Characterisation of Membrane Glycoproteins that are Essential for Flagellar Attachment in the Bloodstream Form of *Trypanosoma brucei*

by

Louise McAteer

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

This thesis was submitted to the University of Dublin
in candidature for the degree of PhD.

Supervisor: Dr Derek Nolan

School of Biochemistry and Immunology
Trinity College Dublin

Submitted June 2020

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work, with the following exception:

- qRT-PCR investigations and initial generation of a FLA2-specific RNAi vector were performed in collaboration with Dr Derek Nolan, Trinity Biomedical Sciences Institute.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish.

Louise McAteer

June 2020

Abstract

The African trypanosome, *Trypanosoma brucei*, is a unicellular parasitic protist that causes a severe disease, trypanosomiasis, in humans and livestock in sub-Saharan Africa. The trypanosome has a single flagellum that performs important roles in motility, environmental sensing and immune evasion. Unusually, the flagellum is attached laterally along the outer surface of the parasite by a complex membrane junction called the flagellar attachment zone (FAZ).

Flagellar attachment is essential for proliferation of the bloodstream form of the trypanosome because the FAZ provides the positional information required for correct localisation of the cleavage furrow during cytokinesis. The FAZ is also of wider interest because it is a novel structure with the potential to provide insights into how cells co-ordinate the spatial and temporal expression of proteins. Most studies of the FAZ have employed detergent-extraction techniques and have therefore focused on the cytoskeletal components of the FAZ. Little is known about the membrane proteins in the extracellular space between the flagellum and the cell body. These surface proteins are likely to play a key role in the physical attachment of the flagellum onto the cell body.

In the procyclic form of the parasite, two membrane glycoproteins, FLA1 and FLA1BP, have been shown to mediate flagellar adhesion. However, the bloodstream-form FAZ is less well understood. Only a single membrane protein, FLA3, has been shown definitely to be essential for flagellar attachment in the bloodstream form, although others have been implicated. This thesis investigates further the localisation and function of FLA3. Additionally, another membrane glycoprotein, FLA2, is demonstrated here to be essential for flagellar attachment and cytokinesis in the bloodstream form.

Multiple *in-situ* tagging approaches were employed to add a variety of epitope tags to the respective C-termini of FLA2 and FLA3, enabling localisation studies, interaction studies and preliminary investigations of the glycan side chains of the two proteins.

Tagged FLA3 was shown to locate to the flagellar membrane of the FAZ. Super-resolution microscopy showed that the protein had a regularly spaced punctate distribution, consistent with the view that FLA3 is a component of the junctional complexes that span the intermembrane region of the FAZ, and which are critical for attachment of the

flagellum to the cell body. Tagged FLA3 was able to bind both RCA-I lectin and tomato lectin, indicating that the glycoprotein possesses galactose residues and, probably, *N*-acetyllactosamine repeats.

Specific knockdown of *FLA2* by RNAi confirmed for the first time that FLA2 is essential for flagellar attachment and cytokinesis in the bloodstream form. Epitope tagging of FLA2 was achieved using a long-primer PCR method. The molecular weight of the mature protein suggested extensive post-translational modification. Tagged FLA2 was able to bind RCA-I lectin, indicating that the protein is glycosylated and possesses galactose residues. The protein was not observed to localise to the FAZ when tagged at the C-terminus with any of a variety of epitope tags. Rather, the tagged protein mislocalised throughout the cell. This mislocalisation precluded precise localisation studies and interaction studies.

It is postulated here that in the bloodstream form, FLA2 performs the adhesion role previously attributed to FLA1. It is highly likely that interaction between FLA2 and FLA3 is required for flagellar attachment.

Acknowledgements

Enormous thanks must go to my supervisor, Derek Nolan, who was unfailingly kind and good-humoured and who managed to create a relatively stress-free research environment. Thank you for always making the effort to make understandable the concepts that are second nature to you; it is much appreciated by me and by hosts of undergrads. All in all, a superlative supervisor (TrypAdvisor?).

Thanks also go to the resident microscope and flow-cytometry experts, Gavin McManus and Barry Moran, who were always generous with their time and who gave me lots of helpful advice.

I am eternally indebted to the Irish Research Council for enabling me to start my career with such a fulfilling and instructive experience; I hope that their confidence in me was justified.

Thanks to all members of the fifth-floor reading room, past and present, – especially to Niki, Nidhi, Peter and Jess – for creating a friendly and welcoming atmosphere. Particular thanks must go to Paul Barry, who showed me the ropes when I was starting off – I honestly don't know how I would have managed without your help and approachability.

Thanks also to the perennial flock of Senior Sophs, for keeping me on my toes (and on my feet…).

I would like to thank my father for enduring my silences and absences during the write-up period, for keeping me fed, and for inquiring after the health of "my little lads" (the tryps).

Finally, I dedicate this thesis to my mother, who would have loved to see it completed.

Table of Contents

List of Figures

List of Tables

Abbreviations

a.a. ..... amino acids

Ab ..... antibody

AP ..... alkaline phosphatase

APEX ..... enhanced ascorbate peroxidase

BCIP ..... 5-bromo-4-chloro-3-indolyl phosphate

BioID ..... proximity-dependent biotin identification

BIP (BiP) ..... binding protein (a soluble ER housekeeping protein)

BirA ..... bacterial biotin ligase

BLAST ..... Basic Local Alignment Search Tool

BLASTn ..... Basic Local Alignment Search Tool (nucleotide)

BLASTp ..... Basic Local Alignment Search Tool (protein)

bp ..... base pairs

BSA ..... bovine serum albumin

BSF ..... bloodstream form

cAMP ..... cyclic adenosine monophosphate

cDNA ..... complementary DNA

CDS ..... coding sequence

CHAPS ..... 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate

CLEM ..... correlative light-electron microscopy

ClpGM6 ..... calpain-like GM6-repeat protein

Co-IP ..... co-immunoprecipitation

ConA ..... concanavalin A

DAB ..... diaminobenzidine

DABCO ..... 1,4-diazabicyclo[2.2.2]octane

DAPI ..... 4′,6-diamidino-2-phenylindole

dH_2O ..... deionised water

DMSO ..... dimethyl sulfoxide

DNDi ..... Drugs for Neglected Diseases initiative

dNTPs ..... deoxyribonucleotide triphosphates

dsRNA ..... double-stranded RNA

EDTA ..... ethylenediaminetetraacetic acid

EGTA ..... ethylene glycol tetraacetic acid

EM ..... electron microscopy

EMA ..... European Medicines Agency

EMBOSS ..... European Molecular Biology Open Software Suite

ER ..... endoplasmic reticulum

ESP8 ..... enhanced in surface proteome protein 8

ET ..... electron tomography

eYFP ..... enhanced yellow fluorescent protein

FAZ ..... flagellar attachment zone

FBS ..... foetal bovine serum

FC ..... flagella connector

FLA ..... flagellar adhesion glycoprotein

FLA1BP ..... FLA1-binding protein

FLAM3 ..... flagellum member 3

FP ..... flagellar pocket or forward primer

FSC ..... forward scatter

Gal ..... galactose

GFP ..... green fluorescent protein

GlcNAc ..... *N*-acetylglucosamine

GPI ..... glycosylphosphatidylinositol

GS ..... glycine-serine linker

(H + L) ..... heavy and light chain

HAT ..... human African trypanosomiasis

HEPES ..... 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HMI-9 ..... Hirumi's modified Iscove's medium 9

HMMP ..... high molecular mass protein

HRP ..... horseradish peroxidase

IF ..... immunofluorescence

IFT ..... intraflagellar transport

Ig..... immunoglobulin

IP..... immunoprecipitation

kb ..... kilobase

kDa ..... kilodalton

LacNAc..... N-acetyllactosamine

LB..... lysogeny broth

MFI..... median fluorescence intensity

MITat 1.1 ..... Molteno Institute Trypanozoon antigenic type

MtQ..... microtubule quartet

Nb ..... nanobody

NBT..... nitro blue tetrazolium

NCBI..... National Center for Biotechnology Information

NG..... mNeonGreen fluorescent protein

NHL repeat ..... NCL-1, HT2A and Lin-41 protein motif

NHS..... N-Hydroxysuccinimide

ORF..... open reading frame

PAGE..... polyacrylamide gel electrophoresis

PBS..... phosphate-buffered saline

PCF..... procyclic form

PCR..... polymerase chain reaction

PFA ..... paraformaldehyde

PFR..... paraflagellar rod

PI..... protease inhibitor

PNGase F..... Peptide: N-glycosidase F

Pol-I..... RNA Polymerase I

Poly(A)..... poly adenosine monophosphate

pPOT..... plasmid for PCR-only tagging

PSG..... phosphate-saline-glucose buffer

pSMOX..... plasmid Single Marker Oxford

PTU..... polycistronic transcription unit

qRT-PCR..... quantitative real-time PCR

RCA-I.....ricinus communis agglutinin lectin

rcf.....relative centrifugal force

RIT-Seq.....RNA interference target sequencing

RNAi.....RNA interference

RNA-Seq.....RNA sequencing

RP.....reverse primer

rpm.....revolutions per minute

r.t......room temperature

SAS.....splice acceptor site

SDM-79.....semi-defined medium 79

SDS.....sodium dodecyl sulfate

S/G.....sucrose & glucose

SIF.....stumpy-inducing factor

SILAC.....stable-isotope labelling of amino acids in cell culture

siRNA.....small interfering RNA

SLRNA.....spliced leader RNA

spp......*species pluralis*

s/n.....supernatant

SR.....mScarlet fluorescent protein

SS.....disulfide

SSC.....side scatter

STED.....stimulated emission depletion microscopy

STEM.....scanning transmission electron microscopy

T......*Trypanosoma*

T7RNAP.....T7 RNA polymerase

TAE.....tris-acetate-EDTA

T. b......*Trypanosoma brucei*

T. brucei.....*Trypanosoma brucei brucei*

TBS.....Tris-buffered saline

TBST.....Tris-buffered saline with Tween-20

TE.....translational efficiency

TEM ..... transmission electron microscopy

TEMED ..... tetramethylethylenediamine

TES.....N-Tris(hydroxymethyl)methyl-2-aminoethane sulfonic acid

TETR..... tetracycline repressor

TFA..... trifluoroacetic acid

TL..... tomato lectin

TM..... transmembrane

TOPCONS ..... topology consensus prediction

U..... enzyme unit

uORF ..... upstream ORF

UTR..... untranslated region

VSG..... variable surface glycoprotein

WHO ..... World Health Organisation

YFP..... yellow fluorescent protein

Chapter 1: Introduction

The morphology of the African trypanosome, *Trypanosoma brucei*, is dominated by a single flagellum that emerges from the posterior end of the parasite, follows a helical path around the cell body and extends out just beyond the cell's anterior tip. The flagellum is attached onto the outer surface of the parasite for most of its length by means of an unusual intracellular membrane-membrane interaction. The mechanism behind this membrane adhesion is incompletely understood. This research project seeks to investigate the glycoproteins that have been implicated to play a role in flagellar attachment in bloodstream forms of the parasite – in particular, FLA2 and FLA3.

1.1 *Trypanosoma brucei*

Trypanosoma brucei is a unicellular eukaryote that belongs to the Kinetoplastea class of Euglenozoa (Kohl and Bastin, 2005). It is an obligate parasite, undergoing a diheteroxenous life cycle in the tsetse fly (*Glossina spp.*) and in any of a wide range of mammalian host species. Trypanosomes are delivered into the mammalian host through the bite of an infected tsetse fly when the insect takes a blood meal. The mammalian-infective metacyclic form of the parasite is in a growth-arrested state, but it soon differentiates into the slender bloodstream form (BSF), which proliferates rapidly in the bloodstream, lymph and interstitium of the host.

A certain proportion of the trypanosome population differentiates further into a morphologically distinct stumpy form, which is non-dividing. This transition protects the host by preventing uncontrolled proliferation of the parasites. The slender BSFs secrete peptidases that degrade host proteins; the resulting oligopeptides act as a paracrine quorum-sensing signal (Stumpy-Inducing Factor (SIF)) that promotes differentiation to the stumpy form when taken up by the surface receptor, TbGPR89 (Rojas et al., 2019). The stumpy-form cells are adapted for survival upon ingestion by a tsetse fly during a blood meal. They show increased sensitivity to environmental cues that signify entry to the tsetse fly and they are relatively resistant to the changes in temperature, pH and proteolytic environment that are encountered in the tsetse midgut. These characteristics enable them to differentiate rapidly into what is termed the procyclic form (PCF) (MacGregor et al., 2012).

The procyclic form has a fundamentally different metabolism to the BSF because of the different nutrient availability in the tsetse vector. Over a period of weeks, the parasite

migrates to the salivary glands, whence it can be passed on into a new mammalian host. Along the way, it goes through the mesocyclic, epimastigote and metacyclic stages, as befits the microenvironment (Sunter and Gull, 2016). The life cycle of the trypanosome is summarised in Fig. 1.1 A.

As members of the kinetoplastid class, trypanosomes possess a distinctive feature called the kinetoplast. This is a condensed granule of mitochondrial DNA, positioned directly adjacent to the basal body of the flagellum (Rotureau et al., 2014). The epimastigote life-cycle stages are defined by the positioning of the kinetoplast (and therefore the base of the flagellum) anterior to the nucleus (Fig. 1.1 B). All other life-cycle stages of *T. brucei* (including the bloodstream- and procyclic forms) have trypomastigote morphology, in which the kinetoplast is posterior to the nucleus. In both epimastigotes and trypomastigotes, the flagellum is attached laterally along the cell body (Kohl and Bastin, 2005).

Since the bloodstream- and procyclic forms are the major proliferative stages, and are readily grown in culture, they are the best-studied forms of *T. brucei*. They are 15 – 30 μm in length and 3 μm in diameter (Sherwin and Gull, 1989). Their long, slender shape, with its tapered ends, is maintained by a helical array of subpellicular microtubules running along the long axis of the cell, just under the plasma membrane. This is a characteristic feature of kinetoplastids. The subpellicular microtubules are approximately 20 nm apart and they have a highly ordered set of side-arms to keep them firmly in position. Whereas most eukaryotic cells disassemble their microtubules prior to mitosis, trypanosomes keep the subpellicular microtubule array intact throughout the cell cycle (Sherwin and Gull, 1989).

The BSF and PCF forms of *T. brucei* reproduce asexually by binary fission and have a tightly-controlled cell division process that ensures that each daughter cell receives exactly one nucleus, kinetoplast and flagellum (Sherwin and Gull, 1989). (Sexual reproduction is possible in the epimastigote life-cycle stage, facilitating exchange of genetic material between different strains, but is not an obligatory part of the parasite's life cycle (Gibson, 2015).) Flagellar attachment is essential for trypanosome reproduction in the bloodstream form. The attachment zone where the flagellum is connected to the cell body has been shown to provide crucial positional information that ensures correct location of the cleavage furrow as the cell splits in two (Robinson et al., 1995, Rotureau et al., 2014).

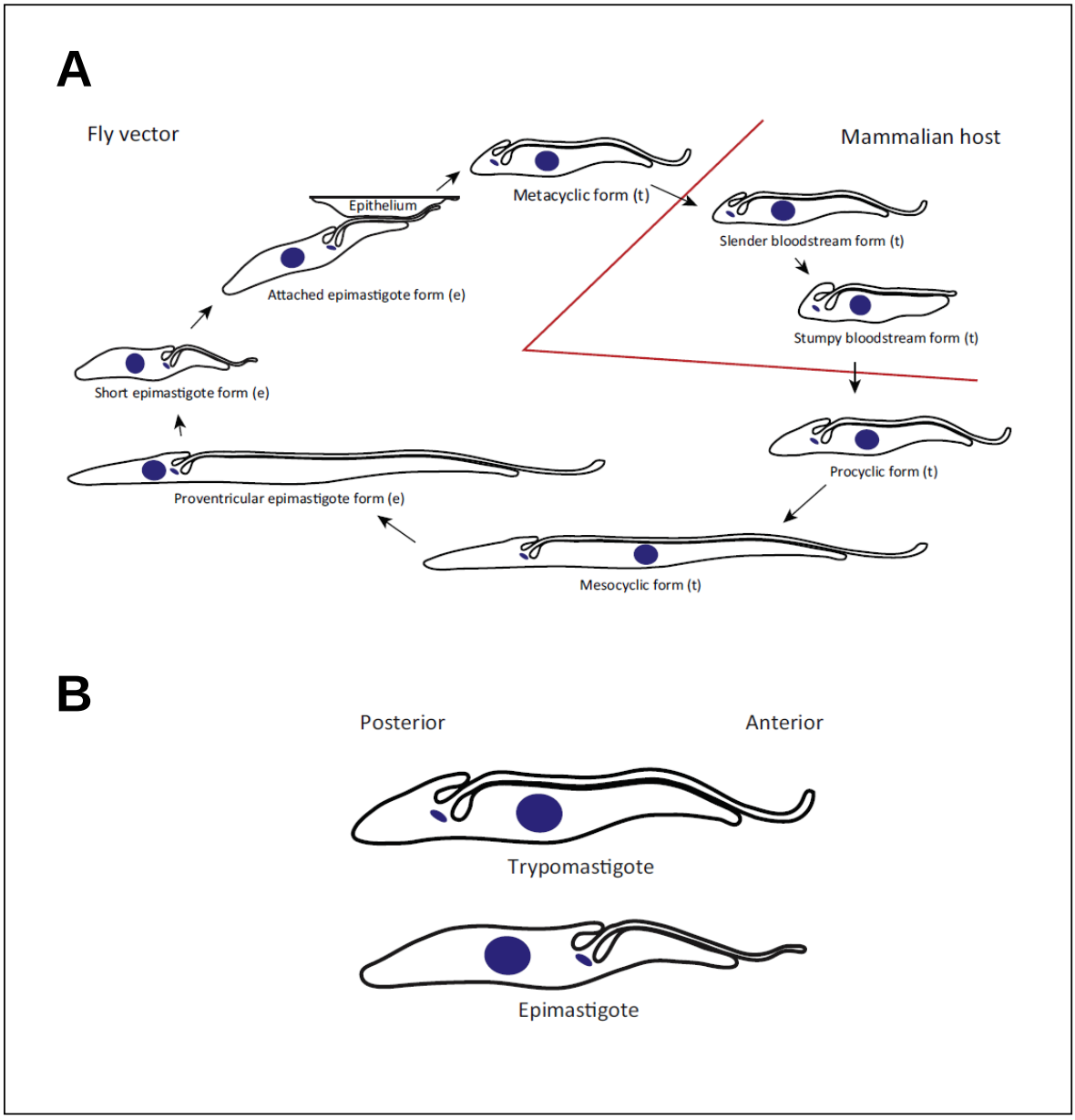


Figure 1.1: The life-cycle stages of *T. brucei* (Taken from Sunter and Gull, 2016)

(A) The trypanosome progresses through a series of different morphological forms over the course of its life cycle. Trypomastigote forms are indicated by (t); epimastigote forms by (e).

(B) A comparison of the trypomastigote and epimastigote forms. Note that in both cases, the flagellum is attached laterally to the cell body for most of its length.

Large blue circles represent the nucleus; small blue ovals represent the kinetoplast.

Since this junctional region is unique to the parasite, it could act as a target for therapeutic drugs without risk of damage to host tissue (Oberholzer et al., 2009).

Indeed, there is a need for new trypanocidal drugs. *T. brucei* is a major pathogen of cattle and other livestock animals across sub-Saharan Africa (Auty et al., 2015). It causes trypanosomiasis, known locally as nagana, a disease that is fatal unless treated. The resulting loss of productivity in East African cattle alone is estimated to cost the economy US$ 2.5 billion over a twenty-year period (Shaw et al., 2014). Two related trypanosome species, *T. congolense* and *T. vivax*, are also causative agents of trypanosomiasis in livestock, but most research has focused on *T. brucei* because it has two subspecies, *T. b. gambiense* and *T. b. rhodesiense*, which are pathogenic to humans (Auty et al., 2015).

