

Faculty of Health and Life Sciences
Oxford Brookes University
Oxford



**Defining the molecular composition and
organisation of transition fibres in
Trypanosoma brucei and their roles in
flagellar assembly and length control**

Manu Ahmed

Supervisor: Professor Sue Vaughan

Date: 25/01/2022

Word count: 37,022

A thesis submitted in partial fulfilment of the requirements of the award of Doctor of Philosophy

Declaration

I declare that the material contained in this thesis has not been used in any other submission for an academic award. I confirm that all the research and findings presented in this thesis are my own work unless otherwise indicated through the use of a clear referencing system.

Abstract

The eukaryotic cilium/flagellum is a slender microtubule-based projection that extends, with a contiguous membrane, out from the cell body. They possess a highly conserved core structure, composed of a microtubule-based axoneme that runs through its core. The axonemal microtubules are nucleated by a barrel shaped microtubule-organising basal body. Transition fibres (TFs) are appendage structures surrounding the distal end of the mature basal body that enable membrane docking and regulate the trafficking of intraflagellar transport (IFT) proteins required for flagellar assembly. Transition fibres are present in all ciliated organisms yet only a handful of TF proteins have been identified in mammalian cells. We hypothesise that there are a number of additional components to be identified.

We have identified 35 putative TF proteins and used automated image analysis to map their spatial arrangement based on their diameter and offset from the mature basal body. We confirmed the localisation of these proteins at the transition fibre by co-localisation with known transition fibre proteins and investigated transition fibre dynamics and assembly by analysing the recruitment tagged proteins over the cell cycle. We assessed the function of 20 transition fibre proteins by RNAi and found six proteins that are indispensable for flagellar assembly. We also identified 2 proteins that appear to be involved in flagellar length control and a further 2 proteins that appear to act as regulators of length control. We identified two transition fibre proteins that exhibit a phenotype relating to cell cycle progression but not related to flagellum formation. Ciliopathies are a class of diseases related to the dysfunction of cilia. We also investigated the mechanism of four human ciliopathy orthologues (OFD1, CEP90, RABL2B and CEP19) in the import and loading of IFT onto the flagellar axoneme.

The study provides evidence that the protein network of the TFs is far more complex and intricate than what has been previously described in mammals. We also show that trypanosome transition fibres are indispensable for both flagellar assembly and length control whilst also potentially playing further roles in cell cycle progression. We have described a novel mechanism of human ciliopathy protein OFD1 in ciliary assembly that has not been discovered previously. We conclude that transition fibres are essential structures in the context of cilium/flagellum assembly and length control that contribute directly to human ciliopathies and we have leveraged the trypanosome model effectively to address broad biological questions that relates closely to human disease.

Acknowledgements

I would firstly like to thank my supervisor Prof Sue Vaughan for the many opportunities she has given me. My first lab experiences as an undergraduate, a rewarding and interesting PhD project, visits and conferences during my time here and of course the support and guidance throughout the entire process - I am extremely grateful for this.

I would also like to thank my co-supervisor Dr Jack Sunter for being a fountain of knowledge and always sharing with us his creative, critical and meticulous style of thinking. From our many conversations, I have truly learned a great deal.

I would like to also thank Dr Richard Wheeler for his support in many of the technical aspects of the project. Much of this work wouldn't have been possible without his knowledge and skills.

I am grateful to all of my colleagues who were always kind and supportive. In particular, I would like to acknowledge Shahaan Shafiq for all of his support throughout my time here.

I want to thank my partner Zee who spent many late nights patiently alongside me while I worked, she was truly a huge support in many of the most challenging moments.

I would like to thank all my family and friends for their support. In particular, my Mum who supports my endeavours with love and kindness. My Grandmother Lindy who always nurtured my interests and curiosities throughout my life— she is the reason I am here today. And lastly, I would like to thank my father who has always served as a beacon of guidance and wisdom, teaching me that hard work and resilience always paves the way.

“The impediment to action advances action.

What stands in the way becomes the way”

Marcus Aurelius – Meditations

Table of contents

Declaration.....	2
Abstract.....	3
Acknowledgements.....	4
List of Figures	9
List of Tables	11
Abbreviations.....	12
1. Introduction.....	13
1.1 Cilia and flagella in eukaryotic biology.....	13
1.1.1 Cilia and flagella in human health and disease	13
1.1.2 Many eukaryotic pathogens require a flagellum	14
1.2 Structure of cilia/flagella.....	14
1.3 Transition fibres: background and known structure	16
1.3.1 Historical background and discovery	16
1.3.2 Structure and composition	17
1.4 Transition fibres in flagellum assembly and length control.....	19
1.4.1 Roles of transition fibres in cilium biogenesis	19
1.4.2 Balance point model	20
1.4.3 Growth and lock model.....	21
1.5 Studying transition fibres in <i>Trypanosoma brucei</i>	22
1.5.1 <i>Trypanosoma brucei</i> is a unicellular eukaryotic pathogen	22
1.5.2 The flagellum is essential for the pathogenicity of <i>Trypanosoma brucei</i>	22
1.5.3 Organelle position and segregation marks the stages of the trypanosome cell cycle	23
1.5.4 Basal body biogenesis and transition fibre assembly	26
1.5.5 Molecular tools are well developed in trypanosomes.....	27
1.6 Project aims and objectives	29
1.6.1 Aims.....	29
1.6.2 Objectives.....	29

2. Materials and methods	30
2.1 Trypanosome cell culture	30
2.2 Gene tagging at the endogenous locus	30
2.2.1 PCR	30
2.2.2 Transfection	31
2.3 Light microscopy	32
2.3.1 Live cell microscopy	32
2.3.2 Methanol fixed cytoskeletons	33
2.3.3 Formaldehyde-methanol fixation	33
2.4 Transmission electron microscopy	33
2.5 Bioinformatics	34
2.6 Automated image analysis	34
2.7 RNAi	34
2.7.1 RNA interference plasmid generation	34
2.7.2 Transfection	35
2.7.3 Growth curve	35
2.8 Immunofluorescence	35
2.9 Solutions	36
2.9.1 Phosphate buffered saline (PBS)	36
2.9.2 Voorheis modified PBS (vPBS) (12)	37
2.9.3 PEME buffer	37
2.9.4 Hoechst	37
2.9.5 4% Formaldehyde solution	37
2.9.6 0.2M Phosphate buffer solution	37
2.9.7 EM primary fixative solution	37
2.9.8 2% uranyl acetate	37
2.9.9 4% osmium tetroxide stock solution	38
3. Transition fibre composition, structure and spatial arrangement	38

3.1 Identification of 35 putative transition fibre proteins	38
3.2 Co-expression of putative basal body proteins with basal body maker SAS-6	39
3.3 Localisation of 35 putative transition fibre proteins.....	41
3.4 Quantitative analysis putative transition fibre protein localisation	51
3.5 Ultrastructural analysis of transition fibre structure	56
3.6 Co-localisation of proteins with a known transition fibre marker	58
3.6.1 Generation of a transition fibre marker cell line	58
3.6.2 Generation of double tagged cell lines for co-localisation	60
3.6.3 Co-localisation of putative transition fibre proteins with CEP164A.....	61
3.6.4 Co-localisation of putative transition fibre proteins with RP2	64
4. Cell cycle regulation, assembly and conservation of transition fibres	70
4.1 Cell cycle regulated transition fibre proteins at the old and new flagellum	70
4.1.1 Cell cycle regulated localisation patterns brighter on the old flagellum	70
4.1.2 Cell cycle regulated localisation pattern brighter on the new flagellum	74
4.2 Dissecting timing of transition fibre assembly.....	77
4.2.1 Transition fibre recruitment in relation to SAS6 duplication	77
4.2.2 Transition fibre recruitment in relation to the arrival of CEP164A.....	80
4.3 Evolutionary analysis of trypanosome transition fibres by reciprocal best blast.....	84
4.3.1 Widely conserved transition fibre proteins	84
4.3.2 Transition fibre proteins only conserved within the excavata super group.....	87
4.3.3 Trypanosome specific transition fibre proteins	89
5. Functional analysis of transition fibre proteins by RNAi.....	92
5.1 RNAi plasmid generation	92
5.2 Selection of 20 transition fibre proteins for functional analysis	93
5.3 6 transition fibre proteins are required for flagellum assembly	94
5.3.1 TFP68 is required for flagellum assembly and localisation of CEP164A	94
5.3.2 Ciliopathy protein CEP90 is required for flagellum assembly and localisation of CEP164A	97

5.3.3 Rabl2b is required for flagellum assembly but not localisation of CEP164A	99
5.3.4 Ciliopathy protein OFD1 is required for flagellum assembly	101
5.3.5 CEP19L is required for flagellar assembly and localisation of CEP164A	103
5.3.6 FBF1 is required for flagellum assembly	105
5.4 Two transition fibre proteins are required for flagella to assemble to their final length	107
5.4.1 CFAP52 is required for flagella to reach their final length	107
5.4.2 TFP49 is required for flagella to assemble to the correct length	109
5.5 Transition fibre proteins required for flagellum length control	111
5.5.1 TFP122 is a regulator of flagellum length	111
5.5.2 TFP15 is a regulator of flagellum length	113
5.6 Transition fibre proteins with a phenotype not related to flagellar function	115
5.6.1 Cullin4b is required for cell division but does not affect flagellar length	115
5.6.2 SPB1 may be required for normal cell cycle progression	117
5.7 Ciliopathy proteins and their effect on intraflagellar transport	119
5.8 No discernible phenotype following knockdown of eight transition fibre proteins	122
6. Discussion	125
6.1 The molecular composition and structure of transition fibres	125
6.1.1 Molecular composition	125
6.1.2 Comparison of the structural dimensions	126
6.2 G1 to S-phase: modelling transition fibre assembly	127
6.3 The roles of transition fibres in flagellum assembly and length control	130
6.3.1 Transition fibres are central drivers of flagellum length control	130
6.3.2 Transition fibres are critical for flagellum assembly in trypanosomes	134
6.3.3 Transition fibres and IFT – insights into human ciliopathies	136
6.4 Conservation of transition fibre proteins across eukaryotic organisms	137
6.4.1 Localisation of conserved orthologues	138
6.4.2 Kinetoplastid specific transition fibre components	138
6.5 Conclusions	139

Bibliography.....	140
-------------------	-----

List of Figures

Chapter 1: Introduction

Figure 1.1: Schematic representing major ciliary structures.....	12
Figure 1.2: Examples of the first observations of transition fibres.....	13
Figure 1.3: Cross-section of the distal end of the mature basal body in <i>Trypanosoma brucei</i> and model showing 3D structure of transition fibres including the position of known components....	15
Figure 1.4: Schematic showing the key roles of transition fibres in ciliogenesis.....	16
Figure 1.5: Position of structures at flagella base and major organelles marking cell cycle progression in <i>Trypanosoma brucei</i>	21
Figure 1.6: Key stages of the basal body life cycle in <i>Trypanosoma brucei</i>	23

Chapter 2: Methods

Figure 2.1: PCR-only method for endogenous tagging, PCR components and PCR cycling conditions.....	27
--	----

Chapter 3: Transition fibre composition, structure and spatial arrangement

Figure 3.1: Examples of basal body proteins searched into tryptag.org.....	34
Figure 3.2: Map of the pPOTv7 plasmid.....	35
Figure 3.3: Map of the pPOTv6 plasmid.....	36
Figure 3.4: Localisation of known transition fibre proteins.....	38
Figure 3.5: Localisation of conserved putative transition fibre proteins.....	40
Figure 3.6: Localisation of putative transition fibre proteins conserved only in the excavata super group.....	42
Figure 3.7: Localisation of trypanosome specific putative transition fibre proteins.....	45
Figure 3.8: Dimension of transition fibres measured by automated analysis.....	47
Figure 3.9: Quantitative analysis of 29 putative transition fibre proteins.....	50
Figure 3.10: Trypanosome transition fibres by thin sections transmission electron microscopy...53	
Figure 3.11: CEP164A mScarlet amplicon loaded on an agarose gel.....	54
Figure 3.12: Live cell images of CEP164A tagged with mScarlet.....	55
Figure 3.13: Amplicon of each putative transition fibre gene loaded onto an agarose gel.....	56

Figure 3.14: Transition fibre proteins tagged with mNeonGreen co-expressed with CEP164A mScarlet.....	57
Figure 3.15: Transition fibre proteins tagged with mNeonGreen with immunolabelled RP2.....	60

Chapter 4: Cell cycle regulation, assembly and conservation of transition fibres

Figure 4.1: Cell cycle regulated Rabl2b brighter at the old flagellum.....	65
Figure 4.2: Cell cycle regulated TFP104 brighter at the old flagellum.....	66
Figure 4.3: Cell cycle regulated TFP39 brighter at the old flagellum.....	67
Figure 4.4: Cell cycle regulated TFP122 brighter at the old flagellum.....	68
Figure 4.5: Cell cycle regulated TFP72 brighter at the new flagellum.....	69
Figure 4.6: Cell cycle regulated TFP59 brighter at the new flagellum.....	70
Figure 4.7: Schematic showing basal body duplication.....	72
Figure 4.8: Timing of transition fibre protein recruitment in relation to basal body duplication...	73
Figure 4.9: Timing of transition fibre protein recruitment relative to arrival of CEP164A.....	75
Figure 4.10: Schematic of timing of transition fibre protein recruitment.....	77
Figure 4.11: Widely conserved transition fibre proteins.....	81
Figure 4.12: Transition fibre proteins conserved only within excavata super group.....	83
Figure 4.13: Trypanosome specific transition fibre proteins.....	84

Chapter 5: Functional analysis of transition fibre proteins by RNAi

Figure 5.1: Schematic of RNAi vector.....	86
Figure 5.2: Depletion of TFP68 RNAi.....	90
Figure 5.3: Depletion of CEP90 RNAi.....	92
Figure 5.4: Depletion of Rabl2b RNAi.....	94
Figure 5.5: Depletion of OFD1 RNAi.....	96
Figure 5.6: Depletion of CEP19L RNAi.....	98
Figure 5.7: Depletion of FBF1 RNAi.....	100
Figure 5.8: Depletion of CFAP52 by RNAi.....	102
Figure 5.9: Depletion of TFP49 by RNAi.....	104
Figure 5.10: Depletion of TFP122 by RNAi.....	106
Figure 5.11: Depletion of TFP15 by RNAi.....	108
Figure 5.12: Depletion of Cullin4b by RNAi.....	110
Figure 5.13: Depletion of SPB1 by RNAi.....	112
Figure 5.14: IFT172 labelling with RNAi of conserved ciliopathy proteins.....	115

Figure 5.15: RNAi of transition fibre proteins with no discernible phenotype.....	117
---	-----

Chapter 6: Discussion

Figure 6.1: Model of transition fibre assembly and the arrangement of proteins along the transition fibre.....	123
Figure 6.2: Model of transition fibre proteins in flagellum locking.....	127
Figure 6.3: Model of transition fibre proteins in flagellum assembly.....	129

List of Tables

Table 2.1: Solutions, drug concentrations and electroporation conditions used for transfection of procyclic form <i>Trypanosoma brucei</i>	28
Table 2.2: PCR components and conditions used for RNAi plasmid generation and transfection.....	31
Table 2.3: List of primary and secondary antibodies for immunofluorescence.....	32
Table 3.1: Localisation in different organisms of known transition fibre proteins.....	39
Table 3.2: Localisation in different organisms of conserved putative transition fibre proteins.....	41
Table 3.3: Reported localisations of excavata specific putative transition fibre proteins.....	44
Table 3.4: Summary table of putative transition fibre proteins.....	61

Abbreviations

F – Flagella

NF – New flagellum

OF – Old flagellum

N – Nucleus

K – Kinetoplast

Post M – Post mitotic

aa – amino acids

BB – basal body

MBB – Mature basal body

PBB – Probasal body

nMBB – Mature basal body of new flagellum

nPBB – Probasal body of new flagellum

TZ – Transition fibre

FP – Flagellar pocket

FPM – Flagellar pocket membrane

BLAST – Basic local alignment search tool

Reciprocal best blast (RBB)

bp – base pairs

Cas9 - CRISPR-associated gene 9

Tet – Tetracycline

RNAi - RNA interference

CRISPR – Clustered regularly interspaced short palindromic repeats

DMSO – Dimethyl sulphoxide

FCS – Foetal calf serum

mNG – mNeonGreen

mSc – mScarlet

IFT – Intraflagellar transport

kDa – Kilodalton

PBS – Phosphate buffered saline

vPBS – Voorheis modified phosphate buffered saline

PCR- polymerase chain reaction

TEM – transmission electron microscopy

1. Introduction

1.1 Cilia and flagella in eukaryotic biology

1.1.1 Cilia and flagella in human health and disease

Eukaryotic cilia and flagella are slender microtubule-based projections that extend, with a contiguous membrane, out from the cell body. The terms cilia and flagella are often used interchangeably, although flagella tend to be associated with cell motility. The cilium is composed of a set of substructures that will be discussed in a later section. The vast majority of cell types in complex organisms have the capacity to grow a cilium or flagellum and the core structure has been well conserved throughout eukaryotic evolution (Mitchell, 2017, Satir et al., 2008, Wan and Jekely, 2020).

Primary cilia are non-motile structures that possess a 9+0 microtubule configuration (Fig. 1.1) and act as sensory organelles, functioning in a number of critical signalling pathways. Most notably are the receptor tyrosine kinase (RTK), Hedgehog (Hh) and Wntless (Wnt) signalling pathways that drive fundamental cellular processes such as cell cycle control, differentiation, migration and cell polarity among others (Christensen et al., 2012, Goetz and Anderson, 2010, Wallingford, 2010).

Often described as 'nanomachines', motile cilia have a 9+2 microtubule conformation and are known to play a more mechanical role in biological systems (Fig. 1.1). Epithelial surfaces, such as the trachea, are decorated with cilia to sweep particulate matter from the mucus secreting tissues (Ganesan et al., 2013). The highly co-ordinated beating of motile cilia also generates transport networks in the brain by driving the flow of cerebrospinal fluid into the ventricles, regulating neurogenesis during embryonic development (Faubel et al., 2016, Nelson and Wright, 1974, Olstad et al., 2019).

In addition, cilia and flagella also play a key role in sexual reproduction, a critical step in the life cycle of complex organisms. In females, the synchronous beating of cilia in the fallopian tube faithfully directs the ovum towards the uterus while in male sperm, the flagellum is used for motility, propelling the sperm towards the ovum for fertilisation (Eddy et al., 1978, Inaba, 2003, Lyons et al., 2006). Their diverse cellular and physiological functions mean that cilia and flagella are highly relevant organelles in the context of human health and disease (Christensen et al., 2007).

Ciliopathies are a class of diseases that arise from the dysfunction of primary and motile cilia. Ciliopathies include diseases such as Bardet-Biedl syndrome, Joubert syndrome, nephronophthisis and retinitis pigmentosa (Hildebrandt et al., 2011). Mutations in genes involved in ciliary function

result in a range of developmental defects, most commonly affecting the heart, brain, kidneys, retina and the skeletal system.

Ciliary dysfunction is also linked with fertility defects in both males and females (Ding et al., 2020, Inaba, 2003, Wallmeier et al., 2020). Although many ciliopathies are rare genetic disorders, cilia also play a role in the development of more common diseases like obesity and cancer (Fliegauf et al., 2007, Goetz and Anderson, 2010, Veland et al., 2009). Unsurprisingly, ciliopathies are one of the fastest growing families of genetic disease, with over 50 new causal loci identified between 2005 and 2012 (Wei et al., 2015, Davis and Katsanis, 2012). Despite this, the molecular composition of many ciliary sub-structures is still poorly understood.

1.1.2 Many eukaryotic pathogens require a flagellum

The eukaryotic flagellum is fundamental to the life cycle of many eukaryotic pathogens. Giardiasis is a common intestinal disease that plagues many of the world's most impoverished regions. The disease is caused by the parasite *Giardia intestinalis* which possesses eight flagella that are used for motility, host attachment and cell division within the host (Nosala and Dawson, 2015). The malaria parasite, *Plasmodium falciparum*, contributes to an immensely high disease burden and mortality throughout the developing world (Guinovart et al., 2006, Liu et al., 2021). *Plasmodium* spp. assemble a flagellum during sexual differentiation of male gametocytes and is therefore essential to the life cycle of the parasite. Kinetoplastid organisms such as *Trypanosoma brucei*, *Trypanosoma cruzi* and *Leishmania* spp. are the causative agents of trypanosomiasis (also known as sleeping sickness), Chagas' disease and leishmaniasis in humans and animals (Filardy et al., 2018, Stuart et al., 2008). Kinetoplastid pathogens also depend on their flagella for motility, attachment, cell division and other functions that are essential to their pathogenicity in the host.

1.2 Structure of cilia/flagella

Cilia/flagella have a highly conserved core structure characterised by a microtubule-based axoneme that runs through its core in a 9+2 conformation in motile cilia and a 9+0 in non-motile (primary) cilia (Fig. 1.1) (Carvalho-Santos et al., 2011, Reiter, 2012). A single microtubule is composed of 10-15 linear protofilaments that are themselves made up of heterodimers of α and β -tubulin (Mohri, 1968, Wade and Hyman, 1997). The basal body is a microtubule-based barrel shaped organelle that sits at the ciliary base and nucleates the axonemal microtubules from which the flagellum is assembled (Fig. 1.1). Basal bodies are structurally equivalent to centrioles; however basal bodies are distinctly associated with cilia and flagella assembly. It should be noted that centrioles can also translocate to the cell surface membrane to form a cilium.

The flagellum/cilium is nucleated by a mature basal body (also known as mother centriole) with an immature probasal body (also known as daughter centriole) positioned adjacent to the mature basal body (Fig. 1.1). This probasal body will go on to mature and nucleate a flagellum as the cell progresses through to the next cell cycle. A hallmark of basal body maturation is the acquisition of appendages called transition fibres (Reiter, 2012).

Transition fibres are a set of blade-like structures that protrude from the distal end of the mature basal body in a 9-fold radial symmetry (Fig. 1.1, Fig. 1.3). Transition fibres and distal appendages are analogous structures and the two terms are often used interchangeably; we will refer these appendages as transition fibres. Transition fibres are essential for connecting the mature basal body to the ciliary membrane and also facilitate the recruitment and import of intraflagellar transport (IFT) complexes into the ciliary compartment (Graser et al., 2007, Kumar et al., 2021, Reiter, 2012, Tanos et al., 2013, Wei et al., 2015, Wei et al., 2013). Transition fibres are the ciliary sub-structures that will be the focus of this project.

The transition zone (Fig. 1.1) is a well characterised ciliary structure that is located distal to the transition fibres. It is composed of microtubule to ciliary membrane connectors known as y-links, that form pores restricting the passage of proteins (Goncalves and Pelletier, 2017). The transition zone can therefore act as a gate to control entry and exit of proteins, compartmentalising the ciliary matrix from the cytoplasm of the cell body (Dean et al., 2016, Reiter and Leroux, 2017). Many ciliopathy genes correspond to proteins that localise to the transition zone region. Most notable are the Barted-Beidl syndrome (BBsome) and Meckel syndrome complexes. The BBsome complex is made up of seven proteins that are involved in import and export of molecules to the ciliary compartment by interacting with the IFT complexes (Lechtreck et al., 2009, Nachury et al., 2007). It has been proposed that transition fibres may also play a role in the ciliary gate, as they act as a molecular checkpoint at the border of the cytoplasm and the ciliary compartment, regulating recruitment and import of proteins involved in ciliary assembly (Reiter, 2012).

Flagellum assembly is driven by IFT machinery that shuttles cargo containing flagellum assembly proteins, such as tubulin, up and down the axonemal microtubules (Fig. 1.1) (Qin et al., 2004). IFT protein complexes are recruited and imported into the flagellar compartment by transition fibres (Goncalves and Pelletier, 2017, Reiter, 2012, Wei et al., 2015, Wei et al., 2013). These IFT complexes then assemble into polymers called trains which are bound to kinesin-II and IFT-dynein motors that shuttle the cargo from the flagellar base to the distal tip for assembly and from the tip to the base for disassembly. Current models indicate that the flagellum is constructed and maintained by regulating IFT import and loading at the flagellar base and striking a balance between assembly and

disassembly at the distal tip (Marshall and Rosenbaum, 2001, Qin et al., 2004, Reiter, 2012, Rosenbaum, 2002, Wei et al., 2015, Wei et al., 2013).

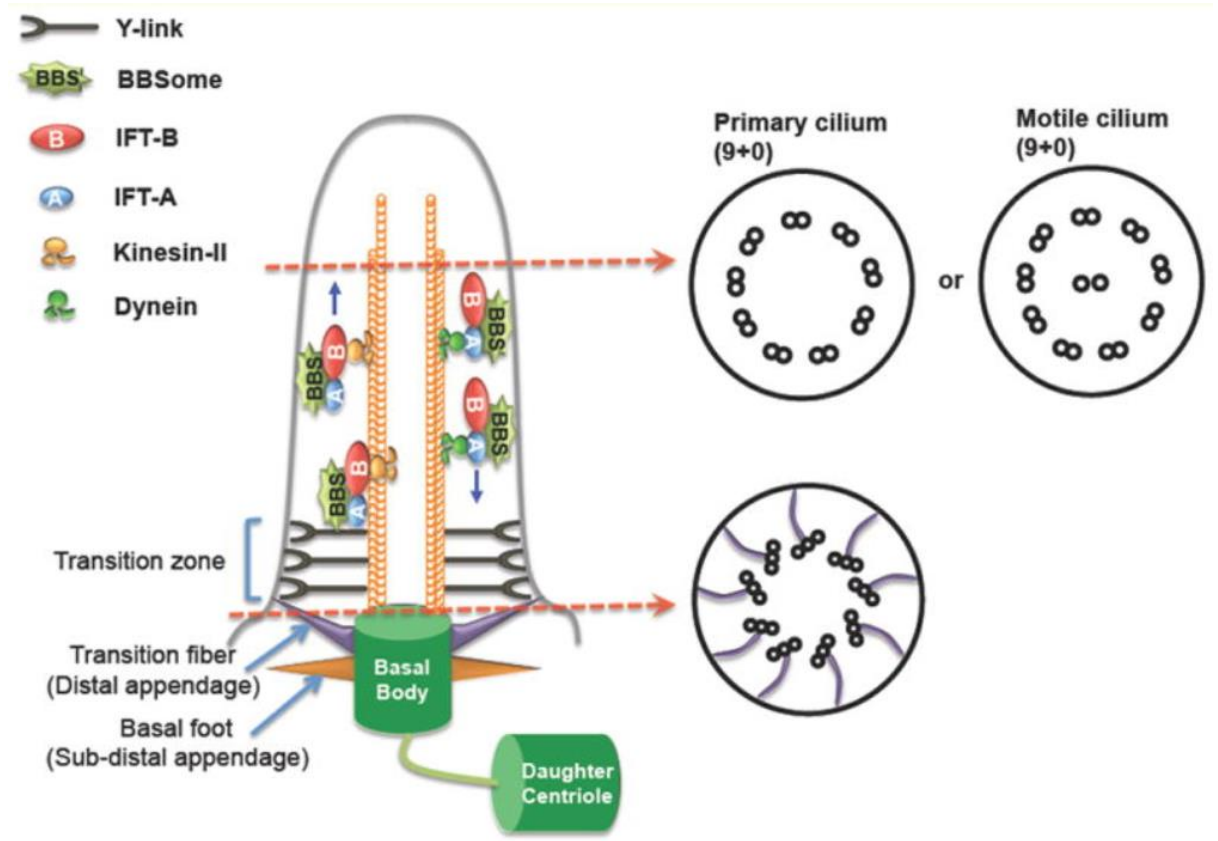


Figure 1.1 Schematic representing major ciliary structures including mature basal body, probasal body (daughter centriole), transition zone, transition fibres, IFT and the microtubule conformation of primary and motile cilia (from Wei et al. 2013).

1.3 Transition fibres: background and known structure

1.3.1 Historical background and discovery

The structures now known as transition fibres/distal appendages were first observed by electron microscopy (EM) in 1958. A model of the leukocyte centriole was proposed which depicted two sets of 9 symmetrical appendages first called 'massules' that were physically connected to the centriole surface (Fig. 1.2a) (Bessis and Breton-Gorius, 1958). The two sets of appendages would later be termed distal and sub-distal appendages; however, the proximal-distal polarity of a centriole or basal body was still unknown.

The same distal structures were recognised in jellyfish spermatids using thin section transmission electron microscopy (TEM). The author described 9 'satellites' emanating from the nine triplets of the centriole that were specifically found on the mother centriole and noted that these supportive structures may anchor a flagellum (Fig. 1.2b) (Szollosi, 1964).

The now called transition fibres/distal appendages were later named 'alar sheets' after being observed on the basal body of the rhesus monkey oviduct (Fig. 1.2c, Fig. 1.2d) and were recognised again on the centrioles of human leukocytes where the term transitional fibre was first coined (Anderson, 1972, O'Hara, 1970). Early EM data suggested that the presence of transition fibres on basal bodies and centrioles is ubiquitous and the symmetrical 9-fold conformation is conserved. The term transition fibre was established and is currently the commonly used term along with distal appendages, which are the two names for these structures.

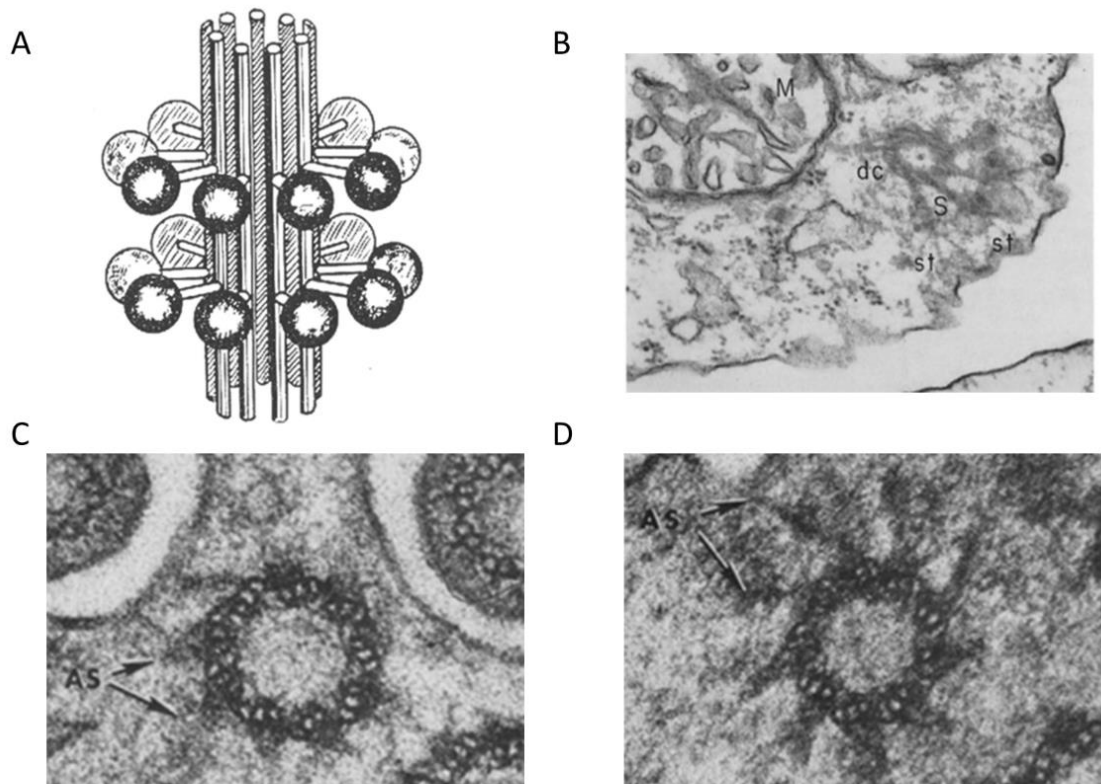


Figure 1.2 (A) Model of 9-symmetrical appendages surrounding the leukocyte centriole first called 'massules' (Bessis and Breton-Gorius, 1958). (B) TEM section of 9 'satellites' surrounding the mother centriole in Jellyfish spermatids (Szollosi, 1964). (C) TEM of 'Alar sheets' surrounding basal body of Rhesus monkey oviduct (Anderson, 1972). (D) TEM of 'Transitional fibres' surrounding centriole of human leukocytes (O'Hara, 1970).

1.3.2 Structure and composition

Transition fibres surround the distal end of the mature basal body in a 9-fold radial symmetry (Fig. 1.3a). Recent EM of mammalian transition fibres revealed a head density at the end of the transition fibre blade that extends around 110nm from the basal body and from a lateral view attach to the basal body at an angle of 34° (Bowler et al., 2019). In trypanosomes, a study has shown that the

transition fibres extend 75nm from the basal body - although head densities were not reported in this study (Trepout et al., 2018).

Five core transition fibre proteins have so far been identified in mammalian cells using SILAC (stable isotope labelling by amino acids in cell culture) in conjunction with mass spectrometry. The technique uses isotopic labelling to detect differences in protein abundance. Due to the specificity of transition fibres to the mother centriole, this was used to differentiate transition fibre proteins from other centriolar components. Candidates were examined by fluorescence microscopy for their subcellular localisation and five transition fibre proteins were identified and later mapped based on their diameter around the basal body (Fig. 1.3b) (Tanos et al., 2013, Yang et al., 2018). Three of these proteins were unknown (CEP83, SCLT1 and FBF1) and two (CEP164 and CEP89) had been characterised in previous studies (Graser et al., 2007, Sillibourne et al., 2011).

Depletion of the core mammalian transition fibre components by RNA interference (RNAi) revealed a hierarchy of assembly. Each protein was depleted and an adjacent protein was visualised with immunofluorescence (IF) – a loss of localisation following depletion of a neighbouring protein shows a dependence of one protein for the localisation of another and thus, the order of assembly can be deduced (Tanos et al., 2013). Using this method, it has been shown that CEP83 (Centrosomal protein of 83 kDa) lies at the root of the appendage ‘blade’, followed by CEP89, SCLT1, FBF1 and CEP164 – which occupies the space furthest from the centriole (Fig. 1.3b).

Each protein within the structure was immunolabeled and visualised with super-resolution direct stochastic optical reconstruction microscopy (dSTORM). The diameter of each transition fibre protein around the mature basal body was measured based on their fluorescent signal. The measurements made were in concordance with the hierarchical model of assembly proposed previously (Tanos et al. 2013) – although further analysis using dSTORM indicated that FBF1 is likely to sit in the matrix between the blades and not on the fibre itself (Fig. 1.3b) (Yang et al., 2018).

Transition fibres are required for ciliary assembly, yet some of the core structural proteins appear to be lost or divergent in other ciliated eukaryotes (Tanos et al., 2013). CEP164 and FBF1 are the only core components conserved in non-vertebrate ciliated/flagellated organisms. This suggests that either transition fibre composition varies significantly between ciliated eukaryotes or that there are additional conserved structural components that have not yet been identified.

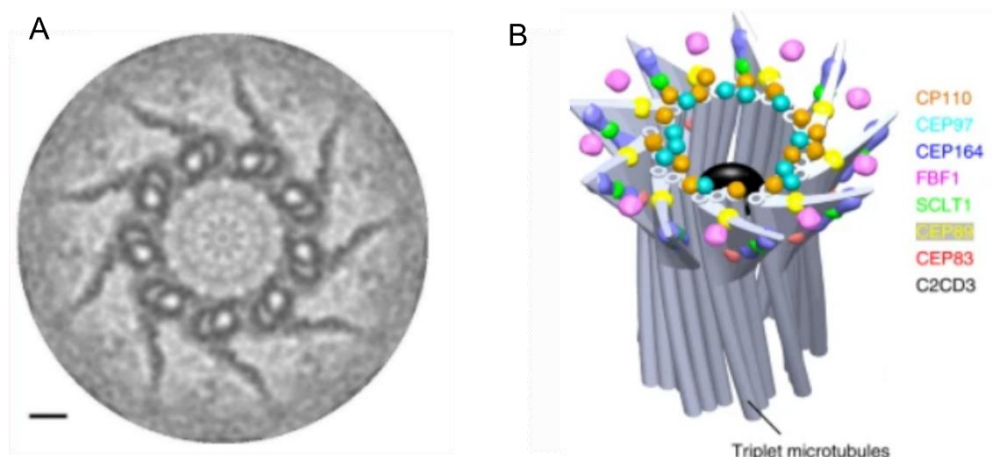


Figure 1.3 (A) Cross section of the distal end of the mature basal body Trypanosoma brucei by scanning transmission electron microscopy (STEM). The cartwheel hub of the basal body is shown at the centre with the outer doublet microtubules of the basal body surrounding. Transition fibres protrude 75nm from the outer doublet microtubules at an angle of 50° in a 9-fold symmetry. Scale bar = 40nm (Trepout et al. 2018). (B) Model showing the positions of each protein in the transition fibre structure represented in 3-dimensional space. CP110, CEP97 and C2CD3 are bona fide centriolar/basal body proteins (Yang et al. 2018).

1.4 Transition fibres in flagellum assembly and length control

1.4.1 Roles of transition fibres in cilium biogenesis

Transition fibres are essential for the formation of a cilium or flagellum and conserved transition fibre components CEP164 and FBF1 are both crucial. In the earliest stages of ciliogenesis, distal appendage vesicles (DAVs) dock to the distal appendages surrounding the mother centriole (Fig.4) (Schmidt et al., 2012) and the membrane remodelling proteins EHD1 and SNAP29 enable fusion of DAVs to form a primary ciliary vesicle (PVC)(Fig.4) (Lu et al., 2015).

Studies have shown that CEP164 links machinery for ciliary membrane biogenesis to the mother centriole (Fig. 1.4) (Chaki et al., 2012, Graser et al., 2007, Liu et al., 2021). CEP164 also recruits and forms a complex with a conserved coiled-coil protein Chibby (Cby) at the transition fibres. Cby goes on to interact with membrane trafficking protein Rabin8 and subsequently recruits vesicles containing the small GTPase Rab8, which supports the fusion of the PVC to the cell surface forming a pre-ciliary membrane (Fig. 1.4) (Lu et al., 2015).

In addition to this, CEP164 induces the de-repression of ciliogenesis. CEP164 firstly recruits and is phosphorylated by Tau-tubulin kinase II (TTBK2), which results in the removal of a key suppressor of ciliogenesis, CP110 (Centriolar coiled-coil protein 10), from the distal end of the mother centriole (Cajane and Nigg, 2014). The removal of CP110 and ciliary membrane formation triggers axoneme extension (Fig. 1.4) (Cajane and Nigg, 2014, Yadav et al., 2016). It has been shown that mutations

in CEP164 are linked to nephronophthisis-related ciliopathies (NPHP-RC) (Chaki et al., 2012, Devlin et al., 2020).

It is known that transition fibres are essential docking sites for IFT complexes at the ciliary base (Deane et al., 2001). In *C. elegans*, knockdown of FBF1 by RNAi affected the localisation of several IFT particles including IFT52 and IFT140, which in turn lead to the formation of truncated cilia (Wei et al., 2013). A number of IFT components were found to accumulate at the ciliary base, indicating that transition fibre protein FBF1 is required for the entry of assembled IFT complexes into the ciliary compartment. Localisation of IFT88, IFT81 and IFT20 are also dependent on the presence of CEP164 at the transition fibres (Cajane and Nigg, 2014, Schmidt et al., 2012). This data indicates that transition fibres play a role in ciliary compartmentalisation as well as assembly.

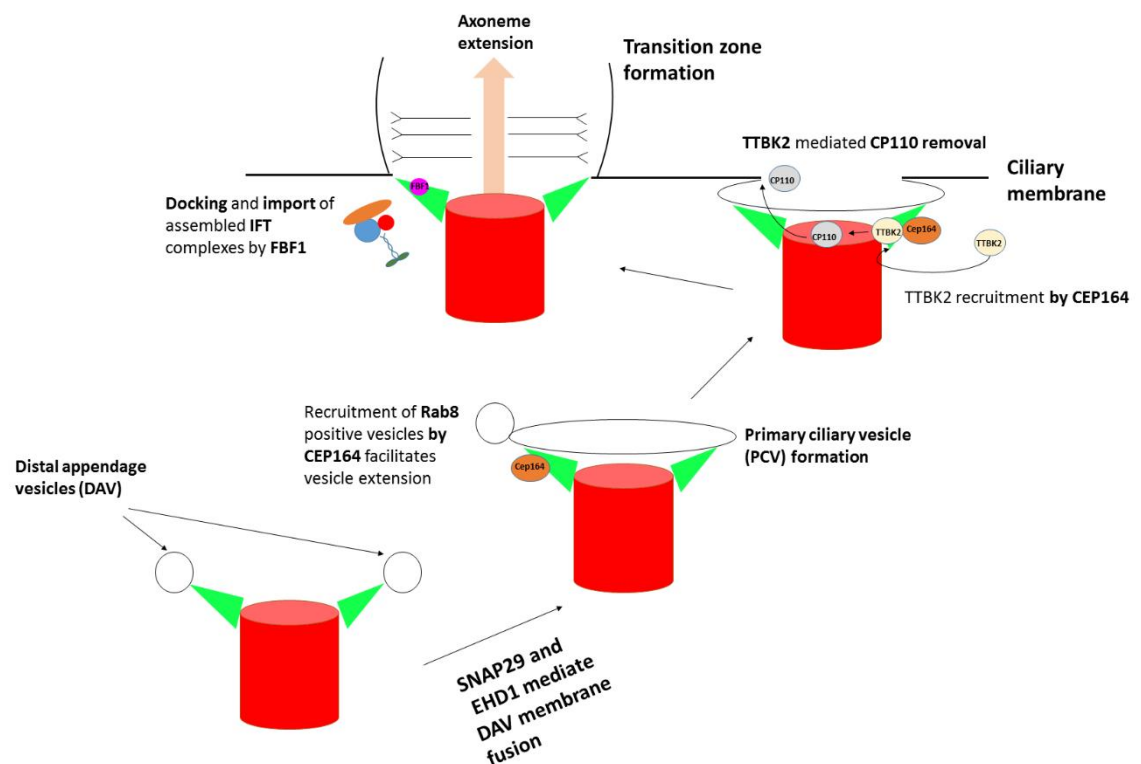


Figure 1.4 Schematic showing key roles of transition fibres in ciliogenesis. Distal appendage vesicles are recruited to the transition fibres of the mature basal body. Conserved transition fibre component CEP164 (orange) mediates vesicle recruitment for membrane docking. CEP164 also drives CP110 (grey) removal via TTBK2 (Yellow) recruitment which is critical for the de-repression of cilium biogenesis. FBF1, indicated in pink, is required for the docking and selective import of IFT complexes to support axoneme assembly.

1.4.2 Balance point model

Some ciliates, such as the green algae *Chlamydomonas reinhardtii*, assemble two flagella in a synchronous manner where tubulin and other flagellum assembly proteins are incorporated into

each flagellum simultaneously (Cavalier-Smith, 1974, Hoog et al., 2014). Pools of flagellum assembly proteins can therefore enter the flagellum at the same rate and distinct regulation of the two flagella is not required for their assembly. Thus, flagellum length can be regulated in both flagella by striking a balance between the assembly and disassembly at the distal tip. This model of assembly indicates that flagellum length is controlled by IFT entry at the base and IFT turnover at the distal tip (Marshall and Rosenbaum, 2001, Qin et al., 2004). This balance is regulated by the recruitment, import and loading of IFT proteins at the flagellar base. Transition fibres serve as a docking site for IFT proteins at the flagellar base and are therefore likely involved in the regulation of flagellum assembly and length control. A model of equal access fits well with a system that undergoes synchronous assembly of two flagella. However, in a cell that assembles flagella asynchronously or maintains an existing flagellum separately, another mechanism may be required that requires greater level of regulation at the flagellar base.

1.4.3 Growth and lock model

In other ciliated eukaryotes, there is a need to assemble flagella whilst simultaneously maintaining the length of existing flagella within the same cell. Such is the case for flagellates such as *Trypanosoma brucei*, *Leishmania* Spp. and *Giardia* spp. - the latter of which must assemble four flagella and while four flagella inherited from the mother cell remain (Dawson and House, 2010). In the case of *Trypanosoma brucei*, a second new flagellum is assembled at the start of S-phase alongside an existing old flagellum. A model of equal access would be unlikely to explain this as the two flagella are distinct in terms of the timing of their assembly.

Recently, a growth and lock model was proposed in trypanosomes which states that a lock is placed on the old flagellum to maintain its length while the cell undergoes assembly of a new flagellum alongside (Bertiaux et al., 2018). In this study, Bertiaux and colleagues (2018) showed that depletion IFT kinesins affected the assembly of a new flagellum as expected. However, an increase in IFT rate did not lead to further elongation of the old flagellum. This indicates that the old flagellum is locked at a given length, unaffected by further entry of IFT.

Proteins at the transition fibres have been shown to be essential drivers of flagellum assembly in ciliated eukaryotes (Graser et al., 2007, Singla et al., 2010, Wei et al., 2015, Wei et al., 2013). It was recently shown that transition fibre protein CEP164C contributes to old flagellum locking in trypanosomes. CEP164C is a protein found exclusively at the transition fibres of the old flagellum in dividing cells. Upon depletion of CEP164C by RNAi, cells assembled a longer old flagellum and ectopic expression revealed that cells assemble shorter flagella (Atkins et al., 2021). Studies have shown transition fibres play a role in both flagellum assembly and maintenance. It can be concluded that transition fibres are key structures that facilitate flagellum assembly and

maintenance. For this to be viable, it is reasoned that there must be some molecular differences between the base of the old and new flagellum to regulate assembly at the new flagellum and simultaneously activate old flagellum locking.

In conclusion, it is clear that transition fibre proteins play important roles in flagellum assembly and length regulation in various ciliary systems. However, it is likely that the molecular composition is not fully understood. The aim of this PhD is to identify novel transition fibre components and assess their roles in flagellum assembly and maintenance in trypanosomes.

1.5 Studying transition fibres in *Trypanosoma brucei*

1.5.1 *Trypanosoma brucei* is a unicellular eukaryotic pathogen

Trypanosoma brucei is a flagellated unicellular eukaryotic pathogen that causes African Trypanosomiasis or sleeping sickness in humans. *T. brucei* belongs to the order Kinetoplastida, which includes other human pathogens such as *Trypanosoma cruzi* and Leishmania Spp. which, as mentioned previously, are the causative agents of Chagas disease and Leishmaniasis. Diseases caused by kinetoplastid organisms create a significant economic and social burden in many developing countries (Fildard et al., 2018, Stuart et al., 2008). Human African trypanosomiasis (HAT) is caused by the parasites *Trypanosoma brucei gambiense* and *Trypanosoma brucei rhodesiense* and is prevalent in more than 30 African countries, although fewer than 1000 cases were reported in endemic countries in 2018 (Gao et al., 2020, Simarro et al., 2010). Animal African trypanosomiasis (AAT) or Nagana is caused by the parasites *Trypanosoma brucei brucei*, *Trypanosoma vivax* and *Trypanosoma congolense*. The prevalence of AAT in livestock animals is approximately 16% in affected regions and remains a significant constraint to livestock health in some of the poorest communities across sub-Saharan Africa (Muhanguzi et al., 2017, Odeniran and Ademola, 2018).

1.5.2 The flagellum is essential for the pathogenicity of *Trypanosoma brucei*

T. brucei alternates between a procyclic form within tsetse fly insect vector and a bloodstream form within the mammalian host, where it relies on a flagellum for motility. The flagellum of *T. brucei* is attached along the length of the cell body (Fig. 1.5) and acts as the sole driver of motility. This allows the parasite to traverse through a range of tissues in the tsetse from the midgut through to the salivary glands where it attaches as an epimastigote form. From the salivary glands, the parasite can then be transferred to the bloodstream of a mammalian host following a blood meal (Griffiths et al., 2007, Langousis and Hill, 2014, Matthews, 2005, Vickerman, 1985).

The flagellum is essential for survival in the hostile bloodstream of a mammalian host. Endocytosis and exocytosis occur exclusively at the flagella pocket and therefore it is essential for the exchange of macromolecules that influence host-parasite interactions (Allen et al., 2003, Field and Carrington,

2009). Trypanosomes have also been shown to express a variable surface glycoprotein (VSG) on their cell surface which aids immune evasion by VSG coat switching. Glycosylphosphatidylinositol phospholipase C (GPI-PLC) is located on the flagellar membrane and cleaves the GPI anchor of the VSG, facilitating VSG switching and thus helping the cell evade immune attack (Hanrahan et al., 2009). The flagellum has also been shown to play mechano-physical roles in immune evasion. For example, hydrodynamic drag, caused by the positioning of the flagellum along the cell body, forces immune complexes in the opposite direction of the cell, protecting the cell against immune attack (Engstler et al., 2007).

The flagellum is an essential site for signalling pathways that regulate a number of critical cellular processes and contribute towards parasite pathogenicity. For example, flagellar cyclic adenosine monophosphate (cAMP) signalling has been shown to regulate parasite progression through host tissues (Shaw et al., 2019).

The trypanosome flagellum is also an integral part of cell division and morphogenesis (Vaughan, 2010). Cells lacking a flagellum lose their polarity, have compromised morphology and are unable to divide (Kohl et al., 2003). This is in part because the flagellum controls the inheritance of subcellular organelles, likely due to the physical connections between them.

In summary, the trypanosome flagellum is fundamental to processes like cell division, motility, morphogenesis and host-parasite interactions and is therefore essential for survival in the insect vector as well as its pathogenicity in the mammalian host.

1.5.3 Organelle position and segregation marks the stages of the trypanosome cell cycle
Organelle position and segregation is highly coordinated throughout the trypanosome cell cycle which can be used to trace the cell cycle stages by microscopy. At its proximal base, the trypanosome flagellum is nucleated by a mature basal body (MBB) which is linked by transition fibres to the flagellar pocket (Fig 1.5a). The flagellar pocket (FP) can be defined as an invagination of the plasma membrane where the flagellum emerges from the cell body (Fig. 1. 5a) (Field and Carrington, 2009). The basal body links to the mitochondrial DNA known as the kinetoplast (K) (Fig. 1. 5a, b) via a structure known as the tripartite attachment complex (TAC) (Ogbadoyi et al., 2003). As a result, basal body movement can directly drive the segregation of the mitochondrial genome (Robinson and Gull, 1991). The position and duplication of the kinetoplast (K), nucleus (N), basal bodies (BB) and flagella (F) can be used to mark the different stages of the trypanosome cell cycle (Sherwin and Gull, 1989).

As with other eukaryotes, the trypanosome cell cycle broadly consists of four stages – G0/G1, S, G2 and M. The post mitotic cells (post M) is also included as a final stage following nuclear division

and prior to cytokinesis. In G0/G1, the cell possesses a nucleus, kinetoplast, and a single flagellum which is attached along the length of the cell body, subtended by a mature basal body at its proximal end. Adjacent to the mature basal body is a single probasal body. G1 can therefore be characterised by one nucleus, one kinetoplast, one flagellum and a single basal body pair at the flagellar base (Fig. 1.5b, G1).

At the G1/S border, the probasal body matures to form a new mature basal body and each mature basal body duplicates to each form a new probasal body. S-phase can therefore be defined by the duplication of basal bodies from two to four which is followed by the assembly of a new flagellum (Fig. 1.5b, S-phase). Studies have indicated that CYC2 (the trypanosome orthologue of the metazoan cell cycle initiator Cyclin E) may be involved in regulating the G1/S-phase transition in trypanosomes (Liu et al., 2013).

In S-phase, following the assembly of the new flagellum, the kinetoplast at the posterior end of the cell elongates and divides. G2 can be defined by cells that have undergone kinetoplast division, with two kinetoplasts, two flagella, four basal bodies and a single nucleus (Fig 1.5b, G2). Mitosis (M) is indicated by elongation of the nucleus and subsequent nuclear division (Fig. 1.5b, M). Cyclin b-like proteins CYC6 and CYC8 are known to be involved in initiating the early stages of mitosis (Hammarton et al., 2003, Hayashi and Akiyoshi, 2018).

Following nuclear division, post mitotic (Post M) cells can be observed with two nuclei, two kinetoplasts, two flagella and four basal bodies (Fig. 1.5, Post-M). This is the last stage prior to cytokinesis where two daughter cells in G0 are formed (Fig. 1.5b, G0). Cytokinesis in trypanosomes is a highly coordinated process that involves the extension of microtubules that make up the new cell posterior. The trypanosome polo-like kinase homolog TbPLK and TOEFAZ1 are two proteins required for the formation of a new cell posterior and therefore indispensable for cytokinetic progression in trypanosomes (Hilton et al., 2018).

The ease with which the cell cycle can be tracked makes trypanosomes an attractive model for the study of the flagellum and associated structures. Sub-structures such as the basal body, transition fibres and transition zone may have been shown to be dynamic in their molecular composition throughout the cell cycle (Dean et al., 2016, Bowler et al., 2019, Dang et al., 2017). Such differences can be easily compared in different cell cycle stages by labelling the kinetoplast, nucleus, basal bodies and target proteins. Endogenous gene tagging and DNA staining techniques have become routine in trypanosome research (Dean et al., 2015, Dean et al., 2017).

