destained in several washes of nuclease free water before transfer (see Section 2.5.2.4) and western blotting (see Section 2.5.2.5).

2.5.2.4 Transferring to a membrane

SDS-PAGE gels for use in western blotting were transferred to PVDF membrane (immoblion-FL membrane) using the Hoefer Blot Module in the Hoefer MiniVE Mini Vertical Electrophoresis Unit. The membrane was activated by incubation in methanol for 15 seconds before washing in water for 2 minutes and pre-soaking in transfer buffer (see Table 2.12) for at least five minutes. The membrane was then sandwiched with the polyacrylamide gel between 2 pieces of filter paper and padded with sponges either side. Samples on the gel were transferred at 300 mA for 72 minutes in transfer buffer with ice and cold water in the Electrophoresis Unit to keep the module cold.

2.5.2.5 Western blotting

After transfer, the membranes were blocked by incubation in blocking solution (see table 2.12) for 1 hour at room temperature. Then, primary antibodies were diluted in this blocking solution (for antigens and concentrations see Table 2.13) and incubated with the membrane overnight at 4 °C. The next day the membrane was washed three times for 15 minutes in wash solution (see table 2.12) to remove excess primary

antibody. Then the secondary antibody was diluted in the wash solution (1:50,000) and incubated with the membrane for 1 hour at room temperature in the dark before repetition of the three wash steps. The membrane was then washed once in nuclease free water before visualisation using Licor Odyssey Fc Imager in the 600 nm channel for the ladder and the 700 nm and/or 800 nm channel depending on the secondary antibody used.

2.6 Imaging

2.6.1 Immunofluorescence staining of HEK293 cells

Due to COVID restrictions some parts of this protocol were done by other members of the lab; coverslip preparation by Ashleigh Davey, cell plating by Emma Turner and imaging by Dr. Twelvetrees.

Coverslips were incubated in 20 µg/ml of poly-L-lysine (PLL) solution at 37 °C. The next day, this solution was removed and the coverslips washed twice with nuclease free water before cells were plated (see Section 2.2.1) and left in the incubator overnight. Cells were then transfected (see Section 2.2.2) with the LAMP1-CLIPf-SNAPf-HisTag-FLAG construct (see Section 2.1.1.3). The next day 0.5 µM of SNAP-Cell 647-SiR was added to the wells and left to incubate for two hours before fixation. For colocalisation experiments, LysoTracker™ Red DND-99 (Invitrogen) was also added to the wells for a final concentration ratio of 1:10,000. This was done one hour after addition of SNAP-Cell 647-SiR and left to incubate for one hour. Coverslips were washed three time in phosphate-buffered saline (PBS) before being fixed with 4 % paraformaldehyde for 10 minutes. The PBS washes were repeated before adding 300 nM of DAPI stain solution and incubating in the dark at room temperature for 10 minutes. The stain was then removed and the PBS washes repeated again. Coverslips were mounted onto microscope slides using ProLong® Gold antifade reagent (Invitrogen) and sealed around the edge with nail

varnish. Cells were imaged using a Leica SP5 confocal microscope with the LAS AF software. DAPI staining was imaged with the 405 nm laser, LysoTracker with the 561 nm laser and SNAP-Cell 647-SiR with the 633 nm laser. Image analysis was done in ImageJ.

2.7 *In vitro* reconstitution

2.7.1 Polymerising microtubules

Lyophilised tubulin is resuspended in GBRB80 (see Table 2.14) for a final concentration of 5 mg/ml. Aliquots are flash frozen and stored at -80 °C.

40 µl of 5 mg/ml unlabelled tubulin was added to an ultracentrifuge tube and spun for 10 minutes at 80,000 RCF at 4 °C to remove any insoluble tubulin. The supernatant was then incubated at 37 °C for twenty minutes to allow the tubulin to polymerise into microtubules. To stabilise these microtubules 0.8 µl of 2 mM Paclitaxel was added before swirling gently and leaving to incubate for a further twenty minutes at 37 °C. Microtubules were pelleted by spinning this solution at room temperature for 10 minutes at 80,000 RCF. The supernatant was discarded and 45 µl of TBRB80 (see Table 2.14) was gently added and left to resuspend overnight.

2.8 Protein purification

Bacteria were grown and induced as detailed in Section 2.1.2.1. The bacteria cell pellet was thawed at room temperature and resuspended in B-PER Complete Bacterial Protein Extraction Reagent (Thermo Scientific) and Halt protease inhibitor cocktail (Thermo Scientific). Volumes depend on weight of pellet; 5 ml B-PER and 50 µl protease inhibitor per 1 g of pellet. The resuspended solution was incubated on a roller for 20 minutes before centrifugation at 23,159 RCF for 30 minutes at 4

°C. The supernatant was removed and filtered (0.45 µm). Protein was purified from cleared bacterial lysate using an ÄKTA Pure 25 L system (Cytiva). Briefly, filtered lysate was loaded onto a StrepTactinXT 4Flow high capacity 1 ml column (Thermo Scientific). Column was then washed in 10 column volumes (10ml) of wash buffer (see Table 2.15) before 1 ml of 4.25 µM filtered (0.45 µm) CLIP-Cell TMR-Star and SNAP-Cell 647-SiR was injected onto the column. The column was left to incubate for 18 hours before it was washed again in 10 column volumes of wash buffer (see Table 2.15). Protein was eluted with a gradient of elution buffer (two column volumes) up to 50 mM biotin for 30 column volumes (see Table 2.15). Fractions containing eluted protein were concentrated by centrifugation in a 10,000 kDa molecular weight cut-off filter (Amicon Ultra-15) to 1 ml. Concentrated protein was then separated by size-exclusion chromatography using a HiLoad™ 16/600 Superdex™ 200 pg column (Cytiva) in wash buffer (see Table 2.15). Eluted protein was flash frozen in liquid nitrogen before storage at -80 °C.

2.9 Computional modelling

2.9.1 AlphaFold2

The AlphaFOld Multimer Google Colab notebook was used for all AlphaFold2 predictions (Jumper et al. 2021 and Evans et al. 2022). The amino acid sequence of the region of interest was inputted into the notebook and the default parameters were used to make the models. The highest ranked model was then chosen. MOdels were presented and aligned using ChimeraX software.

	Component	Final concentration
3x sample buffer	Tris (pH 6.8)	50 mM
	SDS	2 %
	DTT	0.1 M
	Bromophenol	0.1 %
	Glycerol	10 %
10 % resolving gel	30 % Acrylamide	10 %
	Tris, pH 8.8	0.38 M
	SDS	0.1 %
	APS	0.1 %
	TEMED	0.04 %
15 % resolving gel	30 % Acrylamide	15 %
	Tris, pH 8.8	0.38 M
	SDS	0.1 %
	APS	0.1 %
	TEMED	0.04 %
5 % stacking gel	30 % Acrylamide	5 %
	Tris, pH 6.8	0.125 M
	SDS	0.1 %
	APS	0.1 %
	TEMED	0.04 %
Running buffer	Tris	25 mM
	Glycine	192 mM
	SDS	0.1 %
Ponceau solution	Ponceau S	0.1 % (w/v)
	Acetic acid	5 %

Table 2.10: SDS-PAGE solutions

	Component	Final concentration
5x non-denaturing sample buffer	Tris (pH 6.8) Bromophenol Glycerol	310 mM 0.05 % 50 %
10 % non-denaturing resolving gel	30 % Acrylamide Tris, pH 8.8 APS TEMED	10 % 0.38 M 0.1 % 0.04 %
5 % non-denaturing stacking gel	30 % Acrylamide Tris, pH 6.8 APS TEMED	5 % 0.125 M 0.1 % 0.04 %

Table 2.11: Native PAGE solutions

	Component	Final concentration
Transfer buffer	Tris Glycine SDS	12.5 mM 96 mM 0.05 %
TBST, pH 7.6	Tris NaCl KCl Tween-20	20 mM 150 mM 50 mM 0.2 %
Blocking	Fish gelatin TBST	1 % -
Wash	BSA TBST	1 % -

Table 2.12: Transfer and western blotting solutions

Antigen	Species	Company	Product numbers	Western blotting
FLAG	Rabbit	Proteintech	20543-1-AP	1:2000
GAPDH	Mouse	Proteintech	6004-1-1g	1:5000
KIF5B	Rabbit	Proteintech	21632-1-AP	1:2000
Tubulin (alpha)	Rabbit	Proteintech	11224-1-AP	1:2000

Table 2.13: Commercial primary antibodies

	Component	Final concentration
BRB80	PIPES $MgCl_2$ EGTA	80 mM 2 mM 1 mM
GRB80	BRB80 GTP	- 1 mM
TRB80	BRB80 Paclitaxel	- 40 µM

Table 2.14: Buffers and solutions for polymerising microtubules

	Component	Final concentration
Wash buffer	Tris (pH 8.0) NaCl EDTA	100 mM 150 mM % 1 mM %
Elution buffer	Tris (pH 8.0) NaCl EDTA Biotin	100 mM 150 mM % 1 mM % 50 mM %

Table 2.15: Protein purification solutions

Chapter 3

Developing the use of self labelling enzymes in smFRET experiments

3.1 Introduction

3.1.1 Self-labelling enzymes and organic dyes

For single molecule work, organic dyes, rather than traditionally used fluorescent proteins, are extremely advantageous. This is down to their greater photostability, brightness and generally smaller size (Xia et al. 2013; Grimm, English, Chen, et al. 2015; Grimm, English, Choi, et al. 2016; L. Wang et al. 2019). Fluorescent proteins can oligomerise causing variability in the fluorescent signal. Cytotoxicity and mislocalisation of labelled protein can then occur. These proteins also have a certain maturation time before fluorescence occurs, which differs from protein to protein, leading to high variability in the fluorescent signal (Remington 2006 and Subach and Verkhusha 2012). But, organic dyes cannot be genetically encoded onto the protein of interest, like fluorescent proteins can, which can result in more complex experimental approaches.

Self-labelling enzymes combine the advantageous spectral and photo properties of organic dyes with the genetic encoding possibilities of fluorescent proteins (Grimm,

English, Choi, et al. 2016). Self-labelling enzymes are derived from naturally occurring enzymes but, are "broken" so the product of the enzymatic reaction is irreversibly bound and not released. The substrate ligand can be attached to many useful biomolecules; including organic dyes (Xia et al. 2013). These ligands can also be cell permeable for live cell imaging (Gautier et al. 2008) and a fusion protein of the protein of interest and the self-labelling enzyme can be made. There are several examples of these self-labelling enzymes, including the commercially available HaloTag (Los et al. 2008), SNAPf-tag (Keppler et al. 2003) and CLIPf-tag (Gautier et al. 2008).

The HaloTag was made to overcome protein purification problems. Product was being lost due to the reversible nature of ligand-affinity tag interactions. There was also a low abundance of affinity tags available for use. The HaloTag, based on haloalkane dehalogenases, and its ligands are specific and amenable (Encell 2013).

The SNAPf-tag is an engineered variant of the human DNA repair protein, O^6- alkylguanine-DNA alkyltransferase (hAGT,) which binds to O^6-benzylguanine derivatives (Xia et al. 2013; Keppler et al. 2003). The CLIPf-tag is derived from the SNAPf-tag but reacts specifically with O^2-benzylcytosine derivatives allowing the two tags to be used orthogonally (Gautier et al. 2008).

3.1.2 Using self-labelling enzymes in smFRET assays

The main aim of this thesis is to develop a single molecule fluorescence resonance energy transfer (smFRET) assay to study kinesin-1 autoinhibition, and for that I need a suitable labelling strategy. The combination of self-labelling enzymes and fluorescent ligands (ligands with organic dyes attached) was chosen due to their genetic encoding and photo properties. The lack of high resolution structural information available for kinesin-1 (see Section 1.3.2) meant fluorophores could not be added within the coiled coil region, for example with site specific cysteine labelling, without potentially altering the structure or function of kinesin-1. Previous ensem-

ble FRET work placed fluorescent proteins at either end of the protein and a change in FRET efficiency was seen due to a conformational change in kinesin-1 (Cai et al. 2007). Therefore it was decided to place the self labelling enzymes on the N- and C- termini of the kinesin-1 heavy chain.

The CLIPf-tag and SNAPf-tag were chosen for this smFRET assay. FRET is highly distant dependant process so the smaller size of these tags (19.4 kDa), compared to the HaloTag (33 kDa), was desirable. This combination of tags have also been used in smFRET experiments before (Ishitsuka et al. 2015). The tag specific fluorescent ligands chosen were CLIP-Cell TMR-Star and SNAP-Cell 647-SiR. Both these are commercially available and comply with the laser and detection parameters of the smFRET confocal instrument. The 515 nm donor laser will excite the CLIP-Cell TMR-Star whereas the 613 nm acceptor laser will only excite the SNAP-Cell 647-SiR. Also the emission spectra of the CLIP-Cell TMR-Star overlaps with the excitation spectra of the SNAP-Cell 647-SiR which is essential for FRET to occur (see Figure 3.1).

CLIP-Cell TMR-Star and SNAP-Cell 647-SiR are cell permeable giving a new option for sample preparation. Labelled purified protein diluted in assay buffer is the typical sample used in smFRET assays (Ishitsuka et al. 2015 and Agam et al. 2022). This minimises any background signal not due to the protein of interest. As both the ligands chosen are cell permeable, diluted cell lysate samples could be used instead. There is a possibility of an increase in background signal due to unspecific labelling and autofluorescence of the lysate, but access to the cellular environment would be retained.

The aim of this chapter is to validate the use of self labelling enzymes and cell permeable ligands in a smFRET based assay. The use of this labelling strategy, and the use of cell lysate samples, will be determined and optimised using a CLIPf-SNAPf fusion protein (see Section 2.1.1.1).

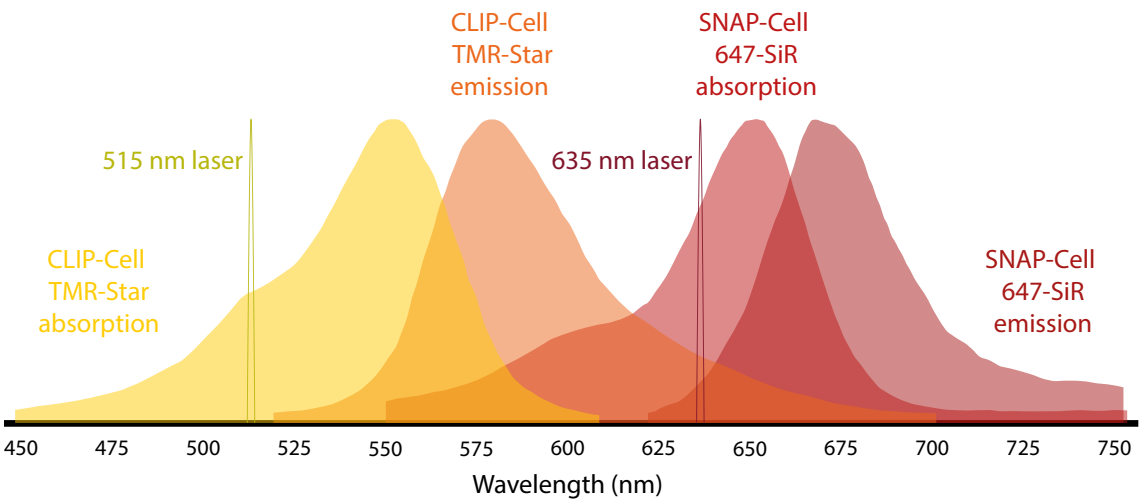

Figure 3.1: **Absorption and emission spectrum of FRET fluorophores**

Absorption (yellow) and emission (orange) spectra of CLIP-Cell TMR-Star. The absorption (red) and emission (dark red) spectra of SNAP-Cell 647-SiR. The donor laser is shown in green at 515 nm. The laser that can excite the acceptor fluorophore is shown in brown at 635 nm. There is an overlap between the emission of CLIP-Cell TMR-Star and the absorption of SNAP-Cell 647-SiR meaning that FRET between the two fluorophores is possible.
Spectra data from fpbase.org

3.2 Results

3.2.1 Validating the use of the self-labelling enzymes and cell lysate samples

It first needs to be established whether this labelling strategy, the combination of the CLIPf- and SNAPf-tags, with CLIP-Cell TMR-Star and SNAP-Cell 647-SiR is a viable approach as it has not been used in a smFRET assay before. To do this a double tagged construct, CLIPf-SNAPf-HisTag-FLAG (see Section 2.1.1.1) was designed by Dr. Alison Twelvetrees and synthesised by GeneArt (invitrogen by Thermo Fisher Scientific). It consists of the two tags (CLIPf and SNAPf) joined by a twelve amino acid flexible linker. This is then followed by another shorter linker, a His-Tag and a FLAG-tag. These final tags were added in case protein purification was needed and for other biochemical assays. The resultant labelled fusion protein should give a high FRET signal due to the close proximity of the tags and the bound fluorescent ligands.

To check if the CLIPf-SNAPf fusion protein resulted in a high FRET signal, HEK293 cells, plated in a 6-well plate, were transfected with the CLIPf-SNAPf-HisTag-FLAG construct. Both the CLIP-Cell TMR-Star and SNAP-Cell 647-SiR were added dropwise to the cells for a final concentration of 0.25 µM. This is much lower then the concentration recommended by the manufacturer (3 µM) as previous lab members had positive cell labelling at this concentration. The cells were left to incubate in the ligands overnight to maximise labelling. The next day cells were washed twice with 500 µl of fresh supplemented media before incubation for 1 hour in 1 ml of media to ensure removal of unbound ligand. Cells were collected using trypsin and the lysed using detergent (see Section 2.2.4 and 2.2.6). The lysate was diluted in initial smFRET assay buffer (see Table 2.7) and 50 µl of the diluted sample was then placed on the coverslip in the custom-built smFRET confocal instrument (see 2.3.1). Photons were collected for 15 minutes.

The output of each smFRET experiment is a table of photon arrival times and the detector (either the donor or acceptor avalanche photodiodes (APD)) that detected those photons. The laser duty cycle for the experiment is then matched to these results so that for each detected photon the laser which excited it, the APD that detected it, and the time this occured, is recorded.

The photons can then be separated by excitation and emission into the three FRET channels; donor-only (DD), donor-acceptor (DA) and acceptor-only (AA). This is only possible due to the use of alternating-laser excitation (ALEX). Donor-only (DD) photons are those that were excited by the donor laser and were detected by the donor APD. It is the same for acceptor-only (AA), but with the acceptor laser and APD rather then donor. The donor-acceptor (DA) channel consists of photons that were excited by the donor laser, but detected by the acceptor APD. Photon traces for these three channels can be plotted over the duration of the experiment (see Figure 3.2 A). Background correction is then carried out followed by burst selection where photons are grouped into bursts.

A burst is defined as many photons in a small amount of time. An example burst can be seen circled in red in Figure 3.2 B. In these experiments the initial smFRET data analysis pipeline was used (see Appendix A). The threshold for burst selection in the Jupyter notebook was set to 20 photons so only bursts with 20 photons or more above background were selected. No specific correction factors for either ligand were used. Bursts were then plotted (using the R notebook) as a function of their stoichiometry and FRET efficiency. This is calculated from the amount of photons, from each of the three channels, in each burst (see Equation 1.1 and 1.2).

The CLIPf-SNAPf fusion protein resulted in two FRET efficiency populations; a smaller high FRET signal (E value of around 0.75-1.0) and a larger lower FRET signal (see Figure: 3.3). The high FRET signal is the population predicted by the close proximity of the two fluorophores. The lower FRET population is probably caused by unbound fluorophores, partial labelled protein or background photons

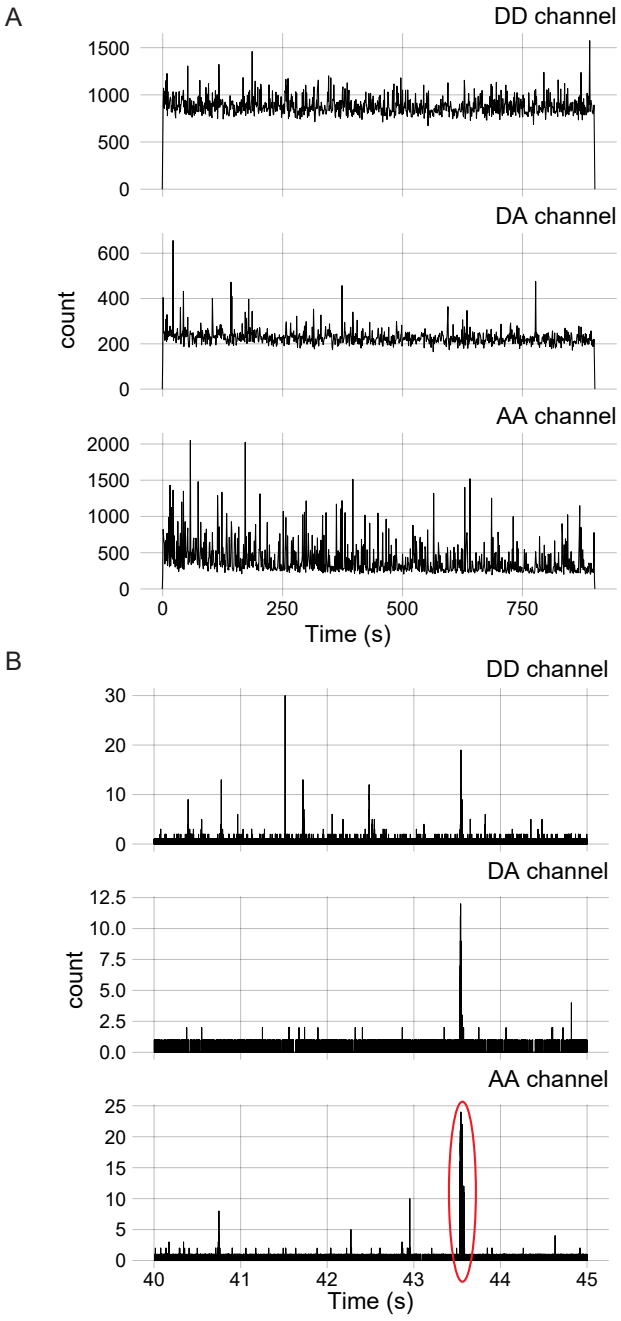

Figure 3.2: **Raw photon trace of an smFRET experiment**

Photons are split into three channels, donor-only (DD), donor-acceptor (DA) and acceptor-only (AA), based on which laser excited the photon and which APD detected the photon. A: Example photon trace over a 15 minute experiment. B: Photon trace over 5 seconds. Red circle indicates a burst of photons - many photons in a short space of time.

from the buffer, as the majority of bursts are a result of donor-only excitation and detection.

Clear smFRET events were visible showing the CLIPf- and SNAPf-tags, in combination with the CLIP-Cell TMR-Star and the SNAP-Cell 647-SiR fluorescent ligands, worked well in this assay. Also that purification of the fusion protein is not necessary and the cell lysate samples can be used directly, after dilution in the assay buffer. The expected high FRET population was seen but the large lower FRET population needs to be drastically reduced. This should be possible by a thorough optimisation of the assay, especially the labelling conditions of the two self-labelling enzymes, to ensure a good signal to noise ratio and a robust reliable assay.

3.2.2 Optimisation of the smFRET assay

3.2.2.1 Optimising the labelling conditions

It is known that the tag to which the fluorescent ligand is bound can affect its photophysical properties (Erdmann et al. 2019 and Presman et al. 2017). For example, silicon rhodamine (SiR) is brighter bound to the HaloTag then to the SNAPf-tag (Erdmann et al. 2019). Also Janelia Fluor 549 showed better photostability and less unspecific binding bound to the HaloTag then to either the SNAPf-tag or CLIPf-tag (Presman et al. 2017). Therefore, each ligand and tag combination needs to be tested and optimised separately. However, there is no systematic, standardised method to optimise the labelling approach for self-labelling enzymes in single molecule experiments. To ensure a high labelling efficiency and a low background signal I wanted to develop a thorough optimisation protocol that could be used for any tag or fluorescent ligand combination.

As the CLIPf-tag is derived from the SNAPf-tag, there is the potential of cross reactivity between the fluorescent ligands and the two tags (Gautier et al. 2008). This needs to be minimised when optimising the labelling conditions of the ligands and tags, to reduce variability and maximise the double labelled population

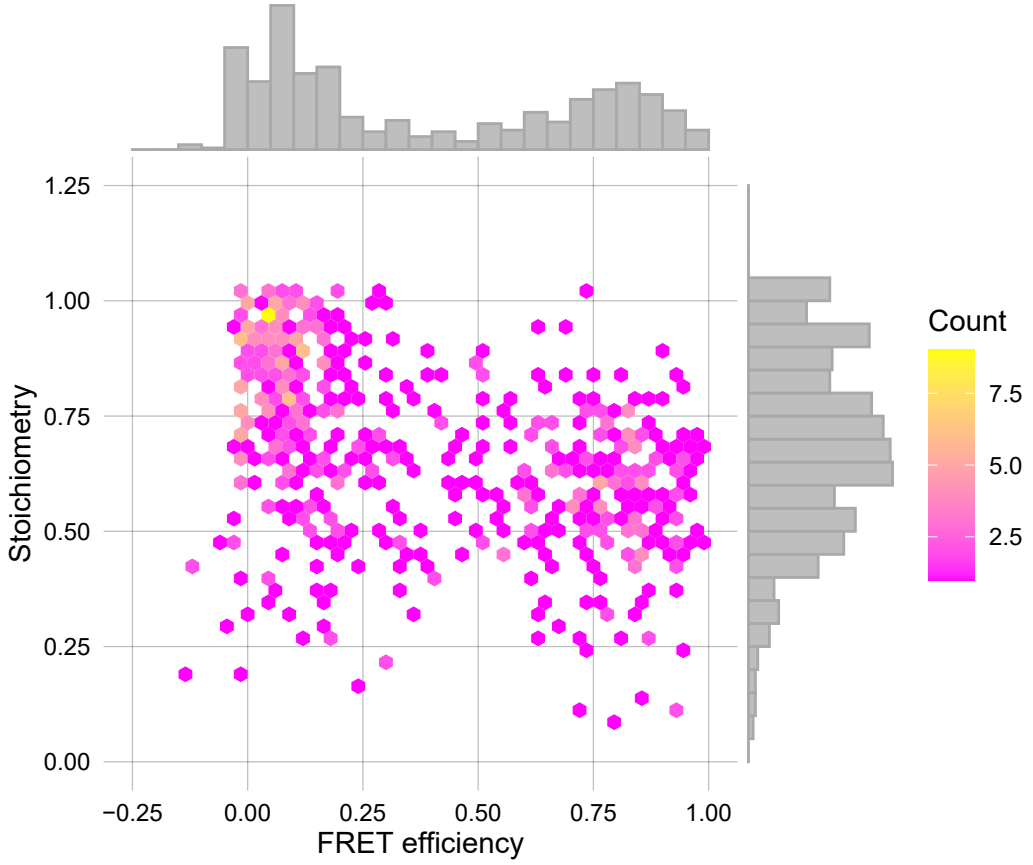

Figure 3.3: **Initial smFRET results for the CLIPf-SNAPf fusion protein**

HEK293 cells were transfected with the CLIPf-SNAPf construct and the initial smFRET protocol and analysis were followed. The stoichiometry vs FRET efficiency hex plot (with histograms) of analysed data is shown. Colour indicates number of bursts in each hexagonal pixel; n=1, bursts = 636.

that could result in a FRET signal. To remove this problem of cross reactivity when optimising the labelling efficiency, single tag constructs were used. These constructs (CLIPf-HisTag-FLAG construct and SNAPf-HisTag-FLAG) were made by deletion mutagenesis of the CLIPf-SNAPf-HisTag-FLAG construct (see Section 2.1.1.1). They were used to optimise the concentration, incubation and cross reactivity by in gel fluorescence. The washout conditions of unbound fluorescent ligands were optimised, using untransfected cells, by flow cytometry.

Washout conditions

Contamination from free unbound ligand being in the cell lysate could be a major contributor to the low FRET population seen in the initial smFRET results (see Figure 3.3). Removal of unbound ligand is facilitated by a washout step; ligand media on cells is removed and replaced with fresh media. By increasing the duration of the wash incubation step, the unbound ligand would have more time to diffuse out of the cells and can then be easily removed. Optimisation of this step was done by flow cytometry as it allowed a much higher throughput approach than smFRET.

To optimise the washout, each fluorescent ligand was added separately to untransfected HEK293 cells and left for 1 hour to incubate. Then two quick washes in media were performed before addition of 3 ml of fresh media testing different wash incubation times; quick wash only, 1 hour, 2 hours, 3 hours and 4 hours. Four hours was deemed the maximum wash incubation period to allow time for the onward smFRET experiments. Cells were lifted with trypsin, washed and resuspended in PBS. Their fluorescence intensity was then measured by flow cytometry.

The normalised density of the transformed fluorescence intensity for each sample was plotted and compared to the negative control (see Figure 3.4). This was a sample where no fluorescent ligand was added to the cells so any fluorescence would be due to autofluorescence the cells. As they were untransfected for all samples, a complete washout should return the fluorescence to that of the negative control. The median fluorescence intensity for each sample was taken to compare all biological repeats.

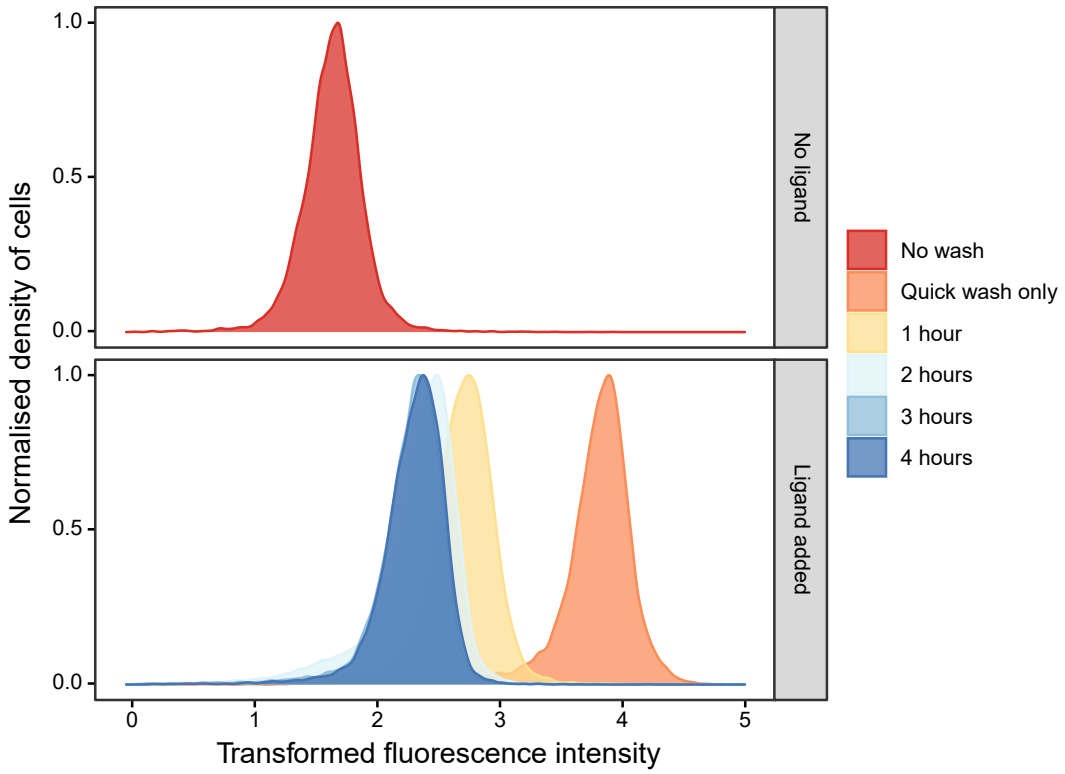

Figure 3.4: **Example flow cytometry data for optimising washout steps**

SNAP-Cell 647-SiR was added for one hour to untransfected HEK293 cells. After a quick wash (two replacements of media) different wash incubation times were tested; quick wash only, 1 hour, 2 hours, 3 hours and 4 hours. The normalised fluorescence density for each sample was plotted including the negative control of untransfected cells where no ligand was added.

For both CLIP-Cell TMR-Star and SNAP-Cell 647-SiR, a washout procedure of two quick washes followed by a 2 hour wash incubation was chosen. This was a compromise between the duration and removing the most ligand, as there was a negligible decrease with longer wash procedures (see Figure 3.5). The experimental data suggests the quick washes are able to remove any excess ligand present in the media, whereas the longer wash incubation is needed to remove ligand that has accumulated non-specifically inside the cell. A longer wash incubation gives more time for the ligand to diffuse out of the cell into the media, to be removed.

Incubation conditions

Partially labelled fusion proteins, from inefficient dual labelling, will lead to large donor-only or acceptor-only populations in the smFRET results. Therefore both the concentration of the ligand added, and the time the ligand is left to incubate with the cells, needs to be optimised.

To determine the optimal incubation time of the ligands, HEK293 cells were transfected with either the CLIPf-HisTag-FLAG construct or the SNAPf-HisTag-FLAG construct. The corresponding fluorescent ligand was added and left to incubate for; 15 minutes, 30 minutes, 1 hour, 2 hours or overnight. To control for differences in the transfection efficiency between conditions, a 10 cm dish was transfected with the construct and the cells were then split into a 6-well plate the next day. The concentration of the ligand was kept consistent at 0.25 µM for each sample, and the previously optimised washout conditions (two quick washes followed by a two hour washout incubation) were used. Unspecific binding of the fluorescent ligand was checked by adding the fluorescent ligand to untransfected cells. Another negative control of transfected cells where no ligand was added was also used. To quantify labelling, cells were lysed and the samples run on a 15 % SDS-PAGE polyacrylamide gel before in gel fluorescence analysis (see Sections 2.2.6 and 2.5.2.2). Gels were imaged on the Licor Odyssey Fc Imager in both the 600 nm and 700 nm channel. Both channels were imaged, even though the fluorophores were tested

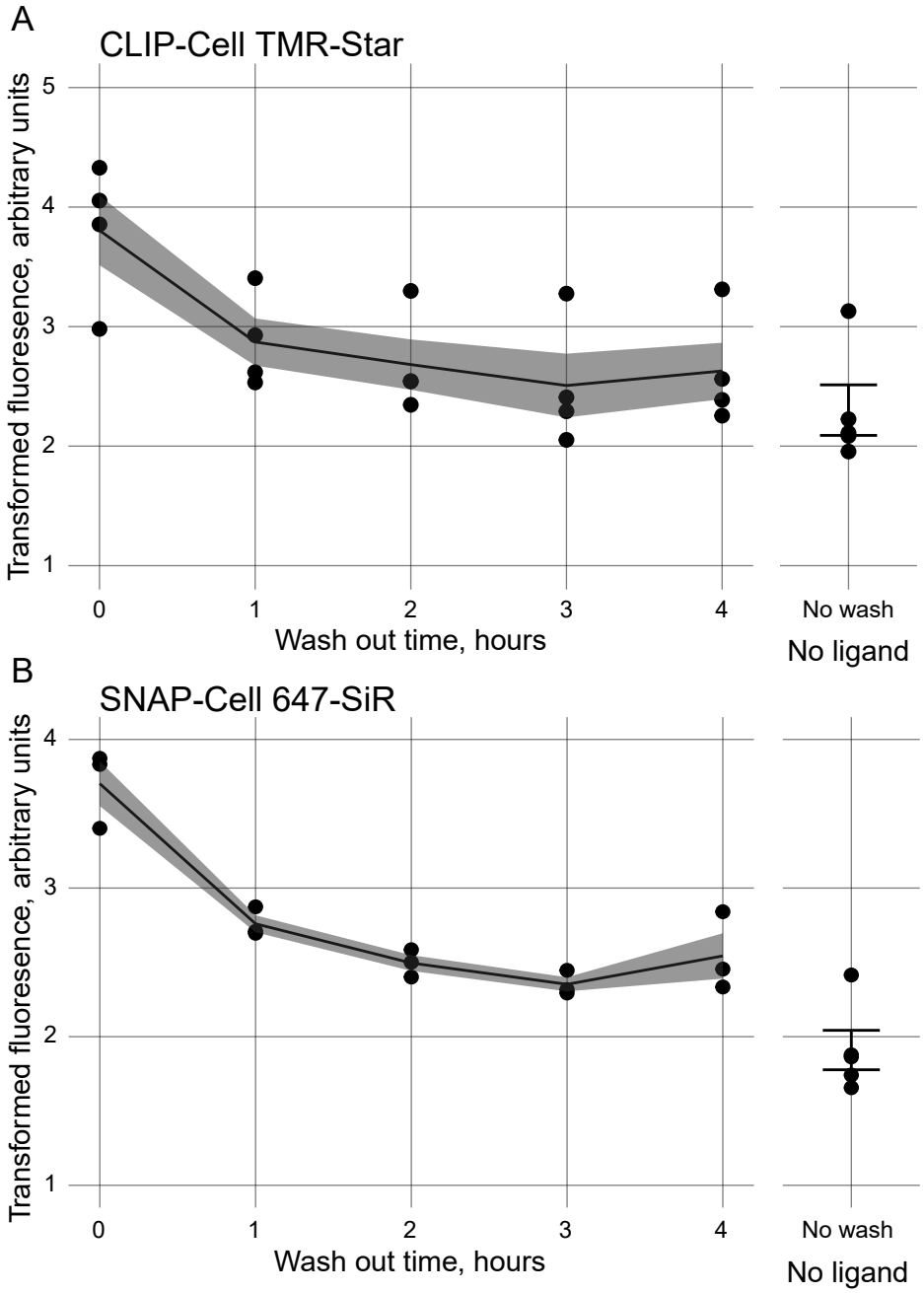

Figure 3.5: **Optimisation of the CLIP and SNAP washout conditions**

Median fluorescence for each biological repeat is plotted as circles. The line is the median of all repeats ± standard error of the mean. A: Washout optimisation of CLIP-Cell TMR-Star. B: Washout optimisation of SNAP-Cell 647-SiR. n was at least 3 for each condition tested.

separately, to check for bleed through between the channels and to fully image the standard protein ladder.