Human African Trypanosomiasis (HAT), which is also called sleeping sickness, is a disease of two stages. The haemolymphatic stage, in which the parasite is confined to the vasculature, is characterised by fever, headache, pruritus and swollen lymph glands. Passage of the trypanosomes across the blood-brain barrier triggers the meningoencephalitic stage, which is considerably more difficult to treat. This stage is characterised by sleep disruption, tremor, paralysis and psychosis and is followed, if unchecked, by coma and death. Until very recently, the drugs used to treat HAT were administered intravenously in a prolonged, complex regimen that was difficult to facilitate in underdeveloped areas. The Stage-2 drugs (melarsoprol and eflornithine) were prone to severe and occasionally lethal side effects, and there was evidence of increasing drug resistance in the trypanosome population. Early screening for the disease, in combination with tsetse vector control, proved effective in reducing the impact of sleeping sickness, with only 9,000 fatalities in 2010 compared to 34,000 in 1990. In 2002, the WHO set up a global alliance for the elimination of HAT. However, the impracticable financial cost of complete eradication of the tsetse fly, coupled with the fact that animals act as a reservoir for the parasite, means that the disease remains a significant threat (Brun et al., 2010).

The Drugs for Neglected Diseases initiative (DNDi) has developed a drug, fexinidazole, that promises to revolutionise the treatment of HAT. Fexinidazole is the first all-oral treatment for HAT. A 5-nitroimidazole derivative, fexinidazole is metabolised into reactive amine species that inhibit DNA synthesis in trypanosomes without being toxic or mutagenic to mammalian cells. The drug is suitable for treatment of Stage 1 and Stage 2

of both chronic (*T. b. gambiense*) and acute (*T. b. rhodesiense*) trypanosomiasis, and its adverse side effects are relatively mild. In 2018, the EMA gave fexinidazole approval to be registered for use in endemic countries. The drug is currently in the closing stages of Phase-3 clinical trials in Guinea and the Democratic Republic of the Congo. Fexinidazole will be manufactured by Sanofi and donated to the WHO for free distribution to affected regions (Deeks, 2019).

Furthermore, another oral drug, acoziborole, has also reached Phase-3 trials (Wall et al., 2018) and shows promise as a single-dose treatment for both forms of HAT (De Rycker et al., 2018). In addition, the discovery of a selective inhibitor of the kinetoplastid proteasome complex has potential to be developed as a treatment for trypanosomiasis and the other kinetoplastid diseases, Chagas disease and leishmaniasis (Khare et al., 2016).

However, there are no plans to use fexinidazole/acoziborole in a veterinary context, due to the risk of outgrowth of drug-resistant trypanosomes, so nagana remains a major issue for livestock farmers. The current drugs, diminazene aceturate and isometamidium chloride, are over fifty years old and their usefulness is threatened by increasing drug resistance (Auty et al., 2015).

Furthermore, the negative impacts of *T. brucei* are not limited to the tsetse region. Two subspecies, *T. b. equiperdum* and *T. b. evansi*, have extended their geographical range far beyond Africa (Auty et al., 2015). For clarity and ease of nomenclature, these are often classified as separate species, grouped with the parental *T. brucei brucei* under the subgenus Trypanozoon (Desquesnes et al., 2013). They have lost most of their kinetoplast DNA and hence cannot differentiate into the procyclic form or survive inside the gut of the tsetse fly (Lai et al., 2008). Instead, they have adapted to mechanical transmission from host to host. *T. equiperdum* is a venereal parasite of horses. *T. evansi* is spread by haematophagous biting flies such as tabanids and stomoxes, without any specific biological interaction with the insect. It infects a wide range of mammalian hosts, especially camels, horses, cattle and water buffalo. It is the causative agent of a disease called surra. Symptoms of surra can be acute or chronic, and vary according to host species and circumstances, but they include fever, anaemia, cachexia, abortion and death. *T. evansi* is found throughout South America, northern Africa, the Middle East and south-

east Asia, and has been prevented from entering North America, Europe and Australia only by stringent quarantine measures (Desquesnes et al., 2013).

No prophylactic is available against any of the various forms of trypanosomiasis (Auty et al., 2015). The pervasiveness of the disease is due to the parasite's remarkable immune evasion strategies: antigenic variation, immunosuppression and hydrodynamic clearance of surface-bound antibodies – the last of which is mediated by the attached flagellum (Engstler et al., 2007).

Antigenic variation is the trypanosome's most powerful weapon against the host's immune system. A monolayer of a variable surface glycoprotein (VSG) forms a uniform surface coat over the cell body and flagellum. This shields all other potential antigens on the trypanosome's surface, and presents a single, distinctive target to the immune cells. The host rapidly mounts a humoral response that clears most of the parasite population. However, an apparently stochastic change in VSG gene expression in an individual trypanosome (which occurs approximately once per 1000 divisions *in vivo*) results in proliferation of an antigenically novel population. The host mounts a fresh antibody response, but once again, a VSG gene-switching event means that some of the parasites escape. This repeated cycle causes the waves of increasing and decreasing parasitaemia that are characteristic of a chronic trypanosome infection (MacGregor et al., 2012). It is in the parasite's own interest to prolong the life of the host, in order to increase its chance of being taken up by a tsetse fly. It therefore releases a range of immunosuppressive molecules including *T. brucei*-derived kinesin heavy chain (TbKHC1), *T. brucei* adenylate cyclase (TbAdC) and indolepyruvate (Stijlemans et al., 2016, Barry, 2017). These dampen the innate immune response, protecting both host and pathogen from a hyperinflammatory reaction.

Nonetheless, the trypanosome encounters formidable opposition from the immune system, especially in the form of antibody-mediated complement lysis. The proliferative population must resist this for long enough for VSG gene-switching to occur; stumpy forms must survive it until ingested by the insect vector. Resistance to low levels of VSG-binding antibodies is achieved by hydrodynamic clearance of the anti-VSG through the action of the flagellum (Engstler et al., 2007). The flagellum of the trypanosome is in constant motion, propelling the cell at an average speed of 20 μm s^{-1} through serum.

Because of its attachment all along the cell body, its motion is accompanied by continuous, vigorous contortion of the entire cell. The directional swimming of the trypanosome creates a substantial drag force towards the posterior end of the cell – especially in the confined space of capillaries. When antibodies bind to the homogenous, smooth surface formed by the VSG layer, they protrude out from the trypanosome and are subjected to this hydrodynamic flow. The force drags the bound VSG laterally through the fluid-mosaic of the plasma membrane in the posterior direction so that the Ig-VSG complex ends up at the rear of the cell. Here, the complex is endocytosed in a unique organelle called the flagellar pocket (FP), which is a large invagination of the cell membrane around the base of the flagellum. The flagellar pocket is the only site of endocytosis in the trypanosome. This is an immune-evasive strategy, since it means that all receptor and transport proteins that could potentially act as antigens are hidden away in this pouch, protected from the host serum by a tight collar region through which immune factors do not pass. Once the Ig-VSG complex is endocytosed, the antibody is degraded in the lysosome and the VSG is recycled back to the plasma membrane. This process is remarkably efficient. When measured experimentally, the half-life of surface Ig-VSG was about 40 seconds in slender BSF cells, and 20 seconds in stumpy cells (Engstler et al., 2007). The *N*-glycan side chains of VSG contribute to the high lateral mobility of the protein, probably by reducing the attraction between neighbouring VSG dimers (Hartel et al., 2016). However, the process only functions at low to moderate antibody concentrations. When the immunoglobulin levels rise until most of the VSG is bound, the drag force reduces exponentially and the system collapses.

It is clear, then, that the flagellum plays a crucial role in the pathogenicity of the trypanosome. However, scientific interest in the trypanosome is by no means limited to its disease-causing attributes. *T. b. brucei* has emerged as a useful model organism for basic molecular-biology research, especially for the study of processes such as mitosis, Golgi duplication and phospholipid biosynthesis (Akiyoshi and Gull, 2013, Zhou et al., 2014, Serricchio and Butikofer, 2011). As members of Excavata, an early-branching supergroup of Eukarya, trypanosomes are only very distant relatives of fungi, plants and animals. They have unique features such as membrane-enclosed glycolytic compartments (glycosomes) (Docampo, 2011) and unique processes such as Pol-I transcription of some protein-coding genes (Haile and Papadopoulou, 2007). Contrasting their biological processes with those

of more familiar species reveals what mechanisms are central to eukaryotic biology, and yields insight into the evolution of these processes (Akiyoshi and Gull, 2013, Sohail, 2005). No matter how unusual the trypanosome's physiology and biochemistry may seem on some levels, they nonetheless share many important features with multicellular species – even humans. Indeed, the area in which *T. brucei* has proven itself most invaluable to research is in facilitating the discovery of fundamental biological mechanisms. For example, extranuclear DNA, glycosylphosphatidylinositol (GPI) anchoring, RNA editing, *trans*-splicing and acidocalcisomes were first observed in the trypanosome, but soon proved to be widespread phenomena throughout the eukaryotic domain (Docampo, 2011).

It is the immense tractability of *T. brucei* in culture that lends itself to basic research. Its compact genome (which has less than 10,000 genes, most of which lack introns) is available online at TriTrypDB (Sohail, 2005, Akiyoshi and Gull, 2013). Trypanosomes are very amenable to a wide range of genetic manipulations, including stable and transient transfections, gene knockdown, ectopic overexpression of genes, tandem affinity purification, quantitative mass spectrometry and high-throughput RNAi screening (Zhou et al., 2014). They have efficient homologous recombination that needs only short regions of homology, which is ideal for gene knockout, gene replacement and *in-situ* epitope tagging (Dean et al., 2015, Zhou et al., 2014). However, complete gene knockout requires two rounds of transfection, each with a different selection marker, because of the parasite's diploidy (Docampo, 2011). Further benefits of working with *T. brucei* include a short doubling time of 6–9 h, which facilitates rapid culture expansion for use in experiments, and ease of cryopreservation (Akiyoshi and Gull, 2013, Vickerman, 2009). Since it is non-pathogenic to humans, it can be safely handled in the laboratory (Sohail, 2005).

Finally, the fact that the trypanosome is unicellular gives the parasite a degree of simplicity that makes it relatively easy to study, even though it is quite sophisticated in its architecture. It also means that the parasite *in vitro* is a closer representative of its *in vivo* counterpart than are, for example, most tissue-culture lines to their *in vivo* mammalian counterparts. It is likely, therefore, that study of the ultrastructure of the *T. brucei* flagellar

attachment zone will produce novel and meaningful results and may even be applicable to the membrane-membrane junctions of more complex species.

1.2 The Flagellar Attachment Zone

1.2.1 The *T. brucei* flagellum: function

As mentioned above, the laterally attached flagellum is required both for antibody clearance (Dean and Matthews, 2007) and for directional movement from place to place (Wheeler, 2017). Translocation is essential during the trypanosome's life cycle – especially in the insect vector, where the parasite must make its own way from the midgut to the salivary glands. It has long been postulated that the trypanosome flagellum also has a sensory role, responding to mechanical or chemical signals in the environment (Maric et al., 2010); a recent study has demonstrated that a cAMP signal transduction pathway in the flagellum is required for migration of the parasite out of the midgut lumen of the tsetse fly (Shaw et al., 2019). Furthermore, the trypanosome's flagellum facilitates physical attachment to the insect host in the epimastigote life-cycle stage. The anterior tip of the flagellum becomes embedded in the inner surface of the tsetse fly's digestive tract. Extensive accumulation of intraflagellar material forms a hemidesmosome-like plaque that anchors the parasite securely in place (Bastin et al., 2000, Kohl and Bastin, 2005).

However, the most interesting function of this versatile organelle is its role in cytokinesis in the bloodstream-form cell. The attachment of the flagellum to the cell body along the flagellar attachment zone provides the positional information needed for cell division to occur correctly (Sunter and Gull, 2016). From a pragmatic point of view, this is noteworthy because it suggests a possible way to stop proliferation of the disease-causing form of the parasite: detaching the flagellum *in vivo*. From a more theoretical standpoint, this is a fascinating – and perplexing – example of precise spatial coordination in a cell. It is easy to see why the elongated and highly structured trypanosome cell needs a precisely oriented, longitudinal division plane to ensure inheritance of a single nucleus and flagellum in each daughter cell. But how is this brought about by the attached flagellum – and how, indeed, is the flagellum attached to the cell body in the first place? The answers to these questions are starting to come to light as more is learned about the structure of the flagellum and its attachment zone.

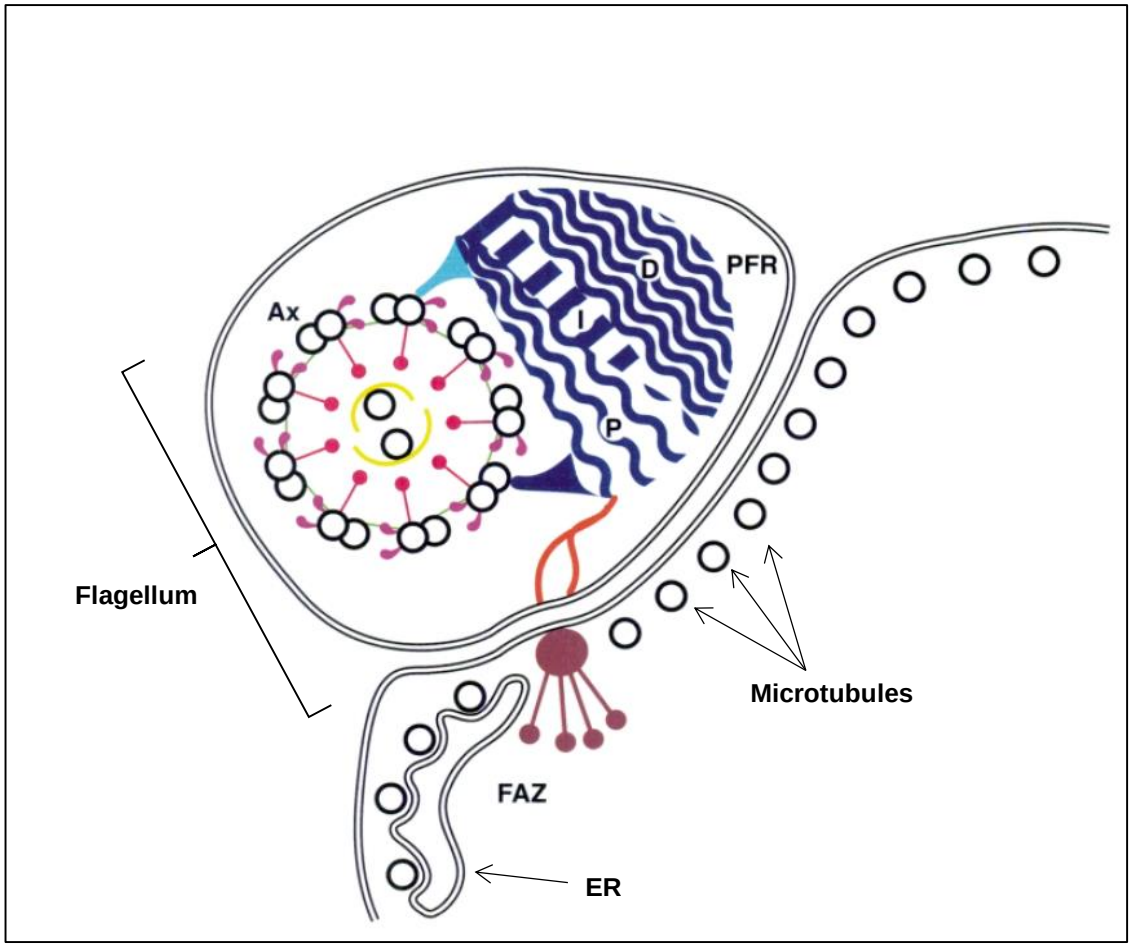


Figure 1.2: Schematic diagram of transverse section of the *Trypanosoma brucei* flagellum (adapted from Bastin, 2000, Fig. 5 C)

The axoneme (Ax) is composed of the classic 9+2 microtubule structure, with dynein arms (pink) and radial spokes (red). The PFR (blue) may be divided into three domains: proximal (P), intermediate (I) and distal (D), defined by their positions relative to the axoneme. Filaments (orange) connect the PFR to the FAZ filament (brown), which is on the cell-body side of the flagellar attachment zone. Four microtubules of the subpellicular corset are associated with the smooth endoplasmic reticulum (ER).

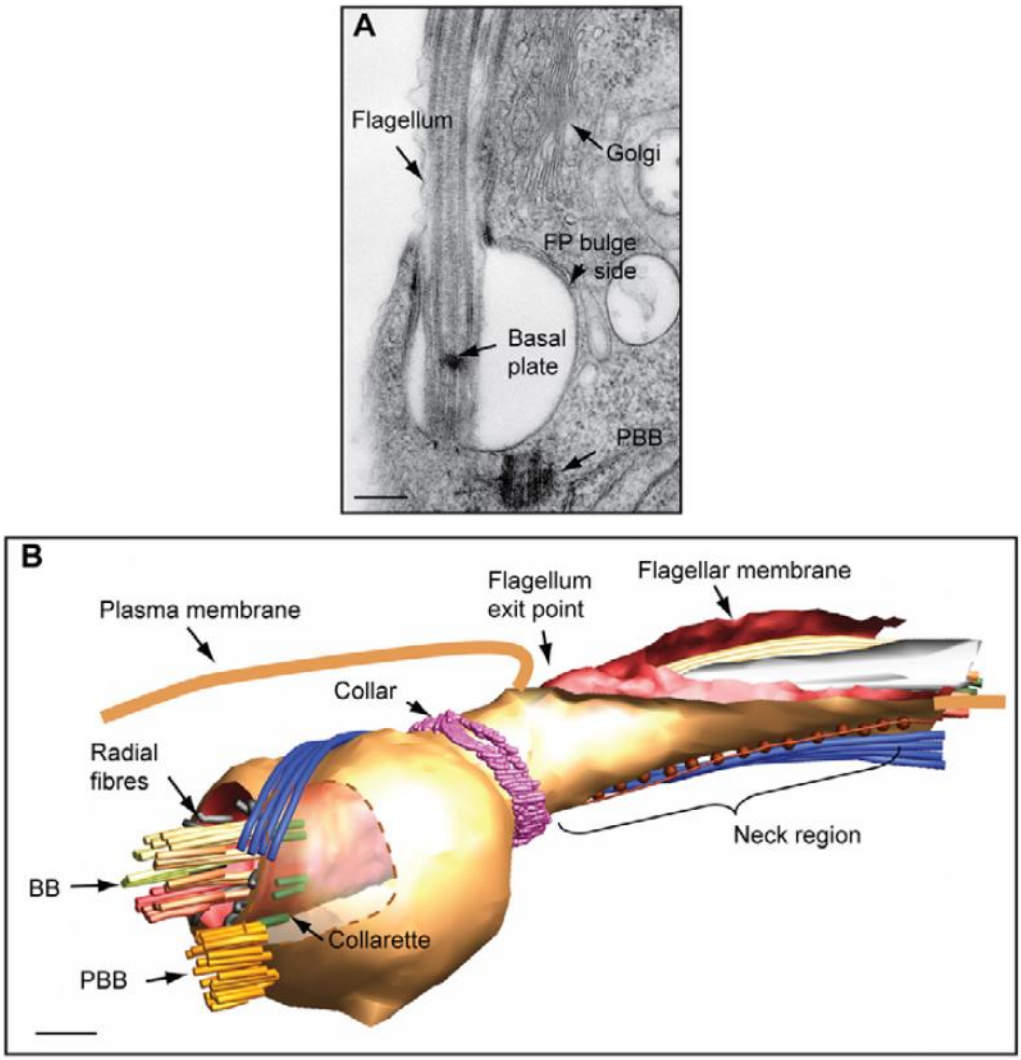


Figure 1.3: The flagellar pocket (Lacomble *et al.*, 2009, Fig. 3)

(A) Longitudinal section of the flagellar pocket (FP), showing the basal plate that anchors the axoneme, and the probasal body (PBB) from which the daughter flagellum extends.

(B) 3D reconstruction of the flagellar pocket, based on electron tomography. Note the extension of the flagellum from the basal body (BB), through the FP and out onto the exterior surface of the cell body. The FP membrane is bounded on one side by the collar and on the other by the collarette.

Scale bars are 200 nm.

1.2.2 The *T. brucei* flagellum: structure

The flagellum of *T. brucei* is homologous to the flagella and cilia found in animals and other protists, not to the flagella of bacteria (Kohl and Bastin, 2005). As is the case for virtually all motile eukaryotic flagella, it has a structurally-conserved 9+2 axoneme as its core – i.e. a set of nine microtubule doublets surrounding an asymmetric pair of microtubules. The associated machinery – radial spokes and dynein arms – enables the sliding of the microtubules against one another. The resulting flagellar beating occurs from tip to base, rather than from base to tip as in most flagellated organisms. The axoneme is anchored by a basal body that is composed of nine microtubule triplets. Although homologous to a centriole, the basal body plays no role in mitosis (Bastin et al., 2000).