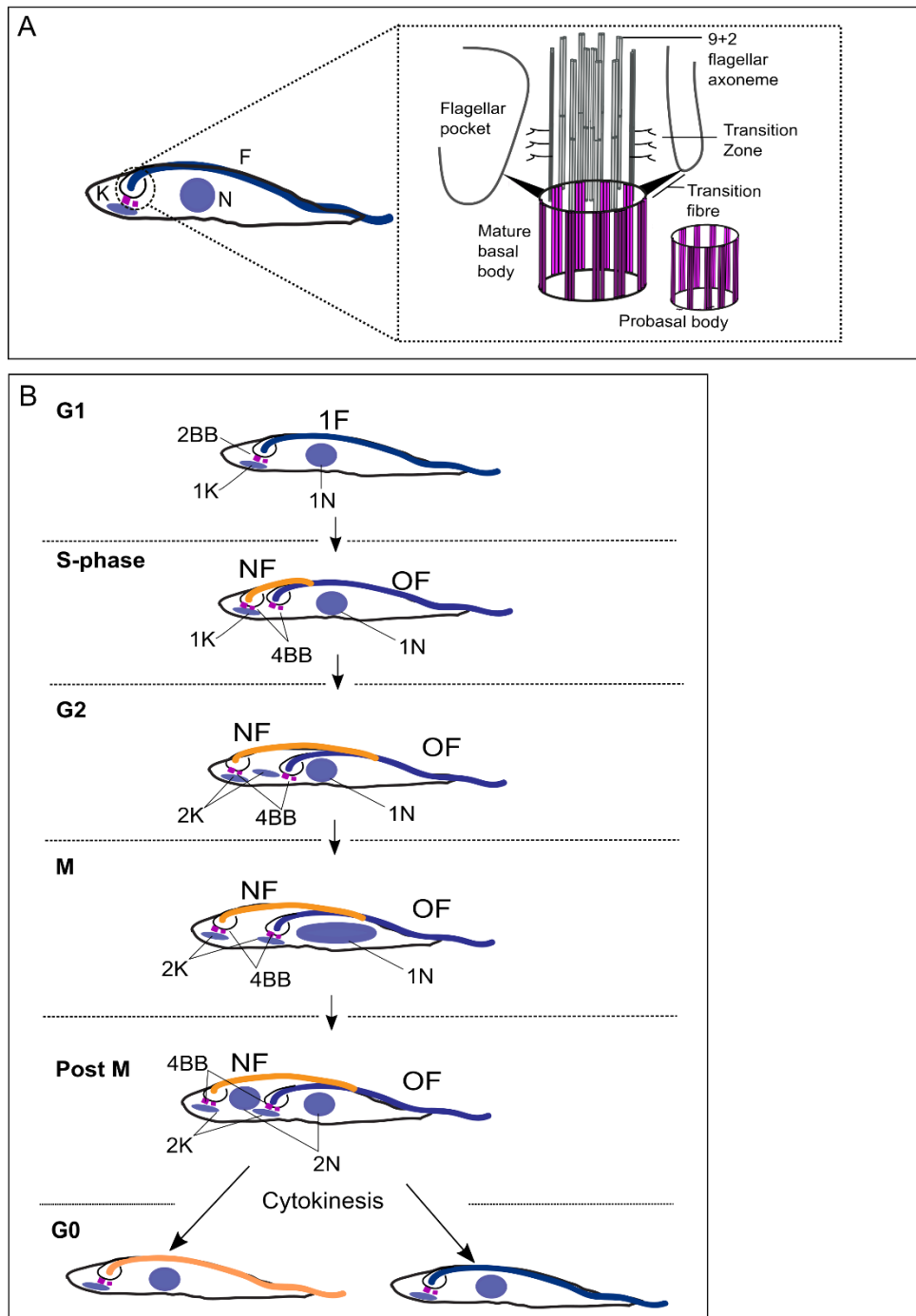


Figure 1.5 (A) Position of structures at the flagellar base in *Trypanosoma brucei*. Mature basal body shown at the base of the 9+2 flagellar axoneme and the probasal body adjacent. Transition fibres are shown at the distal end of the mature basal body docked to the flagellar pocket. The transition zone is shown further distal to the transition fibres. (B) Cell cycle progression (G1, S-phase, G2, M, Post M and G0) in *Trypanosoma brucei* marked by organelle position and number. The flagellum (F) is attached along the length of the cell body. The old flagellum and new flagellum are labelled NF and OF respectively. Basal body indicated by BB, kinetoplast indicated by K and nucleus represented by N.

1.5.4 Basal body biogenesis and transition fibre assembly

The basal body life cycle in trypanosomes has a defined set of steps that are crucial to understanding the timing of probasal body maturation/transition fibre assembly, basal body docking, and subsequent new flagellum assembly (Lacomble et al., 2010, Vaughan and Gull, 2015). In G1, cells have a single mature basal body (MBB) at the proximal end of the flagellum alongside a probasal body (PBB) which is positioned parallel to the mature basal body (Fig. 1.6, G1). At the G1/S border, the probasal body matures and docks to the membrane of the existing flagellar pocket (Fig. 1.6, G1/S).

It is likely that transition fibre assembly begins in G1 prior to basal body docking as it is known that transition fibres are required for basal body to membrane docking (Schmidt et al., 2012, Tanos et al., 2013, Wei et al., 2015). By S-phase, the existing mature basal body (MBB) and newly matured basal body (nMBB) have duplicated to form two probasal bodies (PBB, nPBB) so that the cell has a total of four basal bodies (Fig. 1.6, S). At this point in the basal body cycle, the new probasal bodies are positioned orthogonal to the corresponding mature basal body. A new flagellar pocket forms **from** the existing flagellar pocket and the new flagellum assembles from the new mature basal body (nMBB) (Fig. 1.6, S). The new flagellum then rotates around the old flagellum where new flagellum assembly can continue (Fig. 1.6, S) (Lacomble et al., 2010). The two probasal bodies reposition to become parallel to their respective mature basal body prior to mitosis (Gluenz et al., 2011).

In mammals, assembly of the primary cilium begins in G0 and continues assembly until G1. However, the resorption of cilia typically begins during S-phase and is completed by G2 prior to mitosis (Sanchez and Dynlacht, 2016). Basal body maturation, membrane docking and subsequent cilium assembly therefore occur consecutively following cytokinesis. This contrasts with the trypanosome basal body cycle, where basal body maturation begins at the G1/S border and flagellar assembly does not begin until cells have reached S-phase (Fig. 1.6). In the case of the green algae *Chlamydomonas reinhardtii*, new probasal bodies are formed during the prophase of mitosis, where transition fibres are assembled and basal bodies can be subsequently anchored to the cell membrane (O'Toole and Dutcher, 2014).

In trypanosomes it is likely that transition fibre assembly begins in G1 prior to basal body docking as they are required for the docking process. It is hypothesised that either transition fibre assembly continues through S-phase to support axoneme elongation or is completed in G1 prior to basal body docking. The trypanosome basal body cycle is highly traceable and can be achieved

by labelling basal body proteins such as SAS-6 (Dang et al., 2017). This means the timing of transition fibres can be assessed as probasal body maturation, transition fibre acquisition, docking and flagellum assembly all occur from G1 to S-phase. The G1/S-phase transition is easily defined by basal body duplication from two to four and provides a good marker to study the assembly of transition fibres.

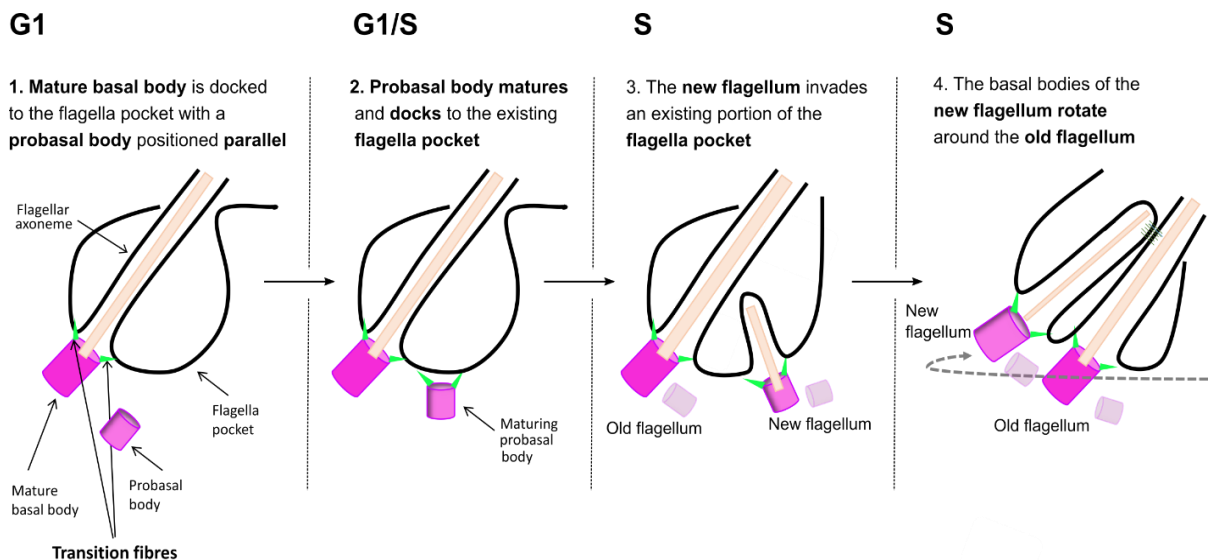


Figure 1.6 Key stages of the basal body life cycle in *Trypanosoma brucei*. (1. G1) The mature basal body (magenta) is docked to the flagella pocket by transition fibres (green). (2. G1/S) border probasal body matures acquiring transition fibres and docking to the existing flagellar pocket. (3. S) In S-phase both basal bodies assemble a new probasal body and the new flagellum assembles. (4. S) Basal body pair of new flagellum rotates around the old basal body pair of the old flagellum.

1.5.5 Molecular tools are well developed in trypanosomes

T. brucei is highly amenable to genetic engineering methods. Homologous recombination (HR) in trypanosomes repairs double strand breaks (DSBs) faster than any other repair mechanism as damage detection is tightly co-ordinated with the recruitment of HR machinery (Marin et al., 2018). Error-prone repair mechanisms, notably non-homologous end joining (NHEJ), which are prevalent in mammalian cell types, have not been reported in Trypanosomes (Glover et al., 2011). This enables more easily the integration of molecular constructs to study protein localisation and function. Endogenous gene tagging is a protein tagging method that labels a protein by inserting a sequence encoding a fluorescent protein or epitope tag at either the C or N terminus of the gene of interest. This method has been optimised in trypanosomes and is more scalable than most other tagging methods. As an example, PCR-only endogenous tagging was successful in labelling almost every gene in the trypanosome genome (Dean et al., 2015, Dean et al., 2017).

In addition to efficient gene tagging to study localisation, there are multiple tools available to study protein function. RNAi knockdown techniques are well established and a CRISPR-Cas9 toolkit has

been developed for kinetoplastids (Bellofatto and Palenchar, 2008, Beneke et al., 2017). The range and applicability of genetic tools makes *T. brucei* an attractive model for investigating transition fibre function in addition to protein localisation. Despite the critical roles of transition fibres in the assembly and length control of cilia and flagella, the molecular composition and functions of transition fibres is not well understood. Using *T. brucei* as a unique and accessible model, the molecular composition and function of transition fibre components may be examined at a scale that has not been achieved previously.

In summary, trypanosomes make an attractive model for the study of transition fibres, flagellum assembly and flagellum length control. The trypanosome flagellum is a metronome of the cell cycle and defects in flagellar assembly or function result in severe defects in cell viability more commonly than most other organelles, making the trypanosome flagellum an important focus of this model. Trypanosomes and related kinetoplastids are also eukaryotic pathogens that can be studied in the context of infectious disease. Understanding the role of transition fibres in flagellum assembly could provide an avenue to identify novel therapeutic targets, as it is known that flagellum plays an integral role in parasite pathogenicity.

The cell cycle of trypanosomes is traceable due to the coordination of organelle position and segregation and thus, can be used to easily monitor cell cycle progression. This traceability extends to the basal body cycle, where endogenous tagging of the basal bodies can be used to mark the timing of basal body maturation and new flagellum formation from G1 to S-phase. This G1/S border can be used to analyse the timing and sequence of transition fibre assembly. Studies show that only a handful of transition fibre proteins have been identified in other ciliated organisms. In this thesis, it is hypothesised that there are a number of additional components to be identified. Trypanosomes have a comprehensive molecular toolkit that includes high-throughput endogenous gene tagging, RNAi knockdown and CRISPR-Cas9 knockout methods.

The accessibility of such tools in trypanosomes means that identification of transition fibre proteins and examination of their arrangement and function can be carried out at a scale that is unachievable in other ciliary models such as mammalian cells or *Chlamydomonas reinhardtii* - where stable transgene expression is more difficult. It is concluded that *Trypanosoma brucei* is an optimal model system for investigating the following three questions:

1. What is the molecular composition of transition fibres?
2. What is the spatial and temporal arrangement of these components?
3. How do they function in the context of flagellar assembly and length control?

1.6 Project aims and objectives

1.6.1 Aims

The overarching aims of this PhD project is to use *Trypanosoma brucei* as a ciliary model to:

1. Characterise the molecular composition of transition fibres
2. Examine the spatial and temporal arrangement of transition fibre components
3. Assess the role of transition fibres in flagellum assembly and length control in trypanosomes

1.6.2 Objectives

1. Identify a cohort of transition fibre proteins by using Tryptag.org, endogenous gene tagging and quantification of their spatial arrangement using automated image analysis
2. Determine the timing and sequence of transition fibre assembly by fluorescent tagging with known marker proteins at the basal body and transition fibres
3. Assess the conservation of transition fibre proteins in eukaryotic organisms by reciprocal best blast
4. Examine the function of transition fibre components in flagellar assembly and maintenance by using RNAi and immunolabelling the flagellar axoneme. Orthologues of ciliopathy genes will be assessed in more detail to gain further mechanistic insight into human diseases.

2. Materials and methods

2.1 Trypanosome cell culture

SmOxP9 procyclic trypanosomes were cultured at 28°C in SDM-79 media supplemented with 10% v/v heat-inactivated fetal calf serum (FCS) (Sigma-Aldrich) (heat-inactivate for 30 min at 60) (Dean et al. 2015). Briefly, the powder was hydrated in 5L Mili-Q water, supplemented with 7.5 mg/L of haemin (from a 10 mg/ml stock in 0.1 M NaOH) and 2 g/L sodium bicarbonate. The pH was adjusted to 7.3 with NaOH and sterile filtered using Stericups 500 (Millipore). Cells were sub-cultured to approximately 2×10^6 cells/mL each working day (checked on Beckman Coulter V1 cell counter) and kept in culture for no longer than 6 weeks. Cells were selected and maintained on drugs according to Table 2.1b.

2.2 Gene tagging at the endogenous locus

2.2.1 PCR

N-terminal tagging primers for PCR-only tagging plasmids (pPOT) were designed using Leishgedit.org. The plasmid (pPOTv7) (Dean et al. 2015) that was used to generate the TbSAS-6 mScarlet and TbCep164a mScarlet marker cell lines contains a sequence encoding the mScarlet fluorescent protein and a G418 resistance cassette (Chapter 3, Fig. 3.2). pPOTv6 contains a 10 x ty epitope, the mNeonGreen fluorescent protein and a Blasticidin resistance gene which was used to tag each putative transition fibre protein (Chapter 3, Fig. 3.3) All components below excluding polymerase were combined into a PCR tube. PCR tubes were heated to 94°C in the thermocycler and polymerase was added. The cycle was run using the conditions described below (Table 2.1b). PCR reaction products were loaded onto a 0.8% agarose gel (80V for 35 mins) to verify the amplicon size.

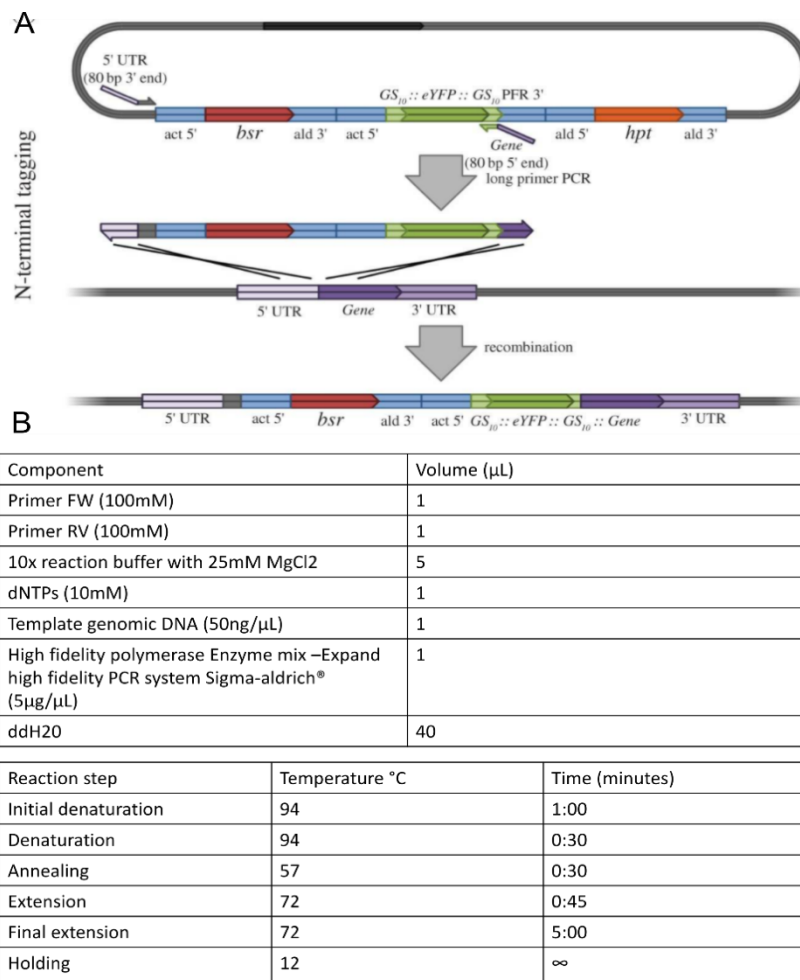


Figure 2.1 (A) PCR-only method for endogenous tagging at the N terminus of a protein (Dean et al. 2015). The construct contains from 5' to 3' a unique forward primer binding site, an actin 3' UTR, resistance resistance gene, an aldolase 3' UTR, an actin 5' UTR and a unique primer binding site (B) Volumes of each component used for each PCR reaction and cycling conditions used for PCR-only method for endogenous tagging.

2.2.2 Transfection

Parental cell line SmOxP927 was sub-cultured to 2×10^6 cells/mL every 24h for two days. Cell density was measured using a Beckman Coulter V1 cell counter. A volume of cell culture containing 1×10^7 cells transfection was centrifuged at 800 x g for 4 minutes. The supernatant was removed, and the pellet was resuspended in 500 μL transfection buffer (Table 2.1a). The suspension was transferred to a 1.5mL microcentrifuge tube containing 50μL of the PCR amplicon and then to a 2mL cuvette and electroporated in a BTX square wave electroporator using the conditions described (Table 2.1c). A negative control containing no amplicon was also electroporated. Electroporated cells were pipetted into a 50mL culture flask containing 11mL of SDM-79 media. After 6-8 hours, cells were treated the drug corresponding to the resistance cassette of the PCR amplicon of the resistance gene already incorporated in the parental cell line (for concentration see Table 2.1b). Cells were then pipetted into 96 well plates in volumes of 200μl per well. Wells populated with cells in 8-14

days were imaged by wide-field microscopy (see section 2.3, Light microscopy) to visualise fluorescence. Cell cultures successfully expressing fluorescent signal were sub-cultured and 1×10^7 cells were frozen at -80°C for 12 hours then transferred to liquid nitrogen. Freezing solution was made up of 50% glycerol. Cells with frozen in 1mL containing 80 μl cell culture and 20 μl freezing solution.

Table 2.1 (A) Components and concentrations of solutions used buffer in transfection. (B) Drug concentrations for selection following transfection. (C) Conditions used for electroporation with the BTX geminiX2 electroporation generator

A	Transfection buffer component		Volume (μL)
	3 x Roditi buffer containing 600 mM Na_2HPO_4 , 210 mM NaH_2PO_4 , 45 mM KCl, 450 mM HEPES (pH 7.3) and 4.5 mM CaCl_2		175
	CaCl_2 (10mM)		50
	Sterile ddH ₂ O		275

B	Drug	Selection concentration ($\mu\text{g}/\text{mL}$)
	Puromycin	5
	G418 (Geneticin)	40
	Blasticidin	20
	Phleomycin	7.5

C	Conditions for BTX electroporator	
	1800 V	
	3 pulses	
	0.1 ms	
	0.5 s interval	

2.3 Light microscopy

All live cells and cytoskeletons were imaged using a Zeiss Axio Imager.Z2 wide-field microscope (490016-0001-000) with an ORCA-flash 4.0 Hamamatsu camera (Okerkochen, Germany, C13440-20CU) and 1.4NA objective. Images were taken at 100x magnification. Exposure time was 30ms.

2.3.1 Live cell microscopy

1×10^7 cells were harvested from culture in log phase (5×10^6 - 1×10^7) by centrifugation for 4 minutes at 800 x g. The pellet of cells was washed once in vPBS (137 mM NaCl, 3 mM KCl, 16 mM Na_2HPO_4 , 3mM KH_2PO_4 , 46 mM sucrose), resuspended in $1 \mu\text{g}/\text{mL}$ Hoechst for 5 minutes and washed twice more and resuspended in vPBS. $1 \mu\text{L}$ of suspension was settled on an 8-well diagnostic glass slide (Epredia™ - X2XER203B) and mounted with a #0 coverslip (Fisher Scientific – 10011913).

2.3.2 Methanol fixed cytoskeletons

1×10^7 cells were harvested in log phase (5×10^6 - 1×10^7) and centrifuged at 800 x g for 4 minutes. Cells were washed in vPBS twice and resuspended in volume of vPBS to that makes a density of 1×10^7 cells/mL. 2cm by 2cm square wells were drawn onto Thermo™ superfrost microscopy slides (Epredia - ZT1268239S) using a DAKO pen (Agilent - S2002). 40µL of cell suspension was pipetted onto each well and left for 2 minutes until the bound cells have reached a density of 70%. Cell suspension was then removed, replaced with 50µL of vPBS and after 1 minute replaced with 100µL 1% IGEPAL-NP40 (Sigma-Aldrich - I8896) in PEME buffer (100 mM PIPES-NaOH pH 6.9, 1 mM MgSO₄, 2 mM EGTA, 0.1 mM EDTA). After 45 seconds the solution was removed and slides were placed in a coplin jar containing ice-cold methanol and left for 20 minutes at -20°C. Slides were rehydrated in PBS at room temperature for 5 minutes, stained with 1µg/mL Hoescht for two minutes and washed three times in sterile PBS. Slides were mounted with 5µL glycerol phosphate and sealed with nail polish prior to imaging.

2.3.3 Formaldehyde-methanol fixation

Cells were harvested in log phase (5×10^6 - 1×10^7) by centrifugation at 800 x g for 4 minutes. The pelleted cells were washed twice in vPBS. Cells were pipetted onto superfrost™ glass slides and left to settle for 10 minutes before being treated with 4% formaldehyde solution (see materials 2.9.5). The slides were then washed in PBS for 5 minutes and placed in a coplin jar of 100% methanol at -20°C for 5 minutes and were then subsequently rehydrated in PBS for 5 minutes before being processed for immunofluorescence (see methods 2.8).

2.4 Transmission electron microscopy (TEM)

Wild type SmOx p927 strain of *T. brucei* were harvested in log phase and fixed in 2.5% (w/v) glutaraldehyde (Sigma - G5882) and centrifuged for 8 minutes at 800 x g. The cells were then washed 5 times (3 minutes per wash) in ddH₂O before being post-fixed in 1% Osmium tetroxide (Sigma - O5500) for 1 hour at 4°C. The fixate was stained with 1% (w/v) aqueous uranyl acetate (TAAB - U006) for 16 hours at 4°C. The sample was washed three times in ddH₂O before being dehydrated in a series of ethanol concentrations (30%, 50%, 70%, 90% and 3x 100%) for 10 minutes each. Cells were embedded in TAAB 812 resin (TAAB – TO30) (Jiri Tyc, Vaughan lab). This was first as a 3:1 ratio of ethanol to resin (1h) followed by a 1:1 and then changed twice in 100% TAAB812 resin and polymerised for 24 hours at 60°. Blocks were sectioned using a DIATOMEUltra® diamond knife in 70nm sections. Sections were post-stained with 5% uranyl acetate for 20 minutes and Reynolds lead citrate (Reynolds, 1963) (TAAB – L037) for 5 minutes. All

images of 70nm thin sections were acquired using a JOEL 1200Kv Transmission electron microscope (TEM) with a Megaview III camera (12 bits).

2.5 Bioinformatics

Conservation of putative transition fibre proteins were assessed using reciprocal best blast (RBB) in a range of ciliated and non-ciliated eukaryotes using the uniprot proteome for each organism. Top hits for each protein were blasted against the proteome of the SmOx TREU927 strain of *T. brucei* (source Trityp.db). A reciprocal top hit would be considered an ortholog where $e \leq 10^{-5}$. Data for reciprocal best blast of the trypanosome proteome shared by Richard Wheeler, University of Oxford. Confirmation of these orthologues was carried out using the method above.

2.6 Automated image analysis

A FIJI macro was developed with Dr Richard J Wheeler, University of Oxford that measured the length of each putative transition fibre protein from a lateral view and the offset from the centre of a red mScarlet signal (TbSAS-6 or TbCEP164A) (<https://github.com/imagej/ImageJ>, <https://imagej.net/software/fiji/downloads>). The centre of each mNeonGreen signal was selected manually or by using the 'find maxima' function and verified visually that it was selecting the correct region of fluorescent signal. A defined number of angle steps (45) were taken and the distance from the green point selected to the centre point of the red signal perpendicular to the longest dimension of the green signal was measured. This therefore measured the length of the green signal and the offset of the green signal from the marker. The mean values were used and outliers (>500nm for distance from BB and pseudo-diameter) were removed. These values were selected based on TEM analysis of trypanosome basal bodies/transition fibres (Chapter 3, Fig. 3.10).

2.7 RNAi

2.7.1 RNA interference plasmid generation

Primers were designed using RNAit (<https://dag.compbio.dundee.ac.uk/RNAit/>). All PCR components were combined, and PCR was carried out according to conditions described in Table 2.2 in a total volume of 50µL. All 50µL of product was run on a 0.8% agarose gel (80V for 40 mins). The pQuadra (pQ1) and pQuadra3 (pQ3) plasmids (Gifted by Gull lab, University of Oxford) were digested with BstXI (NEB - R0113S) and run on the agarose gel. A 500bp band for the amplicon was excised from the gel along with a 1500bp fragment of pQ3 and a 500bp fragments of pQ1. The DNA was extracted from the gel using the Monarch® DNA gel extraction kit (NEB - T1020S) and quantified using a nanodrop spectrophotometer. The extracted amplicon was digested with BstXI and ligated with the pQ3 backbone and pQ1 plasmid in a 4:2:1 ratio (insert: pQ1: pQ3) at 4°C overnight using a T4 DNA ligase (NEB - M0202S). The ligation product was transformed into SURE2 strain of *E. coli*

(Agilent – 200152) and plasmid was harvested using the Monarch® miniprep plasmid extraction kit (NEB - T1010S). Plasmids were sequenced using the mix2seq® (EUROFINS) sequencing kits to confirm correct ligation.

Table 2.2 PCR components and conditions used for RNAi plasmid generation

Component	Volume (μL)
Primer FW (100mM)	1
Primer RV (100mM)	1
10x reaction buffer with 25mM MgCl ₂	5
dNTPs (10mM)	1
Template genomic DNA (50ng/μL)	1
High fidelity polymerase Enzyme mix –Expand high fidelity PCR system Sigma-aldrich® (5μg/μL)	1
ddH ₂ O	40

Reaction step	Temperature °C	Time (minutes)
Initial denaturation	94	1:00
Denaturation	94	0:30
Annealing	57	0:30
Extension	72	0:45
Final extension	72	5:00
Holding	12	∞

2.7.2 Transfection

2μg of plasmid was linearized by digesting with NotI (NEB – R3189) overnight at 37°C in a volume of 30μL. Prior to transfection the NotI was deactivated by heating the sample to 65°C for 20 minutes. Transfection was carried out as described above (3.2.2) but instead cells were selected using a concentration 7.5μg/mL of phleomycin. Cells were kept in flasks following transfection and recovered after approximately 12 days.

2.7.3 Growth analysis

Cell density was measured using a Beckman Coulter V1 cell counter and cells were diluted to 2×10^5 cells/mL in a total volume of 10mL. This volume was split into two flasks, each containing 5mL of culture to ensure the cell densities were equal. In one flask, cells were induced by the addition of 1μg/mL of doxycycline. Cells were split to 2×10^5 cells/mL at the same time every 24 hours. Growth was measured up to 72 hours in induced and uninduced cells.

2.8 Immunofluorescence

Cells used for immunofluorescence were methanol-fixed cytoskeletons and formaldehyde-methanol fixed cells (see 2.3.2 and 2.3.3) Slides were removed from a coplin jar containing PBS and liquid was blotted using a soft tissue before being placed into a humidity chamber. 100μL of

1% bovine serum albumin (BSA) in PBS was pipetted to each well for 30 minutes to block non-specific protein binding sites. Block solution was removed with a pipette and 100 μ L of primary antibody diluted in PBS containing 1% BSA was added and left in the humidity chamber for 1 hour at room temperature. Slides were placed back into the coplin jar and filled with PBS. After 5 minutes the PBS was emptied and filled with fresh PBS. This process was repeated 5 times. 100 μ L of secondary antibody diluted in PBS containing 1% BSA was added to each well and left in a humidity chamber at room temperature for another hour. Slides were then processed and mounted for imaging according to 2.3.2. Concentrations of primary and secondary antibodies used can be found in Table 2.3.

Table 2.3 List of primary and secondary antibodies for immunofluorescence

Primary antibody	Target	Dilution for immunofluorescence
mAB25	Flagellar axoneme	1:40
YL1/2	Tyrosinated tubulin	1:50
BBA4	Basal body and pro-basal body	1:25
BB2	TY epitope	1:10
TAT-1	Alpha tubulin	1:10
IFT172	IFT172	1:500
Secondary antibody		Dilution for immunofluorescence
Flourescein (FITC) AffiniPure F(ab') ₂ Fragment Donkey Anti-Mouse IgG (H+L)		1:200
Rhodamine (TRITC) AffiniPure Goat Anti-Rat IgG		1:200
Rhodamine (TRITC) Goat Anti-Mouse IgM		1:200

2.9 Solutions

2.9.1 Phosphate buffered saline (PBS):

137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄. For a 10X stock, approximately 100 mL of water was added to a 1 L measuring cylinder. 80 g NaCl, 1 g KCl 14.4 g Na₂HPO₄ and 2.4 g KH₂PO₄ was then weighed and transferred to a measuring cylinder. Water was then added to make volume up to 1 L. 10x stock could be stored at room temperature and diluted as needed.

2.9.2 Voorheis modified PBS (vPBS):

137 mM NaCl, 3 mM KCl, 16 mM Na₂HPO₄, 3mM KH₂PO₄, 46 mM sucrose. For a 1X stock, approximately 100 mL of water was added to a 1 L measuring cylinder. 8 g NaCl, 0.22 g KCl, 2.27 g Na₂HPO₄, 0.41 g KH₂PO₄, and 15.7 g sucrose was weighed and transfer to the measuring cylinder. Water was then added to make up to 1 L and then sterile filtered through a 0.22 µm filter.

2.9.3 PEME buffer:

100mM PIPES-NaOH pH 6.9, 1mMMgSO₄, 2mM EGTA, 0.1mM EDTA. For a 1x stock 50mL water was added to a measuring cylinder. 10mL 1M PIPES-NaOH was added along with 1mL 200mM EGTA and 50µL 200mM EDTA. Water was then added to make a total volume of 100mL.

2.9.4 Hoechst:

5 mg Hoechst (98% bisBenzimide G 33342 trichlorid) was dissolved in 1mL ddH₂O and stored at - 20°C.

2.9.5 4% Formaldehyde solution:

A 10mL ampule containing 16% methanol free formaldehyde (Thermo Scientific™ 28908) was added to 40mL ddH₂O in a 50mL falcon conical centrifuge tube.

2.9.6 0.2M Phosphate buffer solution:

Sodium dihydrogen phosphate was made up by dissolving 69g NaH₂PO₄ in 500mL ddH₂O. Disodium hydrogen phosphate was made by dissolving 8.9 Na₂HPO₄ in 250mL ddH₂O. 23.4mL sodium dihydrogen phosphate was combined with 91.5mL Disodium hydrogen phosphate and 5M NaOH added until pH reached 7.4.

2.9.7 EM primary fixative solution:

Fixative made up of 2% paraformaldehyde, 2.5% Gluteraldehyde (TAAB) and 0.1M phosphate buffer solution. 2mL 10% paraformaldehyde, 1mL 25% Gluteraldehyde and 5mL (0.2M) phosphate buffer were combined and made up to 10mL with ddH₂O.

2.9.8 2% uranyl acetate

0.4 g uranyl acetate (Agar Scientific) was dissolved in 20 mL ddH₂O by sonicating for 30mins. The solution was kept in the dark at room temperature and aliquots were taken from this stock solution.

2.9.9 4% osmium tetroxide stock solution

5mL ampule containing 4% osmium tetroxide (TAAB) was added to a glass vile and kept in the dark. Aliquots were taken from this solution.

3. Transition fibre composition, structure and spatial arrangement

3.1 Identification of 35 putative transition fibre proteins

Tryptag is a genome-wide protein localisation resource that has labelled every protein encoded in the *Trypanosoma brucei* genome (excluding genes encoding variable surface glycoprotein) (Dean et al., 2017). Genes were tagged at the endogenous locus with the mNeonGreen (mNG) fluorescent protein using the long primer PCR-only tagging (POT) method (see methods 2.2) (Dean et al., 2015, Dean et al., 2017). *T. brucei* has a well-defined cellular architecture, with a clear anterior-posterior polarity. Most major organelles, such as the flagellum, have a consistent location within the cell and therefore protein localisation can be annotated with relative ease (Dean et al., 2017, Hammarton, 2007).

In *T. brucei*, the flagellum extends along the length of the cell body and is nucleated by a microtubule-organising basal body located at its proximal end. In collaboration with Shahaan Shafiq and Jiri Tyc (Vaughan lab), a screen of proteins localised to the basal body was carried out. Proteins localising to the basal body region were identified by searching the term 'basal body' into the Tryptag database (www.tryptag.org). A total of 312 proteins were identified as proteins localising to the basal body region and were taken forward for further analysis. Figure 3.1 shows

examples of basal body proteins as identified by TrypTag (Shahaan Shafiq, Jirka Tyc, Vaughan lab).

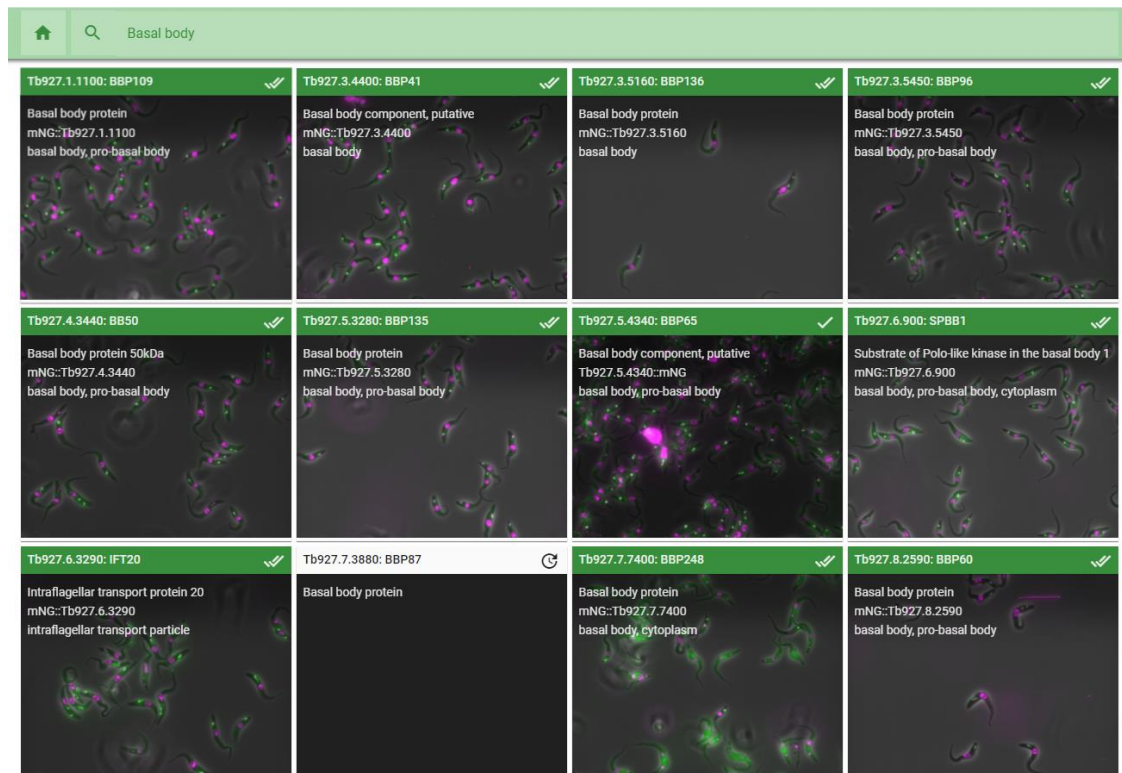


Figure 3.1 Examples of cell lines expressing proteins tagged at the N-terminus with the mNeonGreen (mNG) fluorescent protein that were annotated as basal body in TrypTag.org. Live cell images were taken at 63x magnification and the DNA in the nucleus and kinetoplast was labelled with Hoescht (magenta). www.tryptag.org

3.2 Co-expression of putative basal body proteins with basal body maker SAS-6

To further investigate the localisation of the cohort of basal body proteins, we aimed to co-express each of the 312 proteins with a basal body marker to assess the localisation of each protein in more detail. SAS6 is a highly conserved basal body protein that establishes the 9-fold symmetry at the base of the centriole known as the cartwheel hub (Nakazawa et al., 2007). In trypanosomes, SAS6 also makes up the cartwheel and is present on the mature and probasal body throughout the cell cycle (Hu et al., 2015, Dang et al., 2017). This allows both the mature and probasal body to be visualised throughout the cell cycle and therefore qualifies as a suitable marker protein for the annotation of basal body proteins.

SAS6 was tagged at the endogenous locus using the PCR-only tagging method (see methods 2.2) (Dean et al., 2015). Briefly, the method works by coupling long primer PCR and the PCR-only tagging (pPOT) plasmid to produce a gene tagging amplicon with a homology region long enough to be integrated at the endogenous locus of a gene of interest via homologous recombination. The pPOTv7 plasmid selected contains a sequence encoding the mScarlet fluorescent protein and

a G418 resistance cassette (Fig. 3.2). The amplicon was generated by PCR (see methods 2.2.1) and transfected (see methods 2.2.2) into the SmOxp927 strain of *T. brucei* (cell line made by Shahaan Shafiq and Jirka Tyc, Vaughan lab).

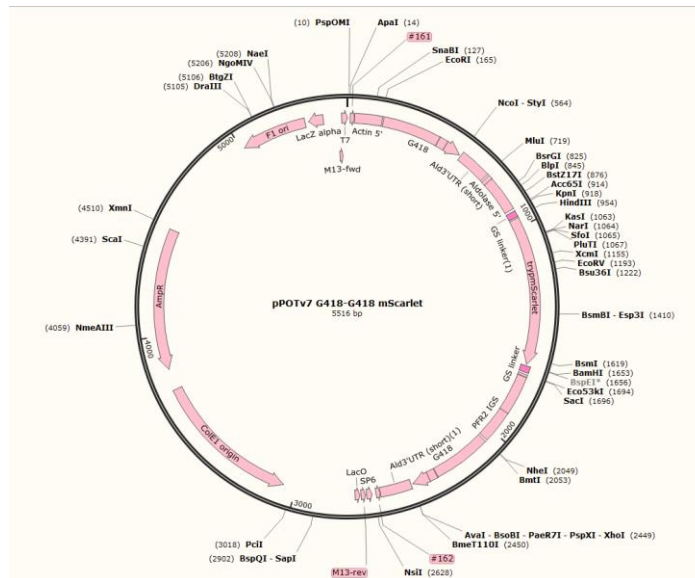


Figure.3.2 Plasmid map of the pPOTv7 plasmid containing a sequence encoding the mScarlet fluorescent protein, the G418 resistance cassette, GS linkers, an ampicillin resistance gene and a range of restriction enzyme sites so that regions can be easily replaced using cloning methods (Dean et al., 2015).

Putative basal body proteins were tagged using the pPOTv6 plasmid which instead contains a sequence encoding the mNeonGreen fluorescent protein flanked by two 3xTY epitope tags and a blasticidin resistance cassette (Fig. 3.3). Candidate proteins were transfected into a cell line expressing the SAS6::mScarlet fusion protein (Cell lines also made by Shahaan Shafiq and Jiri Tyc, Vaughan lab).

Proteins were categorised based on their localisation at the basal body as part of a separate project (Shahaan Shafiq, Vaughan lab). The focus of this project is the transition fibres that surround the distal end of the mature basal body (chapter 1, Fig. 1.1-1.3). Therefore, proteins that localised to the distal region of the mature basal body were taken forward for analysis.

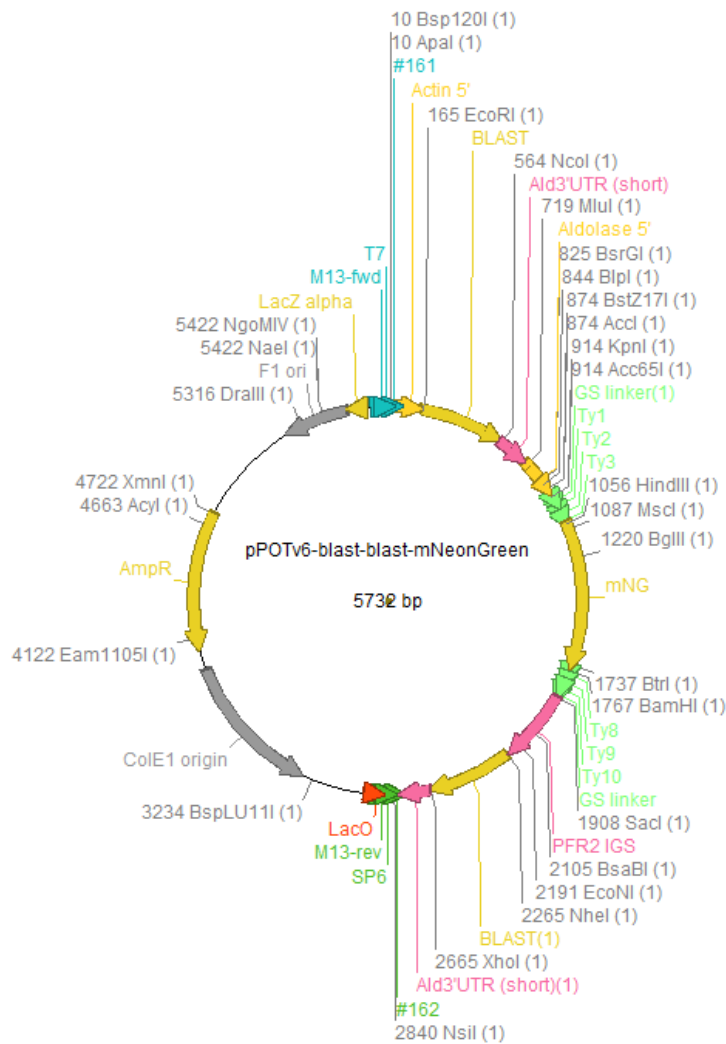


Figure 3.3 Plasmid map of the pPOTv6 plasmid containing a sequence encoding the mNeonGreen fluorescent protein flanked by two 3x TY epitope tags, a blasticidin resistance cassette, GS linkers, an ampicillin resistance gene and a range of restriction enzyme sites so that sites can be easily replaced using cloning methods (Dean et al., 2015).

3.3 Localisation of 35 putative transition fibre proteins

A total of 35 proteins from the 309 basal body proteins in section 3.2 were found to localise to the distal end of the mature basal body and were taken forward for analysis. All 35 proteins exhibited a cloud shaped signal that is distal to SAS6 and found specifically on the mature basal body (MBB) consistent with the localisation expected for a transition fibre protein. Candidates were considered putative transition fibre proteins based on fluorescent signals that are specific to the mature basal body, distal to the mature basal body and wider than the mature basal body - as transition fibres are structures that surround the basal body and are not part of the barrel itself. CEP164 is a highly conserved transition fibre protein found in a wide range of ciliated organisms that has three orthologues in *T. brucei* that have been named CEP164A, CEP164B and CEP164C (Hodges et al., 2010, Atkins et al., 2021). The localisation of CEP164A, CEP164B and CEP164C is

consistent with other putative transition fibre proteins and the localisation and function of CEP164C has been well examined in trypanosomes (Atkins et al., 2021). This gave us confidence that the design of our screen could correctly identify transition fibre proteins.

For each putative transition fibre protein, an mNeonGreen signal of the candidate protein is shown co-expressed with the basal body marker SAS6 tagged with the mScarlet fluorescent protein. A G1 and Post-mitotic (Post M) cell (that has undergone nuclear division) is shown for each protein (Fig. 3.4, Fig. 3.5, Fig. 3.6, Fig. 3.7). Known proteins were given their original name or orthologue names and hypothetical proteins were named TFP (transition fibre protein) plus their predicted molecular weight.

Conservation of putative transition fibre proteins was assessed by reciprocal best blast (RBB) in 45 different ciliated and nonciliated eukaryotes (chapter 4). Briefly, this involved searching the amino acid sequence into the proteome of each organism using the BLAST tool and a reciprocal top hit is considered an orthologue. A more in depth evolutionary analysis will be discussed in chapter 4. Putative transition fibre proteins were grouped based on their conservation as either conserved transition fibre (conserved outside the excavata supergroup and known transition fibre protein), conserved unknown (conserved outside excavata supergroup and localisation not transition fibre or unknown), excavata specific (orthologues only present within excavata) and trypanosome specific (protein only present in trypanosomes). Five conserved known transition fibre proteins FBF1, CEP164A, CEP164B, CEP164C (Tanos et al., 2013) and OFD1 (Kumar et al., 2021) are shown in Figure 3.4.

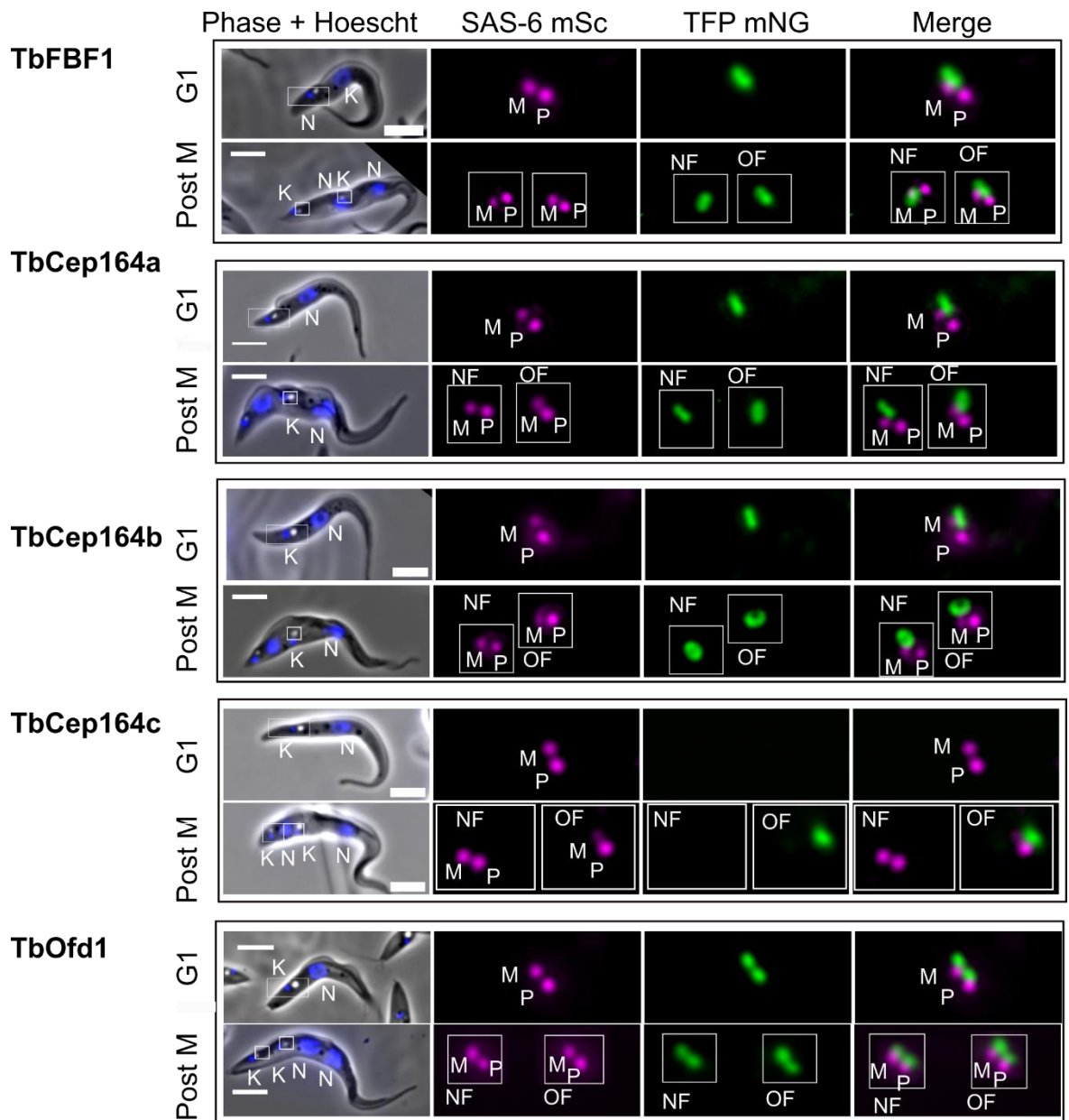


Figure 3.4: Live cell images of known transition fibre proteins (green) conserved in organisms outside excavata tagged at the N-terminus with the mNeonGreen fluorescent protein co-expressed with the basal body protein SAS6 (magenta) tagged with mScarlet. Mature basal body and probasal body are labelled M and P respectively. DNA in the nucleus (N) and kinetoplast (K) was labelled with Hoescht (blue). All images were taken at 100x magnification.

All five proteins in Figure 3.4 exhibit the typical transition fibre localisation of a cloud shaped signal across the distal end of the basal body. All proteins are shown from a lateral view of the basal body and transition fibre. OFD1 appears closest to the mature basal body, followed by FBF1. CEP164A, CEP164B and CEP164C appear to be further from the mature basal body (MBB) in G1 and post-mitotic cells. CEP164C is absent in 60% of G1 cells and the new flagellum of post-mitotic cells. FBF1 appears to have a narrower diameter, CEP164A, CEP164B and CEP164C appear wider than FBF1 but similar to each other and OFD1 appears wider than the other proteins. All proteins

have a signal present in both G1 and post-mitotic cells and present in equal intensity on the old and new flagellum. In mammals, FBF1 is narrower than CEP164 (Yang et al., 2018). The position of OFD1 at the transition fibres has not yet been examined. Table 3.1 summarises the reported localisation of conserved known transition fibre proteins in other organisms.

Table 3.1 Conserved known transition fibre proteins in other organisms. Trypanosome putative transition fibre protein (left). Orthologue and organism in which it has been localised (middle). Reported localisation of the orthologue in other organism (right).

Trypanosome putative transition fibre protein	Orthologue/organism	Localisation
CEP164A, CEP164B, CEP164C	CEP164, Human	Distal appendage/Transition fibre (Graser et al. 2007)
FBF1	FBF1, Human DYF-19, <i>C elegans</i>	Distal appendage/Transition Fibres (Tanos et al. 2013) Distal appendage/Transition Fibres (Wei et al. 2013)
OFD1	OFD1, Human	Distal appendage/Transition fibre (Singla et al. 2010)

Putative transition fibre proteins that are conserved outside excavata, but their orthologues are not known transition fibre proteins are shown in Figure 3.5. At the top of Figure 3.5, TFP123 shows a narrowest mNG signal relative to the SAS6 foci but appears to be further away from the mature basal body than other proteins. The signal appears present in equal intensity on the new and old flagellum in post-mitotic cells. This is also true for all other putative transition fibre proteins in Figure 3.5, with the exception of Rabl2b, which appears to have a brighter signal on the old flagellum of post-mitotic cells (further analysis shown in chapter 4).

CEP19L exhibits a similar localisation to TFP123 but appears to have a slightly wider signal than TFP123. TFP101 exhibits a localisation very similar to that of CEP19L and TFP123. In mammals, CEP19 is known to localise to the distal end of the mature basal body and in mammalian cells is known to form a complex with Rabl2b and facilitates intraflagellar transport (IFT) recruitment (Nishijima et al., 2017). However, studies have not yet localised these two proteins to the transition fibres specifically (Table 2). Tb927.3.4470 has a signal wider than the mature basal body and shows a very weak signal around the probasal body. However, the signal of Tb927.3.4470 is mostly overlapping with the mature basal body and is not clearly distal to the basal body. CEP90, SAS6L and CFAP52 clearly exhibit a wider signal across the mature basal body, however in both cases appear to be closer to the mature basal body than other proteins with the exception of Tb927.3.4470. Table 3.2 summarises the reported localisation of conserved putative transition

fibre proteins in other organisms.