The fluorescence intensity for sample was recorded and each biological repeat was normalised to the lowest incubation period; 15 minutes. No signal was observed in any of the negative control samples. This indicates there was little unspecific binding of the fluorescent ligands within the cells.

Overnight incubation for CLIP-Cell TMR-Star and 2 hour incubation for SNAP-Cell 647-SiR resulted in the maximum normalised fluorescence signal (see Figure 3.6). SNAP-Cell 647-SiR seems to label the SNAPf-tag faster than CLIP-Cell TMR-Star labels the CLIPf-tag as saturation of labelling occurred earlier. Longer incubation resulted in a decrease in labelling. Therefore CLIP-Cell TMR-Star will be left overnight to incubate on the cells, whereas SNAP-Cell 647-SiR will only be left for two hours.

Concentration conditions

Concentration of the fluorescent ligands was optimised in a similar way to incubation conditions, with HEK293 cells being transfected with either the CLIPf-HisTag-FLAG construct or the SNAPf-HisTag-FLAG construct. Again the previously optimised washout conditions were used, and the same negative controls. Transfection efficiency of the cells was also controlled in the same way. For these experiments the previously optimised ligand incubation conditions were used; 2 hours for SNAP-Cell 647-SiR and overnight for CLIP-Cell TMR-Star. The different concentrations tested were: 0.1 µM, 0.25 µM, 0.5 µM and 1 µM. SNAP-Cell 647-SiR was also tested at 2 µM as saturation at 1 µM was not seen in the first experiment. Cell lysate samples were again run on 15 % SDS-PAGE gels before in gel fluorescence analysis.

Data collection and analysis was done in the same way as for optimising the incubation conditions, with each biological repeat normalised to the 0.1 µM condition. Negative controls all resulted in no signal in the in gel fluorescence analysis again indicating little unspecific binding of the fluorescent ligands within the cells.

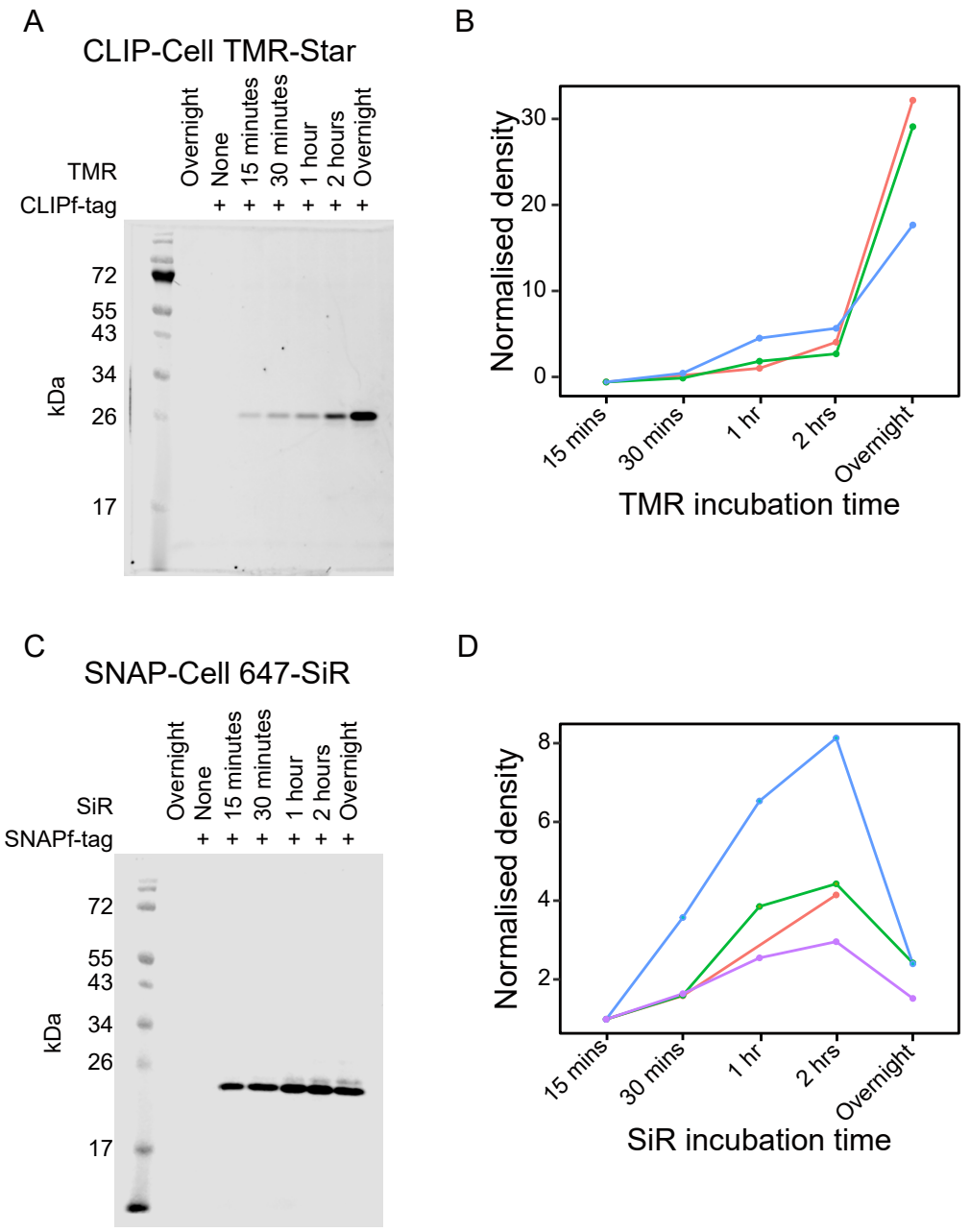

Figure 3.6: **Optimisation of the CLIP and SNAP ligand incubation time**

Each fluorescent ligand was added to HEK293 cells transfected with either CLIPf-HisTag-FLAG or SNAPf-HisTag-FLAG for: 15 mins, 30 mins, 1 hour, 2 hours and overnight. Concentration was kept constant at 0.25 µM. Cells were lysed and analysed by in gel fluorescence. Colours indicate different biological repeats. Each repeat data set was normalised to the 15 minute condition. A: Example in gel for CLIP-Cell TMR-Star. B: Graph of normalised fluorescence integrated density for incubation of CLIP-Cell TMR-Star (n=3). C: Example in gel for SNAP-Cell 647-SiR. D: Graph of normalised fluorescence integrated density for incubation of SNAP-Cell 647-SiR (n=4).

A concentration of 0.5 µM was decided as the best concentration for both ligands (see Figure 3.7). For CLIP-Cell TMR-Star this was due to the ligand reaching saturation as no significant increase was seen with 1 µM. For SNAP-Cell 647-SiR, the 1 µM and 2 µM conditions did show an increase in fluorescent signal, however this was not large enough to overcome the increased cost of using these concentrations going forward.

Cross reactivity

The CLIPf-tag is derived from the SNAPf-tag, and although they have previously been used orthogonally (Gautier et al. 2008), the possibility of cross reactivity between the two tags and their fluorescent ligands has been overlooked and not investigated. To test cross reactivity of the CLIP-Cell TMR-Star ligand for the SNAPf-tag, and visa versa, HEK293 cells were transfected with either the CLIPf-HisTag-FLAG construct or the SNAPf-HisTag-FLAG construct. Again transfection efficiency was controlled by transfecting a 10 cm dish and splitting the cells into a 6-well plate the next day. Both fluorescent ligands were then added to the transfected cells. The previously optimised washout condition were used for each ligand, but a ligand concentration of 0.25 µM was used as concentration had not yet been optimised at this point. Both the two hour and overnight ligand incubation periods were tested, as these were previously found to be the optimal incubation conditions for the SNAP-Cell 647-SiR and the CLIP-Cell TMR-Star ligands respectively. After the washout incubation, cells were lysed and the samples run on 15 % SDS-PAGE gels. In gel fluorescence analysis was again used as no bleed through between the two channels was seen when optimising the incubation and concentration conditions. CLIP-Cell TMR-Star could be seen in the 600 nm channel and SNAP-Cell 647-SiR in the 700 nm channel.

Data was collected and analysed as previously described for the optimisation of the incubation conditions. Each biological repeat contained samples from: CLIPf-tag transfected cells treated with both ligands for two hours or overnight and SNAPf-

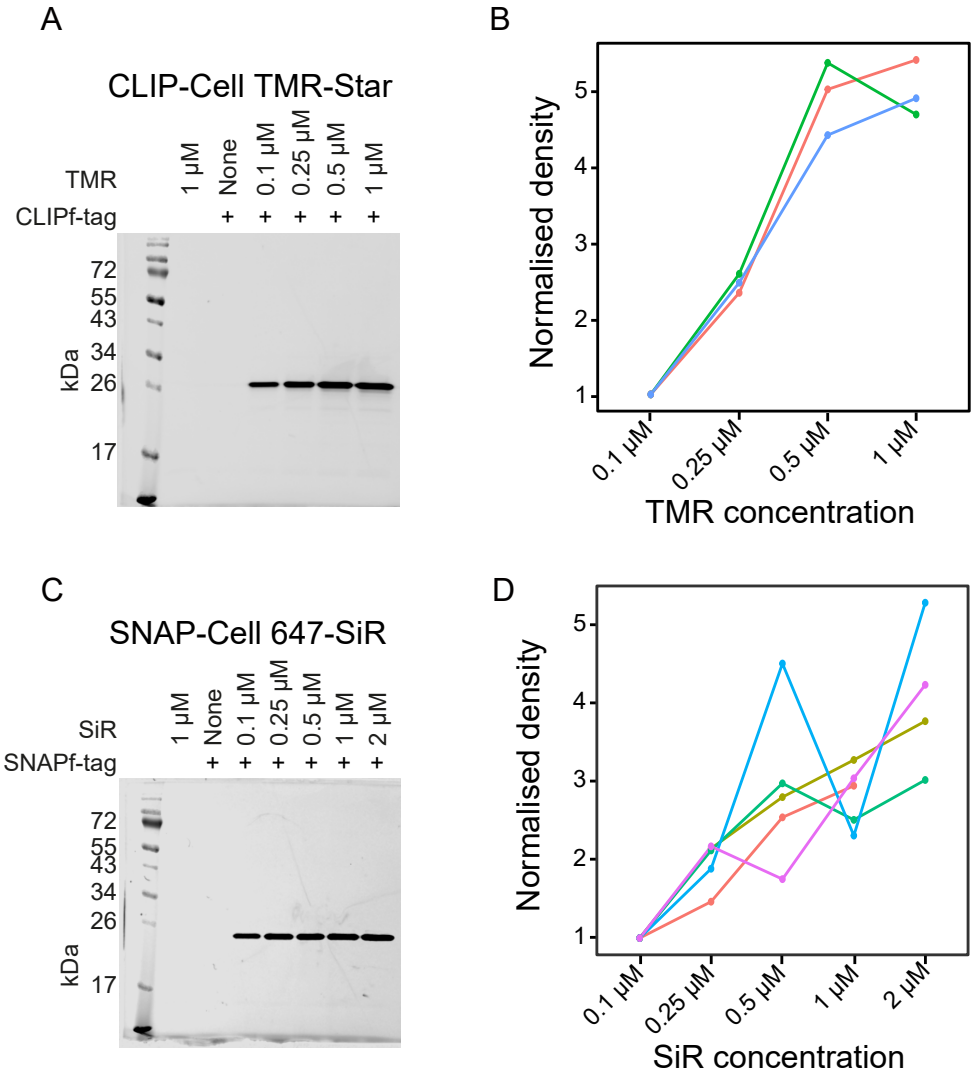

Figure 3.7: **Optimisation of the CLIP and SNAP ligand concentration**

Each fluorescent ligand was added to HEK293 cells transfected with either CLIPf-HisTag-FLAG or SNAPf-HisTag-FLAG at a range of concentrations. Cells were lysed and analysed by in gel fluorescence. Colours indicate different biological repeats. Each repeat data set was normalised to the 0.1 µM condition. A: Example in gel for CLIP-Cell TMR-Star. B: Graph of normalised fluorescence integrated density for concentration of CLIP-Cell TMR-Star. C: Example in gel for SNAP-Cell 647-SiR. D: Graph of normalised fluorescence integrated density for concentration of SNAP-Cell 647-SiR.

tag transfected cells treated with both ligands for the same periods. These repeats were was then split into four groups depending on the time the ligands were left to incubate on the cells and which channel the signal was recorded in. This gave: 600 nm channel with overnight incubation, 700 nm channel with overnight incubation, 600 nm channel with 2 hour incubation and 700 nm channel with 2 hour incubation. Each group had a pair of results; one for the signal due to the correct ligand and the other due to the incorrectly bound ligand. The cross reactivity signal was normalised to the correct signal to determine the relative extent of cross reactive labelling.

For both ligands, cross reactive labelling was minimal, but greater cross reactivity was seen for the SNAP-Cell 647-SiR ligand then for the CLIP-Cell TMR-Star (see Figure 3.8). At the previously optimised incubation times for each ligand, overnight for CLIP-Cell TMR-Star (600 nm channel) and two hours for SNAP-Cell 647-SiR (700 nm channel), cross reactivity was below 2 % ans so is unlikely to have a major impact in the smFRET assay.

3.2.2.2 Optimising the buffers for smFRET

Background photons that are recorded in the smFRET experiment, are any photons that have not originated from the fusion protein being investigated in the experiment. This could be due to either unbound ligand, the cell lysate or components in the buffer. The optimisation of the washout conditions for both ligands should result in very little unbound ligand being present, but the buffer components and cell lysate still need to be tested to ensure they do not result in a high background signal.

The initial smFRET assay buffer (see Table 2.7) was based on a DNA smFRET buffer and therefore not specific to kinesin-1. A new assay buffer, the kinesin-1 smFRET assay buffer, based on the conditions used in single molecule walking assays, will work better for kinesin-1 constructs. This buffer can be seen in Table 2.7. Bovine serum albumin (BSA) and adenosine triphosphate (ATP) were also

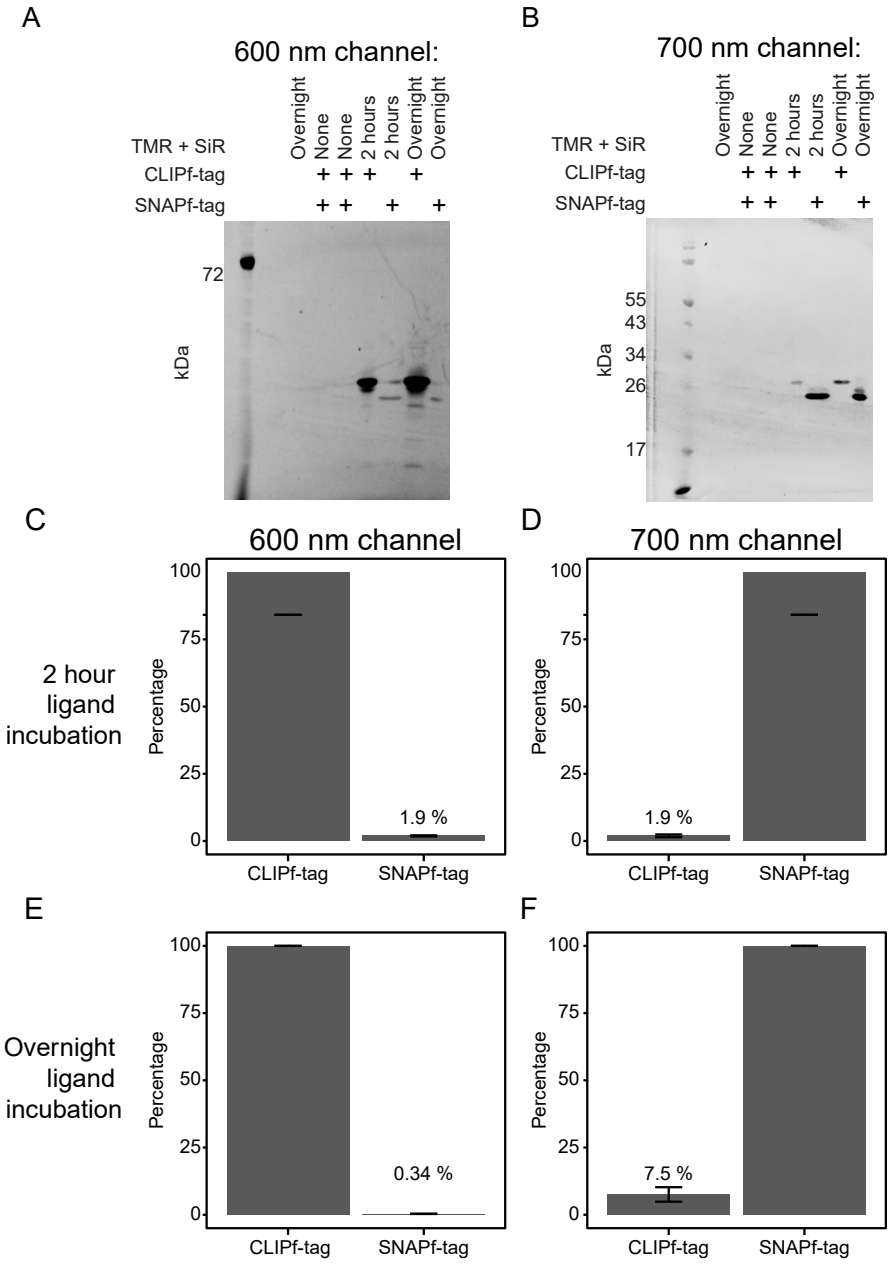

Figure 3.8: **Cross reactivity of SNAP and CLIP fluorescent ligands** The SNAP and CLIP fluorescent ligands were added to HEK293 cells transfected with either CLIPf-HisTag-FLAG or SNAPf-HisTag-FLAG at a concentration 0.25 µM for either 2 hours or overnight (around 16 hours). Cells were lysed and analysed by in gel fluorescence. Median fluorescence from independent experiments were normalised as a percentage of labelling of the correct tag and ligand combination. Error bars are ± standard error of the mean; n=3. A: Example in gel for cross reactivity imaged in the 600 nm channel. B: Example in gel for cross reactivity imaged in the 700 nm channel. C: Cross reactivity of CLIP-Cell TMR-Star after 2 hour incubation. D: Cross reactivity of SNAP-Cell 647-SiR after 2 hour incubation. E: Cross reactivity of CLIP-Cell TMR-Star after overnight incubation. F: Cross reactivity of SNAP-Cell 647-SiR after overnight incubation.

added to the lysis buffer to try and maintain correctly folded kinesin-1 proteins as samples are kept in lysis buffer (on ice) until used (see Table 2.5 - smFRET lysis buffer).

To ensure minimal background from this new assay buffer, each of the components tested separately on the smFRET instrument, and the amount of photons excited in two minutes recorded. Photon rate was then determined by dividing the total photon count by the duration of the experiment in seconds (see Appendices - smFRET background analysis). A high photon rate was considered as a rate above 2000 photons per second. This would result in too much background noise to conduct the smFRET experiments.

The only component to result in a high unusable photon rate was BSA (see Figure 3.9.A). BSA itself is not fluorescent at these wavelengths (515 nm and 613 nm) so therefore the fluorescence is likely due to a contamination of the BSA during the manufacturing process. For that reason purer BSA was purchased which did result in a reduced photon rate, but it was not enough to bring the rate below the 2000 photons per second threshold. Subsequently I developed a procedure to photobleach the purer BSA with a custom built setup (see Section 2.3.2). If it was done for long enough, the BSA signal was brought below the threshold for background photon rate (see Figure 3.9.B).

To test if any background was originating from the cell lysate, double distilled water and untransfected cell lysate (diluted to similar concentrations used in the assay) were run separately on the smFRET instrument. The amount of photons excited over three minutes was recorded. The diltued cell lysate of untransfected cells did not result in a high photon rate, meaning that protein purification isn't necessary in order to eliminate background contamination from the cells themselves (see Figure 3.9.C).

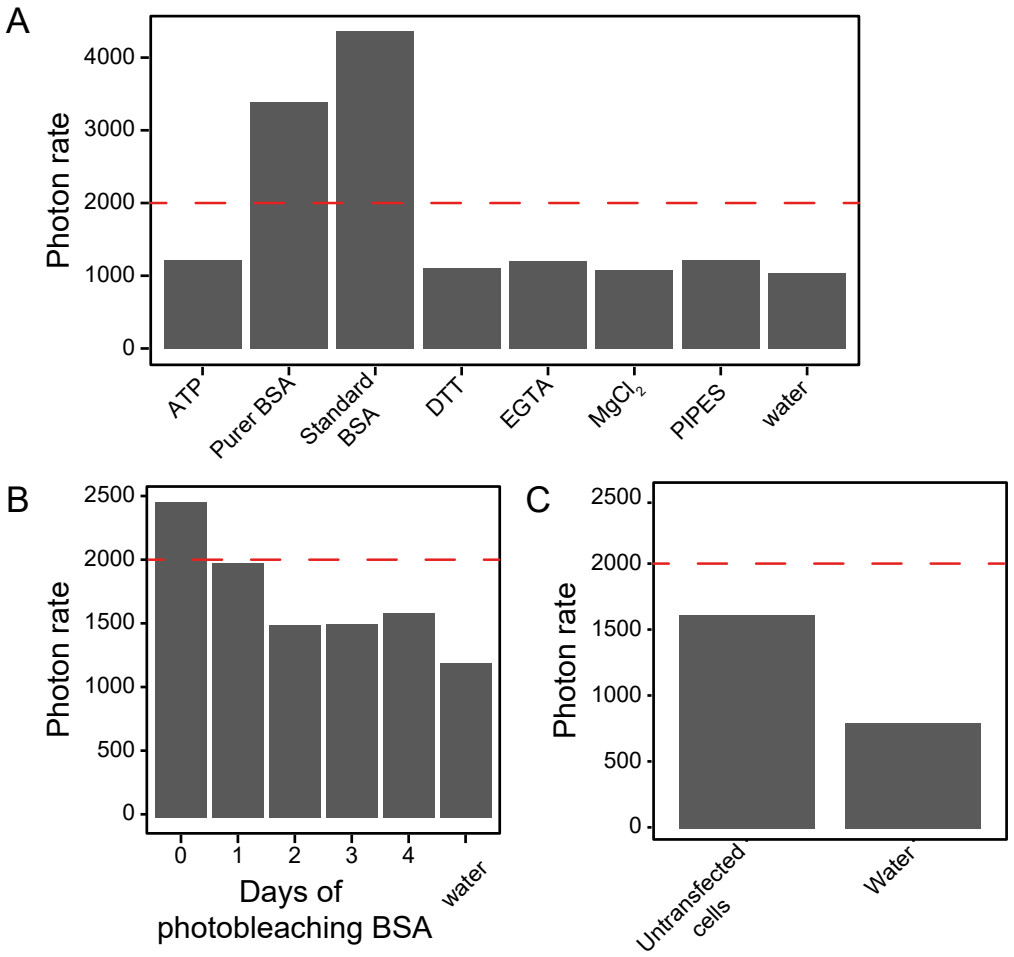

Figure 3.9: **Photon rates of isolated assay components**

Red dotted line marks the threshold for photon rate above which background photons begin to interfere with burst analysis. A: New kinesin-1 smFRET assay buffer components are all below the threshold photon rate, apart from BSA which is too fluorescent to use as is. B: Photobleaching BSA for multiple days reduces the photon rate to below the threshold photon rate. C: Cell lysate (dilution in assay buffer is equivalent to those done to result in single molecule measurements on the fusion proteins) does not add considerably to the background when compared to water alone.

3.2.2.3 Determining correction factors

Specific correction factors for each fluorescent dye are needed for accurate smFRET measurements. Spectral overlap between the emission of the donor fluorophore and the excitation of the acceptor fluorophore is essential for smFRET, but the amount of overlap is specific to the pair of fluorescent dyes used and the experimental setup. Due to alternating-laser excitation (ALEX) these correction factors can be easily determined, as the laser resulting in excitation is recorded when each photon is detected.

There are four smFRET correction factors that need to be determined: alpha, beta, gamma and delta. Each one describes different components of the ligand excitation and emission in the experimental system. Gamma and beta describe how well the fluorophores are detected and excited by the experimental setup, respectively. These correction factors are only needed for calculating accurate distances from the FRET efficiencies so will not be determined here.

The alpha correction factor, also known as the leakage factor, results from the tail of the donor emission spectrum. This tail means that photons emitted by the donor fluorophore can be mistakenly detected by the acceptor APD. Therefore, the alpha correction factor describes the amount of photons, as a result of donor fluorophore emission, that are detected by the acceptor APD, compared to the donor APD. Alpha is calculated from the mean FRET efficiency of the donor only population (E_D) after background correction (see Equation 3.1).

$$\alpha = \frac{E_D}{1 - E_D} \tag{3.1}$$

The beginning tail of acceptor absorption spectrum also needs to be considered. This is done by the direct excitation, or delta correction factor, which describes the emission of the acceptor, relative to the donor, under excitation by the donor laser.

Delta is calculated from the mean stoichiometry of the acceptor only population (S_A) after background correction (see Equation 3.2).

$$\delta = \frac{S_A}{1 - S_A} \tag{3.2}$$

To calculate the correction factors, HEK293 cells were transfected with the CLIPf-SNAPf-HisTag-FLAG construct and labelled with only one of the fluorophores (CLIP-Cell TMR-Star or SNAP-Cell 647-SiR). Optimised labelling conditions (see Table 2.4) were used along with the new smFRET assay buffer (kinesin smFRET assay buffer - see Table 2.7). The alpha correction factor was determined to be 0.03572 and the delta correction factor to be 0.03718. These factors are included in the Jupyter notebook analysis.

3.2.2.4 Optimisation of data analysis

Whilst I was optimising the labelling conditions for smFRET some of the software for acquisition was also updated by Elliot Steele in Prof. A Cadby's lab. Instead of producing a .txt file and then having to use a Jupyter notebook to convert it to a .HDF5 file for analysis, the software now produces a .h5 file directly. Therefore the Jupyter conversion notebook was no longer needed in the data analysis pipeline.

The main contributor to the time taken to analyse the data was manually creating a new Jupyter notebook for each data file and then making sure it: was named correctly, called the right .h5 file and produced a correctly named .csv file. Therefore, I automated this process using the papermill Python package. This new Python script would now: find all .h5 files in a certain folder, create and run a separate Jupyter notebook for each file and then produce a correctly named .csv file of the selected bursts (see Appendices - optimised smFRET data analysis). This cut the data analysis time from most of an afternoon to about half an hour. Also if any changes are made to the data analysis only the template Jupyter notebook has

to be changed. Rerunning the automated Python script will then process all the data using this new notebook. Change could include adding dye specific correction factors or changing the burst threshold, which is currently set to 20 photons above the background threshold.

An extra layer of background filtering was also added using a custom written R notebook. This looks at the three photon channels (donor-only (DD), acceptor-only (AA) and donor-acceptor (DA)) and removes bursts where the photon count is below a certain threshold (see Appendices - optimised smFRET data analysis). Any photons due to contamination should have a lower photon count compared to the sample. This process removes a lot of the donor-only and acceptor-only bursts which were previously present in the lower FRET population (see Figure 3.3). Overall the smFRET data analysis is now faster and more reproducible.

3.2.3 Post optimisation smFRET data

The combination of all these optimisation steps has resulted in a more reliable and robust assay. Only the optimisation of the assay buffer, switching to a more particular kinesin-1 buffer, is specific to kinesin-1. The rest of the buffer optimisation (for example the need to photo bleach BSA) and the data analysis optimisation is, not specific to the protein being studied so can be used widely with any smFRET assay. The optimisation of the labelling and correction factors is specific to the fluorescent ligands used, but the protocol developed should be easily applied to others. Therefore, any protein labelled with these two fluorophores, CLIP-Cell TMR-Star and SNAP-Cell 647-SiR, can be studied with no additional optimisation of these parameters. These optimisation steps were tested using the smFRET assay.

The washout and background buffer optimisations were tested by calculating the photon rate for: double distilled water, kinesin smFRET assay buffer, untransfected cell lysate and the lysate of untransfected cells where both fluorophores were added and washed out according to optimised conditions. Samples were placed on the

smFRET instrument and data was collected for three minutes. None of the four samples had a photon rate above 2000 photons per second and only a marginal difference was seen between the assay buffer and the untransfected cell lysate samples (see Figure 3.10). This means that the new assay buffer (using photo bleached BSA) and any unbound ligand will not contribute significantly to the background photon rate of this assay. Samples are taken and tested with every experiment to monitor background photon levels.

3.2.3.1 Optimised smFRET result for the CLIPf-SNAPf fusion protein

To see if the optimisation altered the smFRET result for the CLIPf-SNAPf fusion protein, HEK293 cells were transfected with the CLIPf-SNAPf-HisTag-FLAG construct. Both the CLIP-Cell TMR-Star and SNAP-Cell 647-SiR were added following the optimised smFRET protocol (see Section 2.3.3). The optimised data analysis pipeline was used (See Appendices - optimised smFRET data analysis).

Only one large high FRET population (see Figure 3.11), which was reproducible across biological repeats (see Figure 3.12), was observed for the CLIPf-SNAPf fusion protein this time. The lower FRET population was drastically reduced compared to the initial results (see Figure 3.3). The stoichiometry remained centred around 0.5, which is expected in smFRET assays. This indicates good equal labelling of the CLIPf- and the SNAPf-tag. This is also seen in the high burst count of the experiment, showing that most molecules are labelled with both ligands. Optimisation has resulted in an accurate highly reproducible assay which can now be used to study conformational changes in kinesin-1 autoinhibition.

3.3 Discussion

It was chosen to base the kinesin-1 conformation assay on a single molecule confocal FRET approach. This would allow the resolution of multiple conformations without

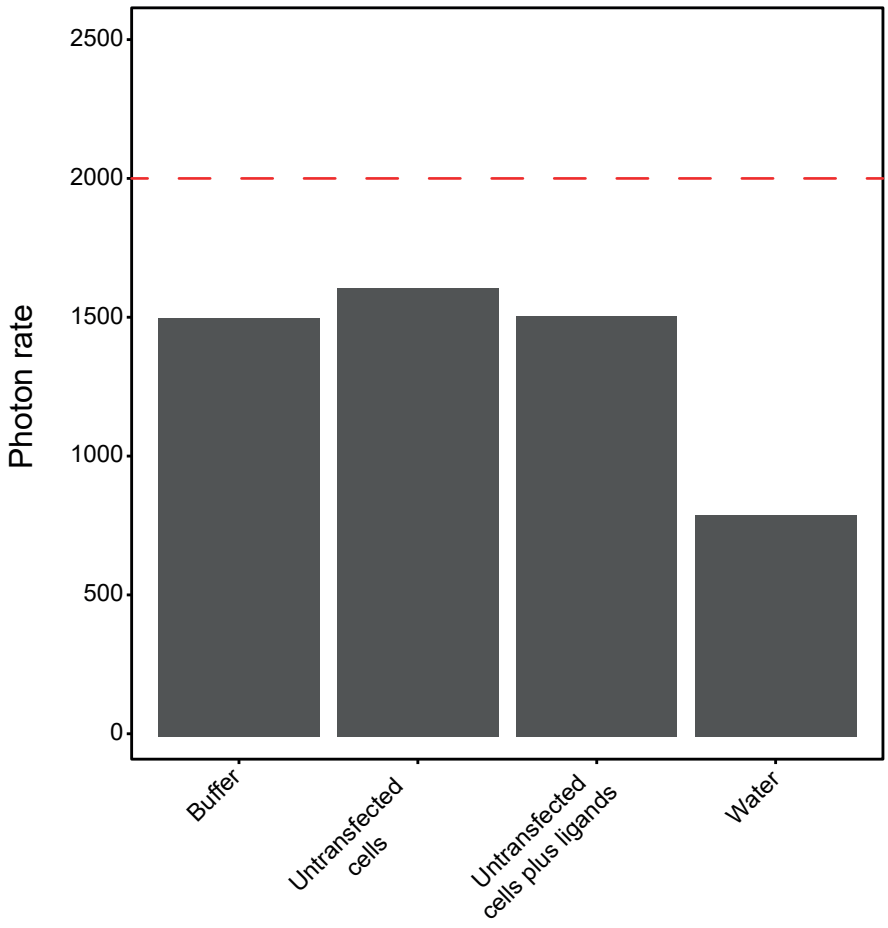

Figure 3.10: **Background photon rate after optimisation**

The total number of photons collected for each sample was divided by the 180 second acquisition time to determine a photon rate. Red dotted line marks the threshold for photon rate. Water was used as a baseline for minimum expected photon rate. This was compared against: the kinesin smFRET assay buffer (made with photo bleached BSA), untransfected HEK293 cell lysate and untransfected cell lysate where both fluorophores had been added and washed out according to the optimised labelling conditions for each. Cell lysate samples were diluted 10,000 fold which is comparable to typical assay conditions where dilution is needed for single molecule concentration.

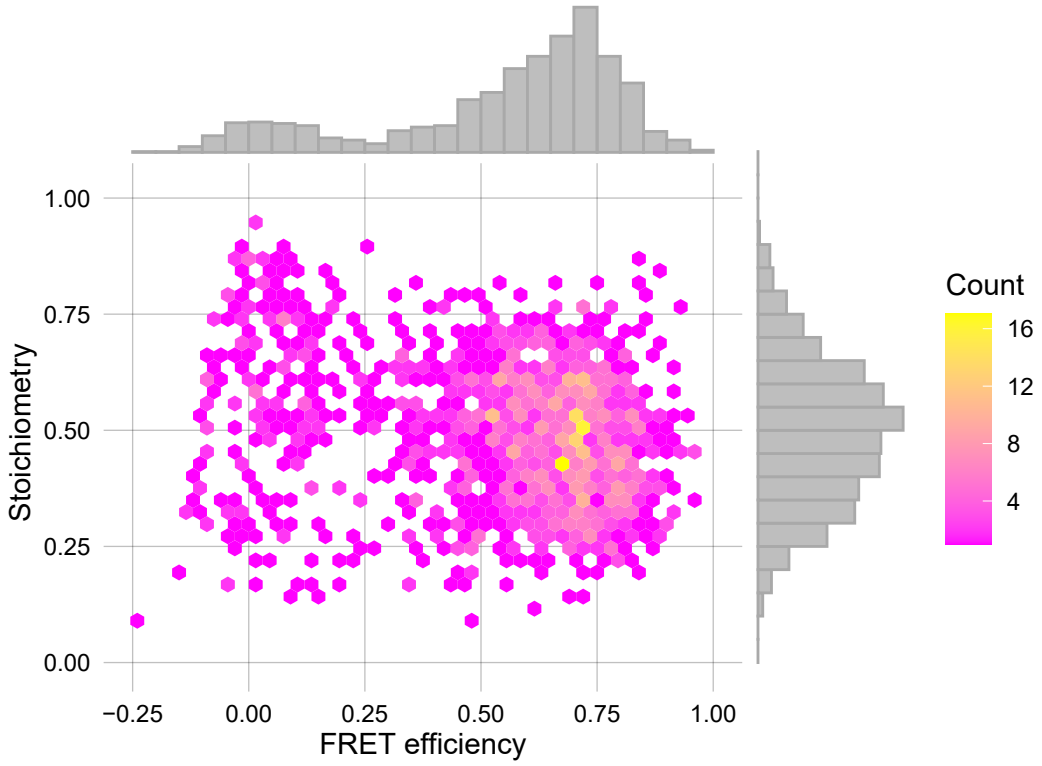

Figure 3.11: **Optimised CLIPf-SNAPf fusion protein smFRET results**

The optimised smFRET protocol was followed (see Section 2.3.3) for CLIPf-SNAPf-HisTag-FLAG transfected HEK293 cells. Stoichiometry vs FRET efficiency hex plot with histograms. Colour indicates number of bursts in each hexagonal pixel. Data is combined from n=3 biological repeats; bursts = 1680

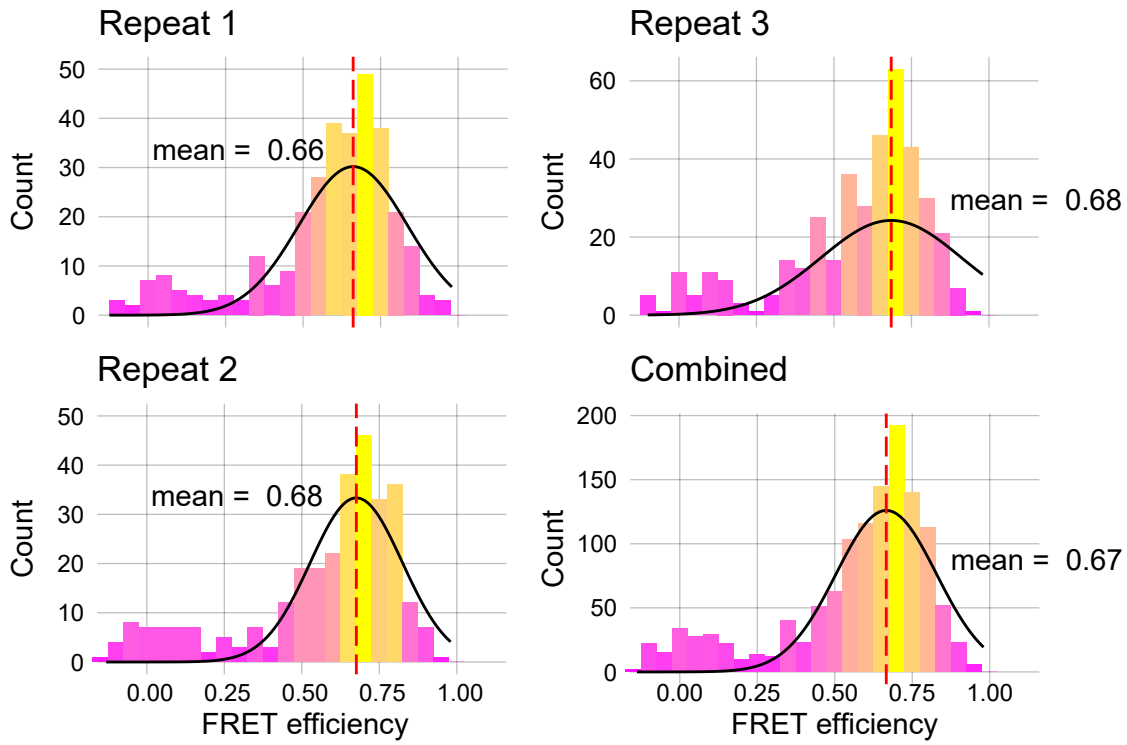

Figure 3.12: **Biological repeats of the CLIPf-SNAPf fusion protein smFRET results**

The optimised smFRET protocol was followed (see Section 2.3.3) for CLIPf-SNAPf-HisTag-FLAG transfected HEK293 cells. The FRET efficiency histogram for each biological repeat is plotted, as well as the data for the all three repeats combined. For each histogram, the data was fitted to two gaussian populations, with that describing the largest population shown. The mean of this gaussian is indicated by the dashed red line and the value written beside. The mean FRET efficiency is consistent across all biological repeats, with no repeats showing a large low FRET population.

tethering any components of the assay to a surface. Therefore, proteins can have conformational freedom as they freely diffuse across the confocal volume (Ambrose et al. 2020).