Adjacent to the axoneme is the paraflagellar rod (PFR), a structure unique to the kinetoplastids and euglenoids. The PFR has a complicated lattice structure and is thought to have many protein components, only some of which have been characterised. It has three distinct regions called the proximal, intermediate and distal domains; the proximal domain is connected by filaments to doublets 4 to 7 of the axoneme. The PFR lies between the axoneme and the cell body and has a roughly crescent shape in cross section (Fig. 1.2). The function of the PFR is to aid motility. Knockdown of PFR2, one of the main components of the paraflagellar rod, greatly reduces flagellar beating. The trypanosomes stop swimming and sediment at the bottom of the culture flask. They remain viable but would probably not survive *in vivo*. It is not clear why the axoneme alone is insufficient for trypanosome motility (Bastin et al., 2000).

Together, the axoneme and PFR are enclosed by a membrane that is contiguous with the membrane of the flagellar pocket (FP), which in turn is contiguous with the plasma membrane that encases the cell body (Lacomble et al., 2009). The FP, as mentioned above, is a large invagination of the cell membrane, and is the only site of endo- and exocytosis in the trypanosome. It is about 1 μm^3 in volume (Kohl and Bastin, 2005). The basal body of the axoneme is located at the base of the FP. The flagellum extends right through the pocket, emerging through its mouth. Thus, the FP is toroid in topology, rather than spherical (Fig. 1.3). A tight collar ensures that the FP-membrane and flagellar membrane are in close contact with each other at the mouth of the FP, restricting entry to the pocket.

Once the flagellum emerges from the FP, its neck region is bent back towards the cell body until it touches the pellicular membrane. For the rest of its length, the flagellum is attached to the cell body by means of a complicated structure called the flagellar attachment zone (FAZ) – except for a short stretch of free flagellum that protrudes beyond the anterior tip of the cell (Lacomble et al., 2009).

Thus, the trypanosome's outer membrane has three distinct domains, the flagellar, cell-body (or pellicular) and flagellar-pocket membranes (the flagellum neck is often considered a fourth domain). Each has a distinct proteome. The sorting of proteins to the correct membrane compartment is believed to occur in the FP. Meanwhile, transitional fibres that join the basal body to the base of the flagellar membrane create a selective structural barrier separating the interior of the flagellum from the cytoplasm. As a result, every protein that enters the lumen or membrane of the flagellum must be specifically targeted there (Lacomble et al., 2009).

The flagellum is constructed by addition of proteins onto the distal tip (i.e. the end farthest away from the basal body). There are no ribosomes in the flagellum, so flagellar proteins are not made *in situ*. Instead, they are synthesised in the cytosol and brought into position by an intraflagellar transport (IFT) system, as is the case in almost all flagella and cilia. IFT complexes, or "rafts", bind to the proteins in question and shuttle back and forth along the flagellum by means of kinesin and dynein motors. This enables outgrowth and continual maintenance of the flagellum (Kohl and Bastin, 2005).

Construction of a new flagellum is a fundamental part of the trypanosomal cell cycle (Fig. 1.4). Initiation of the cell cycle is marked by maturation of the pro-basal body that flanks the basal body of the flagellum. A daughter flagellum starts to elongate on this new basal body. Meanwhile, the mitochondrial genome replicates inside the kinetoplast. The kinetoplast is always in tight association with the basal body. The old and new basal bodies start to separate as the daughter flagellum extends. This causes the attached kinetoplast to split in two, providing an elegant and efficient way to ensure that each flagellum has an associated kinetoplast. Mitosis then begins in the nucleus – without dissolution of the nuclear membrane. The elongate nucleus straddles the axis of the old flagellum and divides in two. One daughter nucleus lies between the two basal bodies (anterior to the one subtending the daughter flagellum and posterior to the one

subtending the mature flagellum). The other daughter nucleus lies anterior to both basal bodies (Sherwin and Gull, 1989).

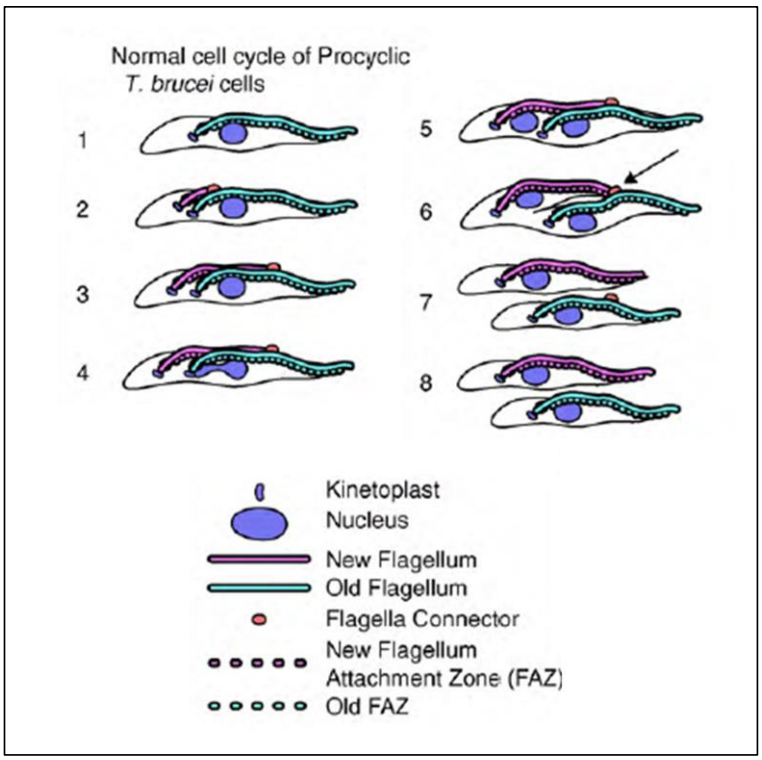


Figure 1.4: The cell cycle of *T. brucei* (Vaughan, 2010, Fig. 2) At the start of the cell cycle, the trypanosome has one kinetoplast and one nucleus (1). A new flagellum starts to grow from the probasal body, and the kinetoplast divides in two (2). The daughter flagellum extends along the path of the old flagellum, to which it is attached by the flagella connector (3). The flagellar connector halts at a point ~0.6 the length of the old flagellum. Mitosis occurs in the nucleus (4). The nucleus divides in two. The daughter flagellum extends in the posterior direction (5). Cleavage-furrow ingression occurs between the two flagella, starting at the distal tip of the new FAZ (6). The cells separate (7). There is post-division growth of the daughter cell with the new flagellum (8).

Cell division in the bloodstream form is very similar, but there is no flagella connector.

Having thus duplicated and segregated its major organelles, the trypanosome is ready to undergo cytokinesis. The cleavage furrow starts at the anterior tip of the FAZ, then follows a precise path between the two flagella such that there is a single nucleus, kinetoplast and basal body on either side. It has been shown that the position and length of the cleavage furrow depend on the position and length of the new flagellum. The division plane is helical, because the attached flagella follow the subpellicular microtubules, whose path forms a left-handed (anticlockwise) helix along the longitudinal axis of the cell. The two daughter cells end up facing in opposite directions, joined only at their posterior end. Flagellar beating provides the force needed separate them completely (Sherwin and Gull, 1989).

Clearly, then, the daughter flagellum must be attached onto just the right part of the cell surface if cytokinesis is to occur correctly. The assembling flagellum takes its positional cues from the existing cytoskeletal structures in the cell, i.e. by cytotaxis. In procyclic cells, the tip of the new flagellum has a novel structure, the flagella connector (FC), that tethers it very firmly to the axoneme of the old flagellum (Moreira-Leite et al., 2001). As the daughter flagellum elongates, the flagella connector moves along the existing flagellum in the distal direction. This ensures that the new flagellum follows the correct path.

In bloodstream-form cells, the tip of the new flagellum is embedded in an indentation called the flagellar groove (Hughes et al., 2013). The groove seems to have the same cytotactic function as the FC, but it follows the intracellular part of the old FAZ (the MtQ, see below), rather than the axoneme of the old flagellum. It is highly probable that the groove is an adaptation to conceal invariant connector proteins from the host immune system.

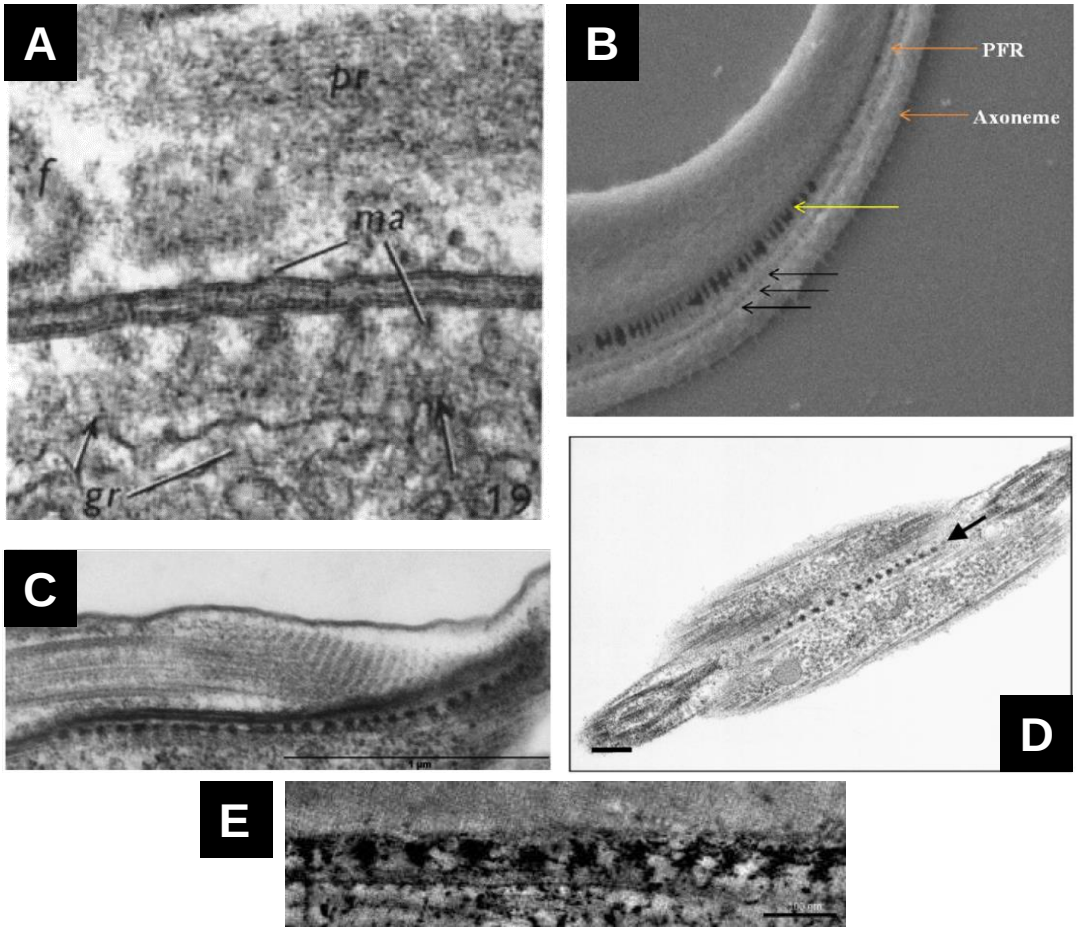


Figure 1.5: Longitudinal sections of the flagellar attachment zone.

(A) In longitudinal section, regularly spaced structures (*macula adherens*, ma) traverse the flagellar attachment zone of BSF *T. congolense*. Taken from Vickerman (1969). **(B)** Scanning electron microscopy of a procyclic *T. brucei* cell shows regularly spaced filamentous structures (yellow arrow) joining the PFR and the cell body. Taken from Tsang (2012). **(C)** In longitudinal section, a row of punctate structures on the cell-body side of the FAZ are easily visible in PCF *T. brucei* by transmission electron microscopy. Taken from Buisson and Bastin (2010). **(D)** In coronal section, it can be seen that the punctate junctional complexes of the FAZ form a single row in PCF *T. brucei*. Taken from Bastin et al. (2000). **(E)** Tomographic slice of junctional complexes in a high-pressure frozen procyclic cell. Taken from Höög (2012).

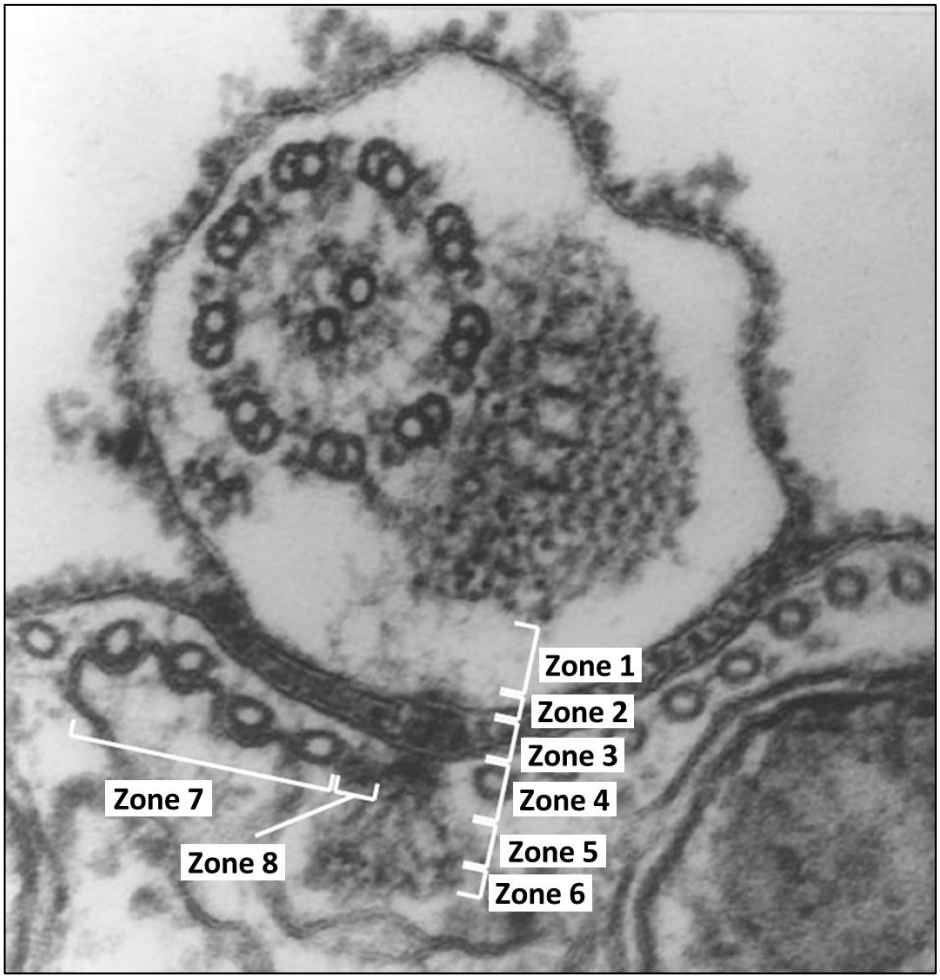


Figure 1.6: The Zones of the FAZ (Sunter and Gull, 2016, Fig. 2 B)

Zone 1: Fibres joining the PFR and PFR-axoneme linking-filaments to Zone 2.

Zone 2: Flagellar-membrane junctional complexes.

Zone 3: Intermembrane staples.

Zone 4: Cell-body-membrane junctional complexes.

Zone 5: Fibres radiating from Zone 4.

Zone 6: Filamentous structure where Zone 5 terminates.

Zone 7: Microtubule Quartet and associated ER.

Zone 8: Fibres linking Zones 7 and 4.

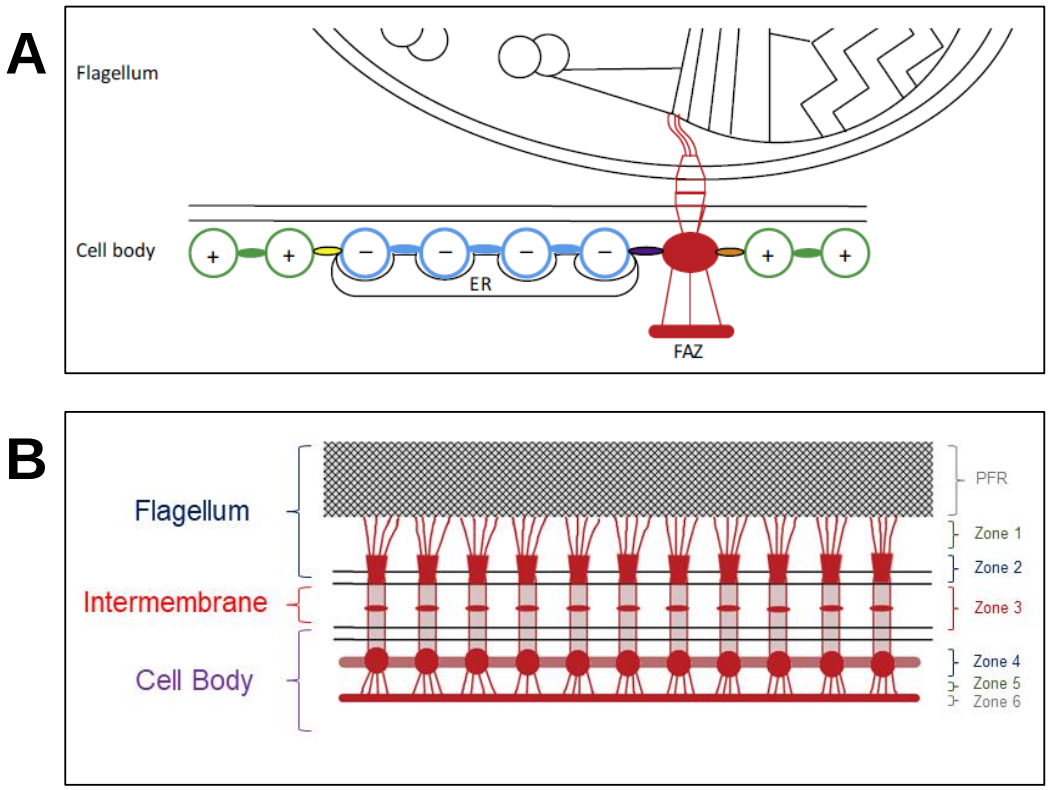


Figure 1.7: Schematic diagrams of the FAZ

(A) Schematic diagram of the FAZ in transverse section, showing the FAZ in red, microtubule quartet (MtQ) in blue and subpellicular microtubules in green. The +/- signs represent the polarity at the posterior end of the cell. Taken from Sunter and Gull (2016).

(B) Schematic diagram of the FAZ in longitudinal section, based on (A) and on the published images of the FAZ. Labels refer to the zones of the FAZ as designated by Sunter and Gull (2016).

1.2.3 The structure of the FAZ

The flagellar attachment zone connects the flagellum skeleton to the cytoskeleton through two membranes. It therefore has a flagellar region, a cell-body region and an extracellular intermembrane region (Sunter and Gull, 2016). It has long been known that, when viewed by transmission electron microscopy, the FAZ is visible as a single row of regularly spaced electron-dense punctate structures, ~95 nm apart (Vickerman, 1969). In longitudinal section, these structures appear to traverse the two membranes, extending filaments into both the flagellum and the cell body (Fig. 1.5) and it has therefore been postulated that these structures are junctional complexes (also called *macula adherens* or staples) that mediate flagellar attachment by acting as molecular rivets (Vickerman, 1969, Höög et al., 2012). However, recent ultrastructural investigations have revealed further intricacies. Indeed, the FAZ is one of the most complex membrane junctions known. Thirty protein components have been identified thus far, and there are expected to be more. To simplify description of the FAZ in transverse section, the region has been divided into eight overlapping zones (Fig. 1.6, 1.7 A) (Sunter and Gull, 2016). Fig. 1.7 (B) shows how these zones are probably arranged in longitudinal section, although there is still some uncertainty in this regard.