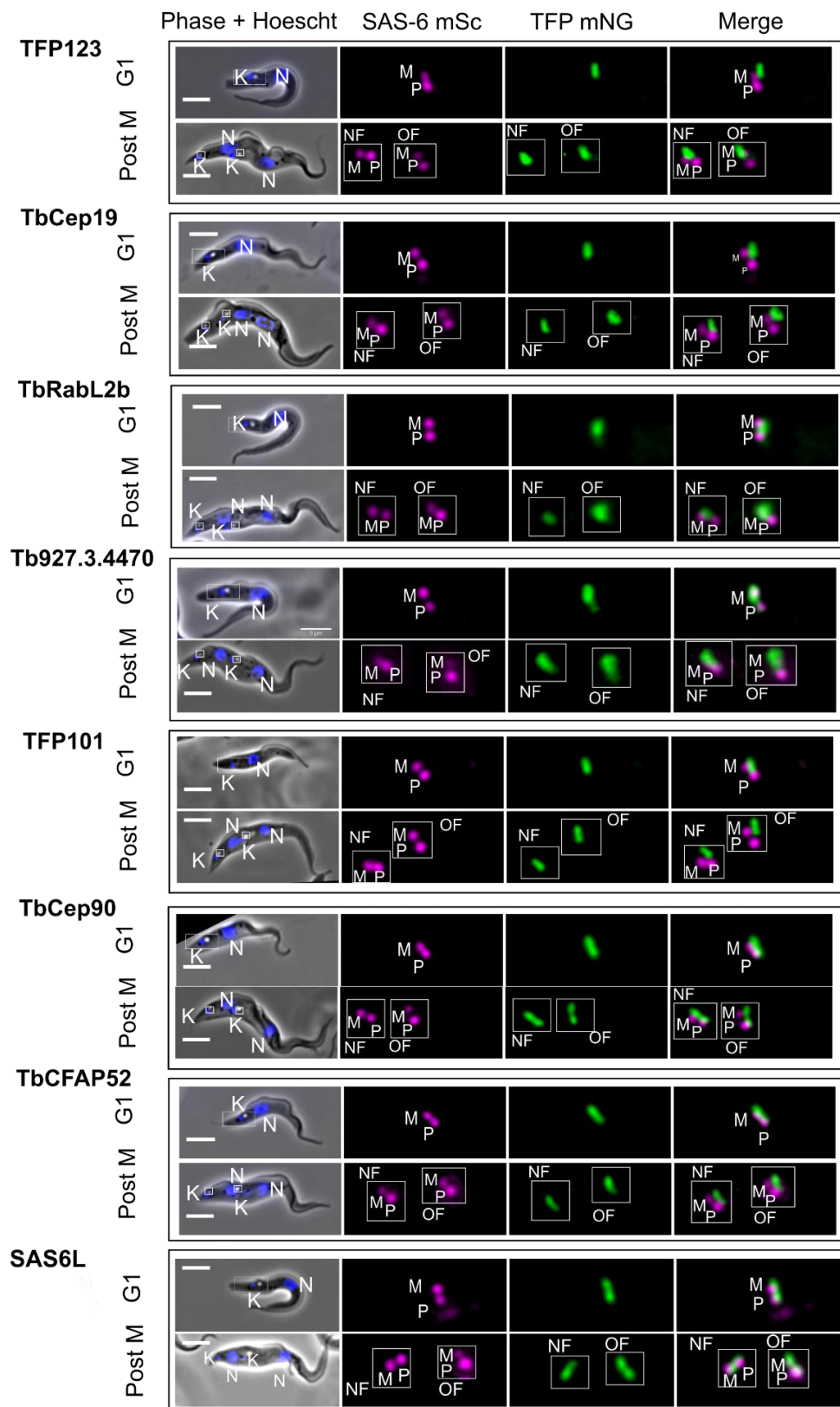


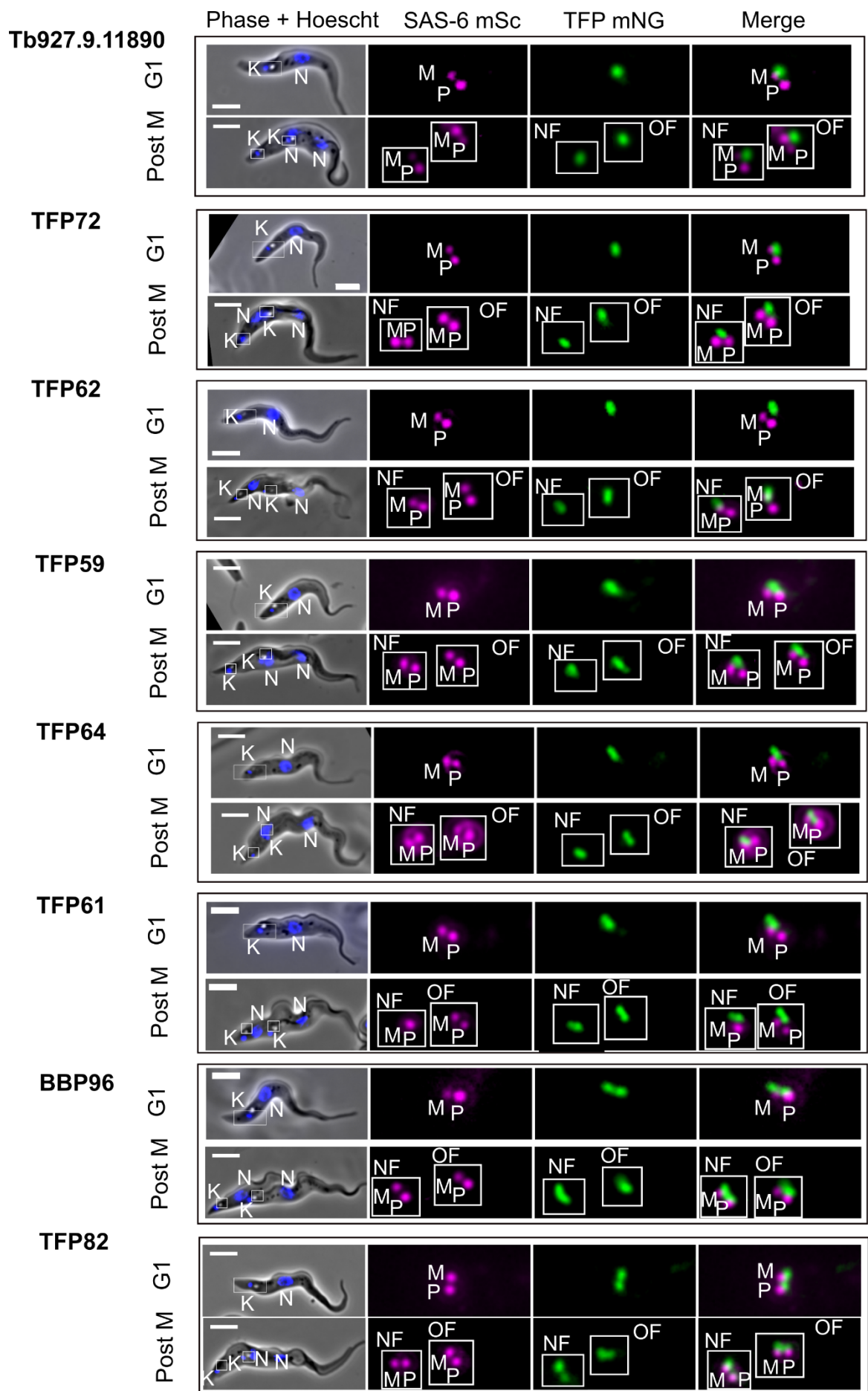
Figure 3.5 Live cell images of putative transition fibre proteins (green) conserved in organisms outside excavata tagged at the N-terminus with the mNeonGreen fluorescent protein co-expressed with the basal body protein SAS-6 (magenta)

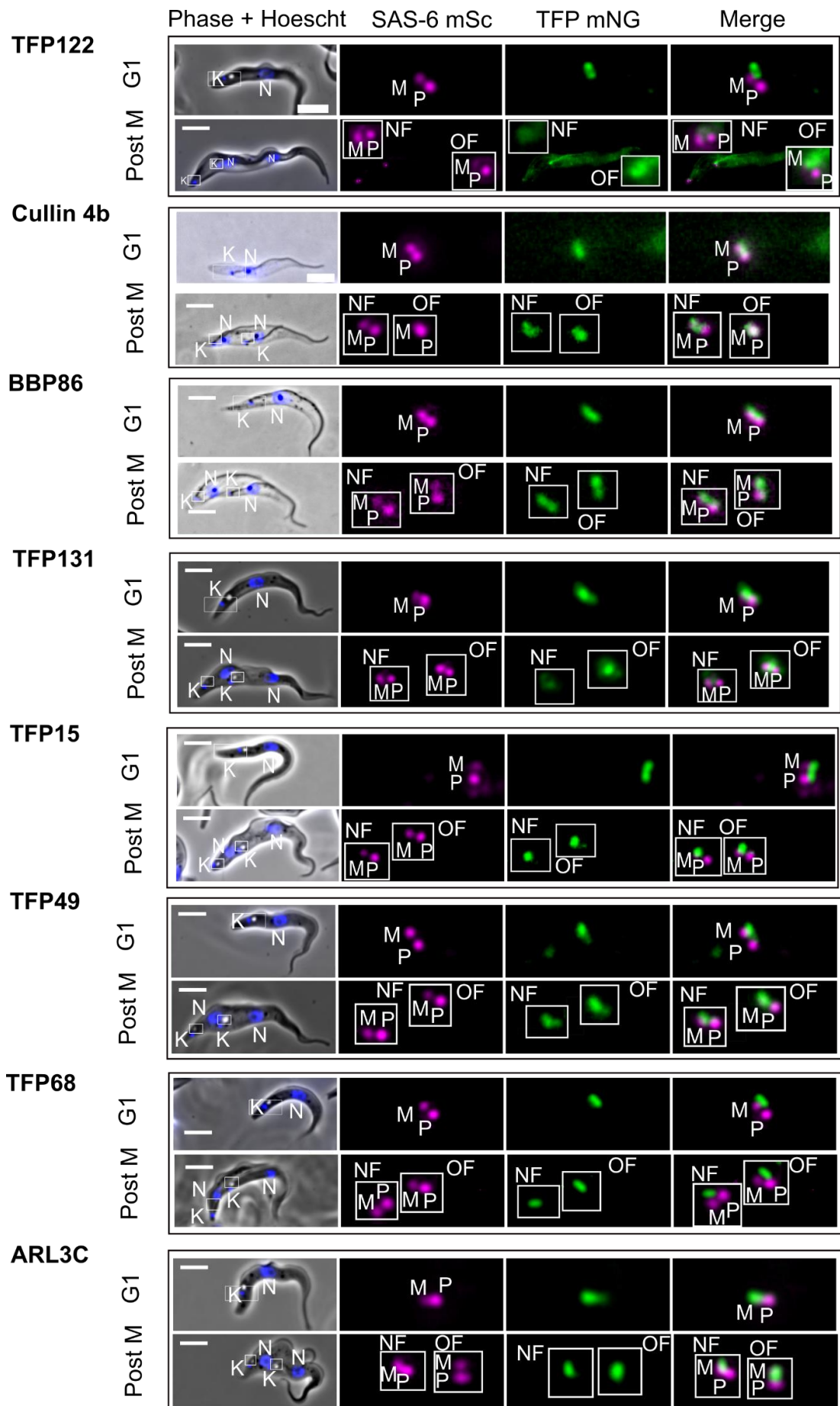
tagged with mScarlet. Mature basal body and probasal body are labelled M and P respectively. DNA in the nucleus (N) and kinetoplast (K) was labelled with Hoescht (blue). All images were taken at 100x magnification.

Table 3.2 Conserved putative transition fibre protein in *Trypanosoma brucei* (left). Orthologue name and organism in which it has been localised (middle). Reported localisation in other organism(s) (right).

Trypanosome putative transition fibre protein	Orthologue/organism(s)	Localisation
CEP19L	CEP19, Human	Mature basal body, RABL2 interactor (Nishijima et al., 2017, Majorad et al., 2017)
CEP90	CEP90, Human	Centriolar satellite (Kim et al. 2010) Mature basal body (Kumar et al., 2021)
RABL2B	RABL2B, Human	Mature basal body, CEP19 interactor (Nishijima et al., 2017)
CFAP52/CMF15	CFAP52/WDR16, Mice	Sperm manchette mature and probasal body (Contreras et al., 2020)
SAS6L	SAS6L, <i>Plasmodium berghei</i> , <i>Toxoplasma gondii</i> , <i>Trypanosoma brucei</i>	<i>Plasmodium berghei</i> - cytosol in female gametocyte and foci in ookinete (motile zygote) apical end (Wall et al., 2016) <i>Toxoplasma gondii</i> – Conoid region (Cruz de Leon et al., 2013) <i>Trypanosoma brucei</i> – Basal plate region (Cruz de Leon et al. 2013)

Figure 3.6 shows the localisation of 17 putative transition fibre proteins that are not conserved outside of the excavata supergroup. Tb927.9.11890 exhibits the narrowest localisation of all the proteins in Figure 3.6. The width/diameter does not appear much wider than the basal body, although it is located distal to the mature basal body. TFP15 shows the widest localisation, however the wide localisation in TFP15 is only apparent in G1 cells and in post-mitotic cells the signal appears narrower. TFP122 localises to the flagellum as well as the transition fibres and exhibits a brighter signal at the old flagellum when compared to the new flagellum in post-mitotic cells. TFP49, ARL3C, TFP68, TFP72, TFP62, TFP59, TFP64, TFP61, BBP96, TFP82, Cullin4b and BBP86 exhibit the typical transition fibre localisation of a cloud-shaped signal across the distal end of the mature basal body. SPB1 exhibits a transition fibre localisation and also localises to the mitotic spindle and spindle pole during mitosis (Fig. 3.6, bottom). The nuclear localisation of SPB1 has been described in previous studies but localisation at the transition fibre/basal body region was not reported (Table 3) (Zhou et al., 2018). Table 3.3 summarises the reported localisations of excavata specific putative transition fibre proteins reported in other studies.





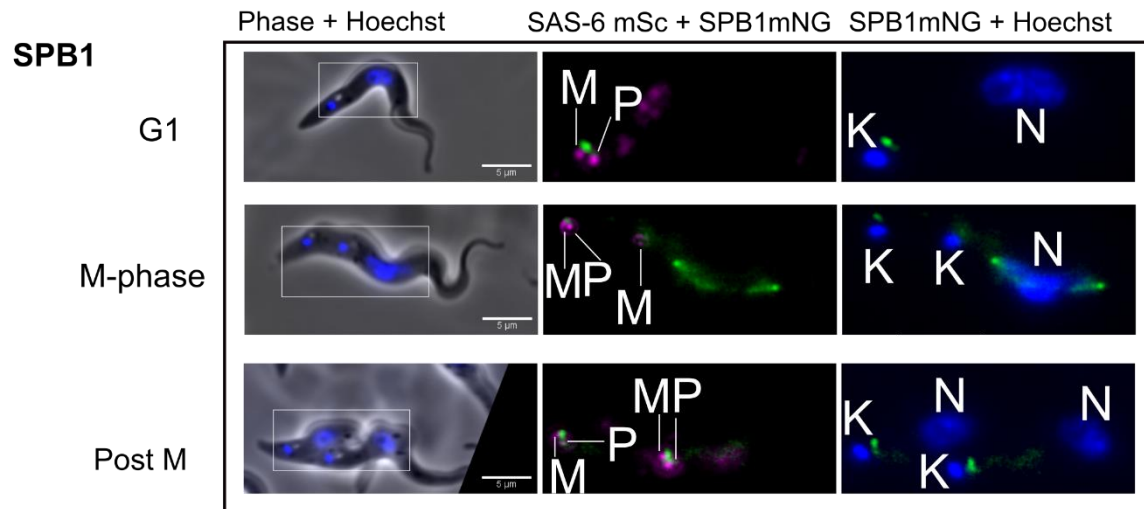


Figure 3.6: Live cell images of putative transition fibre proteins (green) found only in excavata that were tagged at the N-terminus with the mNeonGreen fluorescent protein co-expressed with the basal body protein SAS-6 (magenta) tagged with mScarlet. **Methanol fixed cytoskeletons were used for Cullin4b and BBP86.** Mature basal body and probasal body are labelled M and P respectively. DNA in the nucleus (N) and kinetoplast (K) was labelled with Hoescht (blue). Images were taken of methanol-fixed cytoskeletons for BBP86 and Cullin4b for better clarity. All images were taken at 100x magnification.

Table 3.3 Kinetoplastid specific putative transition fibre protein in *Trypanosoma brucei* (left). Orthologue name and organism in which it has been localised (middle). Reported localisation in other studies (right).

Trypanosome putative transition fibre protein	Orthologue/organism	Localisation
CEP164A, CEP164B, CEP164C	CEP164, Human	Distal appendage/Transition fibre (Graser et al. 2007)
FBF1	FBF1, Human DYF-19, <i>C. elegans</i>	Distal appendage/Transition Fibres (Tanos et al. 2013) Distal appendage/Transition Fibres (Wei et al. 2013)
OFD1	OFD1, Human	Distal appendage/Transition fibre (Singla et al. 2010)

The localisation of five trypanosome specific putative transition fibre proteins are shown in Figure 3.7. TFP52, TFP17, TFP39 and TFP104 all exhibit a typical transition fibre localisation of a cloud shaped signal across the distal end of the mature basal body, with TFP52 showing the narrowest mNG signal. All proteins except for TFP32 have no consistent differences in the intensity of signal at the old and new flagellum. TFP32 exhibits a brighter signal at the old flagellum in post-mitotic cells as well as a secondary localisation at the flagella pocket region.

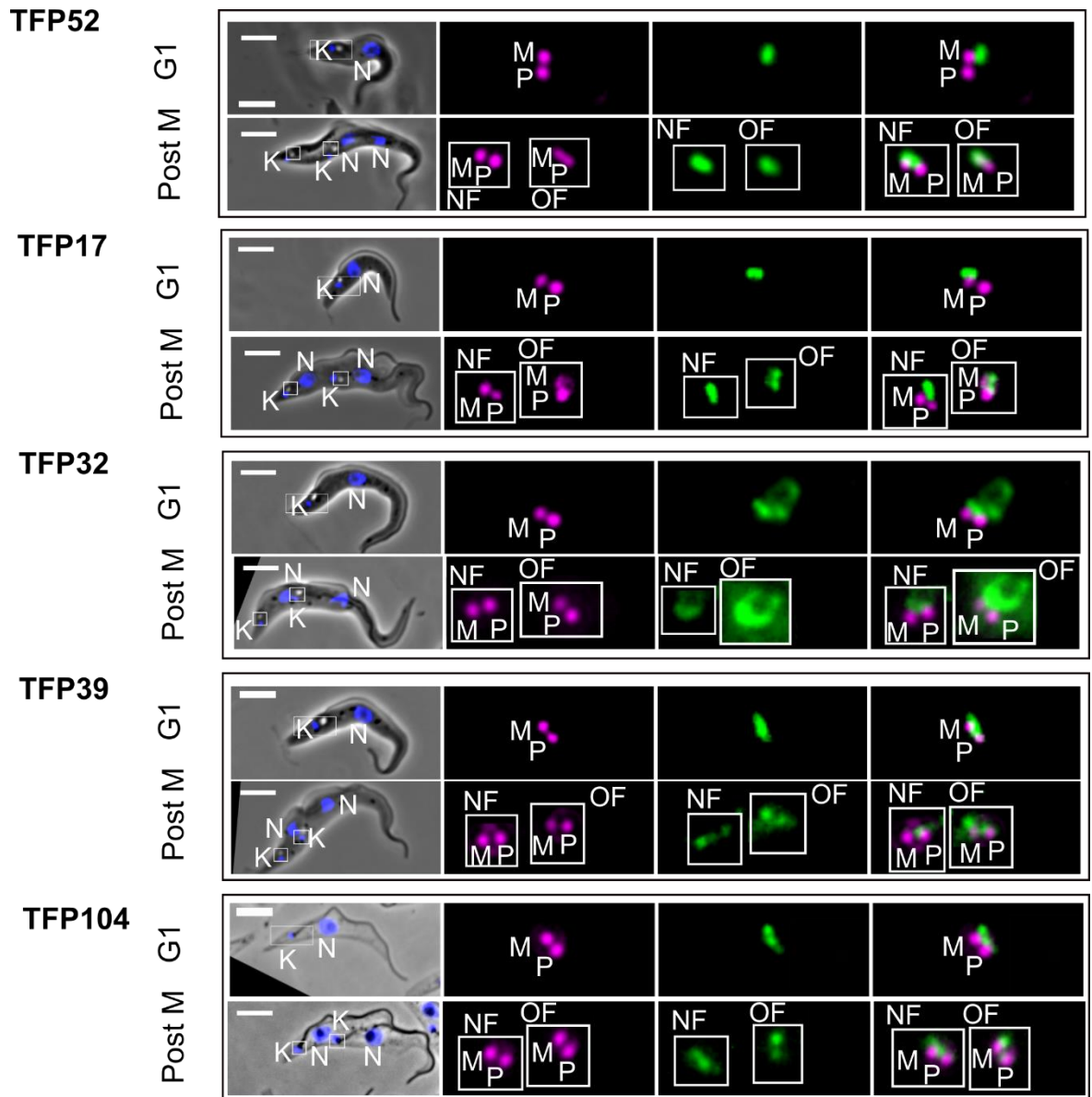


Figure 3.7: Live cell images of putative transition fibre proteins (green) found only in Trypanosomes that were tagged at the N-terminus with the mNeonGreen fluorescent protein co-expressed with the basal body protein SAS-6 (magenta) tagged with mScarlet. **Methanol-fixed cytoskeletons were used for TFP104.** Mature basal body and probasal body are labelled M and P respectively. DNA in the nucleus (N) and kinetoplast (K) was labelled with Hoescht (blue). Images were taken of methanol-fixed cytoskeletons for TFP104 for better clarity. All images were taken at 100x magnification.

In summary, we have identified 35 putative transition fibre proteins based on the criteria that they are specific to the mature basal body, distal to the mature basal body and wider than the mature basal body. We conclude that it is difficult to classify and exclude candidates based on visual inspection alone. Therefore, quantitative assessment of transition fibre localisation will be required to a more accurate set of inclusion criteria. Proteins that exhibit a cell cycle regulated localisation, with a brighter signal on the new or old flagellum will be discussed in more detail in chapter 4.

3.4 Quantitative analysis putative transition fibre protein localisation

Candidates were considered putative transition fibre proteins based on an mNG signal that is at the distal end of the mature basal body and wider than the mature basal body. Again, this is because transition fibres are structures that surround the basal body and therefore when viewed laterally the signal should be wider than the mature basal body. We aimed to quantitatively assess the width of those candidate proteins, as well as their distance from the basal body to decide which proteins should be considered as transition fibre proteins and to gain an insight into the spatial arrangement of proteins along the transition fibre.

Previous studies in mammalian cells have determined the spatial organisation of proteins along the transition fibre blade based on their diameter using super-resolution microscopy such as dSTORM (Direct stochastic optical reconstruction microscopy) (Yang et al., 2018). However, this analysis was only carried out on five transition fibres proteins. To determine the location of each protein in a larger cohort, we created an automated FIJI script that was designed in collaboration with Dr. Richard Wheeler (University of Oxford) to measure the length of the transition fibre signal from a lateral view and the distance from the mature basal body (Fig. 3.8, Fig 2.8 methods) (Richard Wheeler, University of Oxford). We reason that the length of the transition fibre signal from a lateral view would be equivalent to the diameter from a cross-sectional view and so termed this the 'pseudo-diameter' (Fig. 3.8). The FIJI script determines the pseudo-diameter of the transition fibre signal by measuring across the mNG signal in a series of angle steps until the length of the widest dimension is determined. The distance to the basal body signal perpendicular to the widest dimension could be determined, which would be the mature basal body due the orientation of the transition fibre signal and mature basal body (Fig. 3.8). The script therefore can

measure the pseudo-diameter and the distance of the signal from the mature basal body. A range between 500 and 3000 individual signals was measured for each transition fibre protein.

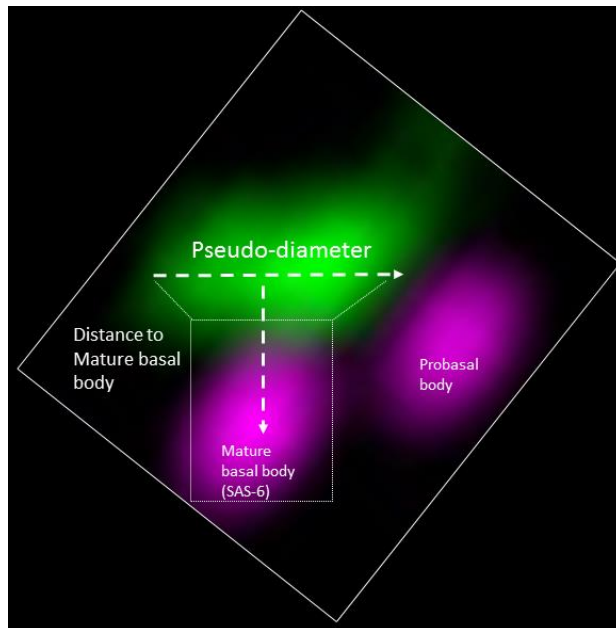


Figure 3.8 Measurement of pseudo-diameter and distance from the mature basal body (SAS-6) of transition fibre proteins by automated analysis using a Fiji script. Mature basal body, probasal body and transition fibre shown from a lateral view by light microscopy.

The distance from the centre of the mature basal body to the transition fibre signal was plotted in 29 proteins that were stably associated with the cytoskeleton (Fig. 3.9a). Methanol-fixed cytoskeletons were used for all measurements as extracted cytoskeletons tend to fix onto the slide in flatter, more lateral orientation and therefore automated measurements of the lateral dimensions is more consistent in fixed cytoskeletons than in live cells. ARL3C and TFP122 do not stably associate with the cytoskeleton and therefore analysis was not carried out on these proteins. Methanol-fixed cytoskeletons were not obtained successfully for SBP1, CFAP52 and Rabl2b and they were also removed from this analysis. CEP164C was also not included in this analysis as it has already been confirmed as a bona fide transition fibre protein (Atkins et al., 2021). Hence, we analysed the location of 29 proteins.

The well characterised distal basal body barrel protein POC5 was used as a control for the distance from the mature basal body (Fig. 3.9a), as it is localised at the distal end of the basal body barrel (Azimzadeh et al., 2009) (Shahaan Shafiq, personal communication, Vaughan lab). Transition fibre proteins should be distal to the basal body and therefore a candidate protein with a distance that was proximal POC5 would not fit the criteria as a transition fibre protein. Tb927.3.4470 (Fig. 3.9a, 3.4470, blue rectangle) had a distance from the basal body that was more proximal than POC5 and was removed. TFP15 was more proximal than POC5, however TFP15 had a greater standard error than other proteins in terms of its distance from the basal body. TFP15

also exhibited the widest pseudo-diameter of all proteins in Figure 3.9. Based on this, we decided to keep TFP15 in our cohort as we know that the widest transition fibre protein in mammals, CEP164, is critical for basal body to membrane docking and cilium assembly (Tanos et al., 2013, Graser et al., 2007). We will confirm the localisation of TFP15 at the transition fibre based on co-localisation with a known transition fibre marker in a later section (Fig. 3.15).

Two characterised transition zone proteins (TZP150, TZP100) (Dean et al., 2016) were used as controls for the region distal to the expected location of transition fibres, where the transition zone would begin (Fig. 3.9a, 3.9c, yellow rectangle). Based on these controls, Tb927.9.11890 (formerly TFP100) was removed as it was located in the transition zone region (Fig. 3.9, yellow rectangle). The remaining 25 transition fibre proteins were between 150-225nm from the centre of the SAS6 signal (Figure 3.9a; red rectangle). Known transition fibre proteins FBF1, CEP164A, CEP164B and OFD1 all sit within the expected region that is between the distal end of the basal body and the transition zone.

We next wanted to confirm the position of each protein around the basal body by measuring the width of the transition fibre protein signal from a lateral view (pseudo-diameter). The pseudo-diameter of the mature basal body protein BBP136 was measured as a control for basal body diameter as it has a single small circular foci only on the mature basal body and is unlikely to be a transition fibre protein but instead a component of the mature basal body itself (Dang et al., 2017). BBP136 had a mean pseudo-diameter of 149nm, which is narrower than other proteins in this analysis. This was expected for a protein that is part of the basal body.

Proteins with a pseudo-diameter greater than the axonemal diameter of 200 nm were investigated further. This is because the diameter of the basal body trypanosomes is approximately 200 nm (Vaughan and Gull, 2015). Anything wider than 200 nm is likely to be an appendage protein surrounding the basal body and not part of the mature basal body itself. All but one protein exhibited a pseudo-diameter greater than that of the axoneme and BBP136. This includes known transition fibre proteins CEP164A, CEP164B, FBF1 and OFD1 (Fig. 3.9a, green rectangle). Tb927.9.11890 (formerly TFP100) had a diameter less than the 200 nm diameter of the axoneme and therefore was excluded (Fig. 3.9b, orange rectangle). This protein exhibited a narrow pseudo-diameter as well as a distance from SAS6 that was in the transition zone control area, so together this illustrates the accuracy of our approach (Fig. 3.9a, b). We infer from this analysis that Tb927.9.11890 is part of the transition zone.

Our results show a range of pseudo-diameter measurements, consistent with there being transition fibre proteins located along the length of each transition fibre blade. Transition fibre

proteins were grouped as inner TF or outer TF based on their pseudo-diameter. Cullin4b, TFP39, TFP104 and Tb927.6.2220 were all classed as outer TF proteins as they exhibited the widest pseudo-diameters measuring between 300 and 345nm (Fig. 3.9b; purple rectangle). The remaining 24 proteins had a pseudo-diameter measuring between 225 and 275 nm and were therefore classed as inner TF proteins (Fig. 3.9b, green rectangle).

In summary, 29 putative transition fibre proteins associated with the cytoskeleton were mapped using an automated FIJI script. The script measured pseudo-diameter and distance from the mature basal body (Fig. 3.8, Fig. 3.9c). Two putative transition fibre proteins (Tb927.3.4470 and Tb927.9.11890) were excluded from the cohort as each of their localisation on the map in Figure 3.9 was either (i) not between the transition zone and the distal end of the basal body (Fig. 9c) or (ii) had a pseudo-diameter that was narrower than the 200 nm of the basal body (Fig. 3.9c).

Proteins were categorised as inner (Fig. 3.9c; green box) or outer (Fig. 3.9c; green box) based on their pseudo-diameter. Of the 34 putative transition fibre proteins identified (excluding CEP164C), 32 candidates remain following this analysis.

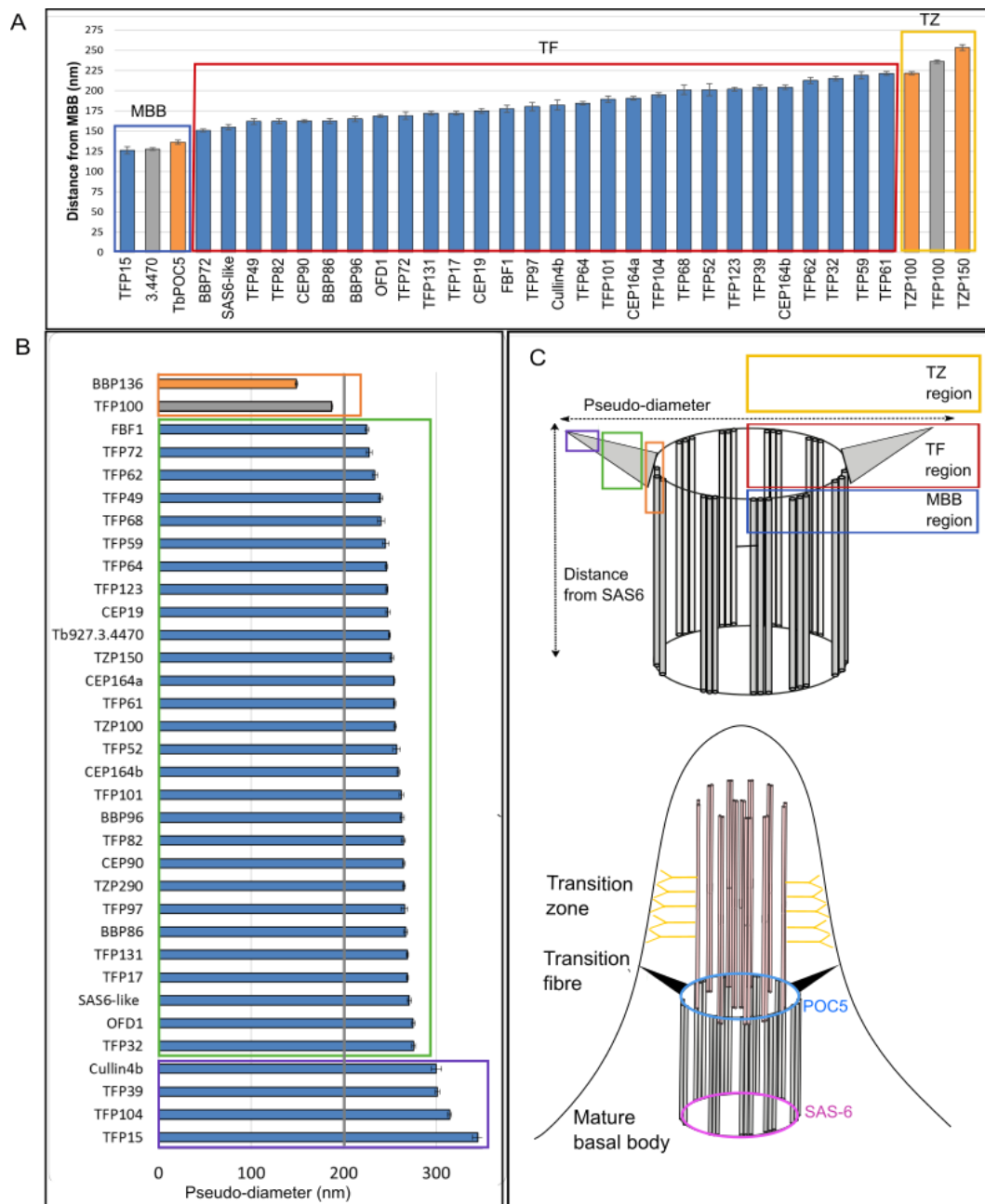


Figure 3.9 Spatial organisation of 29 putative transition fibre proteins based on pseudo-diameter and distance from the mature basal body (SAS-6) measured using an automated FIJI script. (a) Range of measurements ordered by their distance from SAS-6 (nm). Distal end basal body control POC5 is shown in orange (blue rectangle), transition zone controls shown also in orange (yellow rectangle) and excluded proteins are shown in grey. (b) The range measurements ordered by pseudo-diameter (nm). Basal body control (BBP136) shown in orange (orange rectangle), theoretical axonemal diameter indicated by the grey line at 200nm. Excluded proteins are also shown in grey. (c) Model representing the location of each transition fibre protein group based on diameter and distance from basal body (SAS6).

Only orthologues of CEP164A, CEP164B, OFD1 and FBF1 have been localised to the transition fibre structure in other organisms. FBF1, CEP164 and OFD1 are all bona fide transition fibre proteins in

mammals (Kumar et al., 2021, Singla et al., 2010, Tanos et al., 2013, Yang et al., 2018). As mentioned previously, studies have shown that FBF1 is narrower than CEP164 (Yang et al., 2018). Our automated analysis shows that FBF1 has a narrower pseudo-diameter than CEP164A and CEP164B in trypanosomes.

CEP19 has been localised to the mature basal body in previous studies (Nishijima et al., 2017). Our analysis indicates CEP19L is a component of the transition fibres in trypanosomes. CEP90 is a conserved centriolar protein that has been localised to the centriolar satellites in previous studies in mammals (Kim et al., 2019). However, our analysis indicates CEP90 localises to the transition fibres around the mature basal body based on its pseudo-diameter and distance from SAS6. Conserved proteins Rabl2b and CFAP52 were not analysed using this method and therefore their localisation will be assessed primarily based on co-localisation with a known transition fibre protein. Rabl2b, like CEP19, has been localised to the mature basal body in mammals and is known to form a complex with CEP19 (Nishijima et al., 2017). CFAP52 has been shown to localise to the mature and probasal body in mammalian sperm (Tapia Contreras and Hoyer-Fender, 2020) (see Table 2).

3.5 Ultrastructural analysis of transition fibre structure

Although the analysis carried out in Figure 3.9 has provided a useful insight into spatial organisation of transition fibre proteins, we understand that the theoretical limit of light microscopy is an obstacle when assessing the arrangement of proteins at the nanometre scale. To gain a more detailed understanding of transition fibre architecture, structure and size, we analysed the ultrastructure of transition fibres in *T. brucei* using thin section Transmission Electron Microscopy (TEM). TEM analysis would provide a far greater resolution and allow us to compare the measurements made from the automated analysis with the ultrastructure. 70nm thin sections of both longitudinal and cross-sections of mature basal bodies were captured for analysis.

In longitudinal sections, transition fibres are seen infrequently as faint lines (TF) (Fig. 3.10a, 3.10d, 3.10e). Densities (D) are visible where the proximal region of the flagella pocket membrane would meet the transition fibre (Fig. 3.10a –f). The densities (D) indicated in Figure 3.10 represents the end of the transition fibre blade, which can also be observed from a cross-section (Fig. 3.10g, inset). Studies have shown that mammalian transition fibres possess a ‘head density’ at the outer part of the structure beyond the blade (Bowler et al., 2019). We show that these densities represent the outer head region of the transition fibre. The distance between the two densities would therefore indicate the maximum diameter of the transition fibre.

We focused on longitudinal sections so that measurements could be compared with the lateral pseudo-diameter measurements in Figure 3.9b. All suitable longitudinal images acquired of the mature basal body and flagella pocket had at least one density visible ($n=20$). Both densities were visible in 40% of images ($n=8$). In order to determine whether the maximum pseudo-diameter in Figure 3.9b by light microscopy aligns with the ultrastructure by EM, the distance was measured between the two densities (mD). This indicates the maximum diameter (mD) (measured laterally) of transition fibres (Fig. 3.10k). The average distance between the two densities (mD, Fig. 3.10k) was measured to be $369\text{nm} \pm 10.6$ ($n=8$). The widest pseudo-diameter by light microscopy in the quantitative analysis had a pseudo-diameter of 345nm (Fig. 3.9b). The data shows that the ultrastructural measurements of transition fibre diameter aligns well with analysis by light microscopy.

We then decided to make measurements of cells with only one density visible to increase the sample size and compare measurements. We therefore measured the distance from the centre of one density to the centre of the basal body or axoneme, which instead would be the ultrastructural equivalent of the 'pseudo-radius'. This distance had an average value of $171.8 \pm 4.1\text{nm}$ ($n=20$). The distance between the distal end of the mature basal body and the most proximal region of the flagella pocket membrane was also measured as an indicator of transition fibre blade length. This distance had an average length of $111.4 \pm 4.6\text{nm}$ ($n=40$). The data shows that these measurements also align with the maximum transition fibre pseudo-diameter measured by light microscopy (Fig. 3.9b).

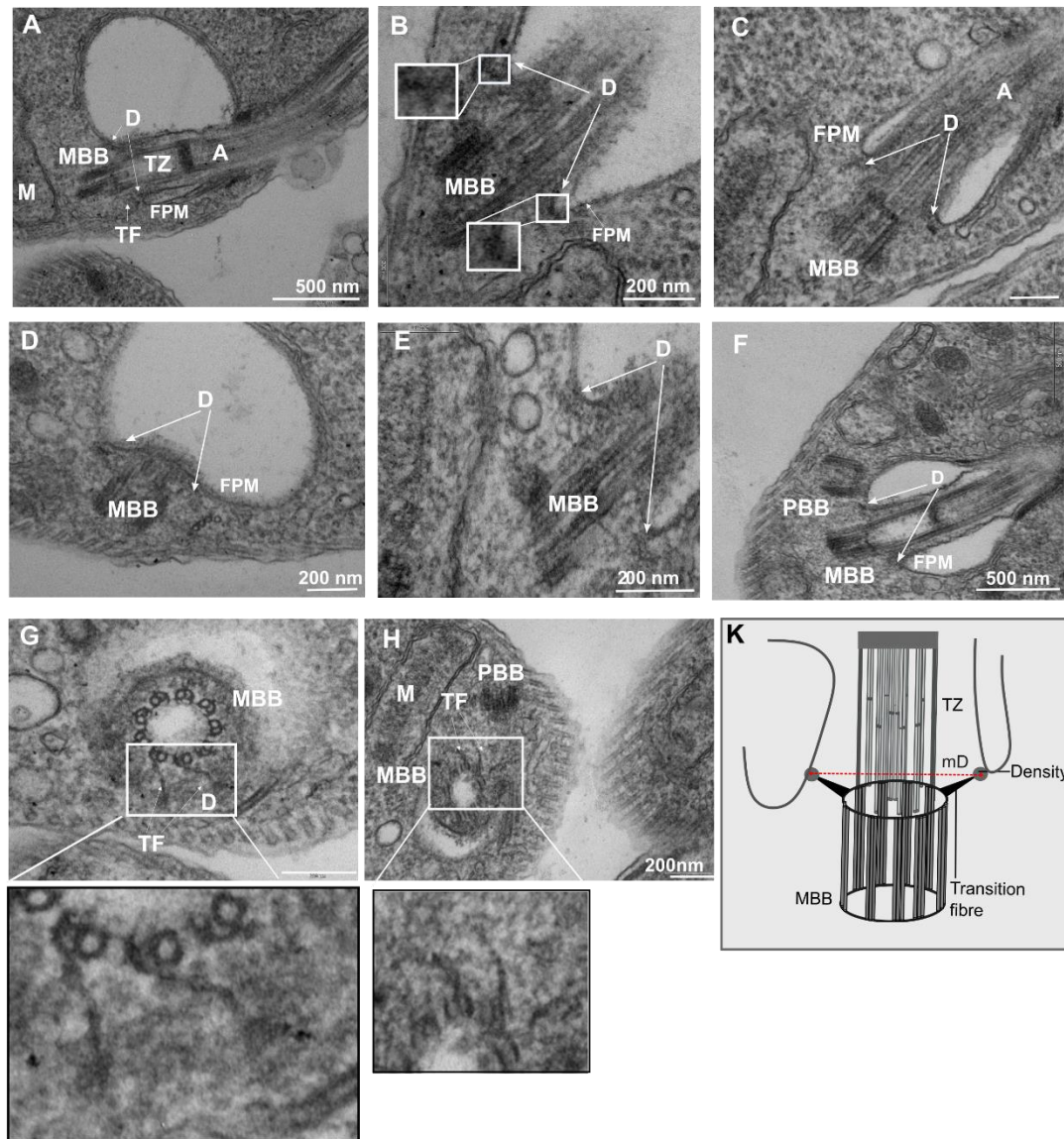


Figure 3.10 70nm section TEM images of the mature basal body and the flagella pocket of the SmOx p927 strain of *T. brucei*. (A, B, C, D, E, F) Longitudinal sections of the mature basal body (MBB), probasal body (PBB), flagella pocket membrane (FPM), axoneme (A), Transition zone (TZ), mitochondrion (M) and transition fibres (TF). Densities (D) observed near the flagella pocket membrane are also indicated. (G, H) Cross-section of the distal end of the mature basal body, where two transition fibres can be seen protruding from the triplet microtubules. (K) Diagram showing the measurement made between the two densities (mD). Inset shows 3x zoom of the region. All images were taken on a Joel1200 kV TEM. The scale bar is visible in the bottom right corner of each image.

3.6 Co-localisation of proteins with a known transition fibre marker

3.6.1 Generation of a transition fibre marker cell line

To further confirm the localisation of each of the putative transition fibre proteins, we aimed to co-localise the remaining 32 proteins with a known transition fibre marker. CEP164 is a highly conserved protein that is present on the transition fibres in a wide range of ciliated/flagellated organisms and it is the most well characterised transition fibre protein (Atkins et al., 2021,

Cajanek and Nigg, 2014, Chaki et al., 2012, Graser et al., 2007, Schmidt et al., 2012, Sillibourne et al., 2011, Tanos et al., 2013, Yang et al., 2018, Hodges et al., 2010). As mentioned previously, CEP164 has three orthologues present in trypanosomes: CEP164A, CEP164B and CEP164C. CEP164A and CEP164B both have a signal that is consistent throughout the cell cycle and is present on the mature basal body of the new and old flagellum alike (Dang et al., 2017). Both proteins would be suitable candidates so we arbitrarily chose CEP164A to take forward as a marker for the transition fibres.

CEP164A was tagged at the N-terminus using the pPOT methodology (methods, 2.2) (Dean et al., 2015). The pPOTv7-G418 plasmid that was used to generate the SAS6::mSc cell line was used again to generate a CEP164A mScarlet cell line (Fig. 3.2). A PCR was carried out according to the methodology described in chapter 2 (methods, 2.2) to generate an amplicon that can be transfected into the cell. The amplicon generated via PCR was loaded onto an agarose gel to confirm the size (Fig. 3.11).

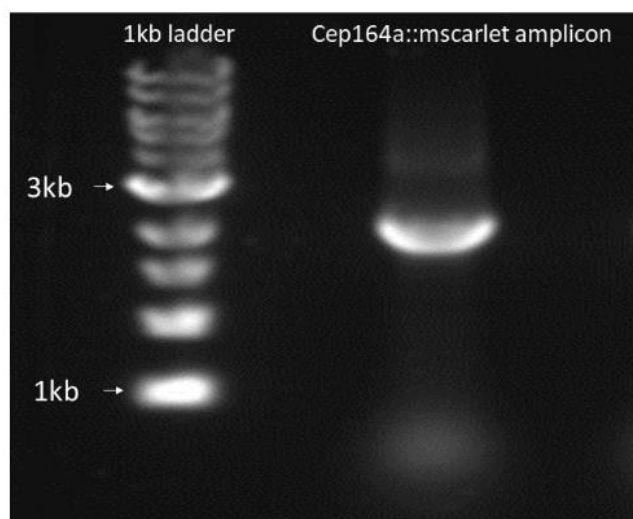


Figure 3.11 CEP164A mScarlet amplicon loaded onto a 0.8% agarose gel to confirm amplicon size of ~1700bp. 1kb DNA ladder shown to the left.

The CEP164A mScarlet PCR amplicon was transfected and a stable G418 resistant cell line was generated expressing CEP164A::mScarlet (see methods 2.2). A single CEP164A foci was observed at the basal body region during G1 which is represented by a single nucleus and a single kinetoplast (Fig. 3.12). S-phase is represented by basal body duplication from two to four foci consisting of 2 mature basal bodies and 2 probasal bodies. As transition fibres are on the mature basal body, S-phase can be also identified by the duplication of the TbCEP164A signal from 1 to 2 foci. During S-phase the CEP164A signal is equal at the new and old flagellum. A single CEP164A foci remains present in equal intensity on both flagella during G2 and M-phase (Fig. 3.12) Figure 3.12 shows CEP164A is present at the basal body region and exhibits a consistent fluorescent

signal throughout the cell cycle. We concluded from this that CEP164A is a suitable marker protein for co-localisation of our cohort of putative transition fibre proteins.

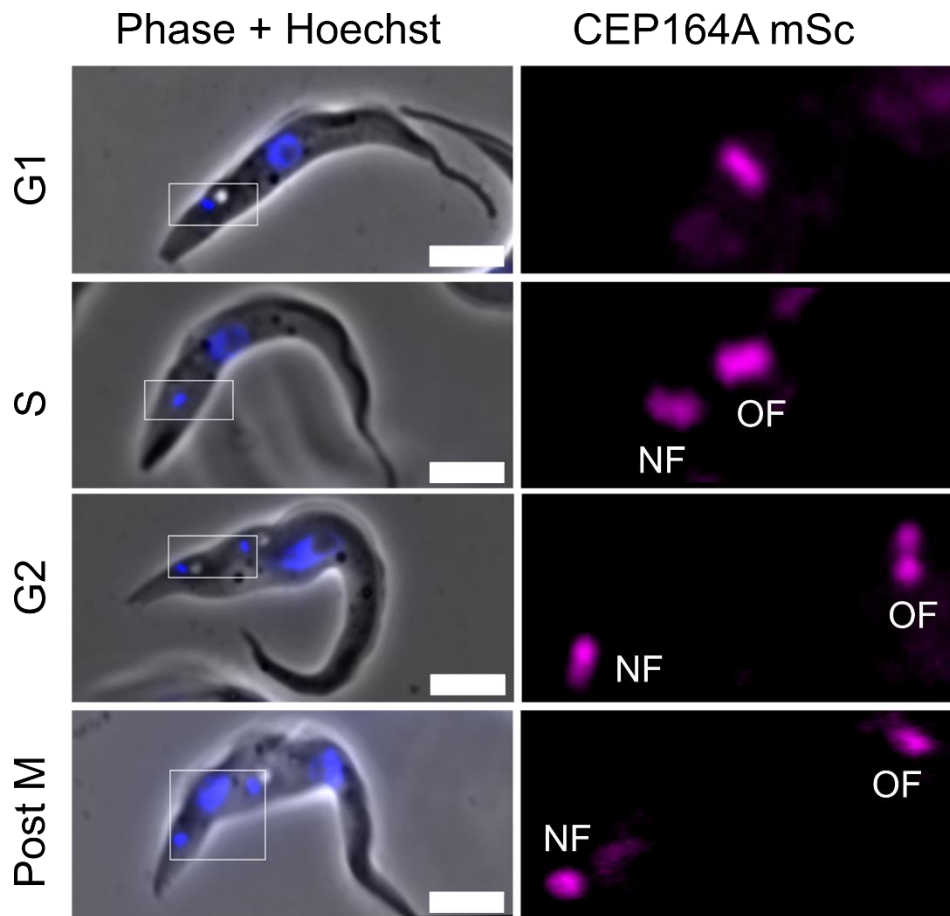


Figure 3.12 Live cell images of TbCEP164A endogenously tagged with the mScarlet fluorescent protein. TbCEP164A signal is shown throughout the cell cycle. G1-phase is characterised by 1 kinetoplast, 1 nucleus (1K1N) and 2 basal bodies represented by a single CEP164A signal on what would be the mature basal body in a pair of basal bodies. G2 is indicated by the division of the kinetoplast (2K1N) and Post-M phase by nuclear segregation (2K2N). **New flagellum and old flagellum are indicated as NF and OF.** All images were taken on a Zeiss wide-field at 100x magnification. Scale bar represents 5 μ m.

3.6.2 Generation of double tagged cell lines for co-localisation

The remaining 32 putative transition fibre proteins (excluding CEP164A) were co-expressed in the cell line expressing CEP164A::mScarlet to confirm their localisation at the transition fibres. A PCR was carried out for all 32 transition fibre proteins using the pPOTv6 mNG plasmid, which included 2 3xTY epitope tags that could potentially be targeted for immunofluorescence or western blotting (Fig. 3.3). The PCR amplicon for all 32 proteins are shown in Figure 3.13. Each amplicon had an expected size of approximately 1700bp. These were taken forward and stable cell lines expressing both the CEP164A marker and the putative transition fibre protein (see methods, 2.2).

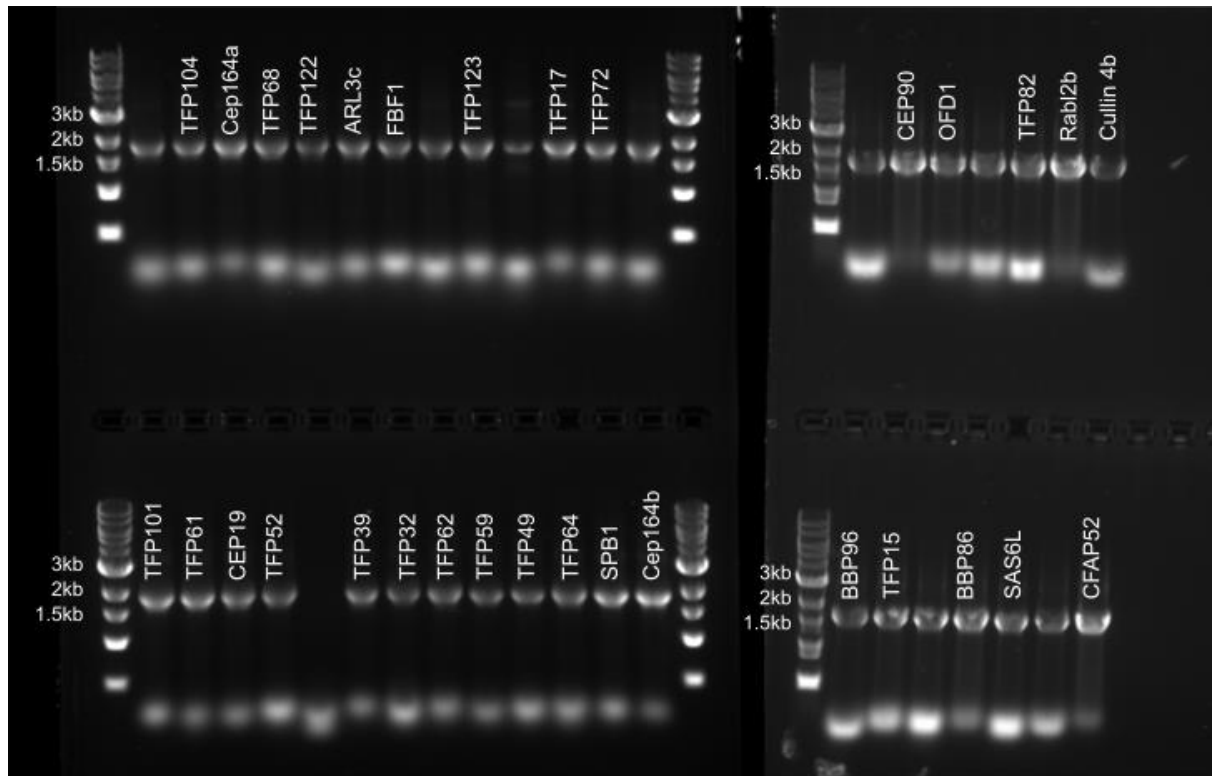
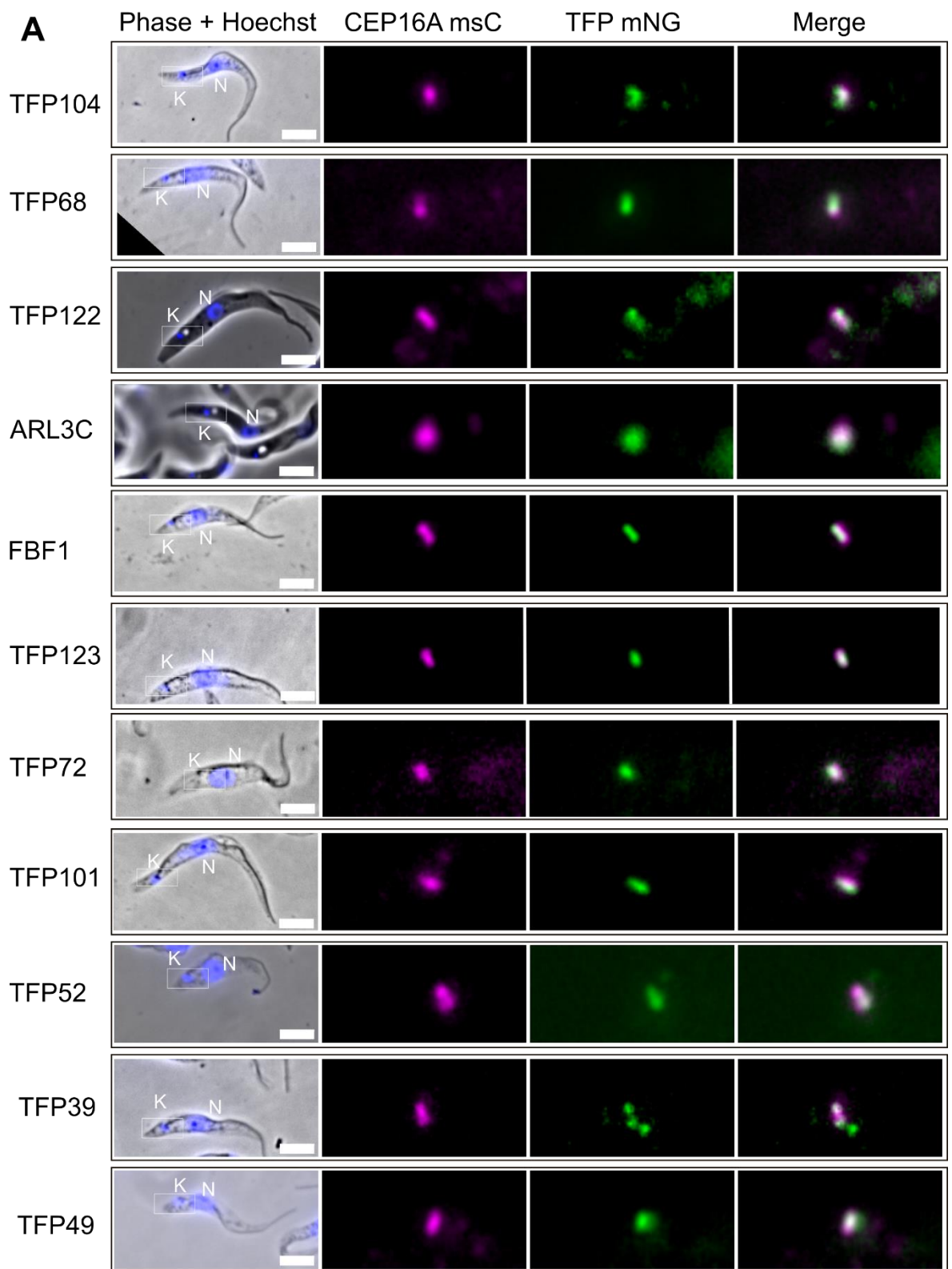


Figure 3.13 The amplicon corresponding to each putative transition fibre protein shown in the white text above each well with a 1kb DNA ladder (left). Samples were run on a 0.8% agarose gel.