The chosen labelling strategy, CLIPf-tag and SNAPf-tag combined with CLIP-Cell TMR-Star and SNAP-Cell 647-SiR, worked well especially after optimisation of labelling conditions. The study of protein using smFRET has usually involved site specific cysteine labelling or fluorescent proteins (Agam et al. 2022 and Cai et al. 2007). Here, self-labelling enzymes were chosen as, along with fluorescent ligands, they combine the genetic encoding abilities of fluorescent proteins with the advantageous spectral and photo properties of organic dyes (Grimm, English, Choi, et al. 2016). A comprehensive optimisation approach (optimising the washout, incubation, concentration and cross reactivity of both dyes) was done to maximise the dual labelling of the fusion protein but minimise potential background signal from unbound dye. This protocol could be easily applied to other self-labelling enzymes and their fluorescent tags to optimise labelling for many different single molecule applications.

A high throughout approach was needed to test multiple conditions for the washout, concentration, incubation and cross reactivity optimisation. The approach also needed to be able to distinguish clearly between the two fluorophores. It was decided to use flow cytometry (using a BD FACS Melody cytometer) and in gel fluorescence (imaged by a Licor Odyssey Fc Imager). Both these instruments had filers and lasers at the correct wavelengths for these two fluorophores with minimal bleed thorough between the channels.

A huge difference between the initial and optimised labelling conditions was seen. They also differ greatly to those recommended by the manufacturer (see Table 3.1). On the whole, SNAP-Cell 647-SiR is more efficient at labelling the SNAPf-tag then CLIP-Cell TMR-Star is at labelling the CLIPf-tag. This can be seen in the shorter fluorophore incubation period needed as well as the slightly higher

cross reactivity results (see Figure 3.6 and 3.8). smFRET assays studying protein conformational changes usually use diluted, labelled, purified protein samples, but these cell permeable ligands meant cell lysate samples can be used.

Fluorescent ligand	Manufacturer's recommendations	Initial conditions	Optimised conditions
CLIP-Cell TMR-Star	30 minute washout 1 hour incubation 3 µM	1 hour washout Overnight incubation 0.25 µM	2 hour washout Overnight incubation 0.5 µM
SNAP-Cell 647-SiR	30 minute washout 30 minute incubation 3 µM	1 hour washout Overnight incubation 0.25 µM	2 hour washout 2 hour incubation 0.5 µM

Table 3.1: Comparison between maunfacturer's (New England Biolabs) recommendations and the pre- and post-optimisation labelling conditions used in the smFRET assay.

Purification of the fusion proteins was not necessary as diluted cell lysate samples did not contribute to the background of the assay (see Figure 3.9). This preserved experimental access to the cellular environment. This also meant new constructs could be tested quickly without potential purification protocol changes for each one. HEK293 cells were the chosen cell line as after transfection they can produce a large amount of the overexpressed protein.

A large contributor to a high background photon rate was the addition of bovine serum albumin (BSA) to the smFRET assay buffer. It is a known problem within the smFRET field and could be solved by photobleaching the BSA for a few days (see Figure 3.9). Aliquots of this photobleached BSA only lasted around one month in -80 °C storage before high background photon rates were again observed. BSA is needed in the assay buffer as it has been shown to increase the efficiency of kinesin-1

movement and is therefore used in most single molecule experiments (Block et al. 1990). BSA is not fluorescent at the wavelengths excited in this assay so the high photon rate must be due to a contaminant in the BSA. Commercially bought purer BSA did have a lower photon rate, but it was still too high to be used without photobleaching (see Figure 3.9). Purification of BSA within the lab or purifying commercially purchased BSA may prevent the need of photobleaching and be a better long term solution for using BSA within the smFRET community.

BSA is usually added to confocal smFRET buffers to help passivate the surface of the coverslip. But, as it causes such a problem with background a pre-blocking of the coverslip with BSA could eliminate the need for its incorporation in the buffer. This passivation of the coverslip could be with BSA or other blocking reagents like the polmer PEG. In this case, when designing a confocal smFRET experiment the elmination of BSA from the assay buffer isn't possible. The addition of BSA, or casein (but this showed similar fluorescent problems), to any assay buffer is standard in the study of kinesin-1 as it has been shown to increase the efficiency of kinesin-1 movement (Block et al. 1990).

After optimisation the CLIPf-SNAPf fusion protein still resulted in a high FRET population, which was consistent across biological repeats (see Figure 3.11 and 3.12). The low FRET population was drastically reduced with the remaining bursts probably due to coincidence events. These are when two labelled proteins come into confocal volume at the same time and result in a FRET signal. Coincidence events can be removed by lowering the concentration of the sample. This was not done due to the consequent reduction in bursts collected and the low amount of these coincident bursts. More bursts could be recorded with longer acquisition times, but as the samples are only viable for a few hours. If purified protein was being used this would be less of a problem as time constraints of the sample and experiment would mean there would be more time for data acquisition.

These results show that with optimisation the CLIPf- and SNAPf-tags can be

used in combination with the CLIP-Cell TMR-Star and SNAP-Cell 647-SiR ligands in smFRET experiments. Also protein purification of the labelled fusion protein is not necessary and cell lysate samples, diluted in assay buffer, can be used. The next chapter will focus on studying kinesin-1 constructs using this new optimised assay.

Chapter 4

Studying kinesin-1 conformation by smFRET

4.1 Introduction

4.1.1 Two known conformations of the kinesin-1 heavy chain

As outlined in the introduction, kinesin-1 can be autoinhibited resulting in two main conformations; an elongated active form and a folded inactive form (see Section 1.5.2). There are two main hypotheses around kinesin-1 autoinhibition. Firstly, that kinesin-1 exists predominately in a pool of inactive folded kinesin-1 waiting for the right cargo to come along (Hackney, Levitt, and Suhan 1992). This will be discussed further in this chapter. Secondly, that multiple kinesin-1 molecules are always attached to their cargo and are either active or inactive upon them (R D Vale, Schnapp, et al. 1985). This will be discussed later in Section 5.1.3. However, there is currently insufficient evidence to support either hypothesis.

Although autoinhibition was suggested by: higher ATPase activity for smaller fragments (Kuznetsov, Y. A. Vaisberg, et al. 1989 and Hackney, Levitt, and Wagner 1991), predicted hinge regions based on the DNA sequence (Yang et al. 1989 and Kosik et al. 1990) and electron microscopy (Amos 1987 and Hirokawa, Pfister, et al.

1989), the first direct evidence for two conformations of kinesin-1 came in 1992. Hackney *et. al.* found the conformation of the kinesin-1 heavy chain can be altered from a compact to an extended form by changing the ionic strength. This was identified by a change in sedimentation coefficients from 9.4 S, for the extended active conformation, to 6.5 S for the folded inactive form. Low ionic conditions resulted in a compact conformation and higher ionic conditions resulted in a change in conformation to an extended form. This is the origin of the theory around a pool of inactive kinesin-1, although the authors themselves only state that this may be a possibility. They say that the likely form of kinesin-1 to exist *in vivo* is the folded form, and that this conformation is inhibited and "may represent a soluble pool of the enzyme", but *in vivo* the transition to the unfolded conformation "may not be too great" (Hackney, Levitt, and Suhan 1992).

By N- and C-terminally tagging the kinesin-1 heavy chain, with the CLIPf- and SNAPf-tags, the conformation of the molecule could be studied by smFRET. This double-tagged kinesin construct will take around 1-2 ms to diffuse through the FRET confocal volume. This means it will be exposed to the instruments laser duty cycle around 10-20 times (see Figure 4.1). In the elongated active conformation the two fluorophores, covalently attached to the tags, will be too far apart for a FRET signal to occur; around 60 nm. Therefore a zero FRET efficiency population will be seen. However, the folded inactive conformation should result in a FRET signal as the two fluorophores are now within 3-10 nm of each other and a high FRET efficiency population should be seen.

The ability to resolve multiple conformations within a population of molecules is only possible by undertaking a single molecule approach. Ensemble measurements would not be able to solve the heterogeneous nature of the overall population, as only a average FRET efficiency is reported. Two distinct FRET efficiency populations for kinesin-1 are expected; one for the elongated and one for the folded conformation. If the inactive pool of kinesin-1 theory is correct then the majority of the smFRET

bursts should have a higher FRET efficiency and only a small zero FRET efficiency population will be seen. The distribution between these two populations should be changeable by addition of different mutations and conditions known to alter the conformation or activity of kinesin-1.

4.1.2 Altering the conformation of kinesin-1

4.1.2.1 Changing the ionic strength

The conformation of kinesin-1 can be altered by changing the ionic strength. This was first shown indirectly in 1986, just one year after the discovery of kinesin-1, but then again in 1991 (Kuznetsov and Gelfand 1986 and Urrutia et al. 1991). In both of the papers the conformation of the kinesin-1 heavy chain was not directly studied but instead the ATPase activity of the molecule. The movement of latex beads by kinesin-1 and the microtubule-activated ATPase activity was studied at increasing KCl concentrations ranging from 0 - 100 mM. They showed that with increased KCl concentration there was a decrease in the ATPase activity and also the amount of moving kinesin-1 molecules (Urrutia et al. 1991). However in 1992, and then again with the *Drosophila* kinesin-1 heavy chain in 1999, it was shown by the use of sedimentation coefficients that addition of NaCl, or KCl, transitions the kinesin-1 heavy chain conformation from a compact to a more extended conformation. A low salt concentration of around 0.2 M resulted in the folded conformation; shown by the lower (6.5 S) sedimentation coefficient. When this concentration was increased to around 1 M, a transition to an elongated conformation was inferred from an increase in the sedimentation coefficient to 9.4 S (Hackney, Levitt, and Suhan 1992 and Stock et al. 1999).

Based on these sedimentation coefficient experiments, changing the salt concentration in the smFRET assay should result in a shift in the proportions between the expected two FRET efficiency populations. Addition of NaCl to the assay buffer should reduce the higher FRET efficiency population and increase the lower FRET

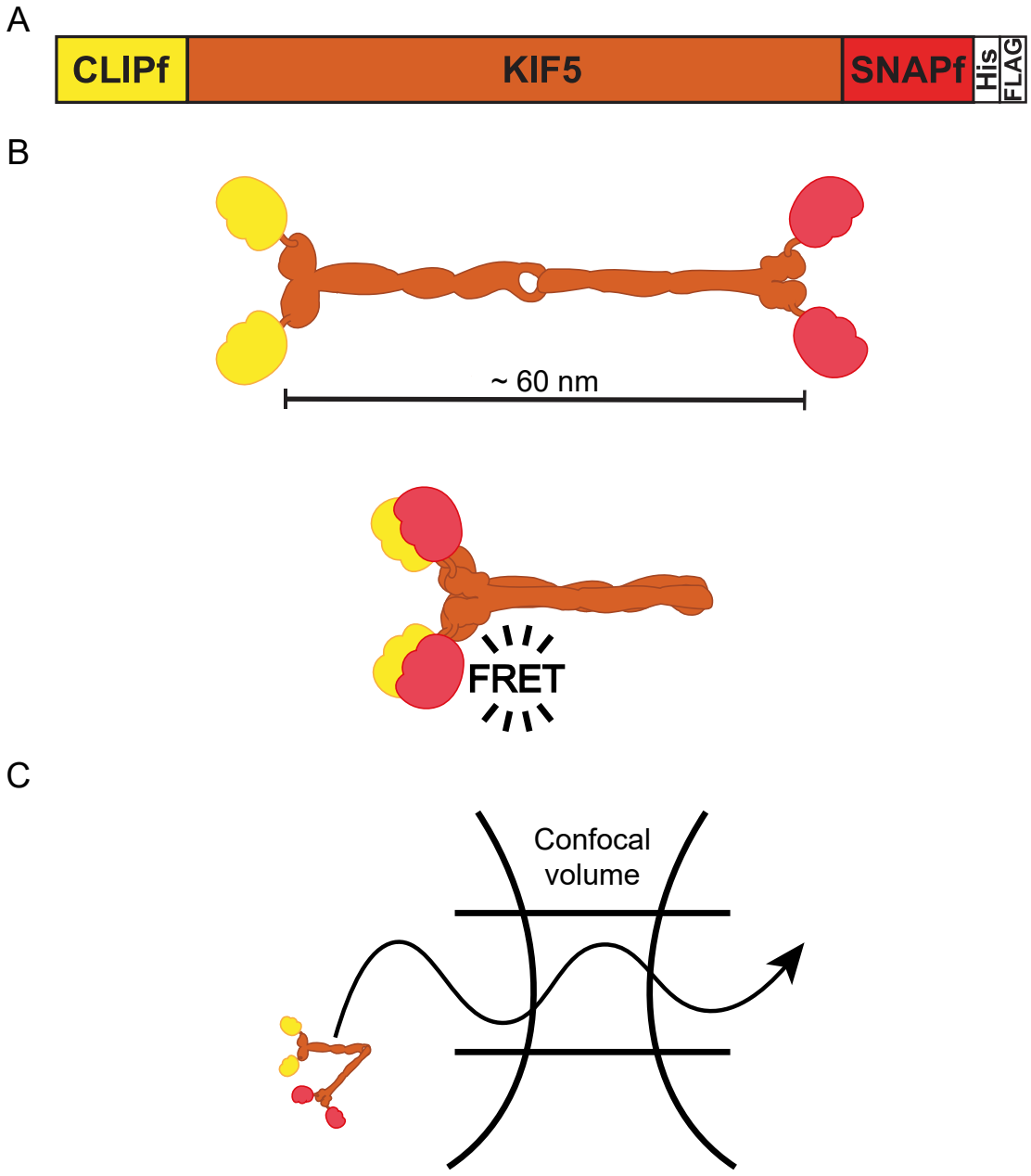

Figure 4.1: **Using smFRET to study kinesin-1 autoinhibition**

Cartoon schematics of a double tagged kinesin heavy chain. A: The double tagged kinesin construct used. The CLIPf- and SNAPf-tags bind to specific fluorophores and the polyhistidine tag and the FLAG tag can be used for purification or immuno assays. B: The two main predicted conformations of the kinesin-1 heavy chain. The two fluorophores will be too far apart for FRET to occur in the elongated active conformation but should be close enough to result in a FRET signal in the closed inactive form. C: The double tagged kinesin-1 molecule will diffuse through the confocal volume of the smFRET instrument as the two lasers cycle on and off.

efficiency population, as a result of switching from the folded to the elongated conformation.

4.1.2.2 Mutations of kinesin-1 heavy chain

Another potential way to alter the conformation of the kinesin-1 heavy chain is by mutations known to increase the ATPase activity of the molecule, or deleting regions known to be important to autoinhibition. The central hinge, sometimes known as "hinge 2", is the region where the kinesin-1 heavy chains are thought to bend in half resulting in autoinhibition. This region was first observed by electron microscopy (Amos 1987 and Hirokawa, Pfister, et al. 1989). DNA sequence analysis then saw a breakdown in the hydrophobic periodicity of the coiled coil stalk region of kinesin-1 (Yang et al. 1989). This region is broadly described as amino acids 505-610 and deletion results in an increase in ATPase activity (Friedman and R D Vale 1999). This mutant has only been shown to affect kinesin-1 activity and not autoinhibition directly.

Autoinhibition of the kinesin-1 heavy chain occurs due to an interaction between the tail region, specifically the conserved QIAK region (amino acids 917-920 in the sequence of the human KIF5A), and the motor domains (Kaan et al. 2011). Before the specifics of this interaction were discovered, evidence that the tail played a role in autoinhibition was seen in differences of sedimentation coefficients. The *Drosophila* kinesin-1 heavy chain showed a similar change in sedimentation coefficient (from 6.9 S to 5.4 S) on addition of high concentrations of salt (1 M). Mutants of *Drosophila* kinesin-1 heavy chain, where increasingly larger regions of the C-terminus were deleted, showed a shift in the transition between sedimentation coefficients to lower salt concentrations. For example, the DKH937 construct required a lower concentration of salt to result in the change of sedimentation coefficient, then the full length construct. The DKH927 construct, and any shorter constructs did not result a lower sedimentation coefficient on addition of salt, indicating that

these constructs are unable to form the compact conformation and are always in the elongated form. This shows some proportion of the tail domain of the kinesin-1 heavy chain is necessary to form the compact conformation and therefore kinesin-1 autoinhibition (Stock et al. 1999).

Mutants that are known to affect the conformation or ATPase activity of kinesin-1, should alter the proportion of FRET efficiency populations seen in the smFRET assay when compared to wildtype kinesin-1 heavy chain. Mutants which increase the ATPase activity (for example deleting "hinge 2") should increase the lower FRET efficiency population and decrease the higher. Whereas, mutants that show no change in sedimentation coefficient (tail deletions) should only result in the lower FRET efficiency population due to their inability to form the inactive folded conformation.

The aim of this chapter is to study the conformation of the kinesin-1 heavy chain by smFRET. Results can be validated by comparing mutants or changing the salt concentration, which should cause a shift in the proportion of the two predicted FRET efficiency populations.

4.2 Results

4.2.1 smFRET analysis of CLIPf-mKIF5A-SNAPf-HisTag-FLAG

Having established that the SNAPf- and the CLIPf-tags, along with the corresponding cell permeable fluorescent ligands (CLIP-Cell TMR-Star and SNAP-Cell 647-SiR), can be used for smFRET, next it needed to be determined whether the assay was sensitive enough to observe the different conformations of the kinesin-1 heavy chain. The same experimental protocol optimised in the previous chapter (Section 3.2.3.1) was used, including the use of cell lysate based samples.

The KIF5A isoform was used first as it is only endogenously expressed in neurons and so is not expressed in HEK293 cells. These cells endogenously express

KIF5B so I wanted to avoid this isoform to prevent problems with expression and formation of mixed populations between the overexpressed tagged protein and any endogenous kinesin-1. Cells were transfected with a double-tagged KIF5A construct: CLIPf-mKIF5A-SNAPf-HisTag-FLAG construct (see Section 2.1.1.2). The resulting KIF5A fusion protein had the CLIPf-tag attached to the N-terminus (motor domain) and the SNAPf-tag attached to the C-terminus (tail domain) with no amino acid linkers between the tags and KIF5A. This was to prevent flexibility of the tags resulting in a broad FRET efficiency population.

A positive FRET signal can only result from the KIF5A being in the autoinhibited closed conformation. This is because in the active open conformation the ligands would be around 60 nm apart, the length of KIF5A, and therefore too far apart for FRET to occur (see Figure 4.1). So, two distinct populations were expected; a low FRET population around E = 0, for the open unfolded active conformation of the kinesin-1 heavy chain and a higher FRET population (E >0.5) for the closed folded inactive conformation. HEK293 cells were transfected with CLIPf-mKIF5A-SNAPf and the diluted cell lysate tested on the smFRET instrument.

One large FRET efficiency population, that is distinctively different from the CLIPf-SNAPf-HisTag-FLAG positive control, was observed (see Figure 4.2). No distinct peak at zero FRET efficiency was seen, instead there is a low FRET population with a broad range of FRET efficiencies from zero to around 0.3. No distinctive higher FRET efficiency population was observed. This result was consistent across multiple biological repeats (see Figure 4.3). Also, the stoichiometry of the population was much lower then expected; at around 0.4 rather then 0.5. Before interpretation of these results can occur, investigations into these unpredicted results needs to be carried out. These will include:

- investigation into heterodimer formation between the overexpressed mKIF5A fusion protein and the endogenously expressed hKIF5B isoform (endogenous dimerisation),

- improving the stoichiometry of the smFRET results for the mKIF5A fusion protein and seeing if this effects the FRET efficiency,
- increasing the salt concentration of the assay buffer to see if this changes the FRET efficiency distribution,
- determining the FRET efficiency for mutants of the mKIF5A fusion protein which are predicted to alter the conformation or activity of the molecules.

4.2.2 Heavy chain heterodimer formation

Although the KIF5A isoform was chosen, endogenous dimerisation with the over-expressed protein could occur resulting in an unexpected smFRET result. The heavy chain isoforms exist as homodimers (Kuznetsov, E. A. Vaisberg, et al. 1988) with HEK293 cells endogenously expressing the hKIF5B isoform. It is possible that the overexpressed tagged mKIF5A construct could dimerise with the endogenous hKIF5B resulting in a heterodimer; one heavy chain being the tagged mKIF5A and the other being the endogenous hKIF5B. The smFRET results may be as a result of this heterodimer, which does not represent physiological kinesin-1 heavy chain structure. This maybe why the two predicted FRET efficiency populations were not seen.

To test this possibility, immunoprecipitation was used to identify any co-complexes of endogenous and over expressed kinesin-1 heavy chain isoforms. HEK293 cells were transfected with either the CLIPf-mKIF5A-SNAPf-HisTag-FLAG construct or the CLIPf-hKIF5B-SNAPf-HisTag-FLAG construct (see Section 2.1.1.2). This human KIF5B construct was to act as a positive control for the α-KIF5B antibody. A negative control of untransfected cells was also used. Co-immunoprecipitation using the α-FLAG rabbit antibody was then performed and the samples (input, flow through and IP for each of the three cell samples) were prepared for western blot (see Sections 2.5.1 and 2.2.8). Samples were run on two 10 % resolving polyacrylamide gels and

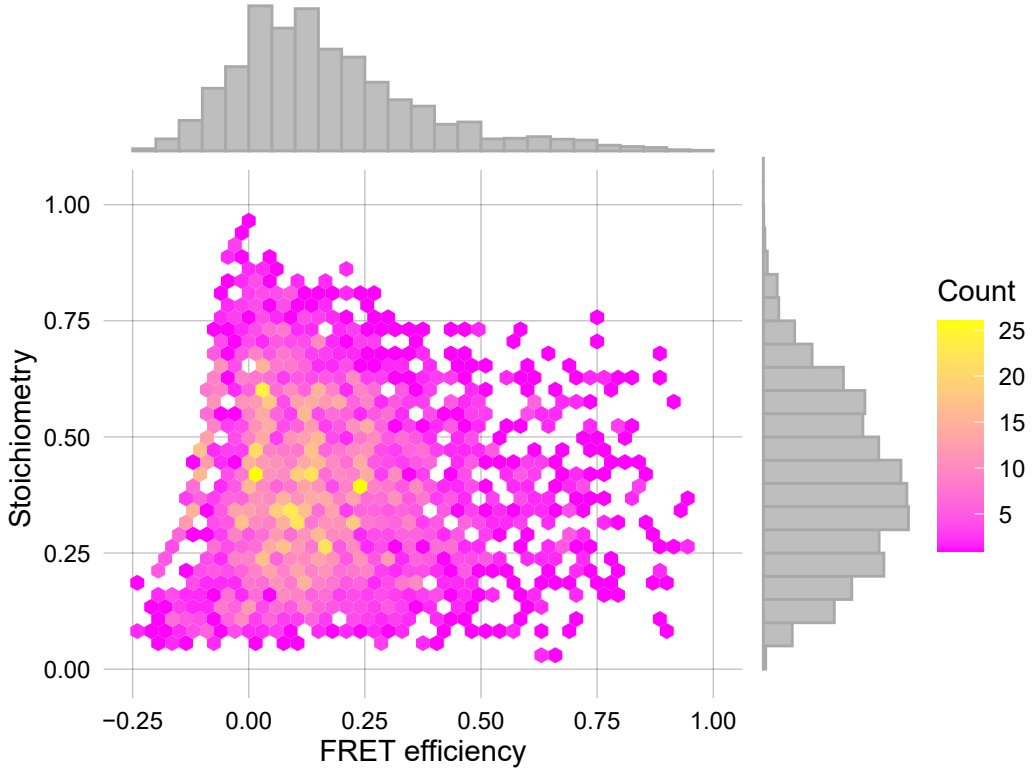

Figure 4.2: **smFRET results for CLIPf-mKIF5A-SNAPf fusion protein**

The optimised smFRET protocol was followed (see Section 2.3.3) for CLIPf-mKIF5A-SNAPf-HisTag-FLAG transfected into HEK293 cells. The stoichiometry vs FRET efficiency hex plot (with histograms) of analysed data is shown. Colour indicates number of bursts in each hexagonal pixel. Data is combined from n = 15 biological repeats; bursts = 3461

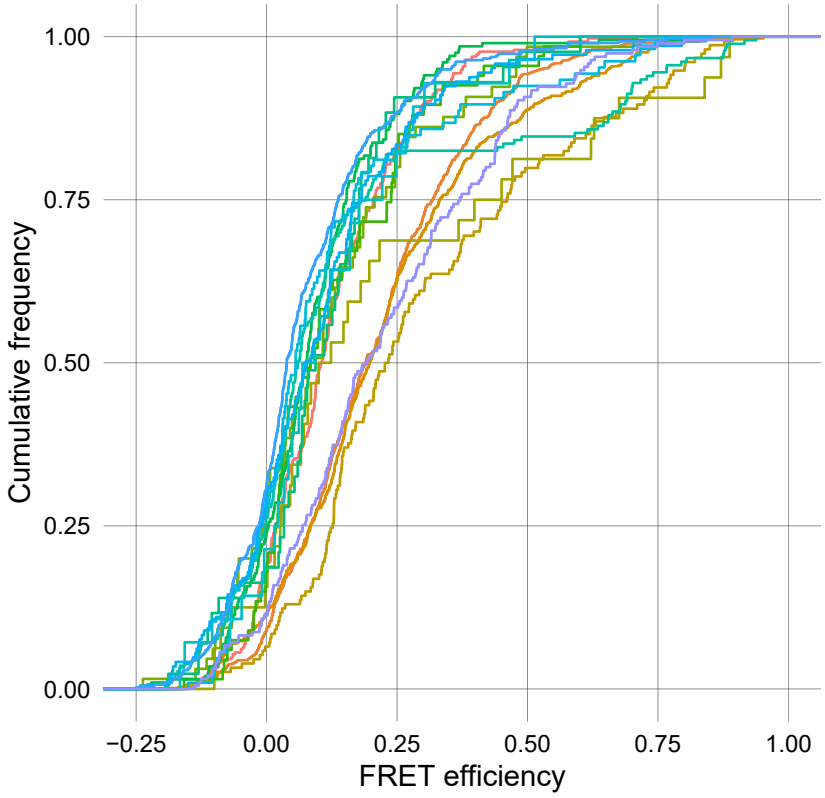

Figure 4.3: **Biological repeats of the smFRET results for the CLIPf-mKIF5A-SNAPf fusion protein**

The optimised smFRET protocol was followed (see Section 2.3.3) for CLIPf-mKIF5A-SNAPf-HisTag-FLAG transfected into HEK293 cells. Cumulative frequency plot of the FRET efficiency comparing the 15 different biological repeats which were combined in Figure 4.2.

transferred to membranes (see Sections 2.5.2.1 and 2.5.2.4). One was blotted with α-KIF5B and α-GAPDH and the other with α-FLAG and α-GAPDH (see Section 2.5.2.5).

No evidence of heavy chain heterodimer formation was seen by co-immunoprecipitation (see Figure 4.4). Probing membranes with α-FLAG showed good transfection and co-immunoprecipitation of the mKIF5A construct. The human KIF5B construct did work as a positive control for the antibody, but transfection was poor; as seen by the fainter α-FLAG band for this construct compared to that of the mKIF5A. This transfection was repeated multiple times, including using different DNA preparations, and low protein expression of the hKIF5B fusion protein was always seen. α-KIF5B antibody showed endogenous KIF5B in all three lanes of the input and flow through but none in the IP lanes. Therefore it is highly unlikely that heterodimers between the overexpressed mKIF5A fusion protein and the endogenously expressed hKIF5B isoforms are a significant population in the smFRET experiments.

4.2.3 Low stoichiometry of the CLIPf-mKIF5A-SNAPf fusion protein

In smFRET, stoichiometry is defined as the proportion of signal resulting from the donor compared to the acceptor fluorophore (see Equation 1.1). A S value of around 1 means the burst only contains signal from the donor fluorophore, whereas a S value of around 0 means the burst only contains signal from the acceptor fluorophore. Double labelled bursts are the ones of interest as they have the potential for FRET to occur; these have a stoichiometry of around 0.5. It is only possible to calculate stoichiometry due to the use of alternating-laser excitation (ALEX) in the smFRET instrument as this means the acceptor occupancy, along with the donor occupancy, is probed independently.

An unexpected difference in stoichiometry was seen between the CLIPf-SNAPf

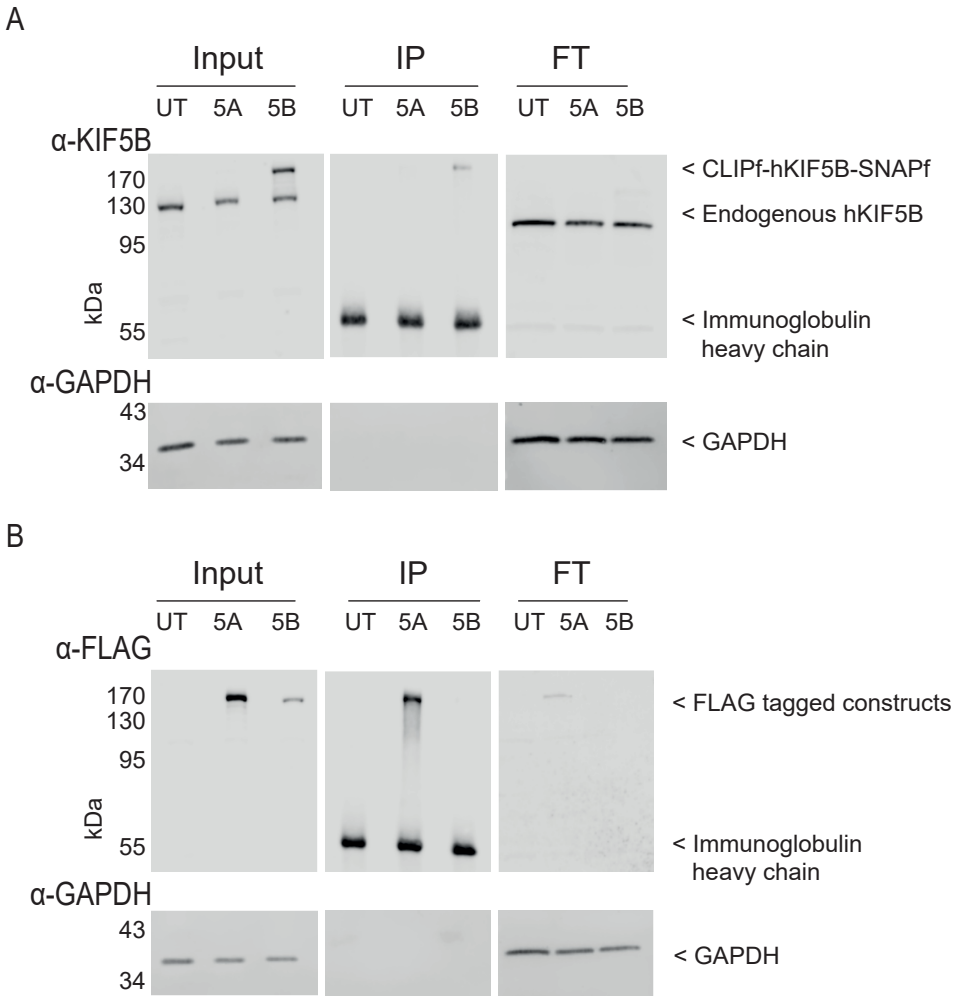

Figure 4.4: **Co-immunoprecipitation of heavy chain constructs to check for heterodimers**

HEK293 cells were transfected with CLIPf-mKIF5A-SNAPf-HisTag-FLAG construct (5A) or the CLIPf-hKIF5B-SNAPf-HisTag-FLAG (5B). UT is the untransfected cell negative control. The overexpressed proteins were immunoprecipitated with α-FLAG antibody. Samples were separated by SDS-PAGE and probed by western blot using A: α-KIF5B rabbit and α-GAPDH mouse or, B: α-FLAG rabbit and α-GAPDH mouse; n =3.

and the CLIPf-mKIF5A-SNAPf fusion proteins (see Figure 4.5). The CLIPf-SNAPf fusion protein has the desired stoichiometry distribution centred around 0.5 (see Figure 3.11). This means that most of the bursts result from double labelled protein and any donor-only or acceptor-only populations have been removed in the data analysis. However the CLIPf-mKIF5A-SNAPf fusion protein has a shifted stoichiometry centred around 0.4, with the most bursts having a stoichiometry between 0.25 and 0.5 (see Figure 4.2). This is above the stoichiometry for acceptor only labelled bursts (0-0.25), but means that the majority of observed bursts have a greater contribution from the acceptor fluorophore then the donor.

This difference in stoichiometry may to be due to the labelling efficiencies of the CLIPf-tag and SNAPf-tag. A protein tagged at either end with a self-labelling enzyme that exists as a dimer, gives rise to nine possible combinations that could result in FRET (see Figure 4.6). The majority of these possibilities result in bursts where S is around 0.5 but some options result in a lower or higher stoichiometry. These possibilities are removed when studying a monomer as only one labelling possibility results in FRET and that gives a stoichiometry of 0.5. It was seen previously that SNAP-Cell 647-SiR needed a considerably shorter ligand incubation period to maximise binding to the SNAPf-tag compared to CLIP-Cell TMR-Star and the CLIPf-tag. This suggests that the labelling efficiency of the SNAPf-tag is greater then the CLIPf-tag in these conditions using these ligands. Therefore it is likely that a mix of labelled populations exists for the CLIPf-mKIF5A-SNAPf fusion protein. This leads to a bigger range of possible stoichiometries with a tendency towards more labelled SNAPf-tags then CLIPf-tags and therefore lower stoichiometries. This tendency towards lower stoichiometries also helps confirm the previous co-immunoprecipitation results (see Section 4.2.2) as if the tagged protein was dimerising with the endogenous protein then the stoichiometry would be 0.5.

Instead of transfecting a double tagged construct, co-transfection of two single tagged constructs was tried to fix the lower stoichiometry. The resultant dimer

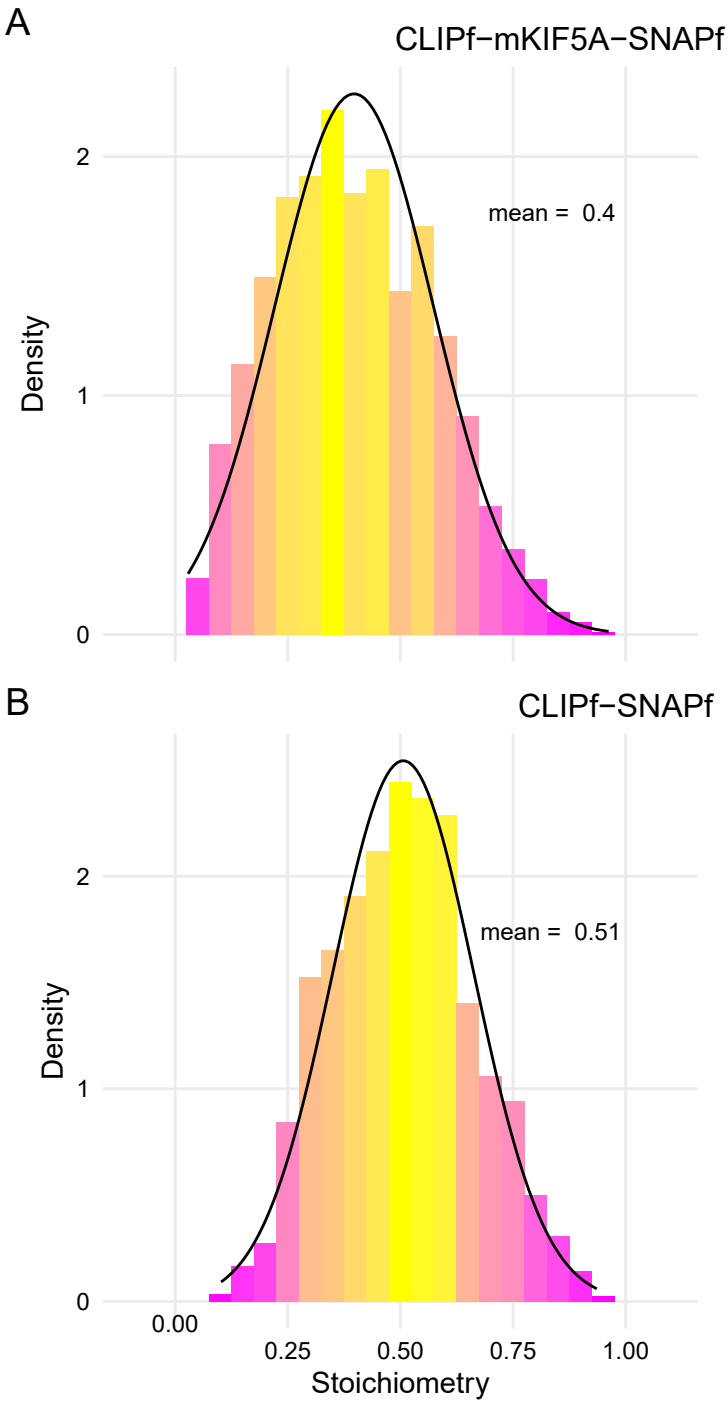

Figure 4.5: **Stoichiometry of CLIPf-SNAPf and CLIPf-mKIF5A-SNAPf fusion proteins**

Histograms of FRET stoichiometry for the CLIPf-mKIF5A-SNAPf fusion protein (A, n = 12, bursts = 3461) and the CLIPf-SNAPf fusion protein (B, n = 3, bursts = 1680). Black line indicates a normal distribution with the mean, or centre of the distribution, written beside.