The junctional complexes comprise Zone 2 (a row of punctate structures along the flagellar membrane), Zone 3 (a corresponding row of structures in the extracellular intermembrane region) and Zone 4 (a corresponding row of structures just inside the cell-body membrane) (Sunter and Gull, 2016). Filaments (Zone 1) extend from each punctum of Zone 2 and seem to reach the PFR and the PFR-axoneme linking filaments. When Zone 3 was first observed by EM, it was thought to be a row of "staples" that was distinct from the junctional complexes (Höög et al., 2012). It has since become evident that the "staples" are simply the middle part of the junctional complexes, so the two terms are now used interchangeably (Moreira et al., 2017, Sun et al., 2018). Aside from the junctional complexes, there does not seem to be any other connection between the flagellar membrane and the cell-body membrane. In other words, there is no continuous join between the two membranes along the FAZ.

In contrast, the Zone 4 punctae appear to be connected to each other, like pearls in a necklace (Sunter and Gull, 2016). These punctae seem to be the most easily visible part of

the FAZ (Fig. 1.5 C). Zone 5 consists of fibres that radiate from the Zone 4 punctae. Zone 6 is a filamentous structure where Zone 5 ends. It is possible that Zone 6 forms a continuous line along the FAZ, but this is not clear. Collectively, Zones 4, 5 and 6 are termed the FAZ filament domain. Although many proteins have been localised to "the FAZ filament" (see below), it is not clear whether these proteins belong in Zone 4, 5 or 6.

To the left of the FAZ filament (when the cell is viewed from the posterior end), there is a distinctive structure called the microtubule quartet (MtQ) (Zone 7). This is a set of four microtubules that are antiparallel to all the other microtubules of the subpellicular corset (i.e., they have their + end at the anterior end of the cell instead of at the posterior end). The MtQ is tightly associated with the smooth ER. Zone 8 consists of fibres linking the MtQ to Zone 4 (Sunter and Gull, 2016).

1.2.4 The proteins of the FAZ

All the FAZ proteins discovered thus far are kinetoplastid-specific. Except for the FLA glycoproteins, which are the subject of this thesis, all known FAZ proteins are constitutively expressed in both PCF and BSF life-cycle stages. Most have only been investigated in the procyclic form, however.

Most known FAZ proteins are intracellular, apart from FAZ5, FS179, the FLA glycoproteins and, possibly, FAZ3 and FAZ9. The FAZ5 protein has eight predicted transmembrane domains; since it was discovered by Co-IP with intracellular FAZ proteins, it probably belongs to the cell-body membrane of the FAZ, rather than the flagellar membrane (Sunter et al., 2015b). FAZ3 and FAZ9 have predicted signal peptides but no TM domains. FS179 is a putative calcium channel that locates to the flagellar membrane of the FAZ (Oberholzer et al., 2011). Knockdown of FS179 causes detachment of the newly synthesised flagellum (at least in BSFs), consistent with a previously observed requirement for Ca^{2+} in flagellar attachment (Vickerman, 1969). The FLA glycoproteins are discussed in detail in Section 1.3.

Of the known intracellular FAZ proteins, most locate to the cell-body side of the FAZ (i.e. the FAZ filament domain), including FAZ1–10, FAZ12–14, CC2D and TbVAP (Sunter and Gull, 2016, Moreira et al., 2017). Some of these proteins locate mainly to one end of the FAZ filament (e.g. FAZ1, 3 and 9 towards the proximal end and FAZ4, 6 and 7 towards the

distal end), while others are found all along the filament, such as FAZ2, 5 and 8 (Sunter et al., 2015b). FAZ2, CC2D and KMP11 are known to form a complex with one another (Zhou et al., 2015). The known intracellular FAZ proteins are all large cytoskeletal proteins. Most have no recognisable structural or functional domains: a notable exception is FAZ9, which has armadillo repeats like those of plakoglobin, an intracellular component of the desmosomes of multicellular organisms (Sunter et al., 2015b). Furthermore, a recent study has suggested that FAZ1 and FAZ10 have desmoplakin domains, which are also a feature of the cytoplasmic plaques of desmosomes (Trépout, 2020). It has been observed previously that the junctional complexes of the FAZ seem to be structurally analogous to desmosomes (Sun et al., 2013). However, no homologues to the cadherin proteins, which form the extracellular region of the desmosome, have been found in the *T. brucei* genome (Sunter and Gull, 2016).

The intracellular FAZ proteins have all been investigated by immunofluorescence in PCF cells. Ordinary confocal microscopy gives insufficient resolution to determine whether the distributions of these proteins along the FAZ are punctate (which would suggest localisation to the junctional complexes) or continuous, although in some cases a punctate distribution seems likely from the published images (e.g. FAZ9, FAZ17) (McAllaster et al., 2015). FAZ1 and CC2D have been investigated by electron microscopy and appear to form a continuous line along the FAZ (Zhou et al., 2011). The only intracellular protein that has been explicitly stated to belong to the junctional complexes is FAZ10 (Moreira et al., 2017). Immunogold electron microscopy of FAZ10 showed the protein to have a punctate distribution on the cell-body side of the FAZ. FAZ10 is remarkable for its size. At ~2000 kDa, FAZ10 is a high molecular mass protein (HMMP) that may act as a scaffold for other proteins. Knockdown of FAZ10 leads to flagellar detachment, misplacement of the cleavage furrow, and cytokinesis defects.

Of the known intracellular FAZ proteins, only two, FLAM3 and ClpGM6, locate to the flagellar side of the FAZ (probably Zone 1) (Rotureau et al., 2014, Hayes et al., 2014, Sunter et al., 2015a). At 468 kDa and ~1500 kDa, respectively, FLAM3 and ClpGM6 are cytoskeletal HMMPs (Moreira et al., 2017). It has been hypothesised that the two proteins form a complex with one another as FLAM3 is required for ClpGM6 stability and localisation (Sunter et al., 2015a).

It has been observed in procyclic cells that knockdown of flagellar FAZ proteins has a different effect to that of FAZ-filament proteins. Knockdown of FAZ-filament proteins tends to interfere with either flagellar attachment or nuclear positioning (Sunter and Gull, 2016). For example, FAZ1 RNAi causes disorganisation of the new FAZ filament and, hence, partial flagellar detachment (Vaughan et al., 2008). This, in turn, leads to cytokinesis errors such as formation of multinucleate cells and anucleate cytoplasts, which shows the importance of the FAZ for cell division. FAZ2 RNAi causes complete flagellar detachment, extreme shortening of the new FAZ, and cell death (Zhou et al., 2015). (Note that FAZ2 depletion destabilises other FAZ-filament proteins, even those that do not bind directly to FAZ2, so these effects are probably due to general defects in FAZ assembly, rather than to specific loss of FAZ2 activity). The FAZ-shortening effect is very interesting because it highlights the interdependence between flagellum assembly and FAZ assembly. It has been shown that an intact flagellum is necessary for full-length FAZ assembly (at least in the procyclic form). Initial nucleation of the FAZ filament occurs independently, but when flagellum assembly is blocked (by knockdown of intraflagellar transport, for example), the FAZ filament does not extend beyond a short stub (Kohl et al., 2003, Sunter and Gull, 2016). Of course, since the FAZ specifies the site of ingression of the cleavage furrow, a short FAZ means a truncated daughter cell (Vaughan, 2010).

Interestingly, knockdown of the proteins of the flagellum side of the FAZ produces a very different effect – at least in procyclic cells. Knockdown of either FLAM3 or ClpGM6 produces an unusual epimastigote-like phenotype (Sunter and Gull, 2016). The kinetoplast (and hence, the adjacent basal body) is anterior to the nucleus, which means that the flagellar attachment zone is much shorter, and the region of free flagellum is much longer. This is not a true transition to epimastigote form, since normal epimastigote markers are not expressed. The effect on cell growth varies, but certain epimastigote-like ClpGM6-RNAi cells proliferate indefinitely at a normal rate (Hayes et al., 2014). Note that this morphology switch does not occur in BSF cells. FLAM3 RNAi in BSFs causes flagellar detachment and cytokinesis defects, leading to cell death (Sunter et al., 2015a).

It is clear, then, that FAZ proteins from both sides of the membrane junction are necessary for correct FAZ function. However, the most informative insights into the mechanism of flagellar attachment have come with the discovery of the flagellar-adhesion glycoproteins

(FLA), which are the subject of this thesis. The FLA glycoproteins are largely extracellular and are mounted on the membranes of the cell body or flagellum. They are likely to be important components of the intermembrane region of the junctional complexes. The current model of flagellar attachment proposes that FLA proteins on the two juxtaposed membranes bind to each other across the intermembrane space.

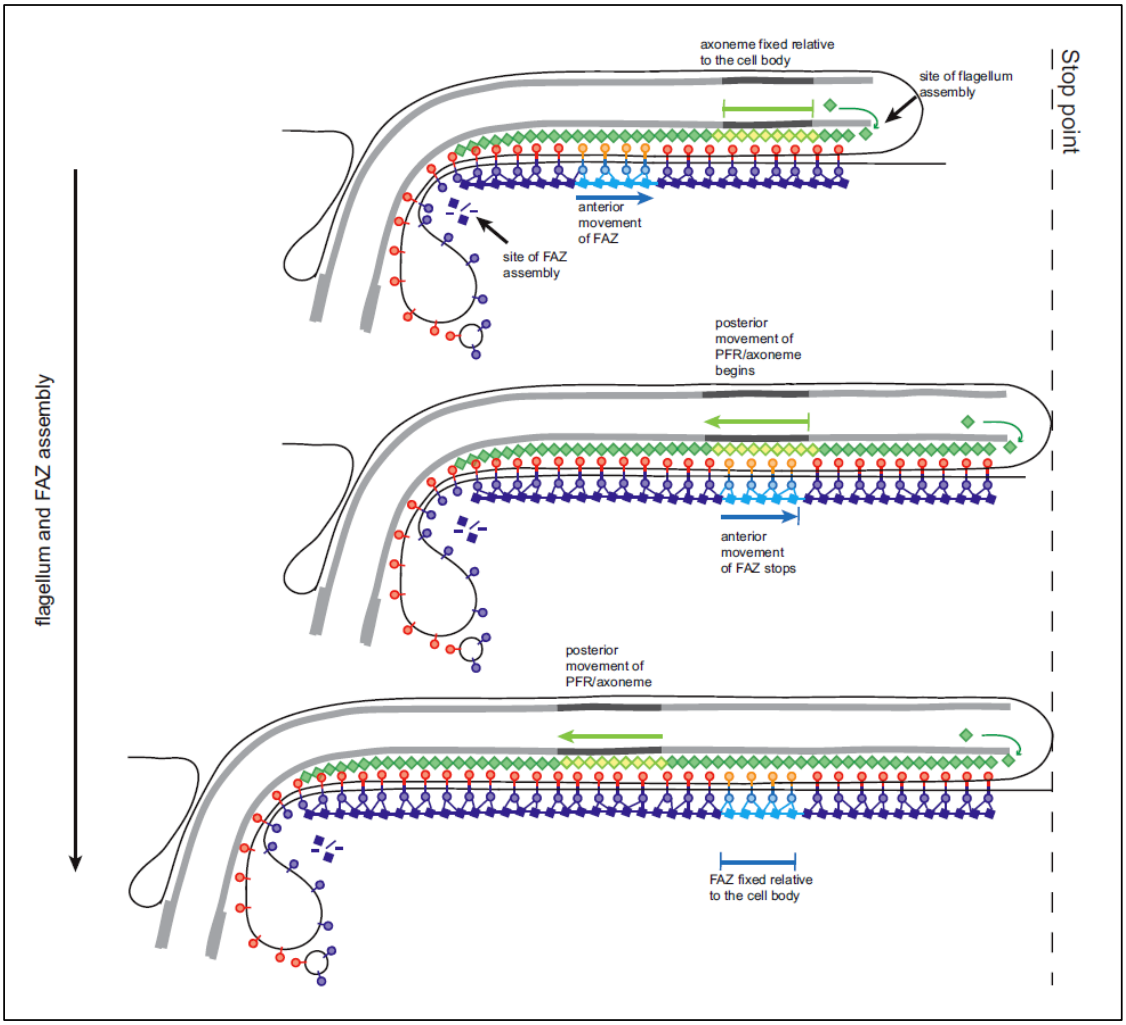


Figure 1.8: Co-ordinated assembly of flagellum and FAZ (Sunter *et al.*, 2015b, Fig. 7) Growth of the new flagellum involves addition of proteins to the distal end. Growth of the new FAZ (on both sides of the membrane junction) involves addition of proteins to the proximal end. These two processes are co-ordinated as shown. Initially, the growing flagellum is anchored in position, while the growing FAZ moves forward as a unit. When the flagellum reaches the "stop point", the growing FAZ remains anchored in position while the growing flagellum shifts backwards as a unit, causing elongation of the posterior end of the cell.

1.2.5 Assembly of the FAZ

Before the roles of the individual FLA proteins can be understood, though, it is necessary to consider in a more generalised way how an intermembrane junction can be constructed. Flagellar-membrane components and cell-body-membrane components must be initially inserted into the flagellar-pocket membrane, since that is the only site of exocytosis. They must then be trafficked separately to their respective locations, which seems to imply that no interactions can occur between them until they are *in situ*. This was a puzzling point at first, because it was assumed that addition of new proteins to the FAZ must occur at the distal end, to match addition of proteins to the growing flagellum. Curiously, though, it was found that addition to the new FAZ occurs solely at the proximal end (Zhou et al., 2015, Sunter et al., 2015b). Several FAZ proteins (including FAZ1–3, FAZ5, FAZ8, KMP11) were tagged with fluorescent tags and localised at early timepoints during the cell cycle. The tagged proteins always localised to the proximal end of the nascent FAZ.

Addition of proteins to the proximal end of the FAZ simplifies the problem of how the junctional complexes bridge the gap between the flagellar and cell-body membranes. Presumably, after sorting occurs in the FP, the flagellar-membrane components and cell-body-membrane components of the junctional complexes can meet in the neck region and their extracellular regions can snap together as appropriate. However, it is difficult to understand how FAZ assembly and flagellum assembly are coordinated, since new FAZ proteins add on to the proximal end of the FAZ while new flagellar proteins add on to the distal end of the flagellum.

The current model to explain this apparent discrepancy proposes a two-stage process (Fig. 1.8). (This model was based on investigations of PCFs but may be relevant to BSFs also.) At first, the growing flagellum extends by addition at its distal end such that a given point on the axoneme stays at a fixed position relative to the cell body, while the proximally assembling FAZ moves its whole length towards the anterior of the cell, to follow the flagellum. The impetus behind this FAZ movement is either a pulling force (because the distal tips of the new flagellum and FAZ are connected by the flagella connector (PCF) or flagellar groove (BSF)) or a pushing force (because the addition of components to the proximal end of the FAZ pushes the rest of the FAZ forwards). Then, when the new

flagellum reaches the stop point, which is a point about 60% along the length of the old flagellum, it does not extend any further in the anterior direction. Instead, the continuing addition of proteins at the distal end pushes the new flagellum in the posterior direction, and the posterior of the cell elongates. During this stage, it is the FAZ that is anchored in position relative to the cell body. Addition onto the proximal end of the FAZ therefore causes the FAZ to elongate in the posterior direction as well (Sunter and Gull, 2016).

This model implies that the new flagellum and FAZ can slide against each other during their formation – while remaining sufficiently well attached that the necessary positional information is not lost. The sliding would have to occur within the flagellar matrix (probably Zone 1 of the FAZ). It is possible that this sliding connectivity works analogously to the actin and myosin fibres in muscles, with multiple weak connections being made and broken consecutively. These connections are probably consolidated once flagellum/FAZ assembly has completed (Sunter and Gull, 2016). Interestingly, there is some evidence that additional FLAM3 protein is integrated into the mature FAZ after assembly of the flagellum has been completed (Sunter et al., 2015a). Thus, it is possible that FLAM3 strengthens the connection between flagellum and FAZ when the sliding process is over.

A recent (pre-published) study in BSF cells has proposed a slightly modified version of the "pull" model, called the "alternative pull" model (Trépout, 2020). This study states that FAZ1 and FAZ10 are predicted to have dynein motor domains, which means that the force required to pull the nascent FAZ in the anterior direction could come from the movement of these proteins along the nearest microtubule of the subpellicular corset, rather than from the movement of the flagella connector/groove. (The microtubules of the subpellicular corset are oriented with the plus ends at the posterior of the cell; the microtubules of the axoneme and the MtQ are oriented with their plus ends at the anterior of the cell.) It is perhaps relevant, too, that on the flagellar side of the FAZ, FLAM3 is known to interact with a kinesin protein, KIN-E (An and Li, 2018).

Once assembly of the flagellum and FAZ are completed, the mature structure is very stable. Maintenance of the composition of the flagellum occurs by intraflagellar transport, but the overall length of the flagellum does not change because it is locked by a mechanism that may involve FLAM8 (Fort et al., 2016, Bertiaux et al., 2018). Unlike desmosomal proteins, which are subject to turnover post-assembly, the proteins of the

FAZ seem to undergo little or no turnover (Sunter et al., 2015b). Consequently, knockdown of FAZ proteins does not usually affect the mature FAZ.

In summary, then: the flagellar attachment zone plays a crucial role in cell division in the African trypanosome by controlling the positioning of the cleavage furrow. The parasite's single-copy organelles are all associated to the FAZ region in some way. For example, the kinetoplast – and hence, the single mitochondrion – is attached to the basal body of the flagellum. The ER is attached to the MtQ, the Golgi apparatus is connected to the bilobe structure at the collar of the FP, and the nucleus is suspected to have some link to the FAZ filament (Sunter et al., 2015b, Vaughan, 2010). Thus, when a division plane forms between the old and new attached flagella, each daughter cell has the correct quota of organelles. However, the mechanism by which the trypanosome's flagellum is attached onto the cell body is still not fully understood, especially in the bloodstream form. Further knowledge of the flagellar-adhesion glycoproteins is required.

1.3 The Discovery of the Flagellar Adhesion Glycoproteins

1.3.1 Discovery of FLA1 (Flagellum Adhesion Glycoprotein 1)

The first flagellar-adhesion glycoprotein to be characterised in *T. brucei* was named flagellum adhesion glycoprotein 1 (FLA1). It was discovered by chance when a plasmid from a *T. b. rhodesiense* cDNA library was sequenced in the course of an unrelated investigation (Nozaki et al., 1996). A gene fragment in the plasmid attracted attention immediately because of its similarity to a known surface glycoprotein, GP72, that mediates flagellar attachment in the closely related South American trypanosome, *Trypanosoma cruzi*. To find the complete gene, the cDNA was used to probe *Bam* HI-digested *T. brucei* genomic DNA on a Southern blot. The cDNA probe cross-hybridised with an 8-kb band under conditions of high stringency. This 8-kb region of the *T. brucei* genome was cloned and sequenced, revealing an ORF (*FLA1)* that was 63% similar to GP72.

The FLA1 protein was predicted to be largely extracellular, with a single transmembrane domain and a short (16 a.a.) cytoplasmic tail. The sequence further suggested that it had a signal peptide (30 a.a.) which, when cleaved, would leave a mature protein of 57 kDa (516 a.a.). There were seven putative *N*-glycosylation sites, and a proline/threonine-rich region

suggestive of *O*-glycosylation. These features all resembled those of GP72, indicating that FLA1 might have a similar function – flagellar attachment (Nozaki et al., 1996).

Further investigations suggested that there was only a single copy of *FLA1* in the *T. brucei* genome. Genomic DNA was digested with several restriction enzymes, Southern blotted and probed with a short stretch of DNA from the middle of *FLA1*: each time, the probe recognised a single band only. However, upon *Hin* dIII digest, a double band was observed. This was interpreted to mean that there was a *Hin* dIII polymorphism downstream of the gene, and that the cells in question (Lister 427) had alleles that differed at that site (Nozaki et al., 1996). Subsequent sequencing and publishing of the entire *T. brucei* genome (Berriman et al., 2005) has confirmed that there is a single copy of *FLA1*, annotated as Tb927.8.4010 and situated on chromosome 8.

To find out whether *FLA1* was transcribed, the polyadenylated RNA of procyclic and bloodstream-form trypanosomes was northern blotted and probed with a segment of *FLA1*. A streaky band of approximately 3.5 kb confirmed expression of mature *FLA1* mRNA in both PCFs and BSFs. (Recent results that challenge this statement will be discussed in Section 5.3.) Phosphorimager quantitation measured seven times more *FLA1* mRNA in PCFs compared to BSFs (Nozaki et al., 1996).