3.6.3 Co-localisation of putative transition fibre proteins with CEP164A

From an original cohort of 35 proteins, two proteins were removed based on quantitative analysis in Figure 3.9. We also left out CEP164C for further analysis as this protein has been examined in trypanosomes in a recent study which showed that it co-localises with RP2 (Atkins et al., 2021). Out of 33 transition fibre proteins in our cohort, 31 proteins were taken forward for double tagging (CEP164C and CEP164A not included). Of these putative transition fibre proteins taken forward in our cohort, 22 of these were successfully double tagged with CEP164A. All 22 proteins that were co-expressed with CEP164A were found to fit the criteria as a transition fibre protein (TFP) based on sufficient co-localisation determined by the amount of overlap between the two signals. Proteins were classed as either entirely overlapping, overlapping distal, overlapping proximal, entirely distal or entirely proximal to CEP164A. Of the 22 transition fibre proteins (Fig. 3.14, green) successfully co-localised with TbCEP164A (Fig. 3.14, magenta), 15 exhibited a signal that was entirely overlapping with TbCEP164A (Fig. 3.14a). CEP19L, TFP17 and TFP61 were classed as overlapping proximal (Fig. 3.14b) and four proteins (TFP32, TFP62, TFP59 and BBP86) were classed as overlapping distal (Fig. 3.14c). No proteins were entirely distal or proximal to CEP164A. All 22 proteins are considered bona fide transition fibre proteins based on analysis in Figure 3.9 and co-localisation in Figure 3.14.



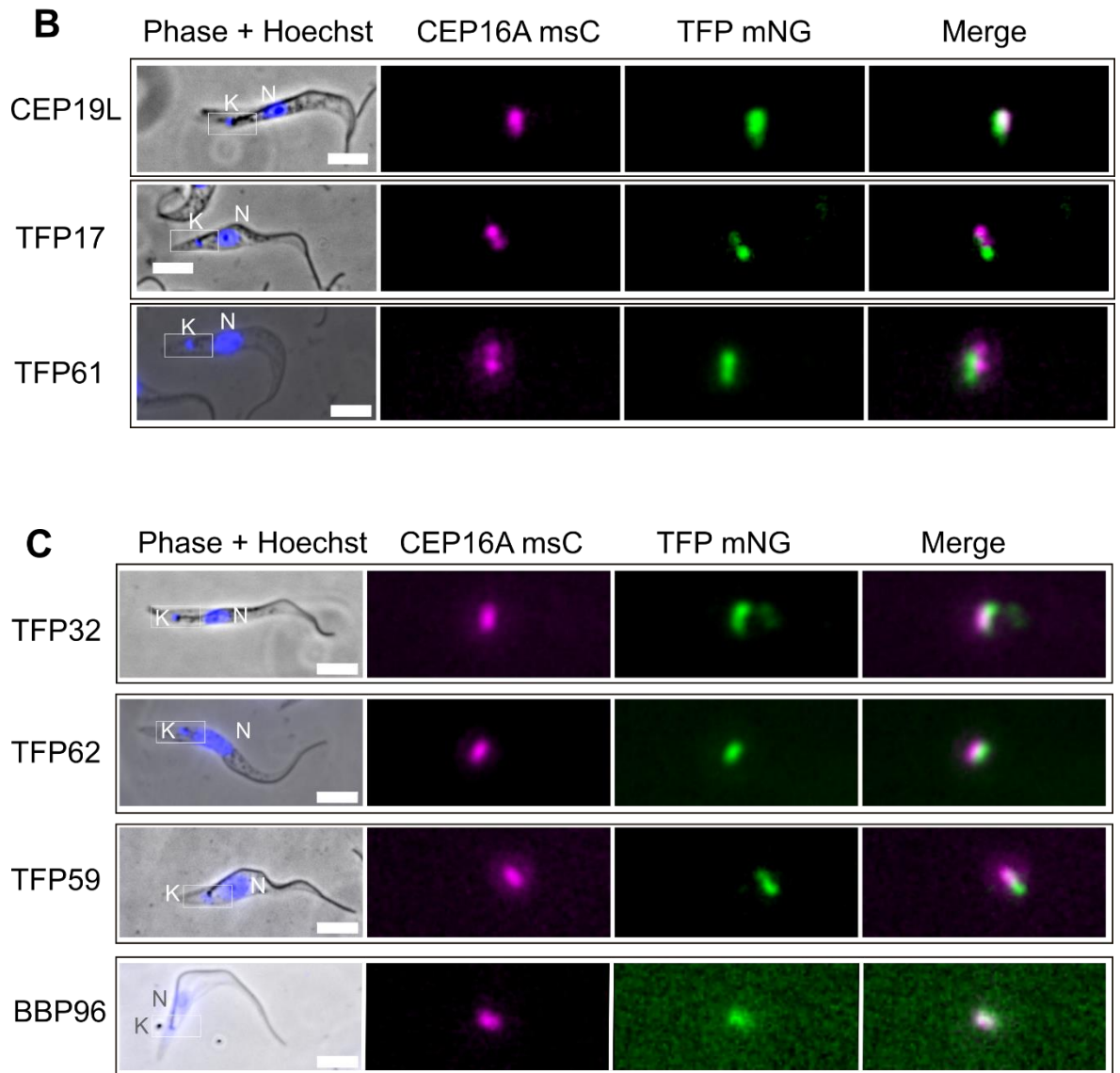


Figure 3.14 29 Transition fibre proteins (green) tagged at the N-terminus with the mNeonGreen (Green) fluorescent protein co-expressed with the bona fide transition fibre protein CEP164A (magenta) tagged with mScarlet. DNA in the nucleus (N) and kinetoplast (K) is labelled with Hoescht (blue). (a) Localisation of 15 transition fibre proteins that are classed as ‘entirely overlapping’ with CEP164A. (b) Localisation of transition fibre proteins classed as ‘overlapping proximal’ with CEP164A. (c) Localisation of transition fibre proteins classed as ‘overlapping distal’ with CEP164A. Images were taken of methanol-fixed cytoskeletons all proteins that were stably associated with the cytoskeleton and live cell images for 2 soluble proteins could not survive cytoskeletal extraction (Arl3c, TFP122). All images were taken at 100x magnification.

3.6.4 Co-localisation of putative transition fibre proteins with RP2

Out of 32 putative transition fibre proteins taken forward, only 22 were successfully co-localised with CEP164A by endogenous double tagging with the mScarlet and mNeonGreen fluorescent proteins. We aimed to generate single tagged cell lines for the remaining proteins and immunolabel a known transition fibre protein. Four proteins (SPB1, SAS6L, Rab12b and TFP15) were successfully single tagged with the mNG fluorescent protein according to the same methodology described in chapter 2 (see methods Fig. 2.2). The remaining five proteins (CFAP52, BBP86, Cullin 4b, TFP82 and TFP131) were unsuccessful in multiple attempts.

RP2 is a protein that localises to transition fibres in *T. brucei* (Stephan et al., 2007). The YL1/2 antibody recognises the carboxy terminus of alpha tubulin when tyrosinated (Wehland et al., 1984, Wehland and Willingham, 1983). However, the C-terminal region of the RP2 protein contains a classic YL1/2 recognition epitope and can therefore be labelled with YL1/2 (Andre et al., 2014). We took the single tagged cell lines and immunostained for RP2 using the YL1/2 antibody.

Figure 3.15 shows the co-localisation of SBP1 (green), SAS6L (green), TFP15 (green) and Rab12b (green) with RP2 (magenta) in G1 cells. TFP15 was classed as 'entirely overlapping' with RP2 (Fig. 3.15a). SPB1 and Rab12b were classed as 'overlapping distal' to RP2 as the signal was slightly offset from RP2 but still overlapping considerably (Fig. 3.15b). However, SAS6L exhibited a signal that was entirely distal to RP2 (Fig. 3.15c) and showed no overlap with the marker protein. SAS6L would therefore be excluded from the cohort due to a lack of co-localisation with the bona fide transition fibre protein RP2. However, it should be noted that transition fibre protein CEP164C has been shown to have a localisation that would be classed as overlapping distal to RP2 (Atkins et al., 2021). In summary, out of 26 proteins co-localised with CEP164A or RP2, all but SASL are considered transition fibre proteins. Based on this analysis we conclude 25 candidate proteins are bona fide transition fibre proteins.

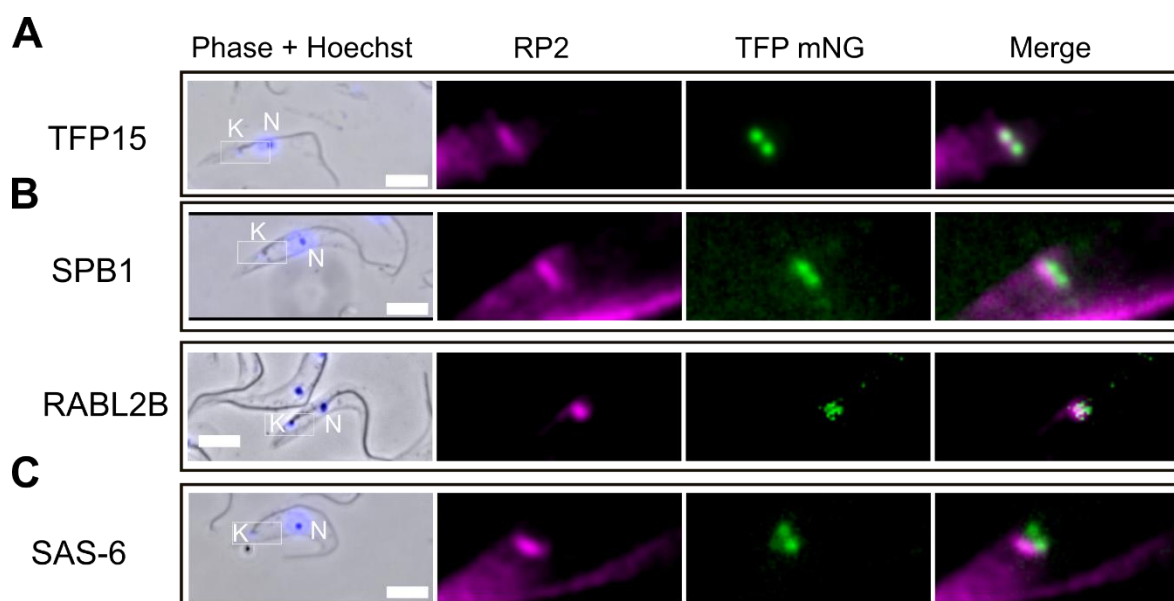


Figure 3.15 Cells expressing endogenously tagged TFP15, SPB1, RABL2B, and SAS6L tagged at the N-terminus with the mNeonGreen fluorescent protein co-localised with the bona fide transition fibre protein RP2 (magenta) using the YL1/2 antibody. DNA in the nucleus (N) and kinetoplast (K) is labelled with Hoescht (blue). (a) Co-localisation of TFP15 (green) with RP2 (magenta). (b) Co-localisation of SPB1 (green) and Rabl2b with RP2 (magenta) at the flagella base of a G1 cell. (c) Co-localisation of SAS6L (green) with RP2 (magenta) at the flagella base of a G1 cell. Images were taken of methanol-fixed cytoskeletons. Images were taken at 100x magnification using the Zeiss wide-field microscope.

In summary, we have identified 35 putative transition fibre proteins based on their localisation with known basal body protein SAS-6. We defined proteins as putative transition fibre proteins based on three criteria (a) Localisation is specific to the mature basal body, (b) localisation is distal to the mature basal body and (c) The signal of the tagged protein is wider than the basal body. We quantified this in more detail by creating a FIJI script that measures pseudo-diameter (diameter from a lateral view) and distance from the mature basal body. Out of 29 proteins from 2 proteins (Tb927.9.11890 and Tb927.3.4470) were excluded based on this analysis as they did not fit the criteria of a transition fibre protein (Fig. 3.9). TEM was also used to measure the ultrastructural diameter of transition fibres and we found the EM data aligns with the measurements made by light microscopy.

Co-localising proteins with a known transition fibre marker revealed that 1 protein (SAS6L) did not co-localise with the marker for the transition fibres (RP2) and therefore was excluded. We did not carry out further analysis on CEP164C as it has been well examined in a previous study (Atkins et al., 2021).

We therefore take forward 31 transition fibre proteins as a set of proteins for further analysis. Information of the cohort of transition fibre proteins is summarised in Table 3.4. This includes their gene ID, protein name, association to the cytoskeleton, pseudo-diameter, distance from

SAS6 and their classification based on co-localisation. All transition fibre proteins in Table 3.4 are ordered based on their pseudo-diameter (where applicable) and those that were excluded are highlighted in orange.

Table 3.4 Cohort of 31 transition fibre proteins taken forward (white) and three excluded proteins (orange) based on pseudo diameter or co-localisation. From left: Gene ID, Given name or existing name of protein, association to the cytoskeleton, pseudo-diameter (Fig. 3.9), distance from SAS6 (Fig. 3.9), and co-localisation with transition fibre marker (Fig. 3.14, 3.15).

Gene ID	Given name	Cytoskeletal	Pseudo-diameter (nm)	Distance from MBB (nm)	Co-localisation
Tb927.9.11890	Tb927.9.11890/TFP100	✓	187	236	Excluded as distal to TFP100
Tb927.6.4100	FBF1	✓	225.2	177.7	Entirely overlapping with CEP164A
Tb927.9.11830	TFP72	✓	227.5	169.2	Entirely overlapping with CEP164A
Tb927.11.5630	TFP62	✓	233.8	212.7	Overlapping distal to CEP164A
Tb927.11.8810	TFP49	✓	239.9	161.9	Entirely overlapping with CEP164A
Tb927.5.3730	TFP68	✓	240.4	201.1	Entirely overlapping

					with CEP164A
Tb927.11.7110	TFP59	✓	245.5	219	Overlapping distal to CEP164A
Tb927.11.9540	TFP64	✓	246.3	184.5	Entirely overlapping with CEP164A
Tb927.7.5370	TFP123	✓	246.8	201.9	Entirely overlapping with CEP164A
Tb927.10.10690	CEP19	✓	247.8	175	Overlapping proximal to CEP164A
Tb927.3.4470	Tb927.3.4470	✓	249.2	127.8	Excluded as more proximal than POC5
Tb927.5.2440	CEP164A	✓	254.6	172.1	Known TF protein
Tb927.10.620	TFP61	✓	255.1	221.3	Overlapping proximal to CEP164A
Tb927.10.9060	TFP52	✓	257.2	201.3	Entirely overlapping with CEP164A
Tb927.11.11650	CEP164b	✓	259.2	204.2	Entirely overlapping

					with CEP164A
Tb927.9.12310	TFP101	✓	262.5	189.5	Entirely overlapping with CEP164A
Tb927.3.5450	BBP96	✓	263.1	165	Overlapping distal to CCEP164A
Tb927.11.9620	TFP82	✓	264.7	162.2	n/a
Tb927.7.5690	CEP90	✓	264.9	162.4	Entirely overlapping with CEP164A
Tb927.10.5910	BBP86	✓	267.3	162.5	n/a
Tb927.8.2430	TFP131	✓	268.9	172.1	n/a
Tb927.8.5550	TFP17	✓	268.9	201.3	Overlapping proximal to CEP164A
Tb927.10.7920	SAS6-like	✓	271	155	Distal to RP2
Tb927.10.3000	OFD1	✓	275.3	168.7	Entirely overlapping with CEP164A
Tb927.11.3460	TFP32	✓	275.7	215.2	Overlapping distal to CEP164A
Tb927.3.1290	Cullin4b	✓	300.1	182.2	n/a
Tb927.11.3050	TFP39	✓	301.4	204.2	Entirely overlapping

					with CEP164A
Tb927.4.4060	TFP104	✓	314.4	194.9	Entirely overlapping with CEP164A
Tb927.6.2220	TFP15	✓	345	126.3	Entirely overlapping with RP2
Tb927.6.2880	TFP122	X	X	X	Entirely overlapping with CEP164A
Tb927.6.3650	ARL3C	X	X	X	Entirely overlapping with CEP164A
Tb927.11.9690	SPB1	✓	X	X	Overlapping distal to RP2
Tb927.11.7560	CFAP52	✓	X	X	n/a
Tb927.11.15030	RabL2b	✓	X	X	Overlapping distal to RP2

4. Cell cycle regulation, assembly and conservation of transition fibres

In the previous chapter, we identified 31 transition fibre proteins and determined their spatial arrangement along the transition fibre based on their pseudo-diameter and distance from the basal body. We also confirmed their localisation to the transition fibre by co-localising with one of the known transition fibre proteins CEP164A or RP2. Following this analysis, we aim to analyse the localisation of cell cycle regulated transition fibre proteins in more detail and assess the timing of recruitment of proteins to the transition fibre to better understand the assembly of transition fibres. Lastly, the following chapter will assess the orthology of each of the 31 transition fibre proteins in a range of ciliated and nonciliated eukaryotic organisms to gain an insight into the conservation of transition fibres in eukaryotic cilia.

4.1 Cell cycle regulated transition fibre proteins at the old and new flagellum

Many unicellular eukaryotes possess multiple flagella within the same cell. In the majority of these eukaryotic organisms, there is a need to assemble a new flagellum and simultaneously maintain the length of an existing old flagellum. Trypanosomes possess a full length old flagellum while simultaneously assembling a new flagellum within the same cell. It is clear that there must be a regulatory system to allow proteins to enter the new flagellum for assembly whilst limiting growth on the old flagellum. A growth and lock model was proposed in trypanosomes which states that a lock is placed on the old flagellum once it reaches its final length (Bertiaux et al., 2018). A central component of this mechanism is CEP164C which localises exclusively to the transition fibres of the old flagellum (Atkins et al., 2021, Bertiaux et al., 2018). We assessed the localisation of our 31 transition fibre proteins in more detail to see if any proteins were specific to either the new or old flagellum in a similar manner to CEP164C.

We found that all proteins localise to both the new and old flagellum (Chapter 3, Fig. 3.4 –Fig. 3.7). However, upon visual inspection, several proteins showed a difference in the intensity of the mNG signal at the base of the new or old flagellum in dividing cells (Chapter 3, Fig. 3.4- Fig. 3.7). Therefore, we visually assessed the differences in intensity between the old and new flagellum throughout the cell cycle to identify candidate transition fibre proteins that could be involved in mechanisms related to old flagellum locking or new flagellum assembly.

4.1.1 Cell cycle regulated localisation patterns brighter on the old flagellum

A total of four transition fibre proteins had a signal that is brighter at the transition fibres of the old flagellum. These proteins were Rab12b, TFP39, TFP104 and TFP122. G1 is shown at the beginning of

the cell cycle when the cells possess a single nucleus accompanied by a single kinetoplast towards the posterior end of the cell and only one basal body pair (SAS6). A single transition fibre signal can be observed in cells in G1. S-phase is defined by the duplication of the basal bodies in trypanosomes. Two basal body pairs can be observed and two transition fibre protein signals on each basal body pair of the old and new flagellum. G2 is characterised by the duplication of the kinetoplast. Cells in G2 possess a single nucleus accompanied by two kinetoplasts. Post-mitotic cells (post-M) have undergone nuclear division and kinetoplast division and therefore are characterised by two nuclei accompanied by two kinetoplasts (Hammarton, 2007, Sherwin and Gull, 1989).

Figure 4.1 shows the localisation of Rabl2b throughout the cell cycle stages described above. The Rabl2b signal is shown to be brighter at the old flagellum in S-phase and is maintained at a brighter level at the old flagellum in G2 and post-mitosis (Fig. 4.1). We conclude that Rabl2b maintains a stronger signal at the old flagellum throughout the cell cycle and this disparity is maintained in G1 cells, with around half of cells exhibiting a weaker signal and half with a stronger signal. This indicates that expression of the protein is restored at the start of the next cell cycle.

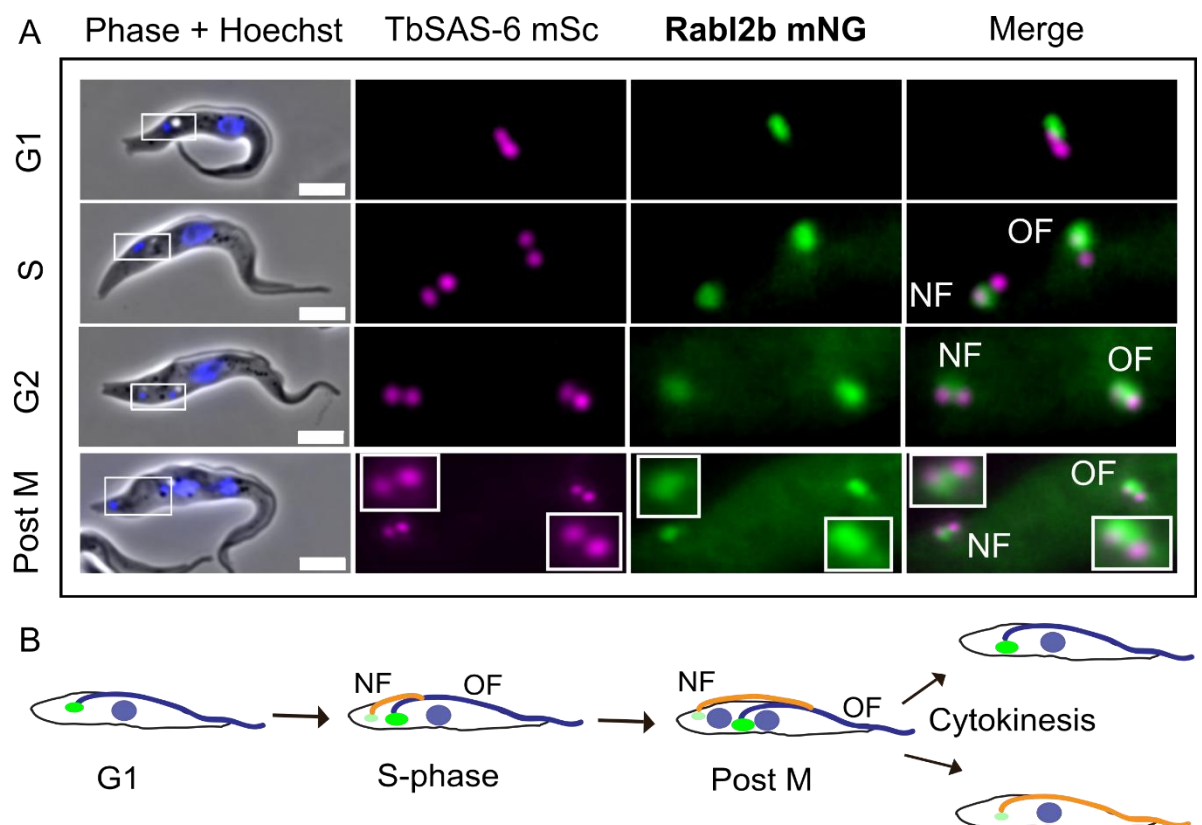


Figure 4.1 (A) Live cell images of transition fibre protein Rabl2b (green) tagged with mNeongreen. Images show a Rabl2b foci at the new and old flagellum across the cell cycle. Images are shown for cells in G1, S-phase, G2 and post-mitotic (post-M). The basal bodies of the new and old flagellum are represented by SAS-6 mScarlet. The new flagellum and old flagellum are labelled NF and OF respectively. All images were taken at 100x magnification on a Zeiss wide-field microscope. Scale bar represents 5µm. The inset shown over the phase + hoechst images represents a 5x zoom in the G1,

*S-phase and G2 cells and a 4x zoom in post-M cells. The second inset **within** the colour image of post-M cells represents a 10 x zoom. (B) Schematic of Rabl2b (green) at the old flagellum (orange) and new flagellum (blue) throughout the cell cycle from G1 (left) to cytokinesis (right).*

The second transition fibre protein that exhibits a brighter mNG signal at the base of the old flagellum is TFP104. The TFP104 signal is brighter at the old flagellum in S-phase and is maintained at a brighter level at the old flagellum in G2 and post-M cells (Fig. 4.2a). However, following visual inspection, there did not appear to be two populations of G1 cells with bright and weak mNG signal, suggesting that the protein is likely restored shortly after cytokinesis.

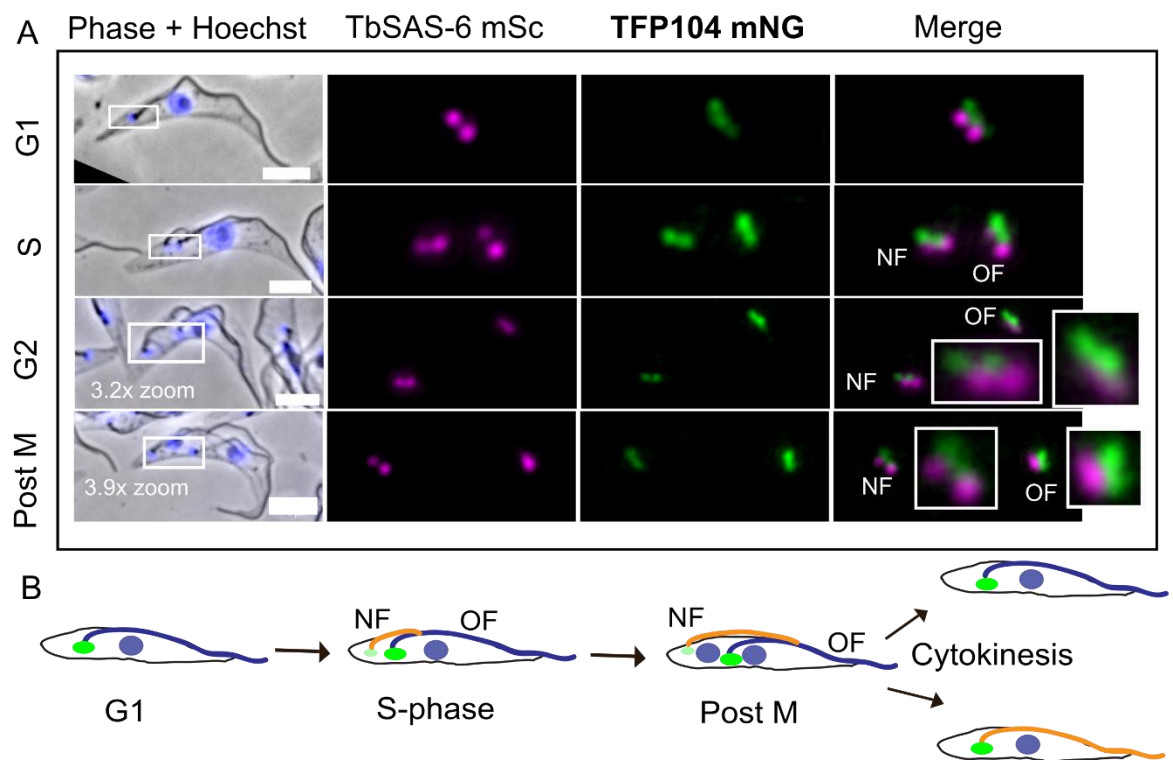


*Figure 4.2 (A) Live cell images of TFP104 (green) tagged with mNeongreen. Images show a TFP104 foci at the new and old flagellum across the cell cycle. Images are shown for cells in G1, S-phase, G2 and post-mitotic (post-M). The basal bodies of the new and old flagellum are represented by SAS-6 mScarlet. The new flagellum and old flagellum are labelled NF and OF respectively. All images were taken at 100x magnification on a Zeiss wide-field microscope. Scale bar is equal to 5µm. The inset shown over the phase + hoechst images represents a 5x zoom in G1 and S-phase cells, a 3.9x zoom in G2 cells and a 3.9x zoom in post-M cells. The second inset **within** the merge images of G2 and post-M cells represents a 10 x zoom. (B) Schematic of TFP104 (green) at the base of the old flagellum (orange) and new flagellum (blue) throughout the cell cycle from G1 (left) through to cytokinesis (right).*

TFP39 also exhibits a stronger signal at the base of the old flagellum compared to that of the new flagellum. The TFP39 signal is brighter at the old flagellum in S-phase and is maintained at a brighter level at the old flagellum throughout the cell cycle in G2 and post-M cells (Fig. 4.3). In G1 cells there

appears to be two populations of G1 cells with brighter and weaker signals at the basal body, indicating that TFP39 expression is restored at the start of the cell cycle.

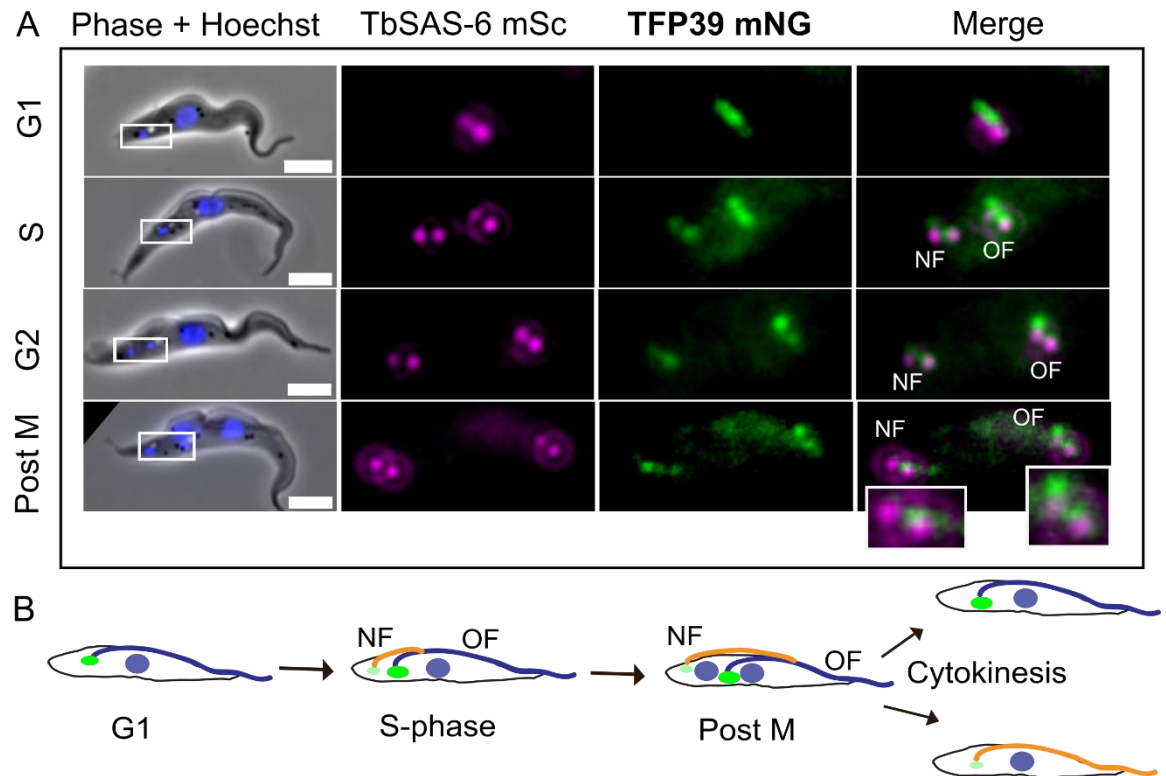


Figure 4.3 (A) Live cell images of TFP39 (green) tagged with mNeongreen. Images show a TFP39 foci at the new and old flagellum across the cell cycle. Images are shown for cells in G1, S-phase, G2 and post-mitotic (post-M). The basal bodies of the new and old flagellum are represented by SAS-6 mScarlet. The new flagellum and old flagellum are labelled NF and OF respectively. All images were taken at 100x magnification on a Zeiss wide-field microscope. Scale bar represents 5 μ m. The inset shown over the phase + hoechst images represents a 5x zoom. The second inset **within** the merge images of post-M cells represents a 10 x zoom. (B) Schematic of TFP39 (green) at the base of the old flagellum (orange) and new flagellum (blue) throughout the cell cycle from G1 (left) through to cytokinesis (right).

The last protein that exhibits a cell cycle regulated signal pattern that is brighter at the mature basal body of the old flagellum compared to the new flagellum is TFP122. TFP122mNG signal is also brighter at the base of the old flagellum throughout the cell cycle in S-phase, G2 and post-M cells (Fig. 4.4). There is some background visible in S-phase and post-M images. This is due to the fact the signal on the new flagellum is very weak. We consider the signal at the new flagellum to be marginally above the background level and so we class it as present on the new flagellum. Cells in G1 appear to have two populations with weak and strong signals, indicating that like Rabl2b and TFP39, that the protein is restored to the flagellum at the start of the cell cycle.

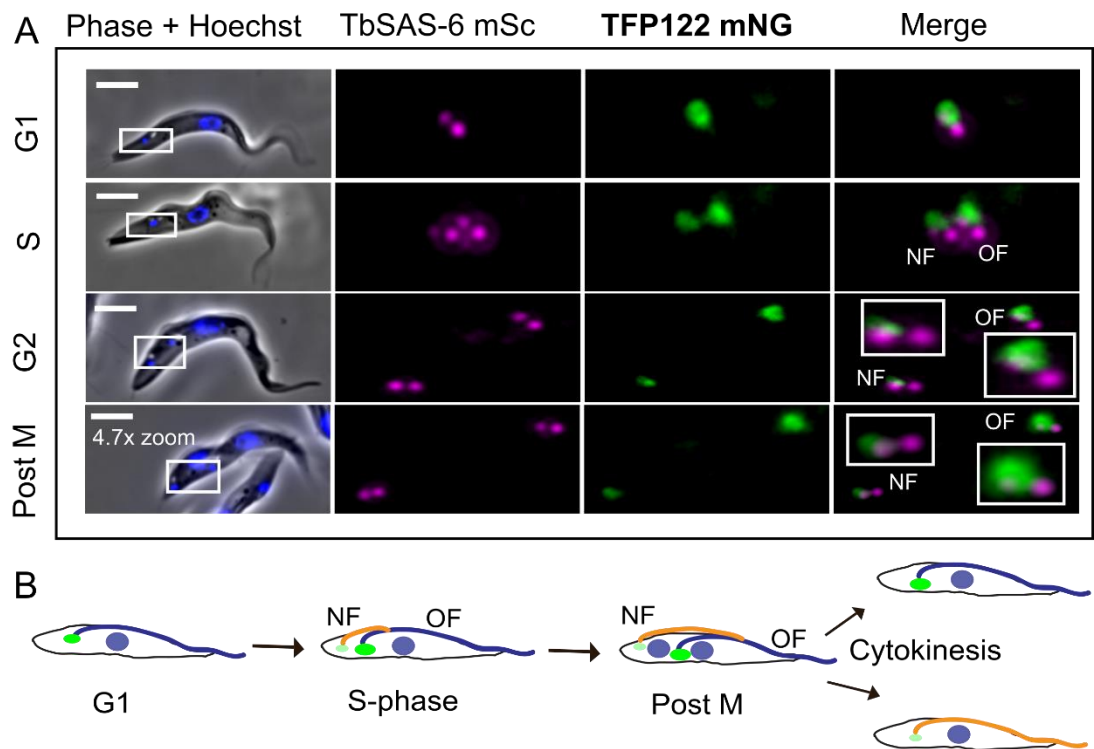


Figure 4.4 (A) Live cell images of TFP122 (green) tagged with mNeogreen. Images show a TFP122 foci at the new and old flagellum across the cell cycle. Images are shown for cells in G1, S-phase, G2 and post-mitotic (post-M). The basal bodies of the new and old flagellum are represented by SAS-6 mScarlet. The new flagellum and old flagellum are labelled NF and OF respectively. All images were taken at 100x magnification on a Zeiss wide-field microscope. Scale bar represents 5 μ m. The inset shown over the phase + hoechst images represents a 5x zoom in G1 and S-phase cell and a 4.7x zoom for G2 and post-M cells. The second inset **within** the merge images of G2 and post M cells represents a 10 x zoom. (B) Schematic of TFP122 (green) at the base of the old flagellum (orange) and new flagellum (blue) throughout the cell cycle from G1 (left) through to cytokinesis (right).

Figure 4.1, 4.2, 4.3 and 4.4 show all four transition fibre proteins (Rabl2b, TFP104, TFP39 and TFP122) exhibit a cell cycle regulated localisation pattern that is brighter at the basal body region of the old flagellum. For all four proteins, this disparity in signal intensity is maintained through the entire cell cycle from S-phase, through G2 and post-mitosis. The data suggest that for Rabl2b, TFP39 and TFP122, the protein is restored at the start of the cell cycle. In the case of TFP104, the protein is restored shortly following cytokinesis.

4.1.2 Cell cycle regulated localisation pattern brighter on the new flagellum

Out of the 31 transition fibre proteins taken forward, a total of two proteins exhibit a cell cycle regulation localisation pattern that is brighter at the base of the new flagellum, which has not yet been shown before. TFP72 is the first protein that is brighter on the new flagellum (Fig. 4.5). In G1, a single mNG foci can be observed on the distal end of the mature basal body. In S-phase, basal body duplication has occurred and the TFP72 signal also duplicates. At this stage the TFP72 signal at the new flagellum is clearly brighter than that of the old flagellum. The TFP72 signal remains

brighter on the new flagellum in G2 and post-mitotic cells. We therefore conclude that TFP72 exhibits a brighter signal at the base of the new flagellum throughout the cell cycle (Fig. 4.5). Additionally, there appears to be two populations with a weak and bright signal in G1 cells, suggesting expression is restored at the start of the cell cycle (G1).

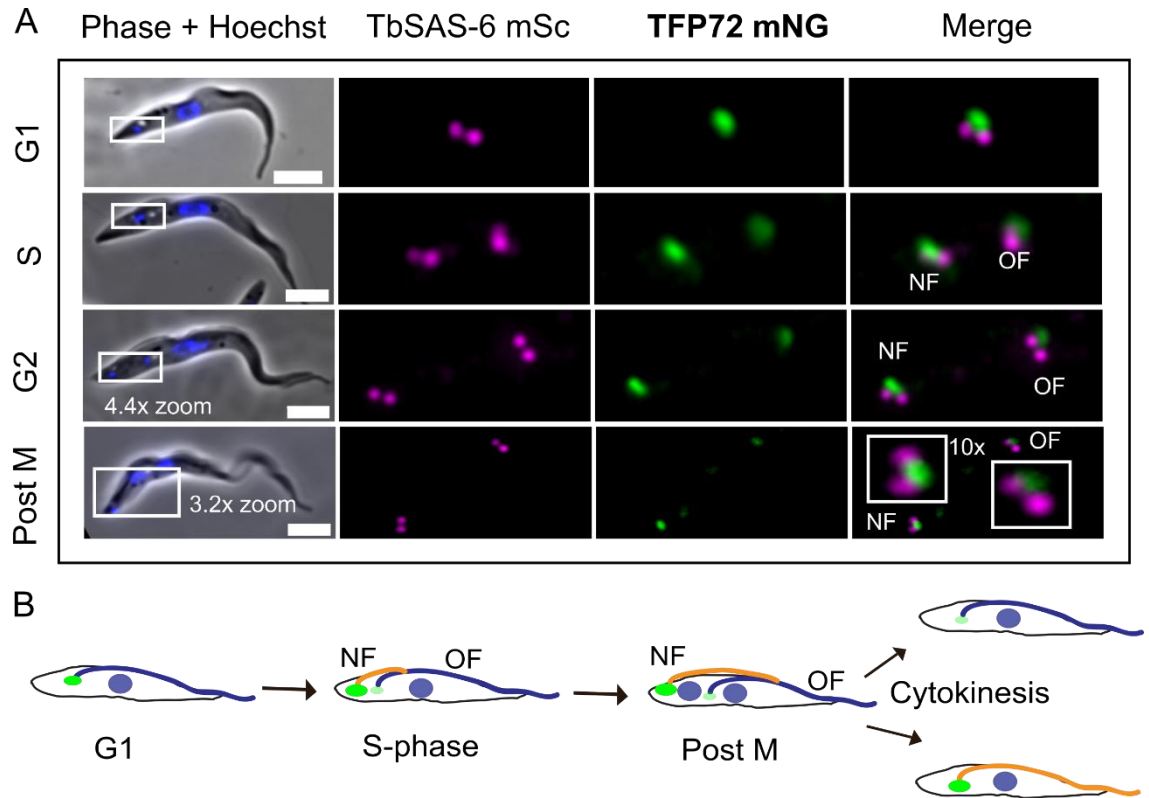


Figure 4.5 (A) Live cell images of TFP72 (green) tagged with mNeongreen. Images show a TFP72 foci at the new and old flagellum across the cell cycle. Images are shown for cells in G1, S-phase, G2 and post-mitotic (post-M). The basal bodies of the new and old flagellum are represented by SAS-6 mScarlet. The new flagellum and old flagellum are labelled NF and OF respectively. All images were taken at 100x magnification on a Zeiss wide-field microscope. Scale bar represents 5 μm. The inset shown over the phase + hoechst images represents a 5x zoom in G1 and S-phase cells, a 4.4x zoom for G2 cells and a 3.2x zoom for post-M cells. The second inset **within** the merge images of post M cells represents a 10 x zoom. (B) Schematic of TFP72 (green) at the base of the old flagellum (orange) and new flagellum (blue) throughout the cell cycle from G1 (left) through to cytokinesis (right).

In contrast, TFP59 is brighter at the new flagellum following basal body duplication during S-phase. However, as the cell progresses through the cell cycle, the signal between the old and new flagellum becomes even (Fig. 4.6). In Figure 4.6, the S-phase cell clearly shows a difference in signal intensity between the old and new flagellum. However, by G2 the signal begins to return to the mature basal body of the old flagellum. Following mitosis (post-M), the signal has returned to the old flagellum so that the intensity of TFP59 at the old and new flagellum is near even.

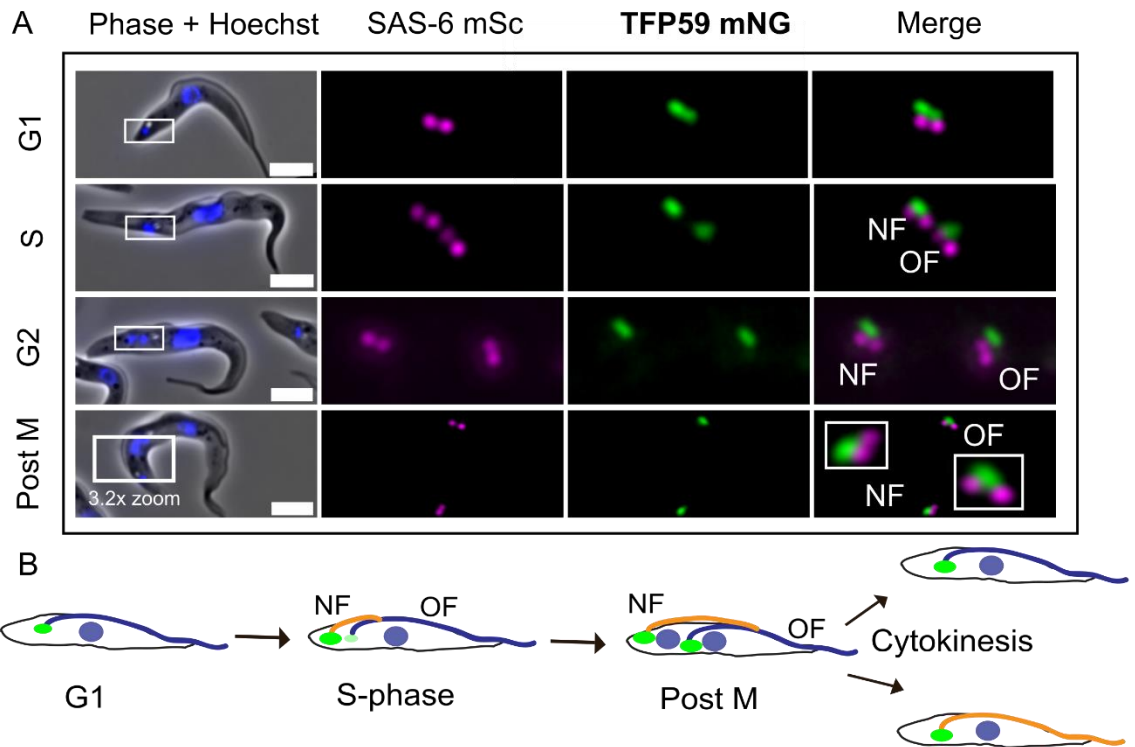


Figure 4.6 (A) Live cell images of TFP59 (green) tagged with mNeongreen. Images show TFP59mNG signal at the new and old flagellum across the cell cycle. Images are shown for cells in G1, S-phase, G2 and post-mitotic (post-M). The basal bodies of the new and old flagellum are represented by SAS-6 mScarlet. The new flagellum and old flagellum are labelled NF and OF respectively. All images were taken at 100x magnification on a Zeiss wide-field microscope. Scale bar represents 5µm. The inset shown over the phase + hoechst images represents a 5x zoom in G1, S-phase and G2 cells. A 3.2x zoom is shown for post-M cells. The second inset **within** the merge images of post M cells represents a 10 x zoom. (B) Schematic of TFP59 (green) at the base of the old flagellum (orange) and new flagellum (blue) throughout the cell cycle from G1 (left) through to cytokinesis (right).

Trypanosomes are able to assemble a new flagellum whilst simultaneously maintaining the length of an existing flagellum. We hypothesise there must be a unique set of proteins in trypanosomes that are targeted to (or excluded from) each flagellum to support its function. Figure 4.1-4.6, we show 6 transition fibre proteins that have a brighter mNG signal at the old or new flagellum at some point during the cell cycle. We conclude these are promising candidates that could explain how trypanosomes simultaneously assemble a new flagellum while maintaining an existing flagellum. A schematic of the cell cycle shows the localisation and relative brightness of each transition fibre protein throughout the cell cycle (Fig. 4.1b, 4.2b, 4.3b, 4.4b, 4.5b, 4.6b)

It should also be noted that transition fibres have traditionally been considered stable structures. However, more recently there have been suggestions that transition fibres are more dynamic in their structure and molecular composition (Bowler et al., 2019). This data shows that trypanosome

transition fibres are dynamic in the composition with several proteins being targeted to or removed from the new and old flagellum throughout the cell cycle.

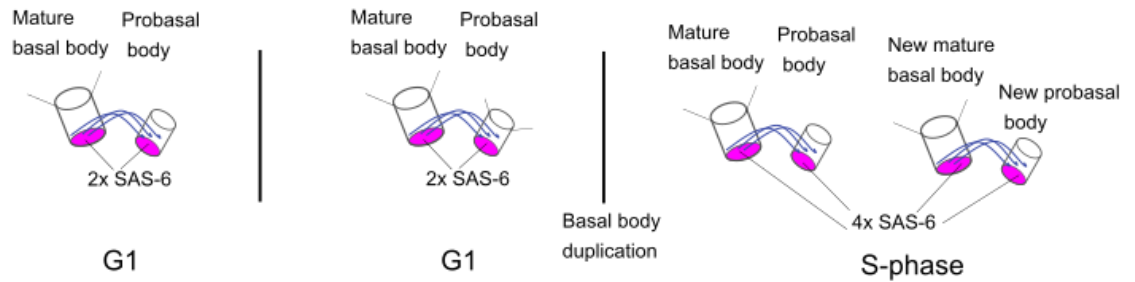
4.2 Dissecting timing of transition fibre assembly

4.2.1 Transition fibre recruitment in relation to SAS6 duplication

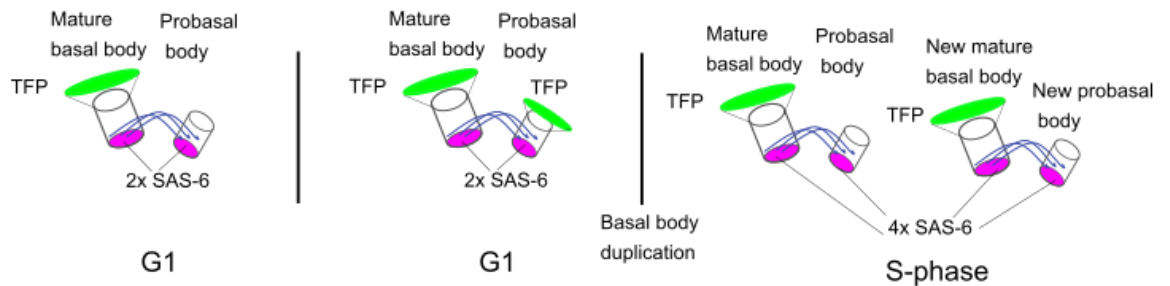
It has been shown that in mammals, transition fibre proteins are recruited to the basal body in a hierarchical manner as the cilium is assembled (Tanos et al., 2013). We have already elucidated the spatial arrangement of transition fibre proteins in Chapter 3. We now aim to better understand the assembly of transition fibres based on the timing of transition fibre protein recruitment to the basal body. To achieve this, a cellular event that can be used as a temporal marker is required. We used the timing of basal body duplication (SAS6) at the G1/S-phase border as a temporal marker as SAS6 represents one of the earliest points in basal body biogenesis as it forms the cartwheel stack at the proximal end of the basal body (Hu et al., 2015, Nakazawa et al., 2007).

Basal body duplication can easily be determined by the doubling of the SAS6 signal from two to four foci (Fig 4.7a). Cell lines expressing SAS6::mScarlet (magenta) and a transition fibre protein tagged with mNG (green) were used to assess whether proteins arrived before or after SAS6 duplication. If a second transition fibre foci is observed on the probasal where there are 2x SAS6 foci (prior to basal body duplication), it would be classed as arriving before basal body duplication (Fig. 4.7b). Where a second foci is only observed on the SAS6 of the mature basal body of cells with 4x SAS6 foci (post basal body duplication), the protein would be classed as arriving after basal body duplication (Fig. 4.7c).

A Basal body duplication occurs in S-phase to form a new pair of basal bodies



B Arrival to the probasal body before basal body duplication in S-phase



C Arrival to the new mature basal body after basal body duplication in S-phase

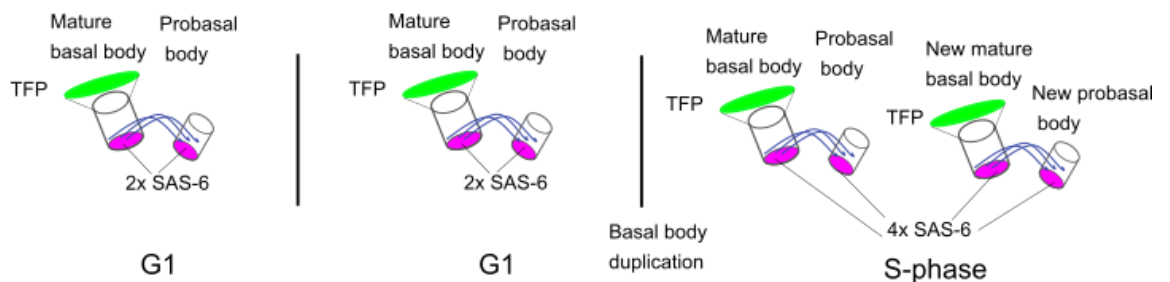


Figure 4.7 (A) Diagram showing basal body duplication event during S-phase which is represented by the SAS-6 at the base of the basal body barrel (magenta) going from two to four foci (B) An early arriving transition fibre protein (TFP) (green) arriving at the probasal body prior to basal body duplication where SAS-6 goes from two to four foci (magenta). (C) A later arriving transition fibre protein (TFP) (green) arriving to the new mature basal body after basal body duplication represented by SAS-6 going from two to four foci.

FBF1 is an example of a transition fibre protein that was classed as arriving prior to basal body duplication (Fig. 4.8a). In early G1, a single FBF1 foci can be observed on the mature basal body (Fig. 4.8a). In a proportion of G1 cells, prior to basal body duplication (two SAS6 foci), a second FBF1 foci can be observed on the maturing probasal body (P) (Fig. 4.8a). By S-phase, when SAS6 has duplicated from two foci to four foci, two FBF1 foci can be observed on each of the mature basal bodies (M) of the old basal body pair and the newly duplicated basal body pair (Fig. 4.8a).

Where a second foci was not observed on the probasal body prior to basal body duplication (two SAS6 foci), the protein would be classed as arriving after basal body duplication. For example,

TbCEP164A did not exhibit a signal on the maturing probasal body prior to S-phase, confirming that recruitment of TbCEP164A occurs after basal body duplication (Fig. 4.8b). In G1 stage of Figure 4.8b, a single CEP164A foci can be observed at the distal end of the mature basal body. However, a second foci is never observed on the probasal body in G1 cells. It is only in S-phase where there are four SAS6 foci that a second CEP164A signal can be observed on the mature basal body of the new and old basal body pair (Fig. 4.8b).

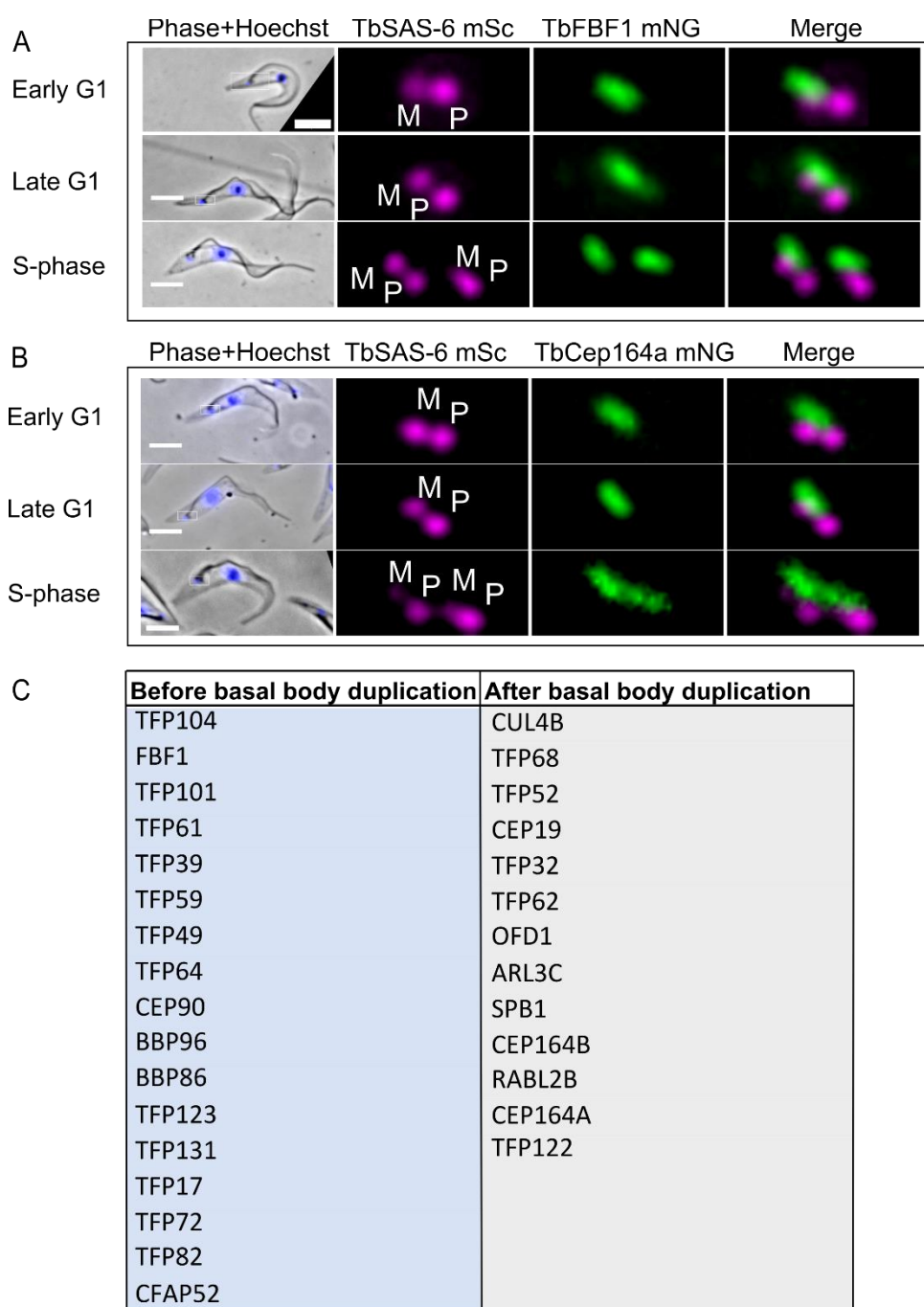


Figure 4.8 (A) Methanol-fixed cytoskeletons of FBF1 tagged with the mNeonGreen fluorescent protein co-expressed with SAS-6 tagged with mScarlet as an example protein that arrives to the basal body before SAS-6 duplication (B) Methanol-fixed cytoskeletons of CEP164A tagged with mNeonGreen co-expressed with SAS-6mScarlet as an example protein that

arrives after basal body duplication. Nucleus and kinetoplast (blue) are stained with Hoechst. Mature and probasal body are labelled M and P respectively. All images are taken at 100x magnification. (C) Table of 31 transition fibre proteins taken forward and the timing of their arrival to the basal body relative to SAS-6 duplication. 17 early transition fibre proteins that arrive before SAS-6 duplication (left, blue). 14 late transition fibre proteins arrive after SAS-6 duplication (right, grey).