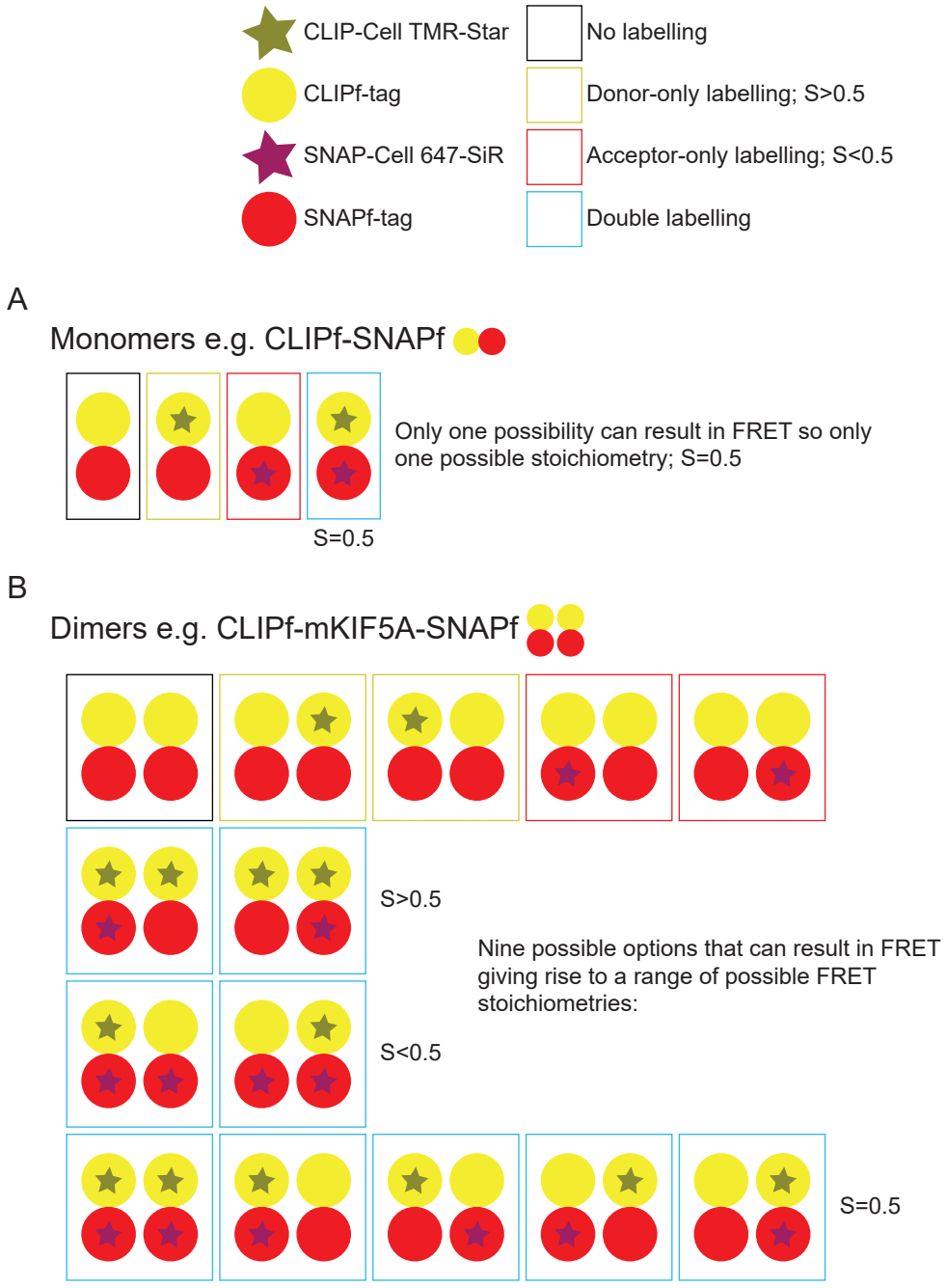

Figure 4.6: **Possible stoichiometries for monomers and dimers due to differences in labelling efficiency**

Partial labelling of dimers can lead to multiple possible configurations that can still result in FRET, but have different stoichiometry values.

would then only have one possible labelling combination that could result in FRET; which has a stoichiometry of 0.5 (see Figure 4.7). Single tagged versions of the CLIPf-mKIF5A-SNAPf-HisTag-FLAG construct were made by deletion mutagenesis resulting in the CLIPf-mKIF5A and the mKIF5A-SNAPf-HisTag-FLAG constructs (see Section 2.1.1.2).

Stoichiometry was shifted more towards 0.5 when HEK293 were cotransfected with the single tagged constructs (CLIPf-mKIF5A and mKIF5A-SNAPf) compared to transfection with the double tagged construct (CLIPf-mKIF5A-SNAPf), with minimal change in FRET efficiency (see Figure 4.8 and 4.9). The increase in stoichiometry also came with a dramatic decrease in the amount of data collected over the same time period. This was expected due to the decrease in the proportion of molecules resulting in FRET; from nine to one. The double tagged constructs will therefore be used from now on as no change in FRET efficiency was observed despite the lower stoichiometry.

4.2.4 Addition of salt to the smFRET assay buffer

To try and explain the unexpected smFRET results for the mKIF5A construct, known conditions that affect kinesin-1 activation were tested to see if they would result in the expected two FRET efficiency populations or a predicted shift in FRET efficiency.

It has been shown that the conformation of the kinesin-1 heavy chain can be altered with increasing salt concentration. At high salt conditions (1 M NaCl) the heavy chains exist in the open conformation whereas at low salt conditions (<0.2 M NaCl) the molecule is in a more compact conformation (Hackney, Levitt, and Suhan 1992 and Stock et al. 1999). So addition of salt to the smFRET assay buffer should result in a shift towards zero FRET efficiency, due to disrupted interactions in the folded inactive conformation.

To test this, HEK293 cells were transfected with the CLIPf-mKIF5A-SNAPf-

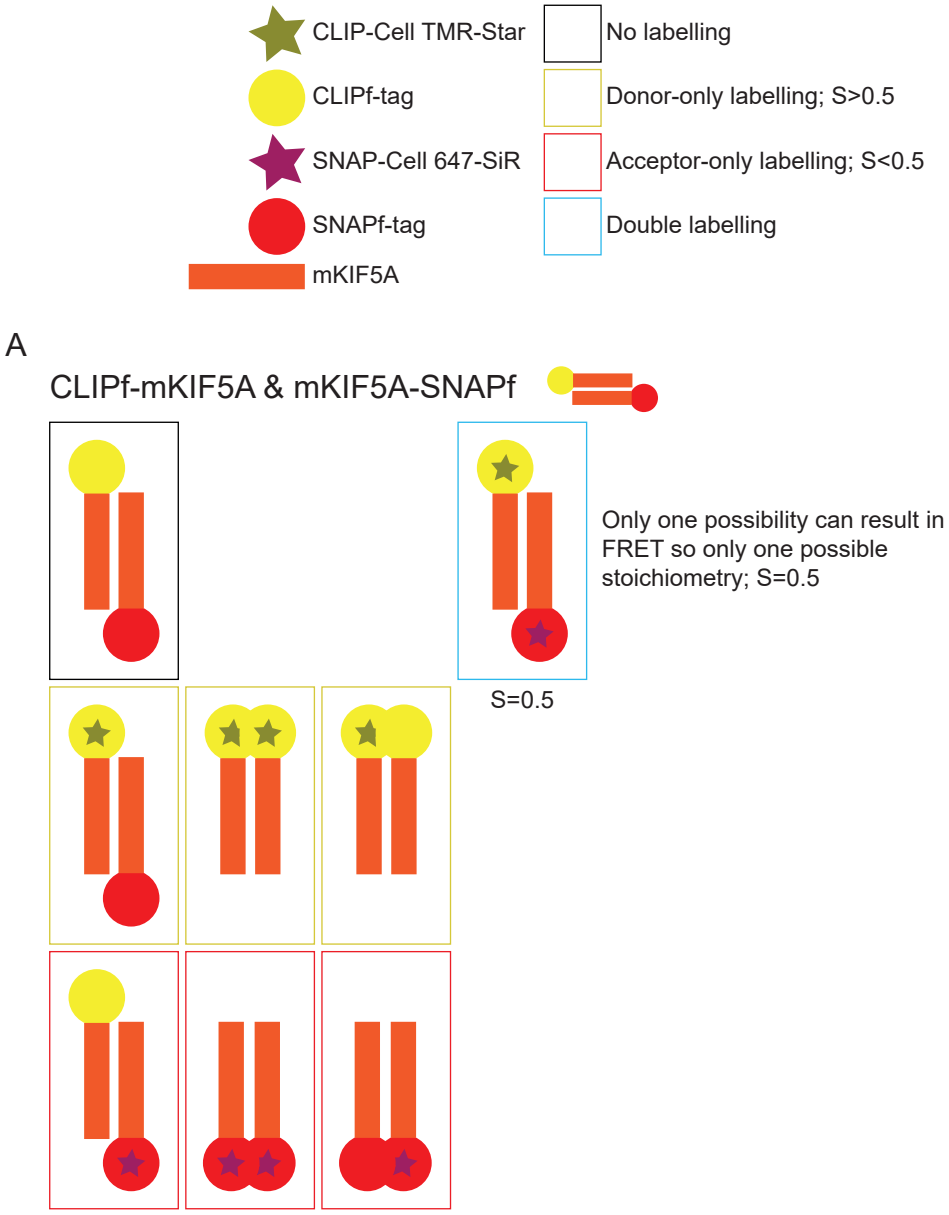

Figure 4.7: **Possible stoichiometries for co-transfection of CLIPf-mKIF5A and mKIF5A-SNAPf-HisTag-FLAG**

Co-transfection of CLIPf-mKIF5A and mKIF5A-SNAPf-HisTag-FLAG (single tagged versions of the CLIPf-mKIF5A-SNAPf-HisTag-FLAG construct) behave similarly to monomeric species with only one possible combination resulting in FRET.

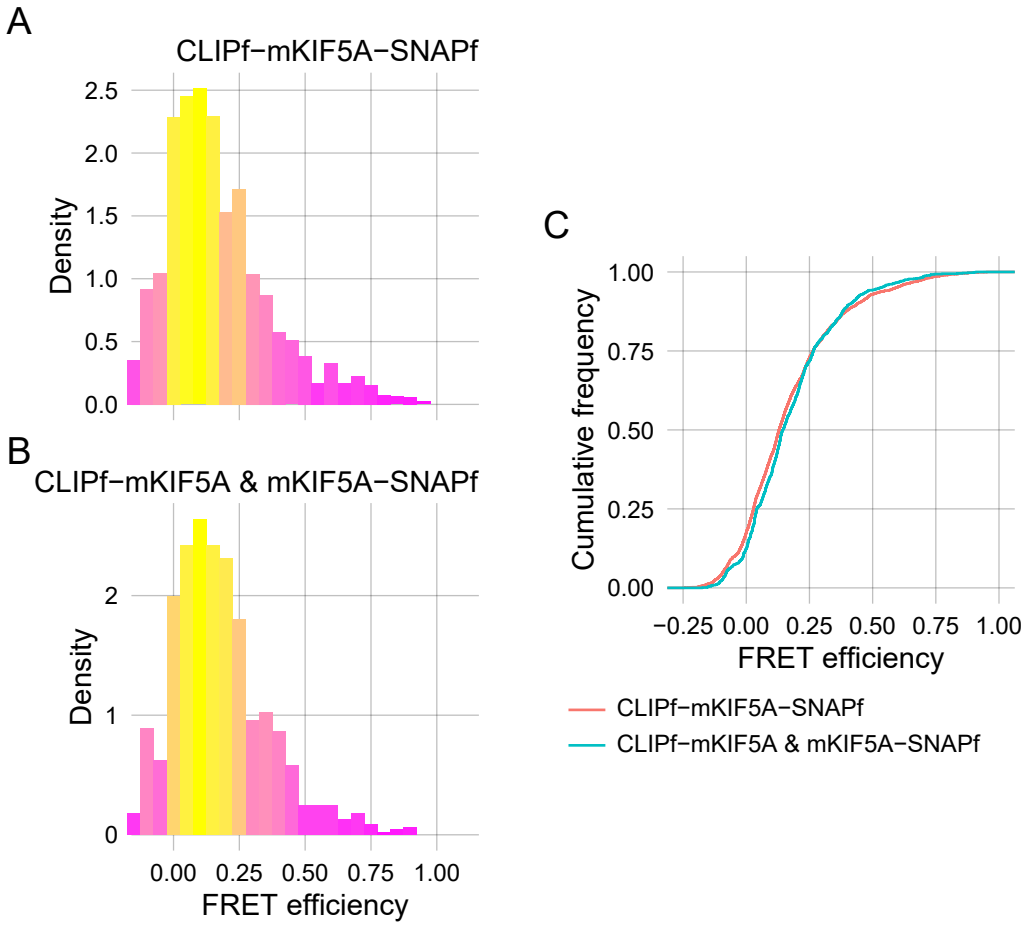

Figure 4.8: **FRET efficiency results for single tagged constructs compared to the double tagged constructs**

Histograms of FRET stoichiometry for the CLIPf-mKIF5A-SNAPf fusion protein (A, n = 15, bursts = 3461) and the co-transfection of CLIPf-mKIF5A and mKIF5A-SNAPf (B, n = 3, bursts = 901). Very small differences in FRET stoichiometry can be seen in cumulative frequency graph (C) when comparing the two data sets.

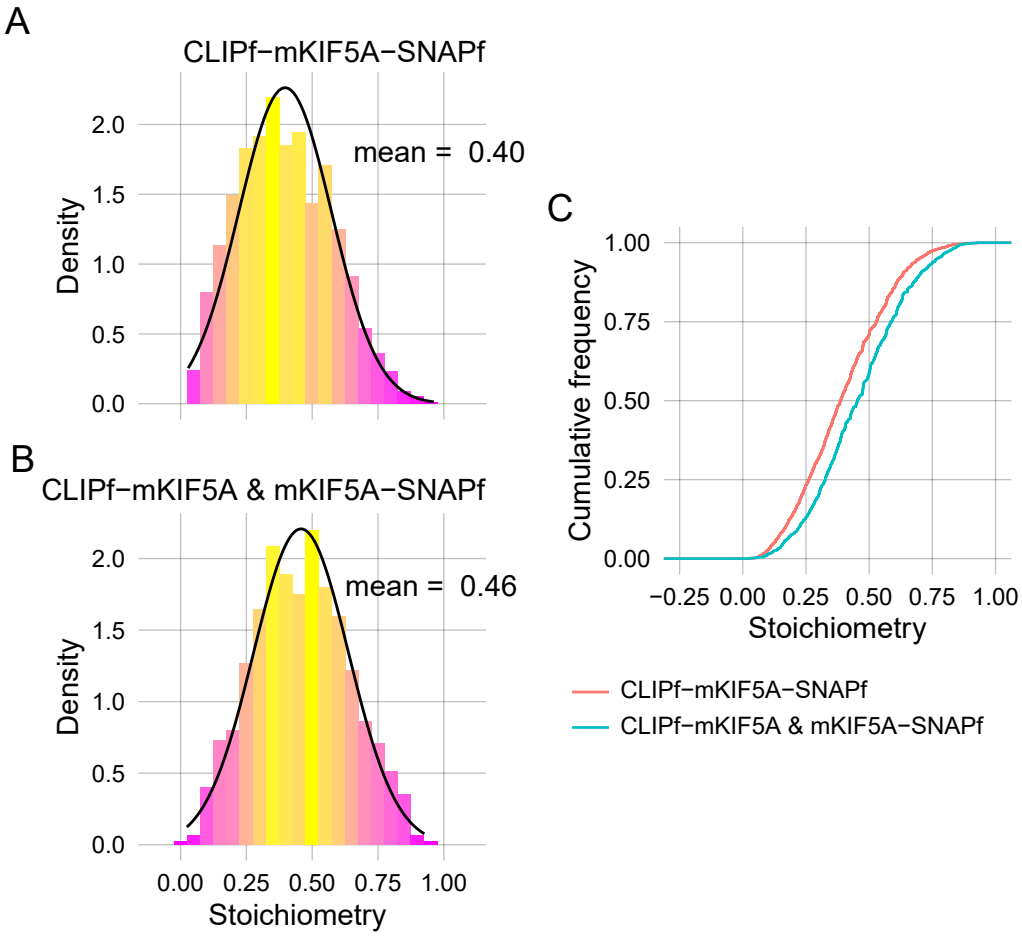

Figure 4.9: **FRET stoichiometry results for single tagged constructs compared to the double tagged constructs**

Histograms of FRET stoichiometry for the CLIPf-mKIF5A-SNAPf fusion protein (A, n = 15, bursts = 3461) and the co-transfection of CLIPf-mKIF5A and mKIF5A-SNAPf (B, n = 3, bursts = 901). Black line indicates a normal distribution with the mean written beside. Differences in FRET stoichiometry can be seen in cumulative frequency graph (C), where CLIPf-mKIF5A-SNAPf is shifted to the left relative to co-transfection of CLIPf-mKIF5A and mKIF5A-SNAPf.

HisTag-FLAG construct. Cell lysate was then diluted in kinesin smFRET assay buffer containing either 0 mM, 150 mM or 800 mM NaCl. Addition of 800 mM NaCl resulted in very inconsistent results (see Figure 4.10). This concentration was chosen to mimic the high salt conditions used in Stock et al. 1999. Any higher concentrations resulted in no FRET signal, likely due to protein precipitation.

A shift towards higher FRET efficiencies was seen on addition of 150 mM NaCl (see Figure 4.10). This result was consistent across biological repeats but the shift was not in the expected direction. Due to disruption in protein interactions it was expected that addition of salt would result in the elongated active form of the kinesin-1 heavy chains resulting in a shift towards lower FRET efficiencies.

4.2.5 Deletion mutants of the mKIF5A fusion protein

The KIF5A isoform has a much longer (64 residues) disordered tail compared to the other two heavy chain isoforms of kinesin-1. This could be the cause of the high FRET efficiency population not previously being seen, as the two fluorophores are not coming in close enough contact. The extra length of the disordered tail also adds flexibility into the system, which might be why the low FRET peak for CLIPf-mKIF5A-SNAPf fusion protein was so broad (see Figure 4.2). Therefore a construct was made, CLIPf-mKIF5A(Δ954-1027)-SNAPf-HisTag-FLAG, by deletion mutagenesis, with the extra tail length deleted. The construct was then the same length as the other two kinesin-1 isoforms; KIF5B and KIF5C (see Section 2.1.1.2). It was expected that this construct should result in the original predicted KIF5A smFRET results, that is two distinct FRET populations; one at zero FRET efficiency and the other at higher FRET efficiency.

No difference in FRET efficiency is seen between the tail deletion construct, CLIPf-mKIF5A(Δ954-1027)-SNAPf-HisTag-FLAG, and the full length KIF5A construct, CLIPf-mKIF5A-SNAPf-HisTag-FLAG (see Figure 4.11). There a still a single broad population at low FRET indicating the KIF5A molecule is still very flexible

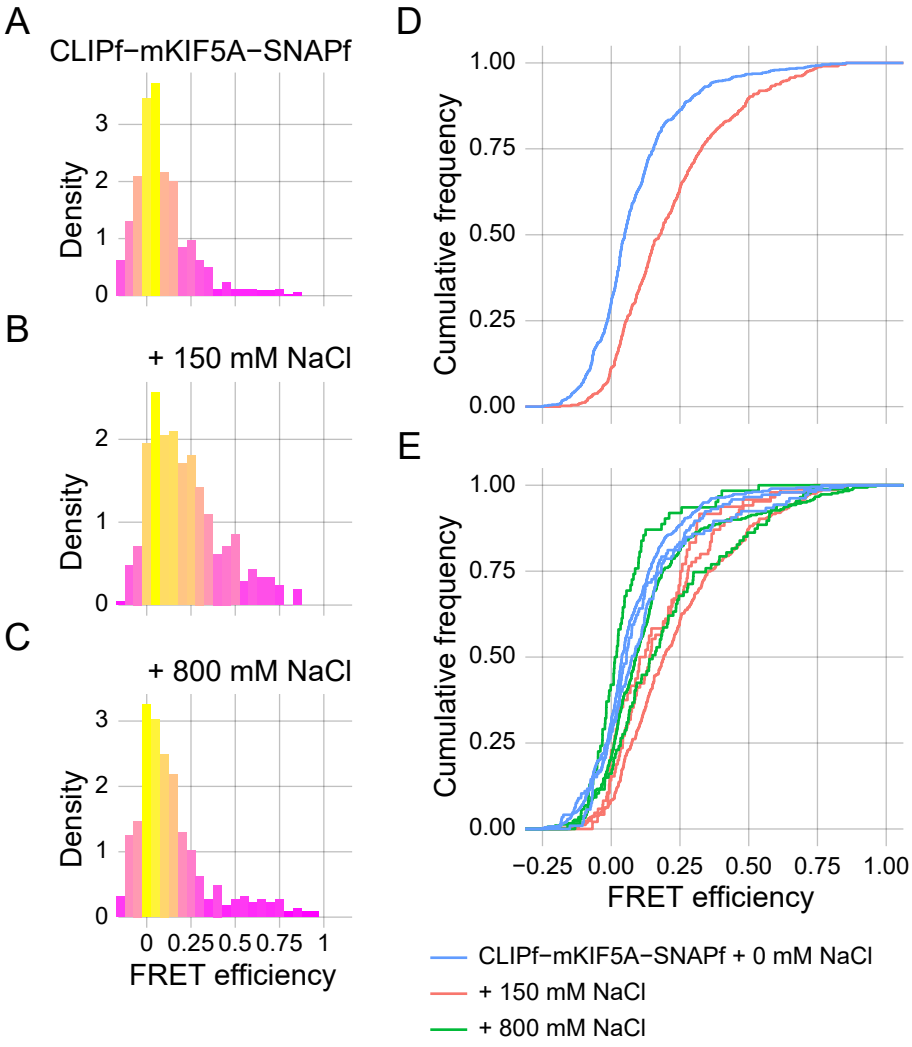

Figure 4.10: **FRET efficiency results for addition of NaCl to the smFRET assay buffer**

Cell lysate with overexpressed CLIPf-mKIF5A-SNAPf fusion protein was diluted in assay buffer with 0 mM (A, n = 3, bursts = 677), 150 mM (B, n = 3, bursts = 421), or 800 mM (C, n = 3, bursts = 449) additional NaCl added. Differences in FRET efficiency between 0 mM and 150 mM NaCl can be seen in cumulative frequency graph (D). Differences in FRET efficiency across biological repeats can be seen in cumulative frequency graph (E), where the highly variable response of kinesin-1 heavy chains to 800 mM NaCl is seen

and adopts many conformations. Also there is still no high FRET population showing that the folded inactive conformation appears to be transient and not a stable population.

The CLIPf-mKIF5A(1-915; Δ505-610)-SNAPf-HisTag-FLAG was made as a zero FRET efficiency, constitutively, active control. The central hinge region, as described by Friedman and R D Vale 1999, was deleted to prevent the folding of the molecule into the inactive folded conformation. The extended tail of KIF5A along with the QIAK motif, the region known to interact with the motor domain in the inactive conformation, was also deleted to again prevent the formation of this conformation.

No significant difference was seen between the FRET efficiency of this double deletion construct, CLIPf-mKIF5A(1-915; Δ505-610)-SNAPf-HisTag-FLAG, and the full length KIF5A construct, CLIPf-mKIF5A-SNAPf-HisTag-FLAG (see Figure 4.11). It was expected for the FRET efficiency of this double deletion mutant to shift to the left and be centred around zero. Instead the FRET efficiency is very similar to the tail deletion mutant and full length wild type KIF5A, with actually a small shift to higher FRET efficiencies. This shows that the flexibility of the KIF5A molecule is still present even with the disordered tail deletion and the central hinge region being deleted. This suggests that there is much more flexibility within the structure of KIF5A then previously thought. The broad range in FRET efficiency suggests the molecule is sampling multiple conformations, which all result in a low FRET efficiency.

4.2.6 AlphaFold2 predicts a new structure for the KIF5A stalk

There is little experimental structural information for the heavy chains of kinesin-1 except for the motor domain. This domain of the KIF5B heavy chain has a solved crystal structure (Kaan et al. 2011). Due to the size and dimerisation state

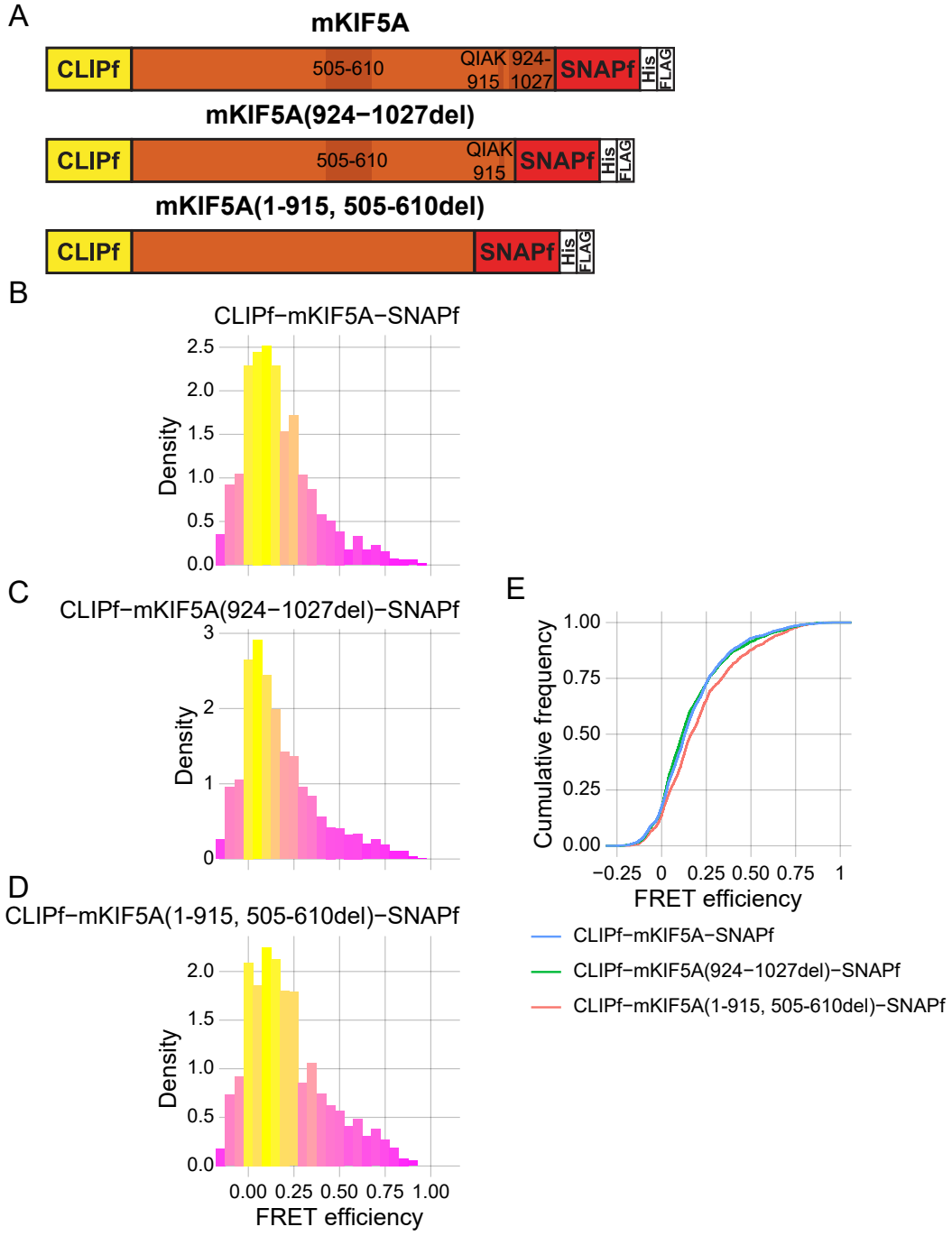

Figure 4.11: **FRET efficiency results for mutants of the CLIPf-mKIF5A-SNAPf fusion protein**

A: Cartoon schematics of CLIPf-mKIF5A-SNAPf mutants. FRET efficiency histograms for the CLIPf-mKIF5A-SNAPf (B, n = 15, bursts = 3461), CLIPf-mKIF5A(Δ954-1027)-SNAP (C, n = 4, bursts = 3800), CLIPf-mKIF5A(1-915; Δ505-610)-SNAPf (D, n = 4, bursts = 1475). Slight differences in FRET efficiency can be seen in cumulative frequency graph (E).

of the molecule, computer based structure predictions have also been limited to disorder/coiled coil or monomeric-based predictions. In July 2021, AlphaFold2 was released allowing the prediction of protein structures by neural networks with no previous or similar structure needed (Jumper et al. 2021).

An AlphaFold2 prediction of mKIF5A(401-1027), motor domain removed, shows more areas of flexibility than has previously been identified through sequence analysis (see Figure 4.12) (Seeger and S. E. Rice 2013). Previous electron microscopy data (Amos 1987 and Hirokawa, Pfister, et al. 1989) along with biochemical analysis (Friedman and R D Vale 1999) suggested that the stalk of KIF5A, anything C-terminal to the motor and neck domain, has a rod-like structure due to dimerisation of the coiled coil. This is only broken by a central hinge region (505-610). The AlphaFold2 prediction shows two more flexible hinge regions (681-690 and 813-823). These new hinge regions may suggest why the previous mutant constructs (CLIPf-mKIF5A(Δ954-1027)-SNAPf-HisTag-FLAG and CLIPf-mKIF5A(1-915; Δ505-610)-SNAPf-HisTag-FLAG) showed no change in FRET efficiency when compared to the full length wildtype KIF5A construct, as not all the flexible regions were removed by these mutants.

Another new prediction in the AlphaFold2 model is the interlocking barrel structure at the central hinge region. It was thought that the structure of the hinge region (505-610) would be a flexible loop; more like the structure predicted for the two new hinge regions. But AlphaFold2 predicts an interlocking hinge region suggesting less general flexibility in this region.

These new regions of flexibility were consistently predicted with shorter constructs of mKIF5A and across the other two isoforms. They were also present in the AlphaFold2 prediction for all three of the human kinesin-1 heavy chain isoforms. In every case, AlphaFold2 predicted five different models and ranked them in order of confidence in the model. Only the top ranked model for each construct was looked at and compared with the rest. In every case the new regions of flexibility were

present. All predictions were dimeric using the ALphaFold-Multimer google colab notebooks.

Figure 4.12: **Alphafold2 prediction of mKIF5A (401-1027)**
Two new hinge regions are predicted at (681-690 and 813-823) which were not thought to be present previously. The central hinge region (505-610) is predicted as an interlocking barrel structure and not a short flexible region like the other hinges now predicted. Structure coloured by pLDDT score: dark blue = 90-100, light blue = 70-90, yellow = 50-70 and orange = 0-50. A high pLDDT indicates high confidence in the structure.

4.3 Discussion

The first kinesin-1 heavy chain studied by this new smFRET assay was mKIF5A. The KIF5B isoform was avoided initially due to its endogenous expression in HEK293 cells. The mouse and human KIF5A genes are very similar at a protein level (Navone et al. 1992). Therefore, as the myc-mKIF5A construct had been used previously (Twelvetrees, Yuen, et al. 2010) this construct was used to make the CLIPf-mKIF5A-SNAPf-HisTag-FLAG construct used in the smFRET assay. The FRET efficiency distribution for the CLIPf-mKIF5A-SNAPf fusion protein was different to that of the CLIPf-SNAPf fusion protein. This shows that this assay allows the conformation of the heavy chains of kinesin-1 to be studied, at a single molecule level, without any surface tethers or protein purification, for the first time.

One limitation of this current work is the lack of evidence that the two tags (CLIPf and SNAPf) are not affecting the conformational freedom of the kinesin-1 heavy chain. Previous studies into the kinesin-1 heavy chain conformation have used similar sized tags (Cai et al. 2007). They were able to see a higher FRET efficiency population indicating that addition of the tags did not prevent the formation of the compact folded autoinhibited conformation. Electron microscopy or atomic force microscopy studies of the CLIPf-mKIF5A-SNAPf fusion protein might help ensure the tags are not affecting autoinhibtion. Also, preliminary data within the Twelvetrees lab suggests that the CLIPf-mKIF5A-SNAPf fusion protein has similar activity to that of other full length mKIF5A tagged constructs again suggesting that the fusion protein is still autoinhibited. Other assays could be done including:

- ATPase activity, to ensure the N-terminal CLIPf-tag is not altering the ability of the kinesin-1 construct to hydrolyse ATP in the presence of microtubules,
- co-immunoprecipation with the light chains, to check the C-terminal SNAPf-tag is not affecting the kinesin-1 construct's ability to form the kinesin-1 heterotetramer,

- microtubule gliding and kinesin walking, to ensure the N-terminal CLIPf-tag is not preventing microtubule binding or walking of the kinesin-1 construct.

The FRET efficiency result for the CLIPf-mKIF5A-SNAPf fusion protein wasn't the expected two distinct peaks; one at low and one at a higher FRET efficiency (see Figure 4.2). One low FRET population was much larger and broader then predicted, and not centred around zero. A higher FRET efficiency population was not present at all. Also the stoichiometry was not centred around 0.5 which is expected in smFRET results. Reasons behind these unexpected results needed to be ruled out before detailed analysis of this result could occur.

The stoichiometry measurement can be shifted predictably by altering the tagging of the overexpressed mKIF5A homodimer (see Figure 4.9). The lower stoichiometry of the CLIPf-mKIF5A-SNAPf fusion protein is due to incomplete labelling of the two CLIPf-tags, resulting in bursts with a higher acceptor contribution then the donor. Stoichiometry can be shifted back towards 0.5, without altering the FRET efficiency, by co-transfection of single tagged mKIF5A constructs (see Figure 4.8 and 4.9). Although the stoichiometry is lower than 0.5, the usual desired stoichiometry in smFRET experiments, double tagged constructs will still be used. This is because they resulted in a much higher burst count per experiment due to the higher ratio of molecules that can result in FRET. This predicted shift in stoichiometry adds to the robustness of the assay by showing that we can shift the stoichiometry of a molecule without altering the FRET efficiency.

The difference in labelling efficiency between the two tags and their corresponding dyes had an unexpected advantage of being able to distinguish between monomeric and dimeric proteins. The labelling of the CLIPf-tag with the CLIP-Cell TMR-Star fluorescent ligand is much slower and less efficient then SNAP-Cell 647-SiR labelling the SNAPf-tag. This was seen in the shorter incubation time and higher concentrations of ligand needed to full saturate the tag (see Figure 3.6 and 3.7). Due to the lower labelling efficiency of the CLIPf-tag compared to the SNAPf-tag a

lower stoichiometry was observed for the double labelled mKIF5A construct. This is because, in dimeric proteins, most had two labelled SNAPf-tags, but only one labelled CLIPf-tag meaning most bursts had a stoichiometry lower then 0.5. This labelling combination could still result in a FRET signal as both ligands were present (see Figure 4.6.B). For monomeric proteins, the stoichiometry of bursts was centred around 0.5 as single labelling of the fusion protein would not result in a FRET signal; a signal is only possible when both fluorophores are present (see Figure 4.7). Therefore by looking at the mean stoichiometry for a fusion protein in the smFRET assay, the dimerisation state of the protein can be determined. This gives confidence that the overexpressed double tagged mKIF5A fusion protein is folding correctly as a homodimer. If the CLIPf-SNAPf fusion protein was undergoing unexpected oligomerisation, the stoichiometry would be even lower.

Co-immunoprecipitation showed no evidence of heterodimer formation between the endogenously expressed hKIF5B isoform and the tagged mKIF5A isoform (see Figure 4.4). The only band present in the mKIF5A immunoprecipitation sample was at around 55 kDa. This is much lower then the expected molecular weight of either the endogenous hKIF5B heavy chain or the tagged construct. It was also present in the negative control immunoprecipitation sample from untransfected cells. This band is therefore likely to be due to the heavy chain of the α-FLAG antibody as it and the α-KIF5B antibody are both from rabbit.

The co-immunoprecipitation samples were also probed with the α-FLAG antibody to check transfection levels. The CLIPf-hKIF5B-SNAPf fusion protein was used as a positive control in the co-immunoprecipitation experiment but very low expression of the protein was seen in all three repeats (see Figure 4.4). New DNA preparations also resulted in the same low expression levels. As this is the isoform endogenously expressed in these HEK293 cells, a regulatory mechanism within the cell might be preventing overexpression by degrading excess protein.

Changing the salt concentration, and using mutants of mKIF5A, did not alter

the FRET efficiency distribution for CLIPf-mKIF5A-SNAPf as expected. A general shift to higher FRET efficiency was seen on addition of 150 mM NaCl to the assay buffer, showing that the conformation of the kinesin-1 heavy chain is salt sensitive. smFRET is much more sensitive to multiple short lived conformations then the sedimentation coefficient experiments previously used (Hackney, Levitt, and Suhan 1992 and Stock et al. 1999). This is due to the sampling of many single molecules rather then the averaging of the result over the entire population. The increase in FRET efficiency, and the broadness of the peak, on addition of 150 mM NaCl suggests an increase in flexibility of the molecule and the sampling of multiple conformations.