To generate an antibody against FLA1, a region of *FLA1* (encoding amino acids 81–312) was amplified by PCR using primers that introduced *Bam* HI sites. The amplicon was inserted into a *Bam* HI-digested plasmid that added a histidine tag. The plasmid was transformed into *E. coli* to amplify the peptide. The peptide was purified on a nickel column and injected into rats. Antiserum was affinity-purified with the histidine-tagged FLA1 peptide (Nozaki et al., 1996). (The specifics of the protocol used to create this anti-FLA1 antibody have implications that will be discussed in Section 5.3.)

Western blotting with the polyclonal anti-FLA1 antibody revealed a 100-kDa band in BSF lysates and an 80-kDa band in PCF lysates. The substantial increase in molecular weight compared to the predicted 57 kDa was hypothesised to signify extensive glycosylation; the marked difference between life-cycle stages was interpreted to mean that FLA1 is differentially glycosylated in BSFs and PCFs. To verify this hypothesis, cell lysates were treated with PNGase F, to remove *N*-glycans and trifluoroacetic acid (TFA), to remove *O*-glycans. In PCF lysates, PNGase F reduced the FLA1 band from 80 kDa to 60 kDa, but TFA

had no effect. In BSFs, treatment with *either* PNGase F or TFA reduced the FLA1 band from 100 kDa to 80 kDa; treatment with *both* reduced the band to 60 kDa. This clear-cut result implied that procyclic FLA1 contained conventional *N*-glycans only, while bloodstream-form FLA1 contained acid-labile *O*-glycans as well (Nozaki et al., 1996).

Localisation of FLA1 was determined by immunofluorescence using the anti-FLA1 antibody. In both PCFs and BSFs, signal was concentrated along the attached flagellum and in the flagellar pocket. This localisation, combined with the fact that FLA1 acted as a membrane protein (i.e. soluble by detergent treatment, but insoluble upon freeze-thaw fractionation), strengthened the hypothesis that FLA1 played a role in flagellar attachment (Nozaki et al., 1996). However, initial functional studies encountered difficulties.

FLA1 knockout constructs were generated using a plasmid that contained the 1.7-kb *FLA1* CDS, flanked by 2.2 kb of the upstream region and 0.9 kb of the downstream region. Outward-facing primers were used to amplify the whole plasmid except the *FLA1* CDS. An antibiotic-resistance gene was ligated into the resulting amplicon, creating a knockout vector. Prior to transfection in trypanosomes, the vector was restriction-digested to release a linear stretch of DNA consisting of the 2.2-kb upstream region, the resistance gene and the 0.9-kb downstream region. Upon insertion into the trypanosome genome by homologous recombination, the endogenous *FLA1* gene was replaced by the resistance gene, but (in theory, at least) the sequence of the flanking regions remained the same as before, to avoid disruption to neighbouring genes (Nozaki et al., 1996).

When a single *FLA1* allele was knocked out in this manner, however, no phenotypic effect was observed in either PCFs or BSFs. Efforts to knock out both alleles simultaneously were unsuccessful. As mentioned above, when *T. brucei* genomic DNA is digested with *Hin* dIII and probed for *FLA1*, a double band is observed. When *FLA1* was knocked out, loss of either band could be observed – but never both. It was concluded that *FLA1* is an essential gene in both PCFs and BSFs, meaning that double knockout of the two alleles is catastrophic to cell proliferation (Nozaki et al., 1996).

In short, the first study of FLA1 showed that this membrane glycoprotein localised to the FAZ and was, apparently, important for both proliferative life-cycle stages, but the function of the protein could not be stated with certainty. The advent of RNAi technology facilitated the first confirmation of the role of FLA1 in flagellar attachment. Nucleotides

318–684 of the *FLA1* cDNA were amplified by PCR and inserted into a transient knockdown vector in procyclic-form cells (LaCount et al., 2000). It was observed that approximately half of the cells had at least one detached flagellum. However, transient RNAi was too brief to see whether there was an effect on cell proliferation.

A follow-up study generated a stable FLA1-RNAi system, and its findings shed a whole new light on the role of FLA1. A 1007-bp fragment of the *FLA1* gene and 5′ UTR (nucleotides -94 to 913) was amplified by PCR and inserted into an RNAi vector, p2T7 (LaCount et al., 2002). The vector inserted stably into the ribosomal DNA of PCF and BSF trypanosomes, enabling tetracycline-induced conditional knockdown of FLA1. In both life-cycle stages, the induction of *FLA1* dsRNA caused flagellar detachment and growth arrest. Procyclic-form cells doubled in cell density about twice after addition of tetracycline, and then ceased to divide, confirming that FLA1 is essential in PCFs. Bloodstream-forms cells failed to double even once after addition of tetracycline. In both life-cycle stages, induced cells developed multiple nuclei (sometimes more than ten), multiple kinetoplasts and, probably by consequence, an abnormal rounded morphology. It was concluded that although FLA1 is required for cytokinesis, it is not required for mitosis or for kinetoplast replication.

Given the apparent similarity between FLA1 and the *T. cruzi* protein, GP72, in terms both of structure and function, it was postulated that GP72 might be able to substitute for FLA1 if introduced ectopically to *T. brucei* (LaCount et al., 2002). Surprisingly, GP72 caused flagellar detachment when expressed in normal procyclic *T. brucei*, and it failed to rescue the FLA1-knockdown phenotype in RNAi lines. Furthermore, these cells continued to divide despite having detached flagella (although at a 2/3-fold reduced rate). This was a very interesting result, because it showed that flagellar attachment is not essential for the viability of procyclic *T. brucei* in culture. Why, then, does knockdown of FLA1 cause total growth arrest? FLA1 must have another role in cytokinesis apart from its role in flagellar attachment. It is not clear if these separable functions represent discrete biochemical properties or distinct concentration dependences (i.e. high levels of FLA1 may be needed for flagellar attachment, but lower levels may be required for cytokinesis). The mechanism by which GP72 causes flagellar detachment in *T. brucei* remains to be determined. GP72 may oligomerise with FLA1 and, by acting as a dominant negative, prevent proper

localisation or function of FLA1. Alternatively, it may bind to FLA1-interacting proteins and prevent their binding to FLA1 (LaCount et al., 2002). No experiments have been performed to investigate the effect of GP72 in BSFs.

It was speculated that the role of FLA1 in cytokinesis may be in specifying the location of the FAZ. The FAZ is improperly formed when FLA1 is knocked down (Moreira-Leite et al., 2001). Since only 16 amino acids of FLA1 are predicted to be intracellular, it is unlikely to play an extensive structural role in the FAZ filament, but it could provide it with positional information (LaCount et al., 2002).

1.3.2 Discovery of FLA2 (Flagellum Adhesion Glycoprotein 2)

An unexpected consequence of the FLA1-RNAi experiments just described was the discovery of another flagellar-adhesion glycoprotein, which was named flagellum adhesion glycoprotein 2 (FLA2).

When knockdown of *FLA1* in bloodstream-form cells was verified by northern blot, it was observed that two transcripts were ablated by the *FLA1* dsRNA (LaCount et al., 2002). The smaller transcript (~3 kb) corresponded in size to the transcript that was knocked down in procyclic-form cells, evidently corresponding to *FLA1* mRNA. The larger transcript (~3.5 kb), however, was not detected during northern blotting of procyclic-form cells. A search of the genome revealed a gene predicted to encode a protein 61% identical (and 73% similar) to FLA1, which was thus named *FLA2*.

FLA1 and *FLA2* are virtually identical over their first 300 nucleotides. This region was included in the *FLA1* RNA probe used in the northern blot, which explains why the *FLA2* bands were visible. It was also included in the original *FLA1* fragment used in the RNAi experiments of LaCount *et al*. Thus, the phenotypes generated by FLA1 RNAi in BSFs, as described above, may be due to loss of FLA2 alone, loss of FLA1 alone, or concomitant loss of both proteins (LaCount et al., 2002). Indeed, the discovery of FLA2 calls into question whether all the results pertaining to FLA1 in the bloodstream form might more accurately be ascribed to FLA2. This possibility is discussed in Section 5.3.

There are two copies of the *FLA2* gene (Tb927.8.4060/4110), which are 98% identical. *FLA2* appears to have arisen as a gene-duplication event on chromosome 8 that also included *FLA1* and *FLA1BP* (Fig. 1.9) (LaCount et al., 2002).

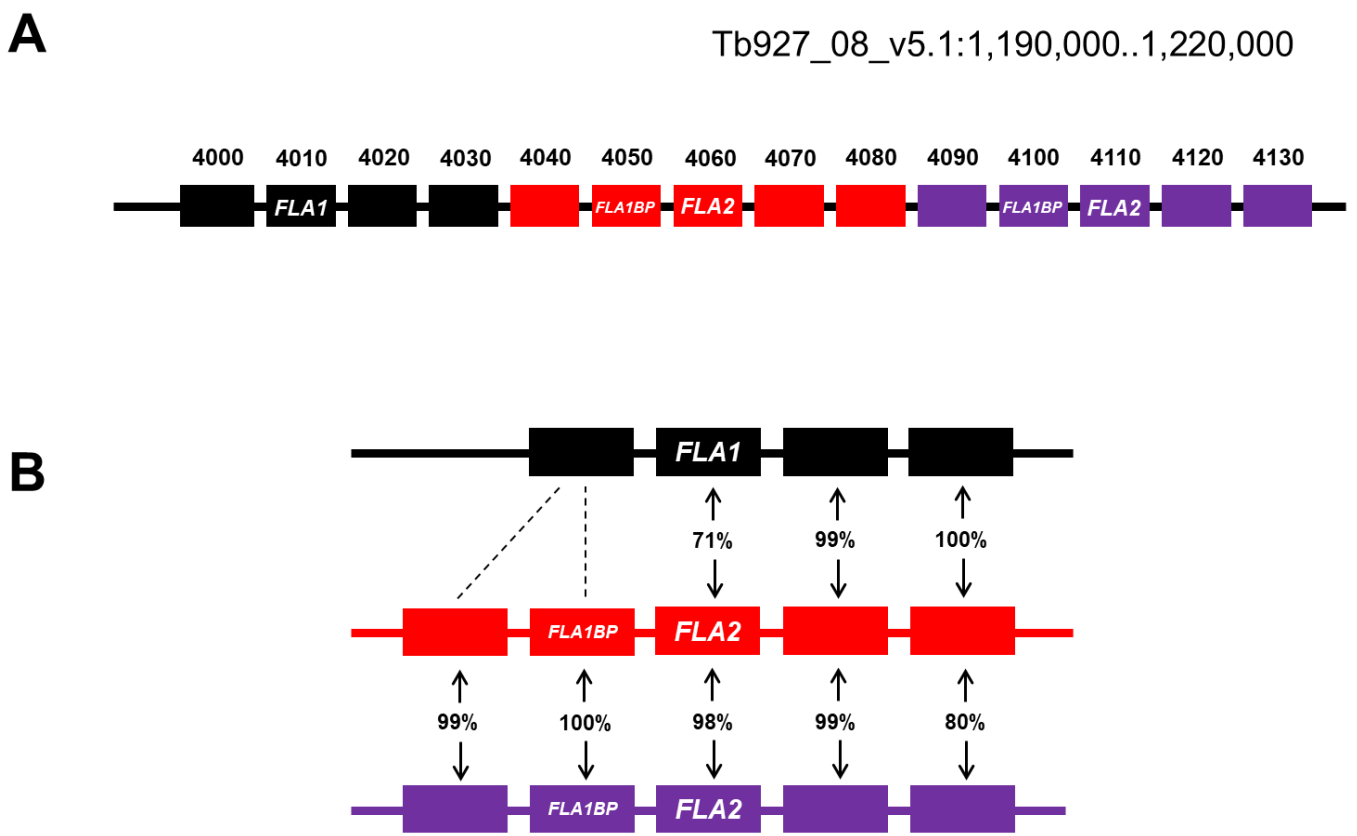


Figure 1.9: *FLA1*, *FLA2* and *FLA1BP* form part of a gene duplication site on chromosome 8.

(A): The indicated region of *T. brucei* chromosome 8 contains the genes Tb927.8.4000 – Tb927.8.4130 (labelled here as 4000–4130), including *FLA1*, *FLA2* and *FLA1BP.* **(B):** Alignment of the same genes, with sequence identities given as percentages, indicates that gene duplication has occurred in this region. The divergence between *FLA1* and *FLA2* is interesting because it suggests adaptation to slightly different roles. The Tb927.8.4000 gene contains a 92-bp region that is 100% identical to a region of *FLA1BP,* and a 201-bp region that is 97% identical to a region of Tb927.8.4040/4090.

The two copies of the *FLA2* gene are sometimes referred to as *FLA2* and *FLA3*, respectively. This is unnecessary, since they encode virtually identical forms of the protein. Furthermore, it risks confusion with another gene named *FLA3*, Tb927.5.4570/80. BLAST analysis shows that the *FLA2* genes have homologues in other trypanosome species, but not in more distantly related taxa (Table 1.1).

Both genes for *FLA2* predict a type I membrane protein of 64 kDa, with a putative N-terminal signal sequence, a large N-terminal extracellular domain and a short C-terminal transmembrane domain. A striking feature is a 44-amino-acid proline-rich insertion in the middle of the extracellular domain, which is not present in *FLA1* (LaCount et al., 2002). The localisation of FLA2 remains unknown, as no specific reagents or cell lines have been developed to address this question.

1.3.3 Discovery of FLA3 (Flagellum Adhesion Glycoprotein 3)

The fortuitous discovery of a third essential glycoprotein in the flagellar attachment zone, during a study of the trypanosome's endocytic pathway, led to fresh speculation as to how the intermembrane region of the FAZ mediates flagellar attachment.

It had long been known that in *T. brucei*, glycoproteins containing linear poly-*N*-acetyllactosamine (poly-LacNAc) side chains $(\beta\text{-Gal-}(1\rightarrow4)\text{-}\beta\text{-GlcNAc-}(1\rightarrow3))_n$ were confined to the flagellar pocket and endocytic pathway, and were restricted to the bloodstream form of the parasite (Nolan et al., 1999). An attempt was made to characterise this set of glycoproteins in order to elucidate the function of the poly-LacNAc glycans, which were hypothesised to act as a sorting signal for endocytosis. The set of proteins in question were isolated as a tomato-lectin-binding fraction (tomato lectin binds specifically to *N*-acetyllactosamine). Antibodies raised against this lectin-binding fraction were used to probe a trypanosome cDNA library, thus identifying the cDNAs of the complete set of linear-poly-LacNAc glycoproteins (Nolan et al., 1999, Rubotham et al., 2005).

One cDNA attracted attention because it appeared to encode a protein with two transmembrane domains – a highly unusual configuration. Upon characterisation, the protein was found to be essential, not for endocytosis, but for flagellar attachment. It was

therefore named FLA3, despite bearing no sequence similarity to FLA1/2 (Woods, 2007, Woods et al., 2013).

Species	Gene(s) related to *FLA2*	E Value	Identity
Trypanosoma brucei brucei	Tb927.8.4060 (*FLA2)* Tb927.8.4110 (*FLA2)* Tb927.8.4010 (*FLA1)*	0 0 1.4e-174	100% 98% 71%
Trypanosoma brucei gambiense	Tbg972.8.3760 Tbg972.8.3810	0 1.9e-171	99% 70%
Trypanosoma congolense	TcIL3000_8_3780.1 TcIL3000_8_3810.1 TcIL3000_0_31510.1	6.6e-115 1.5e-55 9.8e-55	62% 66% 66%
Trypanosoma vivax	TvY486_0803430	1.4e-42	60%
Trypanosoma cruzi	TcCLB.509561.20 TcCLB.503571.10	2.0e-25 1.1e-23	59% 58%

Table 1.1: BLASTn of Tb927.8.4060 (*FLA2*) A BLASTn search using the GeneDB BLAST tool shows the genes related to *FLA2* in *T. brucei* and other related trypanosomatids. The Identity column shows the percentage of nucleotides that match Tb927.8.4060, while the E value gives the probability that the perceived identities are due merely to chance.

It is almost certain that *FLA1* is a paralogue of *FLA2.*

It also seems that *T. brucei gambiense* has one close orthologue of *FLA2* (Tbg972.8.3760) and another of *FLA1* (Tbg972.8.3810). The other trypanosomes have increasingly dissimilar versions of the genes, according to their phylogenetic distance from *T. brucei*.

A further BLAST was performed using the NCBI BLASTn function, which extended the search to all species in the database. This did not reveal any other highly similar genes, indicating that *FLA2* is kinetoplastid-specific.

Species	Gene(s) related to *FLA3*	E Value	Identity
Trypanosoma brucei brucei	Tb927.5.4570 (*FLA3)* Tb927.5.4580 (*FLA3)* Tb.927.8.4050 (*FLA4)* Tb927.8.4100 (*FLA4)*	0 0 2.9e-71 2.9e-71	100% 93% 61% 61%
Trypanosoma brucei gambiense	Tbg972.5.6180 Tbg972.5.6160 Tbg972.8.3750 Tbg972.8.3800	0 0 8.7e-70 1.9e-65	94% 98% 60% 60%
Trypanosoma congolense	TcIL3000_0_35140.1 TcIL3000_0_17090.1 TcIL3000_0_31520.1	8.6e-27 8.6e-27 6.6e-19	69% 69% 65%
Trypanosoma vivax	TvY486_0803420	2.8e-12	61%
Trypanosoma cruzi	TcCLB.478283.10 pseudogene	1.4e-13	59%

Table 1.2: BLASTn of Tb927.5.4570 (*FLA3*) A BLASTn search using the GeneDB BLAST tool shows the genes related to *FLA3* in *T. brucei* and other related trypanosomatids.

It is very probable that *FLA1BP* (*FLA4*) is a paralogue of *FLA3.*

It seems that *T. b. gambiense* has two orthologues of *FLA3* (Tbg972.5.6180/6160) and two of *FLA1BP* (Tbg972.8.3750/3800). The other trypanosomes have increasingly dissimilar versions of the genes, according to their phylogenetic distance from *T. brucei*. Indeed, *T. cruzi* appears to have lost the ability to express its homologue of *FLA3*, as the gene is annotated as being a pseudogene.

A further BLAST was performed using the NCBI BLASTn function. This did not reveal any other highly similar genes, indicating that *FLA3* is kinetoplastid-specific.

Encoded by a pair of genes (Tb927.5.4570/4580) on chromosome 5, FLA3 is 818 a.a. in length, with a predicted molecular weight of 89 kDa (Woods et al., 2013). BLAST analysis shows that the *FLA3* genes have homologues in other trypanosome species, but not in more distantly related taxa (Table 1.2). The genes encode two versions of the protein that are 91% identical (94% similar).

Analysis of the primary structure suggests that most of the protein is extracellular, with many *N*-glycosylation sites. The protein is not predicted to have a cleavable N-terminal signal sequence. The putative TM domains are at residues 22–45 and 750–773, respectively. The putative C-terminal membrane-spanning region is followed by four basic residues, suggesting that the C-terminal tail (45 a.a.) is located on the internal side of the membrane (Woods et al., 2013). Woods et al. proposed that both transmembrane domains are embedded in the same membrane such that, topologically, FLA3 forms an extracellular U-shape with two intracellular tails. However, it remains possible that the predicted "N-terminal TM domain" is actually an atypical signal sequence. In that case, FLA3 would be entirely extracellular apart from its C-terminal tail, analogously to FLA1/2 (discussed further in Section 3.3.8).

Expression analysis confirmed that FLA3 is bloodstream-stage-specific. Northern blots of total BSF and PCF mRNA were probed with a fragment (nucleotides 1174–1792) common to both *FLA3* genes; a ~3.4-kb band was observed in the BSF sample, but no bands were observed in the PCF sample. Furthermore, qRT-PCR analysis gave a BSF:PCF relative expression ratio of approximately 8 for *FLA3* mRNA (Woods et al., 2013).

A polyclonal antibody was raised against the C-terminal 18 amino acids of FLA3 (a region common to both *FLA3* genes) and affinity-purified with the peptide antigen. Western blots probed with the antibody showed a single ~160-kDa band in BSF lysates and no specific bands in PCF lysates. The large difference between the observed (160 kDa) and predicted (89 kDa) molecular weights was suggestive of extensive post-translational modification. Deglycosylation experiments (using PNGase F) demonstrated that removal of *N*-glycans reduced the FLA3 band to 100-120 kDa, indicating that the protein is heavily *N*-glycosylated. Fractionation by mechanical disruption with glass beads followed by detergent solubilisation confirmed that FLA3 behaved as a membrane protein (Woods et al., 2013).