This analysis was carried out on all 31 transition fibre proteins that were taken forward. Out of 31 transition fibre proteins, 17 transition fibre proteins (TFP104, FBF1, TFP101, TFP61, TFP39, TFP59, TFP49, TFP64, CEP90, BBP96, BBP86, TFP123, TFP131, TFP17, TFP72, TFP82 and CFAP52) were found to have a signal present on the maturing probasal body and therefore were classed as arriving before basal body duplication (Fig. 4.8c). This suggests that these are the early recruited proteins to the transition fibre of the newly maturing basal body. For the remaining 14 transition fibre proteins, a signal was never observed on the maturing probasal body of G1 cells and therefore we consider these as proteins that were recruited to the mature basal body of the new flagellum after basal body duplication. These 'late' transition fibre proteins are CEP164A, TFP68, TFP52, CEP19L, TFP32, TFP62, Ofd1, ARL3C, SPB1, CEP164b, RabL2b, Cullin4b, TFP15 and TFP122 (Fig. 4.8c). The data is consistent with the transition fibres having some hierarchical assembly where some proteins are recruited earlier and some later. We hypothesise that the proteins arriving earlier are involved in a more structural role while those arriving late may be more functional (for example IFT loading or basal body docking).

4.2.2 Transition fibre recruitment in relation to the arrival of CEP164A

Following this analysis, we aimed to further refine our understanding of when each protein arrives at the transition fibre by using the arrival of a transition fibre protein as a second temporal marker. CEP164A was considered a suitable marker for previous analysis, and so we decided to assess whether proteins arrive before or after CEP164A. This analysis was carried out on 21 transition fibre proteins that were tagged with mNG in a cell line expressing CEP164A::mScarlet. Cells in early S-phase with recently duplicated transition fibre signals were analysed. If a second mNG foci was observed at the new flagellum only accompanied by a single CEP164A foci at the old flagellum, this protein would be classed as arriving to the new mature basal body before CEP164A. However, if a second CEP164A foci is observed at the new flagellum but is not accompanied by an mNG signal at the new flagellum, this protein would be classed as arriving after CEP164A. Figure 4.9a shows an example of a transition fibre protein (TbFBF1) arriving before CEP164A. In G1 cells, there is a single CEP164A foci overlapping with the FBFmNG signal. In Early S-phase, the FBF1 has duplicated and now has a second foci on the basal body of the new flagellum (Fig. 4.9a, green, orange arrow). However, CEP164A has not yet been duplicated as there is only one foci present at the old flagellum

(Fig. 4.9a, magenta, white arrow) and we therefore conclude that FBF1 arrives at the transition fibres before CEP164A.

Figure 4.9b shows the TFP122mNG and CEP164AmSc signals in G1 and S-phase. In G1 cells, a single CEP164A foci can be observed overlapping with a single TFP122 foci at the base of the flagellum (Fig. 4.9b). In S-phase, a second CEP164A foci appears at the new flagellum (Fig. 4.9b, magenta, orange arrow). However, TFP122 only has a single foci at the old flagellum (Fig. 4.9b, green, white arrow). Later in S-phase, there is a TFP122 and CEP164A foci at the new and old flagellum. From this, we reason that TFP122 arrives after CEP164A.

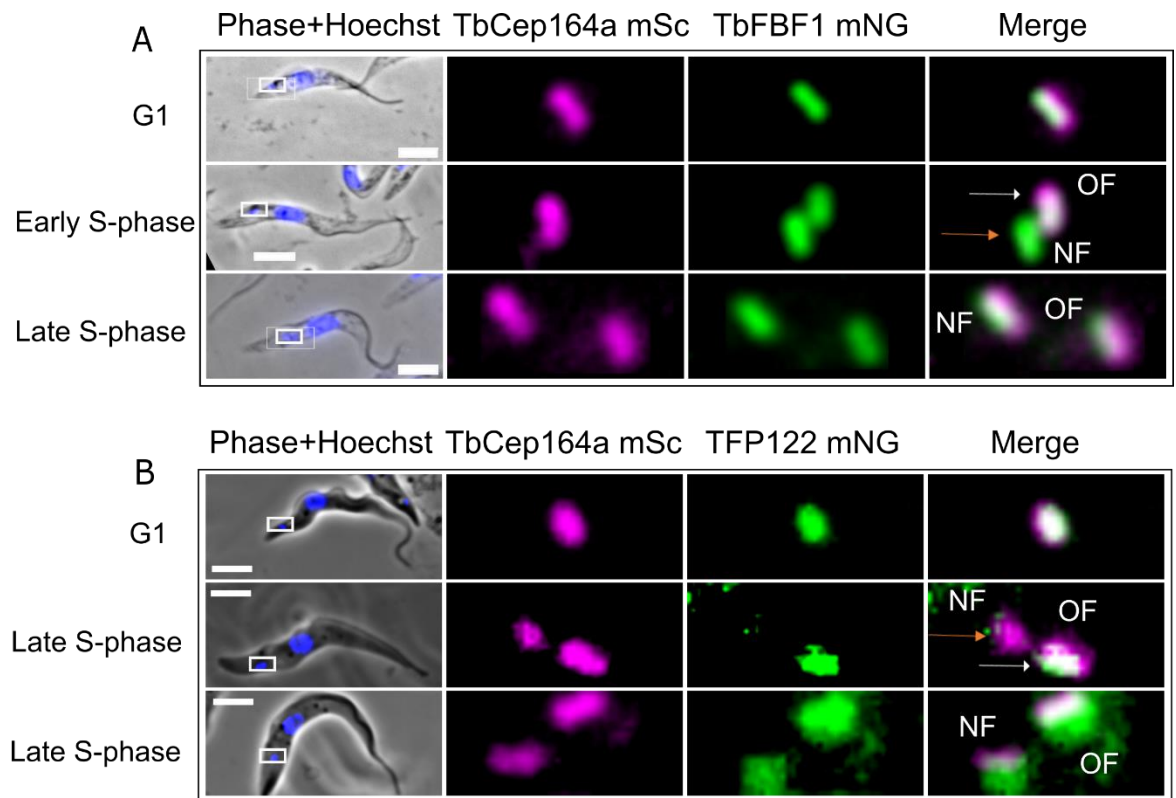


Figure 4.9 (A) Methanol-fixed cytoskeletons of FBF1 tagged with the mNeonGreen fluorescent protein co-expressed with CEP164A tagged with mScarlet as an example protein that arrives to the basal body before SAS-6 duplication (B) Live cell images of FBF1 tagged with mNeonGreen co-expressed with CEP164AmScarlet as an example protein that arrives after basal body duplication. Nucleus and kinetoplast (blue) are stained with Hoechst. All images are taken at 100x magnification using a Zeiss wide-field microscope.

A total of 22 proteins were successfully tagged with CEP164A. ARL3C had an mNG signal that was too weak to carry out this analysis reliably. Therefore, this analysis was carried out on 21 transition fibre proteins. The signal CEP164A arrives to the basal body after basal body duplication so we would expect all proteins that arrive before basal body duplication to also arrive before CEP164A. This was found to be the case as all proteins except for TFP122 were found to arrive before CEP164A.

In summary, we have determined the order of recruitment of transition fibre components based on the timing of basal body duplication and CEP164A recruitment. Figure 4.10 summarises this data. We show that there are 17 early arriving transition fibre proteins that are recruited to the basal body prior to basal body duplication (Fig. 4.10). A total of seven proteins (TFP68, TFP52, CEP191, TFP32, TFP62, Ofd1 and CEP164b) arrive to the basal body after basal body duplication but are recruited before CEP164A arrives to the basal body. These are classed as 'later' proteins. We determined CEP164A arrives after these seven proteins (Fig. 4.10; magenta ring). This is consistent with CEP164 being the last protein to arrive to the basal body in mammalian distal appendage assembly (Tanos et al., 2013, Yang et al., 2018).

Lastly, we determined that TFP122 is the latest protein to be recruited to the transition fibres as it arrives after basal body duplication and CEP164A (Fig. 4.10). For the remaining seven 'later' proteins that arrived to the basal body after SAS6 duplication, 6 were unable to be tagged with CEP164 (BBP86, Cullin4b, Rabl2b, TFP82, SPB1, TFP15) and 1 (ARL3C) did not have a sufficiently bright mNG in enough cells signal to carry out this analysis reliably. For these seven proteins, we determine that they are later proteins and arrive at some point after basal body duplication but it could not be determined whether they arrive before or after CEP164A (Fig. 4.10).

The data shows that transition fibre assembly is a highly regulated process that spans from G1 to S-phase of the cell cycle, with proteins arriving at the transition fibres at different points during this time. Transition fibre assembly begins in G1, with a number of proteins arriving to the probasal body in the early stages of its maturation but before basal body duplication. By S-phase, following basal body duplication, several proteins arrive to the newly matured basal body (nMBB, Fig. 4.10) and assembly continues until CEP164A is recruited to the basal body later in S-phase. Following this event, our data show that TFP122 is the only protein to arrive. We hypothesise TFP122 is the most downstream protein in a chain of regulatory events at the transition fibres.

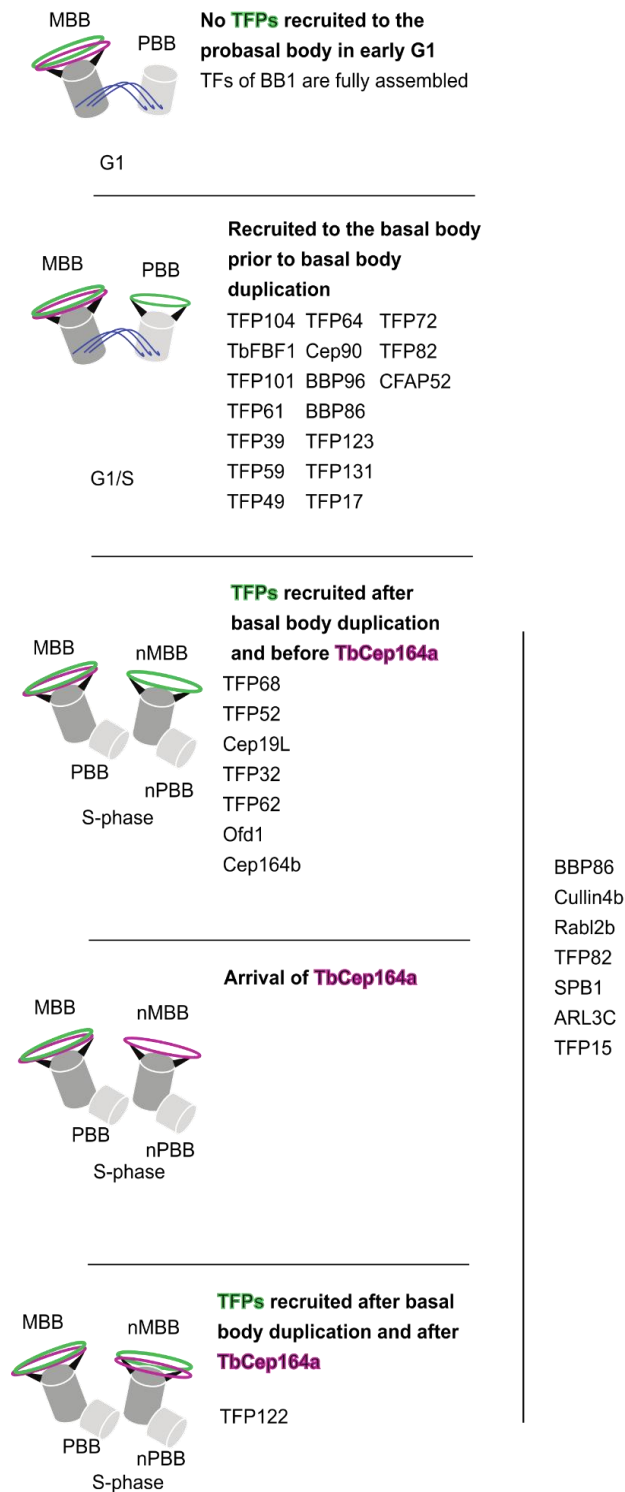


Figure 4.10 Model representing the arrival of transition fibre components based on the timing of basal body duplication and CEP164A arrival in S-phase. Transition fibre proteins are represented with a green circle around the top of the basal body and CEP164A is represented by a magenta circle. Early proteins (top, green circle) are shown arriving to the probasal body (PBB) (light grey) prior to basal body duplication and CEP164A. Late proteins are shown arriving to the new mature basal body (nMBB) (dark grey) which was formed from the maturing of the original probasal body (PBB) following basal body duplication. TFP122 is shown arriving at the newly duplicated mature basal body (nMBB) after basal body duplication and CEP164A. The new probasal body is labelled nPBB.

4.3 Evolutionary analysis of trypanosome transition fibres by reciprocal best blast

Out of 34 putative transition fibre proteins initially taken forward, 31 of these were considered transition fibre proteins based on their diameter from a lateral view (pseudo-diameter), distance from the mature basal body and/or co-localisation with a known transition fibre marker. A total of 31 proteins were carried forward for further analysis.

We aimed to examine the conservation of these transition fibre components across a range of ciliated and non-ciliated organisms. Reciprocal best blast is a method used to rapidly identify orthologs in different organisms. Briefly, the amino acid sequence of the trypanosome protein of interest is searched into the proteome of each organism of interest and where the best scoring protein is a top hit in the reciprocal search into the trypanosome proteome, we consider this protein an orthologue **if the E-value is lower than 1×10^{-5}** (see Methods 2.5). This analysis was carried out in 45 ciliated and nonciliated eukaryotic proteomes and orthology of transition fibre proteins was later refined by literature searching. Reciprocal best blast analysis was carried out on 30 transition fibre proteins (CEP164A and CEP164B have one orthologue in most other organisms and so it was considered as one protein for this analysis).

4.3.1 Widely conserved transition fibre proteins

A reciprocal best blast carried out on all transition fibre proteins revealed a total of 9 transition fibre proteins that are widely conserved across eukaryotes. These proteins are CFAP52, Rabl2b, CEP164, CEP19L, Ofd1, CEP90, FBF1, TFP123 and TFP101 (Fig. 4.11). Of these, 6 proteins (CFAP52, Rabl2b, CEP164, Ofd1, CEP90 and FBF1) are conserved in Metazoa. All 6 proteins conserved in Metazoa are present in Humans, Chickens (*Gallus gallus*) and Zebrafish (*Danio rerio*). However, of these 6, only CFAP52 and CEP164 are present in insects such as the honey bee (*Apis mellifera*) and the fruit fly (*Drosophila melanogaster*). All but one of the widely conserved proteins (FBF1) are absent from the *C. elegans* proteome. This indicates that *C. elegans* may have evolved a unique appendage structure that is divergent in its composition.

The most widely conserved protein is CFAP52 (Cilia and flagella associated protein 52). This protein has orthologues in plants, Chromista, Metazoa, fungi and Excavata (Fig. 4.11). This protein has been implicated in motile cilia and was shown to have orthologues humans, *C. reinhardtii* and *D. melanogaster* which is in concordant with our analysis of CFAP52 (Baron et al., 2007). CFAP52 has also been examined in the context of mammalian sperm and has been localised to the mature and probasal body at the base of the sperm flagellum which conflicts with our localisation of mature basal body transition fibres (Tapia Contreras and Hoyer-Fender, 2020). Interestingly, CFAP52 has orthologues in the plant species *Psycomitrella patens*, *Chlamydomonas reinhardtii* and

Ostreococcus tauri, the latter of which does not assemble a cilium or flagellum in any stage of its life cycle (Chretiennot-Dinet et al. 1995). CFAP52 is 629 amino acids in length and contains multiple WD40 repeats between amino acids 109 – 620. The WD40 repeat protein family is found widely across eukaryotes and is known to function in cytoskeletal rearrangement and vesicular trafficking (Smith et al., 1999). CFAP52 stands as an interesting candidate for functional analysis based on its widespread conservation.

Another well conserved protein is Rabl2b, which is present across a range of species in Excavata and Metazoa however appears not to be found in insects such as *Drosophila melanogaster*, *Apis Mellifera* (honey bee) and *C. elegans*. Rabl2b is a relatively small protein of 199 amino acids and contains a Rab-GTPase-like family domain that spans most of the protein. It contains the switch II and switch I domains required for GTP hydrolysis and subsequent GDP stability, indicating Rabl2b is a functioning GTPase at the ciliary base. Based on this, Rabl2b also stands as a good candidate for functional analysis.

CEP164 is widely conserved across ciliated eukaryotes (Hodges et al., 2010). CEP164A and CEP164b in Trypanosomes are of similar amino acid length to each other (1177 and 1006 respectively) and both contain an N terminal WW domain. WW domains are known to be involved in organisation and assembly of cellular structures. Functional analysis by RNAi however has shown that they have no discernible phenotype (Jiri Tyc, Vaughan lab, personal communication).

CEP19L protein is conserved in well characterised ciliated species such as *Tetrahymena thermophila* and *Chlamydomonas reinhardtii*. The CEP19L protein is not conserved in most Metazoans, however an orthologue is present in *Trichoplax adherans*. The CEP19 protein found in mammals and other ciliated species exhibits a similar localisation to the CEP19L protein in Trypanosomes, however it is not similar enough at the sequence level to be considered an orthologue by reciprocal best blast. CEP19L protein is 383 amino acids in length and possesses a CEP19-like protein family domain spanning from amino acids 186-380.

The Oral-facial digital syndrome protein 1 (Ofd1) is conserved in *Homo sapiens*, *Gallus gallus* and *Danio rerio* but is absent in insects such as *Drosophila melanogaster*, *Apis mellifera* and *C. elegans*. An orthologue of Ofd1 is present in the fungal species *Batrachotrium dendrobatis* but is absent in other organisms outside Excavata. Mutations in Ofd1 are associated with Oral-facial digital syndrome in humans. The TbOfd1 protein is 785 amino acids in length and contains a domain of unknown function (DUF5401) between amino acids 227- 406.

The CEP90 protein is a well characterised conserved centriolar component that we have identified as part of the transition fibres in Trypanosomes. Orthologues of CEP90 were only identified in *Homo*

sapiens, *Gallus gallus* and *Danio rerio* outside oChlaf the Excavata. TbCEP90 is 608 amino acids in length and contains a SMC (structural maintenance of chromosomes) across the N-terminal region of the protein between amino acids 68-333. FBF1 is also conserved in *Homo sapiens*, *Gallus gallus* and *Danio rerio*, however FBF1 also has an orthologue in *C. elegans*. FBF1 is a protein of 797 amino acids in length and also contains an SMC superfamily domain between amino acids 426 and 753. The SMC domain is known to be involved in chromosome partitioning and cell division.

TFP123 is only conserved in *Aureococcus anophagefferens* outside of the Excavata and is a larger protein of 1147 amino acids and contains a dermokine superfamily domain between amino acids 80-320. Lastly, TFP101 is only conserved in *Leishmania Mexicana* and *Plasmodium falciparum* out of all the eukaryotic organisms shown in Figure 4.11. TFP101 is the only transition fibre protein in this cohort that is conserved in *Plasmodium* spp. TFP101 is 865 amino acids in length and has a SMC prok B domain between amino acid 94 and amino acid 697. SMC prok B represents the structural maintenance of chromosome common bacterial types.

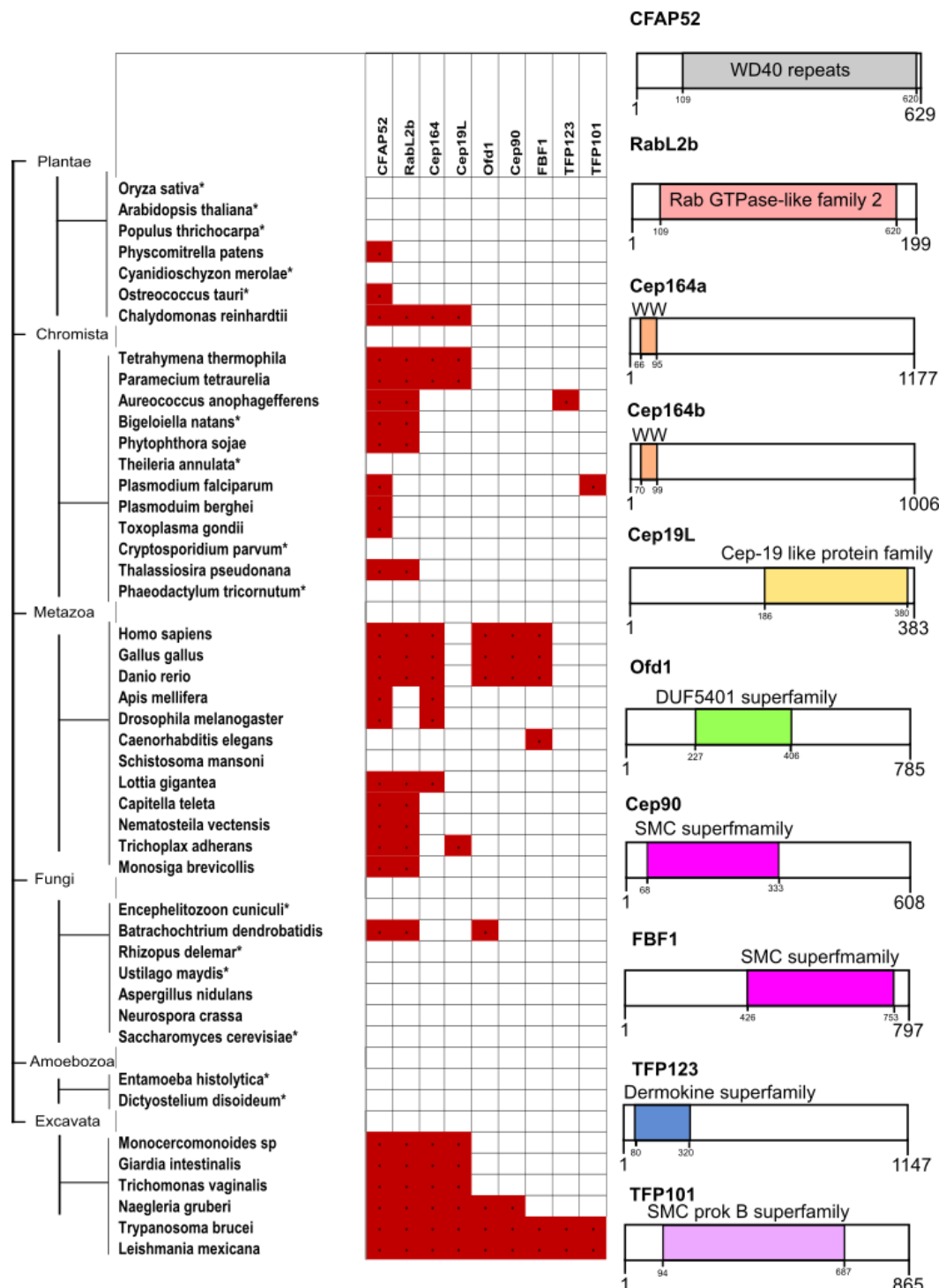


Figure 4.11 Conservation pattern 9 transition fibre proteins that are widely conserved across eukaryotic organisms (left). 45 model eukaryote proteomes were used for analysis and organisms with an asterisk indicate those without cilia/flagella. Those without an asterisk indicate organisms that possess cilia/flagella. Schematic of protein domains (right) searched into the NCBI database. Domains were represented only where $E < 1 \times 10^{-5}$. Amino acid position of each domain is indicated in smaller font and the total amino acid length of the protein is indicated in larger font.

4.3.2 Transition fibre proteins only conserved within the excavata super group

Out of 16 proteins conserved in Excavata, 15 proteins (TFP122, TFP131, TFP72, TFP61, TFP62, TFP59, TFP64, Cullin4b, BBP96, BBP86, TFP68, ARL3C, TFP49, SPB1, TFP15) are only found in

Leishmania Mexicana and *Trypanosoma brucei* out of the eukaryotic organisms in Figure 4.12. One protein (TFP82) is present in *Leishmania Mexicana*, *Trypanosoma brucei* and *Naegleria gruberi* (Fig. 4.12).

TFP82 is 750 amino acids in length and contains multiple Leucine rich repeat domains (LRR) between amino acids 111 to 341. TFP122 is 1085 amino acids in length and contains an N-terminal motor domain (2-360aa) and a Plant-like phospholipase C plekstrin homology (PLC PH) domain at the C terminal region (831-960 aa). TFP131 has a total length of 1179 amino acids and contains a SMC prok B domain between amino acids 391 and 952. TFP72 is smaller in length (653 aa) and contains a SMC domain between amino acids 139 and 320. TFP64 contains an SMC prok B domain (36-370 aa) and a DUF (domain of unknown function) between amino acids 239 and 511. TFP64 contains the same DUF5401 domain that was identified in TbOfd1. Cullin 4b is a protein of 947 amino acids in length and contains a Cullin family domain across the mid region of the protein and a Cullin neddylation domain at the C-terminus between amino acids 879 and 943. TFP68 has a total length of 577 amino acids and contains a MAEBL domain between amino acids 97 and 563. ARL3C is a smaller protein of 183 amino acids and contains a P-loop NTPase superfamily spanning from amino acid 1 to 180. Lastly, SPB1 is 325 amino acids in length and contains a Phosphatidylinositol phosphate kinase domain from amino acid 43 to amino acid 285. The remaining 6 proteins (BBP86, TFP61, TFP62, TFP59, BBP96, BBP86 and TFP15) produced no significant domain hits when the amino acids sequences were searched into NCBI, therefore the schematic for these five proteins was not included in Figure 4.12.

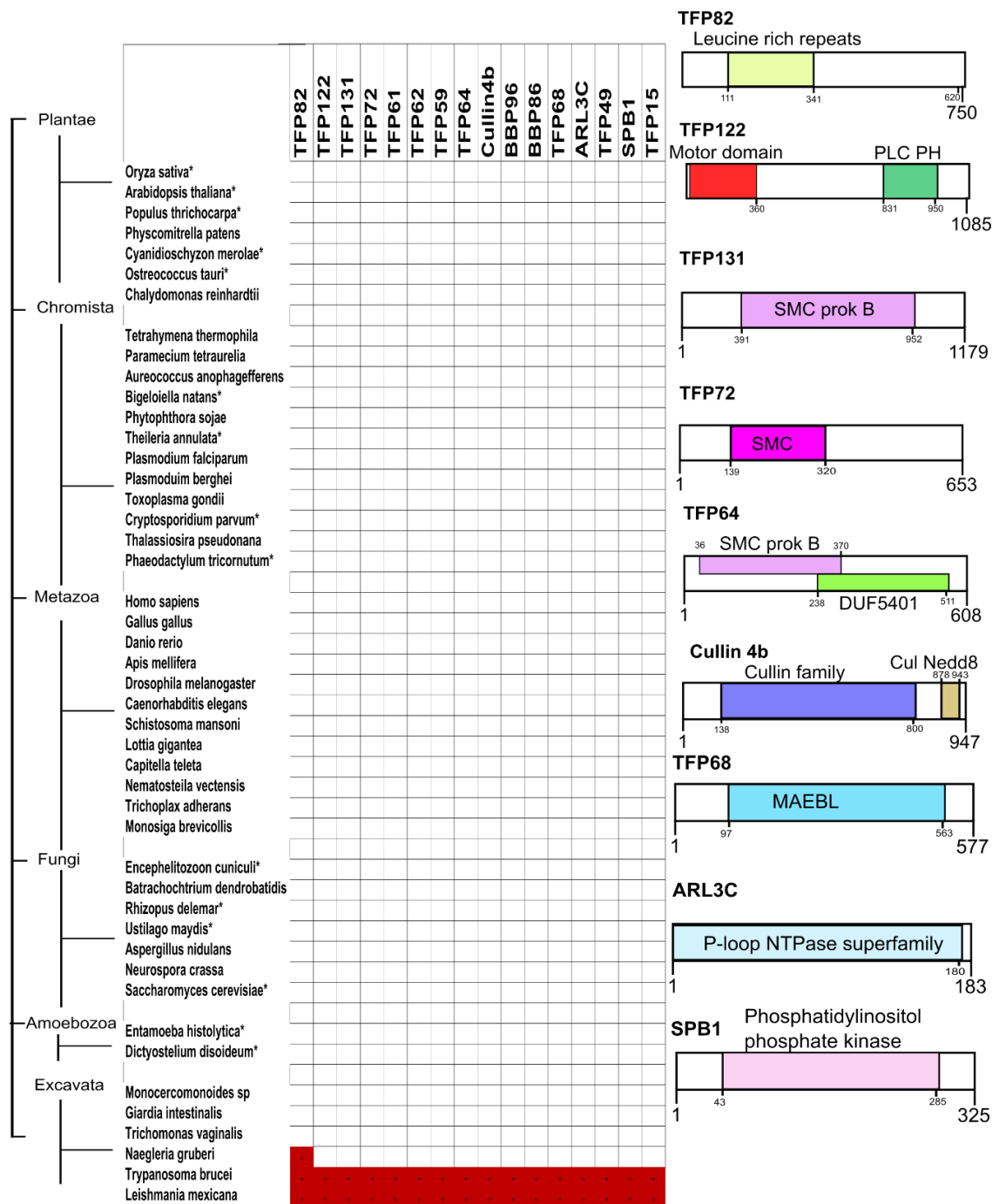


Figure 4.12 Conservation pattern of 16 transition fibre proteins that are only conserved within excavata (left). 45 model eukaryote proteomes were used for analysis and organisms with an asterisk indicate those without cilia/flagella. Those without indicate organisms that possess cilia/flagella. Schematic of protein domains (right) searched into the NCBI database. Domains considered only where $E < 1 \times 10^{-5}$. Amino acid position of each domain is indicated in smaller font and the total amino acid length of the protein is indicated in larger font.

4.3.3 Trypanosome specific transition fibre proteins

Of the 31 transition fibre proteins analysed in this study, five proteins (TFP104, TFP17, TFP52, TFP39, TFP32) had no orthologues present in any of the eukaryotic organisms in Figure 4.13. This suggests these proteins are trypanosomes specific. Searching the amino acid sequence of all five

trypanosome specific proteins had produced no significant domain hits.

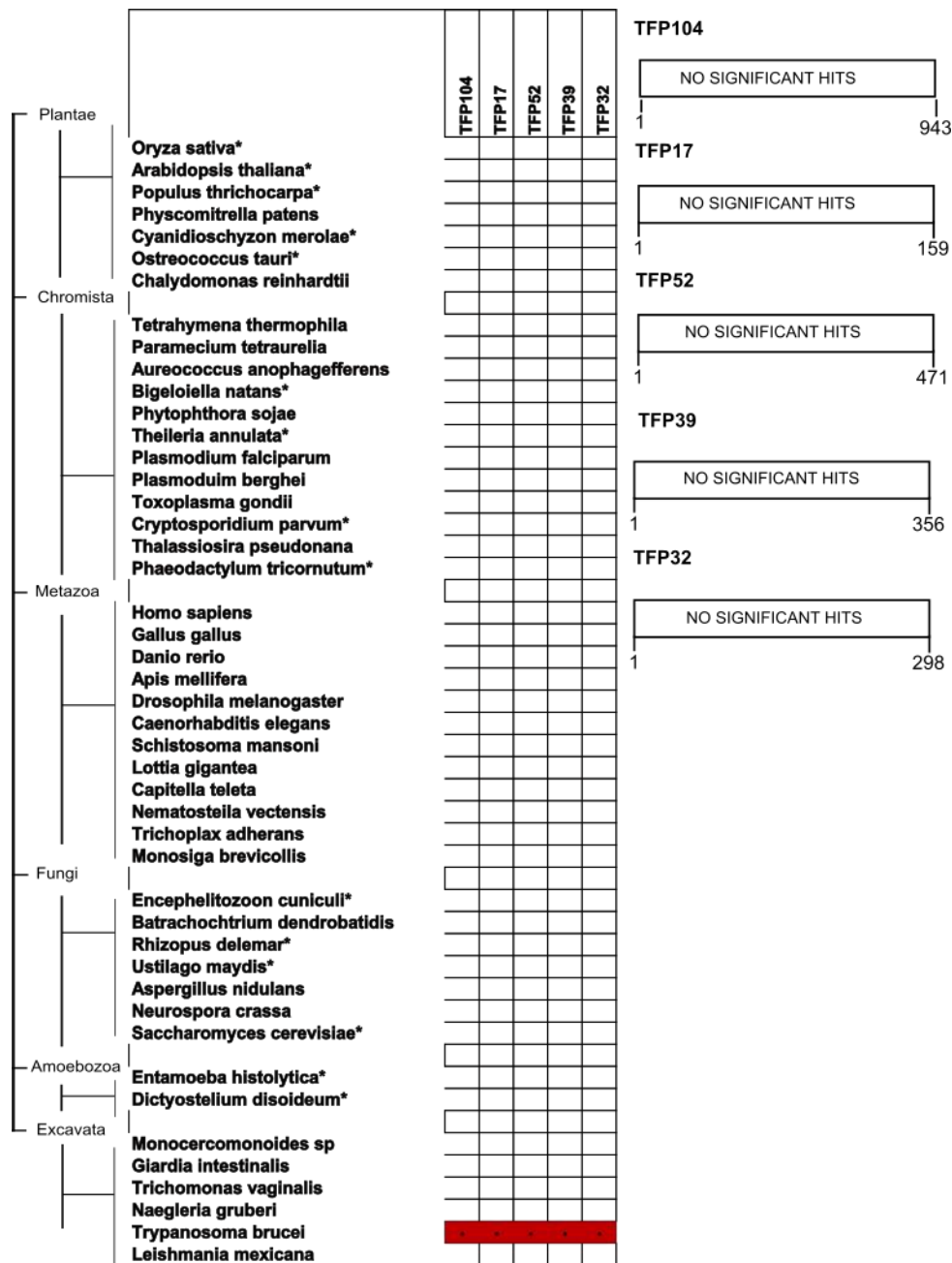


Figure 4.13 Five transition fibre proteins that are only present in *Trypanosoma brucei* out of all 45 eukaryotic organisms (left). 45 model eukaryote proteomes were used for analysis and organisms with an asterisk indicate those without cilia/flagella. Those without indicate organisms that possess cilia/flagella. Schematic of protein domains (right) searched into the NCBI database. Domains considered only where $E < 1 \times 10^{-5}$. Amino acid position of each domain is indicated in smaller font and the total amino acid length of the protein is indicated in larger font.

In summary, the conservation of 31 trypanosome transition fibre proteins was assessed in 45 ciliated and nonciliated eukaryotic organisms using reciprocal best blast. A total of 9 proteins were conserved outside of the Excavata, including seven orthologues in metazoan organisms and 6 orthologues present in humans but only a single transition fibre protein (FBF1) was found to have

an orthologue in *C. elegans*. These seven proteins all contained protein domains. CFAP52 contained a WD40 domain which is associated with cytoskeletal arrangement and vesicular trafficking. Rabl2b is a small GTPases which are known to act as key molecular switches in the context of ciliary assembly and function (Li and Hu, 2011). Ofd1, FBF1 and CEP90 are known centriolar proteins that have mammalian orthologs and the CEP19L protein shares sequence similarity with the known centriolar protein CEP19. The protein domains, literature and conservation of these proteins makes them attractive candidates for functional analysis.

Of all the proteins conserved only in excavata, a total of 15 proteins were found to have orthologues in kinetoplastids with 1 protein (TFP82) also conserved in *Naegleria gruberi*. Of these, 9 proteins contained conserved protein domains. Notably, TFP122 contained a motor domain at the N-terminus and a pleckstrin homology domain at the C-terminus while Cullin4b contained a Cullin family domain associated with ubiquitination. These two proteins also make good candidates to assess transition fibre function. Of the remaining five proteins that are only conserved in trypanosomes, none produced a significant domain hit. Therefore, out of the trypanosome specific proteins, we selected TFP104 for functional analysis as it has a brighter signal at the old flagellum as well as TFP52 and TFP17. We have therefore determined a set of proteins that make good candidates for functional characterisation based on their conservation, known literature in other organisms and the presence of conserved protein domains.

5. Functional analysis of transition fibre proteins by RNAi

To understand the roles of transition fibres in flagellum assembly and length control, we took forward a set of transition fibre proteins forward for functional analysis. The CRISPR Cas-9 knockout system and inducible RNA interference (RNAi) are two molecular tools available in trypanosomes to study gene function. CRISPR-cas9 works by using a guide RNA to target the gene of interest and the endonuclease activity of the Cas9 enzyme to remove the gene of interest from the genome. One disadvantage of this technique is that genes essential to trypanosome survival cannot be assessed. As the flagellum is an integral to trypanosome biology, we predict many of these genes to be essential. Therefore, we decided to study transition fibre function using inducible RNA interference (RNAi). Depletion of a target protein can be induced by the addition of tetracycline, which in turn allows for the study of transition fibre components that are essential in trypanosomes. In the following chapter, we will examine the function of 20 transition fibre proteins in relation to flagellar assembly and length regulation in trypanosomes by using the inducible RNAi system.

5.1 RNAi plasmid generation

The inducible RNA interference (RNAi) system enables the knockdown of a gene of interest by degradation of the mRNA. Briefly, this system works by expressing dsRNA which is cleaved into shorter fragments by the endoribonuclease DICER. This results in the production of small interfering RNAs (siRNA) which enters the RNA induced silencing complex (RISC) where the dsRNA can be unwound. The antisense strand can then target the mRNA transcribed by the gene of interest which can be degraded by RISC. The RNAi vector for dsRNA expression contains a procyclin acidic repetitive protein promotor (PARP), a tetracycline operator and 2 x targeting sequences that sandwich a stuffer sequence that allows for the formation of the stem loop RNA (Figure 5.1). The addition of tetracycline induces the expression of the targeting sequence, which initiates gene silencing via the mechanism described above.



*Figure 5.1 Simple schematic of the vector used for dsRNA expression in *T. brucei*. PARP promotor (red) is controlled by the TET operator for induction of the RNAi. Gene targeting sequences (pink) flank a stuffer sequence (grey) generating the stem loop dsRNA.*

To generate the plasmid vector containing the features described in Figure 5.1, we used the pQuadra stem-loop vector (Inoue et al., 2005). This involved a four-part quadratic ligation of the pQuadra 1 plasmid, pQuadra 3 backbone and 2 x inserts containing the targeting sequence (x2). Plasmids were generated and subsequently confirmed by sequencing (see methods, Fig. 2.7.1). The plasmid construct was transfected (see methods, Fig. 2.7.2) into cell lines expressing the protein on interest tagged with mNeonGreen and CEP164A tagged with mScarlet where possible.

5.2 Selection of 20 transition fibre proteins for functional analysis

It was decided that a total of 20 transition fibre proteins was a suitable number to take forward for functional analysis. Proteins were selected for RNAi based on their conservation, localisation pattern, presence of conserved domains and literature searching. Widely conserved proteins were candidates for functional analysis with the exception of CEP164A, CEP164B and CEP164C as their function has already been studied in trypanosomes (Jiri Tyc, Vaughan lab) (Atkins et al., 2021).

A total of seven widely conserved proteins were taken forward (CFAP52, Rab12b, CEP19L, CEP90, OFD1, FBF1, TFP123). One of these proteins (Rab12b) also exhibited a cell cycle regulated localisation that was brighter at the old flagellum (Chapter 4, Fig. 4.1). We also took forward three more cell cycle regulated that are brighter on the old flagellum (TFP122, TFP104, TFP39) and one cell cycle regulated protein that was brighter on the new flagellum (TFP72). In addition, we selected 6 proteins conserved within excavata (Cullin 4b, TFP49, TFP62, ARL3C, SPB1 and TFP68) and 2 trypanosome specific proteins (TFP52, TFP17). Cullin4b and ARL3C were selected due to the presence of a conserved Cullin family domain and an ARL (ADP-ribosylation factor-like) protein domain. ARL domain containing proteins have been associated with cilia formation in previous studies (Li and Hu, 2011).

TFP15 was originally excluded as it was more proximal than POC5 (Chapter 3, Fig. 3.9b). However, later analysis showed TFP15 to co-localise with known transition fibre protein RP2. Interestingly, TFP15 exhibited the widest pseudo-diameter (Fig. 3.9b Chapter 3). For this reason, we decided to take this protein forward for functional analysis as the outer most protein of mammalian transition fibres, CEP164, plays an integral role in assembly basal body docking and ciliary assembly (Cajanek and Nigg, 2014, Graser et al., 2007, Schmidt et al., 2012, Tanos et al., 2013, Yang et al., 2018).

Several studies have shown that transition fibres are required for flagellum/cilium formation (Garcia-Gonzalo and Reiter, 2017, Graser et al., 2007, Reiter, 2012, Singla et al., 2010, Tanos et al., 2013, Wei et al., 2015, Wei et al., 2013). Therefore, we decided to assess effect of knockdown by RNAi of transition fibre proteins on assembly and length control of the flagella. This was achieved

by immunolabelling the flagellar axoneme (mAb25) (Dacheux et al., 2012) and measuring flagellum length in G1 cells and the old and new flagella of post mitotic cells. In addition to this, growth was also assessed up to 72h post induction of RNAi and where appropriate, basal body duplication and segregation was assessed immunostaining the basal body (BBA4) (Woods et al., 1989).

5.3 6 transition fibre proteins are required for flagellum assembly

A total of 6 proteins (TFP68, CEP90, OFD1, FBF1, CEP19L and Rabl2b) exhibited a flagellum growth defect with cells assembling significantly shorter flagella (Mann-Whitney U test, $p < 0.05$) accompanied by a growth defect and changes in cell morphology following depletion by RNAi. The data suggests these 6 transition fibre proteins are required for new flagellum assembly.

5.3.1 TFP68 is required for flagellum assembly and localisation of CEP164A

TFP68 is a trypanosome specific transition fibre protein that contains a MAEBL domain spanning across the mid-region of the protein (Chapter 2, Fig. 4.13). MAEBL proteins are associated with erythrocyte binding by merozoites of *Plasmodium* spp (Ocampo et al., 2004). TFP68 arrives to the basal body after basal body duplication in S-phase (Chapter 2, Fig. 4.8c). Literature searching for TFP68 revealed that this protein has only been identified in large scale genomic, transcriptomic and proteomic studies and no specific literature on the function of TFP68 exists.

We show for the first time that depletion of TFP68 by RNAi severely perturbed flagellum assembly. G1 flagella were significantly shorter 24 hours post RNAi induction (Mann-Whitney U test, $p = 1.13 \times 10^{-14}$) (Fig. 5.2a, Fig. 5.2b). 30% of G1 cells assembled a very short flagellum/axoneme that was between 0 and 1 μm in length and this distribution was mirrored in the new flagella (NF) of post mitotic cells ($n=100$, $p < 2.2 \times 10^{-16}$) (Fig. 5.2a, Fig. 5.2b). The old flagella of post mitotic cells were not significantly different to that of the uninduced cells (Mann-Whitney U test, $p = 0.29$). We can conclude from this that TFP68 is required for the assembly of the new flagellum but not for the maintenance of old flagellum (Fig. 5.2b).

In addition, the CEP164A signal was very weak or absent from the base of the flagellum at 24 hours post induction of RNAi (Fig. 5.2a). Induced cells also exhibited a severe growth defect at 72 hours compared to the uninduced population (Fig. 5.2c). The number of basal bodies were counted to assess whether basal body duplication was affected. In a G1 cell we normally observe one pair of basal bodies (1 mature and 1 probasal body) at the base of the flagellum. In a post mitotic cell, we would expect to have a total of four basal bodies in the cell, two at the old flagellum (1 mature basal body, 1 probasal body) and two at the new flagellum (1 mature basal body and 1 probasal body). The total number of basal bodies in each cell was not significantly

different from the uninduced population. However, at 24 hours post induction, 15% of cells had no basal body detectable where the new flagellum should be located and at the old flagellum, four basal bodies were observed in the same percentage of cells (Fig. 5.2d, n=100). Figure 5.2d shows that basal body duplication was not affected by depletion of TFP68. However, basal body segregation was affected as the basal bodies of the new flagellum were clustered together with those of the old flagellum in a proportion of cells. We reason that this is likely due to the inability of cells to assemble a new flagellum correctly as basal body segregation is likely driven by the assembly of the new axoneme (Vaughan and Gull, 2015). In conclusion, we have shown that TFP68 is an essential component of the transition fibres that is required for the localisation of CEP164A and for the formation of a new flagellum in trypanosomes.

TFP68 RNAi

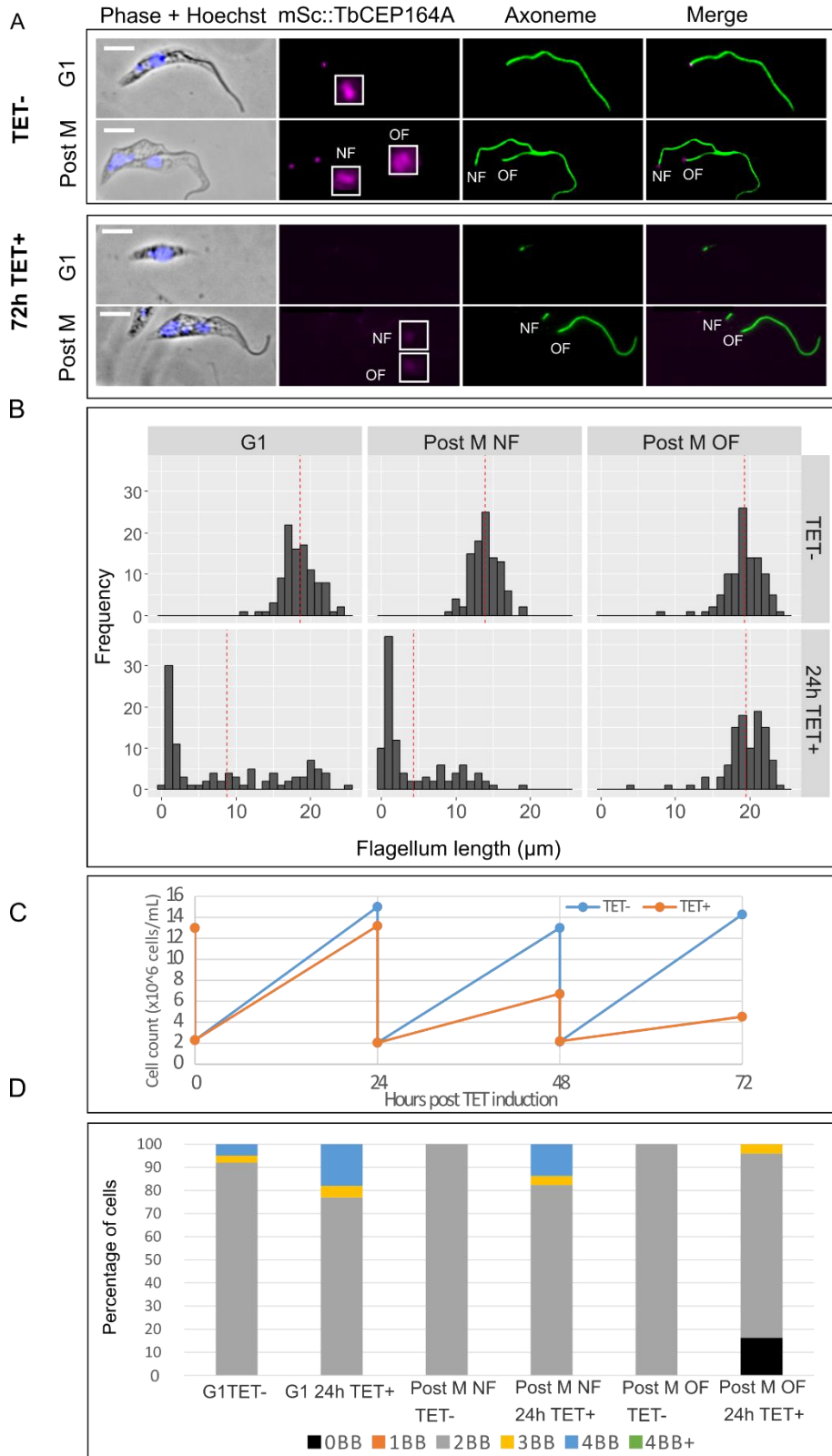


Figure 5.2 Depletion of TFP68 by RNAi. (a) Methanol-fixed cytoskeletons of uninduced cells that have not been depleted for TFP68 (TET-) and cells following 24 hours' induction of RNAi (24h TET+) in both G1 and post mitotic cells. DNA within the nucleus and kinetoplast is labelled with Hoechst. Insert shown within mSc::CEP164A (magenta) column represents a 5x zoom. Immunolabelling of the flagellar axoneme (green) with the mAb25 antibody. New flagellum and old flagellum

are labelled NF and OF respectively. Images taken at 100x magnification. Scale bar represents 5µm. (b) Length of flagellar axoneme (µm) based on mAb25 labelling comparing uninduced (TET-, top) and 24 hours induced (24h TET+, bottom) in G1 cells (n=100) (Mann-Whitney U test, $p=1.13e-14$), the new flagellum of post mitotic cells (Post M NF) (n=100) (Mann-Whitney U test, $p<2.2e-16$) and old flagella of post mitotic cells (post M OF) (n=100) (Mann-Whitney U test, $p=0.29$). (c) Growth over 72 hours in cells induced for RNAi (TET+) and uninduced (TET-) cells. Cells were cultured back to 2×10^6 cells/mL every 24 hours. (d) Basal body count in cell in induced (TET+) and uninduced (TET-) cells at G1 and post M phase of the cell cycle. Basal body number (BB) is indicated by colour (n=100).

5.3.2 Ciliopathy protein CEP90 is required for flagellum assembly and localisation of CEP164A

CEP90 is a widely conserved centriolar protein that contains an SMC superfamily domain at its N-terminal region (Chapter 4, Fig. 4.11). Studies have shown CEP90 to localise to centriolar satellites in mammals (Kim et al., 2012). However, more recently CEP90 has been described as part of a ciliopathy complex required for distal appendage formation (Kumar et al., 2021). Recessive mutations in the CEP90 gene in humans (also known as PIBF1) have been associated with ciliopathies Jeune syndrome and Joubert syndrome (Shen et al., 2020, Wheway et al., 2015). Given the link to human disease, CEP90 was taken forward for functional analysis.

Following depletion of the trypanosome CEP90 protein by RNAi, flagella assembly was severely affected by 48 hours post induction of RNAi. The new flagellum of post mitotic (post M) cells was most severely affected with 98% of the new flagella in post M cells having a length between 0-1 µm. Flagella were between 0-1µm in 85% of G1 cells and 80% of the old flagella in post M cells (Fig. 5.3a, Fig. 5.3b). The significant change in the length of old flagella after 48 hours suggests that either old flagella are remodelling to become shorter or as with other examples, the inheritance of the shorter new flagellum over several cell cycles leaves cells with a significantly shorter old flagellum.

Like TFP68, CEP90 is also required for the localisation of CEP164A. CEP164A signal was lost in all CEP90 depleted cells by 48 hours. This suggests that CEP90 is also required for the correct assembly of transition fibres. A severe growth defect was also observed by 72 hours post induction of RNAi (Fig. 5.3c). The data demonstrates that like TFP68, CEP90 is essential for new flagellum assembly and for the localisation of conserved transition fibre component CEP164A. We group TFP68 and CEP90 together based on the rapid and severe effect on growth, flagellum assembly and the absence of CEP164A at the base of the flagellum following knockdown.