The kinesin-1 heavy chain structure predicted by AlphaFold2 supports the possible greater flexibility of the molecule with the prediction of two new hinge regions. These hinge regions explain why the double KIF5A mutant, with the central hinge and the tail (including the QIAK motif) regions deleted, still resulted in a broad low FRET population. The two newly predicted hinge regions were still present in the mutant so the molecule was still highly flexible. A new mutant where these regions are also deleted could provide the zero FRET efficiency expected by the double mutant. Although AlphaFold2 does do dimer structural predictions there are limitations. The size of the protein predicted is limited meaning the kinesin-1 heavy chain has to be split into two sections. Also it is only a prediction and requires experimental validation before any conclusions are conclusive. When no other detailed structural information is available, it does help to analyse the unexpected smFRET results.

These results suggest that not only is the kinesin-1 heavy chain much more flexible then suggested by previous experimental results (Hackney, Levitt, and Suhan 1992) but, that the folded inactive conformation is very transient and not stable. This is indicated by the broadness of the low FRET efficiency distribution and the lack of a higher FRET efficiency population respectively. More detailed analysis, for example burst variance analysis, could result in the resolution of multiple con-

formations from this broad FRET efficiency peak, but currently these results do not agree with the theory of a stable inactive pool of kinesin-1. Rather they suggest a more intricate complex mechanism of activation and autoinhibition.

The lack of two distinct FRET efficiency populations for the kinesin-1 heavy chain could be due to the current simplicity of the smFRET assay. I have so far studied kinesin-1 in isolation, but in cells it works with other proteins to transport cargo. Therefore, the component needed to cause kinesin-1 autoinhibition could be currently missing from the assay. In the next chapter the complexity of the smFRET assay will be increased by adding in more components of axonal transport.

Chapter 5

Increasing the complexity of the kinesin-1 smFRET assay

5.1 Introduction

5.1.1 Axonal transport

Neurons are extremely large highly complex cells that have to last your entire lifetime. Kinesin-1 is the main antereograde motor protein, transporting newly synthesised material from the cell body of neurons, along the axon, to the synapse. The axon is the part of the neuron that separates the cell body from the synapses and in humans it can be up to a metre in length (Maday et al. 2014). Kinesin-1 can only take 8 nm steps so traveling the length of the axon is an extreme endeavour (Schnitzer and Block 1997). Signalling molecules and old proteins are taken back to the cell body by another motor protein; cytoplasmic dynein. This distribution of material, caused by motor proteins travelling up and down the axon, is called axonal transport. It must be highly efficient to not compete with the high energy demands of neurons (Maday et al. 2014).

Kinesin-1 activation and function must be highly controlled in axonal transport. This is achieved mainly through autoinhibition preventing unnecessary hydrolysis

of ATP, and accumulation of the motor protein in the cell periphery. As kinesin-1 autoinhibtion is so crucial in axonal transport, adding in other components of the process to the smFRET assay will help understand the activation of kinesin-1 and axonal transport regulation.

5.1.2 Microtubule tracks

In the axon, motor proteins, like kinesin-1 and cytoplasmic dynein, travel along microtubule tracks. They are made up of tubulin monomers that dimerise. These dimers come together to form the hollow cylindrical structure of the microtubule resulting in usually thirteen protofilaments per structure. Protofilaments are long rod-like polymers consisting of tubulin dimers (Downing and Nogales 1998). Due to formation through tubulin dimer polymerisation, the microtubule structure has inherent polarity with a minus-end and a fast growing plus-end. In the axon all microtubules are aligned to result in uniform polarity across the axon with all minus-ends being directed towards the cell body (Baas et al. 1988). This ensures the correct movement of material in axonal transport; kinesin-1 is a plus-end directed motor whereas cytoplasmic dynein travels towards the minus-ends of microtubules.

Kinesin-1 must be active in the presence of microtubules, in order to walk along them, but research has focused on this motility and not on the pre-motile activation step. Therefore it is not clear whether the presence of microtubules alone are enough to activate kinesin-1 or alter its conformation. The type of single molecule assays used to study kinesin-1 has prevented investigation of this as they either do not look at the conformation of the molecule or must contain microtubules to work. For example, total internal reflection fluorescence (TIRF) microscopy or optical trap experiments both contain microtubules and don't provide any information without them. Therefore any information about the activation step is lost within the experiment.

The smFRET assay I have developed allows the investigation of this pre-motile

step. This is due to the fact that microtubules aren't required for the assay so kinesin-1 conformation can be studied with and without them. Therefore whether microtubules are necessary for kinesin-1 activation, or whether they alter the conformation of the molecule, can be studied for the first time.

5.1.3 The conformation of kinesin-1 attached to cargo

One of the main theories about kinesin-1 activation is that kinesin-1 is always attached to the cargo it is transporting and exists in either the active or inactive conformation upon them. This was first proposed in the second Vale *et. al.* paper in 1985 detailing the co-discovery of kinesin-1 (R D Vale, Schnapp, et al. 1985) but still no conclusive evidence has been shown.

The effect of cargo on kinesin-1 activation has been long investigated but currently with no conclusive answers. In 1992, Schapp *et. al.* looked at the transport of endogenous organelles using kinesin-1 and microtubules purified from squid optical lobes. They found that kinesin-1 seemed to be "firmly attached to organelles under physiological conditions" but it wasn't clear if kinesin-1 was in the active or inactive conformation (Schnapp et al. 1992). More recently, a pre-print by Canty *et. al.*, when studying the transport of mitochondria, showed that TRAK1 and TRAK2 recruited kinesin-1 but did not seem to fully activate the molecule. The addition of these adaptor proteins did increase landing rate and mobile fraction of kinesin-1 but the presence of an addition factor, MAP7, resulted in a more dramatic increase in activation. Again, the conformation of kinesin-1 was not directly studied with activation being inferred through increased landing on microtubules and increased movement of kinesin-1 (Canty et al. 2021). As the smFRET assay can be used with cell lysate samples the conformation of kinesin-1 when attached to endogenous cargo can be studied for the first time with only small alterations to the current assay.

The aims of this chapter are to increase the complexity of the smFRET assay and study kinesin-1 conformation in the presence of cargo and microtubules. As

molecules are allowed to freely diffuse through the confocal volume, without any constraints on their conformational freedom, the conformation of kinesin-1 bound to other molecules can be measured. Then the necessity for the presence of either cargo or microtubules for the inactivation of kinesin-1 can be investigated. The absence of these components might be why the two predicted FRET efficiency populations were not seen in the previous chapter.

5.2 Results

5.2.1 Studying heavy chain kinesin-1 conformation when attached to endogenous cargo

To be able to investigate the conformation of kinesin-1 when bound to endogenous cargo, some small changes need to be made to the current smFRET sample preparation. Firstly the isoform of kinesin-1 heavy chain studied needs to change to KIF5B and the species needs to swap from mouse to human. Secondly, a positive control for the presence of endogenous cargo is needed and lastly the cells need to lysed without the use of detergent. These points will be be discussed in greater detail in the following sections.

5.2.1.1 Changing the isoform, and species, of the kinesin-1 heavy chain construct

The heavy chain construct studied by the smFRET assay needs to be CLIPf-hKIF5B-SNAPf-HisTag-FLAG instead of CLIPf-mKIF5A-SNAPf-HisTag-FLAG (see Section 2.1.1.2). This is because HEK293 cells only express the hKIF5B isoform of the kinesin-1 heavy chain. Therefore to study endogenous cargo the endogenously expressed kinesin-1 isoform needs to be tagged to prevent potential problems with cargo recognition. Before changing the smFRET protocol, HEK293 cells were transfected with the CLIPf-hKIF5B-SNAPf-HisTag-FLAG construct to see if there were

any changes in the smFRET results for hKIF5B compared to mKIF5A. The optimised smFRET assay protocol was used to ensure high labelling and low background.

Again two distinct FRET efficiency populations, representing the open and closed conformations of the heavy chain, were not seen. The results were comparable to those of the CLIPF-mKIF5A-SNAPf fusion protein where only one broad FRET efficiency population is observed (see Figure 5.1). This suggests that hKIF5B behaves similarly to mKIF5A in this experiment with the closed inactive form not being stable, and the whole molecule being much more flexible then previously thought.

One difference between the isoform constructs is the expression levels of the resulting fusion proteins. CLIPf-hKIF5B-SNAPf fusion protein is expressed at much lower levels then CLIPf-mKIF5A-SNAPf. This was seen in the co-immunoprecipitation results discussed earlier (see Section 4.2.2), but was seen again in these experiments with the amount of bursts per experiment being drastically reduced. In gel fluorescence analysis of the smFRET samples was carried out to confirm low burst count was due to expression levels. Signal for both fluorophores was drastically reduced for the CLIPf-hKIF5B-SNAPf fusion protein when compared to the CLIPf-mKIF5A-SNAPf (see Figure 5.2).

5.2.1.2 A positive control for the presence of endogenous cargo

Lysosomes were chosen as the endogenous cargo to investigate the conformation of kinesin-1 when bound upon them. This was because a large population of lysosomes, of different sizes, should be present within any cell lysate sample. Lysosomal-associated membrane protein 1 (LAMP1) is a known lysosome marker as it primarily resides across lysosomal membranes. It was chosen as the positive control protein as should be present in lysosomes of different sizes and isn't specific for a certain cargo of kinesin-1.

A LAMP1 construct was tagged with CLIPf and SNAPf at the C-terminus resulting in the LAMP1-CLIPf-SNAPf-HisTag-FLAG construct (see section 2.1.1.3).

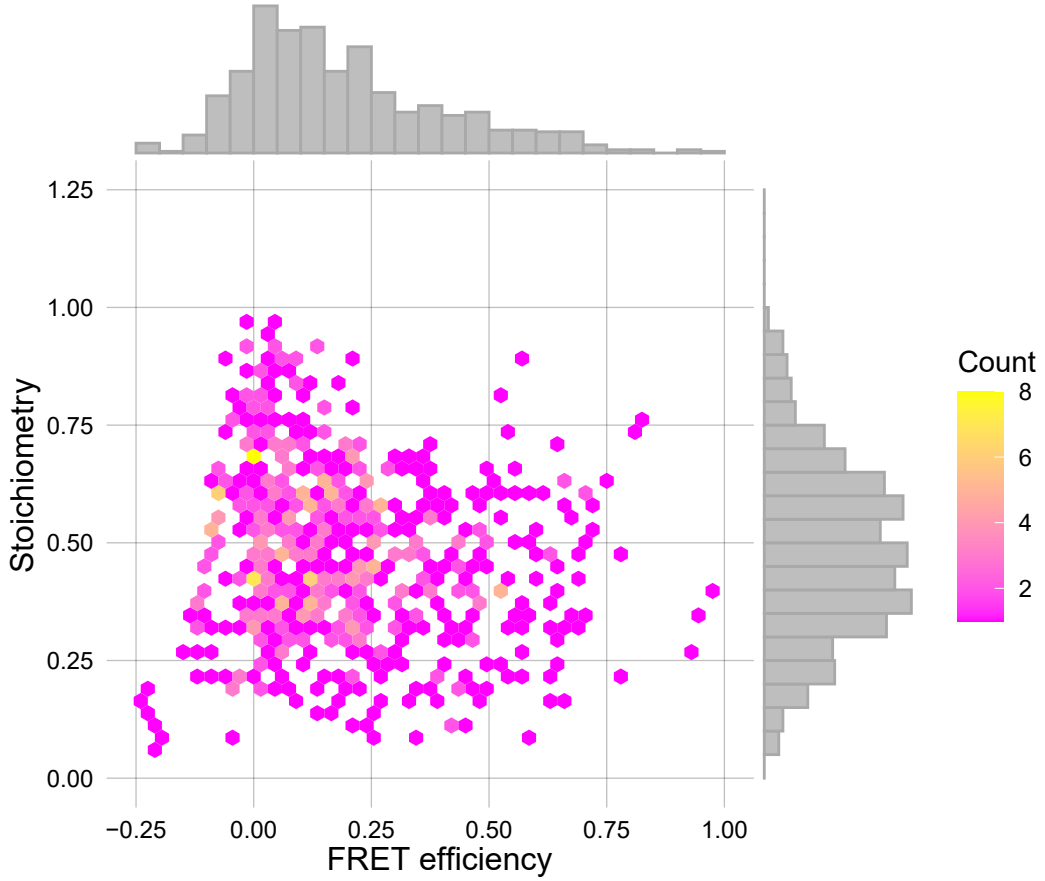

Figure 5.1: **smFRET results for CLIPf-hKIF5B-SNAPf fusion protein**

The stoichiometry vs FRET efficiency hex plot (with histograms) of analysed data is shown. Colour indicates number of bursts in each hexagonal pixel. Data is combined from n = 6 biological repeats; bursts = 666.

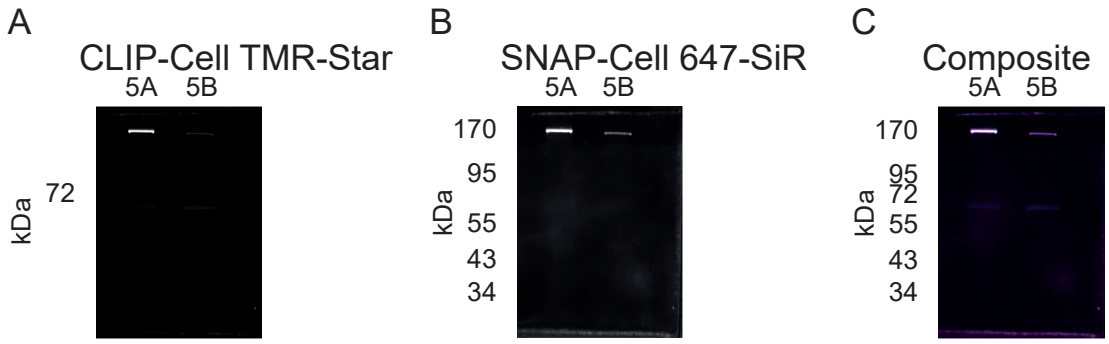

Figure 5.2: **In gel fluorescence for the CLIPf-mKIF5A-SNAPf and the CLIPf-hKIF5B-SNAPf fusion proteins**

smFRET samples for the CLIPf-mKIF5A-SNAPf (5A) and the CLIPf-hKIF5B-SNAPf (5B) fusion proteins were run on a SDS-PAGE gel and analysed by in gel fluorescence. A: 600 nm channel imaging CLIP-Cell TMR-Star fluorescence. B: 700 nm channel imaging SNAP-Cell 647-SiR fluorescence. C: Composite image of both channels with the 600 nm channel shown in cyan and the 700 nm channel in magenta.

This construct should result in a similar FRET efficiency to the previous positive control (the CLIPf-SNAPf fusion protein) as the distance between the two tags is unchanged. Therefore, this LAMP1 positive control could be used to indicate the presence of intact lysosomes in the smFRET samples. This would be done using the high FRET efficiency of the positive control along with the large burst duration of an intact lysosome. The duration of a FRET burst is proportion to the size of the causative molecule. A larger molecule will take longer to diffuse across the confocal volume of the smFRET instrument. This allows for a longer sampling time and therefore a longer duration for the resultant FRET bursts. If the new LAMP1 positive control was contained within a lysosome the burst would have a much longer duration then if the fusion protein was on its own.

The CLIPf- and SNAPf-tag were placed on the C-terminus to minimise mislocalisaion of the LAMP1 protein. The N-terminus of LAMP1 is located in the lumen of lysosomes and is highly conserved within the LAMP protein family. A a transmembrane helix then spans the outer membrane of the lysosomes with the C-terminus being a cytosolic tail (Wilke et al. 2012). Therefore the addition of tags to the C-terminus was less likely to cause mislocalisation of the LAMP1 protein.

To check for mislocalisation, HEK293 cells (plated on coverslips) were transfected with the LAMP1 positive control construct and incubated in SNAP-Cell 647-SiR using the previously optimised conditions. Unbound ligand was not removed through a washout incubation step as cells were fixed with paraformaldehyde. This step increases the permeability of the cells meaning unbound ligand is removed in later wash steps without needing another elongated washout period. The coverslips were then mounted on microscope slides and imaged with a Leica SP5 confocal microscope (see Section 2.6.1 for more details). Due to COVID restrictions, coverslip preparation, cell plating and imaging were done by other members of the lab (Ashleigh Davey, Emma Turner and Alison Twelvetrees respectively). LAMP1-CLIPf-SNAPf fusion protein transfected and expressed well in the HEK293 cells (see Figure 5.3).

Staining resulted in distinct puntate as expected from a lysosomal marker but overlap with another lysosomal marker would ensure mislocalisation is not occurring; especially due to overexpression.

This experiment was repeated with the addition of LysoTracker™ Red DND-99 (Invitrogen) to the cells. LysoTracker consists of a fluorophore connected to a weak base making it sensitive to pH. It is therefore selective for acidic organelles like lysosomes. Again due to COVID restrictions, the same parts of this protocol were done by the same lab members. Good overlap between the LysoTracker and the SNAP-Cell 647-SiR signal was seen in cells transfected with the LAMP1 construct (see Figure 5.4). This indicates that the construct is not being mislocalised due to the addition of the CLIPf- and SNAPf-tags or overexpression and can be used as a positive control for intact endogenous lysosomal cargo in future smFRET experiments.

5.2.1.3 Lysing HEK293 cells without detergent

Homogenisation, the process of lysing cells but keeping individual components within the cell intact, will be used to break apart cells instead of a detergent based lysis buffer. In the current smFRET sample preparation, HEK293 cells are lysed by incubation in a detergent based lysis buffer (see Table 2.5) followed by a high speed centrifugation step to remove the cellular debris. The supernatant then contains the solubilised proteins and lipids. The detergent in the lysis buffer breaks up the cell membrane and therefore lyses the cells. But this can lead to further disruption of other membranes within the cells, like lysosomes. In order to study kinesin-1 attached to endogenous cargo (in this case lysosomes) a large concentrated population of intact lysosomes is needed meaning detergent can not be used to lyse the cells.

Homogenisation by mechanical force can be used. The homogenisation protocol used is based on the "Homogenization of Mammalian Cells" protocol by de Araújo *et. al.* (Araújo et al. 2015). They use mechanical force from a syringe and needle

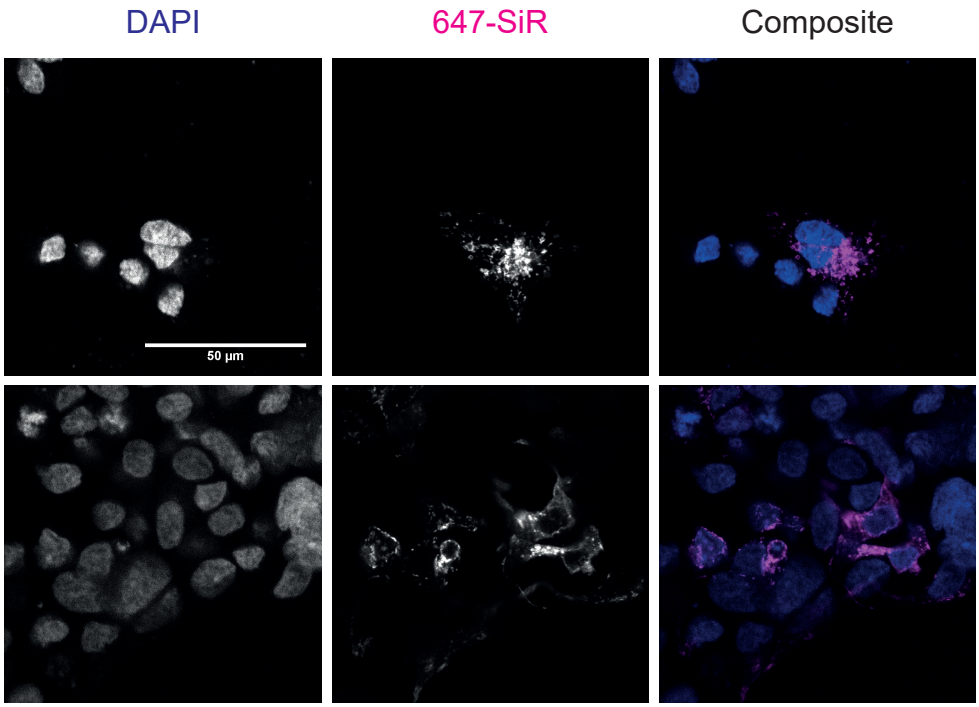

Figure 5.3: **Confocal microscopy images of the LAMP1-CLIPf-SNAPf fusion protein**

LAMP1-CLIPf-SNAPf transfected cells were stained with DAPI for nuclei (blue) and SNAP-Cell 647-SiR for LAMP1-CLIPf-SNAPf transfection (magenta). Cells were fixed with paraformaldehyde before imaging on Leica SP5 confocal microscope. Due to COVID restrictions coverslip preparation, cell plating and imaging were done by other members of the Twelvetrees lab (see Section 2.6.1)

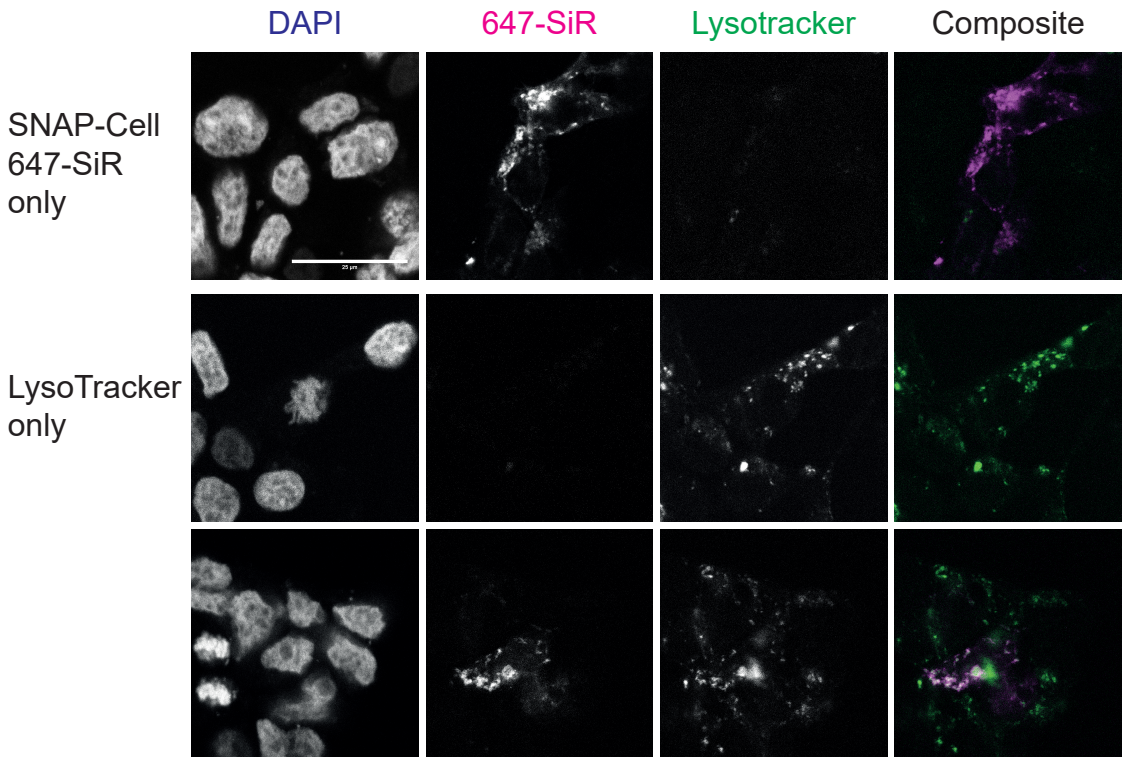

Figure 5.4: **Overlap of LysoTracker with the LAMP1-CLIPf-SNAPf fusion protein signal**

LAMP1-CLIPf-SNAPf transfected cells were stained with DAPI for nuclei (blue), SNAP-Cell 647-SiR for LAMP1-CLIPf-SNAPf transfection (magenta) and LysoTracker for lysosomes (green). Overlap appears white in composite. Cells were fixed with paraformaldehyde before imaging on Leica SP5 confocal microscope. Due to COVID restrictions the same parts of this protocol were done by other members of the Twelvetrees lab (see Section 2.6.1)

to break apart the plasma membranes. This can also be done with a homogeniser; in this case a Isobiotec Cell Homogenizer and two gas tight syringes. The cell suspension is repeatedly past from one syringe to the other through a small chamber containing a tungsten ball bearing only just smaller then the chamber. This forces the cells through the narrow gap resulting in mechanical lysis of the cells. The size of the ball bearing varies giving a clearance size ranging from 4 to 18 µm. The ball bearing resulting in a clearance size of 4 µm was chosen to ensure good lysis of the cells. The homogenisation buffer from Aguado et al. 2016 was used as their homogenisation protocol was specific for the isolation of lysosomes.

To check the overexpressed hKIF5B fusion protein was still present in the supernatant after homogenisation, a 10 cm dish of HEK293 cells were transfected with the CLIPf-hKIF5B-SNAPf-HisTag-FLAG construct. They were homogenised as described in Section 2.2.7 and the resultant samples run on an SDS-PAGE gel before analysis by western blot. The α-tubulin(alpha) rabbit antibody was used as a marker of cell lysis as the pellet should not contain much free tubulin after complete lysis.

Expression of the hKIF5B fusion protein looks high and it is present in both the lysis pellet and the supernatant (see Figure 5.5). The general expression of the fusion protein is higher then in the co-immunoprecipitation (in Section 4.2.2) as more cells were used. This was to overcome low expression, but also to have larger volumes making homogenisation easier. The tubulin marker suggests lysis by homogenisation was efficient and the CLIPf-hKIF5B-SNAPf fusion protein being present in the supernatant shows that this supernatant can be used for future smFRET experiments.

5.2.1.4 Detecting intact lysosomes by smFRET

The LAMP1 positive control was used to determine if intact lysosomes were present in the resultant homogenisation supernatant and were then detectable by smFRET.

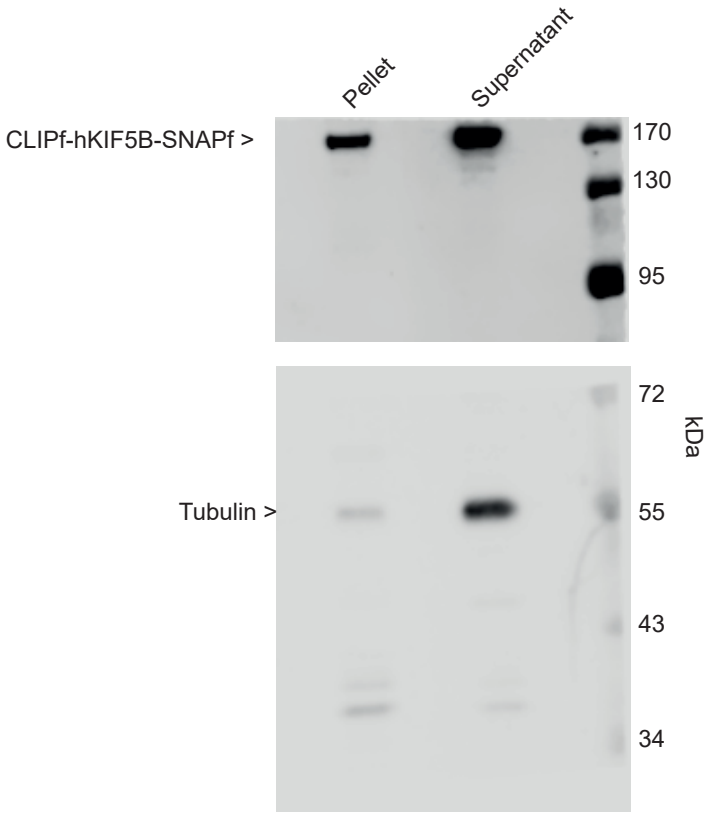

Figure 5.5: **Cell lysis by homogenisation**

CLIPf-hKIF5B-SNAPf-HisTag-FLAG transfected cells were lyse by homogenisation. The resultant pellet and supernatant were separated by SDS-PAGE and probed by western blot using α-KIF5B rabbit and α-tubulin rabbit.

HEK293 cells were transfected with the LAMP1-CLIPf-SNAPf-HisTag-FLAG construct and the fluorescence ligands were added and excess washed away as per the optimised conditions (see Section 2.4). Cells were then lysed following the optimised homogenisation protocol (see Section 2.2.7). The resultant supernatant was diluted in kinesin smFRET assay buffer and data recorded from the smFRET instrument as usual.

Intact lysosomes were detected by smFRET but they were only labelled with the acceptor fluorophore (see Figure 5.6). This shows that the new LAMP1 positive control works in smFRET conditions and can be used to identify intact lysosomes. The large duration of the lysosomal bursts, especially when compared to the length of the bursts of CLIPf-SNAPf fusion protein, indicate that the lysosomes are still intact and were not broken up by the cell lysis process. The photon count for each bursts was much higher then that of CLIPf-SNAPf fusion protein bursts. This suggests that more then one labelled LAMP1 fusion protein is present in the lysosome.

In total less then ten bursts were recorded over the whole 15 minute experiment. Usually the number of bursts per 15 minutes is around 200. Also all bursts were only labelled with the acceptor but with some changes to the sample preparation double labelled intact lysosomes should be able to be detected by smFRET. Due to time constraints these experiments weren't continued but they do show the possibility of studying the conformation of the kinesin-1 heavy chain when it is attached to endogenous lysosomes by smFRET. Homogenisation of the cells was shown to result in intact lysosomes with the hKIF5B construct still present in the resultant supernatant.

5.2.2 Studying heavy chain kinesin-1 conformation in the presence of microtubules

Autoinhibition is only necessary when microtubules are present but there is no evidence that they are required. Therefore addition of microtubules into the smFRET

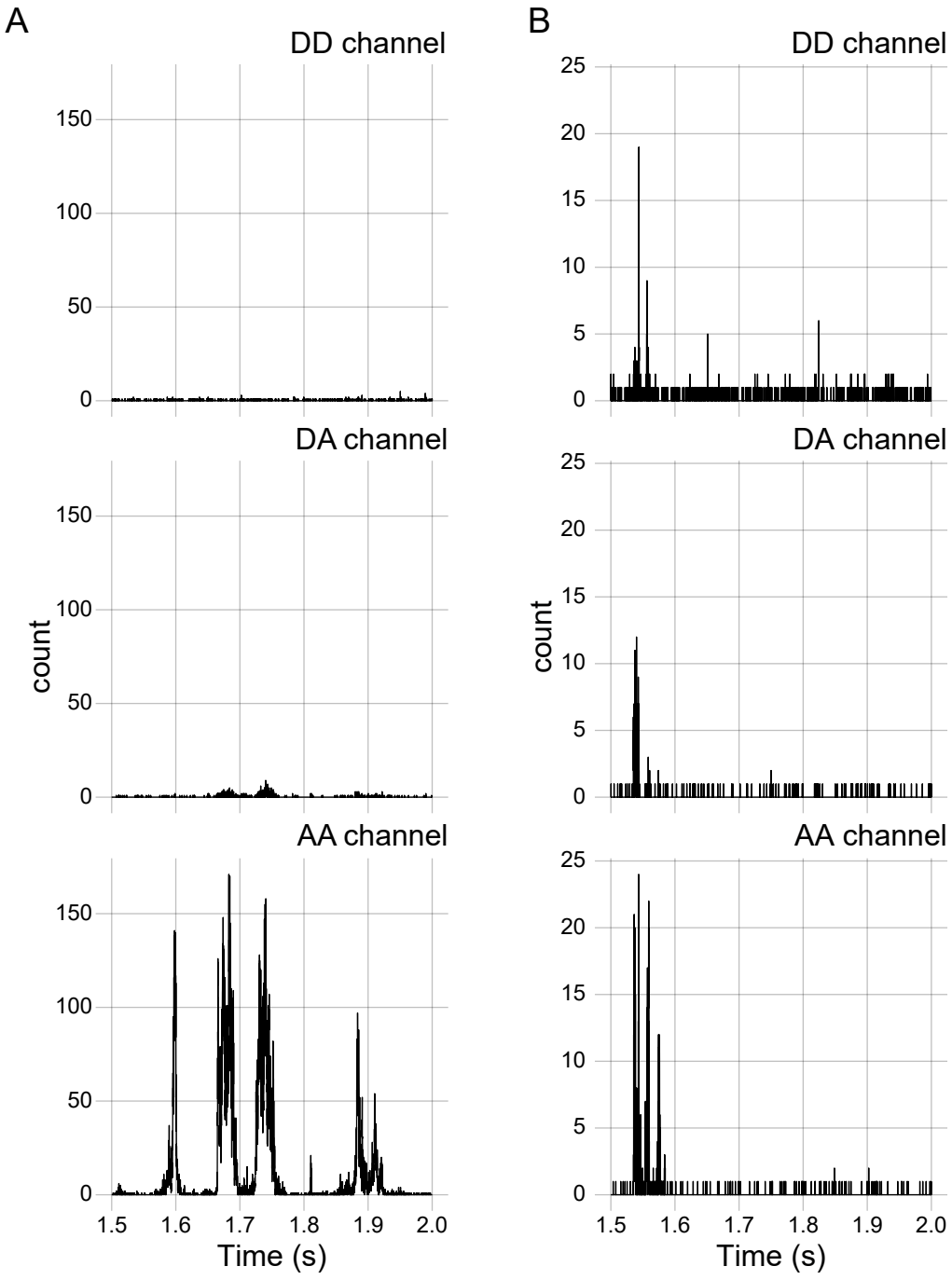

Figure 5.6: **Photon trace for the LAMP1-CLIPf-SNAPf and CLIPf-SNAPf fusion proteins**

Photons are split into three channels, donor-only (DD), donor-acceptor (DA) and acceptor-only (AA), based on which laser excited the photon and which APD detected the photon. A: Example photon trace for the LAMP1-CLIPf-SNAPf fusion protein. B: Example photon trace for the CLIPf-SNAPf fusion protein.

experiment might result in the two distinct FRET efficiency populations currently not observed in the previous experiments. This would indicate the stability of the folded inactive conformation of the heavy chain of kinesin-1 only on the addition of microtubules and provide evidence that microtubules are necessary for the autoinhibition of kinesin-1.

Porcine microtubules were polymerised *in vitro* before addition to the kinesin smFRET assay buffer (see Section 2.7.1). Although HEK293 cells do produce microtubules they are at a much lower expression level to the kinesin-1 fusion protein and are depolymerised or pelleted along with other cellular debris in the cell lysis process. Therefore, *in vitro* microtubules need to be added to the assay buffer at high concentrations to ensure the high probability of the kinesin being attached to the microtubule during the assay. Due to poor expression levels of the CLIPf-hKIF5B-SNAPf fusion protein the CLIPf-mKIF5A-SNAPf fusion protein will be studied with and without the addition of microtubules.

Burst duration should be able to distinguish between bound and unbound kinesin-1. This is because bursts from unbound kinesin-1 should be similar in size and shorter then kinesin-1 bound to microtubules. The kinesin-microtubule complex will be much larger then just the kinesin-1 fusion protein so will take longer to diffuse through the confocal volume, in the smFRET instrument, resulting in a longer burst duration. Therefore data was filtered removing any bursts with a duration less then 0.75 ms.

Change in the FRET efficiency distribution were seen when microtubules were present (see Figure 5.7). There seems to be a slight polarisation of the data with the larger lower FRET population shifted closer to zero. The distribution of this population isn't as broad as the large low FRET efficiency population when microtubules aren't present. Also a small high FRET population seems to be emerging indicating that microtubules are needed to form the inactive closed conformation. More experiments need to be done including more detailed data analysis before a

new model of kinesin-1 autoinhibition, involving microtubules, can be put forward.

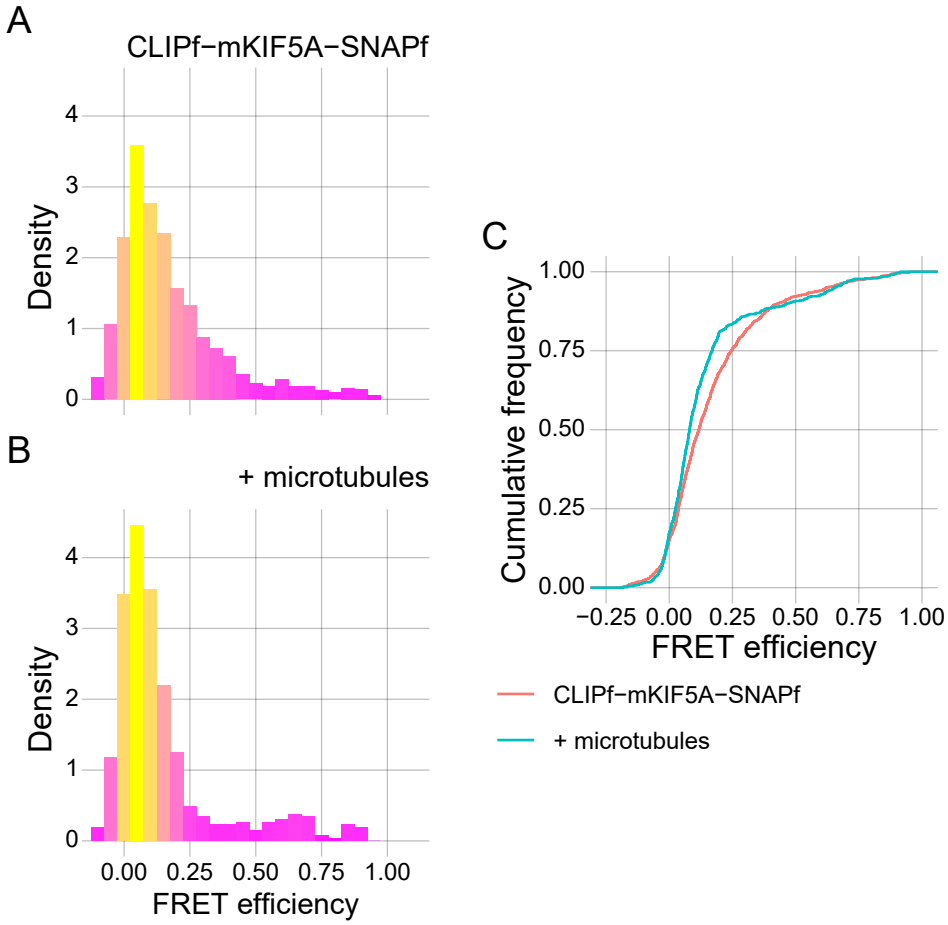

Figure 5.7: **smFRET results for CLIPf-mKIF5A-SNAPf fusion protein in the presence of microtubules**

Histograms of FRET efficiency for the co-transfection of CLIPf-mKIF5A-SNAPf fusion protein (A, n = 15, bursts = 1247) and in the presence of microtubules (B, n = 7, bursts = 529). Differences in FRET efficiency can be seen in cumulative frequency graph (C). Burst duration $\geq$= 0.75 ms.