Immunofluorescence with the polyclonal anti-FLA3 antibody showed strong signal along the flagellar attachment zone, with a weaker and more diffuse background signal throughout the cell body. The FLA3 signal along the FAZ appeared punctate, although the resolution was inadequate for certainty on this point. In many cells, FLA3 signal was most intense at the posterior and anterior extremities of the flagellar attachment zone (Woods et al., 2013).

In cells where the flagellum became partially detached during fixation, the FLA3 signal was always found on the cell body rather than on the detached region of the flagellum (Woods et al., 2013). This led Woods et al. to suppose that FLA3 is anchored in the cell-body side of the FAZ, rather than in the flagellar membrane. However, recent results have called this conclusion into question (discussed in Section 3.3.6).

To determine the function of FLA3, a conditional knockdown vector was generated (Woods et al., 2013). Nucleotides 1174–1792 were amplified by PCR and inserted into a p2T7.177 RNAi vector. Upon transfection of the vector into BSFs, *FLA3* dsRNA was induced by addition of tetracycline. Knockdown of FLA3 resulted in flagellar detachment. Even after 6 h, 25% of the cells had a detached flagellum. After 24 h, practically all the cells had multiple detached flagella. Only newly constructed flagella were affected; pre-existing flagella remained attached as normal. Detached flagella continued to beat rapidly, but directional motility was lost. Growth arrest occurred 8–12 h after knockdown was induced, and cell death occurred after 2–3 days – confirming that FLA3 is essential in the bloodstream form. The effect on growth seemed to be at the level of cytokinesis: there was a progressive increase in large, multinucleate cells with aberrant morphology (Woods et al., 2013, Barry, 2013).

Unsurprisingly, induction of *FLA3* dsRNA in procyclic cells had no effect (Woods et al., 2013).

The effect of FLA3 on the FAZ filament was tested by probing FLA3-RNAi cells for the FAZ1 protein (Woods et al., 2013, Barry, 2013). When FLA3 was knocked down, a single line of FAZ1 remained associated with the attached flagellum. However, there was no corresponding line of FAZ1 alongside the detached flagellum – instead, the FAZ1 signal was mislocalised at the posterior of the cell. Furthermore, electron microscopy revealed that in FLA3-RNAi cells, detachment of the flagellum was accompanied by loss of the

regularly spaced FAZ structures. These findings suggested that a new FAZ filament could not be formed in the absence of FLA3.

It was postulated that FLA3 might influence construction of the FAZ filament via its short cytoplasmic tail (Woods et al., 2013). However, this is not necessarily the case. It is well known that perturbation of normal growth or attachment of the flagellum can have knock-on implications for the cytoplasmic FAZ filament (Kohl et al., 2003, Vaughan, 2010, Sunter and Gull, 2016). For example, in BSFs, knockdown of the known intraflagellar FAZ protein, FLAM3, causes flagellar detachment and cytokinesis failure, even though it can have no direct interaction with the FAZ filament (Sunter et al., 2015a). Therefore, the mislocalisation of FAZ1 upon FLA3 knockdown could be due directly to flagellar detachment, rather than to a specific interaction between FLA3 and the FAZ filament.

To summarise, then: Woods et al. demonstrated that FLA3, a large glycosylated membrane protein, was essential in the bloodstream form of *T. brucei* because it played a crucial, though as yet uncharacterised, role in flagellar attachment. The purpose of the extensive glycosylation of FLA3 remained unclear, but it was hypothesised that the glycans could play a direct role in mediating flagellar attachment (Woods et al., 2013), a possibility discussed in Section 4.3.4. Another major unanswered question was the orientation of FLA3 (and, indeed, FLA1/2) within the FAZ. Considerable light was shed on this perplexing point with the discovery of a fourth flagellar adhesion glycoprotein – FLA1BP, a putative homologue of FLA3.

1.3.4 Discovery of FLA1BP (FLA1-Binding Protein)

A BLAST search of FLA3 in the *T. brucei* genomic database revealed a protein sharing 45% identity (63% similarity) with FLA3 (Woods, 2007). This protein was initially given the name "FLA4", but subsequent research has dubbed it FLA1BP (Sun et al., 2013), so that is the name that will be used here. There are two identical copies of the *FLA1BP* gene, annotated as Tb927.8.4050 and Tb927.8.4100, respectively. Interestingly, *FLA1BP* is found on chromosome 8 with *FLA1* and *FLA2*, rather than on chromosome 5 with *FLA3*. In fact, the two *FLA1BP* genes evidently arose during the same gene-duplication event that generated the two copies of *FLA2* (Fig. 1.9).

Analysis of the *FLA1BP* sequence predicts a 750-a.a. protein of 83 kDa (Woods et al., 2013). The protein lacks several regions of insertion found in FLA3. Like FLA3, FLA1BP is annotated in the database as having two transmembrane domains, one near the N-terminus (residues 20–43) and one near the C-terminus (residues 687–710), with a large extracellular domain and a short (40 a.a.) intracellular C-terminal tail. However, the putative N-terminal TM domain has also been interpreted to be a cleavable signal sequence (Sun et al., 2013). As was the case for the other FLA glycoproteins, FLA1BP has putative homologues in other trypanosomatid species, but not in more distantly-related taxa (Table 1.2) (Woods et al., 2013).

Analysis of *FLA1BP* mRNA levels by qRT-PCR showed that transcription of *FLA1BP* is upregulated in the procyclic form, with a BSF:PCF expression ratio of 0.2 (Woods et al., 2013, Schöttler, 2010).

To determine the function of FLA1BP, a fragment of the *FLA1BP* ORF (nucleotides 173–979) was amplified by PCR, ligated into the p2T7.177 RNAi vector, amplified in *E. coli* and transfected into PCF trypanosomes (Woods, 2007, Tsang, 2012). An 85% knockdown of *FLA1BP* mRNA in procyclic cells caused the doubling rate to halve. New flagella detached from the cell body, but no multinucleate cells were observed, indicating that cytokinesis was not affected. This result contrasted with the effects of FLA1 knockdown (flagellar detachment with major cytokinesis defects), but corroborated the finding that in procyclic cells, flagellar detachment can be induced without blocking cytokinesis (LaCount et al., 2002). It was concluded that FLA1BP is essential for flagellar attachment in PCFs; the possibility remained that low levels of FLA1BP might be necessary for cytokinesis to occur (Tsang, 2012).

Using the same method, FLA1BP knockdown was induced in bloodstream-form cells. No effect was observed on morphology or proliferation (Woods, 2007, Tsang, 2012). The knockdown rate achieved was only 55% (as measured by qRT-PCR) but given that the wildtype level of *FLA1BP* mRNA in BSFs was already very low, it seemed highly probable that FLA1BP does not play a role in the bloodstream form.

FLA1BP was also discovered, quite independently, during an investigation into the localisation of FLA1 within the FAZ (Sun et al., 2013). This study was the key to determining how FLA1 and FLA1BP mediate flagellar attachment in the procyclic form.

1.3.5 Localisation of FLA1 within the FAZ

At this point, the critical question with regards to all four flagellar-adhesion glycoproteins was where they belonged within the complex FAZ structure. It was clear that they were all largely extracellular, but were they anchored in the flagellar membrane, the cell-body membrane – or both?

To determine the precise localisation of FLA1, YFP-tagged FLA1 mutants were stably overexpressed in PCFs and examined by fluorescence microscopy (Sun et al., 2013). Firstly, it was verified that FLA1 truncation mutants localised correctly to the surface-exposed region of the FAZ: the N-terminal signal sequence ensured that correct localisation would occur even if the extracellular domain or C-terminal tail of FLA1 were deleted. Secondly, codon-scrambling was used to make these FLA1 truncation mutants resistant to RNAi. While the flagellum was still attached, it was not possible to work out in which membrane FLA1 was anchored. Therefore, RNAi was used to knock down wildtype FLA1 in PCF cells that expressed the RNAi-resistant FLA1 truncation mutants. The flagellum detached and the location of the YFP-tagged truncation mutants was determined.

The FLA1 mutant that lacked the extracellular domain located exclusively to the cell-body side, in a continuous line overlapping with the FAZ filament (Sun et al., 2013). This suggested strongly that the protein was anchored in the cell-body membrane, not the flagellar membrane. In contrast, the FLA1 mutant lacking the 16-a.a. cytoplasmic tail was observed mainly as punctate dots along the detached flagellum. It was hypothesised that when this truncation mutant inserted into the growing FAZ, its extracellular domain formed an attachment with a protein on the flagellar membrane. Without a C-terminal tail, the truncation mutant had little or no anchoring to the FAZ-filament region. When the flagellum detached (as the pool of wildtype FLA1 was depleted), the truncation mutant was therefore left on the flagellar membrane, held in place by the putative FLA1-binding protein.

In short, the data suggested that in procyclic cells, FLA1 anchors in the cell-body side of the FAZ and uses its extracellular domain to bind to a protein or proteins on the flagellar membrane, thereby mediating flagellar attachment (Sun et al., 2013). (An important caveat: this conclusion was based on the localisation of overexpressed truncation mutants,

which could behave differently to the wildtype protein. However, subsequent research has strongly supported the conclusions reached by Sun et al. regarding localisation of FLA1.)

1.3.6 Characterisation of FLA1BP (FLA1-Binding Protein)

To identify the flagellar-membrane protein or proteins to which FLA1 binds, a co-immunoprecipitation approach was taken (Sun et al., 2013). Procyclic cells expressing a YFP-tagged FLA1 mutant with no cytoplasmic tail were biotinylated so that all surface proteins were labelled. They were lysed by sonication and detergent before pulldown of the FLA1 mutant with anti-GFP. The immunoprecipitated proteins were then subjected to affinity purification with streptavidin, to isolate only the surface proteins among them. Silver-staining revealed two bands: the FLA1 mutant itself and an 80-kDa band that could be detected by anti-biotin, but not by anti-GFP. This was evidently a protein that bound to FLA1 and was pulled down with it. Mass spectrometry revealed it to be none other than the protein encoded by Tb927.8.4050/4100 (i.e. the protein described in Section 1.3.4). It was given the name FLA1BP (FLA1-binding protein).

Reverse co-immunoprecipitation, wherein FLA1 was pulled down alongside YFP-tagged FLA1BP, confirmed the interaction between the two proteins (Sun et al., 2013) Curiously, YFP-FLA1BP appears on blots as two bands – one of ~110–120 kDa and one of ~160–170 kDa (inclusive of the 27-kDa tag). It seems probable that the larger FLA1BP band was also present on the original Co-IP blot, but was simply obscured by the YFP-FLA1 band. The larger band could represent a post-translationally modified form of FLA1BP.

To determine the localisation of FLA1BP within the FAZ, YFP-tagged FLA1BP was stably overexpressed in FLA1-RNAi PCF cells, which had detached flagella (Sun et al., 2013). Immunofluorescence showed that YFP-FLA1BP localised exclusively along the detached flagellum, in a punctate pattern. Removal of the extracellular domain of YFP-FLA1BP did not alter the localisation pattern, nor did knockdown of a PFR component. These findings suggested that FLA1BP is anchored in the flagellar membrane by its C-terminal tail, at least in FLA1-RNAi cells. It is probable that the protein locates in the same manner under wildtype conditions, but this cannot be stated with absolute certainty because it remains possible that ablation of FLA1 causes mislocalisation of FLA1BP.

To verify that FLA1BP plays a role in flagellar attachment, the protein was knocked down in procyclic cells (Sun et al., 2013). There was no effect on growth rate, even seven days after induction of *FLA1BP* dsRNA (nucleotides 1385–1902). However, flagellar detachment was observed in >90% of the cells two days after induction. The flagellum was the same length as in control cells, but elongation of the new FAZ was uncoupled from elongation of the new flagellum. The FAZ was 50% of the average control FAZ length, which is 12 µm. As a result, the cells were much shorter than usual, particularly in the anterior region. A 6-µm FAZ was adequate for cell survival *in vitro*. It was not clear if this partial FAZ assembly was due to incomplete depletion of FLA1BP or to an intrinsic length regulation of the other FAZ components. FAZ1 and FLA1 localised correctly to the cell-body side of the FAZ.

To confirm the importance of flagellar attachment for normal FAZ construction, FLA1BP-knockdown cells were treated with lithium chloride. It is known that lithium chloride lengthens flagella and cilia in various cell types, including mammalian cells. Treatment with lithium chloride increased flagellum length in FLA1BP-RNAi and control cells, but a corresponding FAZ lengthening was observed only in control cells, demonstrating that flagellar adhesion is necessary for transmission of flagellum-length information to the intracellular FAZ (Sun et al., 2013).

Interestingly, the kinetoplast was misplaced upon knockdown of FLA1BP (Sun et al., 2013). It was found to be slightly anterior to the nucleus, apparently because of reduced basal-body separation. This phenotype is reminiscent, although not identical, to the epimastigote-like phenotype observed when intraflagellar FAZ proteins FLAM3 and ClpGM6 are knocked down in procyclic cells (Sunter et al., 2015a). By altering the arrangement of its organelles, the PCF can adapt to truncation of the FAZ filament and avoid catastrophic cytokinesis errors. Bloodstream-form cells seem to be incapable of altering their morphology in this drastic way, and are therefore more sensitive to any interference with the FAZ (Sunter et al., 2015a).

1.3.7 Further insights into FLA1 and FLA1BP

The findings described so far seem to indicate that in procyclic *T. brucei*, adhesion of the flagellar membrane to the cell-body membrane is mediated, at least in part, by the binding of FLA1 and FLA1BP. The model suggests that FLA1 locates to the cell-body membrane, with its C-terminal tail inside the cell and its extracellular domain in the

intermembrane space, whereas FLA1BP locates to the flagellar membrane, with its C-terminal tail in the flagellum and its extracellular domain in the intermembrane space (Fig. 1.10 A).

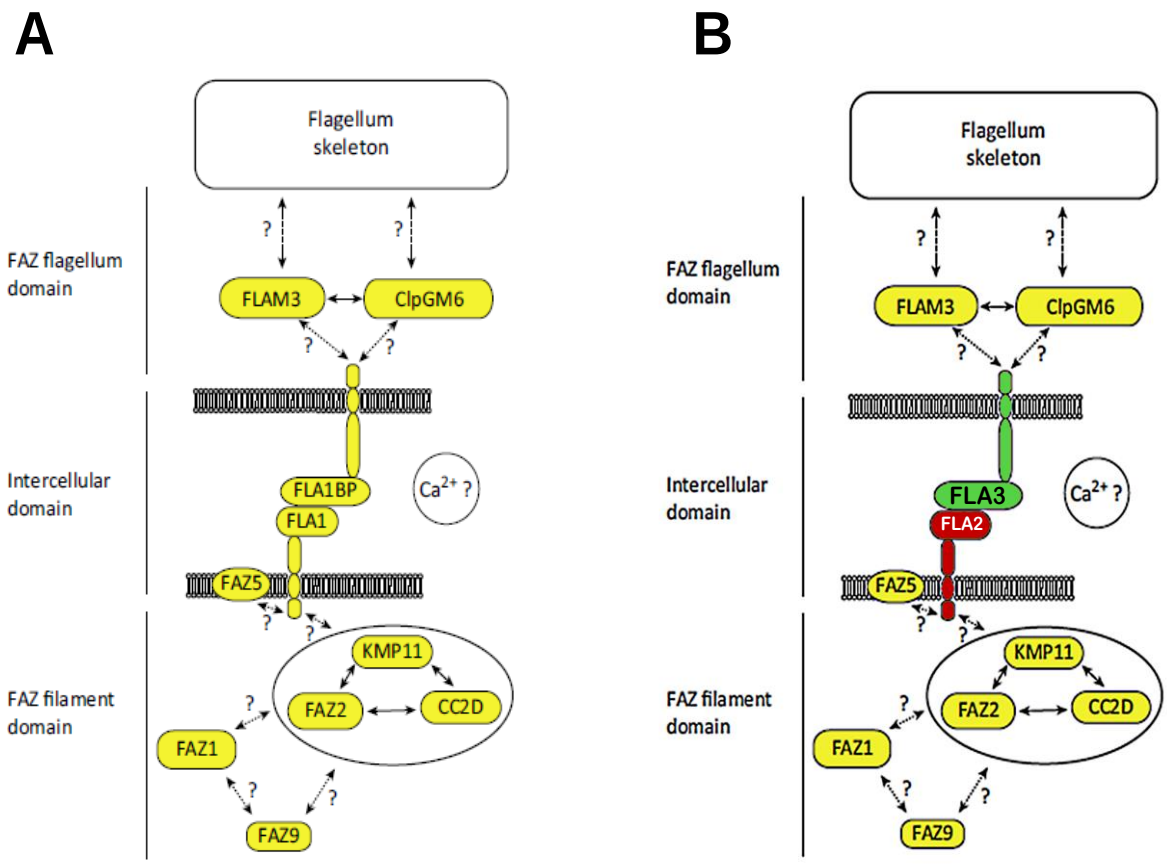


Figure 1.10: Speculative models of the flagellar attachment zone in procyclic and bloodstream-form cells.

(A) Model (from Sunter et al., 2016) describing the current understanding of the layout of the procyclic-form FAZ, with emphasis on the interaction between FLA1 and FLA1BP in the intermembrane zone.

(B) Putative model (adapted from Sunter et al., 2016) of the bloodstream-form FAZ, showing the hypothesised interaction between FLA2 and FLA3.

All four flagellar-adhesion glycoproteins have NHL repeats in their extracellular domains (Sun et al., 2013, Sunter and Gull, 2016). The NHL repeat is a conserved structural motif found in a wide variety of eukaryotic and prokaryotic proteins (Slack and Ruvkun, 1998). It forms a "six-bladed beta-propeller" shape, and it functions by binding to other proteins or, in some cases, to RNA (Good et al., 2004, Kumari et al., 2018). It is postulated that FLA1 and FLA1BP might bind to each other using their NHL repeats (Sun et al., 2013). Alternatively, the proteins could interact via their glycosyl side chains. (Note that glycosylation of FLA1BP has not been explicitly demonstrated. However, the fact that it sometimes appears on blots as a band ~50 kDa higher than its predicted molecular weight suggests that it could be extensively glycosylated – if this is not an artefact of overexpression.)

If the punctate nature of the FLA1BP signal (Sun et al., 2013) is representative of the distribution of the wildtype protein, it is probable that FLA1BP and FLA1 form a series of junctional complexes along the FAZ, rather than a continuous join. The flagellar-adhesion glycoproteins may be important components of the "intermembrane staples" that can be seen to bridge the FAZ when viewed by electron microscopy.

Further research confirmed that the C-terminal tail of FLA1 is anchored to the FAZ filament (Rotureau et al., 2014). Detergent-extracted cytoskeletons were probed by immunogold labelling and viewed by EM; FLA1 could be seen attached to the FAZ filament. The binding partner(s) of FLA1 within the filament were not determined. Notably, alteration of the detergent-extraction conditions occasionally showed a punctate FLA1 signal along the flagellum skeleton instead, supporting the idea that FLA1 binds to the flagellum-associated FLA1BP.

FLA1 and FLA1BP each localise correctly when the other is knocked down (Sun et al., 2013), which implies that the connection between them is established after they are targeted to their respective locations (Rotureau et al., 2014). Interestingly, however, FLA1BP mislocalises to the very base of the flagellum when FLAM3, the intraflagellar protein, is knocked down. This suggests that FLA1BP is anchored, either directly or indirectly, to FLAM3, via its C-terminal tail. FLAM3, a very large protein (468 kDa; 4151 a.a.), is proposed to be component of the fibres that link the FAZ to the axoneme/PFR. It is

thought to act as a counterpart to the FAZ filament, but on the flagellum side of the FAZ (Rotureau et al., 2014).

As the new FAZ is being constructed, both FLA1 and FLA1BP add on to the proximal end only, as is the case for all other FAZ proteins (Sunter and Gull, 2016, Sunter et al., 2015b). Post-assembly addition of proteins, though, is thought to occur along the entire length of the FAZ – probably as a form of maintenance. FLA1BP has been observed to add onto the distal end of the mature FAZ in such a manner (Sunter et al., 2015b).

The interaction between FLA1BP and FLA1 in procyclic cells may drive elongation of the new FAZ as the new flagellum is assembled (Rotureau et al., 2014). This intermembrane connection would thus provide the necessary positional information to align the flagellum and FAZ filament, as well as providing the means of attachment.

1.4 High-Throughput Studies and the FLA Glycoproteins

1.4.1 FLA2 and FLA3 may act as the bloodstream-form counterparts of FLA1 and FLA1BP, respectively.

Table 1.3 summarises what is currently known about the four flagellar-adhesion glycoproteins.