TbCep90 RNAi

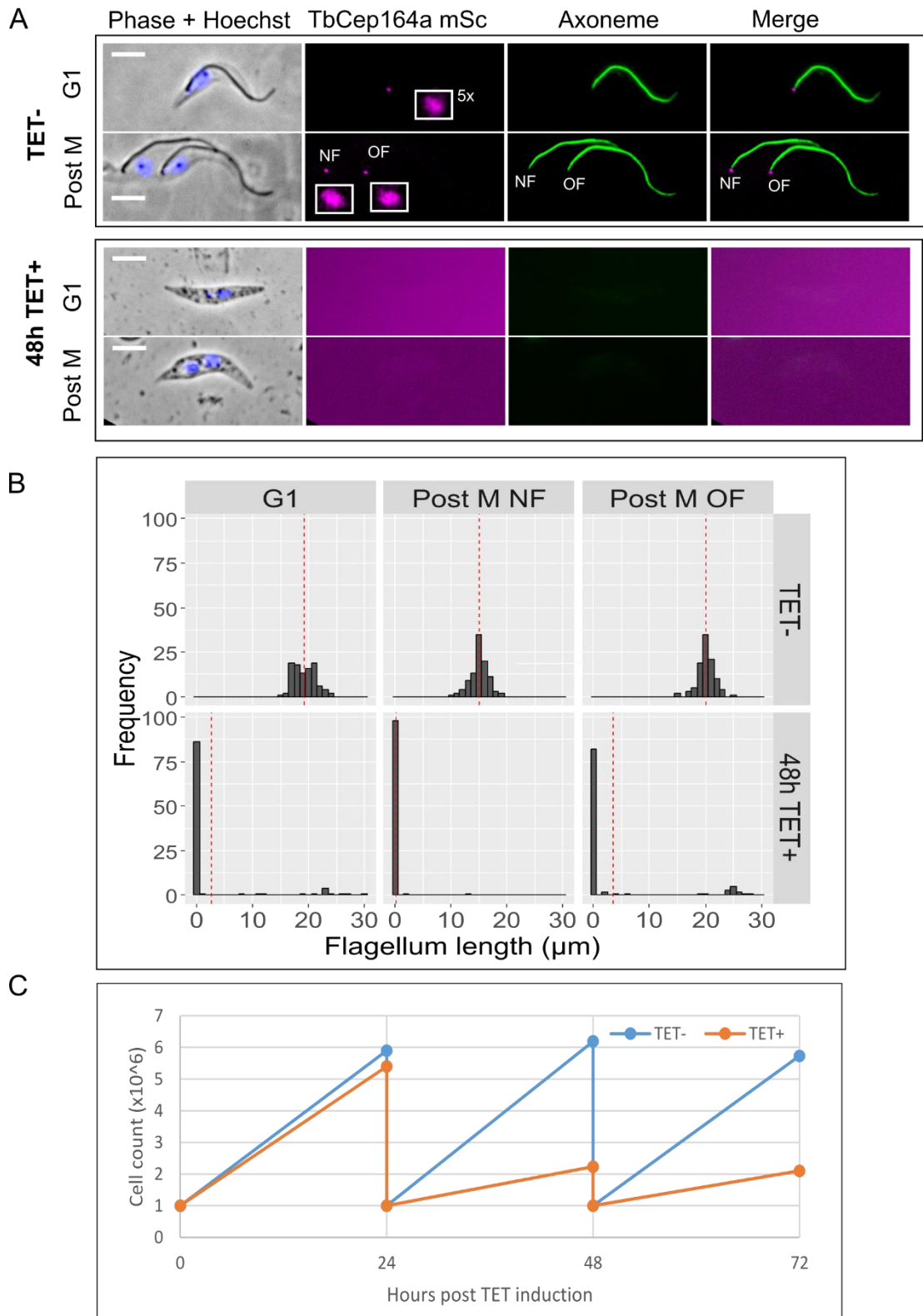


Figure 5.3 Depletion of CEP90 by RNAi. (a) Methanol-fixed cytoskeletons of uninduced cells that have not been depleted for CEP90 (TET-) and cells following 48 hours' induction of RNAi (48h TET+) in both G1 and post mitotic cells. DNA within the nucleus and kinetoplast is labelled with Hoechst. Insert shown within mSc::CEP164A (magenta) column represents a 5x zoom. Immunolabelling of the flagellar axoneme (green) was carried out using the mAb25 antibody. New flagellum

and old flagellum are labelled NF and OF respectively. Images taken at 100x magnification. Scale bar represents 5µm. (b) Length of flagellar axoneme (µm) based on mAb25 labelling comparing uninduced (TET-, top) and 48 hours induced (48h TET+, bottom) in G1 cells (n=100) (Mann-Whitney U test, $p<2.2e-16$), the new flagellum of post mitotic cells (Post M NF) (n=100) (Mann-Whitney U test, $p<2.2e-16$) and old flagella of post mitotic cells (post M OF) (n=100) (Mann-Whitney U test, $p<2.2e-16$). (c) Growth over 72 hours in cells induced for RNAi (TET+) and uninduced (TET-) cells. Cells were cultured back to 2×10^6 cells/mL every 24 hours. (d) Basal body count in cell in induced (TET+) and uninduced (TET-) cells at G1 and post M phase of the cell cycle.

5.3.3 Rabl2b is required for flagellum assembly but not localisation of CEP164A

Rabl2b is a widely conserved small GTPase that is required for ciliary biogenesis in humans via the binding of IFT complex B at the ciliary base (Kanie et al., 2017, Nishijima et al., 2017).

Following depletion of RabL2b by RNAi, we found the flagella were significantly shorter following 48 hours' induction of RNAi in G1 cells (Mann-Whitney U test, $p<2.2e-16$), the new flagella of post mitotic cells (Mann-Whitney U test, $p<2.2e-16$) and the old flagella of post mitotic cells (Mann-Whitney U test, $p<2.2e-16$). In induced cells (48h TET+), a bimodal distribution can be observed in G1 cells and in the old flagellum of post mitotic cells (Fig. 5.4b). New flagella of post mitotic cells are severely affected with a mean length of 3.4µm compared to a mean length of 14.7µm in uninduced (TET-) cells. The data suggests that it is the assembly of the new flagellum that is severely affected and that the old flagella and G1 flagella show a bimodal distribution due to the inheritance of the shorter new flagellum over several cell cycles. Depletion of Rabl2b also results in a severe growth defect, with induced (TET+) cells reaching ~25% of the density of uninduced (TET-) cells by 72 hours. However, unlike TFP68 and CEP90, depletion of Rabl2b did not affect the localisation of CEP164A at the transition fibres (Fig. 5.5a, magenta).

RabI2b RNAi

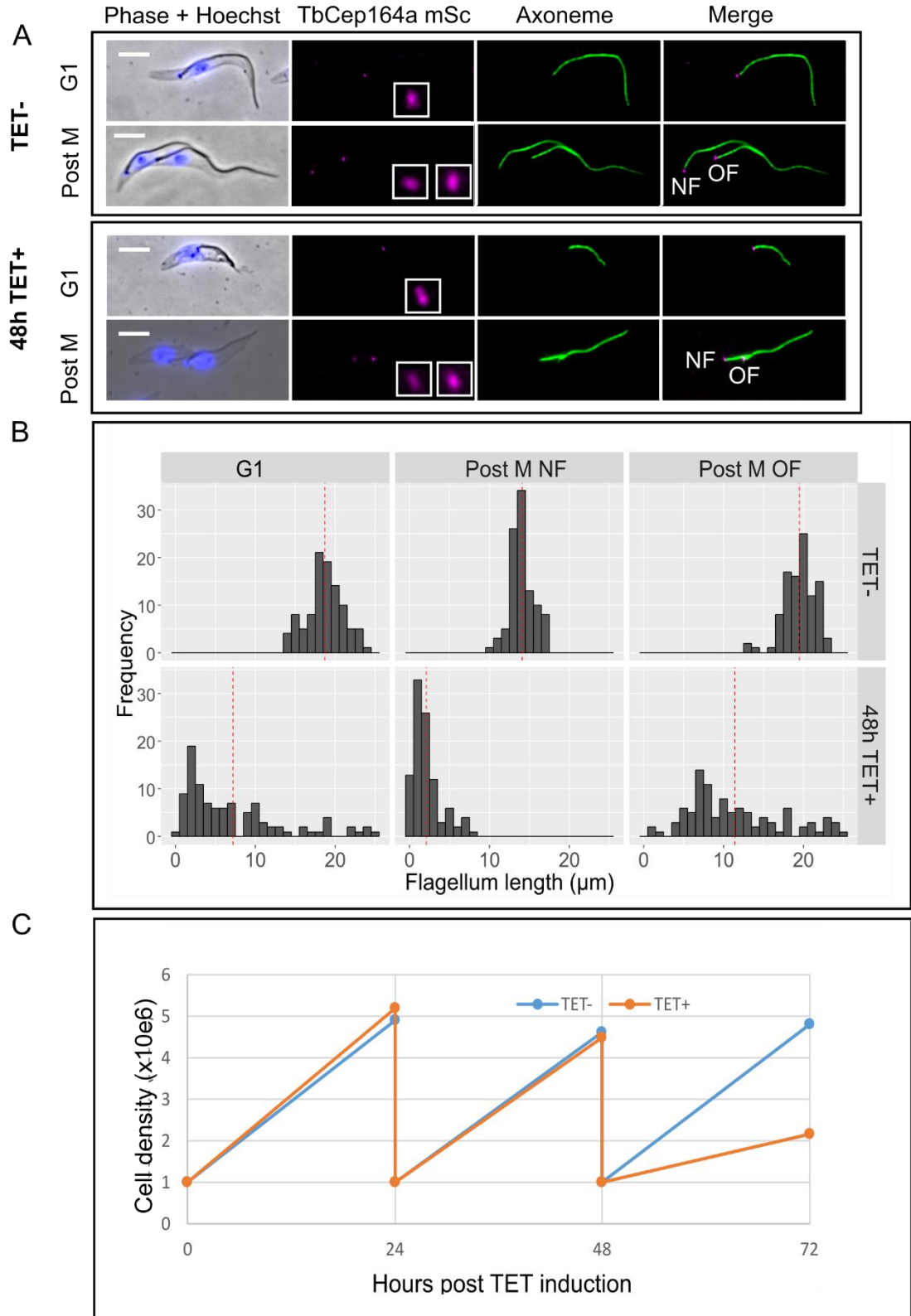


Figure 5.4 Depletion of *RabI2b* by RNAi. (a) Methanol-fixed cytoskeletons of uninduced cells that have not been depleted for CEP90 (TET-) and cells following 48 hours' induction of RNAi (48h TET+) in both G1 and post mitotic cells. DNA within the nucleus and kinetoplast is labelled with Hoechst. Insert shown within mSc::CEP164A (magenta) column represents a 5x zoom. Immunolabelling of the flagellar axoneme (green) was carried out using the mAb25 antibody. New flagellum and old flagellum are labelled NF and OF respectively. Images taken at 100x magnification. Scale bar represents 5μm.

(b) Length of flagellar axoneme (μm) based on mAb25 labelling comparing uninduced (TET-, top) and 48 hours induced (48h TET+, bottom) in G1 cells ($n=100$) (Mann-Whitney U test, $p<2.2\text{e-}16$), the new flagellum of post mitotic cells (Post M NF) ($n=100$) (Mann-Whitney U test, $p<2.2\text{e-}16$) and old flagella of post mitotic cells (post M OF) ($n=100$) (Mann-Whitney U test, $p<2.2\text{e-}16$). (c) Growth over 72 hours in cells induced for RNAi (TET+) and uninduced (TET-) cells. Cells were cultured back to 2×10^6 cells/mL every 24 hours. (d) Basal body count in cell in induced (TET+) and uninduced (TET-) cells at G1 and post M phase of the cell cycle.

5.3.4 Ciliopathy protein OFD1 is required for flagellum assembly

OFD1 is a conserved ciliary protein associated with human ciliopathies such as Oral-facial-digital (OFD1) syndrome type 1 (Sakakibara et al., 2019) and Joubert syndrome (Hammarsjo et al., 2017). Understanding the function and mechanism of OFD1 in trypanosome flagellar assembly may give us a broader understanding of the human disease mechanisms. Therefore, the trypanosome OFD1 protein was taken forward for functional analysis by RNAi.

Following knockdown of OFD1 in trypanosomes, we found flagella were significantly shorter at 48 hours post induction of RNAi in G1 cells (Mann-Whitney U test, $p=0.00032$) and the new flagellum of post mitotic cells (Mann-Whitney U test, $p=4.7\text{e-}11$). The difference in the length of the old flagellum of post mitotic cells was not statistically significant between induced cells (48h TET+) and the control (TET-) (Mann-Whitney U test, $p=0.189$). However, there is a clear change in the distribution of old flagellum lengths (Fig. 5.4b). This suggests again that the distribution in old flagella lengths is changing due to the inheritance of shorter new flagella from previous cell cycles.

Similar to Rabl2b, there is a clear bimodal distribution with shorter and longer flagella in G1 cells at 48 hours post induction. This indicates the proportion of G1 daughter cells have inherited a new flagellum in which assembly was defective and a population of G1 cells that have inherited the already assembled old flagellum from its parent cell. This therefore suggests that OFD1 is also required for the assembly of a new flagellum in trypanosomes. Interestingly, in G1 cells and the old flagellum of post mitotic cells, there are some cells with longer flagella in the induced population compared to the uninduced population (Fig. 5.4b). We hypothesise this may be either due to dysregulation of flagellum assembly mechanisms such as an increased pool of material loaded into the old flagellum or an increase in the length of the cell cycle that allows the old flagellum more time to assemble to a greater length.

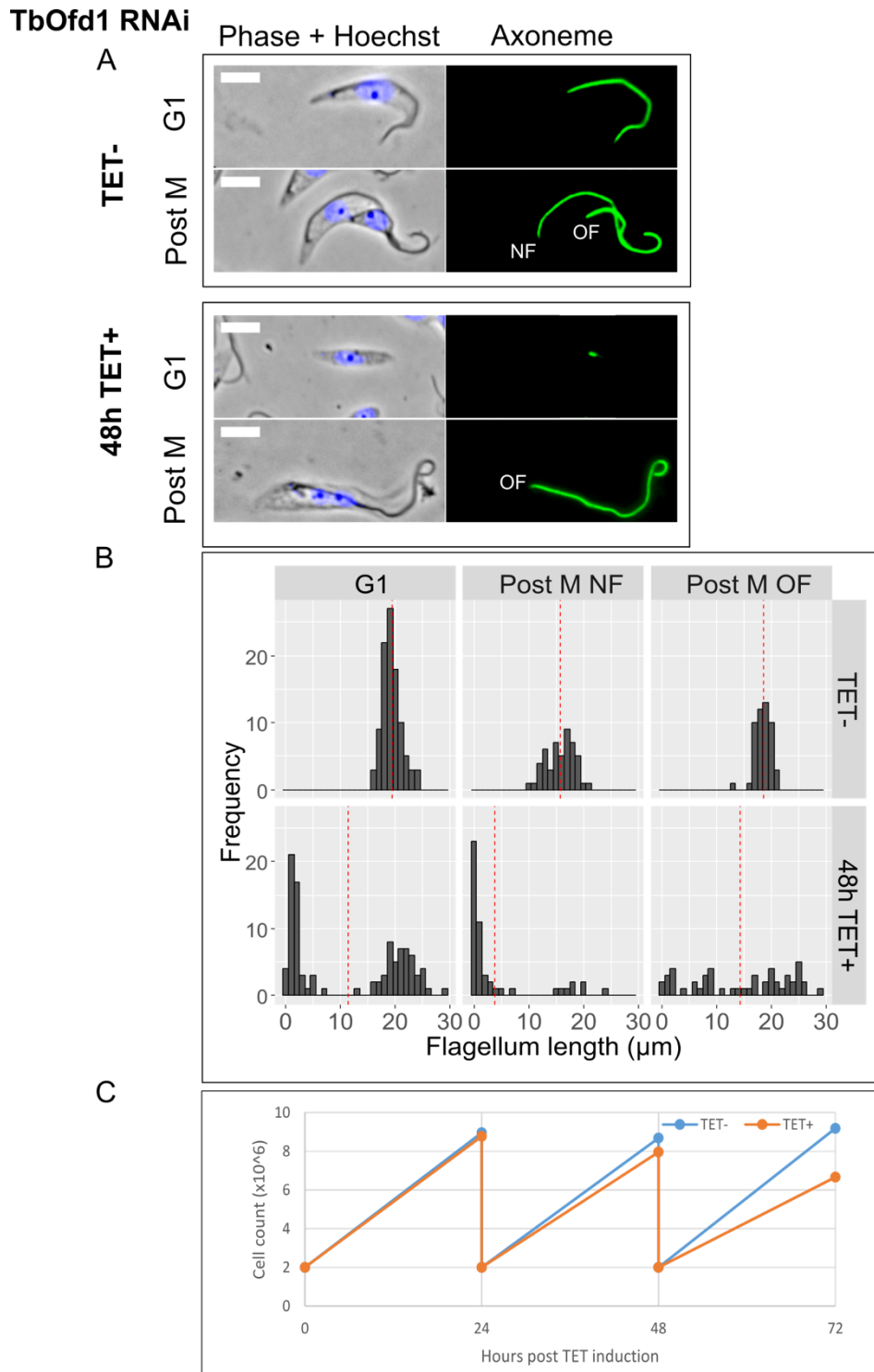


Figure 5.5 Depletion of OFD1 by RNAi. (a) Methanol-fixed cytoskeletons of uninduced cells that have not been depleted for TFP68 (TET-) and cells following 48 hours' induction of RNAi (48h TET+) in both G1 and post mitotic cells. DNA within the nucleus and kinetoplast is labelled with Hoechst. Insert shown within mSc::CEP164A (magenta) column represents a 5x zoom. Immunolabelling of the flagellar axoneme (green) with the mAb25 antibody. New flagellum and old flagellum

are labelled NF and OF respectively. Images taken at 100x magnification. Scale bar represents 5µm. (b) Length of flagellar axoneme (µm) based on mAb25 labelling comparing uninduced (TET-, top) and 24 hours induced (24h TET+, bottom) in G1 cells (n=100) (Mann-Whitney U test, $p=0.000316$), the new flagellum of post mitotic cells (Post M NF) (n=100) (Mann-Whitney U test, $p<2.2e-16$) and old flagella of post mitotic cells (post M OF) (n=100) (Mann-Whitney U test, $p=0.19$). (c) Growth over 72 hours in cells induced for RNAi (TET+) and uninduced (TET-) cells. Cells were cultured back to 2×10^6 cells/mL every 24 hours. (d) Basal body count in cell in induced (TET+) and uninduced (TET-) cells at G1 and post M phase of the cell cycle.

5.3.5 CEP19L is required for flagellar assembly and localisation of CEP164A

CEP19L protein in trypanosomes is not an orthologue of the mammalian CEP19 protein, however it possesses a CEP19-like family domain (Chapter 4, Figure 4.11). Mammalian CEP19 forms a complex with Rabl2b at the ciliary base, facilitating the import of IFT- complex B into the flagellum (Kanie et al., 2017, Nishijima et al., 2017). We examined the function of the trypanosome CEP19L protein by RNAi. At 72h post induction of RNAi, flagella were significantly shorter in G1 cells (Mann-Whitney U test, $p<2.2e-16$), new flagella of post mitotic cells (Mann-Whitney U test, $p<2.2e-16$) and in old flagella of post mitotic cells (Mann-Whitney U test, $p<2.2e-16$) (Fig. 5.6a, 5.6b). CEP19L depleted cells exhibited a modest growth defect by 72 hours. CEP164A was maintained in all cells at 48 hours post induction of RNAi. However, by 72 hours post induction signal was very weak or absent in 10% of G1 cells. In post M cells, the CEP164A signal was missing or very weak from both the new and old flagellum in 23% of cells and absent or very weak at the old flagellum only in 23% of cells. (Fig. 5.6a, magenta). The remaining 54% of post M cells had a CEP164A signal visible on the new and old flagellum.

Cep19L RNAi

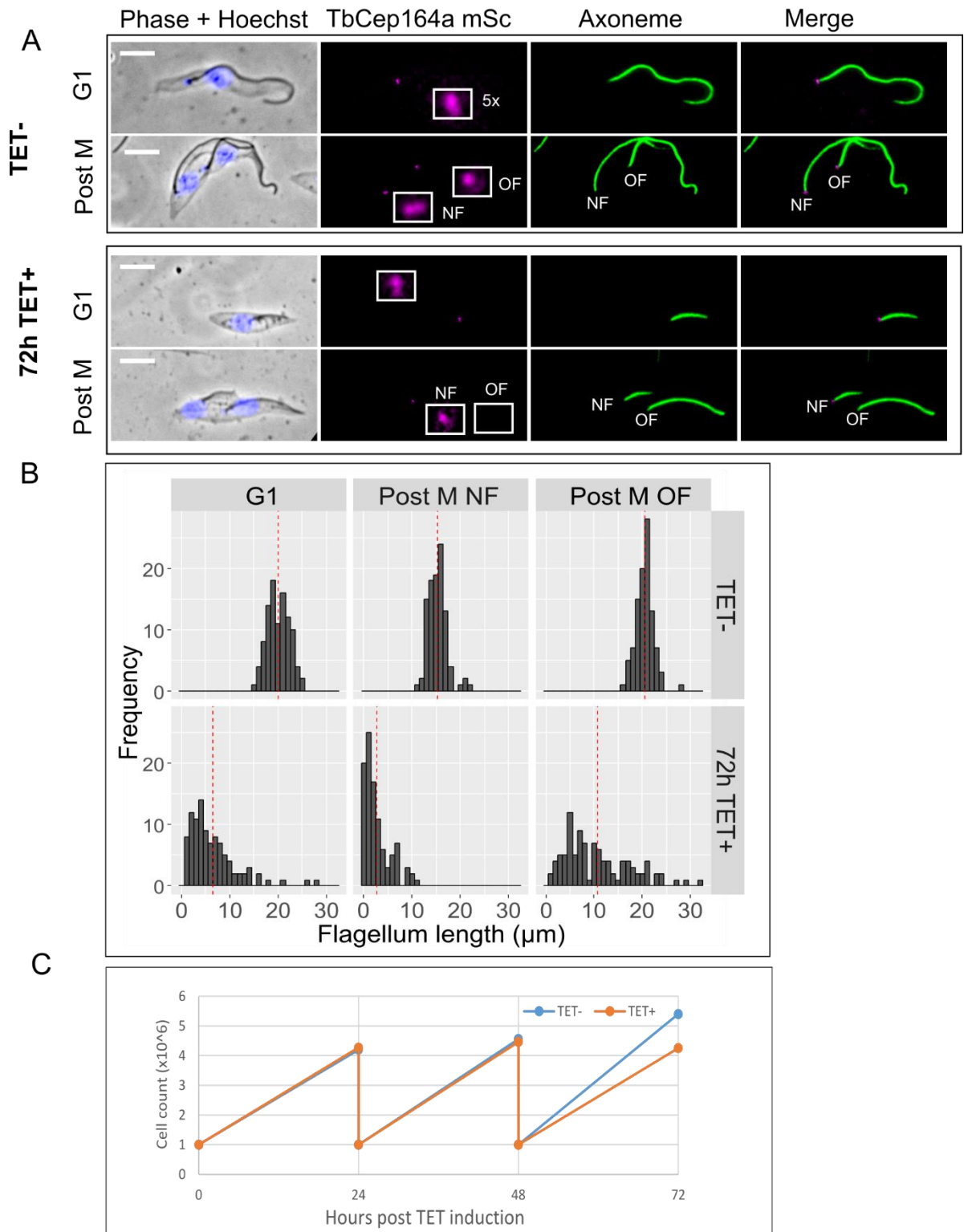


Figure 5.6 Depletion of CEP19L by RNAi. (a) Methanol-fixed cytoskeletons of uninduced cells that have not been depleted for CEP19L (TET-) and cells following 72 hours' induction of RNAi (72h TET+) in both G1 and post mitotic cells. DNA within the nucleus and kinetoplast is labelled with Hoechst. Insert shown within mSc::CEP164A (magenta) column represents a 5x zoom. Immunolabelling of the flagellar axoneme (green) with the mAb25 antibody. New flagellum and old flagellum are labelled NF and OF respectively. Images taken at 100x magnification. Scale bar represents 5 μm . (b) Length of

flagellar axoneme (μm) based on mAb25 labelling comparing uninduced (TET-, top) and 24 hours induced (24h TET+, bottom) in G1 cells ($n=100$) (Mann-Whitney U test, $p<2.2\text{e-}16$), the new flagellum of post mitotic cells (Post M NF) ($n=100$) (Mann-Whitney U test, $p<2.2\text{e-}16$) and old flagella of post mitotic cells (post M OF) ($n=100$) (Mann-Whitney U test, $p<2.2\text{e-}16$). (c) Growth over 72 hours in cells induced for RNAi (TET+) and uninduced (TET-) cells. Cells were cultured back to 2×10^6 cells/mL every 24 hours. (d) Basal body count in cell in induced (TET+) and uninduced (TET-) cells at G1 and post M phase of the cell cycle.

5.3.6 FBF1 is required for flagellum assembly

FBF1 is a conserved transition fibre protein with a known orthologue that has been previously localised to the transition fibres in humans (Tanos et al., 2013, Yang et al., 2018). FBF1 has also been shown to function in IFT recruitment *C. elegans* (Wei et al., 2013). Depletion of FBF1 by RNAi resulted in shorter flagella in G1 cells (Mann-Whitney U test, $p<2.2\text{e-}16$), new flagella of post mitotic cells (Mann-Whitney U test, $p<2.2\text{e-}16$) and old flagella of post mitotic cells (Mann-Whitney U test, $p<2.2\text{e-}16$) (Fig. 5.7a, Fig. 5.7b). CEP164A signal was maintained by 72 hours post induction of RNAi (Fig. 5.5a, magenta) and there was a moderate effect on growth by 72 hours (Fig 5.7c). The data indicate that FBF1 is required for the assembly of the axoneme of the new flagellum and by 72 hours we observe a shortening old flagellum due to inheritance of the shorter new flagellum in daughter cells that goes on to form the old flagellum.

TbFBF1 RNAi

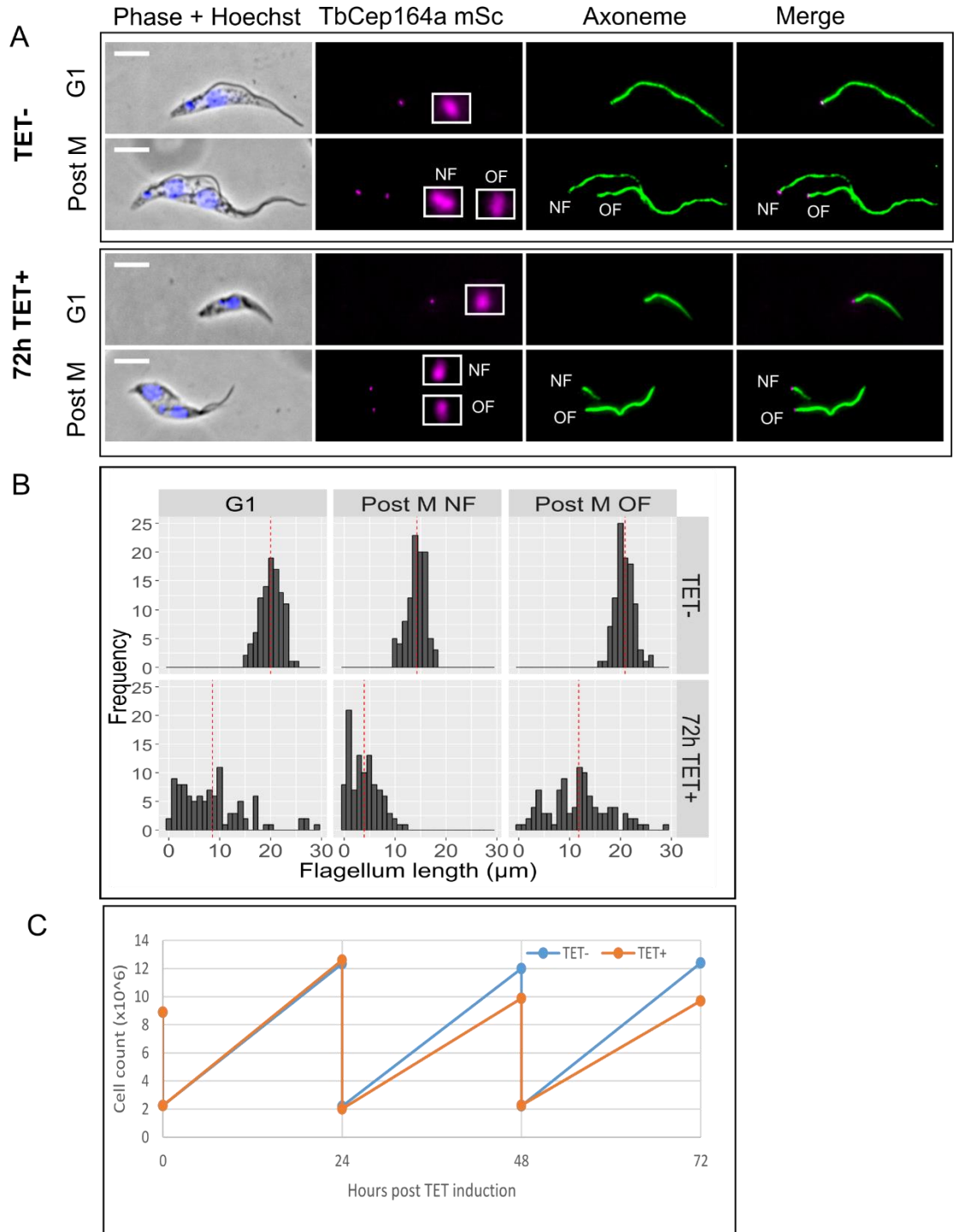


Figure 5.7 Depletion of FBF1 by RNAi. (a) Methanol-fixed cytoskeletons of uninduced cells that have not been depleted for FBF1 (TET-) and cells following 72 hours' induction of RNAi (72h TET+) in both G1 and post mitotic cells. DNA within the nucleus and kinetoplast is labelled with Hoechst. Insert shown within mSc::CEP164A (magenta) column represents a 5x zoom. Immunolabelling of the flagellar axoneme (green) with the mAb25 antibody. New flagellum and old flagellum are labelled NF and OF respectively. Images taken at 100x magnification. Scale bar represents 5μm. (b) Length of flagellar axoneme (μm) based on mAb25 labelling comparing uninduced (TET-, top) and 24 hours induced (24h TET+, bottom) in G1 cells (n=100) (Mann-Whitney U test, $p < 2.2 \times 10^{-16}$), the new flagellum of post mitotic cells (Post M NF) (n=100) (Mann-Whitney U test, $p < 2.2 \times 10^{-16}$) and old flagella of post mitotic cells (post M OF) (n=100) (Mann-Whitney U

test, $p < 2.2 \times 10^{-16}$). (c) Growth over 72 hours in cells induced for RNAi (TET+) and uninduced (TET-) cells. Cells were cultured back to 2×10^6 cells/mL every 24 hours. (d) Basal body count in cell in induced (TET+) and uninduced (TET-) cells at G1 and post M phase of the cell cycle.

In summary, knockdown of 6 trypanosome transition fibre protein resulted in defects in flagellum assembly. TFP68, CEP90 and Rabl2b appeared to **affect** flagellum assembly most severely based on their rapid and severe growth phenotype and the substantial reduction in flagellum length. OFD1, CEP19L and FBF1 also appear to play a significant role in new flagellum assembly, but exhibited a more moderate growth defect which indicated these cells were still able to assemble some flagellar-like structure but assembly was still severely perturbed. We conclude that these 6 proteins are required for the assembly of a new axoneme in trypanosomes.

5.4 Two transition fibre proteins are required for flagella to assemble to their final length

5.4.1 CFAP52 is required for flagella to reach their final length

CFAP52 (Cilia flagella associated protein 52) is the most widely conserved protein in this cohort, with orthologues present in all major taxonomic groups with the exception of Amoebozoa (chapter 4, Fig. 4.11). CFAP52 has never been associated with transition fibres in any organism. Following 72 hours' depletion of CFAP52 by RNAi cells appeared morphologically similar to uninduced cells by visual inspection (Fig 5.8a). However, measurements of the axoneme in post M cells showed that measurements of the new flagella were significantly shorter when compared to uninduced cells with a mean difference of $3.5 \mu\text{m}$ (Mann-Whitney U test, $p = 1.7 \times 10^{-6}$) (Fig 5.8b). However, the length of the old flagella was not significantly different between the induced and uninduced populations (Mann-Whitney U test, $p = 0.0159$, $\alpha = 0.01$). This is likely because the effect on flagellum length was more nuanced and cells had not yet inherited enough shorter flagella by 72 hours. The effect of CFAP52 on growth was mild following 72 hours' induction of RNAi. We conclude from this that CFAP52 may not be required for the initial steps of axoneme assembly, but is required for the flagellum to reach its final length.

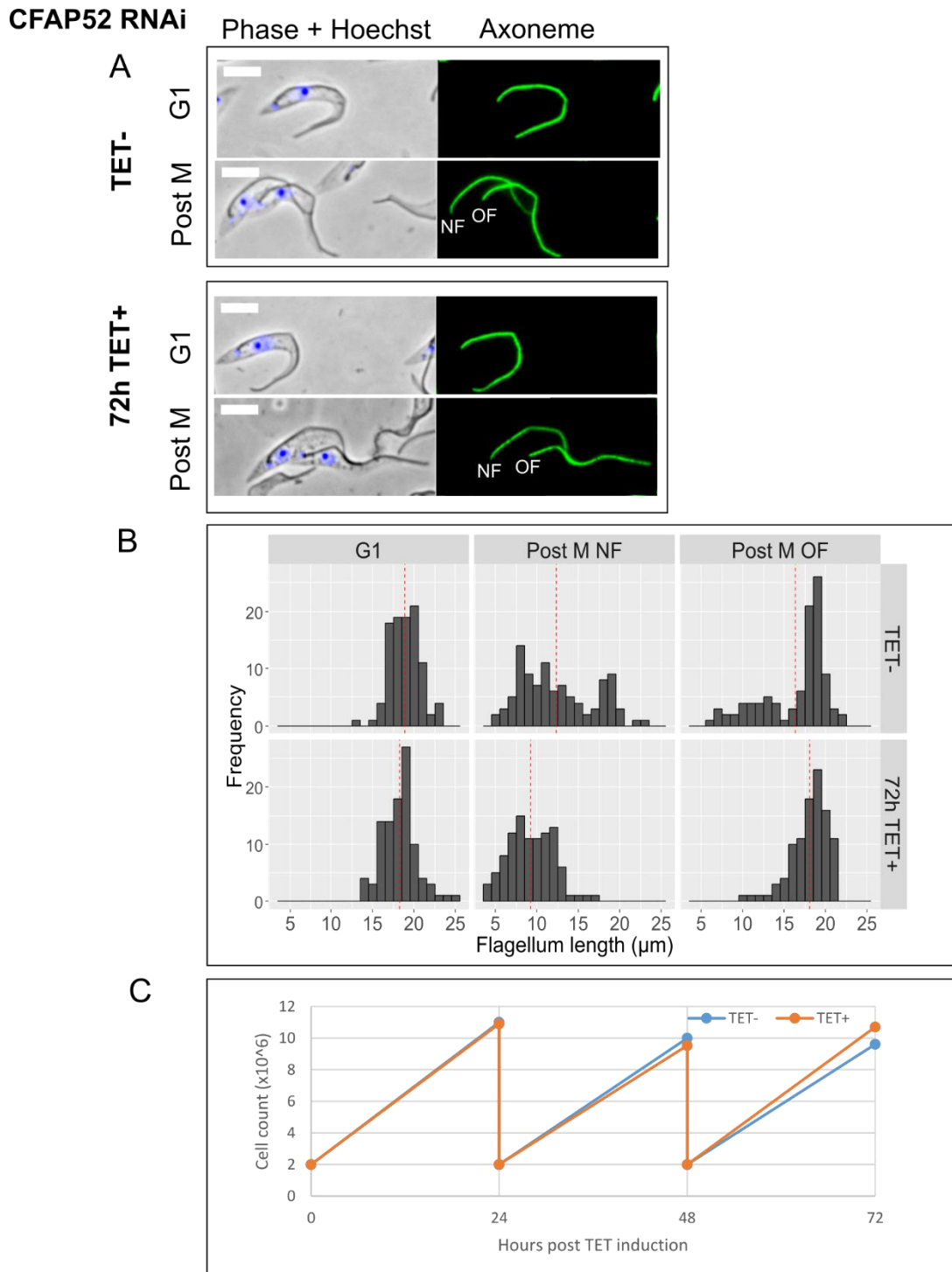


Figure 5.8 Depletion of CFAP52 by RNAi. (a) Methanol-fixed cytoskeletons of uninduced cells that have not been depleted for CFAP52 (TET-) and cells following 72 hours' induction of RNAi (72h TET+) in both G1 and post mitotic cells. DNA within the nucleus and kinetoplast is labelled with Hoechst. Insert shown within *mSc::CEP164A* (magenta) column represents a 5x zoom. Immunolabelling of the flagellar axoneme (green) with the mAb25 antibody. New flagellum and old flagellum are labelled NF and OF respectively. Images taken at 100x magnification. Scale bar represents 5 μm . (b) Length of flagellar axoneme (μm) based on mAb25 labelling comparing uninduced (TET-, top) and 24 hours induced (24h TET+, bottom) in G1 cells ($n=100$) (Mann-Whitney U test, $p=0.019$), the new flagellum of post mitotic cells (Post M NF) ($n=100$) (Mann-Whitney U test, $p=1.7\text{e-}6$) and old flagella of post mitotic cells (post M OF) ($n=100$) (Mann-Whitney U

test, $p=0.0159$). (c) Growth over 72 hours in cells induced for RNAi (TET+) and uninduced (TET-) cells. Cells were cultured back to 2×10^6 cells/mL every 24 hours. (d) Basal body count in cell in induced (TET+) and uninduced (TET-) cells at G1 and post M phase of the cell cycle.

5.4.2 TFP49 is required for flagella to assemble to the correct length

TFP49 is a transition fibre protein that is unique to kinetoplastids. We examined the function of TFP49 by RNAi. Following 72 hours' induction of RNAi, there was no observable difference in cell morphology and flagella length and CEP164A signal was maintained in all cells (Fig. 5.9a).

However, measurements of the flagella at 72 hours revealed that TFP49 depleted cells had significantly shorter flagella in G1 cells (Mann-Whitney U test, $p < 2.2 \times 10^{-16}$), the new flagella of post mitotic cells (Mann-Whitney U test, $p < 2.2 \times 10^{-16}$) and the old flagella of post mitotic cells (1.07×10^{-9}). A moderate growth defect could be observed following 72 hours depletion by RNAi. The data show that (similar to CFAP52) TFP49 depleted cells were still able to initiate assembly and form a flagellum but cells were unable to assemble a flagellum to the correct length.

In summary, we show that knockdown of 2 proteins (TFP49 and CFAP52) results in cells that are able to assemble a flagellum and appear morphologically normal. However, the flagellum was not able to reach the correct length. We conclude that these 2 transition fibre proteins are not required for initial assembly of the new flagellum but speculate they function ensure the flagellum lock is activated at the correct time. If a flagellum lock is activated prematurely we would expect to see cells normal in appearance with a flagellum that is slightly shorter than the control population.

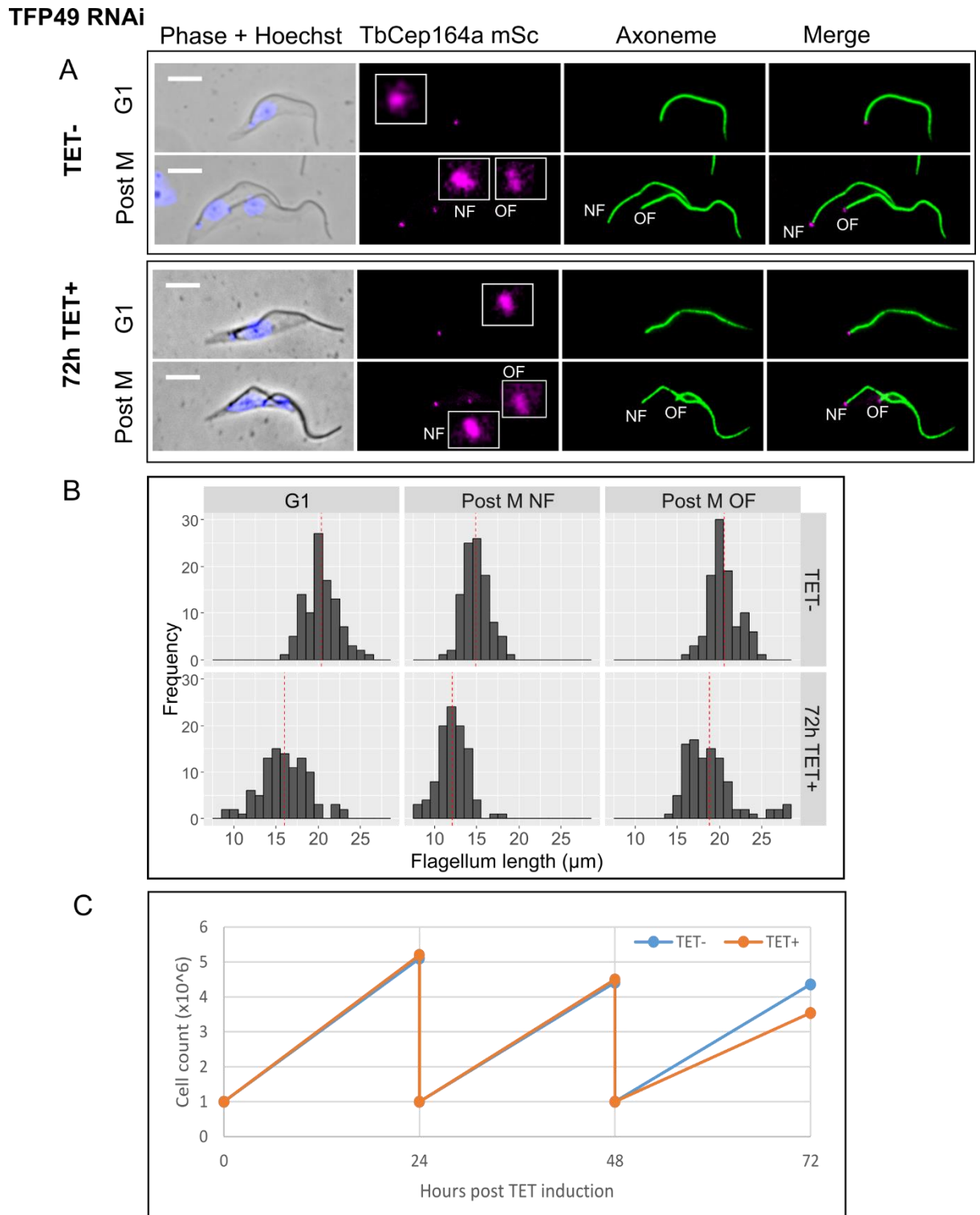


Figure 5.9 Depletion of TFP49 by RNAi. (a) Methanol-fixed cytoskeletons of uninduced cells that have not been depleted for TFP49 (TET-) and cells following 72 hours' induction of RNAi (72h TET+) in both G1 and post mitotic cells. DNA within the nucleus and kinetoplast is labelled with Hoechst. Insert shown within mSc::CEP164A (magenta) column represents a 5x zoom. Immunolabelling of the flagellar axoneme (green) with the mAb25 antibody. New flagellum and old flagellum are labelled NF and OF respectively. Images taken at 100x magnification. Scale bar represents 5 μ m. (b) Length of flagellar axoneme (μ m) based on mAb25 labelling comparing uninduced (TET-, top) and 24 hours induced (24h TET+, bottom) in G1 cells ($n=100$) (Mann-Whitney U test, $p=0.019$), the new flagellum of post mitotic cells (Post M NF) ($n=100$)

(Mann-Whitney U test, $p=1.7e-6$) and old flagella of post mitotic cells (post M OF) ($n=100$) (Mann-Whitney U test, $p=0.0159$). (c) Growth over 72 hours in cells induced for RNAi (TET+) and uninduced (TET-) cells. Cells were cultured back to 2×10^6 cells/mL every 24 hours. (d) Basal body count in cell in induced (TET+) and uninduced (TET-) cells at G1 and post M phase of the cell cycle.

5.5 Transition fibre proteins required for flagellum length control

Many ciliated eukaryotes such as *Trypanosoma brucei*, *Giardia* spp. and *Paramecium* spp. are confronted with the unique biological challenge of having to assemble flagella while simultaneously maintaining a full length existing flagellum within the same cell. Recently, a growth and lock model was proposed in trypanosomes which suggests that the old flagellum in trypanosomes is locked once it reaches a certain length (Bertiaux et al., 2018). One of the proteins involved in this locking mechanism in trypanosomes is the transition fibre protein CEP164C. CEP164C is found exclusively at the old flagellum and when depleted for CEP164C, cells assemble a flagellum that is longer (Atkins et al., 2021). We hypothesise that there are other proteins at the transition fibres that also function as part of the old flagellum 'lock' in trypanosomes.

5.5.1 TFP122 is a regulator of flagellum length

TFP122 is a kinetoplastid specific transition fibre protein that possesses a kinesin motor domain at its N terminal region. TFP122 possess a secondary localization in the flagellum and also exhibits a brighter signal at the transition fibres of the old flagellum of post mitotic cells, which raises an interesting question of whether TFP122 is targeting the old flagellum to support an old flagellum specific mechanism like old flagellum locking. Following 72 hours' depletion of TFP122 by RNAi, cells appeared to have morphology and flagella similar to that of uninduced (TET-) cells upon visual inspection, and the localisation of the CEP164A signal was unaffected (magenta) (Fig. 5.10a). A closer look at measurements of the immunolabelled axoneme (mAb25) revealed that flagella in TFP122 depleted cells were significantly longer in G1 cells (Mann-Whitney U test, $p=2.26e-13$), new flagella of post M cells (Mann-Whitney U test, $p=1.31e-05$) and in old flagella of post M cells (Mann-Whitney U test, $p=5.9e-06$). TFP122 depleted cells also showed a moderate growth defect, suggesting that either longer flagella are restricting the rate of division by 72 hours or there is another mechanism that is affecting the rate of division in TFP122 depleted cells.

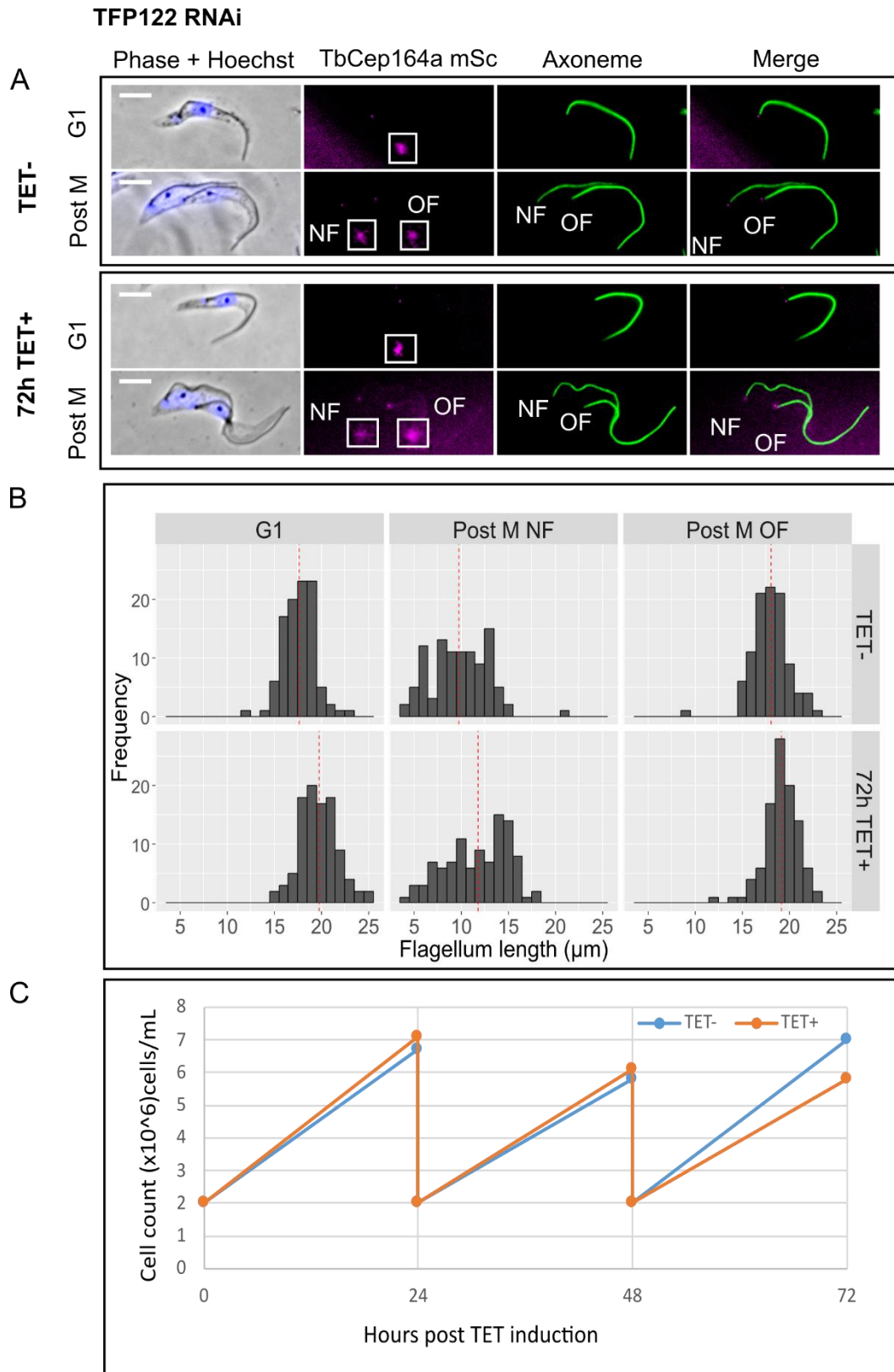


Figure 5.10 Figure 5.9 Depletion of TFP122 by RNAi. (a) Methanol-fixed cytoskeletons of uninduced cells that have not been depleted for TFP122 (TET-) and cells following 72 hours' induction of RNAi (72h TET+) in both G1 and post mitotic cells. DNA within the nucleus and kinetoplast is labelled with Hoechst. Insert shown within mSc::CEP164A (magenta) column represents a 5x zoom. Immunolabelling of the flagellar axoneme (green) with the mAb25 antibody. New flagellum and old flagellum are labelled NF and OF respectively. Images taken at 100x magnification. Scale bar

represents 5 μm . (b) Length of flagellar axoneme (μm) based on mAb25 labelling comparing uninduced (TET-, top) and 24 hours induced (24h TET+, bottom) in G1 cells ($n=100$) (Mann-Whitney U test, $p=2.3\text{e-}13$), the new flagellum of post mitotic cells (Post M NF) ($n=100$) (Mann-Whitney U test, $p=1.31\text{e-}05$) and old flagella of post mitotic cells (post M OF) ($n=100$) (Mann-Whitney U test, $p=4.28\text{e-}11$). (c) Growth over 72 hours in cells induced for RNAi (TET+) and uninduced (TET-) cells. Cells were cultured back to 2×10^6 cells/mL every 24 hours. (d) Basal body count in cell in induced (TET+) and uninduced (TET-) cells at G1 and post M phase of the cell cycle.

5.5.2 TFP15 is a regulator of flagellum length

TFP15 is a kinetoplastids specific protein that exhibits the widest pseudo-diameter out of all transition fibre proteins in chapter 3 (Fig. 3.9). It is known that the outer most transition fibre protein in mammals, CEP164a, is indispensable for basal body to membrane docking and ciliary assembly (Cajanek and Nigg, 2014, Graser et al., 2007, Schmidt et al., 2012, Tanos et al., 2013, Yang et al., 2018). We therefore decided to examine the function of TFP15 by RNAi. Following 72 hours' induction of RNAi, cells appeared similar in their morphology and flagella length compared to uninduced cells upon visual inspection (Fig. 5.11a). Measurements of the immunolabelled axoneme (mAb25, Fig. 5.11a) revealed that flagella were significantly longer in G1 cells (Mann-Whitney U test, $p<2.2\text{e-}16$), new flagella of post mitotic cells (Mann-Whitney U test, $p=4.27\text{e-}07$) and old flagella of post mitotic cells (Mann-Whitney U test, $p=5.9\text{e-}06$) (Fig. 5.11b). However, in contrast to TFP122, growth was unaffected by 72 hours post induction of RNAi (Fig. 5.11c).

In summary, trypanosome flagella require a locking mechanism to control the length of the old flagellum while assembling a new flagellum. CEP164C has been shown to be a central component of this (Atkins et al., 2021). Our analysis has identified two more candidates (TFP122 and TFP15) that may also be involved in the flagellar length control mechanism in trypanosomes and potentially in other kinetoplastids.

TFP15 RNAi

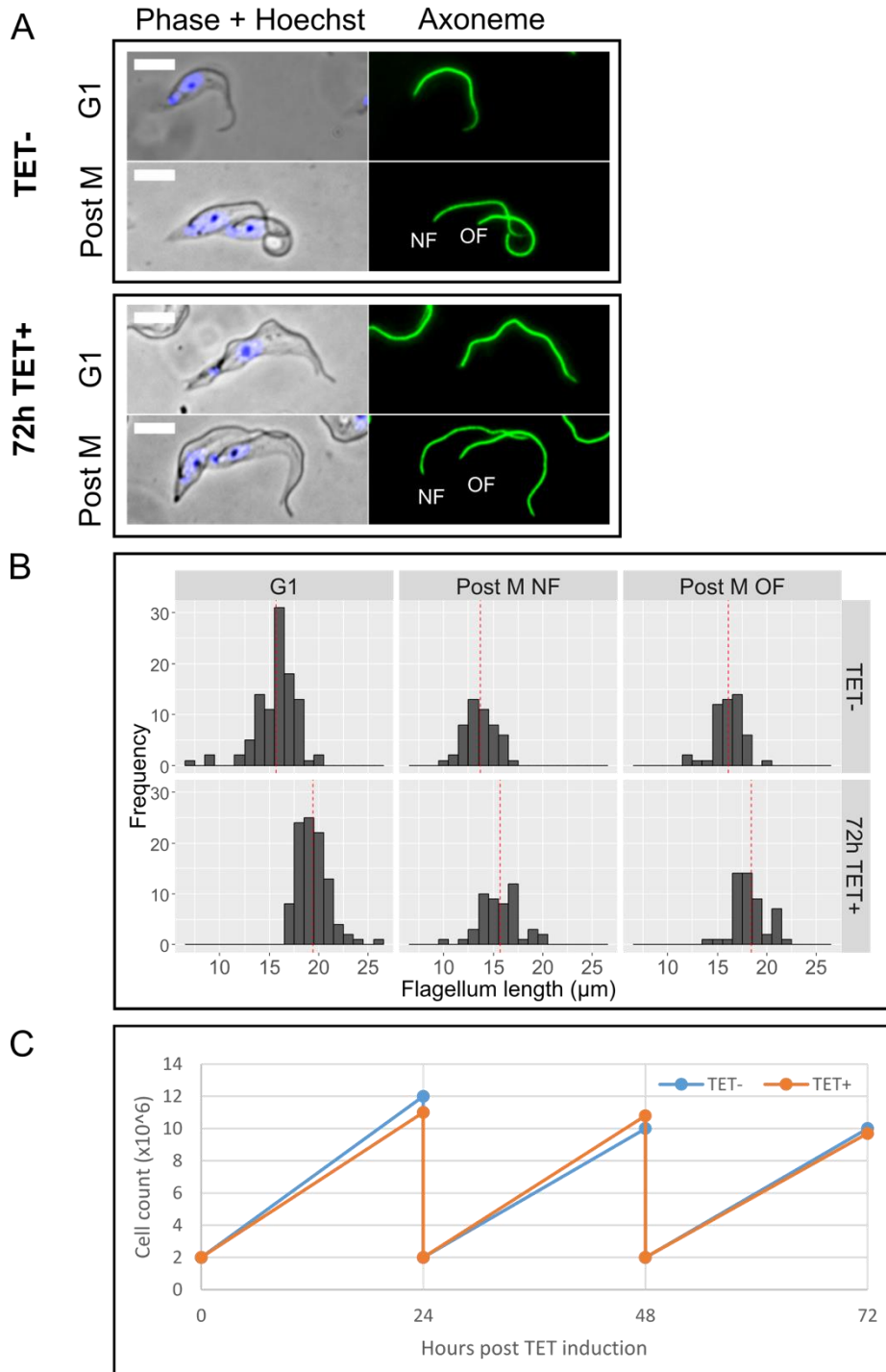


Figure 5.11 Depletion of TFP15 by RNAi. (a) Methanol-fixed cytoskeletons of uninduced cells that have not been depleted for TFP15 (TET-) and cells following 72 hours' induction of RNAi (72h TET+) in both G1 and post mitotic cells. DNA within the nucleus and kinetoplast is labelled with Hoechst. Immunolabelling of the flagellar axoneme (green) with the mAb25 antibody. New flagellum and old flagellum are labelled NF and OF respectively. Images taken at 100x magnification. Scale bar represents 5 μ m. (b) Length of flagellar axoneme (μ m) based on mAb25 labelling comparing uninduced (TET-, top) and 24 hours induced (24h TET+, bottom) in G1 cells ($n=100$) (Mann-Whitney U test, $p<2.2e-16$), the new flagellum of post mitotic cells (Post M NF) ($n=100$) (Mann-Whitney U test, $p=4.27e-07$) and old flagella of post mitotic cells (post M OF) ($n=100$) (Mann-Whitney U test, $p=5.9e-06$). (c) Growth over 72 hours in cells induced for RNAi (TET+) and uninduced

(TET-) cells. Cells were cultured back to 2×10^6 cells/mL every 24 hours. (d) Basal body count in cell in induced (TET+) and uninduced (TET-) cells at G1 and post M phase of the cell cycle.