5.3 Discussion

The minimum requirements for kinesin-1 activation have not been studied. Therefore it might be possible that the addition of cargo or microtubules are needed for kinesin-1 to undergo autoinhibition. This would make sense as unnecessary hydrolysis of ATP and kinesin-1 movement can only occur when microtubules are present. Previous methods of studying kinesin-1 activation have been unable to answer this question as microtubules were always present. The setup of this smFRET based assay allows the effect of individual components of axonal transport on kinesin-1 conformation and activation to be studied for the first time. The lack of a stable folded inactive conformation for the heavy chain of kinesin-1 in previous smFRET experiments could be due to the lack of one of these components.

This smFRET assay shows potential to study kinesin-1 conformation in the presence of microtubules and different cargoes. A new method of cell lysis, mechanical homogenisation, disrupted the cell membrane without also destroying other membranes within the cell. This was shown primarily through large bursts detected by the smFRET assay. Slight change in FRET efficiency were seen on addition of microtubules but, more detailed data analysis is needed (see Figure 5.7). In both cases, minor optimisation to the experimental protocol could result in a robust assay to investigate these effects. Due to time constraints this was not finalised in this thesis, but the basis for both assays has been proven.

A current limitation of this work is the data analysis of the smFRET data as it is currently designed for studying small molecules. Addition of microtubules or cargo, results in a mixed population with the tagged kinesin-1 heavy chain bound or unbound to these components; whilst also potentially being in different conformations. Bound proteins should result in a longer burst duration. This parameter was used when analysing the microtubule experiments to try and remove unbound kinesin-1 heavy chain from the analysis by selecting bursts which had a duration longer, or equal to 0.75 ms (see Figure 5.7). The Jupyter notebook currently prevents greater

filtering of the data using this parameter as the burst selection feature splits larger bursts into many smaller bursts. This is potentially happening to the smFRET data for LAMP1-CLIPf-SNAPf in Figure 5.6. When burst duration is not needed, this isn't a problem as it is a much more accurate way of calculating FRET efficiency and therefore detecting multiple conformations. These smaller bursts need to be combined, to enable differentiation between bound and unbound kinesin-1 heavy chains and allow better analysis of these more complex smFRET assays.

Any expression of the full length human KIF5B heavy chain isoform was extremely low when compared to the full length mouse KIF5A isoform (see Figure 4.4 and 5.2). This is probably due to a regulatory mechanism within the HEK293 cells as they endogenously express this protein. Expression was slightly better when more cells were used in the homogenisation experiments. Therefore protein purification might be necessary for these constructs to acquire accurate smFRET results. It would be beneficial to do a comprehensive comparison between cell lysate and purified proteins samples to determine if there are any differences in the FRET efficiency due to sample preparation.

Addition of the CLIPf- and SNAPf-tags to the LAMP1 protein, and overexpression of the fusion protein, didn't seem to result in mislocalisation of the protein (see Figure 5.3 and 5.4). Repetition of the confocal imaging is needed to ensure reproducibility of this result as currently only one biological repeat was performed for each experiment. Also live cell imaging of lysosomal trafficking with the two lysosomal markers, LysoTracker and the LAMP1 construct, would be helpful to ensure the added tags do not alter binding of kinesin-1 to the lysosomes. Similar experiments with the resultant homogenisation supernatant could be done to prove the movement of cargo by the overexpressed hKIF5B fusion protein. This would also show that the self-labelling enzymes are not affecting either the movement of kinesin-1 or its ability to recognise and bind to cargo.

Lysosomal bursts were seen by smFRET of LAMP1-CLIPf-SNAPf but only

singly labelled and in very low numbers (see Figure 5.6). The lysosomes were presumed to be intact as the duration of the resultant bursts were much longer then that of the CLIPf-SNAPf fusion protein. This is because of the massive size difference between the CLIPf-SNAPf fusion protein and the lysosome containing the LAMP1-CLIPf-SNAPf fusion protein. The larger size results in the tagged lysosome takes much longer to diffuse through the smFRET confocal volume resulting in much longer bursts. The number of detected bursts needs to be increased. This could be done by adding steps after homogenisation to enrich the sample for the lysosomal population; possibly through sucrose gradients. Also the high photon count of the detected bursts suggests that more then one labelled LAMP1-CLIPf-SNAPf fusion protein is present in the lysosome. Ideally lysosomes should only have one labelled kinesin-1 molecule attached as otherwise the results of the experiment would be hard to analyse conclusively. Therefore, the number of labelled kinesin-1 molecules attached to the lysosome needs to be checked and limited in this experiment. Although optimisation is needed these results show the potential of using this new smFRET assay to study the conformation of kinesin-1 when attached to endogenous cargo. In this case the LAMP1-CLIPf-SNAPf fusion protein would just be used as a positive control for intact lysosomes, and the FRET efficiency of the CLIPf-hKIF5B-SNAPf fusion protein would be measure with and without the addition of lysosomes.

The smFRET results for the CLIPf-mKIF5A-SNAPf fusion protein in the presence of microtubules suggests they are necessary for kinesin-1 autoinhibition (see Figure 5.7). Although more detailed data analysis is needed, emergence of a high FRET efficiency population and a general polarisation of the FRET efficiency was observed. The high FRET efficiency population is still relatively small though suggesting that the folded conformation is very transient. Although inactive folded kinesin-1 has been assumed to be the prevalent form in the cytosol (Hackney, Levitt, and Suhan 1992), kinesin-1 is a microtubule stimulated ATPase (Bloom et al. 1988 and Hackney 1994) meaning the folded inactive state is only necessary in the pres-

ence of microtubules. Therefore, it makes sense that the inactive conformation is only observed in the presence of microtubules but this had never been proven before. This autoinhibition model requires the presence of microtubules but also allows for the apparent transient nature of the folded state suggested by the small high FRET population. More in depth analysis of this data, along with additional experiments, are needed before proposing this new model of kinesin-1 autoinhibition, where inactivation only occurs on microtubules.

In the next chapter new smFRET standards will be designed and tested to calculate accurate distances from the FRET efficiencies. This will help in the validation process of the AlphaFold2 structure which needs to occur before the structure can be used.

Chapter 6

Investigating the potential of using smFRET to validate AlphaFold2 predictions

6.1 Introduction

6.1.1 Computational structural predictions

Currently it is not possible to experimentally determine all protein structures at a high resolution. Therefore, the structure of many proteins is unknown. The development of AlphaFold and AlphaFold2, by Deepmind, has finally resulted in an accurate protein structure prediction tool, for many proteins, after decades of work. AlphaFold is the first computational approach to predict the 3D structure of a protein, from their primary protein sequence, with near experimentally determined accuracy(Akdel et al. 2022).

Alphafold2 is a neural network that predicts the 3D structure of a protein based on multiple sequence alignments and training from the Protein Data Bank. Along with the structure, a confidence measure per residue is provided which is known as the predicted local-distance difference test (pLDDT). This confidence score can

be used to colour code the structure based on the expected accuracy of the model. Dark blue indicates regions of high accuracy (a pLDDT score of >90) and light blue shows areas with a pLDDT score between 70 and 90 which is expected to be modelled well. A pLDDT score of below 70 stipulates low confidence in the model. Areas are coloured yellow where the pLDDT score is between 50 and 70 and orange if the score is less than 50. These areas are probably disordered and are not accurately predicted by the model (see Figure 4.12, 6.12 - A and 6.13 - A). Another output of AlphaFold2 is the predicted aligned error (PAE). This shows the predicted error between residues for the prediction, and can be used to define domains within the protein. Although these confidence and error values help to interpret the resultant AlphaFold2 structure it is still only a prediction and needs to validated by experimental characterisation. Also these predictions don't provide information on conformations or dynamics of the protein (Akdel et al. 2022).

6.1.2 Using smFRET results to help validate structural predictions

Traditionally protein structure validation has been done experimentally by techniques like: x-ray crystallography, nuclear magnetic resonance spectroscopy (NMR) and cryo-EM. These have allowed near atomic resolution of protein structures that are rigid and/or can be crystallised. Extremely flexible proteins, for example kinesin-1, have very little structural information gathered by these techniques, meaning the predicted AlphaFold2 structures can't be validated in this way.

smFRET can be used to gain structural information about proteins. This can be for single conformations of proteins or for conformational changes. The FRET efficiency between two fluorophores is indirectly proportional to the distance between them. Therefore FRET efficiency results can be used to accurately calculate distances within the protein structure. An advantage of smFRET over techniques like x-ray crystallography and cyro-EM is the addition of temporal resolution. This

means that different conformations, and the transition between them, can be studied. Until recently wide spread use of smFRET for structural validation has been limited due to the unknown accuracy, reliability and reproducibility of the results. However by using well-characterised samples, known as FRET standards, smFRET results have been proven to be highly accurate, reliable and reproducible (Hellenkamp et al. 2018 and Agam et al. 2022).

6.1.2.1 Interfluorophore distances can be calculated from the FRET efficiency

FRET efficiencies can be used to calculate the precise distance between the two fluorophores. This is done using following equation (Equation 6.1) where E is the FRET efficiency, R_0 is the Föster radius and r is the distance between the two fluorophores (Hellenkamp et al. 2018).

$$r = R_0 \left(\frac{1 - E}{E} \right)^{\frac{1}{6}} \tag{6.1}$$

The Föster radius (R_0), is dependant on the pair of fluorophores used in the FRET experiment and their individual spectral photophysical properties. It is defined as the distance between the two fluorophore when the efficiency of energy transfer between them is 50 %. In order to calculate R_0 for a specific pair of FRET fluorophores four values are needed (see Equation 6.2). These are: the quantum yield of the donor fluorophore (Φ_D), the spectral overlap between the fluorophores (J), the dipole orientation factor (κ) and the refractive index of the medium between the fluorophores (n).

$$R_0 = 0.2108 \left(n^{-4} J \Phi_D \kappa^2 \right)^{\frac{1}{6}} \tag{6.2}$$

6.1.2.2 FRET standards are used across the smFRET community

FRET standards, well-characterized samples, are used within the smFRET community to ensure accuracy and reproducibility (Hellenkamp et al. 2018). This is especially important as most of the instruments are custom built there was difficulties in comparing results meaning smFRET results lacked reproducibility. To find out if the smFRET results could be compared, but also limit discrepancies due to variation in data acquisition and analysis, Hellenkamp *et al.* sent samples to 20 labs around the world, along with a generalised protocol and data analysis pipeline, and asked them to take part in a comparative blind study. These samples were a set of FRET standards comprising of double stranded DNA with either 15 or 23 base pairs between the fluorophores. The 23 base pair separated fluorophores resulted in low FRET whereas the 15 base pair DNA sample gave rise to a mid FRET population. Across the 20 labs a high consistency in FRET efficiency measurements was seen with the standard deviation being between 0.02 and 0.05 (Hellenkamp et al. 2018).

A similar study was recently done with protein samples (Agam et al. 2022)). Here 19 labs worldwide were sent two protein systems which were known to undergo conformational change. These proteins underwent site specific cysteine labelling meaning other naturally occurring solvent accessible cysteines were mutated. Again a high consistency across the labs was seen with a standard deviation of 0.06. Although this study showed the accuracy, reliability and reproducibility of protein based smFRET measurements, protein FRET samples of fixed known length were not used (Agam et al. 2022). The addition of these samples to the field would allow easier comparison across different instruments as the FRET efficiency doesn't result from the current conformation of a protein which can be altered by many conditions such as salt concentration and pH.

The aims of this chapter are to design protein FRET samples of fixed known length across the FRET efficiency range and to investigate whether smFRET has the potential to validate AlphaFold2 predictions.

6.2 Results

6.2.1 inflexible protein FRET standards

6.2.1.1 Design of inflexible protein FRET standards

A simple set of protein based FRET standards are needed to help measure, and ensure, accuracy and reliability of smFRET results across the community. As we had good smFRET results using self-labelling enzymes and organic dyes, this labelling strategy was chosen. This would give the better photophysical properties desired in single molecule measurements, compared to fluorescent proteins, without the need to mutate residues for site specific cysteine labelling. Labelled inflexible protein samples would prevent changes in FRET efficiency due to conformational change, which is currently possible with the standard protein systems used in Agam et al. 2022. These new standards could also be used across multiple systems (TIRF or confocal smFRET) and with different sample preparations methods; whole cell imaging, cell lysate or purfied protein.

The tetratricopeptide repeat (TPR) domain was chosen as a repeatable unit of known length for the protein based FRET standards. This motif naturally occurs in a wide array of different proteins, including the kinesin light chains, and mediates a diverse range of protein-protein interactions. One TPR motif forms two stacked antiparallel α-helices. This motif is repeated, different amounts in different proteins, to form a "spiral staircase" structure. In 2003, Main *et al.* designed a stable α-helical arrays, based on an "idealised TPR motif", to help understand how folding and function of the TPR domain is altered by the amino acid sequence. The consensus TPR motif was designed using the global propensity of each amino acid for each position in the TPR motif based on a database of 1837 naturally occurring TPR sequences.

The structure of novel proteins, with one, two and three repeats of the designed TPR motif were characterised by x-ray crystallography and NMR. All three novel

proteins folded as predicted as well as being stable and monomeric (Main et al. 2003). Therefore this designed TPR motif was used as a basis of inflexible FRET protein standards.

CLIPf/SNAPf fusion proteins of these three proteins (TPR1, TPR2 and TPR3) were designed and codon optimised for bacterial and mammalian expressions systems. This is to allow comparison between purified and cell lysate protein samples. The designed TPR motif by Main *et al.* started with a three amino acid N-cap before the two helices of the repeatable motif. A solvating helix was added at the C-terminus to to help with solubility of the proteins and to mimic naturally occurring TPR domains in proteins. The structure of the TPR3 can be seen in Figure 6.1 - A (Main et al. 2003). With a CLIPf-SNAPf fusion protein which contained no TPR motif, this resulted in four protein FRET standards (CLIPf-SNAPf, CLIPf-TPR1-SNAPf, CLIPf-TPR2-SNAPf and CLIPf-TPR3-SNAPf) which should have increasing FRET efficiency values that span the entire FRET efficiency range. The CLIPf-SNAPf construct was slightly different to the one previously used in Chapter 3. These CLIPf-SNAPf constructs have a N-terminal FLAG tag and a C-terminal StrepTag for protein purification and immuno assays.

The four mammalian expression constructs (CLIPf-SNAPf, CLIPf-TPR1-SNAPf, CLIPf-TPR2-SNAPf and CLIPf-TPR3-SNAPf) were purchased and synthesised by GeneArt (invitrogen by Thermo Fisher Scientific), along with the TPR3 construct for bacterial expression (CLIPf-TPR3-SNAPf). The other three bacterial expression constructs were made by deletion mutagenesis of the purchased CLIPf-TPR3-SNAPf construct. From the N-terminus to the C-terminus this construct contained: a FLAG tag, a CLIPf-tag, the designed N-cap, three repeats of the designed TPR motif, the designed solvating helix, a SNAPf-tag and a double StrepTag. Deletion of one and two TPR motifs resulted in the CLIPf-TPR2-SNAPf and the CLIPf-TPR1-SNAPf constructs respectively. A new CLIPf-SNAPf construct was made by deletion of: the N-cap, all three repeats of the TPR motif and the solvating helix. This meant there

was no linker between the CLIPf- and SNAPf-tag unlike the previous CLIPf-SNAPf construct (see Section 2.1.1.4 for details). The resultant FRET standard constructs can be seen in Figure 6.1 - B.

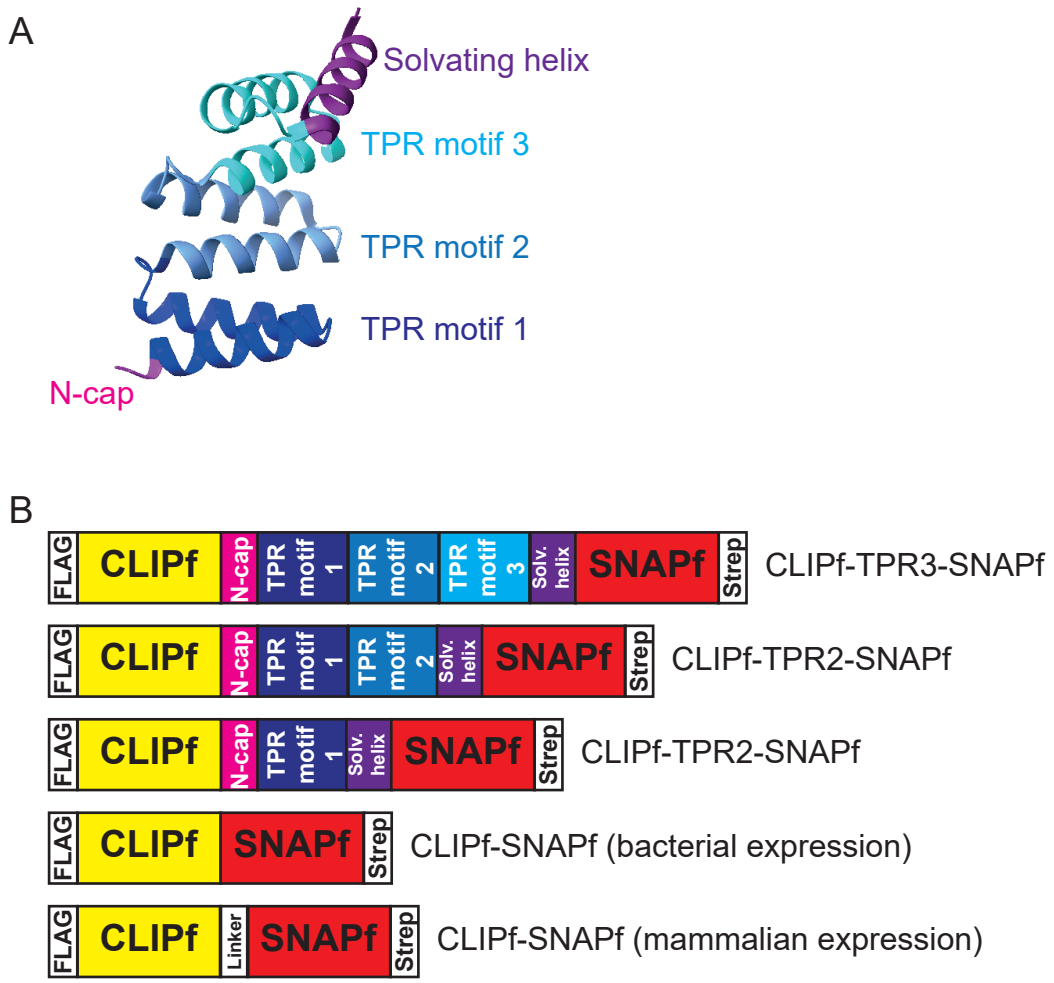

Figure 6.1: **TPR constructs for use as new FRET standards**

A: Protein structure of TPR3 in Main et al. 2003 (PBD ID: 1NA0) B: Cartoon schematics of the four TPR constructs; CLIPf-TPR3-SNAPf, CLIPf-TPR2-SNAPf, CLIPf-TPR1-SNAPf and CLIPf-SNAPf. There are two different CLIPf-SNAPf constructs as the one codon optimised for mammalian expression includes a small linker (GSAGSAAGSGEF, Waldo et al. 1999) whereas the one codon optimised for bacterial expression does not (see Section 2.1.1.4 for details).

6.2.1.2 Protein purification of the TPR constructs from *E. coli*

To maximise protein labelling, the bacterial expressed fusion proteins were purified by StrepTag affinity and then labelled with the CLIP-Cell TMR-Star and SNAP-Cell 647-SiR fluorophores whilst bound to the column (see Section 2.8). Purification and labelling of the protein was shown by in gel fluorescence and the elution peak fractions were pooled. The protein was then concentrated to have a correct volume for size exclusion chromatography. This was needed to remove the biotin (used to elute during StrepTag affinity purification) and any chaperone proteins. Again elution fractions were pooled and purification and labelling was checked by in gel fluorescence.

The new CLIPf-SNAPf fusion protein (see Figure 6.2 and 6.3) along with the TPR1 (see Figure 6.4 and 6.5) and TPR3 (see Figure 6.6 and 6.7) fusion proteins were purified using this approach. All three proteins were purified and doubly labelled but showed an additional band after size exclusion chromatography that wasn't present after StrepTag affinity purification. As this is only seen in the coomassie stain and not in the in gel fluorescence it was thought that it wouldn't cause any problems with the smFRET assay so the samples were used.

6.2.1.3 smFRET results for engineered TPR proteins expressed and purified from *E. coli*

To determine the FRET efficiency for each fusion protein, the labelled purified protein samples were diluted in TPR smFRET assay buffer (see table 2.7). This assay buffer was different to the buffer previously used to dilute cell lysate samples as some components weren't necessary for the TPR fusion protein; for example ATP. The TPR smFRET assay buffer is similar to the initial smFRET assay buffer but is Tris based rather then HEPES. This is because Craggs' lab members had issues with HEPES causing a lot of background fluorescence. In all other aspects the previously optimised data acquisition and analysis protocols were used.

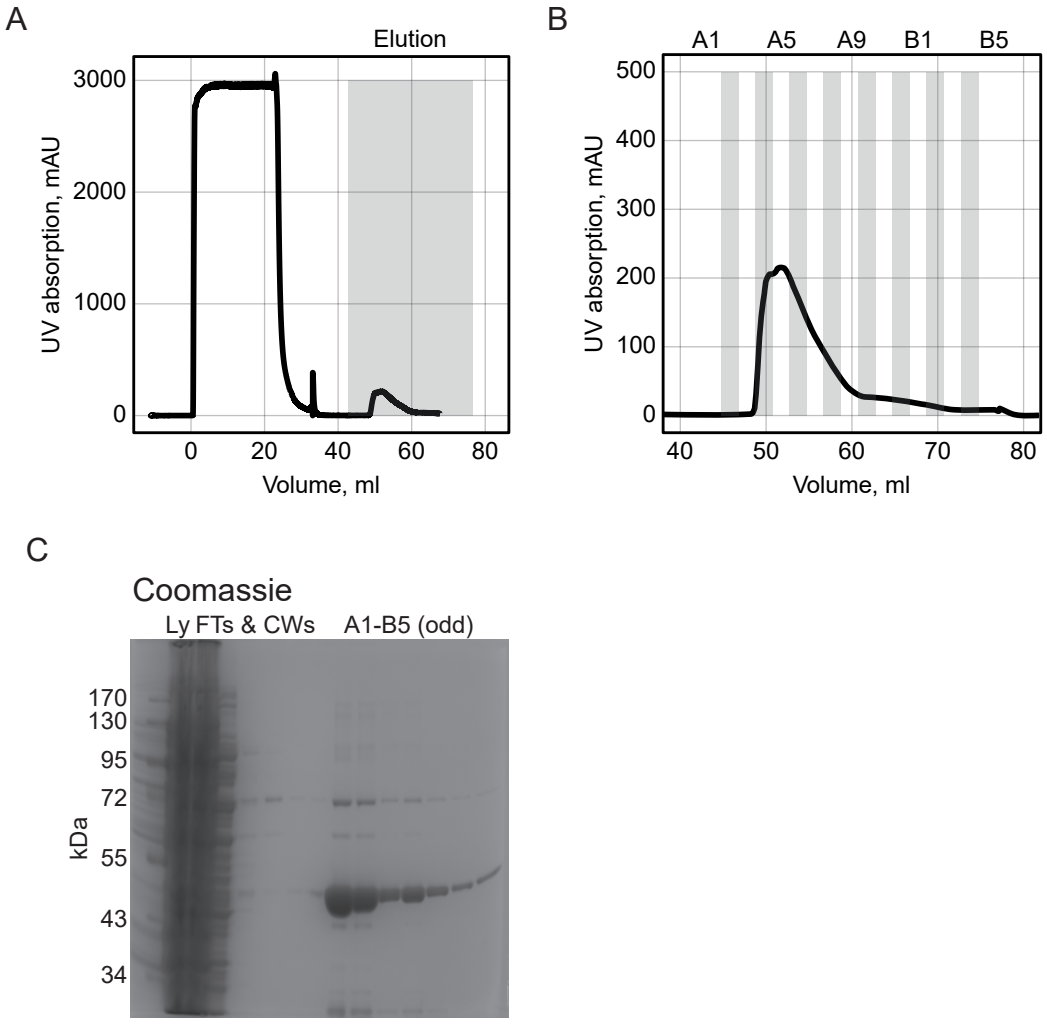

Figure 6.2: **Purification of bacterial expressed CLIPf-SNAPf fusion protein by StrepTag affinity**

A: UV absorption profile of StrepTag affinity purification. B: Expansion of UV absorption profile of elution highlighted in A. C: Coomassie stained SDS-PAGE gel of StrepTag affinity purification fractions. Ly = cell lysate, FTs = the flow through after sample and after ligand injection onto the column and CWs = the wash steps after sample and after ligand injection onto the column. A1-B5 (odd) = the odd number fractions from the elution seen in B.

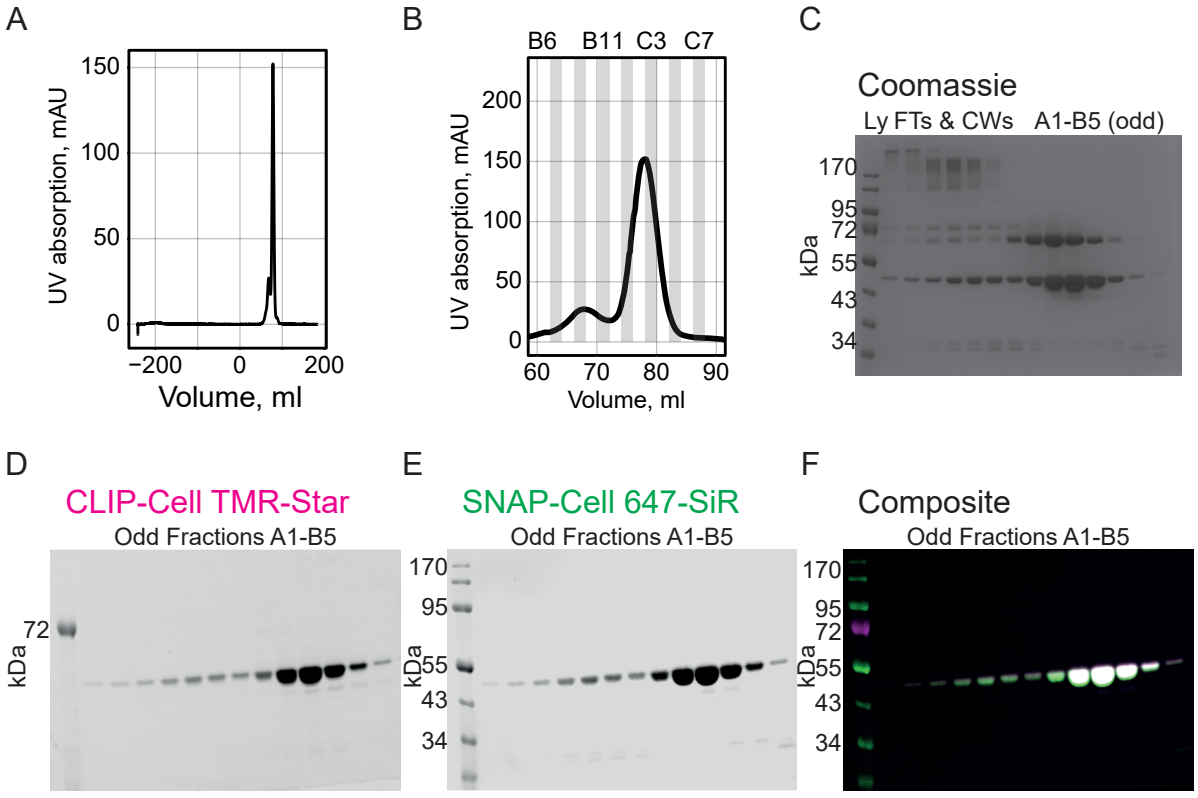

Figure 6.3: **Purification of bacterial expressed CLIPf-SNAPf fusion protein by size exclusion chromatography**

A: UV absorption profile of size exclusion chromatography. B: Expansion of UV absorption profile of elution. C: Coomassie stained SDS-PAGE gel of fractions. Ly = cell lysate, FTs = the flow through after sample and after ligand injection onto the column and CWs = the wash steps after sample and after ligand injection onto the column. A1-B5 (odd) = the odd number fractions from the elution seen in B. D: 600 nm channel imaging CLIP-Cell TMR-Star fluorescence. E: 700 nm channel imaging SNAP-Cell 647-SiR fluorescence. F: Composite image of both channels with the 600 nm channel shown in magenta and the 700 nm channel in green; overlap appears white.

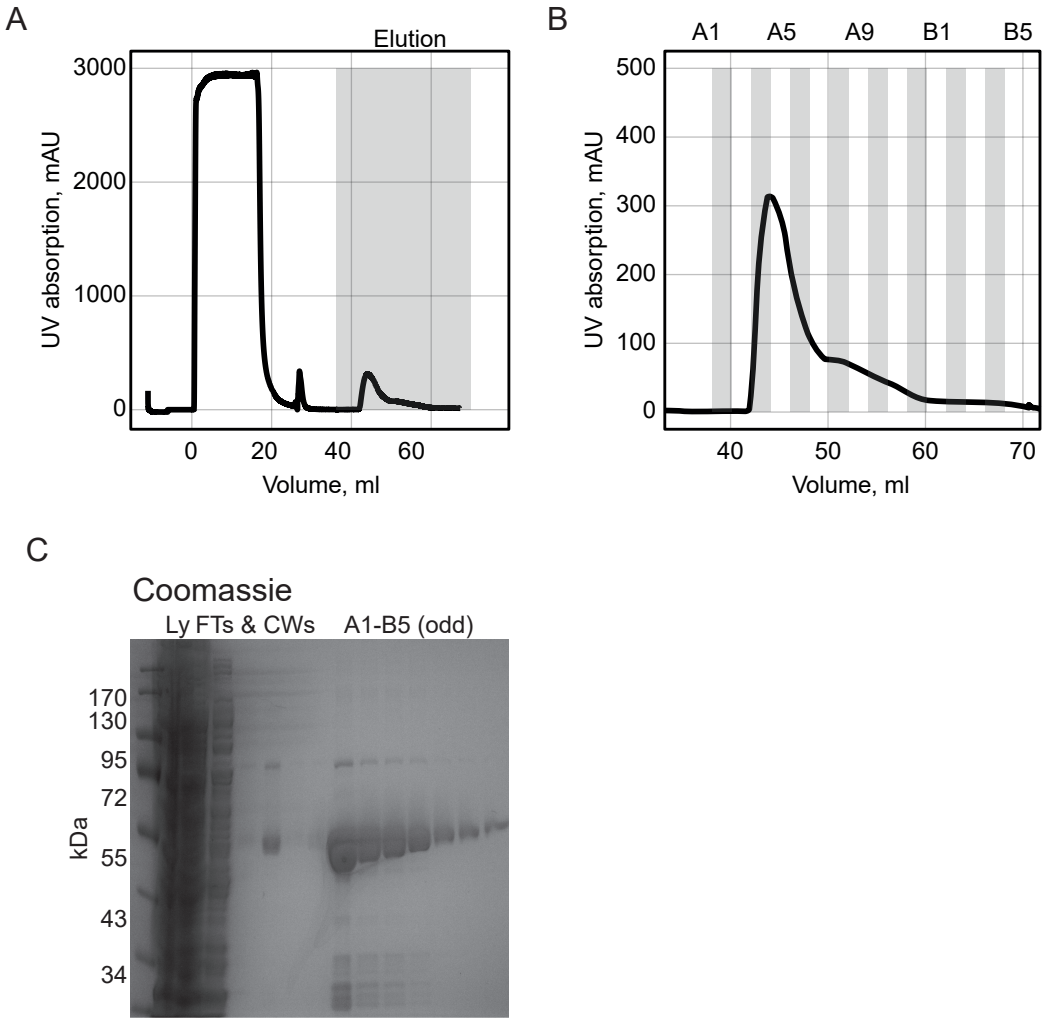

Figure 6.4: **Purification of bacterial expressed CLIPf-TPR1-SNAPf fusion protein by StrepTag affinity**

A: UV absorption profile of StrepTag affinity purification. B: Expansion of UV absorption profile of elution highlighted in A. C: Coomassie stained SDS-PAGE gel of StrepTag affinity purification fractions. Ly = cell lysate, FTs = the flow through after sample and after ligand injection onto the column and CWs = the wash steps after sample and after ligand injection onto the column. A1-B5 (odd) = the odd number fractions from the elution seen in B.

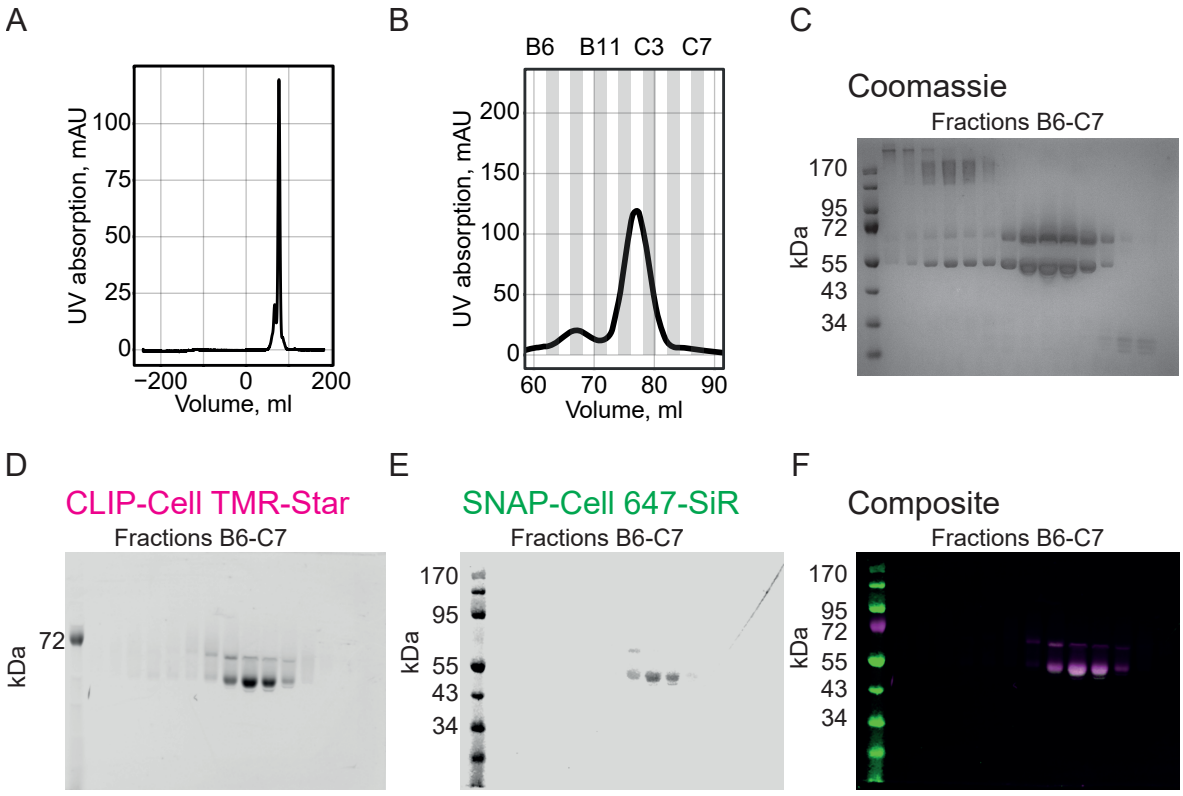

Figure 6.5: **Purification of bacterial expressed CLIPf-TPR1-SNAPf fusion protein by size exclusion chromatography**

A: UV absorption profile of size exclusion chromatography. B: Expansion of UV absorption profile of elution. C: Coomassie stained SDS-PAGE gel of fractions. Ly = cell lysate, FTs = the flow through after sample and after ligand injection onto the column and CWs = the wash steps after sample and after ligand injection onto the column. A1-B5 (odd) = the odd number fractions from the elution seen in B. D: 600 nm channel imaging CLIP-Cell TMR-Star fluorescence. E: 700 nm channel imaging SNAP-Cell 647-SiR fluorescence. F: Composite image of both channels with the 600 nm channel shown in magenta and the 700 nm channel in green; overlap appears white.