All four FLA proteins have orthologues in other kinetoplastids, such as *Leishmania major*, *Trypanosoma cruzi* and *Bodo saltans*, but not in any other taxa (Sunter et al., 2015b). This is the case for all FAZ proteins. Unlike the other FAZ proteins, however, the FLA proteins are developmentally-regulated (Sunter and Gull, 2016). FLA1 and FLA1BP are expressed in PCFs (see Section 1.4.4); FLA2 and FLA3 are expressed in BSFs. This is interesting because developmental regulation usually indicates that a protein is important, and it can provide clues as to function. The reason for the life-cycle stage-specificity of the FLA proteins is probably their need to integrate with different surface coats (Sunter and Gull, 2016). Procyclic cells are coated with procyclin; bloodstream-form cells with VSG. Being predominantly extracellular, the FLA proteins must be adjacent to these surface-coat proteins. It is not known how the FLA proteins can, in BSFs, exist at the cell surface without being immunogenic. It is possible that their extensive glycosylation plays a role.

A

Protein	Gene ID	No. of Genes	M_r (kDa)	Length (aa)	BSF/PCF	Signal Sequence	Domains	TM domains	Location within FAZ
FLA1	Tb927.8.4010	1	59	546	PCF*	Yes	NHL	1	Cell body side
FLA2	Tb927.8.4060; Tb927.8.4110	2 (98% identity)	64	590	BSF	Yes	NHL	1	???
FLA3	Tb927.5.4570; Tb927.5.4580	2 (93% identity)	89	818	BSF	No (?)	NHL	2 (?)	???
FLA1BP	Tb927.8.4050; Tb927.8.4100	2 (100% identity)	83	750	PCF	Yes (?)	NHL	1	Flagellum side

B

% Identity	*FLA1*	*FLA2*	*FLA3*	*FLA1BP*
FLA1	100	71	–	–
FLA2	71	100	–	–
FLA3	–	–	100	61
FLA1BP	–	–	61	100

Table 1.3: The flagellar-adhesion glycoproteins: a summary

(A): A summary of the published information about the FLA proteins. The molecular weights given are those prior to glycosylation. It is not clear from the sequences of *FLA3* and *FLA1BP* whether there is a signal sequence or a transmembrane (TM) domain at the N-terminus. *See Section 1.4.4.

(B): A table showing the relatedness between the *FLA1* and *FLA2* genes, and between the *FLA3* and *FLA1BP* genes.

Most recent research into the flagellar-attachment zone has been done in procyclic cells. Hence, more is known about FLA1 and FLA1BP than about FLA2 and FLA3. However, since FLA2 is homologous to FLA1 and FLA3 is homologous to FLA1BP, it seems very possible that FLA2 and FLA3 form an analogous intermembrane junction to that formed by FLA1 and FLA1BP (Sun et al., 2013). This would mean that FLA2 would locate to the cell-body membrane and FLA3 to the flagellum membrane, as depicted in Fig. 1.10 (B). Note that, in contradiction of this model, Woods et al. (2013) suggested that FLA3 is only on the cell-body membrane. However, this suggestion was based solely on the behaviour of the protein in artefactual conditions (accidental detachment of the flagellum during imaging), so it is not conclusive. The precise localisation of FLA2 is not known, and no interaction has yet been shown between FLA2 and FLA3. This research project seeks to investigate these points.

The membrane junction formed by FLA1 and FLA1BP is structurally analogous to a desmosome (Sunter and Gull, 2016, Sun et al., 2013). In multicellular organisms, desmosomes join adjacent cells using transmembrane proteins called cadherins that bind to each other across the intermembrane space. The cadherins are anchored in cytoplasmic plaques, just as FLA1 and FLA1BP appear to be anchored to the FAZ filament and flagellum skeleton, respectively. No proteins with cadherin domains have been found in the *T. brucei* genome. The mechanism of binding of FLA1 and FLA1BP is suspected to involve their NHL repeats. FLA2 and FLA3 also contain NHL repeats, which supports the hypothesis that they bind to each other. Alternatively, the many glycosyl side chains of the FLA proteins may function in attachment through a potential carbohydrate-carbohydrate interaction (Sunter and Gull, 2016).

If FLA2 and FLA3 are indeed the BSF counterparts of FLA1 and FLA1BP, they may be able to substitute for them, and vice versa. It would be interesting to see, for example, if FLA1BP could rescue FLA3-knockdown morphology when introduced into a FLA3 RNAi cell line.

High-throughput studies of the *T. brucei* transcriptome, translatome and proteome lend weight to the model that depicts FLA2 and FLA3 as the bloodstream-form counterparts of FLA1 and FLA1BP. A transcriptomic study shows that FLA2 and FLA3 share a common expression profile during cell differentiation, which suggests a shared role (Queiroz et al.,

2009). Proteomic studies demonstrate that FLA1 is PCF-specific, which clears up the questions over the essentiality of FLA2 in the bloodstream form. The transcriptomic data on the FLA proteins will be reviewed in Section 1.4.3 and the proteomic data in Section 1.4.4; firstly, however, it is necessary to summarise how post-transcriptional regulation is achieved in the trypanosome.

1.4.2 Post-Transcriptional Regulation in *T. brucei*

Trypanosome genes are arranged in polycistronic transcription units (PTUs) that are each transcribed by RNA polymerase II as a single unit (Haile and Papadopoulou, 2007, Clayton and Shapira, 2007). Each PTU can contain more than 100 genes, which are mostly functionally unrelated and which can show different developmental regulation (Vasquez et al., 2014, Jensen et al., 2014). As the long pre-mRNA is transcribed, it is processed into individual mRNAs by *trans*-splicing and polyadenylation. During *trans* splicing, a 39-nt spliced leader RNA (*SLRNA*) bearing a trimethyl cap is added to a splice acceptor site in the 5′ UTR. This process allows simultaneous polyadenylation of the 3′ UTR of the upstream gene (Clayton and Shapira, 2007). Only two trypanosome genes have introns, so *cis*-splicing is almost never required (Siegel et al., 2011). Individual genes lack promoters; this, in conjunction with the co-dependent processing of adjacent mRNAs and the conserved open chromatin structure around transcription start sites, indicates that there is no transcriptional control of protein-coding genes (Vasquez et al., 2014, Haile and Papadopoulou, 2007, Clayton and Shapira, 2007). Gene regulation therefore occurs post-transcriptionally. (The exceptions to this rule are the genes for the surface-coat proteins, procyclin and VSG, which are transcribed by RNA polymerase I from dedicated promoters, with transcriptional regulation by chromatin-mediated silencing (Clayton and Shapira, 2007).)

Although *T. brucei* is unicellular, gene regulation is still necessary because different proteins are required in the various life-cycle stages of the parasite (Haile and Papadopoulou, 2007). Post-transcriptional control can occur at several levels: choice of alternative splice-acceptor sites and polyadenylation sites, nuclear export, mRNA degradation, protein translation and protein stability. Of these, the most important are control of mRNA degradation and control of protein translation (Siegel et al., 2011, Jensen et al., 2014).

The half-lives of different mRNAs vary over two orders of magnitude (Manful et al., 2011). One of the main contributing factors towards the stability of an mRNA is the secondary structure of the 3' UTR (Clayton and Shapira, 2007). There appear to be two pathways of mRNA degradation. In the regulated pathway, destabilising proteins bind to the 3' UTR of certain mRNAs that are not wanted in a particular life-cycle stage, triggering decapping of the mRNA and digestion by the 5'–3' exoribonuclease, XRNA. In the constitutive pathway, stabilising proteins bind to the 3' UTR of mRNAs, forming a complex with cap-binding proteins and poly(A)-binding proteins and allowing multiple rounds of translation before the mRNA is eventually deadenylated, decapped and digested by XRNA and by the 3'–5' exosome (Clayton and Shapira, 2007, Haile and Papadopoulou, 2007).

However, differential control of mRNA degradation is insufficient to account for the extensive differences in protein production in BSFs and PCFs (Jensen et al., 2014, Vasquez et al., 2014). Both the number of developmentally controlled genes, and the magnitude of their differential expression are greater at the protein level than at the transcript level. This is because mRNAs of similar abundance can have very different translational efficiencies (Jensen et al., 2014, Vasquez et al., 2014). It has been postulated that uORFs (small translated ORFs located upstream of the main coding sequence of an mRNA) can control translation efficiency by competing for translation initiation (Vasquez et al., 2014). This mechanism occurs in other eukaryotes, but it seems to be relevant for at most a tenth of trypanosome genes (Jensen et al., 2014, Clayton, 2019). It is also hypothesised that differences in translational efficiency between life-cycle stages may be due to alternative splicing of pre-mRNAs in the nucleus (Siegel et al., 2011).

Although the mechanisms of translational control in *T. brucei* have not been fully elucidated, it is certain that such control exists and is regulated separately to mRNA stability (Vasquez et al., 2014). It is probable that the fate of an individual mRNA depends on the interaction of multiple regulatory RNA-binding proteins attached to the 3′ UTR (Clayton, 2019). For example, FLA2 has been found to belong to a cohort of developmentally-regulated proteins whose translation in the procyclic form is repressed by the DEAD-box RNA helicase, DHH1 (Kramer et al., 2010).

1.4.3 Transcriptomics, translatomics and the FLA Glycoproteins

Several high-throughput transcriptomics studies have investigated the relative abundance of certain mRNAs in the procyclic form relative to the bloodstream form. All transcriptomic studies that give data on *FLA* gene expression agree that *FLA2* and *FLA3* mRNAs are more abundant in BSFs, whereas *FLA1* and *FLA1BP* mRNAs are more abundant in PCFs (Table 1.4).

Protein	Gene	mRNA is more abundant	PCF:BSF ratio if specified	Measurement technique	Source
FLA1	4010	PCF	–	Microarray	(Brems *et al.*, 2005)
	''	PCF	2.0	Microarray	(Koumandou *et al.*, 2008)
	''	PCF	3.12	Microarray	(Queiroz *et al.*, 2009)
	''	PCF	4.6	RP	(Vasquez *et al.*, 2014)
FLA2	4060	BSF	0.33	Microarray	(Koumandou *et al.*, 2008)
	''	BSF	0.14	Microarray	(Queiroz *et al.*, 2009)
	''	BSF	–	SLT	(Nilsson *et al.*, 2010)
	''	BSF	0.17	RP	(Vasquez *et al.*, 2014)
	4110	BSF	0.26	Microarray	(Queiroz *et al.*, 2009)
	''	BSF	0.19	RP	(Vasquez *et al.*, 2014)
FLA3	4570	BSF	0.26	Microarray	(Queiroz *et al.*, 2009)
	''	BSF	0.22	RP	(Vasquez *et al.*, 2014)
	4580	BSF	0.17	RP	(Vasquez *et al.*, 2014)
FLA1BP	4050/4100	PCF	4.11	Microarray	(Queiroz *et al.*, 2009)
	''	PCF	–	SLT	(Nilsson *et al.*, 2010)
	''	PCF	3.63	RP	(Vasquez *et al.*, 2014)

Table 1.4: High-throughput transcriptomic studies demonstrate the stage-specificity of FLA-glycoprotein mRNA.

SLT = splice leader trapping. RP = ribosome profiling.

Furthermore, Fadda et al. state that *FLA2* and *FLA3* mRNAs are more stable in BSFs, whereas *FLA1* and *FLA1BP* mRNAs are more stable in PCFs (Fadda et al., 2014).

However, as indicated in Section 1.4.2, the lack of transcriptional control in *T. brucei* means that mRNA abundance is only a weak indicator of the extent of protein production. Transcriptomic data is therefore insufficient to resolve the uncertainty about whether the FLA1 protein is expressed in bloodstream-form cells.

In the case of developmentally-regulated proteins, the difference in protein production between life-cycle stages is normally of greater magnitude than the corresponding difference in gene expression (Jensen et al., 2014). This discrepancy is due mainly to the difference in translational efficiency of a given transcript when expressed in different life-cycle stages. Two high-throughput translatomics studies have measured the relative translational efficiencies (TE) (i.e. the relative protein production from a single mRNA) of the *FLA* transcripts (Table 1.5) (Vasquez et al., 2014, Jensen et al., 2014) These studies used ribosome profiling (high-throughput sequencing of ribosome-protected mRNA fragments) to determine the ribosome footprint density of each transcript; the ribosome footprint was divided by the transcript abundance to control for differences in mRNA levels.

The translatomic data in Table 1.5 suggests that each *FLA1* transcript in the procyclic form produces slightly more protein than the same transcript in the bloodstream form. This could indicate a moderate level of transcriptional regulation of FLA1 between life-cycle stages. However, Jensen et al, do not consider the FLA1 TE ratios to be statistically significant. Vasquez et al. warn that direct comparisons of TE values are unreliable anyway; instead, they rank all the proteins in order of translational efficiency and record each protein's change in rank between life-cycle forms. Between PCF and BSF stages, FLA1 drops in rank by 1400 places, which is considered evidence of differential translational control during development.

To answer conclusively the question of whether the FLA1 protein is produced in BSFs, transcriptomics and translatomics are inadequate. Proteomic data has proved much more informative in this regard (reviewed in Section 1.4.4).

Nonetheless, transcriptomic studies performed in trypanosomes during differentiation from BSF to PCF have generated other highly interesting information about the regulation of the FLA glycoproteins, as detailed below.

Protein	Gene	TE is higher	PCF:BSF TE ratio (monomorphic) (Vasquez *et al.*, 2014)	PCF:BSF TE ratio (monomorphic) (Jensen *et al.*, 2014)	PCF:BSF TE ratio (pleomorphic) (Jensen *et al.*, 2014)
FLA1	4010	PCF	2.14	1.37	1.63
FLA2	4060	BSF	0.44	0.08	0.15
	4110	BSF	0.75	0.17	0.31
FLA3	4570	BSF	0.15	0.08	0.15
	4580	BSF	0.22	0.06	0.10
FLA1BP	4050/4100	PCF	6.24	4.43	4.23

Table 1.5: High-throughput translatomic studies indicate the change in translational efficiencies (TE) of FLA-glycoprotein mRNA between life-cycle stages.

Translation efficiency is a measure of the number of times a given mRNA is translated to protein.

Most studies of *T. brucei* are performed in the procyclic and slender bloodstream forms. Little is known about flagellar attachment in the other clinically relevant life-cycle stages, namely the metacyclic form and the stumpy bloodstream form. As described in Section 1.1, the metacyclic form is the form of the trypanosome that first enters the mammalian host via a tsetse bite, and the stumpy form is the form that is predisposed to be taken up by a new tsetse fly and which can differentiate rapidly into the procyclic form. Metacyclic and stumpy cells are both trypomastigotes, and as such, they must have flagellar-attachment zones. However, they are also non-proliferative, which makes them very difficult to culture. Consequently, it is not known whether the FAZ of these cells bears more resemblance to that of the BSF or PCF life-cycle stages. Two high-throughput studies have produced data relating to the expression and function of the *FLA* genes during differentiation of the BSF to the PCF (Alsford et al., 2011, Queiroz et al., 2009). Differentiating cells in culture offer only limited insight into the nature of the true stumpy form, which exists only *in vivo*. Nonetheless, these studies are noteworthy because they reveal striking similarities between the expression profiles of *FLA2* and *FLA3*.

One of these studies (Alsford et al., 2011) used a technique called RIT-Seq to knock down ~7,500 trypanosome genes during the differentiation of BSF cells to the procyclic form. Each gene knockdown was examined to see whether it caused loss of function at specific timepoints of the differentiation process. In this way, it could be seen which genes were essential in BSF, PCF or differentiating cells, respectively. Several *FLA* genes were investigated in this study. The results are summarised in the following table (Table 1.6), where "X" indicates that loss of function occurred upon knockdown of the corresponding gene.

The results obtained as to the essentiality of the *FLA* genes in BSF and PCF cells are as expected. *FLA2* and *FLA3* are required in BSFs, but not PCFs; *FLA1BP* has no significant effect on growth in either life-cycle stage. The fact that knockdown of *FLA2* (Tb927.8.4060) caused loss-of-function in PCFs is probably due to cross-reaction of the FLA2 RNAi vector with *FLA1*. (The RNAi library was generated from randomly-sheared DNA rather than being designed to ensure specificity to individual genes.) More interestingly, both *FLA2* and *FLA3* are shown to be essential in cells that are differentiating from BSF to PCF. This suggests that, during the differentiation process, the FAZ remains similar to the BSF FAZ.

Gene	Gene ID	BSF (Day 3)	BSF (Day 6)	Differentiating	PCF
FLA2	Tb927.8.4060	**X**	**X**	**X**	**X**
	Tb927.8.4110	**X**	**X**	**X**	-
FLA3	Tb927.5.4570	**X**	**X**	**X**	-
	Tb927.5.4580	-	**X**	**X**	-
FLA1BP	Tb927.8.4050	-	-	-	-

Table 1.6: RIT-Seq analysis indicates that FLA2 and FLA3 are essential in monomorphic cells during differentiation from BSF to PCF (data from Alsford et al., 2011).

"X" indicates loss of function upon conditional knockdown of the corresponding gene. Timepoints refer to days after initiation of RNAi by tetracycline.

However, that this study was performed in a monomorphic cell line (most bloodstream-form cultures have, because of long-term passaging *in vitro*, lost the ability to differentiate into the stumpy form, and so are termed monomorphic). Therefore, the differentiating cells tested in this study can give only very limited insight into the probable composition of the stumpy-form FAZ.

To produce more useful data, it would be necessary to observe the expression of the *FLA* genes in pleomorphic cells. Such cells retain the ability to adopt stumpy-like morphology, and so their differentiation from BSF to PCF in culture should reflect the *in vivo* differentiation process more accurately. Queiroz et al. performed a high-throughput study in which the expression of trypanosome genes was measured by microarray at different timepoints while a pleomorphic BSF cell line was induced to differentiate into the procyclic form by addition of *cis*-aconitate. The *FLA* genes were included in this investigation, and the results were striking (Fig. 1.11). The data were normalised to enable comparison of the transcriptomic profiles of different genes. (The mRNA level of each gene at a given timepoint was expressed as a ratio to the mRNA level of the same gene in the procyclic form.) Pairs of genes that underwent a similar regulation pattern during the differentiation process could therefore be visualised easily, even if their absolute mRNA levels were different.

The procyclic-form *FLA* genes (*FLA1* and *FLA1BP*), for example, showed remarkably similar transcriptomic profiles (Fig. 1.11 A). This makes intuitive sense, since the proteins encoded by these genes work in conjunction with each other. The *FLA1* and *FLA1BP* genes are located near to each other on chromosome 8 and are therefore probably transcribed in the same PTU. However, the closeness of the transcriptomic profiles of these genes means that their post-transcriptional regulations must also be very similar. This supports the main conclusion of the study, namely that "trypanosome genes form post-transcriptional regulons in which mRNAs with functions in particular pathways, or encoding components of protein complexes, show almost identical patterns of regulation" (Queiroz et al., 2009). Note that at the time of this study, the interaction between FLA1 and FLA1BP was not known, and this conclusion was based on the observation of other sets of genes.