5.6 Transition fibre proteins with a phenotype not related to flagellar function

Immunolabelling of the axoneme was used to determine the effects of transition fibre protein mutants (RNAi) on flagellar assembly and length control (Fig. 5.2-Fig. 5.11). Using this method, we identified 6 proteins required for flagellum assembly (TFP68, CEP90, OFD1, FBF1, CEP19L, Rabl2b) and 2 proteins required for flagellar length control. However, out of the 20 proteins taken forward for RNAi, 2 of these (Cullin 4b, SPB1) exhibited no change in flagellum length following induction of RNAi but did exhibit a phenotype that did not appear to be related to the flagellum.

5.6.1 Cullin4b is required for cell division but does not affect flagellar length

The human Cullin4b protein is known to be a core component of the E3 ligase complex that mediates proteasomal degradation of target proteins (Chen et al., 2018, Guerrero-Santoro et al., 2008, Higa et al., 2006). The trypanosome Cullin4b protein contains a Cullin family domain and is therefore an interesting candidate to take forward for functional analysis (Chapter 3, Fig 4.12).

Following depletion of Cullin4b, cells appeared to be morphologically distinct to uninduced (TET-) cells, although cells still appeared to have a normal flagellum (Fig. 5.12a). Measurements of the flagella axoneme revealed no significant differences in length between the induced and uninduced cells (Fig. 5.12b). However, growth was severely affected by 72 hours post induction compared to uninduced (TET-) cells with the induced cells reaching around 25% the density of the uninduced (TET-) cells by 72 hours. We hypothesise that ubiquitination at the transition fibres may contribute to cell cycle progression in trypanosomes that is not related to flagellum formation, although further studies would be required to support this.

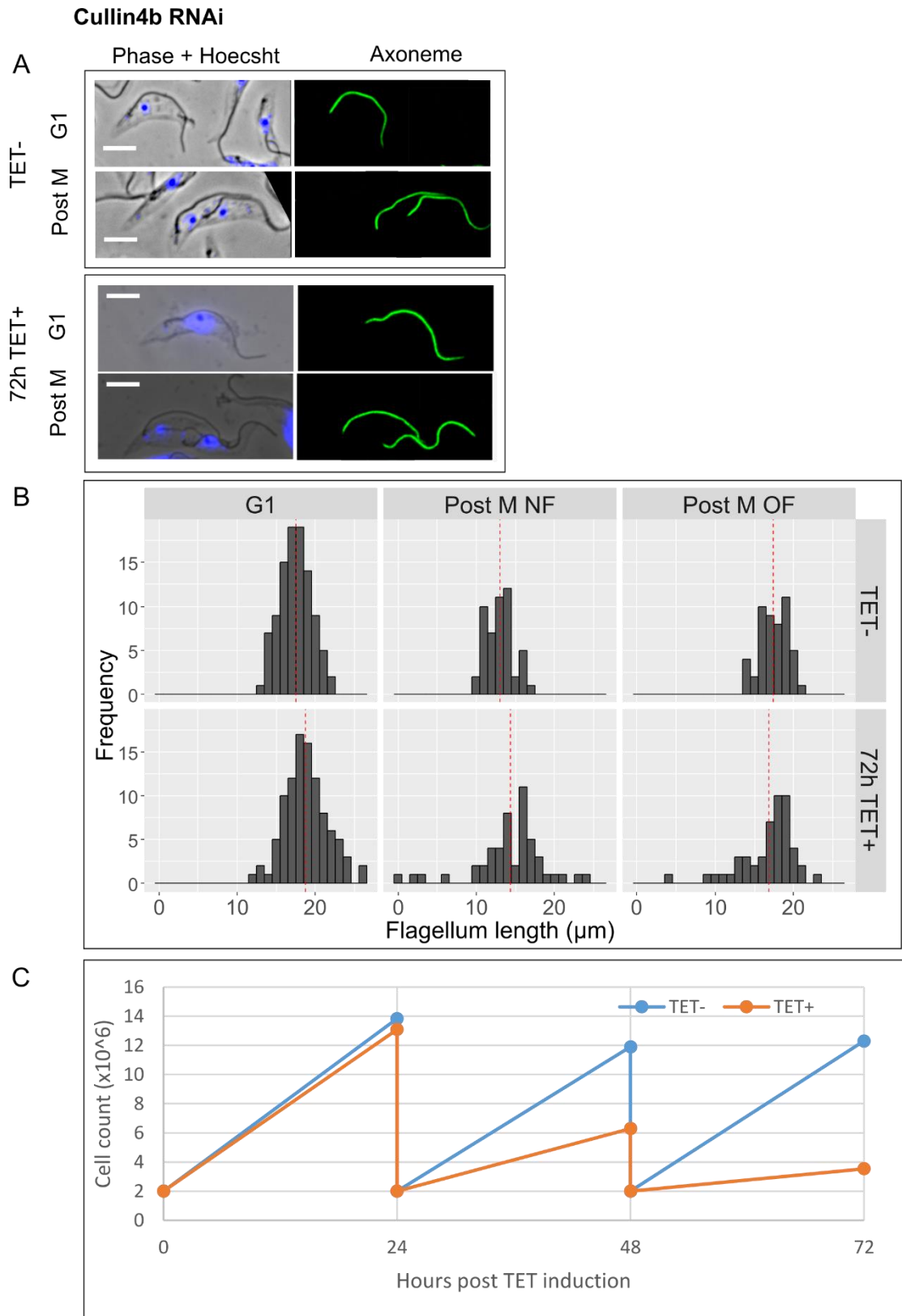


Figure 5.12 Depletion of Cullin4b by RNAi. (a) Methanol-fixed cytoskeletons of uninduced cells that have not been depleted for Cullin4b (TET-) and cells following 72 hours' induction of RNAi (72h TET+) in both G1 and post mitotic cells. DNA within the nucleus and kinetoplast is labelled with Hoechst. Immunolabelling of the flagellar axoneme (green) with the mAb25 antibody. New flagellum and old flagellum are labelled NF and OF respectively. Images taken at 100x magnification. Scale bar represents 5 μ m. (b) Length of flagellar axoneme (μ m) based on mAb25 labelling comparing

uninduced (TET-, top) and 24 hours induced (24h TET+, bottom) in G1 cells (n=100) (Mann-Whitney U test, $p>0.05$), the new flagellum of post mitotic cells (Post M NF) (n=100) (Mann-Whitney U test, $p>0.05$) and old flagella of post mitotic cells (post M OF) (n=100) (Mann-Whitney U test, $p>0.05$). (c) Growth over 72 hours in cells induced for RNAi (TET+) and uninduced (TET-) cells. Cells were cultured back to 2×10^6 cells/mL every 24 hours. (d) Basal body count in cell in induced (TET+) and uninduced (TET-) cells at G1 and post M phase of the cell cycle.

5.6.2 SPB1 may be required for normal cell cycle progression

SPB1 (spindle pole body protein 1) is a trypanosome specific protein that has been shown to localise to the mitotic spindle in trypanosomes, however no transition fibre localisation was reported in this study (Zhou et al., 2018). We have shown that SPB1 localises to the transition fibres, mitotic spindle pole and the mitotic spindle fibre (chapter 1, Fig. 3.7). Given the unusual localisation of the SPB1 protein, we decided to take SPB1 forward for functional analysis. Cells depleted for SPB1 by RNAi at 72 hours showed no differences in their cell morphology or flagella upon visual inspection (Fig. 5.13a). Measurements of the immunolabelled axoneme (mAb25) showed no significant differences between cells induced with RNAi (TET+) and uninduced cells (TET-) (Fig. 5.13b). However, when measuring the flagella of post M cells, it was observed that there was a larger proportion of post M cells in the induced (TET+) population compared to the uninduced population (TET-). We therefore carried out a cell cycle count in induced and uninduced cells to compare the number of cells at different stages of the cell cycle. This count revealed that by 72 hours of induction of RNAi, the number of post M cells was double that of the uninduced population and the corresponding number of cells in G2 was reduced (n=200). We hypothesise that either cells are moving more rapidly through G2 phase or cells are arresting in post M phase for longer prior to cytokinesis. Given the localisation of SPB1 during mitosis, the latter seems more likely.

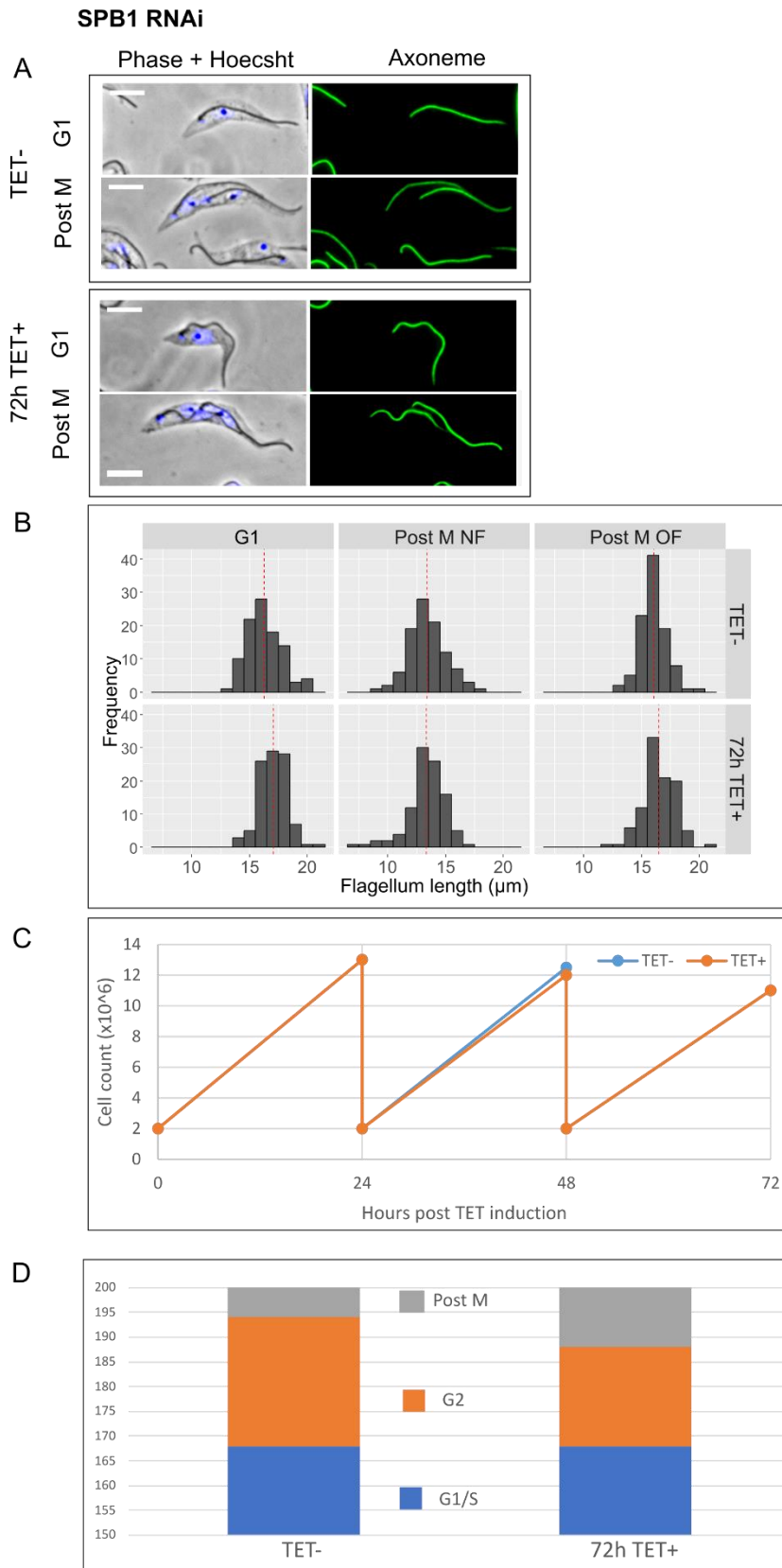


Figure 5.13 Depletion of SPB1 15 by RNAi. (a) Methanol-fixed cytoskeletons of uninduced cells that have not been depleted for SPB1 (TET⁻) and cells following 72 hours' induction of RNAi (72h TET⁺) in both G1 and post mitotic cells. DNA within the nucleus and kinetoplast is labelled with Hoechst. Immunolabelling of the flagellar axoneme (green) with the mAb25 antibody. New flagellum and old flagellum are labelled NF and OF respectively. Images taken at 100x

magnification. Scale bar represents 5µm. (b) Length of flagellar axoneme (µm) based on mAb25 labelling comparing uninduced (TET-, top) and 24 hours induced (24h TET+, bottom) in G1 cells (n=100) (Mann-Whitney U test, $p>0.05$), the new flagellum of post mitotic cells (Post M NF) (n=100) (Mann-Whitney U test, $p>0.05$) and old flagella of post mitotic cells (post M OF) (n=100) (Mann-Whitney U test, $p>0.05$). (c) Growth over 72 hours in cells induced for RNAi (TET+) and uninduced (TET-) cells. Cells were cultured back to 2×10^6 cells/mL every 24 hours. (d) Basal body count in cell in induced (TET+) and uninduced (TET-) cells at G1 and post M phase of the cell cycle.

5.7 Ciliopathy proteins and their effect on intraflagellar transport

Studies on transition fibres in other organisms including mammals and *C. elegans* have revealed that transition fibre proteins are required for the recruitment, import and loading of IFT complexes into the flagella/ciliary compartment to enable flagellum biogenesis (Kumar et al., 2021, Nishijima et al., 2017, Wei et al., 2015, Wei et al., 2013). We decided to take forward transition fibre proteins for IFT analysis that are (i) required for either flagellum assembly in trypanosomes, (ii) are conserved in humans and (iii) have human orthologues associated with ciliopathies. Therefore, we took forward OFD1, CEP90, Rabl2b and CEP19L for analysis. As discussed previously, mutations in the OFD1 and CEP90 genes are associated with ciliopathies in humans (Sakakibara et al., 2019, Wheway et al., 2015). We generated RNAi cell lines for these four transition fibre proteins and labelled IFT172 by immunofluorescence following 24 hours' depletion of each protein (IFT172 antibody gifted by Phillippe Bastin, Institut Pasteur). We considered this a suitable time based on the growth curve of each RNAi induction as the earliest time point before the severe flagella defects would occur. IFT172 is required for ciliogenesis in a range of organisms and is the largest subunit of IFT the complex (Bell et al., 2006, Huangfu et al., 2003, Taschner et al., 2016).

Figure 5.14 shows intensity of IFT172 in the cytoplasm base of the flagellum in uninduced cells (TET-) and cells 24 hours post induction of RNAi (24h TET+). Following depletion of CEP90, basal IFT172 intensity was lower than that of the control in G1 cells and at the new flagellum of post mitotic cells Fig 5.14a). However, the intensity of IFT172 at the flagellar base was slightly greater at the old flagellum of post mitotic cells. In the G1 cell in Figure 5.14a, there appeared to be a bright pool of IFT172 inside the flagellum but the cytoplasmic pool of IFT172 was reduced. We hypothesise that CEP90 is required for the loading of IFT172 onto the flagellar axoneme but is not required for the import of IFT172 into the flagellar compartment.

Following 24 hours of OFD1 knockdown by RNAi, the cytoplasmic pool of IFT172 was reduced in both the new and old flagella of post mitotic cells but not in the flagella of G1 cells (Fig. 5.14b). This indicates OFD1 may be playing a more significant role in dividing cells. As with CEP90, the

data suggests that OFD1 is required for the loading of IFT172 onto the axoneme but not the import into the flagella compartment.

Conversely, following 24 hours' knockdown of Rabl2b, cells showed a modest increase in the intensity of IFT172 at the new and old flagellum of post mitotic cells (Fig. 5.14c). This same effect was also observed with CEP19L depleted cells, however with CEP19L depleted cells, this effect was also measured in G1 cells as well as dividing cells (Fig. 5.14d). The accumulation of Rabl2b and CEP19L suggest that Rabl2b and CEP19L may be required for the import of IFT172 into the flagellar compartment.

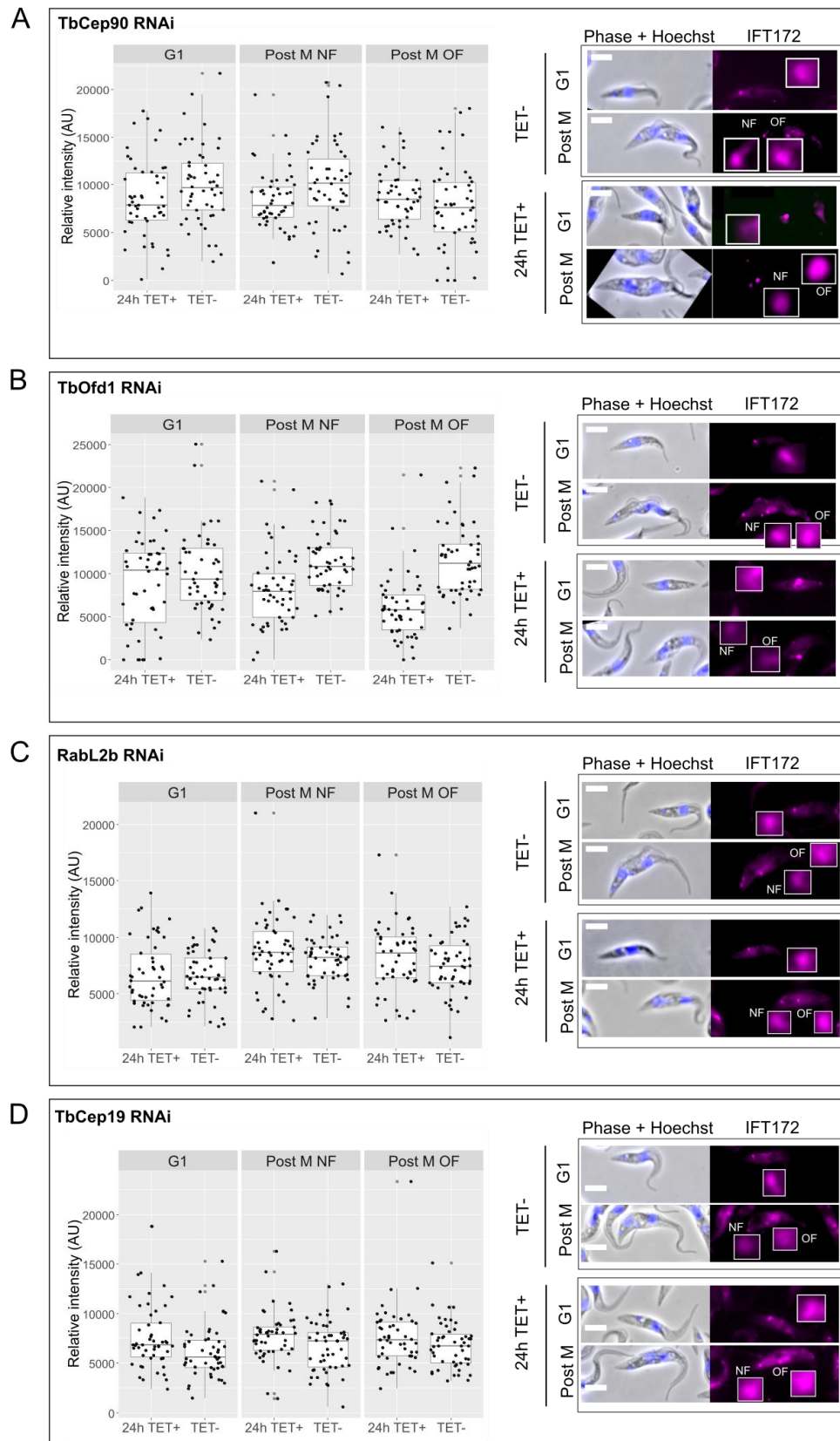
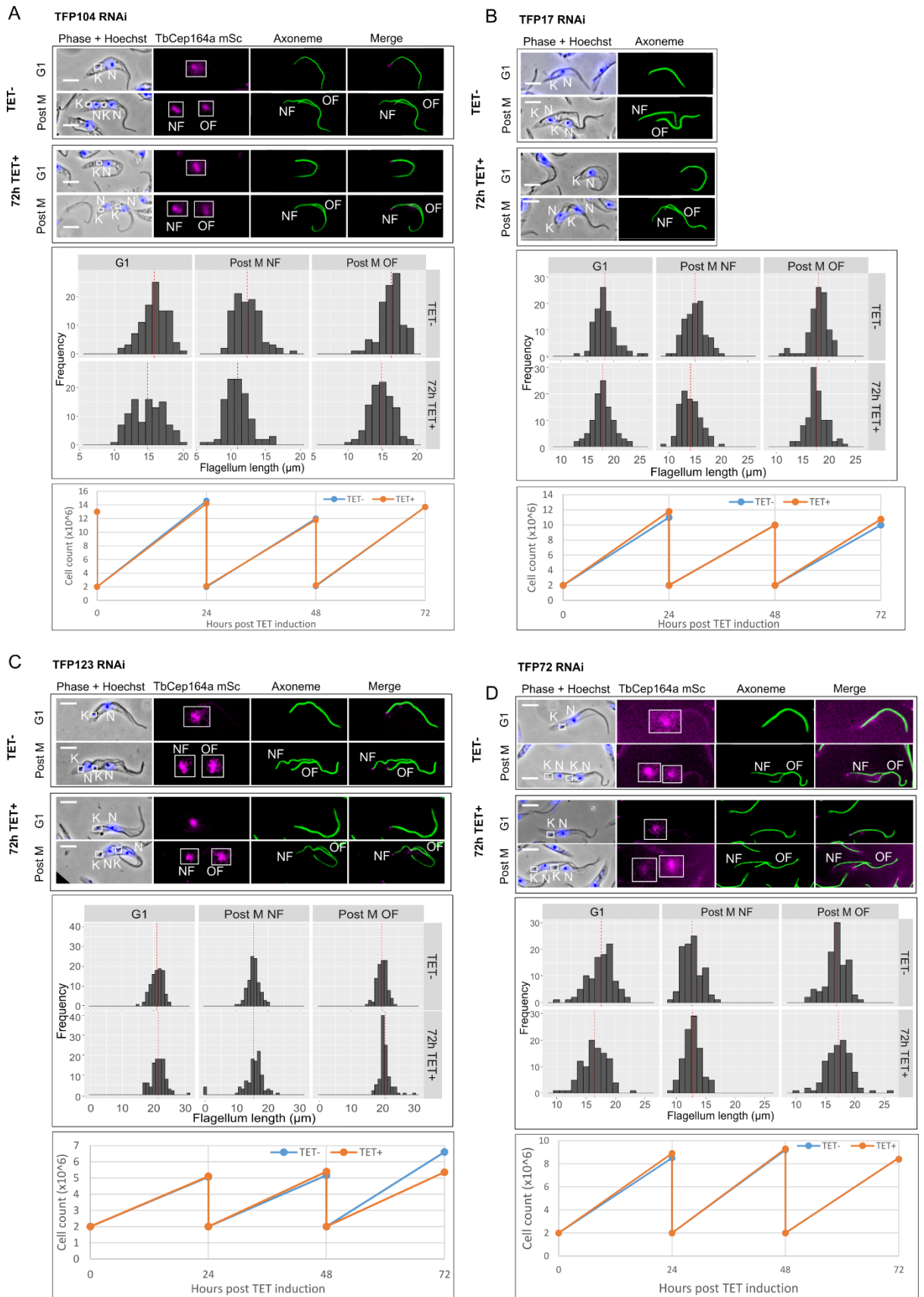


Figure 5.14 Immunolabelled IFT172 (magenta) with RNAi of transition fibre proteins in formaldehyde-methanol fixed procyclic form *Trypanosoma brucei*. Jitter and box plot (left) of the relative intensity of IFT172 at the flagellar base in uninduced cells (TET-) and cells following 24 hours induction of RNAi (24h TET+). (A) RNAi of CEP90 following 24 hours

induction, including jitter/box plot (left) and images of uninduced (TET-) and induced (24h TET+) cells labelled with IFT172 (right) (B) RNAi of OFD1 following 24 hours induction, including jitter/box plot (left) and images of uninduced (TET-) and induced (24h TET+) cells labelled with IFT172 (right). (C) RNAi of RABL2B following 24 hours induction, including jitter/box plot (left) and images of uninduced (TET-) and induced (24h TET+) cells labelled with IFT172 (right). (D) RNAi of CEP19L following 24 hours induction, including jitter/box plot (left) and images of uninduced (TET-) and induced (24h TET+) cells labelled with IFT172 (right). DNA is labelled with Hoechst (blue). Crop box represent a 5x zoom. New flagellum and old flagellum are labelled NF and OF respectively.

5.8 No discernible phenotype following knockdown of eight transition fibre proteins

We conclude that out of 20 transition fibre proteins taken forward for functional analysis, 6 proteins were required for flagellar assembly, a total of 2 proteins were required for the flagellum to reach their final length and 2 proteins showed a phenotype related to flagellar 'locking'. A further 2 proteins exhibited a phenotype that did not relate to flagellar length. For seven of the remaining transition fibre proteins (TFP104, TFP17, TFP123, TFP72, TFP52, TFP62 and TFP72) knockdown caused no significant changes in axoneme length and where applicable localisation of CEP164A and axoneme length (Fig. 5.15a-g). TFP123 (Fig. 5.15c) and TFP52 (Fig. 5.15e) exhibited a mild growth defect but no changes in flagellum length. In the case of ARL3C mAb25 staining of the flagellar axoneme was not clear enough to measure flagella reliably, however there were no clear changes in growth or CEP164A localisation following induction of RNAi (Fig. 5.15h).



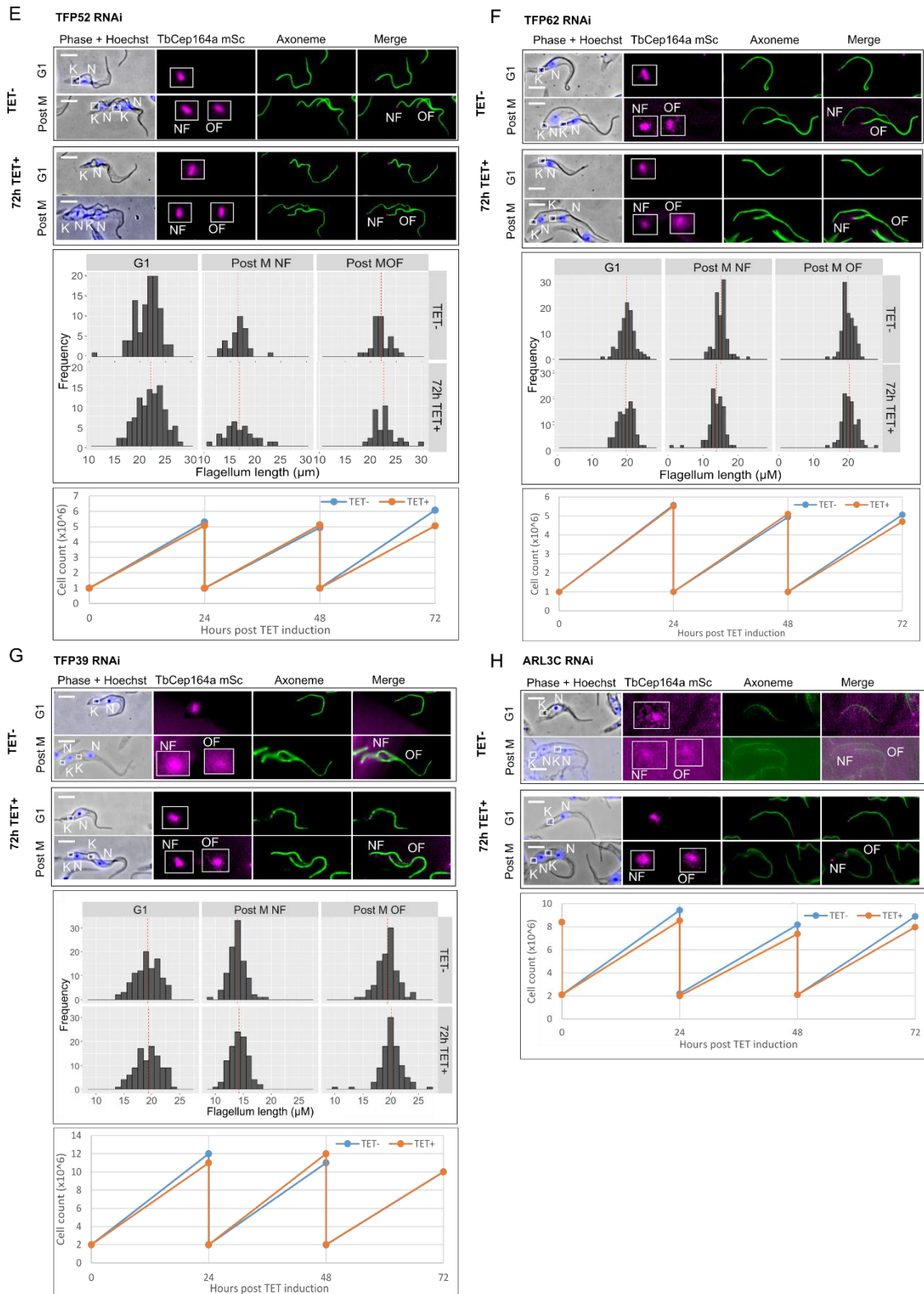


Figure 5.15 RNAi of transition fibre proteins exhibiting no discernible phenotype based on our analysis. Each shows Images of uninduced cells (TET-) and cells 72 hours post induction (72h TET+). CEP164AmSc shown as a 5x zoom (where applicable). Flagellum length analysis in G1 and Post M cells in uninduced cells (TET-) and cells following 72h induction of RNAi (72h TET+). Flagellar axoneme labelled with mAb25 monoclonal antibody. Growth curve shows cell density (2×10^6) measured every 24 hours for 72 hours. (A) TFP104 RNAi (B) TFP17 RNAi, (C) TFP123 RNAi (D) TFP72 RNAi, (E) TFP52 RNAi, (F) TFP62 RNAi, (G) TFP39 RNAi, (H) ARL3C RNAi

6. Discussion

6.1 The molecular composition and structure of transition fibres

6.1.1 Molecular composition

Centrioles and basal bodies are barrel shaped organelles that possess a highly conserved core structure (Hodges et al., 2010, Vaughan and Gull, 2015). However, transition fibres/distal appendages that surround the mature basal body have been predominantly examined in the context of mammalian cilia/centrioles. Transition fibres and distal appendages are equivalent structures and will be referred to as transition fibres in this discussion. To date only eight proteins have been localised to the structure (Bowler et al., 2019, Tanos et al., 2013, Yang et al., 2018). The five core components of so called distal appendages in mammals are recruited in a hierarchical manner as the cilium is assembled. CEP83 is recruited first to the root of the transition fibre blade followed by CEP89, SCLT1, FBF1 and finally CEP164, which is located at the outer region of the transition fibre (Tanos et al., 2013).

Our work has identified and confirmed 31 proteins that localise to the transition fibres in trypanosomes. Our results demonstrate that transition fibres are complex structures and, thus we reason that there are likely a number of additional transition fibre components in mammals that have not yet been identified. In addition to this, it is likely that there are more transition fibre proteins in trypanosomes that have not been identified due to failures (~30%) in co-expression of tagged genes during the screening of basal body components (Shahaan Shafiq, Jiri Tyc, Vaughan lab). We speculate that this may be due to the close proximity of two tagged proteins that prevents successful co-expression.

An automated FIJI script was developed to measure pseudo-diameter and distance from the mature basal body to discern the spatial arrangement of transition fibre proteins (Collaboration with Richard Wheeler, University of Oxford). Orthologues of two proteins within our cohort (CEP164 and FBF1) have also been mapped on mammalian transition fibres based on their diameter. These studies showed that FBF1 is narrower than CEP164. From our analysis, we show that in trypanosomes FBF1 is also narrower than CEP164 orthologues CEP164A and CEP164B (Chapter 3, Fig. 3.9). Based on these conserved known transition fibre proteins, our data is in concordance based their location relative to each other.

In contrast, CEP164 is the widest transition fibre proteins in mammals, but our analysis revealed that CEP164A and CEP164B have a pseudo-diameter in the mid-region of the transition fibre and there are many proteins wider than CEP164A and CEP164B. This may be because there are other

transition fibre proteins in mammals that are wider than CEP164 but have not been identified or simply that CEP164 orthologues in trypanosomes localise to a different part of the transition fibre. The widest protein measured on trypanosome transition fibres is TFP15, which has a pseudo-diameter of 345nm. Conserved proteins CEP90 and OFD1 have a pseudo-diameter greater than CEP164A and CEP164B. CEP90 and OFD1 proteins have not been mapped based on their diameter around the basal body in mammalian cells.

6.1.2 Comparison of the structural dimensions

TEM analysis of trypanosome transition fibres showed the mean distance between transition fibre head densities at the proximal end of the flagella pocket to be $368 \pm 10\text{nm}$ (Diameter measured laterally by EM) (Chapter 3, Fig. 3.10, mD), which was in concordance with our light microscopy data. Indeed, the widest protein in our cohort, TFP15, has a pseudo-diameter of 345nm by light microscopy. Previous scanning transmission electron microscopy (STEM) analysis carried out by Trepout and colleagues, estimated a cross sectional diameter in trypanosomes to be 330nm (Trepout et al., 2018). It is unclear exactly the reason for the slightly smaller diameter reported here. However, it should be noted that head densities were not reported in this study. We conclude from three independent methods that transition fibres in trypanosomes are between 330-370nm in diameter. Furthermore, EM analysis of mammalian transition fibres revealed a similar diameter at 378 nm (Bowler et al., 2019). The length of the transition fibre from the basal was measured to be 110 nm, which aligns closely with our measurement of 111 nm in trypanosomes (Bowler et al., 2019) (Chapter 3, Fig. 3.9).

The mammalian cell light microscopy work does conflict with EM analysis and will need further clarification. Measurements made of CEP164 using STORM (stochastic optical reconstruction microscopy) showed a diameter of $475 \pm 38\text{nm}$ (Yang et al., 2018) and a repeat of this using STORM in another study showed CEP164 diameter as $568 \pm 18\text{nm}$ (Bowler et al., 2019) which is more than 100nm wider than both our data and EM data of mammalian transition fibres.

The theoretical limit to the resolution of light microscopy is just below 200nm and so we acknowledge that mapping protein position at the nanometre scale may not be within the confines of our limit of resolution. However, for distance from the mature basal body (Chapter 3, Fig. 3.9) we are measuring the distance between a red and green point in two separate channels, and therefore the limit of resolution does not apply in this instance. For pseudo-diameter, we measured the length of the mNeonGreen signal in between 500-2000 cells for each protein, using an automated method of measurement. This means that there is no human error or variation in measurement and a larger sample size can be achieved compared to manual measurements. It

should also be noted that this method allows us to determine the approximate position of each protein along the transition fibre and not necessarily the exact location.

Previous studies have assessed the spatial arrangement of transition fibre proteins in mammals based on their cross-sectional diameter or 'radial occupancy' using super-resolution microscopy techniques such as STORM (stochastic optical reconstruction microscopy) and structured illumination microscopy (SIM) (Bowler et al., 2019, Tanos et al., 2013, Yang et al., 2018). However, immunolabelling of target proteins with an antibody is required for these techniques. An antibody has a molecular weight of ~150kDa and a length of ~15nm. The combined length of a primary and secondary antibody could be up to 30nm long (Opazo et al., 2012). This means a fluorophore may be up to 30nm from the target protein and could account for a significant amount of error at the nanometre scale. For accurate imaging at this level, the fluorophore should be as close as possible to the target protein and might explain the differences in the EM data versus STORM data.

The pPOT endogenous tagging method inserts sequence encoding the mNG fluorescent protein at the N-terminus of a gene of interest (Dean et al., 2015). The mNeonGreen protein is a homologue of the autofluorescent protein GFP (Green fluorescent protein) which possesses a cylinder shape with a length of 4.2nm and a diameter of 2.4nm (Hink et al., 2000, Steiert et al., 2018). By endogenous tagging, we deduce that the fluorescent protein may be less than 5nm away from our target protein. We therefore defend the use of our methodology on the premises that it is more scalable, there is minimal human error/variation in measurement and the fluorescent signal is significantly closer to protein of interest when compared to super-resolution techniques.

6.2 G1 to S-phase: modelling transition fibre assembly

In our study, the timing of transition fibre protein recruitment to the basal body was determined based on whether each protein arrived in G1 to the maturing probasal body, or in S-phase to the newly matured basal body following basal body duplication. We found that 17 proteins arrived before basal body duplication in the G1/S-phase transition and 14 proteins arrived after basal body duplication in S-phase.

The timing of arrival does not correlate with diameter in our case, i.e. later proteins don't necessarily have a wide diameter and vice versa. In the case of the five core proteins identified in mammals, narrower proteins are recruited earliest and wider proteins are recruited later in a hierarchical manner until the structure is fully formed (Tanos et al., 2013). However, our more extensive screen of 31 proteins does not support this correlation.

We suggest a different model for transition fibre assembly, with a core of proteins forming the structural base of the transition fibre and later proteins arriving at different positions on the transition fibre to support required functions such as membrane docking, flagellar assembly and length control. This model would indicate that transition fibres would have the capacity to be dynamic, with proteins being recruited and removed from the base structure as required.

CEP164 and FBF1 are the only two proteins within our cohort that have been studied in the context transition fibre assembly in mammals. Studies in mammals show FBF1 is recruited to the transition fibre before CEP164 (Tanos et al., 2013, Yang et al., 2018). Our analysis also shows that FBF1 is recruited to the transition fibre first (e.g. before basal body duplication) and CEP164A and CEP164B are both recruited later in S-phase. Therefore, our data shows that there is some conservation in the assembly process with respect to FBF1 and CEP164.

We also assessed whether proteins arrived before or after CEP164A in trypanosomes. Similar to mammals, CEP164A in trypanosomes was one of the latest proteins to arrive to the basal body with only TFP122 and CEP164C arriving later. We propose a model of assembly in trypanosomes where all identified transition fibre proteins (except CEP164C – see below) are present at the base of the existing G1 flagellum (Fig 6.1A). 17 proteins arrive to the maturing probasal body (Fig. 6.1b, orange) in G1 prior to basal body duplication (Fig 6.1b, TFP blue), followed by 12 more proteins in early S-phase (Fig 6.1c - red). We can conclude that transition fibre assembly occurs from G1 through to early S-phase. Figure 6.1 also clearly shows some of the later arriving proteins in red have a narrower diameter than the early recruited proteins in blue. Later in S-phase CEP164A is recruited (Fig 6.1c, magenta) followed by TFP122 (Fig. 6.1d, yellow).

Our previous work demonstrated that CEP164C is unlike the other transition fibre proteins outlined above. All of these arrive either before or very soon after basal body duplication at the start of the cell cycle. However, CEP164C only arrives in the next cell cycle to the base of the existing flagellum in G1 and thus this is the final transition fibre to be recruited. CEP164C is hypothesised to be important for regulating the growth of the old flagellum after it has reached its final length (Atkins et al., 2021).

(Top, S-phase). CEP16A (magenta) is recruited later in S-phase (Top, S-phase). TFP122 (yellow) arrives in S-phase after CEP164A. Position of each protein is based on pseudo-diameter and distance from the basal body (Chapter 3, Fig. 3.9).

6.3 The roles of transition fibres in flagellum assembly and length control

6.3.1 Transition fibres are central drivers of flagellum length control

One of the most fundamental questions related to ciliary biology is how the cilium or flagellum is able to faithfully assemble to and maintain a specific length. To understand this process, we may look into how ciliated eukaryotes are able to assemble and maintain two or more flagella within the same cell. In the simplest scenario, such is the case of the algae species *Chlamydomonas reinhardtii*, two seemingly identical flagella are assembled synchronously, where there is equal access to flagellum assembly proteins such as tubulin (Cavalier-Smith, 1974). The flagella are able to reach and maintain their final length by striking a balance between assembly and disassembly at the distal tip (Marshall and Rosenbaum, 2001). This is known as the balance point model, as flagella length is controlled by the balance of cargo carrying IFT complexes imported and loaded for assembly and disassembly at the distal tip of the flagellar axoneme.

Other ciliated organisms do not assemble their flagella synchronously, but instead assemble a new flagellum while maintaining the length of an existing flagellum. In the case of *Trypanosoma brucei*, a single new flagellum is assembled at the start of S-phase and continues its assembly until cytokinesis, while the existing old flagellum is maintained (see chapter 1, Fig. 1.5). Recently, a growth and lock model was proposed for trypanosomes which states that a lock is placed on the old flagellum so that increasing IFT rate in a locked flagellum does not lead to further elongation (Bertiaux et al., 2018). For this to be possible, there must be some molecular differences between the old and new flagellum to facilitate differential regulation between the two flagella. Assembly and locking must be regulated at two points; the flagellar base and the distal tip. Recent studies have implicated a role for the transition fibres in flagellum length control.

CEP164C is a protein that localises exclusively to the transition fibres of the old flagellum and functions as part of the old flagellum locking mechanism in trypanosomes (Atkins et al., 2021). Our analysis did not find another transition fibre protein where the signal was completely absent on the new flagellum, but we did identify four proteins (RABL2B, TFP104, TFP39 and TFP122) that were brighter at the base of the old throughout the cell cycle, suggesting they might be part of this mechanism.

Following knockdown of TFP122 by RNAi, the same phenotype was observed as CEP164C, with cells assembling longer flagella (Chapter 5, Fig. 5.10) (Atkins et al., 2021). TFP122 contains a kinesin motor head domain (Chapter 4, Fig. 4.12) and kinesin family proteins have previously been

implicated in flagellar length control in trypanosomes and *Leishmania*. Studies showed that knockdown of trypanosome kinesin-13 family protein, Kin13-2, resulted in cells that assemble longer flagella (Blaineau et al., 2007) and knockout of the Kin13-2 orthologue in *Leishmania major*, also resulted in cells that assembled longer flagella (Blaineau et al., 2007). Interestingly, Kin13-2 localises to the flagellum with a concentration of signal at the flagellar base in *Leishmania* (Blaineau et al., 2007). However, in trypanosomes Kif13-2 localises along the flagellum with an accumulation of signal instead at the distal tip of the flagellum (Chan and Ersfeld, 2010).

TFP122 also has a signal present within the old flagellum as well as the transition fibres (Chapter 3, Fig. 3.6). We conclude from this it is likely TFP122 functions to control flagellar length in trypanosomes, by functioning as part of the flagellum lock. TFP122 has an orthologue in other kinetoplastid species such as *Leishmania Mexicana* and therefore may be involved in flagellum locking in other kinetoplastids. Further studies involving TFP122 would be critical to understand its role in flagellum length regulation in more detail.

Another protein that appears to be involved in flagellum locking is TFP15, which has the widest pseudo-diameter around the basal body and is recruited to the transition fibres in S-phase. Unlike CEP164C and TFP122, the tagged TFP15 protein is present in equal intensity at both the old and new flagellum. However, similar to CEP164C and TFP122, knockdown of TFP15 resulted in cells with longer flagella by 72 hours post induction of RNAi (Chapter 5, Fig. 5.11). The phenotype suggests that TFP15 is another molecular player at the transition fibres driving flagellum length regulation. We therefore have three transition proteins that appear to be promoters of flagellum length control; TFP15, TFP122 and CEP164C. Out of these three proteins, we know that TFP15 is recruited to the basal body first in S-phase as it arrives before transition fibre marker CEP164A. TFP122 is recruited later in S-phase after the arrival of CEP164A. CEP164C returns to the base of the flagellum in G1 where transition fibres are fully assembled and therefore we can reason it would be the last of the three proteins to arrive to the transition fibres.

Our data also suggests that flagellum length control may be in itself regulated by a further set of proteins. Knockdown of TFP49 and CFAP52 by RNAi resulted in cells that were able to assemble a flagellum and cells had little/no defects in growth or cell morphology. However, following 72 hours' induction of RNAi, the flagella were few microns shorter than control cells (Chapter 5, Fig 5.8-5.9). The phenotype of these two proteins differs to the phenotype of flagellum assembly proteins that only form a very short flagellum that results in growth and morphological defects (see later).

One explanation for the flagellum not reaching the correct length is that it is caused by premature locking in CFAP52 and TFP49 depleted cells. This would indicate that CFAP52 and TFP49 are regulators of the locking mechanism, ensuring that the flagellum lock is activated at the correct time. CFAP52 and TFP49 arrive to the maturing probasal basal body in G1 prior to basal body duplication and before the arrival of locking proteins (TFP15, TFP122 and CEP164C). This would make sense if they are in fact required for the regulation of locking as if they arrived after locking proteins, the locking may be activated prematurely.

Based on the phenotype of transition fibre proteins involved in flagellum length regulation, we propose the following model (Fig. 6.2). TFP49 and CFAP52 arrive first to the maturing probasal body (Fig. 6.2, Orange) in G1 to prevent locking as this maturing probasal body will form the new flagellum later in the cycle during S-phase. Locking protein TFP15 then arrives to the outer region of the transition fibres at the mature basal body of the new flagellum (Fig. 6.2, left, green). Marker protein CEP164A arrives to the basal body of the new flagellum after TFP15 (Fig. 6.2, left, magenta) followed by the recruitment of locking protein TFP122 later in S-phase. These proteins remain at the transition fibres of the new flagellum until after cytokinesis in G1, where the daughter cells inherit the new flagellum (Fig. 6.2, G1, orange flagellum). In G1, the daughter cells inheriting the new flagellum recruit CEP164C so that all flagellum locking proteins are present at the transition fibres prior to new flagellum assembly in S-phase. This set of proteins are maintained around the basal body that will form the old flagellum, where locking must be implemented throughout the cell cycle (Fig. 6.2, right).

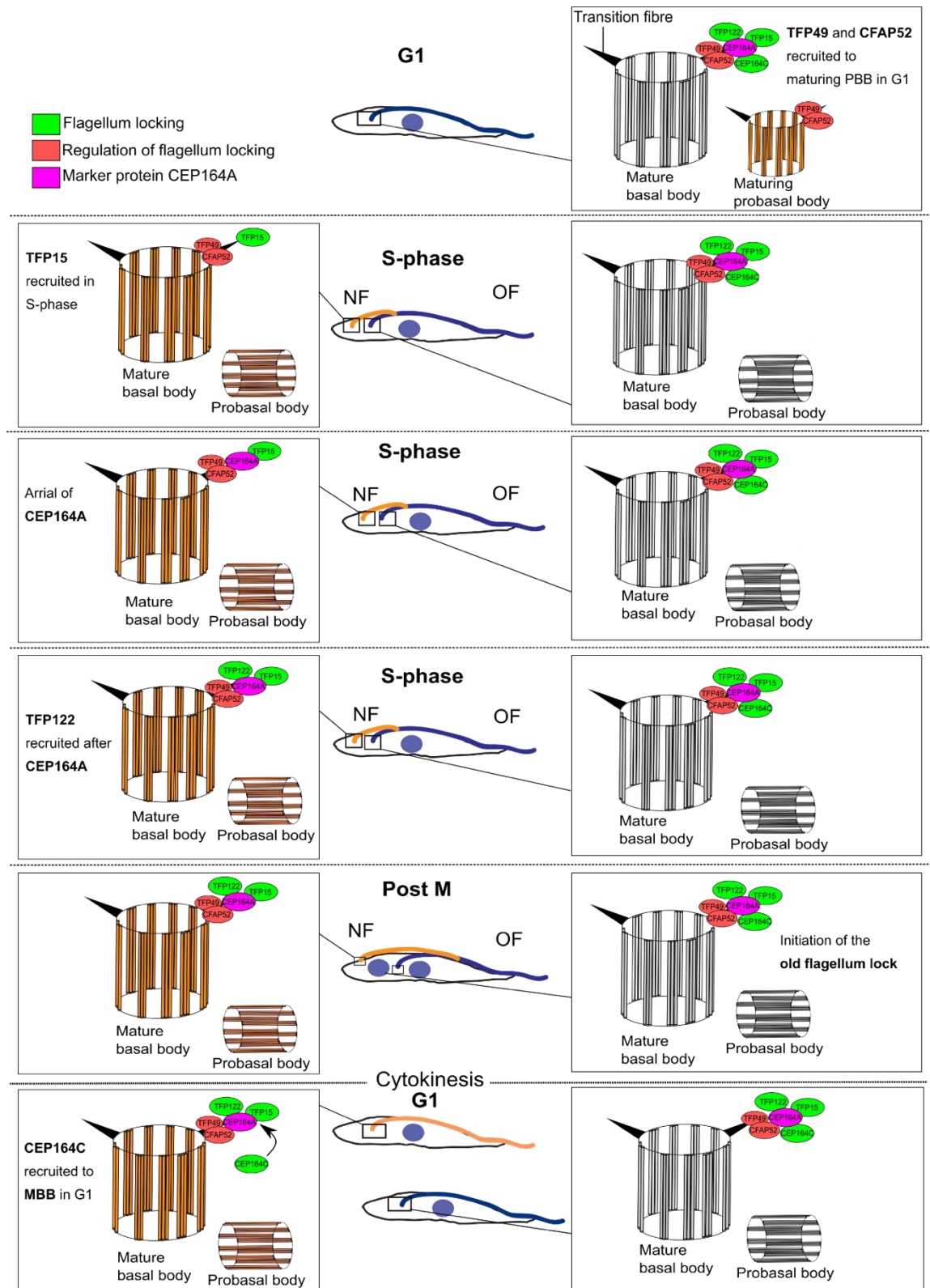


Figure 6.2 Model of transition fibre proteins involved in flagellum locking (green) and regulation of flagellum locking (red) at the new and old flagellum throughout the cell cycle. Transition fibres of the new flagellum are shown to the left (orange) and transition fibres of the old flagellum shown to the right (grey). Regulator proteins TFP49 and CFAP52 arrive to the maturing probasal body in G1. This maturing probasal body will become the mature basal body (MBB) of the new flagellum in S-phase. Following probasal body formation in S-phase, locking protein TFP15 is recruited to the wider

region of the transition fibre. CEP164A arrive after TFP15 in S-phase followed by TFP122. CEP164C arrived to mature basal of the daughter G1 cell inheriting the new flagellum. All locking proteins are present. The locking proteins are all maintained at the old flagellum throughout the cell cycle.

To understand whether these proteins interact, further studies would be required. This could be achieved by creating an RNAi matrix, labelling the possible downstream proteins and depleting by RNAi the upstream proteins. For example, TFP15 could be depleted by RNAi in cells lines tagged with TFP122 and CEP164C to examine whether their localisation is affected. If so, TFP122 could be depleted and the localisation of CEP164C could be examined. Ectopic expression of TFP122 and TFP15 could also be carried out to assess whether an increased abundance can induce premature locking and result in shorter flagella, similar to the phenotype observed with CEP164C (Atkins et al., 2021).

6.3.2 Transition fibres are critical for flagellum assembly in trypanosomes

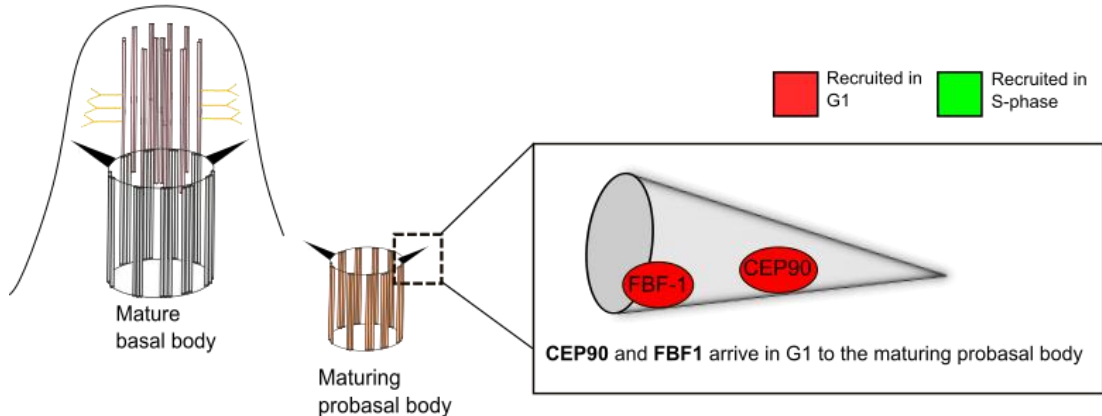
Following RNAi knockdown of 20 trypanosome transition fibre proteins, we identified six proteins (CEP90, TFP68, FBF1, CEP19L, OFD1 and RABL2B) that are critical for flagellum assembly. We determined that these proteins are essential for flagellum formation based on the observation of very short flagella combined with a growth and morphology defect following depletion by RNAi (Chapter 5, Fig. 5.2-5.7). Of these, CEP90, TFP68 and CEP19L were required for the localisation CEP164A suggesting they are also required for correct transition fibre formation. These three proteins all arrive to the transition fibre before CEP164A. Interestingly, out of the six flagellum assembly proteins, TFP68, CEP19L and CEP90 are located closest to CEP164A based on our quantitative analysis (Chapter 3, Fig. 3.9; Fig. 6.1). FBF1 and CEP90 are recruited in earlier in G1 whilst TFP68, CEP19L, OFD1 and RABL2B all arrive to the transition fibres later in S-phase.

Our data shows that these six proteins must be recruited to the transition fibres for successful flagellum formation. FBF1 and CEP90 arrive to the transition fibres in G1 before axoneme assembly. The remaining four flagellum assembly proteins (TFP68, CEP19L, OFD1 and RABL2B) were observed at the new mature basal body in all S-phase cells that have two basal body pairs. This includes the time point very shortly after basal body duplication. This indicates that axoneme assembly occurs in S-phase after the arrival of TFP68, CEP19L, OFD1 and RABL2B.