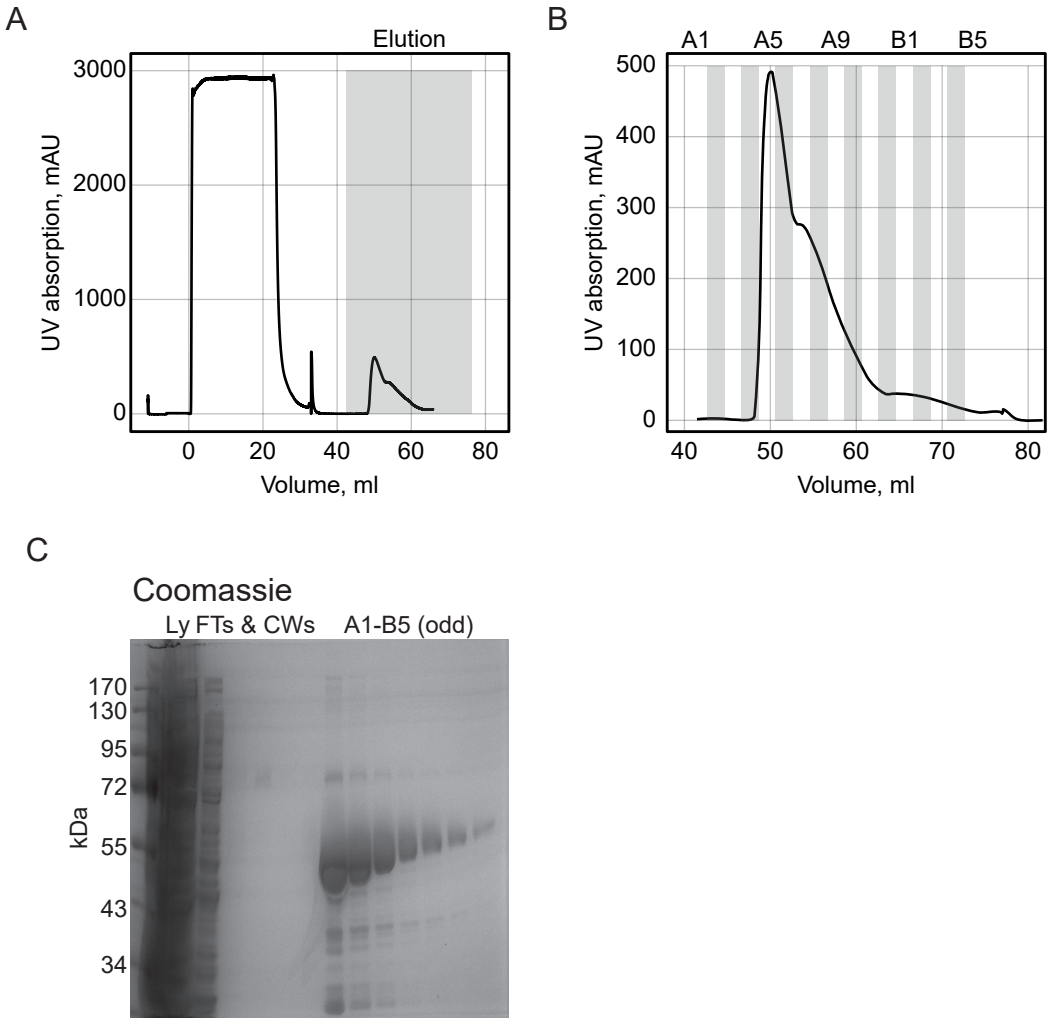

Figure 6.6: **Purification of bacterial expressed CLIPf-TPR3-SNAPf fusion protein by StrepTag affinity**

A: UV absorption profile of StrepTag affinity purification. B: Expansion of UV absorption profile of elution highlighted in A. C: Coomassie stained SDS-PAGE gel of StrepTag affinity purification fractions. Ly = cell lysate, FTs = the flow through after sample and after ligand injection onto the column and CWs = the wash steps after sample and after ligand injection onto the column. A1-B5 (odd) = the odd number fractions from the elution seen in B.

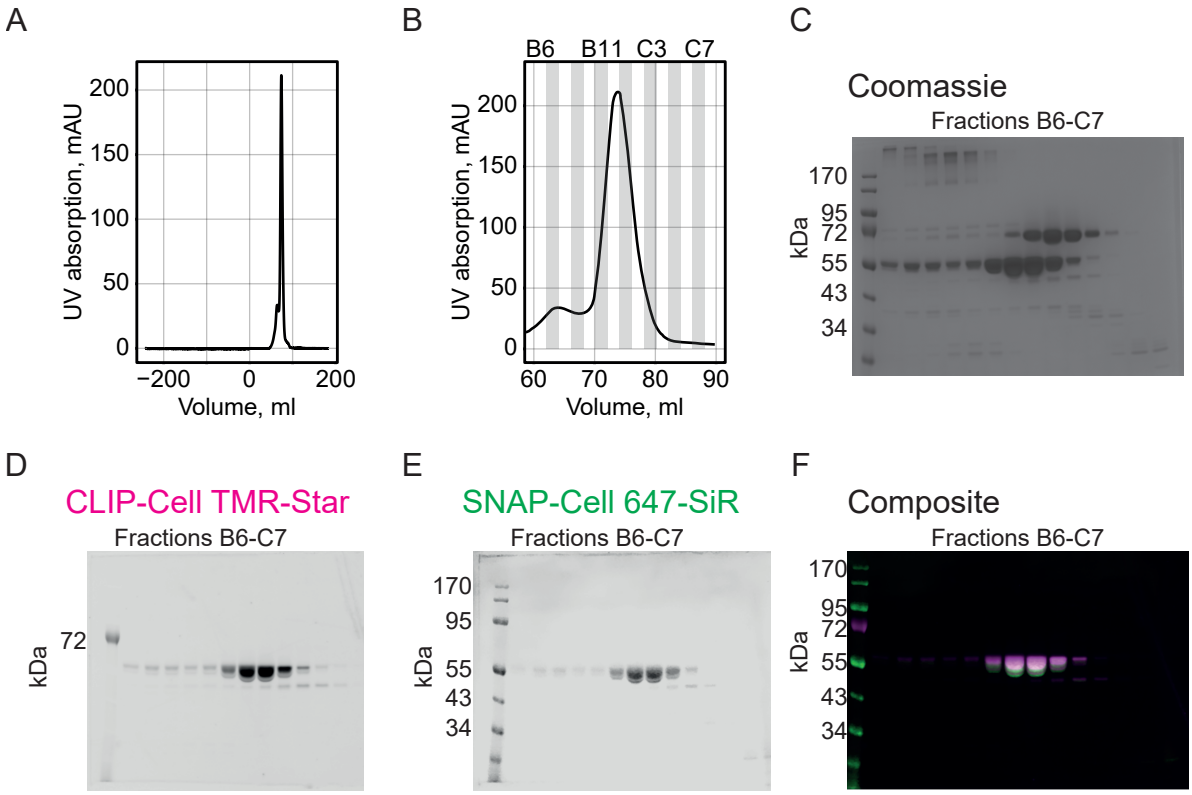

Figure 6.7: **Purification of bacterial expressed CLIPf-TPR3-SNAPf fusion protein by size exclusion chromatography**

A: UV absorption profile of size exclusion chromatography. B: Expansion of UV absorption profile of elution. C: Coomassie stained SDS-PAGE gel of fractions. Ly = cell lysate, FTs = the flow through after sample and after ligand injection onto the column and CWs = the wash steps after sample and after ligand injection onto the column. A1-B5 (odd) = the odd number fractions from the elution seen in B. D: 600 nm channel imaging CLIP-Cell TMR-Star fluorescence. E: 700 nm channel imaging SNAP-Cell 647-SiR fluorescence. F: Composite image of both channels with the 600 nm channel shown in magenta and the 700 nm channel in green; overlap appears white.

All three fusion proteins (TPR3, TPR1 and CLIPf-SNAPf) resulted in a large low FRET population (see Figure 6.8). It was predicted that the both the shorter constructs, CLIPf-SNAPf and CLIPf-TPR1-SNAPf, should give a higher FRET efficiency population with the TPR1 construct being slightly lower in FRET efficiency. Although a small high FRET population can be seen for the CLIPf-SNAPf fusion protein, the vast majority of bursts resulted in a low FRET efficiency population.

Bands on the coomassie stain and the in gel fluorescence did indicate that three different proteins were purified. The molecular weight of each construct seemed correct on the SDS-PAGE, with the CLIPf-SNAPf fusion protein having the lowest molecular weight and the TPR3 fusion protein having the highest. This could also be seen in the size exclusion chromatography as the elution peak for each fusion protein was slightly shifted depending on the molecular weight. The TPR3 fusion protein eluted quicker (fractions B12 and C1) then the TPR1 fusion protein (fractions C1-C3) which in turn eluted quicker then the CLIPf-SNAPf fusion protein (fractions C2 and C3). But this change in size of the protein was not seen in the smFRET experiments.

6.2.1.4 smFRET results for engineered TPR proteins expressed from mammalian cells

It was expected that the increase in TPR motifs in the construct would result in the CLIPf- and SNAPf-tags being further apart and therefore resulting in a lower FRET efficiency. To check the constructs could result in the desired structures the mammalian codon optimised versions of the four constructs were tested. If they gave different results then it is clear something went wrong during the purification process.

To ensure the previous smFRET results were correct for the four fusion proteins, HEK293 cells were transfected with the four constructs that were codon optimised for mammalian expression. These constructs are: FLAG-CLIPf-SNAPf-Strep,

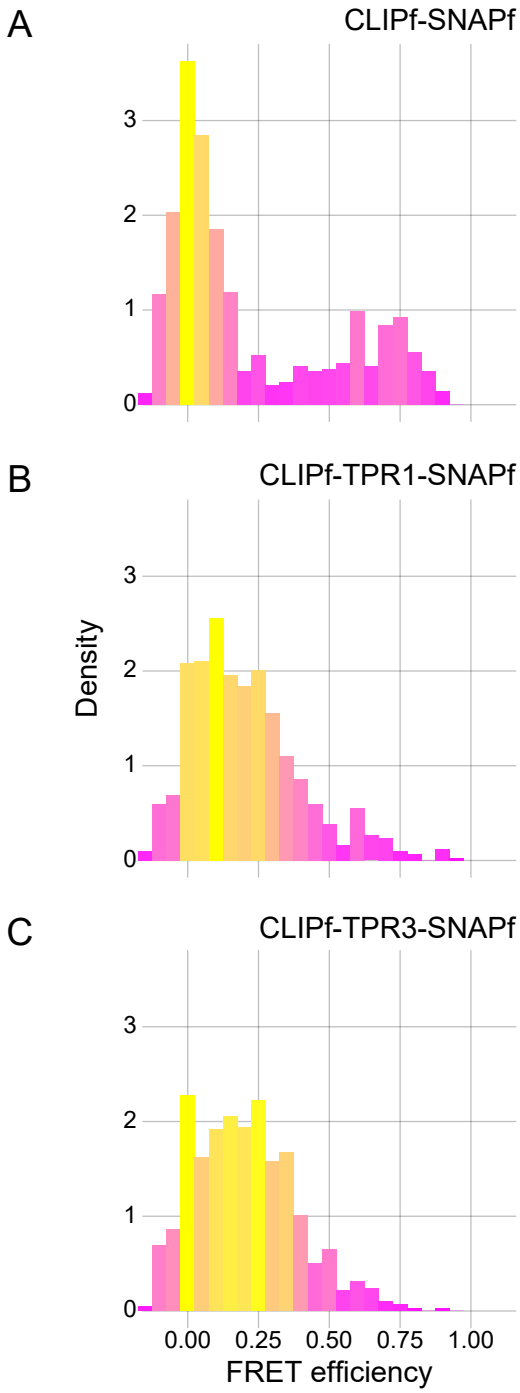

Figure 6.8: **FRET efficiency results for engineered TPR proteins expressed and purified from *E. coli*.**

FRET efficiency histograms for the CLIPf-SNAPf (A, n = 1, bursts = 689), CLIPf-TPR1-SNAP (B, n = 1, bursts = 836), CLIPf-TPR3-SNAPf (C, n = 1, bursts = 837).

FLAG-CLIPf-TPR1-SNAPf-Strep, FLAG-CLIPf-TPR2-SNAPf-Strep and FLAG-CLIPf-TPR3-SNAPf-Strep. The protein sequence is exactly the same as the fusion protein previously purified apart for the CLIPf-SNAPf construct which contains a small linker (GSAGSAAGSGEF, Waldo et al. 1999).

The four fusion proteins expressed by the HEK293 cells resulted in different FRET efficiencies which decreased with increased number of TPR motifs (see Figure 6.9). As predicted the CLIPf-SNAPf fusion protein resulted in the highest FRET efficiency and the TPR3 fusion protein with the smallest. By plotting the cumulative frequency of the FRET efficiencies of the four fusion proteins the difference in FRET efficiency can be seen more clearly, with the gap in FRET efficiencies between the TPR fusion proteins being quite even. All four fusion proteins did show the presence of a small low FRET population, which needs to be reduced through assay optimisation before wide use of these constructs as FRET standards.

The FRET efficiencies covered nearly the whole range of possible FRET efficiencies meaning these proteins could act as good ladder of protein distances and therefore, inflexible protein FRET standards. More biological repeats are needed before this result in conclusive but it shows the possibility of using TPR based proteins for FRET standards.

6.2.2 Investigating the central hinge region in the heavy chains of kinesin-1

6.2.2.1 Design of central hinge mKIF5B constructs

One of the major unexpected structures in the AlphaFold2 structure prediction was the central hinge region of the kinesin-1 heavy chain stalk. This region has been highlighted since 1987 and shown to have an impact on the overall activity of the kinesin-1 molecule by a decrease in ATPase activity (Amos 1987 and Friedman and R D Vale 1999). It is thought to be an unstructured disordered region in the

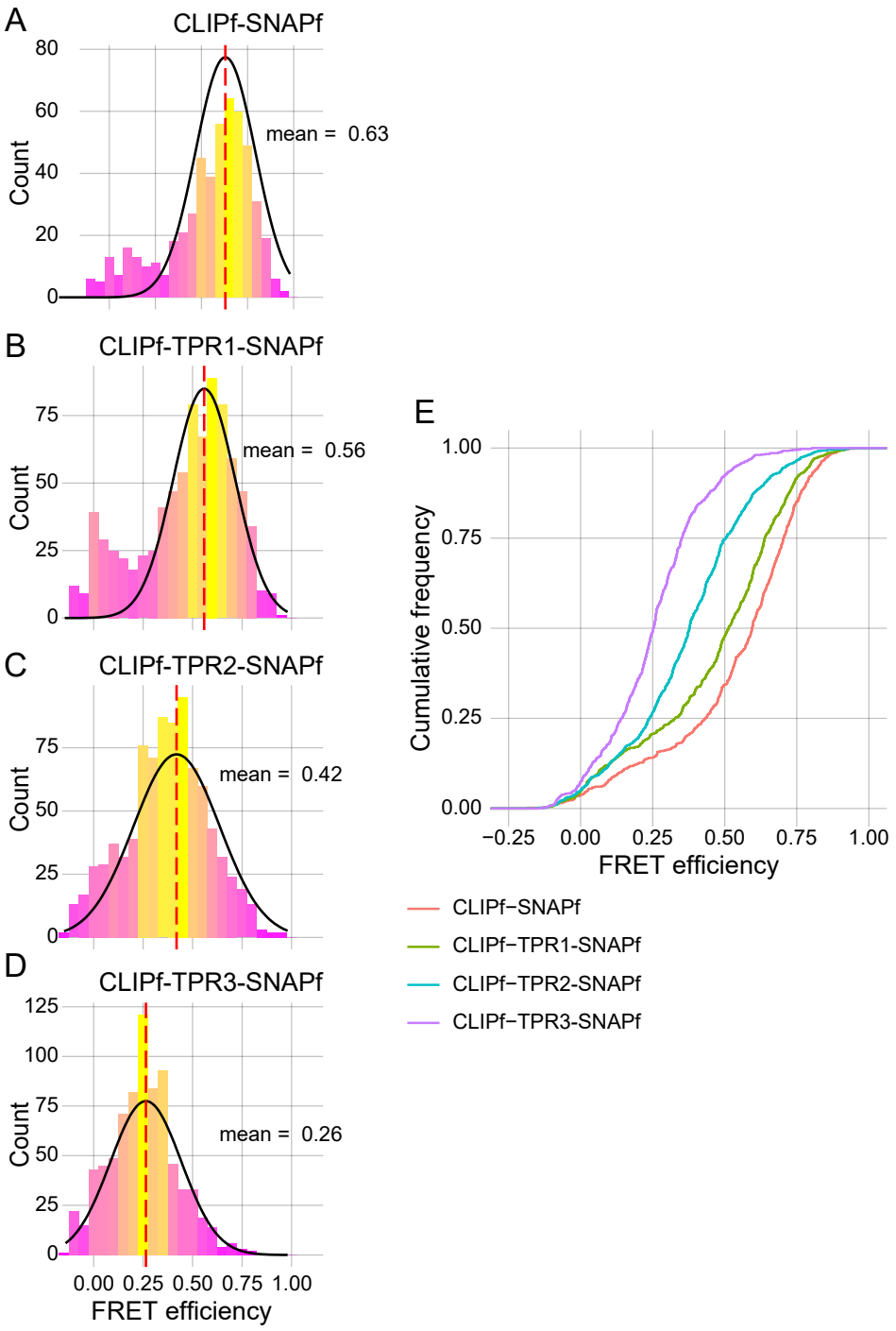

Figure 6.9: **FRET efficiency results for engineered TPR proteins expressed from mammalian cells** FRET efficiency histograms for the CLIPf-SNAPf (A, n = 1, bursts = 528), CLIPf-TPR1-SNAP (B, n = 1, bursts = 820), CLIPf-TPR2-SNAP (C, n = 1, bursts = 876), CLIPf-TPR3-SNAPf (D, n = 1, bursts = 786). Differences in FRET efficiency can be seen in cumulative frequency graph (E). For each histogram, the data was fitted to two gaussian populations, with that describing the largest population shown. The mean of this gaussian is indicated by the dashed red line and the value written beside.

coiled coil stalk, but there has been no in depth structural investigations into this region. AlphaFold2 predicts a interlocking barrel structure which seems to suggest the possibility of a more complicated regulatory mechanism in this region. However, the AlphaFold2 structure needs to be experimentally validated.

Four KIF5B hinge fragments were chosen, using paircoil2, in order to investigate the hinge region of kinesin-1 (see Figure 6.10 - A). Following alignment of the three mouse heavy chain isoforms (KIF5A, KIF5B and KIF5C), a proline was identified as the possible middle of the hinge region. This is present for both the KIF5A (P573) and the KIF5B (P577) isoforms but not in the KIF5C isoform where alignment shows a leucine. This suggests there maybe some isoformal differences in kinesin-1 activation. Paircoil2 was used for disordered structure and heptad structure prediction. For KIF5B, the region 545-593 was predicted to be disordered so this was selected as the first hinge region to be investigated. A region half this size around proline 577 was selected as the second and smallest hinge fragment. Two larger hinge regions were chosen based on the heptad prediction either side of the larger disordered region. A fragment with two heptads either side and one with four heptads either side was chosen. This gave four hinge fragments to be explored which can be seen in Figure 6.10 - B. These four fragments were synthesised by GeneArt with a BamHI and EcoRI restriction sites at the N- and C-terminus respectively.

These four fragments were then inserted into the CLIPf-SNAPf vector (see 2.1.1.1) in between the CLIPf- and SNAPf-tag using double restriction digest of the fragments and vector (see 2.1.2.4) followed by gel extraction of the two fragments and then ligation (see 2.1.2.5 and 2.1.2.6). This resulted in the four hinge fragments tagged at the N-terminus with CLIPf-tag and the C-terminus with SNAPf-tag with no linkers in between. After the SNAPf-tag there are also a polyhistidine-tag and a FLAG-tag.

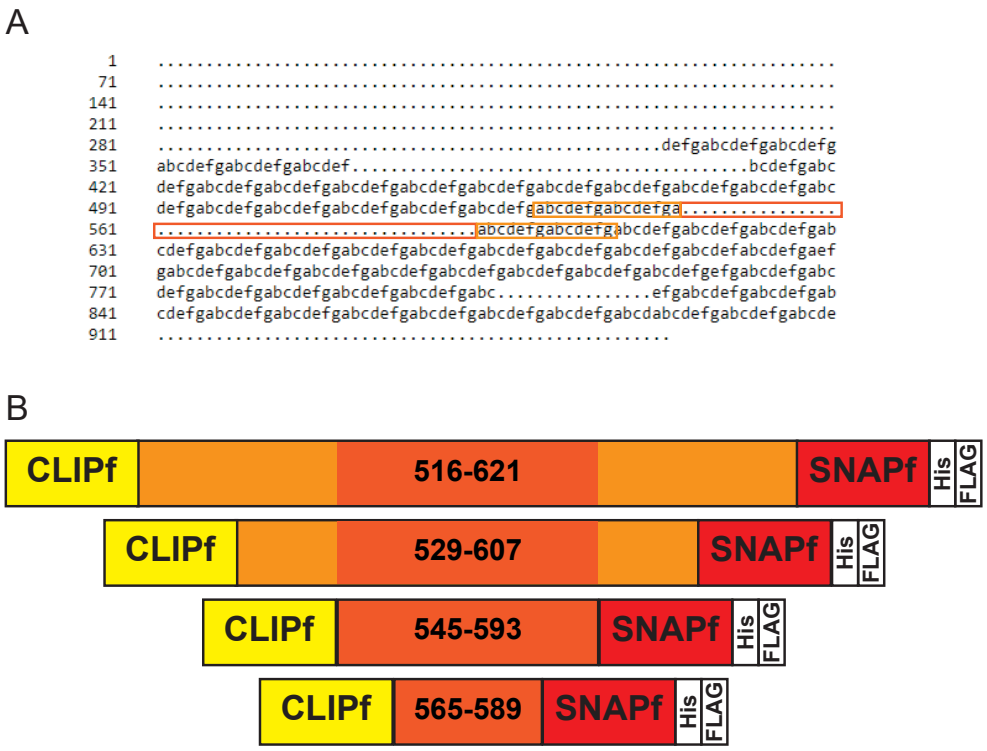

Figure 6.10: **mKIF5B central hinge constructs**

A: Paircoil2 prediction for mouse KIF5B. Dark orange box indicates the disordered region (545-593) and light orange box shows the two complete heptads either side (529-607). B: Cartoon schematics of the four mKIF5B central hinge constructs.

6.2.2.2 Dimerisation state of the mKIF5B central hinge constructs by Native-PAGE

The heavy chains of kinesin-1 exists as a dimer so the state of the hinge fragments, whether they exist as a dimer or monomer, needs to be determined. This was done by native SDS-PAGE followed by western blot analysis using the α-FLAG rabbit antibody. Native-PAGE differs from SDS-PAGE by the lack of denaturation of the protein samples. The samples are not boiled and the sample buffer used does not contain detergent. Therefore the oligomerisation state of the protein complex can be kept resulting in altered migrtation in the gel.

The longest hinge fusion protein, CLIPf-KIF5B(516-621)-SNAPf-HisTag-FLAG, seem to be able to stably exist as a dimer as two bands were clearly seen in the western blot (see Figure 6.11 - A). Dimers band can also be seen for the two next longest hinge fragments but the shortest hinge fusion protein, CLIPf-KIF5B(565-589)-SNAPf-HisTag-FLAG, only resulted in one clear band suggesting it is mainly exists as a monomer. Quantification of relative band intensity showed the increasing likelihood of dimer formation with increasing length (see Figure 6.11 - B).

6.2.2.3 smFRET of the mKIF5B central hinge constructs confirm dimerisation states

It was shown previously that this smFRET assay can distinguish between monomers and dimers based on the resultant stoichiometry of the bursts (see Section 4.2.3). This is due to the labelling efficiency of CLIP-Cell TMR-Star and the CLIPf-tag compared to SNAP-Cell 647-SiR and the SNAPf-tag, resulting in the majority of bursts having more of an acceptor contribution then a donor. When a fusion protein is monomeric these types of bursts can not result in FRET so are filtered out as acceptor-only bursts. This gives monomeric fusion proteins a mean stoichiometry centred around 0.5 whereas dimeric fusion proteins have a lowered mean stoichiometry centred around 0.4 (see Figure 4.5). Therefore this smFRET assay can be used

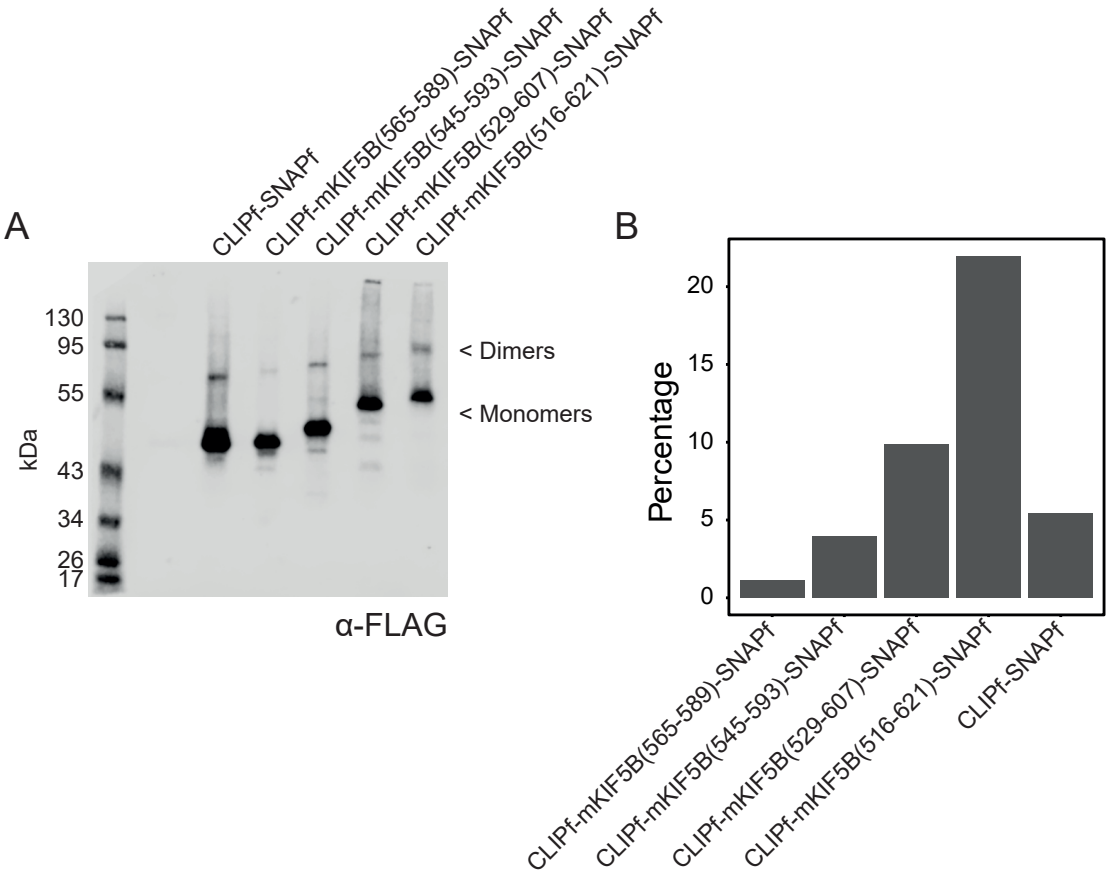

Figure 6.11: **Native PAGE and western blot of central hinge constructs of mKIF5B**

A: Samples were run on a native PAGE gel and analysed by western blotting with the α-FLAG rabbit antibody. B: Quantification of the percentage of band intensity, in A, for the dimer bands compared to the monomers bands. For each sample the signal of the dimer band was divided by that for the monomer.

to confirm the previous native-PAGE results.

To confirm the hinge fragments oligomerisation state, HEK293 cells were transfected with the two of the four hinge; CLIPf-KIF5B(516-621)-SNAPf-HisTag-FLAG and CLIPf-KIF5B(565-589)-SNAPf-HisTag-FLAG. The longest hinge fragment resulted in a lower stoichiometry then the shortest hinge fragment, which was more centred around S = 0.5 (see Figure 6.12). This can also be seen as a shift to the left in the culmulative frequency plot (see Figure 6.12-C). This agrees with the native-PAGE results that CLIPf-KIF5B(516-621)-SNAPf-HisTag-FLAG exists as a dimer whereas CLIPf-KIF5B(565-589)-SNAPf-HisTag-FLAG is monomeric.

There is also a difference in FRET efficiency which was expected due to the difference in size of the two fragments (see Figure 6.13). The larger hinge fragment result in a very broad spread in FRET efficiency over most of the FRET efficiency range. This is probably due to structure of this hinge fragment. The AlphaFold2 predicts the interlocking barrel structure with all tags being within the FRET efficiency range. This leads to a mix of FRET efficiency; one as a result of the two labelled tags on the same kinesin-1 heavy chain and one from the labelled tags on different heavy chains. These tags pair are different distance apart resulting in different FRET efficiencies. As the longest hinge fragment exists as a dimer this could be a good fragment to start investigating the structure of this region and whether it is similar to that predicted by AlphaFold2.

6.3 Discussion

To use smFRET as a technique to validate the AlphaFold2 predicted structure and results need to accurate and reliable. This can be determined by the use of protein FRET standards. The DNA smFRET community already routinely use DNA FRET standards to ensure accuracy and reproduciblity of results across many different instrumental setups. These standards are double labelled double stranded DNA

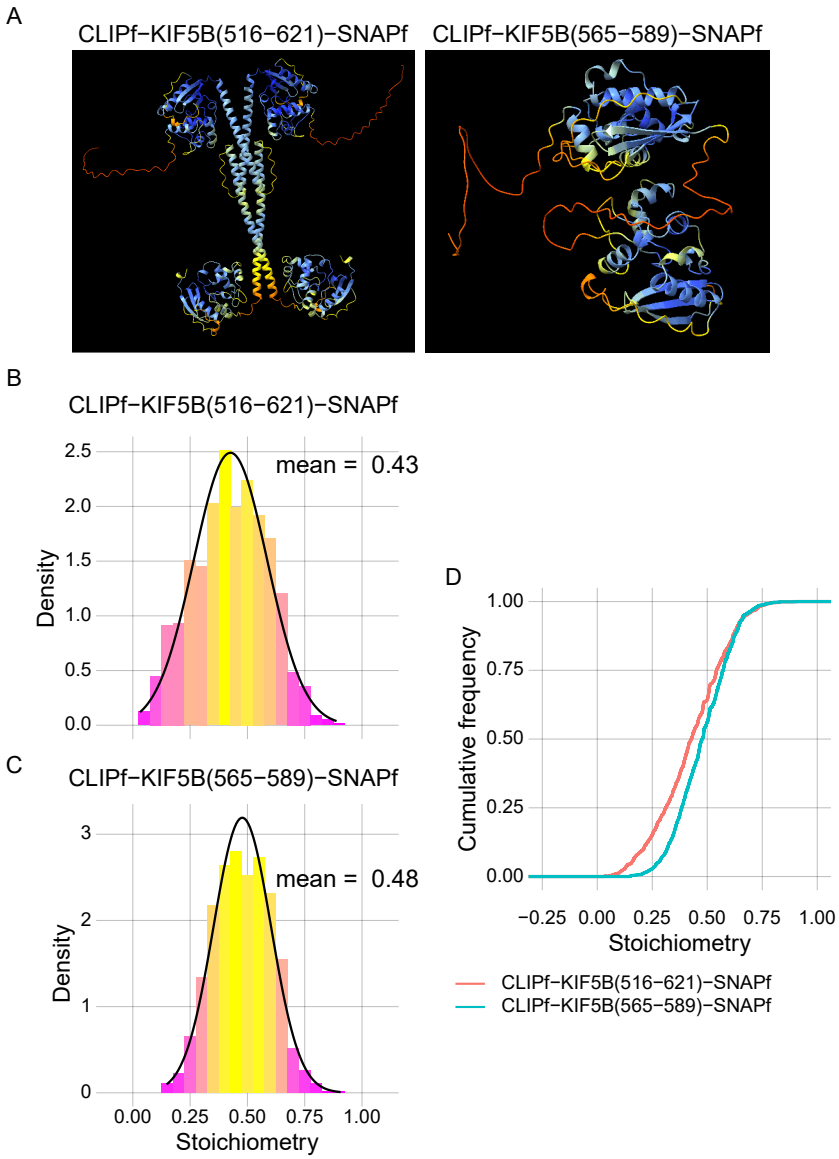

Figure 6.12: **FRET stoichiometry of CLIPf-mKIF5B(516-621)-SNAPf and CLIPf-mKIF5B(565-589)-SNAPf fusion proteins**

A: AlphaFold2 predictions for the two tagged fusion proteins. Structure coloured by pLDDT score: dark blue = 90-100, light blue = 70-90, yellow = 50-70 and orange = 0-50. A high pLDDT indicates high confidence in the structure. Histograms of FRET stoichiometry for the CLIPf-mKIF5B(516-621)-SNAPf fusion protein (B, n = 5, bursts = 1114) and the CLIPf-mKIF5B(565-589)-SNAPf fusion protein (C, n = 2, bursts = 1242). Black line indicates a normal distribution with the mean, or centre of the distribution, written beside. Differences in FRET stoichiometry can be seen in cumulative frequency graph (D), where CLIPf-mKIF5B(516-621)-SNAPf is shifted to the left relative to CLIPf-mKIF5B(565-589)-SNAPf.

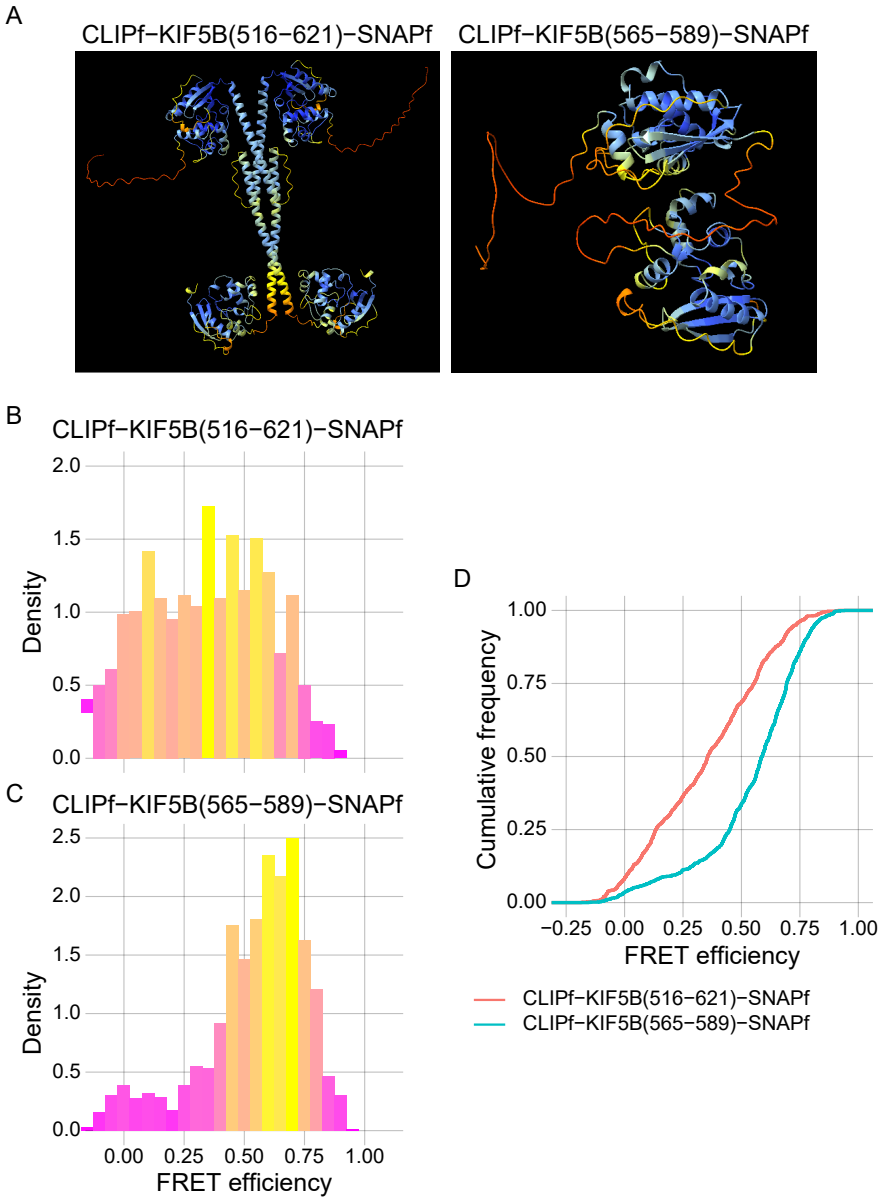

Figure 6.13: **FRET efficiency of CLIPf-mKIF5B(516-621)-SNAPf and CLIPf-mKIF5B(565-589)-SNAPf fusion proteins**

A: AlphaFold2 predictions for the two tagged fusion proteins. Structure coloured by pLDDT score: dark blue = 90-100, light blue = 70-90, yellow = 50-70 and orange = 0-50. A high pLDDT indicates high confidence in the structure. Histograms of FRET efficiency for the CLIPf-mKIF5B(516-621)-SNAPf fusion protein (B, n = 5, bursts = 1114) and the CLIPf-mKIF5B(565-589)-SNAPf fusion protein (C, n = 2, bursts = 1242). Black line indicates a normal distribution with the mean, or centre of the distribution, written beside. Differences in FRET efficiency can be seen in cumulative frequency graph (D), where CLIPf-mKIF5B(516-621)-SNAPf is shifted to the left relative to CLIPf-mKIF5B(565-589)-SNAPf.

constructs resulting in a range of FRET efficiencies from low to high (Hellenkamp et al. 2018). The design and potential effectiveness of using TPR-based fusion proteins with the previously optimised labelling strategy was proven to have potential as protein FRET standards.

The FRET efficiency for each of the four new protein TPR-based FRET standards need to be compared to that of known crystal structures to determine the accuracy and reliability of the results. Lysates from HEK293 cells transfected with the individual protein constructs (codon optimised for mammalian expression) resulted in ruler of different FRET efficiencies across the majority of the FRET efficiency range. The mean FRET efficiency for each fusion protein (CLIPf-SNAPf, CLIPf-TPR1-SNAPf, CLIPf-TPR2-SNAPf and CLIPf-TPR3-SNAPf) could be used to calculate the distance between the fluorophores. From this the length of a TPR motif could be determined and then compared to that of the known crystal structure for these designed motifs (Main et al. 2003).