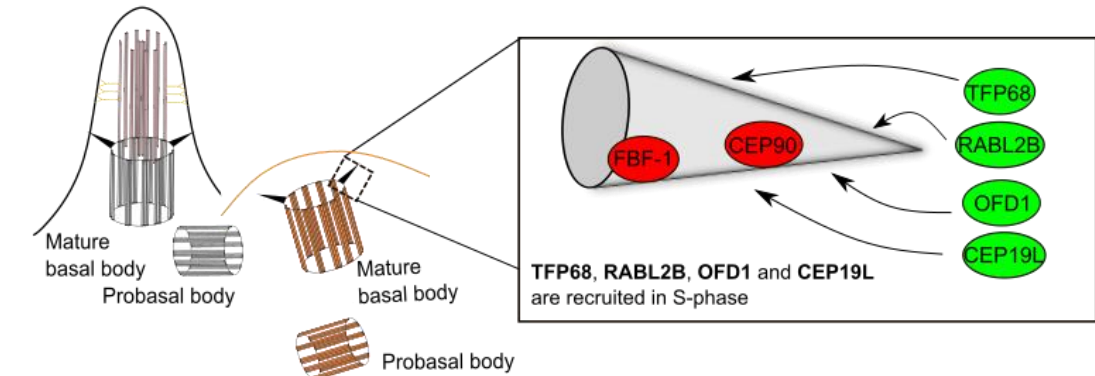
We thereby propose a model which describes the events that occur at the transition fibres that precede the formation of a new flagellum (Fig. 6.3). FBF1 and CEP90 are first recruited to the maturing probasal body in G1. The maturing probasal body then begins to invade the existing flagella pocket which is shortly followed by the formation of two new probasal bodies at the G1/S-phase border (Lacomble et al., 2010, Vaughan and Gull, 2015). Following basal body duplication,

TFP68, CEP19L, OFD1 and RABL2B are recruited to the new mature basal body which is finally followed by the assembly of the new flagellar axoneme (Fig. 6.3). Where applicable, the location of proteins is based on their pseudo-diameter and distance from the mature basal body.

EARLY FLAGELLUM ASSEMBLY PROTEINS ARE RECRUITED IN G1 -----



PROBASAL BODY SYNTHESIS AND POCKET INVASION AT G1/S BORDER -----



AXONEME ASSEMBLY IS INITIATED IN S-PHASE -----

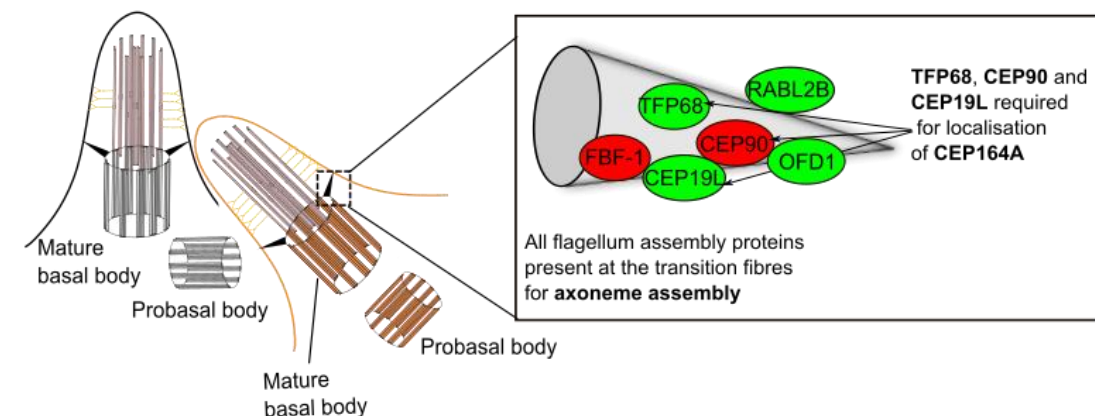


Figure 6.3 Model proposed of proteins driving flagellar assembly. FBF1 and CEP90 (Red) are recruited first to the maturing probasal body (orange) in G1 (top). Probasal body synthesis occurs in S-phase (middle) and the new mature basal body (orange1) invades the flagellar pocket and TFP68, RABL2B, OFD1 and CEP19L (green) are all recruited. All flagella assembly proteins are present at the transition fibres in S-phase (bottom) where axoneme assembly can be initiated.

Studies have shown that transition fibre protein CEP164 links machinery for ciliary membrane biogenesis to the mother centriole and therefore suggests a role for transition fibres in ciliary vesicle formation (Chaki et al., 2012, Graser et al., 2007, Liu et al., 2021). It is not known if trypanosomes also utilise vesicles for flagellum development, or whether the basal body tethers directly to the flagellar pocket membrane following probasal body assembly. Further analysis of basal body docking by TEM or super-resolution microscopy in trypanosomes could provide some insights into this question.

6.3.3 Transition fibres and IFT – insights into human ciliopathies

Studies have shown that intraflagellar transport (IFT) is essential for the correct assembly of a cilium or flagellum (Rosenbaum, 2002). The early stages of flagellum biogenesis depend upon three steps following assembly of the transition fibres. The recruitment of IFT to the flagellar base, the import of IFT into the flagellar compartment and loading of IFT onto the flagellar axoneme (Reiter, 2012, Wei et al., 2015, Wei et al., 2013). We depleted four transition fibre proteins (CEP90, OFD1, RABL2b and CEP19L) by RNAi and assesses the localisation of IFT protein IFT172 by immunofluorescence.

Depletion of CEP90 and OFD1 by RNAi resulted in a reduction in the intensity of IFT172 signal at the flagellar base/transition fibres and an intense IFT172 pool was observed at the proximal region of the axoneme (Chapter 5, Fig. 5.14a, Fig. 5.14b). We showed that IFT172 can be imported from cytoplasm into the flagellar compartment in OFD1 and CEP90 depleted cells, suggesting that OFD1 and CEP90 are not required for IFT recruitment or subsequent import into the flagellar compartment. However, the increase in IFT172 at the proximal region of the flagellum suggests that IFT172 is not loading correctly onto the flagellar axoneme. A recent study in mice showed that IFT88 was depleted at the mother centriole in CEP90 depleted cells, which supports our findings (Kumar et al., 2021).

Recessive mutations in the CEP90 gene in humans (also known as PIBF1) have been associated with the ciliopathies Jeune syndrome and Joubert syndrome (Shen et al., 2020, Wheway et al., 2015). Mutations in OFD1 is also associated with human ciliopathies such as Oral-facial-digital (Ofd1) syndrome type 1 (Sakakibara et al., 2019) and Joubert syndrome (Hammarsjo et al., 2017). Our data indicate a role for CEP90 and OFD1 in IFT loading in trypanosomes and, thus provide potentially novel insights into the mechanism of human ciliopathies. The result is particularly significant for OFD1 which has not been associated with IFT previously.

Depletion of RABL2B and CEP19L resulted in a modest increase in the intensity of IFT172 at the base of the flagellum, suggesting RABL2B and CEP19L may be required for import of IFT into the

flagellar compartment. Studies in mice have shown that RABL2 binds CEP19 and interacts with IFT Complex-B at the ciliary base (Nishijima et al., 2017). The study has suggested that RABL2 and CEP19 may regulate recruitment or assembly of IFT at the ciliary base although the study used a visible immunoprecipitation (VIP) to illustrate interaction but did not look at the intensity of IFT at the flagellar base to assess more specifically whether these proteins function in recruitment, import or loading. Our data suggest that RABL2B and CEP19L may be required for import of IFT172 in trypanosomes. The RABL2B gene has been implicated in male infertility causing immotile short tail sperm defects (Hosseini et al., 2017). While mutations in CEP19 have been associated with obesity, male infertility and intellectual defects in humans (Shalata et al., 2013, Yildiz Bolukbasi et al., 2018).

It is clear that CEP90, OFD1, RABL2B and CEP19 are associated with human ciliopathies or ciliopathy like syndromes. The mechanisms we are proposing may or may not be conserved in humans, however given the assessment of their localisation and function in humans, we think it likely but not certain that these mechanisms are also conserved. In summary, our findings suggest viable potential mechanisms relating to cilium/flagellum defects in human ciliopathies. We can conclude that we have leveraged the trypanosome model effectively to address broad biological questions that relate closely to human disease.

6.4 Conservation of transition fibre proteins across eukaryotic organisms

Cilia and flagella are highly conserved structures assembled from a microtubule-organising basal body. Basal bodies have a core structure that has been conserved throughout eukaryotic evolution (Dutcher and O'Toole, 2016, Hodges et al., 2010, Vaughan and Gull, 2015). However, the conservation of transition fibres across ciliated eukaryotes has not been assessed. We have identified a cohort of 31 transition fibre proteins and therefore carried out an analysis assessing their conservation in 45 eukaryotic organisms by reciprocal best blast (Chapter 4, Fig. 4.11-4.13; see methods, Fig. 2.5). Our data indicate there are seven transition fibre proteins (CFAP52, CEP90, CEP164, OFD1, RABL2B, CEP19 and FBF1) that are widely conserved across different eukaryotes and thereby likely make up a conserved core structure. It should be noted the in the case of TFP101 assigning orthology based on a coiled-coil SMC domain may be challenging, although according to our criteria TFP101 did produce a significant hit in *Plasmodium falciparum* and so this was kept in Fig.4.11.

Given that the ultrastructural description of mammalian transition fibres and trypanosome transition fibres are of a similar size by EM (and even greater by light microscopy), we consider it likely that mammalian transition fibres have not been fully characterised in terms of their

molecular composition. Our study indicates that CEP90, RABL2B, CEP19, CFAP52 may be four components of the mammalian transition fibre/distal appendage that have not been explicitly defined as components of the transition fibres although CEP90 has very recently been described as proteins required for distal appendage assembly (Kumar et al., 2021).

6.4.1 Localisation of conserved orthologues

Previous studies have localised CEP90 to centriolar satellites in mammals (Kim et al., 2012). It is reasonable to speculate that other transition fibre proteins may have been mislocalised, or have multiple localisations. RABL2B and CEP19 have been localised to the mature basal body but have not been explicitly annotated as components of the transition fibres (Kanie et al., 2017, Nishijima et al., 2017). CFAP52 has been localised to the mature and probasal body protein in mammalian sperm (Tapia Contreras and Hoyer-Fender, 2020). We show that CFAP52 is a transition fibre protein in trypanosomes that is required for flagella to assemble to the correct length (Chapter 3, Fig. 3.5; Chapter 5, Fig. 5.8). The mammalian Cullin 4b protein localises to the nucleus in humans, however at the level of the amino acid sequence, the trypanosome and human proteins were not recognised as orthologues, despite being annotated with the same name (Zou et al., 2009). There is no localisation data for the orthologues of TFP101 in the malaria parasite *P. falciparum* or TFP123 in the blooming algae species *A. anophagefferens*.

6.4.2 Kinetoplastid specific transition fibre components

Out of the 31 proteins analysed, 16 proteins had orthologues only in kinetoplastid species whilst five transition fibre proteins were specific to trypanosomes. Trypanosomes and Leishmania are unique in their ability to assemble a single new flagellum while maintaining an existing flagellum within the same cell (Bertiaux and Bastin, 2020, Bertiaux et al., 2018, Gluenz et al., 2010). We speculate that many of the specific proteins found only in trypanosomes and Leishmania could function for this purpose. TFP15 and TFP122 are both specific to kinetoplastid organisms and our data indicate that they are involved in old flagellum locking (Chapter 5, Fig. 5.10-5.11).

We conclude that transition fibres likely have a conserved core structure made up of the seven widely conserved transition fibre proteins in this cohort. However, transition fibres are also specialised structures, with unique proteins present that may support unique functions in kinetoplastids like simultaneous flagellum assembly and maintenance. For this mechanism to be viable, transition fibres must be dynamic in their composition throughout the cell cycle so that the transition fibres of existing flagellum can recruit proteins to regulate length and those of the newly formed flagellum recruit the proteins that can drive assembly.

6.5 Conclusions

In summary, we have identified 31 proteins localising to the transition fibres in trypanosomes and mapped out the position of these proteins by quantitative analysis of the light microscopy data. We then went on to determine the timing of their recruitment and propose that transition fibre assembly is primarily bi-phasic, with a cohort arriving to the maturing basal body in G1 and a second cohort arriving in S-phase. Figure 6.1 summarises the spatio-temporal arrangement of these proteins. Our discoveries have highlighted that the protein network of transition fibres is far more complex and intricate than ever thought previously.

We also showed that transition fibres are dynamic in their molecular composition, with proteins being targeted to the new and old flagellum over the cell cycle. This provides an explanation to how trypanosomes use transition fibre to assemble a flagellum while maintaining the length of an existing flagellum. We have identified a set of proteins involved in length control and a set involved in flagellar assembly. Our work demonstrates the essential roles transition fibres play in flagellum assembly and length regulation. Transition fibres are also directly involved in human ciliopathies and our findings may provide novel insights into the mechanism of some diseases.

We can conclude by saying that there is a fundamental biological question that asks how a cilium is able to assemble to a certain length. We believe this can be largely answered by the role of transition fibres at the ciliary base, dynamically regulating protein trafficking throughout the cell cycle. Many studies, for good reason, have focused attention on dynamics at the distal flagellar tip. But based on our findings, we think more focus should be shifted to transition fibres at the ciliary base. Given their fundamental roles in assembly and maintenance, it seems clear that transition fibres are understudied and underrepresented. Our results support that they are major players in ciliary function. We believe our work unveils new avenues of discussion for the future of ciliary biology and human diseases.

Bibliography

- ABAL, M., KERYER, G. & BORNENS, M. 2005. Centrioles resist forces applied on centrosomes during G2/M transition. *Biol Cell*, 97, 425-34.
- ALLEN, C. L., GOULDING, D. & FIELD, M. C. 2003. Clathrin-mediated endocytosis is essential in *Trypanosoma brucei*. *EMBO J*, 22, 4991-5002.
- ANDERSON, R. G. 1972. The three-dimensional structure of the basal body from the rhesus monkey oviduct. *J Cell Biol*, 54, 246-65.
- ANDRE, J., KERRY, L., QI, X., HAWKINS, E., DRIZYTE, K., GINGER, M. L. & MCKEAN, P. G. 2014. An alternative model for the role of RP2 protein in flagellum assembly in the African trypanosome. *J Biol Chem*, 289, 464-75.
- ATKINS, M., TYC, J., SHAFIQ, S., AHMED, M., BERTIAUX, E., DE CASTRO NETO, A. L., SUNTER, J., BASTIN, P., DEAN, S. D. & VAUGHAN, S. 2021. CEP164C regulates flagellum length in stable flagella. *J Cell Biol*, 220.
- AZIMZADEH, J., HERGERT, P., DELOUVEE, A., EUTENEUER, U., FORMSTECHE, E., KHODJAKOV, A. & BORNENS, M. 2009. hPOC5 is a centrin-binding protein required for assembly of full-length centrioles. *J Cell Biol*, 185, 101-14.
- BARON, D. M., RALSTON, K. S., KABUTUTU, Z. P. & HILL, K. L. 2007. Functional genomics in *Trypanosoma brucei* identifies evolutionarily conserved components of motile flagella. *J Cell Sci*, 120, 478-91.
- BELL, L. R., STONE, S., YOCHEM, J., SHAW, J. E. & HERMAN, R. K. 2006. The molecular identities of the *Caenorhabditis elegans* intraflagellar transport genes *dyf-6*, *daf-10* and *osm-1*. *Genetics*, 173, 1275-86.
- BELLOFATTO, V. & PALENCHAR, J. B. 2008. RNA interference as a genetic tool in trypanosomes. *Methods Mol Biol*, 442, 83-94.
- BENEKE, T., MADDEN, R., MAKIN, L., VALLI, J., SUNTER, J. & GLUENZ, E. 2017. A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. *R Soc Open Sci*, 4, 170095.
- BERTIAUX, E. & BASTIN, P. 2020. Dealing with several flagella in the same cell. *Cell Microbiol*, 22, e13162.
- BERTIAUX, E., MORGA, B., BLISNICK, T., ROTUREAU, B. & BASTIN, P. 2018. A Grow-and-Lock Model for the Control of Flagellum Length in Trypanosomes. *Curr Biol*, 28, 3802-3814 e3.

BESSIS, M. & BRETON-GORIUS, J. 1958. [On an pericentral inframicroscopic structure; electron microscopy of mammalian leukocytes]. *C R Hebd Seances Acad Sci*, 246, 1289-91.

BLAINEAU, C., TESSIER, M., DUBESSAY, P., TASSE, L., CROBU, L., PAGES, M. & BASTIEN, P. 2007. A novel microtubule-depolymerizing kinesin involved in length control of a eukaryotic flagellum. *Curr Biol*, 17, 778-82.

BLENKINSOP, T. A., SALERO, E., STERN, J. H. & TEMPLE, S. 2013. The culture and maintenance of functional retinal pigment epithelial monolayers from adult human eye. *Methods Mol Biol*, 945, 45-65.

BOWLER, M., KONG, D., SUN, S., NANJUNDAPPA, R., EVANS, L., FARMER, V., HOLLAND, A., MAHJOUB, M. R., SUI, H. & LONCAREK, J. 2019. High-resolution characterization of centriole distal appendage morphology and dynamics by correlative STORM and electron microscopy. *Nat Commun*, 10, 993.

CAJANEK, L. & NIGG, E. A. 2014. Cep164 triggers ciliogenesis by recruiting Tau tubulin kinase 2 to the mother centriole. *Proc Natl Acad Sci U S A*, 111, E2841-50.

CAO, S., WALKER, G. B., WANG, X., CUI, J. Z. & MATSUBARA, J. A. 2013. Altered cytokine profiles of human retinal pigment epithelium: oxidant injury and replicative senescence. *Mol Vis*, 19, 718-28.

CARVALHO-SANTOS, Z., AZIMZADEH, J., PEREIRA-LEAL, J. B. & BETTENCOURT-DIAS, M. 2011. Evolution: Tracing the origins of centrioles, cilia, and flagella. *J Cell Biol*, 194, 165-75.

CAVALIER-SMITH, T. 1974. Basal body and flagellar development during the vegetative cell cycle and the sexual cycle of *Chlamydomonas reinhardtii*. *J Cell Sci*, 16, 529-56.

CHAKI, M., AIRIK, R., GHOSH, A. K., GILES, R. H., CHEN, R., SLAATS, G. G., WANG, H., HURD, T. W., ZHOU, W., CLUCKEY, A., GEE, H. Y., RAMASWAMI, G., HONG, C. J., HAMILTON, B. A., CERVENKA, I., GANJI, R. S., BRYJA, V., ARTS, H. H., VAN REEUWIJK, J., OUD, M. M., LETTEBOER, S. J., ROEPMAN, R., HUSSON, H., IBRAGHIMOV-BESKROVNAYA, O., YASUNAGA, T., WALZ, G., ELEY, L., SAYER, J. A., SCHERMER, B., LIEBAU, M. C., BENZING, T., LE CORRE, S., DRUMMOND, I., JANSSEN, S., ALLEN, S. J., NATARAJAN, S., O'TOOLE, J. F., ATTANASIO, M., SAUNIER, S., ANTIGNAC, C., KOENEKOOP, R. K., REN, H., LOPEZ, I., NAYIR, A., STOETZEL, C., DOLLFUS, H., MASSOUDI, R., GLEESON, J. G., ANDREOLI, S. P., DOHERTY, D. G., LINDSTRAD, A., GOLZIO, C., KATSANIS, N., PAPE, L., ABBOUD, E. B., AL-RAJHI, A. A., LEWIS, R. A., OMRAN, H., LEE, E. Y., WANG, S., SEKIGUCHI, J. M., SAUNDERS, R., JOHNSON, C. A., GARNER, E., VANSELOW, K.,

- ANDERSEN, J. S., SHLOMAI, J., NURNBERG, G., NURNBERG, P., LEVY, S., SMOGORZEWSKA, A., OTTO, E. A. & HILDEBRANDT, F. 2012. Exome capture reveals ZNF423 and CEP164 mutations, linking renal ciliopathies to DNA damage response signaling. *Cell*, 150, 533-48.
- CHAN, K. Y. & ERSFELD, K. 2010. The role of the Kinesin-13 family protein TbKif13-2 in flagellar length control of *Trypanosoma brucei*. *Mol Biochem Parasitol*, 174, 137-40.
- CHEN, S. H., JANG, G. M., HUTTENHAIN, R., GORDON, D. E., DU, D., NEWTON, B. W., JOHNSON, J. R., HIATT, J., HULTQUIST, J. F., JOHNSON, T. L., LIU, Y. L., BURTON, L. A., YE, J., REICHERMEIER, K. M., STROUD, R. M., MARSON, A., DEBNATH, J., GROSS, J. D. & KROGAN, N. J. 2018. CRL4(AMBRA1) targets Elongin C for ubiquitination and degradation to modulate CRL5 signaling. *EMBO J*, 37.
- CHRISTENSEN, S. T., CLEMENT, C. A., SATIR, P. & PEDERSEN, L. B. 2012. Primary cilia and coordination of receptor tyrosine kinase (RTK) signalling. *J Pathol*, 226, 172-84.
- CHRISTENSEN, S. T., PEDERSEN, L. B., SCHNEIDER, L. & SATIR, P. 2007. Sensory cilia and integration of signal transduction in human health and disease. *Traffic*, 8, 97-109.
- DACHEUX, D., LANDREIN, N., THONNUS, M., GILBERT, G., SAHIN, A., WODRICH, H., ROBINSON, D. R. & BONHIVERS, M. 2012. A MAP6-related protein is present in protozoa and is involved in flagellum motility. *PLoS One*, 7, e31344.
- DANG, H. Q., ZHOU, Q., ROWLETT, V. W., HU, H., LEE, K. J., MARGOLIN, W. & LI, Z. 2017. Proximity Interactions among Basal Body Components in *Trypanosoma brucei* Identify Novel Regulators of Basal Body Biogenesis and Inheritance. *mBio*, 8.
- DAVIS, E. E. & KATSANIS, N. 2012. The ciliopathies: a transitional model into systems biology of human genetic disease. *Curr Opin Genet Dev*, 22, 290-303.
- DAWSON, S. C. & HOUSE, S. A. 2010. Life with eight flagella: flagellar assembly and division in *Giardia*. *Curr Opin Microbiol*, 13, 480-90.
- DEAN, S., MOREIRA-LEITE, F., VARGA, V. & GULL, K. 2016. Cilium transition zone proteome reveals compartmentalization and differential dynamics of ciliopathy complexes. *Proc Natl Acad Sci U S A*, 113, E5135-43.
- DEAN, S., SUNTER, J. D. & WHEELER, R. J. 2017. TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. *Trends Parasitol*, 33, 80-82.

- DEAN, S., SUNTER, J., WHEELER, R. J., HODKINSON, I., GLUENZ, E. & GULL, K. 2015. A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biol*, 5, 140197.
- DEANE, J. A., COLE, D. G., SEELEY, E. S., DIENER, D. R. & ROSENBAUM, J. L. 2001. Localization of intraflagellar transport protein IFT52 identifies basal body transitional fibers as the docking site for IFT particles. *Curr Biol*, 11, 1586-90.
- DEVLIN, L. A., RAMSBOTTOM, S. A., OVERMAN, L. M., LISGO, S. N., CLOWRY, G., MOLINARI, E., POWELL, L., MILES, C. G. & SAYER, J. A. 2020. Embryonic and foetal expression patterns of the ciliopathy gene CEP164. *PLoS One*, 15, e0221914.
- DING, X., FRAGOZA, R., SINGH, P., ZHANG, S., YU, H. & SCHIMENTI, J. C. 2020. Variants in RABL2A causing male infertility and ciliopathy. *Hum Mol Genet*, 29, 3402-3411.
- DUTCHER, S. K. & O'TOOLE, E. T. 2016. The basal bodies of *Chlamydomonas reinhardtii*. *Cilia*, 5, 18.
- EDDY, C. A., FLORES, J. J., ARCHER, D. R. & PAUERSTEIN, C. J. 1978. The role of cilia in fertility: an evaluation by selective microsurgical modification of the rabbit oviduct. *Am J Obstet Gynecol*, 132, 814-21.
- ENGSTLER, M., PFOHL, T., HERMINGHAUS, S., BOSHART, M., WIEGERTJES, G., HEDDERGOTT, N. & OVERATH, P. 2007. Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes. *Cell*, 131, 505-15.
- FAUBEL, R., WESTENDORF, C., BODENSCHATZ, E. & EICHELE, G. 2016. Cilia-based flow network in the brain ventricles. *Science*, 353, 176-8.
- FIELD, M. C. & CARRINGTON, M. 2009. The trypanosome flagellar pocket. *Nat Rev Microbiol*, 7, 775-86.
- FILARDY, A. A., GUIMARAES-PINTO, K., NUNES, M. P., ZUKERAM, K., FLIESS, L., PEREIRA, L., OLIVEIRA NASCIMENTO, D., CONDE, L. & MORROT, A. 2018. Human Kinetoplastid Protozoan Infections: Where Are We Going Next? *Front Immunol*, 9, 1493.
- FLIEGAUF, M., BENZING, T. & OMRAN, H. 2007. When cilia go bad: cilia defects and ciliopathies. *Nat Rev Mol Cell Biol*, 8, 880-93.
- FRANCIA, M. E., DUBREMETZ, J. F. & MORRISSETTE, N. S. 2015. Basal body structure and composition in the apicomplexans *Toxoplasma* and *Plasmodium*. *Cilia*, 5, 3.

- GANESAN, S., COMSTOCK, A. T. & SAJJAN, U. S. 2013. Barrier function of airway tract epithelium. *Tissue Barriers*, 1, e24997.
- GAO, J. M., QIAN, Z. Y., HIDE, G., LAI, D. H., LUN, Z. R. & WU, Z. D. 2020. Human African trypanosomiasis: the current situation in endemic regions and the risks for non-endemic regions from imported cases. *Parasitology*, 147, 922-931.
- GARCIA-GONZALO, F. R. & REITER, J. F. 2017. Open Sesame: How Transition Fibers and the Transition Zone Control Ciliary Composition. *Cold Spring Harb Perspect Biol*, 9.
- GLOVER, L., JUN, J. & HORN, D. 2011. Microhomology-mediated deletion and gene conversion in African trypanosomes. *Nucleic Acids Res*, 39, 1372-80.
- GLUENZ, E., GINGER, M. L. & MCKEAN, P. G. 2010. Flagellum assembly and function during the *Leishmania* life cycle. *Curr Opin Microbiol*, 13, 473-9.
- GLUENZ, E., POVELONES, M. L., ENGLUND, P. T. & GULL, K. 2011. The kinetoplast duplication cycle in *Trypanosoma brucei* is orchestrated by cytoskeleton-mediated cell morphogenesis. *Mol Cell Biol*, 31, 1012-21.
- GOETZ, S. C. & ANDERSON, K. V. 2010. The primary cilium: a signalling centre during vertebrate development. *Nat Rev Genet*, 11, 331-44.
- GONCALVES, J. & PELLETIER, L. 2017. The Ciliary Transition Zone: Finding the Pieces and Assembling the Gate. *Mol Cells*, 40, 243-253.
- GRASER, S., STIERHOF, Y. D., LAVOIE, S. B., GASSNER, O. S., LAMLA, S., LE CLECH, M. & NIGG, E. A. 2007. Cep164, a novel centriole appendage protein required for primary cilium formation. *J Cell Biol*, 179, 321-30.
- GRIFFITHS, S., PORTMAN, N., TAYLOR, P. R., GORDON, S., GINGER, M. L. & GULL, K. 2007. RNA interference mutant induction in vivo demonstrates the essential nature of trypanosome flagellar function during mammalian infection. *Eukaryot Cell*, 6, 1248-50.
- GUERRERO-SANTORO, J., KAPETANAKI, M. G., HSIEH, C. L., GORBACHINSKY, I., LEVINE, A. S. & RAPIC-OTRIN, V. 2008. The cullin 4B-based UV-damaged DNA-binding protein ligase binds to UV-damaged chromatin and ubiquitinates histone H2A. *Cancer Res*, 68, 5014-22.
- GUINOVAR, C., NAVIA, M. M., TANNER, M. & ALONSO, P. L. 2006. Malaria: burden of disease. *Curr Mol Med*, 6, 137-40.

HAMMARSJO, A., WANG, Z., VAZ, R., TAYLAN, F., SEDGHI, M., GIRISHA, K. M., CHITAYAT, D., NEETHUKRISHNA, K., SHANNON, P., GODOY, R., GOWRISHANKAR, K., LINDSTRAND, A., NASIRI, J., BAKTASHIAN, M., NEWTON, P. T., GUO, L., HOFMEISTER, W., PETTERSSON, M., CHAGIN, A. S., NISHIMURA, G., YAN, L., MATSUMOTO, N., NORDGREN, A., MIYAKE, N., GRIGELIONIENE, G. & IKEGAWA, S. 2017. Novel KIAA0753 mutations extend the phenotype of skeletal ciliopathies. *Sci Rep*, 7, 15585.

HAMMARTON, T. C. 2007. Cell cycle regulation in *Trypanosoma brucei*. *Mol Biochem Parasitol*, 153, 1-8.

HAMMARTON, T. C., CLARK, J., DOUGLAS, F., BOSHART, M. & MOTTRAM, J. C. 2003. Stage-specific differences in cell cycle control in *Trypanosoma brucei* revealed by RNA interference of a mitotic cyclin. *J Biol Chem*, 278, 22877-86.

HANRAHAN, O., WEBB, H., O'BYRNE, R., BRABAZON, E., TREUMAN, A., SUNTER, J.D., CARRINGTON, M. and VOORHEIS, H.P. (2009). The Glycosylphosphatidylinositol-PLC in *Trypanosoma brucei* Forms a Linear Array on the Exterior of the Flagellar Membrane Before and After Activation. *PLoS Pathogens*, 5(6), p.e1000468.

HAYASHI, H. & AKIYOSHI, B. 2018. Degradation of cyclin B is critical for nuclear division in *Trypanosoma brucei*. *Biol Open*, 7.

HIGA, L. A., YANG, X., ZHENG, J., BANKS, D., WU, M., GHOSH, P., SUN, H. & ZHANG, H. 2006. Involvement of CUL4 ubiquitin E3 ligases in regulating CDK inhibitors Dacapo/p27Kip1 and cyclin E degradation. *Cell Cycle*, 5, 71-7.

HILDEBRANDT, F., BENZING, T. & KATSANIS, N. 2011. Ciliopathies. *N Engl J Med*, 364, 1533-43.

HILTON, N. A., SLADEWSKI, T. E., PERRY, J. A., PATAKI, Z., SINCLAIR-DAVIS, A. N., MUNIZ, R. S., TRAN, H. L., WURSTER, J. I., SEO, J. & DE GRAFFENRIED, C. L. 2018. Identification of TOEFAZ1-interacting proteins reveals key regulators of *Trypanosoma brucei* cytokinesis. *Mol Microbiol*, 109, 306-326.

HINK, M. A., GRIEP, R. A., BORST, J. W., VAN HOEK, A., EPPINK, M. H., SCHOTS, A. & VISSER, A. J. 2000. Structural dynamics of green fluorescent protein alone and fused with a single chain Fv protein. *J Biol Chem*, 275, 17556-60.

HODGES, M. E., SCHEUMANN, N., WICKSTEAD, B., LANGDALE, J. A. & GULL, K. 2010. Reconstructing the evolutionary history of the centriole from protein components. *J Cell Sci*, 123, 1407-13.

- HOOG, J. L., LACOMBLE, S., O'TOOLE, E. T., HOENGER, A., MCINTOSH, J. R. & GULL, K. 2014. Modes of flagellar assembly in *Chlamydomonas reinhardtii* and *Trypanosoma brucei*. *Elife*, 3, e01479.
- HOSSEINI, S. H., SADIGHI GILANI, M. A., MEYBODI, A. M. & SABBAGHIAN, M. 2017. The impact of RABL2B gene (rs144944885) on human male infertility in patients with oligoasthenoteratozoospermia and immotile short tail sperm defects. *J Assist Reprod Genet*, 34, 505-510.
- HU, H., LIU, Y., ZHOU, Q., SIEGEL, S. & LI, Z. 2015. The Centriole Cartwheel Protein SAS-6 in *Trypanosoma brucei* Is Required for Probasal Body Biogenesis and Flagellum Assembly. *Eukaryot Cell*, 14, 898-907.
- HUANGFU, D., LIU, A., RAKEMAN, A. S., MURCIA, N. S., NISWANDER, L. & ANDERSON, K. V. 2003. Hedgehog signalling in the mouse requires intraflagellar transport proteins. *Nature*, 426, 83-7.
- INABA, K. 2003. Molecular architecture of the sperm flagella: molecules for motility and signaling. *Zoolog Sci*, 20, 1043-56.
- INOUE, M., NAKAMURA, Y., YASUDA, K., YASAKA, N., HARA, T., SCHNAUFER, A., STUART, K. & FUKUMA, T. 2005. The 14-3-3 proteins of *Trypanosoma brucei* function in motility, cytokinesis, and cell cycle. *J Biol Chem*, 280, 14085-96.
- ISHIKAWA, H. & MARSHALL, W. F. 2011. Ciliogenesis: building the cell's antenna. *Nature Reviews Molecular Cell Biology*, 12, 222.
- KANIE, T., ABBOTT, K. L., MOONEY, N. A., PLOWEY, E. D., DEMETER, J. & JACKSON, P. K. 2017. The CEP19-RABL2 GTPase Complex Binds IFT-B to Initiate Intraflagellar Transport at the Ciliary Base. *Dev Cell*, 42, 22-36 e12.
- KIM, D. H., AHN, J. S., HAN, H. J., KIM, H. M., HWANG, J., LEE, K. H., CHA-MOLSTAD, H., RYOO, I. J., JANG, J. H., KO, S. K., YANG, J. O., LEE, H. G., LEE, S., SONG, E. J., KIM, J. Y., HUH, Y. H., KWON, Y. T., SOUNG, N. K. & KIM, B. Y. 2019. Cep131 overexpression promotes centrosome amplification and colon cancer progression by regulating Plk4 stability. *Cell Death Dis*, 10, 570.
- KIM, K., LEE, K. & RHEE, K. 2012. CEP90 is required for the assembly and centrosomal accumulation of centriolar satellites, which is essential for primary cilia formation. *PLoS One*, 7, e48196.

- KOHL, L., ROBINSON, D. & BASTIN, P. 2003. Novel roles for the flagellum in cell morphogenesis and cytokinesis of trypanosomes. *EMBO J*, 22, 5336-46.
- KUMAR, D., RAINS, A., HERRANZ-PEREZ, V., LU, Q., SHI, X., SWANEY, D. L., STEVENSON, E., KROGAN, N. J., HUANG, B., WESTLAKE, C., GARCIA-VERDUGO, J. M., YODER, B. K. & REITER, J. F. 2021. A ciliopathy complex builds distal appendages to initiate ciliogenesis. *J Cell Biol*, 220.
- LACOMBLE, S., VAUGHAN, S., GADELHA, C., MORPHEW, M. K., SHAW, M. K., MCINTOSH, J. R. & GULL, K. 2010. Basal body movements orchestrate membrane organelle division and cell morphogenesis in *Trypanosoma brucei*. *J Cell Sci*, 123, 2884-91.
- LANGOUSIS, G. & HILL, K. L. 2014. Motility and more: the flagellum of *Trypanosoma brucei*. *Nat Rev Microbiol*, 12, 505-18.
- LECHTRECK, K. F., JOHNSON, E. C., SAKAI, T., COCHRAN, D., BALLIF, B. A., RUSH, J., PAZOUR, G. J., IKEBE, M. & WITMAN, G. B. 2009. The *Chlamydomonas reinhardtii* BBSome is an IFT cargo required for export of specific signaling proteins from flagella. *J Cell Biol*, 187, 1117-32.
- LI, Y. & HU, J. 2011. Small GTPases and cilia. *Protein Cell*, 2, 13-25.
- LIU, Q., JING, W., KANG, L., LIU, J. & LIU, M. 2021. Trends of the global, regional and national incidence of malaria in 204 countries from 1990 to 2019 and implications for malaria prevention. *J Travel Med*, 28.
- LIU, Y., HU, H. & LI, Z. 2013. The cooperative roles of PHO80-like cyclins in regulating the G1/S transition and posterior cytoskeletal morphogenesis in *Trypanosoma brucei*. *Mol Microbiol*, 90, 130-46.
- LU, Q., INSINNA, C., OTT, C., STAUFFER, J., PINTADO, P. A., RAHAJENG, J., BAXA, U., WALIA, V., CUENCA, A., HWANG, Y. S., DAAR, I. O., LOPES, S., LIPPINCOTT-SCHWARTZ, J., JACKSON, P. K., CAPLAN, S. & WESTLAKE, C. J. 2015. Early steps in primary cilium assembly require EHD1/EHD3-dependent ciliary vesicle formation. *Nat Cell Biol*, 17, 531.
- LYONS, R. A., SARIDOGAN, E. & DJAHANBAKHCH, O. 2006. The reproductive significance of human Fallopian tube cilia. *Hum Reprod Update*, 12, 363-72.
- MARIN, P. A., DA SILVA, M. S., PAVANI, R. S., MACHADO, C. R. & ELIAS, M. C. 2018. Recruitment kinetics of the homologous recombination pathway in procyclic forms of *Trypanosoma brucei* after ionizing radiation treatment. *Sci Rep*, 8, 5405.

- MARSHALL, W. F. & ROSENBAUM, J. L. 2001. Intraflagellar transport balances continuous turnover of outer doublet microtubules: implications for flagellar length control. *J Cell Biol*, 155, 405-14.
- MARSHALL, W. F. 2007. What is the function of centrioles? *J Cell Biochem*, 100, 916-22.
- MATTHEWS, K. R. 2005. The developmental cell biology of *Trypanosoma brucei*. *J Cell Sci*, 118, 283-90.
- MITCHELL, D. R. 2017. Evolution of Cilia. *Cold Spring Harb Perspect Biol*, 9.
- MOHRI, H. 1968. Amino-acid composition of "Tubulin" constituting microtubules of sperm flagella. *Nature*, 217, 1053-4.
- MUHANGUZI, D., MUGENYI, A., BIGIRWA, G., KAMUSIIME, M., KITIBWA, A., AKURUT, G. G., OCHWO, S., AMANYIRE, W., OKECH, S. G., HATTENDORF, J. & TWEYONGYERE, R. 2017. African animal trypanosomiasis as a constraint to livestock health and production in Karamoja region: a detailed qualitative and quantitative assessment. *BMC Vet Res*, 13, 355.
- NACHURY, M. V., LOKTEV, A. V., ZHANG, Q., WESTLAKE, C. J., PERANEN, J., MERDES, A., SLUSARSKI, D. C., SCHELLER, R. H., BAZAN, J. F., SHEFFIELD, V. C. & JACKSON, P. K. 2007. A core complex of BBS proteins cooperates with the GTPase Rab8 to promote ciliary membrane biogenesis. *Cell*, 129, 1201-13.
- NAKAZAWA, Y., HIRAKI, M., KAMIYA, R. & HIRONO, M. 2007. SAS-6 is a cartwheel protein that establishes the 9-fold symmetry of the centriole. *Curr Biol*, 17, 2169-74.
- NELSON, D. J. & WRIGHT, E. M. 1974. The distribution, activity, and function of the cilia in the frog brain. *J Physiol*, 243, 63-78.
- NISHIHARA, D., YAJIMA, I., TABATA, H., NAKAI, M., TSUKIJI, N., KATAHIRA, T., TAKEDA, K., SHIBAHARA, S., NAKAMURA, H. & YAMAMOTO, H. 2012. Otx2 is involved in the regional specification of the developing retinal pigment epithelium by preventing the expression of sox2 and fgf8, factors that induce neural retina differentiation. *PLoS One*, 7, e48879.
- NISHIJIMA, Y., HAGIYA, Y., KUBO, T., TAKEI, R., KATOH, Y. & NAKAYAMA, K. 2017. RABL2 interacts with the intraflagellar transport-B complex and CEP19 and participates in ciliary assembly. *Mol Biol Cell*, 28, 1652-1666.
- NOSALA, C. & DAWSON, S. C. 2015. The Critical Role of the Cytoskeleton in the Pathogenesis of *Giardia*. *Curr Clin Microbiol Rep*, 2, 155-162.

- OCAMPO, M., CURTIDOR, H., VERA, R., VALBUENA, J. J., RODRIGUEZ, L. E., PUENTES, A., LOPEZ, R., GARCIA, J. E., TOVAR, D., PACHECO, P., NAVARRO, M. A. & PATARROYO, M. E. 2004. MAEBL Plasmodium falciparum protein peptides bind specifically to erythrocytes and inhibit in vitro merozoite invasion. *Biochem Biophys Res Commun*, 315, 319-29.
- ODENIRAN, P. O. & ADEMOLA, I. O. 2018. A meta-analysis of the prevalence of African animal trypanosomiasis in Nigeria from 1960 to 2017. *Parasit Vectors*, 11, 280.
- OGBADOYI, E.O., ROBINSON, D.R. and GULL, K. (2003). A High-Order Trans-Membrane Structural Linkage Is Responsible for Mitochondrial Genome Positioning and Segregation by Flagellar Basal Bodies in Trypanosomes. *Molecular Biology of the Cell*, 14(5), pp.1769–1779.
- O'HARA, P. T. 1970. Spiral tilt of triplet fibers in human leukocyte centrioles. *J Ultrastruct Res*, 31, 195-8.
- OLSTAD, E. W., RINGERS, C., HANSEN, J. N., WENS, A., BRANDT, C., WACHTEN, D., YAKSI, E. & JURISCH-YAKSI, N. 2019. Ciliary Beating Compartmentalizes Cerebrospinal Fluid Flow in the Brain and Regulates Ventricular Development. *Curr Biol*, 29, 229-241 e6.
- OPAZO, F., LEVY, M., BYROM, M., SCHAFER, C., GEISLER, C., GROEMER, T. W., ELLINGTON, A. D. & RIZZOLI, S. O. 2012. Aptamers as potential tools for super-resolution microscopy. *Nat Methods*, 9, 938-9.
- O'TOOLE, E. T. & DUTCHER, S. K. 2014. Site-specific basal body duplication in *Chlamydomonas*. *Cytoskeleton (Hoboken)*, 71, 108-18.
- QIN, H., DIENER, D. R., GEIMER, S., COLE, D. G. & ROSENBAUM, J. L. 2004. Intraflagellar transport (IFT) cargo: IFT transports flagellar precursors to the tip and turnover products to the cell body. *J Cell Biol*, 164, 255-66.
- REITER, B., LEROUX 2012. The base of the cilium: roles for transition fibres and the transition zone in ciliary formation, maintenance and compartmentalization. *EMBO*, 13.
- REITER, J. F. & LEROUX, M. R. 2017. Genes and molecular pathways underpinning ciliopathies. *Nat Rev Mol Cell Biol*, 18, 533-547.
- ROBINSON, D. R. & GULL, K. 1991. Basal body movements as a mechanism for mitochondrial genome segregation in the trypanosome cell cycle. *Nature*, 352, 731-3.
- ROSENBAUM, J. 2002. Intraflagellar transport. *Curr Biol*, 12, R125.

- SAKAKIBARA, N., MORISADA, N., NOZU, K., NAGATANI, K., OHTA, T., SHIMIZU, J., WADA, T., SHIMA, Y., YAMAMURA, T., MINAMIKAWA, S., FUJIMURA, J., HORINOCHI, T., NAGANO, C., SHONO, A., YE, M. J., NOZU, Y., NAKANISHI, K. & IJIMA, K. 2019. Clinical spectrum of male patients with OFD1 mutations. *J Hum Genet*, 64, 3-9.
- SANCHEZ, I. & DYNLACHT, B. D. 2016. Cilium assembly and disassembly. *Nat Cell Biol*, 18, 711-7.
- SATIR, P., MITCHELL, D. R. & JEKELY, G. 2008. How did the cilium evolve? *Curr Top Dev Biol*, 85, 63-82.
- SCHMIDT, K. N., KUHNS, S., NEUNER, A., HUB, B., ZENTGRAF, H. & PEREIRA, G. 2012. Cep164 mediates vesicular docking to the mother centriole during early steps of ciliogenesis. *J Cell Biol*, 199, 1083-101.
- SHALATA, A., RAMIREZ, M. C., DESNICK, R. J., PRIEDIGKEIT, N., BUETTNER, C., LINDTNER, C., MAHROUM, M., ABDUL-GHANI, M., DONG, F., ARAR, N., CAMACHO-VANEGAS, O., ZHANG, R., CAMACHO, S. C., CHEN, Y., IBDAH, M., DEFRONZO, R., GILLESPIE, V., KELLEY, K., DYNLACHT, B. D., KIM, S., GLUCKSMAN, M. J., BOROCHOWITZ, Z. U. & MARTIGNETTI, J. A. 2013. Morbid obesity resulting from inactivation of the ciliary protein CEP19 in humans and mice. *Am J Hum Genet*, 93, 1061-71.
- SHAW, S., DEMARCO, S.F., REHMANN, R., WENZLER, T., FLORINI, F., RODITI, I. and HILL, K.L. (2019). Flagellar cAMP signaling controls trypanosome progression through host tissues. *Nature Communications*, [online] 10(1). Available at: <https://www.nature.com/articles/s41467-019-08696-y>.
- SHEN, Y., WANG, H., LIU, Z., LUO, M., MA, S., LU, C., CAO, Z., YU, Y., CAI, R., CHEN, C., LI, Q., GAO, H., PENG, Y., XU, B. & MA, X. 2020. Identification of two novel pathogenic variants of PIBF1 by whole exome sequencing in a 2-year-old boy with Joubert syndrome. *BMC Med Genet*, 21, 192.
- SHERWIN, T. & GULL, K. 1989. The cell division cycle of *Trypanosoma brucei brucei*: timing of event markers and cytoskeletal modulations. *Philos Trans R Soc Lond B Biol Sci*, 323, 573-88.
- SILLIBOURNE, J. E., SPECHT, C. G., IZEDDIN, I., HURBAIN, I., TRAN, P., TRILLER, A., DARZACQ, X., DAHAN, M. & BORNENS, M. 2011. Assessing the localization of centrosomal proteins by PALM/STORM nanoscopy. *Cytoskeleton (Hoboken)*, 68, 619-27.

- SIMARRO, P. P., CECCHI, G., PAONE, M., FRANCO, J. R., DIARRA, A., RUIZ, J. A., FEVRE, E. M., COURTIN, F., MATTIOLI, R. C. & JANNIN, J. G. 2010. The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. *Int J Health Geogr*, 9, 57.
- SINDEN, R. E., TALMAN, A., MARQUES, S. R., WASS, M. N. & STERNBERG, M. J. 2010. The flagellum in malarial parasites. *Curr Opin Microbiol*, 13, 491-500.
- SINGLA, V., ROMAGUERA-ROS, M., GARCIA-VERDUGO, J. M. & REITER, J. F. 2010. *Odf1*, a human disease gene, regulates the length and distal structure of centrioles. *Dev Cell*, 18, 410-24.
- SMITH, T. F., GAITATZES, C., SAXENA, K. & NEER, E. J. 1999. The WD repeat: a common architecture for diverse functions. *Trends Biochem Sci*, 24, 181-5.
- STEIERT, F., PETROV, E. P., SCHULTZ, P., SCHWILLE, P. & WEIDEMANN, T. 2018. Photophysical Behavior of mNeonGreen, an Evolutionarily Distant Green Fluorescent Protein. *Biophys J*, 114, 2419-2431.
- STEPHAN, A., VAUGHAN, S., SHAW, M. K., GULL, K. & MCKEAN, P. G. 2007. An essential quality control mechanism at the eukaryotic basal body prior to intraflagellar transport. *Traffic*, 8, 1323-30.
- STUART, K., BRUN, R., CROFT, S., FAIRLAMB, A., GURTNER, R. E., MCKERROW, J., REED, S. & TARLETON, R. 2008. Kinetoplastids: related protozoan pathogens, different diseases. *J Clin Invest*, 118, 1301-10.
- SZOLLOSI, D. 1964. The Structure and Function of Centrioles and Their Satellites in the Jellyfish *Phialidium Gregarium*. *J Cell Biol*, 21, 465-79.
- TANOS, B. E., YANG, H. J., SONI, R., WANG, W. J., MACALUSO, F. P., ASARA, J. M. & TSOU, M. F. 2013. Centriole distal appendages promote membrane docking, leading to cilia initiation. *Genes Dev*, 27, 163-8.
- TAPIA CONTRERAS, C. & HOYER-FENDER, S. 2020. The WD40-protein CFAP52/WDR16 is a centrosome/basal body protein and localizes to the manchette and the flagellum in male germ cells. *Sci Rep*, 10, 14240.
- TASCHNER, M., WEBER, K., MOURAO, A., VETTER, M., AWASTHI, M., STIEGLER, M., BHOGARAJU, S. & LORENTZEN, E. 2016. Intraflagellar transport proteins 172, 80, 57, 54, 38, and 20 form a stable tubulin-binding IFT-B2 complex. *EMBO J*, 35, 773-90.

- TREPOUT, S., TASSIN, A. M., MARCO, S. & BASTIN, P. 2018. STEM tomography analysis of the trypanosome transition zone. *J Struct Biol*, 202, 51-60.
- VAUGHAN, S. & GULL, K. 2015. Basal body structure and cell cycle-dependent biogenesis in *Trypanosoma brucei*. *Cilia*, 5, 5.
- VAUGHAN, S. 2010. Assembly of the flagellum and its role in cell morphogenesis in *Trypanosoma brucei*. *Curr Opin Microbiol*, 13, 453-8.
- VELAND, I. R., AWAN, A., PEDERSEN, L. B., YODER, B. K. & CHRISTENSEN, S. T. 2009. Primary cilia and signaling pathways in mammalian development, health and disease. *Nephron Physiol*, 111, p39-53.
- VICKERMAN, K. 1985. Developmental cycles and biology of pathogenic trypanosomes. *Br Med Bull*, 41, 105-14.
- WADE, R. H. & HYMAN, A. A. 1997. Microtubule structure and dynamics. *Curr Opin Cell Biol*, 9, 12-7.
- WALLINGFORD, J. B. 2010. Planar cell polarity signaling, cilia and polarized ciliary beating. *Curr Opin Cell Biol*, 22, 597-604.
- WALLMEIER, J., NIELSEN, K. G., KUEHNI, C. E., LUCAS, J. S., LEIGH, M. W., ZARIWALA, M. A. & OMRAN, H. 2020. Motile ciliopathies. *Nat Rev Dis Primers*, 6, 77.
- WAN, K. Y. & JEKELY, G. 2020. On the unity and diversity of cilia. *Philos Trans R Soc Lond B Biol Sci*, 375, 20190148.
- WEHLAND, J. & WILLINGHAM, M. C. 1983. A rat monoclonal antibody reacting specifically with the tyrosylated form of alpha-tubulin. II. Effects on cell movement, organization of microtubules, and intermediate filaments, and arrangement of Golgi elements. *J Cell Biol*, 97, 1476-90.
- WEHLAND, J., SCHRODER, H. C. & WEBER, K. 1984. Amino acid sequence requirements in the epitope recognized by the alpha-tubulin-specific rat monoclonal antibody YL 1/2. *EMBO J*, 3, 1295-300.
- WEI, Q., LING, K. & HU, J. 2015. The essential roles of transition fibers in the context of cilia. *Curr Opin Cell Biol*, 35, 98-105.
- WEI, Q., LING, K. & HU, J. 2015. The essential roles of transition fibers in the context of cilia. *Curr Opin Cell Biol*, 35, 98-105.

WEI, Q., XU, Q., ZHANG, Y., LI, Y., ZHANG, Q., HU, Z., HARRIS, P. C., TORRES, V. E., LING, K. & HU, J. 2013. Transition fibre protein FBF1 is required for the ciliary entry of assembled intraflagellar transport complexes. *Nat Commun*, 4, 2750.

WHEWAY, G., SCHMIDTS, M., MANS, D. A., SZYMANSKA, K., NGUYEN, T. T., RACHER, H., PHELPS, I. G., TOEDT, G., KENNEDY, J., WUNDERLICH, K. A., SORUSCH, N., ABDELHAMED, Z. A., NATARAJAN, S., HERRIDGE, W., VAN REEUWIJK, J., HORN, N., BOLDT, K., PARRY, D. A., LETTEBOER, S. J. F., ROOSING, S., ADAMS, M., BELL, S. M., BOND, J., HIGGINS, J., MORRISON, E. E., TOMLINSON, D. C., SLAATS, G. G., VAN DAM, T. J. P., HUANG, L., KESSLER, K., GIESSL, A., LOGAN, C. V., BOYLE, E. A., SHENDURE, J., ANAZI, S., ALDAHMEH, M., AL HAZZAA, S., HEGELE, R. A., OBER, C., FROSK, P., MHANNI, A. A., CHODIRKER, B. N., CHUDLEY, A. E., LAMONT, R., BERNIER, F. P., BEAULIEU, C. L., GORDON, P., PON, R. T., DONAHUE, C., BARKOVICH, A. J., WOLF, L., TOOMES, C., THIEL, C. T., BOYCOTT, K. M., MCKIBBIN, M., INGLEHEARN, C. F., CONSORTIUM, U. K., UNIVERSITY OF WASHINGTON CENTER FOR MENDELIAN, G., STEWART, F., OMRAN, H., HUYNEN, M. A., SERGOUNIOTIS, P. I., ALKURAYA, F. S., PARBOOSINGH, J. S., INNES, A. M., WILLOUGHBY, C. E., GILES, R. H., WEBSTER, A. R., UEFFING, M., BLACQUE, O., GLEESON, J. G., WOLFRUM, U., BEALES, P. L., GIBSON, T., DOHERTY, D., MITCHISON, H. M., ROEPMAN, R. & JOHNSON, C. A. 2015. An siRNA-based functional genomics screen for the identification of regulators of ciliogenesis and ciliopathy genes. *Nat Cell Biol*, 17, 1074-1087.

WOODS, A., SHERWIN, T., SASSE, R., MACRAE, T. H., BAINES, A. J. & GULL, K. 1989. Definition of individual components within the cytoskeleton of *Trypanosoma brucei* by a library of monoclonal antibodies. *J Cell Sci*, 93 (Pt 3), 491-500.

YADAV, S. P., SHARMA, N. K., LIU, C., DONG, L., LI, T. & SWAROOP, A. 2016. Centrosomal protein CP110 controls maturation of the mother centriole during cilia biogenesis. *Development*, 143, 1491-501.

YANG, T. T., CHONG, W. M., WANG, W. J., MAZO, G., TANOS, B., CHEN, Z., TRAN, T. M. N., CHEN, Y. D., WENG, R. R., HUANG, C. E., JANE, W. N., TSOU, M. B. & LIAO, J. C. 2018. Super-resolution architecture of mammalian centriole distal appendages reveals distinct blade and matrix functional components. *Nat Commun*, 9, 2023.

YILDIZ BOLUKBASI, E., MUMTAZ, S., AFZAL, M., WOHLBIER, U., MALIK, S. & TOLUN, A. 2018. Homozygous mutation in CEP19, a gene mutated in morbid obesity, in Bardet-Biedl syndrome with predominant postaxial polydactyly. *J Med Genet*, 55, 189-197.

ZHOU, Q., LEE, K. J., KURASAWA, Y., HU, H., AN, T. & LI, Z. 2018. Faithful chromosome segregation in *Trypanosoma brucei* requires a cohort of divergent spindle-associated proteins with distinct functions. *Nucleic Acids Res*, 46, 8216-8231.

ZOU, Y., MI, J., CUI, J., LU, D., ZHANG, X., GUO, C., GAO, G., LIU, Q., CHEN, B., SHAO, C. & GONG, Y. 2009. Characterization of nuclear localization signal in the N terminus of CUL4B and its essential role in cyclin E degradation and cell cycle progression. *J Biol Chem*, 284, 33320-32.