The smFRET results from the, bacterial expressed, purified fusion proteins did not show any differences in FRET efficiency, with all three that were tried resulting in a large low FRET population. I believe this is due to BSA contamination of the purified protein. BSA was used to block the concentrator and hopefully prevent a huge loss of product. This step was performed before the size exclusion chromatography to help ensure sufficient protein for this step. An additional band, at just below 72 kDa, can be seen in the size exclusion samples in the SDS-PAGE analysis which was not present after StrepTag affinity purification (see Figures 6.3, 6.5 and 6.7). This suggests the addition of a contaminant between these two steps. The additional band was initially ignored as in gel analysis showed it was not labelled by either of the two fluorophores. However, as seen previously in the smFRET assay optimisation (see Section 3.2.2.2) commercially purchased BSA contains a contaminant which is fluorescent in the conditions of this experiment. Therefore the low FRET population is probably due to BSA contamination of the purified proteins

from the concentrating step. This was not a problem in the mammalian expressed constructs as cell lysate, not purified protein, samples were used.

A major limitation of the work with TPR domains presented in this chapter is that the smFRET experiments, either bacteria or mammalian codon optimised, were only performed once. Repetition is needed for more precise mean FRET efficiencies before distances are calculated from them. This would allow investigation into the reliability and reproducibility of the sample, whether cell lysate or purified protein, which is needed before wide spread use in the FRET community as protein standards. The potential major benefit of these standards is that they could be used for confocal- or TIRF-based smFRET.

Single molecule fluorescence resonance energy transfer (smFRET) could provide valuable information to help validate the structure predictions of AlphaFold2. Other traditional structure validation approaches, for example x-ray crystallography and cyro-EM, have not provided high resolution structural information for full length kinesin-1. This is probably due to the inherent flexibility of the molecule and its ability to adopt multiple conformations. These properties were seen in the smFRET results for both the CLIPf-hKIF5B-SNAPf fusion protein and the CLIPf-mKIF5A-SNAPf fusion proteins (discussed in earlier chapters). Conformational states and the flexibility of a molecule can be observed in smFRET experiments as results include the temporal resolution which is missing from the other structural experiments, like cryo-EM. This allows the heterogeneity of the protein population to be investigated as well as the structure of the protein (Agam et al. 2022).

Calculation of distances from FRET efficiency depends on the specific dyes used in the experiment. This is due to photophysical properties, like the spectral overlap between the emission of the donor fluorophore and the excitiation of the acceptor fluorophore, that impact on the value of the Föster radius (R_0). This value needs to be calculated for each pair of dyes used in a smFRET experiments. Currently this has not been done for CILP-Cell TMR-Star and SNAP-Cell 647-SiR so accurate

distances can not be calculated from the mean FRET efficiencies. Also two dye specific correction factors, βand γ, need to be calculated to determine how well each fluorophore is detected and excited by the instrumental setup. Again currently these two factors have not been calculated for this dye pair preventing accurate distant calculations.

For accurate validation of the interlocking barrel hinge region predicted by AlphaFold2, the designed hinge fragments need to exist as dimers. It wasn't known, especially for the shorter fragments, if these protein would exist as dimers or monomers. Two experimental approaches were used to try and determine the dimerisation state of the fragment.

Native-PAGE analysis resulted in variable results due to the difficulty in preventing protein denaturation. Changes to the sample preparation help prevent this issue, but not completely, and sometimes it was hard to see dimer bands. A change from cell lysis using detergent to homogenisation (similar to the process used in Section 5.2.1.3) could help prevent protein denaturation into monomers.

The native-PAGE result was confirmed by smFRET to show the longest hinge fragment exists primarily as a dimer, whereas the shortest fragment is monomeric. This was shown using the stoichiometry of the resultant bursts. Due to inefficient labelling of the CLIP-Cell TMR-Star ligand, the proportion of labelling of most dimers is usually lower then the expected 0.5; which is the value seen for monomers. This gives an easy and reproducible approach to studying the dimerisation state of the fusion protein.

A detailed structural investigation of the heavy chain of kinesin-1 is needed to validate the AlphaFold2 model. This could be done using accurate distance calculations from smFRET efficiency results. More structural information is needed to understand the unexpected smFRET results previously seen as the AlphaFold2 structure indicate more regions of flexibility than were previously thought. This information will also help with understanding the autoinhibition mechanism of the

kinesin-1 heavy chains. Structural information can not be gained by traditional approaches like cryo-EM, but smFRET results could give this and temporal information about different conformations.

Chapter 7

Discussion

7.1 Summary

Kinesin-1 activity needs to be regulated to prevent unnecessary ATP hydrolysis and accumulation of kinesin-1 in the cell periphery. This is done by autoinhibition of kinesin-1, but a detailed mechanism of kinesin-1 activation, including in response to the presence of microtubules and cargo, is currently unknown. A single molecule assay to investigate the conformational changes that take place in autoinhibition and activation is a necessary step to understanding this process. The experiments in this thesis have aimed to develop, optimise and validate a single molecule assay to study kinesin-1 conformation.

Single molecule fluorescence resonance energy transfer (smFRET) was chosen as the single molecule technique to base the new kinesin-1 conformation assay on. It was predicted that the two known conformations of the kinesin-1 heavy chain, folded and unfolded (Hackney, Levitt, and Suhan 1992 and Stock et al. 1999), could be differentiated using this approach. If each terminus of the kinesin-1 heavy chain was fluorescently tagged, the active stretched conformation would not result in a FRET signal due to the two fluorophores being more than 10 nm apart; roughly 60 nm. On the other hand, the closed inactive conformation of the kinesin-1 heavy

chain should result in the two flurorphores being within the 3-10 nm FRET range and therefore result in a FRET signal. The open conformation should result in a zero FRET efficiency population, whereas the closed conformation would give a high FRET population. smFRET allows temporal resolution of the two conformations and the proportions of these populations could be measured in response in different conditions, like ionic strength and the addition of microtubules.

smFRET can be made with specialised confocal or total internal reflection fluorescence (TIRF) microscopes. Tethering of the molecule to a surface, a requirement for TIRF based measurements, would potentially limit the conformational freedom of the molecule. Therefore a confocal based smFRET approach was chosen where kinesin-1 could freely diffuse across the confocal volume. This means kinesin-1 is free to adopt any conformation.

The combination of self-labelling enzymes and fluorescent ligands (containing organic dyes) were chosen to fluorescently tag the kinesin-1 heavy chain. This labelling strategy combines the genetic encoding ability of fluorescent proteins with the advantageous photophysical properties of organic dyes. The CLIPf-tag and the SNAPf-tag self labelling enzymes were selected due to their smaller size, and commercial availability of the sequence and fluorescent ligands. The CLIP-Cell TMR-Star and SNAP-Cell 647-SiR fluorescent ligands were chosen as they were compatible for use with the smFRET instrument and had the desired spectral overlap. They were also cell permeable meaning protein purification wasn't necessary and allowed the possibility of using cell lysate based samples.

As part of this thesis, the validation of the labelling strategy as well as the use of cell lysate samples was carried out, and both were determined to be successful in this smFRET assay. These experiments were done with the CLIPf-SNAPf fusion protein, or single tagged versions, to validate use and also optimise the conditions of the assay. The labelling conditions (washout, dye incubation, dye concentration and dye cross reactivity), data analysis pipeline, correction factors, assay buffer and

assay background were all optimised, resulting in a reproducible smFRET assay.

The smFRET results for the kinesin-1 heavy chains, mKIF5A and hKIF5B, were unexpected. Two distinct FRET efficiency populations were predicted for the kinesin-1 heavy chain based on the sedimentation coefficients differences seen by Hackney, Levitt, and Suhan 1992 and Stock et al. 1999. Instead one large, broad, low FRET population was seen with a mean FRET efficiency of around 0.2 with no higher FRET population. Increasing the salt concentration did result in a slight change in FRET efficiency, but again not in the expected way. Mutations of mKIF5A, predicted to alter the conformation of the molecule, also resulted in no change in FRET efficiency; still one large broad low FRET efficiency population with no higher population. The combination of all these results suggests the heavy chain of kinesin-1 adopts a wide range of conformations, indicated by the broad FRET efficiency population, and is highly flexible. Also the folded inactive conformation does not seem to be stable in the current assay conditions; as seen by the lack of a higher FRET population.

Other components, for example microtubules or cargo, might be required for kinesin-1 autoinhibition and the folded inactive conformation. This was investigated in this thesis, but more experimental work and data analysis is required. At the moment it appears that addition of microtubules cause a slight polarisation of the broad low FRET efficiency population. The emergence of a high FRET population can also be seen. This suggests that microtubules are required for kinesin-1 autoinhibtion, which in agreement with kinesin-1 being a microtubule-stimulated ATPase (Bloom et al. 1988 and Hackney 1994).

AlphaFold2 predicts a new structure for the kinesin-1 heavy chain. This includes two new hinge regions which could explain the unexpected flexibility seen in the smFRET results. Also the central hinge region of the heavy chain was predicted as an interlocking barrel structure instead of an extended disordered region. This new predicted structure needs to be validated through experimental approaches,

but traditional techniques, like cryo-EM, have limited structural resolution due to the flexibility of kinesin-1 (Weijman et al. 2022). smFRET could be used to help validate this structure by calculating distances from the FRET efficiency results. The accuracy and reliability of the smFRET results can be determined by the use of new protein FRET standards.

In conclusion, the data presented suggests kinesin-1 autoinhibition is a more complex mechanism then previously thought. The smFRET assay provides a new single molecule approach to study this mechanism by investigation of the conformation of the molecule as a direct readout of kinesin-1 activation. This assay and technique could also be used to help validate the AlphaFold2 predicted structure and the accuracy and reliability of the results determined by the use of new protein FRET standards.

7.2 smFRET assay using self labelling enzymes and cell lysate based samples

The labelling optimisation of the CLIPf- and SNAPf-tags presented in this work comprises a complete protocol to optimise the; washout, concentration, incubation and cross reactivity of fluorescent ligands. This protocol can be applied to optimise the labelling of any fluorescent ligands with any self-labelling enzyme. It is known that the self-labelling enzyme, that the fluorescent ligand binds to, affects the photophysical properties of the ligand; for example, photostability and brightness. Therefore each ligand and self-labelling enzyme pair needs to be optimised before use. An example of this has been seen recently in the lab. Similar labelling conditions to that of CLIP-Cell TMR-Star with the CLIPf-tag were used for the HaloTag and the HaloTag TMR ligand. This caused high levels of unspecific binding; even when the incubation time was reduced to that of SNAP-Cell 647-SiR with the SNAPf-tag.

In this work the labelling optimisation protocol was used to improve the results for an smFRET assay, but this protocol has the potential to improve the results of any single molecule experiment. In these types of experiment, the signal-to-noise ratio is crucial to success. Therefore, a standard protocol to increase the amount of labelled sample, whilst also reducing background signal from unbound ligands, for any fluorescent ligand and tag combination could be extremely beneficial.

To the best of my knowledge, this is the first example of smFRET results, using alternating-laser excitation (ALEX), from cell lysate samples. This keeps the protein in a more physiological environment until they are used; purified protein samples are stored in buffers and kept at - 80 °C. A cell lysate approach is beneficial to the smFRET community as it allows the technique to be used on a wider array of samples. For example, low abundant proteins, or proteins that can't be purified, can now be transfected, or otherwise engineered, into cells and studied by smFRET. It also allows the quick investigation of many samples. When the TPR constructs in Chapter 6 were being investigated by smFRET, it took three days to get the results from the constructs that were codon optimised for mammalian expression. On the other hand, it took around two months to optimise the purification and labelling strategy for the constructs that were codon optimised for bacterial expression.

7.3 Heavy chain kinesin-1 activation is more complex then previously thought

The smFRET results for the kinesin-1 heavy chains, mKIF5A and hKIF5B (see Chapter 4 and Figure 5.1), did not culminate in two distinct FRET efficiency populations. These populations were expected from previous sedimentation assays (Hackney, Levitt, and Suhan 1992 and Stock et al. 1999). This is probably due to the different experimental approaches and samples. In sedimentation experiments purified kinesin-1 heavy chains (with and without the light chains) were centrifuged

with standard proteins on a exponential sucrose gradient. This technique shows the possible conformations a protein can adopt, but not whether they occur in physiological conditions or if they are stable. The smFRET work presented here does not say that the kinesin-1 heavy chain can not adopt two conformations, but that it doesn't adopt the folded inactive form when it is alone; without cargo or microtubules. Further work needs to be done to show when autoinhibition occurs and then two distinct FRET efficiency populations will be seen.

Previous whole cell FRET experiments showed the conformational change of kinesin-1 through changes in FRET efficiency (Cai et al. 2007). This was an ensemble FRET approach giving the average conformation of the entire kinesin-1 population whereas this smFRET technique studies the proportion of kinesin-1 molecules in each conformation. There "high" FRET efficiency signal was very low compared to their positive control so this might not represent the folded conformation as they suggest. The labelling strategy was also different to that in this work. Fluorescent proteins were used to tag the protein of interest, and consequently required cotransfection of singley labelled constructs (similar to that seen in Section 4.2.3) to prevent aggregation of the four fluorescent proteins. In some of their experiments the high FRET signal looks to be down to protein aggregation at the cell periphery and not actually the fold conformation. Aggregation of the protein was not seen in the smFRET results. This would have been noticeable due to extremely large burst duration and unexpected changes in stoichiometry between biological repeats.

Other then the lack of a higher FRET efficiency population in these results, there is more evidence that the folded inactive state of kinesin-1 is not stable. Recently two groups have independently shown that cross-linking, between the motor and tail domains, is needed to form a stable folded conformation (Weijman et al. 2022 and personal communication with Dr. M Cianfrocco at the GRC cytoskeletal motors conference in July 2022). The cross-linker bis(sulfosuccinimidyl)suberate (BS3) had to be added to the purified protein to result in two conformations of kinesin-1; the

elongated active and the folded inactive forms. Even then that pool still had a mix of open and closed conformations (Weijman et al. 2022).

7.4 Components of axonal transport can affect kinesin-1 autoinhibition

Kinesin-1 is a microtubule-stimulated ATPase (Hackney 1994, Stock et al. 1999 and Hackney and Stock 2000) and therefore it makes sense that autoinhibition, into the inactive folded form, is not necessary when microtubules aren't present. Autoinhibtion, in the presence of microtubules can be seen in single molecule walking assays. Full length kinesin-1 heavy chain alone (in the absence of activators) shows minimal movement (for example Twelvetrees, Lesept, et al. 2019). The preliminary smFRET data with microtubules presented here (see Figure 5.7) supports this idea of kinesin-1 autoinhibtion with the emergence of a higher FRET population and the polarisation of the FRET efficiency distribution. The higher FRET efficiency population was still very small compared to the low FRET efficiency population, suggesting that the folded inactive conformation is still not very stable and only occurs transiently.

7.5 Using smFRET to validate structure predictions

In order to validate AlphaFold2 predictions using smFRET, results need to be proven to be accurate and reproducuible. This can be done thorugh the use of comparative blind studies, but well-characterised samples, known as FRET standards, are needed (Hellenkamp et al. 2018 and Agam et al. 2022). Currently the only available protein FRET standards measure the conformational change in a cysteine labelled protein. It would be beneficial to have fixed length protein FRET standards, that used the

same labelling strategy as the one chosen in this work.

In this work the potential use of TPR based constructs for fixed length protein FRET standards was shown (see Figure 6.9). Four fusion proteins covered the majority of the FRET efficiency range, but one more larger construct (TPR4) might be helpful. It would be beneficial if these FRET standards could be used independently of the chosen sample preparation method and the smFRET system (TIRF or confocal based). It was shown that these fusion proteins could be purified from bacterial expression, but smFRET results were not conclusive due to possible contamination from BSA in the concentrating step of purification. These constructs contain both the FLAG and Strep tags for use in protein purification, but the FLAG tag in particular can also be used in surface tethering for TIRF smFRET experiments.

The possibility of using smFRET to validate AlphaFold2 structures was shown in this work (see section 6.2.2). This has the potential to be extremely useful as some proteins do not result in high resolution structures from traditional experimental approaches like cryo-EM and x-ray crystallography. For example, recently only 2D class average analysis could be achieved from negative-stain EM of kinesin-1 (Weijman et al. 2022). This is probably due to the high flexibility and low electron density of the stalk and tail domain of kinesin-1. Another advantage of smFRET to validate predicted structures is the temporal resolution of the technique. This allows the dynamics and any possible conformations of the molecule to be investigated as well as the structure.

7.6 Future work

This thesis has shown that kinesin-1 autoinhibition and activation is more complex then a simple binary on/off mechanism, with much more work to be done on understanding the regulation of kinesin-1.

This work only focuses on the kinesin-1 heavy chains. Kinesin-1 exists primarily

as a heterotetramer composed of two heavy chains and two light chains. It has been shown that the light chains are not necessary for the autoinhibtion of the heavy chains (Hackney, Levitt, and Suhan 1992 and Verhey et al. 1998), but they do affect have an affect on the overall activity of the molecule (Hackney, Levitt, and Wagner 1991). In order to gain a fully representative detailed picture of kinesin-1 autoinhibition, and the activation mechanism of kinesin-1, the light chains can be added into the smFRET assay. This could be especially important when looking at the effect of different cargoes on kinesin-1 autoinhibition as the light chains are involved in cargo recognition. There is evidence that the light chains themselves can be autoinhibited (Yip et al. 2016). By only studying the heavy chains additional levels of kinesin-1 regulation might be overlooked. Other conditions, like phosphorylation, also need to be studied as there is some evidence of this affecting kinesin-1 activity (Matthies et al. 1993 and Chiba et al. 2017).

Another potential limitation of this smFRET assay is the impact of the self-labelling enzyme on the conformational freedom of the tagged kinesin-1 molecule. In previous studies of kinesin-1, tags of larger size, for example fluorescent proteins or the HaloTag, have been shown to not affect kinesin-1 function or movement (Norris et al. 2015). However the tags, in this system, need to be proven to not change the behaviour of the kinesin-1 heavy chain. Preliminary data from the lab indicts that the tags do not affect the ability of the kinesin-1 heavy chain to walk along the microtubule once it is active. However, there is potential for these tags to affect the autoinhibition process by limiting the conformational flexibility of the molecule. Kinesin-1 function and movement could be confirmed by ATPase assays and TIRF-based walking assays. Conservation of cargo or light chain binding could be shown by immunoprecipitation.

There is currently no evidence of any difference in autoinhibition between the three heavy chain isomers, or the four light chain isomers, due to lack of investigation. This could be studied using this new smFRET assay. The overall structure,

including that predicted by AlphaFold2, is expected to be the same for the each of the heavy chain isoforms. They do differ in their cell specificity; KIF5A and KIF5C are neuronal specific whereas KIF5B is endogenously expressed in all cell types. It is not known if the different isoforms are needed to carry different cargo, but isoform specificity could be possible. For example, mutations in the different isoforms result in different disease phenotypes. KIF5A mutations are associated with neurodegenerative diseases like amyotrophic lateral sclerosisanoside (ALS) (Nicolas et al. 2018), KIF5B mutations with lung cancer (Kohno et al. 2012) and KIF5C mutations with neurodevelopmental disorders (Poirier et al. 2013).

A hugely beneficial experiment, needed in this work, would be to observe where the closed conformation of kinesin-1 should appear on the FRET efficiency scale. At the moment, potentially due to its transient nature, no stable high FRET efficiency population has been observed, apart from a potential small population in the presence of microtubules. By trapping the inactive closed conformation and measuring the FRET efficiency the position of this population could be observed and then compared to the untrapped kinesin-1 heavy chain. If this population is at high FRET efficiency (E >0.5) as expected, it can be shown not to be present in the original smFRET results for the mKIF5A fusion protein. This will help to show that this conformation is not stable in the conditions of the assay. On the other hand if this inactive conformation results in a mid to low FRET then more analysis of the current smFRET results needs to occur to differentiate the populations. This could be done by burst variance analysis (Torella et al. 2011)). The trapping of this inactive conformation could be done by cross-linking the molecule, as used in (Weijman et al. 2022) or by formation of a disulphide bridge between the motor domain and QIAK interacting motif in the tail domain. A cysteine light mutant of the heavy chain would have to be made with cysteine residues added at specific points in the motor and tail domain.

Bibliography

Adio, Sarah et al. (2006). "Review: regulation mechanisms of Kinesin-1". *Journal of muscle research and cell motility* 27.2, pp. 153–160.

Agam, Ganesh et al. (2022). "Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins". *bioRxiv*, DOI: 10.1101/2022.08.03.502619.

Aguado, Carmen et al. (2016). "Isolation of Lysosomes from Mammalian Tissues and Cultured Cells". *Proteostasis: Methods and Protocols*. Ed. by Rune Matthiesen. New York, NY: Springer New York, pp. 299–311.

Akdel, Mehmet et al. (2022). "A structural biology community assessment of AlphaFold2 applications". *Nature structural & molecular biology*, pp. 1–12.

Alushin, Gregory M et al. (2014). "High-resolution microtubule structures reveal the structural transitions in $\alpha\beta$-tubulin upon GTP hydrolysis". *Cell* 157.5, pp. 1117–1129.

Ambrose, Benjamin et al. (2020). "The smfBox is an open-source platform for single-molecule FRET". *Nature communications* 11.1, p. 5641.

Amos, L A (1987). "Kinesin from pig brain studied by electron microscopy". *Journal of cell science* 87 (Pt 1), pp. 105–111.

Araújo, Mariana E G de et al. (2015). "Homogenization of Mammalian Cells". *Cold Spring Harbor protocols* 2015.11, pp. 1009–1012.

Asbury, Charles L et al. (2003). "Kinesin moves by an asymmetric hand-over-hand mechanism". *Science* 302.5653, pp. 2130–2134.

Baas, P W et al. (1988). "Polarity orientation of microtubules in hippocampal neurons: uniformity in the axon and nonuniformity in the dendrite". *Proceedings of the National Academy of Sciences of the United States of America* 85.21, pp. 8335–8339.

Block, S M et al. (1990). "Bead movement by single kinesin molecules studied with optical tweezers". *Nature* 348.6299, pp. 348–352.

Bloom, George S et al. (1988). "Native structure and physical properties of bovine brain kinesin and identification of the ATP-binding subunit polypeptide". *Biochemistry* 27.9, pp. 3409–3416.

Brady, S T (1985). "A novel brain ATPase with properties expected for the fast axonal transport motor". *Nature* 317.6032, pp. 73–75.

Cai, Dawen et al. (2007). "Kinesin-1 structural organization and conformational changes revealed by FRET stoichiometry in live cells". *The Journal of cell biology* 176.1, pp. 51–63.

Canty, John T et al. (2021). "TRAK adaptors coordinate the recruitment and activation of dynein and kinesin to control mitochondrial transport". *bioRxiv*, DOI: 10.1101/2021.07.30.454553.

Cao, Luyan, Soraya Cantos-Fernandes, et al. (2017). "The structural switch of nucleotide-free kinesin". *Scientific reports* 7, p. 42558.

Cao, Luyan, Weiyi Wang, et al. (2014). "The structure of apo-kinesin bound to tubulin links the nucleotide cycle to movement". *Nature communications* 5, p. 5364.

Chiba, Kyoko et al. (2017). "Phosphorylation of KLC1 modifies interaction with JIP1 and abolishes the enhanced fast velocity of APP transport by kinesin-1". *Molecular biology of the cell* 28.26, pp. 3857–3869.

Downing, K H and E Nogales (1998). "Tubulin structure: insights into microtubule properties and functions". *Current opinion in structural biology* 8.6, pp. 785–791.

Ems-McClung, Stephanie C et al. (2013). "Aurora B inhibits MCAK activity through a phosphoconformational switch that reduces microtubule association". *Current biology: CB* 23.24, pp. 2491–2499.

Encell, Lance P (2013). "Development of a Dehalogenase-Based Protein Fusion Tag Capable of Rapid, Selective and Covalent Attachment to Customizable Ligands". *Current chemical genomics* 6.1, pp. 55–71.

Erdmann, Roman S et al. (2019). "Labeling Strategies Matter for Super-Resolution Microscopy: A Comparison between HaloTags and SNAP-tags". *Cell chemical biology*.

Evans, Richard et al. (2022). "Protein complex prediction with AlphaFold-Multimer". *bioRxiv*, DOI: 10.1101/2021.10.04.463034.

Friedman, D S and R D Vale (1999). "Single-molecule analysis of kinesin motility reveals regulation by the cargo-binding tail domain". *Nature cell biology* 1.5, pp. 293–297.

Gautier, Arnaud et al. (2008). "An engineered protein tag for multiprotein labeling in living cells". *Chemistry & biology* 15.2, pp. 128–136.

Grimm, Jonathan B, Brian P English, Jiji Chen, et al. (2015). "A general method to improve fluorophores for live-cell and single-molecule microscopy". *Nature methods* 12.3, 244–50, 3 p following 250.

Grimm, Jonathan B, Brian P English, Heejun Choi, et al. (2016). "Bright photoactivatable fluorophores for single-molecule imaging". *Nature methods* 13.12, pp. 985–988.

Hackney, D D (1988). "Kinesin ATPase: rate-limiting ADP release". *Proceedings of the National Academy of Sciences of the United States of America* 85.17, pp. 6314–6318.

— (1994). "Evidence for alternating head catalysis by kinesin during microtubule-stimulated ATP hydrolysis". *Proceedings of the National Academy of Sciences of the United States of America* 91.15, pp. 6865–6869.

Hackney, D D, J D Levitt, and J Suhan (1992). "Kinesin undergoes a 9 S to 6 S conformational transition". *The Journal of biological chemistry*.

Hackney, D D, J D Levitt, and D D Wagner (1991). "Characterization of alpha 2 beta 2 and alpha 2 forms of kinesin". *Biochemical and biophysical research communications* 174.2, pp. 810–815.

Hackney, D D and M F Stock (2000). "Kinesin's IAK tail domain inhibits initial microtubule-stimulated ADP release". *Nature cell biology* 2.5, pp. 257–260.

He, Yufan et al. (2015). "Single-molecule spectroscopy reveals how calmodulin activates NO synthase by controlling its conformational fluctuation dynamics". *Proceedings of the National Academy of Sciences of the United States of America* 112.38, pp. 11835–11840.

Hellenkamp, Björn et al. (2018). "Publisher Correction: Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study". *Nature methods* 15.11, p. 984.

Hirokawa, Nobutaka and Yasuko Noda (2008). "Intracellular transport and kinesin superfamily proteins, KIFs: structure, function, and dynamics". *Physiological reviews* 88.3, pp. 1089–1118.

Hirokawa, Nobutaka, K Kevin Pfister, et al. (1989). "Submolecular domains of bovine brain kinesin identified by electron microscopy and monoclonal antibody decoration". *Cell* 56.5, pp. 867–878.

Howard, J et al. (1989). "Movement of microtubules by single kinesin molecules". *Nature* 342.6246, pp. 154–158.

Hua, Wei et al. (2002). "Distinguishing inchworm and hand-over-hand processive kinesin movement by neck rotation measurements". *Science* 295.5556, pp. 844–848.

Ingargiola, Antonino et al. (2016). "FRETBursts: An Open Source Toolkit for Analysis of Freely-Diffusing Single-Molecule FRET". *PloS one* 11.8, e0160716.

Ishitsuka, Yuji et al. (2015). "Evaluation of Genetically Encoded Chemical Tags as Orthogonal Fluorophore Labeling Tools for Single-Molecule FRET Applications". *The journal of physical chemistry. B* 119.22, pp. 6611–6619.

Isojima, Hiroshi et al. (2016). "Direct observation of intermediate states during the stepping motion of kinesin-1". *Nature chemical biology* 12.4, pp. 290–297.

Jumper, John et al. (2021). "Highly accurate protein structure prediction with AlphaFold". *Nature* 596.7873, pp. 583–589.

Kaan, Hung Yi Kristal et al. (2011). "The structure of the kinesin-1 motor-tail complex reveals the mechanism of autoinhibition". *Science* 333.6044, pp. 883–885.

Kanai, Y et al. (2000). "KIF5C, a novel neuronal kinesin enriched in motor neurons". *The Journal of neuroscience: the official journal of the Society for Neuroscience* 20.17, pp. 6374–6384.

Keppler, Antje et al. (2003). "A general method for the covalent labeling of fusion proteins with small molecules in vivo". *Nature biotechnology* 21.1, pp. 86–89.

Kohno, Takashi et al. (2012). "KIF5B-RET fusions in lung adenocarcinoma". *Nature medicine* 18.3, pp. 375–377.

Kosik, K S et al. (1990). "The primary structure and analysis of the squid kinesin heavy chain". *The Journal of biological chemistry* 265.6, pp. 3278–3283.

Kull, F J et al. (1996). "Crystal structure of the kinesin motor domain reveals a structural similarity to myosin". *Nature* 380.6574, pp. 550–555.

Kuznetsov, S A and V I Gelfand (1986). "Bovine brain kinesin is a microtubule-activated ATPase". *Proceedings of the National Academy of Sciences of the United States of America* 83.22, pp. 8530–8534.

Kuznetsov, S A, E A Vaisberg, et al. (1988). "The quaternary structure of bovine brain kinesin". *The EMBO journal* 7.2, pp. 353–356.

Kuznetsov, S A, Y A Vaisberg, et al. (1989). "Isolation of a 45-kDa fragment from the kinesin heavy chain with enhanced ATPase and microtubule-binding activities". *The Journal of biological chemistry* 264.1, pp. 589–595.

Lawrence, Carolyn J et al. (2004). "A standardized kinesin nomenclature". *The Journal of cell biology* 167.1, pp. 19–22.

Lee, Nam Ki et al. (2005). "Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation". *Biophysical journal* 88.4, pp. 2939–2953.

Los, Georgyi V et al. (2008). "HaloTag: a novel protein labeling technology for cell imaging and protein analysis". *ACS chemical biology* 3.6, pp. 373–382.

Macaskill, Andrew F et al. (2009). "Miro1 is a calcium sensor for glutamate receptor-dependent localization of mitochondria at synapses". *Neuron* 61.4, pp. 541–555.

Maday, Sandra et al. (2014). "Axonal transport: cargo-specific mechanisms of motility and regulation". *Neuron* 84.2, pp. 292–309.

Main, Ewan R G et al. (2003). "Design of stable alpha-helical arrays from an idealized TPR motif". *Structure* 11.5, pp. 497–508.

Mandelkow, E and E M Mandelkow (1995). "Microtubules and microtubule-associated proteins". *Current opinion in cell biology* 7.1, pp. 72–81.

Matthies, H J et al. (1993). "Calmodulin binding to and cAMP-dependent phosphorylation of kinesin light chains modulate kinesin ATPase activity". *The Journal of biological chemistry* 268.15, pp. 11176–11187.

Mitchison, T and M Kirschner (1984). "Dynamic instability of microtubule growth". *Nature* 312.5991, pp. 237–242.

Navone, F et al. (1992). "Cloning and expression of a human kinesin heavy chain gene: interaction of the COOH-terminal domain with cytoplasmic microtubules in transfected CV-1 cells". *The Journal of cell biology* 117.6, pp. 1263–1275.

Nicolas, Aude et al. (2018). "Genome-wide Analyses Identify KIF5A as a Novel ALS Gene". *Neuron* 97.6, 1268–1283.e6.

Nogales, E et al. (1999). "High-resolution model of the microtubule". *Cell* 96.1, pp. 79–88.

Norris, Stephen R et al. (2015). "Influence of fluorescent tag on the motility properties of kinesin-1 in single-molecule assays". *Biophysical journal* 108.5, pp. 1133–1143.

Poirier, Karine et al. (2013). "Mutations in TUBG1, DYNC1H1, KIF5C and KIF2A cause malformations of cortical development and microcephaly". *Nature genetics* 45.6, pp. 639–647.

Presman, Diego M et al. (2017). "Quantifying transcription factor binding dynamics at the single-molecule level in live cells". *Methods* 123, pp. 76–88.

Remington, J (2006). "Fluorescent proteins: maturation, photochemistry and photophysics". *Current opinion in structural biology* 16, pp. 714–721.

Rice, S et al. (1999). "A structural change in the kinesin motor protein that drives motility". *Nature* 402.6763, pp. 778–784.

Sack, S et al. (1997). "X-ray structure of motor and neck domains from rat brain kinesin". *Biochemistry* 36.51, pp. 16155–16165.

Schnapp, B J et al. (1992). "Kinesin is bound with high affinity to squid axon organelles that move to the plus-end of microtubules". *The Journal of cell biology* 119.2, pp. 389–399.

Schnitzer, M J and S M Block (1997). "Kinesin hydrolyses one ATP per 8-nm step". *Nature* 388.6640, pp. 386–390.

Seeger, Mark A and Sarah E Rice (2013). "Intrinsic Disorder in the Kinesin Superfamily". *Biophysical reviews* 5.3, pp. 233–247.

Shang, Zhiguo et al. (2014). "High-resolution structures of kinesin on microtubules provide a basis for nucleotide-gated force-generation". *eLife* 3, e04686.

Sindelar, Charles V et al. (2002). "Two conformations in the human kinesin power stroke defined by X-ray crystallography and EPR spectroscopy". *Nature structural biology* 9.11, pp. 844–848.

Skoufias, D A et al. (1994). "The carboxyl-terminal domain of kinesin heavy chain is important for membrane binding". *The Journal of biological chemistry* 269.2, pp. 1477–1485.

Stock, M F et al. (1999). "Formation of the compact confomer of kinesin requires a COOH-terminal heavy chain domain and inhibits microtubule-stimulated ATPase activity". *The Journal of biological chemistry* 274.21, pp. 14617–14623.

Su, Qingning et al. (2004). "Syntabulin is a microtubule-associated protein implicated in syntaxin transport in neurons". *Nature cell biology* 6.10, pp. 941–953.

Subach, Fedor V and Vladislav V Verkhusha (2012). "Chromophore transformations in red fluorescent proteins". *Chemical reviews* 112.7, pp. 4308–4327.

Sun, Faneng et al. (2011). "Sunday Driver/JIP3 binds kinesin heavy chain directly and enhances its motility". *The EMBO journal* 30.16, pp. 3416–3429.

Sun, Yuansheng et al. (2011). "FRET microscopy in 2010: the legacy of Theodor Förster on the 100th anniversary of his birth". *Chemphyschem: a European journal of chemical physics and physical chemistry* 12.3, pp. 462–474.

Svoboda, K and S M Block (1994). "Force and velocity measured for single kinesin molecules". *Cell* 77.5, pp. 773–784.

Svoboda, K, C F Schmidt, et al. (1993). "Direct observation of kinesin stepping by optical trapping interferometry". *Nature* 365.6448, pp. 721–727.

Thormählen, M et al. (1998). "The coiled-coil helix in the neck of kinesin". *Journal of structural biology* 122.1-2, pp. 30–41.

Tomishige, Michio et al. (2006). "Single-molecule observations of neck linker conformational changes in the kinesin motor protein". *Nature structural & molecular biology* 13.10, pp. 887–894.

Torella, Joseph P et al. (2011). "Identifying molecular dynamics in single-molecule FRET experiments with burst variance analysis". *Biophysical journal* 100.6, pp. 1568–1577.

Twelvetrees, Alison E (2020). "The lifecycle of the neuronal microtubule transport machinery". *Seminars in cell & developmental biology* 107, pp. 74–81.

Twelvetrees, Alison E, Flavie Lesept, et al. (2019). "The adaptor proteins HAP1a and GRIP1 collaborate to activate the kinesin-1 isoform KIF5C". *Journal of cell science* 132.24.

Twelvetrees, Alison E, Eunice Y Yuen, et al. (2010). "Delivery of GABAARs to synapses is mediated by HAP1-KIF5 and disrupted by mutant huntingtin". *Neuron* 65.1, pp. 53–65.

Urrutia, A et al. (1991). "Purified kinesin promotes vesicle motility and induces active sliding between microtubules in vitro". *Proc. Natl. Acad. Sci* 88, pp. 6701–6705.

Vale, R D, T Funatsu, et al. (1996). "Direct observation of single kinesin molecules moving along microtubules". *Nature* 380.6573, pp. 451–453.

Vale, R D, B J Schnapp, et al. (1985). "Organelle, bead, and microtubule translocations promoted by soluble factors from the squid giant axon". *Cell* 40.3, pp. 559–569.

Vale, Ronald D et al. (1985). "Identification of a novel force-generating protein, kinesin, involved in microtubule-based motility". *Cell* 42.1, pp. 39–50.

Verhey, K J et al. (1998). "Light chain-dependent regulation of Kinesin's interaction with microtubules". *The Journal of cell biology* 143.4, pp. 1053–1066.

Waldo, G S et al. (1999). "Rapid protein-folding assay using green fluorescent protein". *Nature biotechnology* 17.7, pp. 691–695.

Wang, Lu et al. (2019). "Small-Molecule Fluorescent Probes for Live-Cell Super-Resolution Microscopy". *Journal of the American Chemical Society* 141.7, pp. 2770–2781.

Weijman, Johannes F et al. (2022). "Molecular architecture of the autoinhibited kinesin-1 lambda particle". *Science advances* 8.37, eabp9660.

Westermann, Stefan and Klaus Weber (2003). "Post-translational modifications regulate microtubule function". *Nature reviews. Molecular cell biology* 4.12, pp. 938–947.

Wilke, Sonja et al. (2012). "Crystal structure of the conserved domain of the DC lysosomal associated membrane protein: implications for the lysosomal glycocalyx". *BMC biology* 10, p. 62.

Wolff, Jan O et al. (2022). "MINFLUX dissects the unimpeded walking of kinesin-1".

Xia, Tie et al. (2013). "Single-molecule fluorescence imaging in living cells". *Annual review of physical chemistry* 64, pp. 459–480.

Yang, J T et al. (1989). "A three-domain structure of kinesin heavy chain revealed by DNA sequence and microtubule binding analyses". *Cell* 56.5, pp. 879–889.

Yildiz, Ahmet et al. (2004). "Kinesin walks hand-over-hand". *Science* 303.5658, pp. 676–678.

Yip, Yan Y et al. (2016). "The light chains of kinesin-1 are autoinhibited". *Proceedings of the National Academy of Sciences of the United States of America* 113.9, pp. 2418–2423.