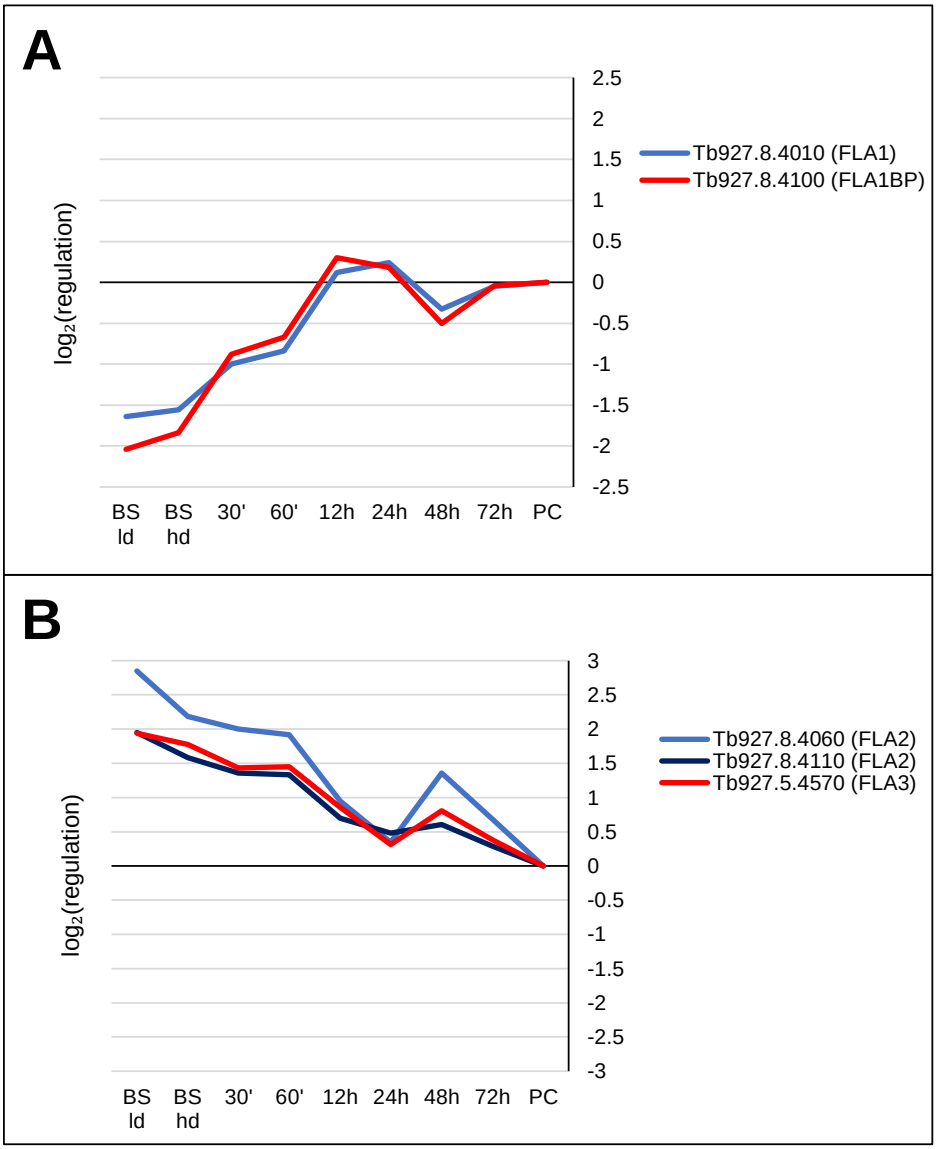


Fig. 1.11: Comparison of the transcriptomic profiles of the *FLA* genes during differentiation (adapted from Queiroz *et al*., 2009, Table S2). The mRNA levels of several *FLA* genes were determined by microarray while BSF→PCF differentiation was induced in pleomorphic cells. The $\log_2$ regulation ratio (sample/procyclic) is plotted on the Y axis for the different time points given on the X axis. [BS ld = low-density BSF; BS hd = high-density BSF; PC = established PCF]

(A) The transcriptomic profiles of the BSF *FLA* genes, *FLA2* and *FLA3*, are remarkably similar. Both show a decrease during BSF→PCF differentiation, with a temporary spike at 48 h. **(B)** The transcriptomic profiles of the PCF *FLA* genes, *FLA1* and *FLA1BP*, are remarkably similar. Both show an increase during BSF→PCF differentiation, with a temporary decrease at 48 h.

For the purposes of the current investigation, however, the most interesting aspect of Fig. 1.11 is the comparison of the transcriptomics profiles of *FLA2* and *FLA3*, the BSF-specific *FLA* genes. It is apparent that the regulation of expression of these two genes is remarkably similar during differentiation from BSF to PCF. The extreme closeness of these two expression profiles is extraordinary because *FLA2* and *FLA3* are located on different chromosomes. They evidently share features (possibly in the 3′ UTR) that cause them to undergo the same post-transcriptional regulation. This is compelling evidence that FLA2 and FLA3 work in close conjunction with one another. Direct interaction may even occur between the two proteins.

1.4.4 Proteomics and the FLA Glycoproteins

Three major proteomic studies have used stable-isotope labelling of amino acids in cell culture (SILAC) to compare quantitatively the respective proteomes of the procyclic form and bloodstream form (Gunasekera et al., 2012, Urbaniak et al., 2012, Butter et al., 2013). Table 1.7 summarises the data that pertain to the FLA glycoproteins.

As expected, the proteomic data confirm that FLA2 and FLA3 are bloodstream-stage-specific, while FLA1BP is procyclic-stage-specific. (Curiously, one of the *FLA3* genes (Tb927.5.4570) is consistently reported to have a significantly lower BSF:PCF ratio than the other (Tb927.5.4580).)

Critically, Urbaniak et al. and Butter et al. both measured almost twelve times more FLA1 protein in PCFs than in BSFs. This is conclusive evidence that FLA1 is not expressed at any meaningful level in the bloodstream-form trypanosome, which has important implications for FLA2. If the proteomics data are correct, the original evidence for FLA1 in the bloodstream form (notably the western blot and the immunofluorescence images (Nozaki et al., 1996)) must in fact represent FLA2. As hypothesised in Section 1.3.2, the 73% similarity between the two homologues caused the FLA1 antibody to cross-react with FLA2. It is therefore FLA2, not FLA1, that performs an essential role in flagellar attachment in BSFs.

Protein	Gene	Protein is more abundant	PCF:BSF ratio	BSF:PCF ratio	Source
FLA1	4010	PCF	11.77	0.08	(Urbaniak *et al.*, 2012)
	''	PCF	11.95	0.05	(Butter *et al.*, 2013)
FLA2	4060/4110	BSF	0.07	14.79	(Gunasekera *et al.*, 2012)
	''	BSF	0.06	16.3	(Urbaniak *et al.*, 2012)
	''	BSF	0.06	19.29	(Butter *et al.*, 2013)
FLA3	4570/4580	BSF	0.09	11.13	(Gunasekera *et al.*, 2012)
	4570	BSF	0.20	4.94	(Urbaniak *et al.*, 2012)
	''	BSF	0.61	1.63	(Butter *et al.*, 2013)
	4580	BSF	0.08	13.09	(Urbaniak *et al.*, 2012)
	''	BSF	0.07	6.01	(Butter *et al.*, 2013)
FLA1BP	4050/4100	PCF	9.70	0.10	(Urbaniak *et al.*, 2012)
	''	PCF	3.88	0.08	(Butter *et al.*, 2013)

Table 1.7: High-throughput proteomic studies demonstrate the stage-specificity of the FLA glycoproteins.

For ease of reference, both PCF:BSF and BSF:PCF ratios are given. Urbaniak et al. measured the PCF:BSF ratio only; the BSF:PCF ratio given here is the inverse of the PCF:BSF value.

Two other high-throughput proteomic studies support the finding that FLA1 is PCF-specific. Shimogawa et al. performed surface labelling followed by shotgun proteomic analyses to describe cell-surface proteomes of BSF and PCF cells (Shimogawa et al., 2015). FLA1 was not present in BSF samples but was found in two of five independent PCF samples. (Only proteins that were present in at least four of five samples were assigned to the high-confidence cell-surface proteome. However, it seems probable that the reason FLA1 was not found in all PCF samples was simply a surface-accessibility issue that could be overcome in a low-throughput investigation.)

Dejung et al. used label-free quantitative mass spectrometry to quantify proteins during differentiation of stumpy pleomorphic cells from BSF to PCF (Dejung et al., 2016). Proteins were grouped into "expression profile classes" if they shared similar patterns of regulation throughout the differentiation process. FLA1 and FLA1BP were grouped in the same profile class (Profile 10) because, along with 70 other proteins, they were significantly upregulated from 4 h after initiation of differentiation onwards. Meanwhile, FLA2 and FLA3 were grouped in the same profile class (Profile 6), along with 107 other proteins, because they were significantly downregulated from 48 h after initiation of differentiation onwards.

As well as confirming the PCF-specificity of FLA1, Dejung's data give an interesting insight into the FAZ composition of the stumpy form. In the case of all four FLA glycoproteins, there was no significant difference in protein level between the stumpy and slender BSF forms (Dejung et al., 2016). In pleomorphic cultured cells, at least, the FAZ therefore appears to resemble that of the bloodstream-form cell. Upon initiation of differentiation, rapid upregulation of FLA1 and FLA1BP was observed, but FLA2/FLA3 levels remained reasonably constant for the first 48 h, as stated above. It seems probable that upon differentiation, the cell produces FLA1 and FLA1BP for insertion into the daughter FAZ, while the stability of the FLA2/FLA3 proteins in the old FAZ means that the latter proteins remain detectable for some time after they have stopped being produced.

Another high-throughput proteomic study produced an interesting result pertaining to FLA3. Oberholzer et al. found FLA3 to be present in the flagellar surface proteome (Oberholzer et al., 2011). The surface localisation of the protein was unsurprising, given that FLA3 is predicted to be largely extracellular. (Furthermore, a more recent high-throughput study has confirmed surface localisation for all four FLA glycoproteins

(Shimogawa et al., 2015).) However, the apparent flagellar localisation of FLA3 is interesting. Oberholzer et al. generated their flagellar proteome by performing FLA1 RNAi to detach the flagellum from the cell body in BSF cells (the RNAi vector used would have knocked down FLA2 also). The detached flagella were sheared mechanically from the cell body, then isolated by sucrose-density sedimentation. It would seem, therefore, that for FLA3 to remain in the flagellar fraction, it must be anchored on the flagellar side of the FAZ – contrary to the assertation of Woods et al. (2013) that FLA3 was anchored on the cell-body side of the FAZ.

However, Oberholzer's result is by no means conclusive as to the localisation of FLA3. Firstly, Oberholzer et al. did not investigate whether FLA3 was also present in the proteome of the cell-body fraction. Secondly, a similar study in procyclic cells was inconclusive about the localisation of FLA1BP within the FAZ (Subota et al., 2014), so this crude high-throughput method may not be an accurate way to analyse the localisation of the proteins of the intermembrane region. Subota et al. performed FLA1 RNAi in procyclic cells, followed by mechanical shearing to separate the flagellum and cell body. Mass spectrometry detected FLA1BP in both the flagellar and cell-body fractions, with no significant enrichment in either (Subota et al., 2014), even though FLA1BP is known to be anchored only on the flagellar side of the FAZ (Sun et al., 2013).

It is likely that the harsh mechanical shearing performed by Oberholzer et al. and Subota et al. broke the FAZ apart in an inconsistent manner. A gentler method, optimised specifically for FLA3, is required to ascertain where the protein is localised within the flagellar attachment zone. The same is true for FLA2.

1.5 Outlook and Aims

1.5.1 Outlook for investigation of the bloodstream-form FLA glycoproteins

Many questions remain about flagellar attachment in the bloodstream form of *T. brucei*. Firstly, it is not clear whether FLA2 and FLA3 are components of the junctional complexes of the FAZ. These junctional complexes traverse the intermembrane space and are presumed to be the structures that mediate physical attachment of the flagellum to the cell body. However, the FLA glycoproteins may locate in a continuous line along the two

membranes of the FAZ, perhaps performing a complementary role to the junctional complexes. This matter can be resolved by high-resolution localisation studies.

To investigate the precise spatial organisation of FLA2 and FLA3 and the control of their expression in different life-cycle stages, new cell lines will need to be generated. In particular, since no suitable antibodies are available, the two proteins will need to be epitope-tagged and expressed stably in trypanosome cells.

Tagging FLA2 and FLA3 will facilitate interaction studies. In their roles as adhesion proteins, it is reasonable to assume that FLA2 and FLA3 bind to other FAZ components and, potentially, to each other. A co-immunoprecipitation approach will be taken here.

The original investigation of FLA3 showed that the protein locates to the intermembrane region of the FAZ, but it remains uncertain whether FLA3 belongs to the cell-body membrane, as suggested by Woods et al. (2013), or to the flagellar membrane, as suggested by the localisation of the homologous protein, FLA1BP. One of the main aims of this thesis is to resolve this issue by localising FLA3 in trypanosomes in which the flagellum is detached from the cell body.

Of all the FLA glycoproteins, FLA2 is the least characterised. The close sequence similarity between FLA2 and FLA1 has meant that no specific investigations of the FLA2 protein have been performed. The product of the *FLA2* gene has never been identified or characterised at the protein level. It is not clear what role, if any, is performed by the protein during the life cycle of the trypanosome. In this thesis, a FLA2-specific RNAi vector will be generated to examine the function and essentiality of the protein.

The sequence similarity between *FLA2* and *FLA1* will be a problem when attempting to add an epitope tag specifically to *FLA2*. This issue can be overcome by means of long-primer PCR tagging, a technique that exploits differences in the respective 3' UTRs of the two genes.

In summary, it is hoped to develop a clearer picture of the role of FLA3 in the bloodstream-form FAZ, and to investigate whether FLA2 likewise has a role in flagellar attachment.

1.5.2 Aims

FLA3:

- To perform RNAi of *FLA3* and verify that the effects on phenotype and proliferation match those described in the literature
- To establish a BSF trypanosome cell line in which FLA3 is epitope-tagged
- To determine the localisation of tagged FLA3 by fractionation experiments and confocal microscopy
- To discover the precise localisation of FLA3 using confocal microscopy in a cell line in which FLA3 is tagged and the flagellum is detached
- To achieve super-resolution localisation of FLA3 by stimulated emission depletion (STED) microscopy
- To establish a PCF trypanosome cell line in which FLA3 is epitope-tagged and to determine the localisation of tagged FLA3 by confocal microscopy
- To investigate the putative interaction partners of FLA3 by immunoprecipitation of tagged FLA3 and subsequent probing by silver-staining or lectin
- To determine whether tagged FLA3 is located at the cell surface by surface-probing with NHS-biotin
- To investigate the glycan side chains of FLA3 by lectin pulldown
- To express two different tagged versions of FLA3 in the same cells and to test for interaction between the two forms

FLA2:

- To generate a conditional-RNAi construct that will selectively knock down *FLA2*
- To perform RNAi of *FLA2* in bloodstream-form cells and to determine the effects on proliferation and morphology
- To perform RNAi of *FLA2* in procyclic-form cells and to determine the effects on proliferation and morphology
- To establish a BSF trypanosome cell line in which FLA2 is tagged by long-primer PCR tagging
- To determine the localisation of tagged FLA2 by fractionation experiments and confocal microscopy
- To estimate the molecular weight of mature wildtype FLA2 based on the migration of tagged FLA2
- To investigate the putative glycan side chains of FLA2 by lectin pulldown

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Buffers/Solutions

AP Buffer: 100 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$, dH_2O, pH 9.5

Blocking Buffer A: 5% w/v BSA, 0.15% v/v Tween-20, PBS

Blocking Buffer B: 2% w/v FBS, 0.05% v/v Tween-20, PBS

Blocking Buffer C: 2% w/v BSA, 0.05% v/v Tween-20, PBS

Cytomix Transfection Buffer: 2 mM EGTA (pH 7.6), 120 mM KCl, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4/KH_2PO_4 (pH 7.6), 25 mM HEPES (pH 7.6), 5 mM $MgCl_2.6H_2O$, 0.5% w/v glucose, 100 µg ml^{-1} BSA, 1 mM hypoxanthine, dH_2O, pH 7.6, sterile-filtered.

Developing Solution: 1.25 g sodium carbonate, 50 µl 37% formaldehyde, make up to 50 ml with dH_2O

Fixer: 20 ml ethanol, 5 ml acetic acid, 25 ml dH_2O

Incubation Solution: 15 ml ethanol, 3.4 g sodium acetate, 250 µl 25% w/v glutaraldehyde, 0.1 g sodium thiosulfate ($Na_2S_2O_3$) [or 0.1 g $Na_2S_2O_3.5H_2O$], make up to 50 ml with dH_2O

Lysis Buffer: 25 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, dH_2O, pH 7.5

Mounting Solution: 5 ml glycerol, 1 ml 10X PBS, 4 ml dH_2O, 2.5% v/v DABCO, 1 µg ml^{-1} Hoechst

PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, dH_2O, pH 7.4

PBS for Immunoprecipitation:.....3 mM KH_2PO_4, 16 mM Na_2HPO_4, 136 mM NaCl, 3 mM KCl, pH 7.5

6% PFA:.....20 ml dH_2O, 5 ml 10X PBS, 3 g paraformaldehyde, add 5 M NaOH until clear, make up to 50 ml with dH_2O, pH 7.4

Protease-Inhibitor Buffer:.....Sigma P8340 PI cocktail diluted 1 in 200 in TES buffer containing 1 mM EDTA

Protein G Slurry:.....Invert Protein-G-agarose to mix. Wash with 1 ml PBS. Spin at 9000 rcf for 30 sec. Resuspend pellet in PBS as a 1:1 slurry.

PSG:.....3 mM NaH_2PO_4, 57 mM Na_2HPO_4, 44 mM NaCl, 5 mM KCl. Initially, make up to 0.8 of the final volume with dH_2O. Check the pH is 8.0. Immediately prior to use, add 0.1 mM adenosine, 10 mM glucose, 70 mM sucrose. Make up to the final volume with dH_2O.

Running Gel (10%):.....5.9 ml dH_2O, 3.8 ml Tris (1.5 M, pH 8.8), 5.0 ml 30% acrylamide, 150 µl 10% SDS, 6 µl TEMED, 150 µl 10% ammonium persulfate

Running Buffer (10X):.....30 g Tris, 144 g glycine, 10 g SDS, make up to 1 L with dH_2O

2X Sample Buffer:.....20% glycerol, 62.5 mM Tris-HCl (0.5 M pH 6.8), 2.5% SDS, 5% β-mercaptoethanol, 0.01% bromophenol blue, dH_2O

Silver Solution:.....50 mg silver nitrate; 10 µl 37% formaldehyde; 50 ml dH_2O

Stacking Gel (5%):.....2.7 ml dH_2O, 500 µl Tris (1 M, pH 6.8), 670 µl 30% acrylamide, 40 µl 10% SDS, 4 µl TEMED, 40 µl 10% ammonium persulfate

Stop Solution: ..... 0.73 g EDTA; 50 ml dH_2O

TBS: ..... 20 mM Tris, 150 mM NaCl, dH_2O, pH 7.5

TBST (10X): ..... 24 g Tris, 88 g NaCl, 10 ml Tween-20, make up to 1 L with dH_2O, pH 7.6

TES Buffer: ..... 50 mM N-Tris(hydroxymethyl)methyl-2-aminoethane sulfonic acid (TES), 150 mM NaCl, pH 7.5

Transfer Buffer: ..... 25 mM Tris-HCl, 192 mM glycine, 20% v/v methanol, dH_2O, pH 8.3

Tris Buffer: ..... 50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, pH 7.5

Tris-Triton Buffer: ..... 50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, 0.1% w/v Triton X-100, pH 7.5

Tris-Triton-NaCl Buffer: ..... 50 mM Tris-HCl, 500 mM NaCl, 1 mM EDTA, 0.1% w/v Triton X-100, pH 7.5

2.1.2 Antibodies

Primary:

α-HA:.....Rat

α-Ty1:.....BB2 Mouse IgG1 Anti-Ty1 Monoclonal Serum, kind gift from Dr Mark Carrington, University of Cambridge (Bastin et al., 1996)

α-VSG: .....Rabbit

α-actin:.....Rabbit Serum

α-BIP: .....Rabbit (Bangs et al., 1993)

Secondary:

α-mouse IgG AP:.....Promega AP S3721 Goat Anti-Mouse IgG (H+L)

α-mouse IgG 568:.....Jackson ImmunoResearch AffiniPure Cy3-conjugate 115-165-003 Goat Anti-Mouse IgG (H+L)

α-mouse IgG 488:.....ImmunoReagents DyLight 488 IMMRIR1889 Goat Anti-Mouse IgG

α-rabbit AP:.....ThermoFisher Scientific Pierce AP 31346 Goat Anti-Rabbit IgG (H+L)

α-rabbit HRP:.....Jackson ImmunoResearch Peroxidase AffiniPure 111-035-144 Goat Anti-Rabbit IgG (H+L)

α-rabbit 488:.....ThermoFisher Scientific Invitrogen Alexa 488 A-21441 Chicken Anti-Rabbit IgG (H+L)

α-rat AP:.....Goat Anti-Rat

α-rat IgG 568:.....ThermoFisher Scientific Invitrogen Alexa 568 A-11077 Goat Anti-Rat IgG (H+L)

2.1.3 Sources of Materials

Acrylamide..... Fisher Scientific

Adenosine..... Sigma

Agarose..... Fisher Scientific

Amaxa Nucleofector Kit..... Lonza VAMI-1011

Ammonium persulfate..... Fisher Scientific

Ampicillin..... Fisher Scientific

β-mercaptoethanol..... Riedel-de-Haën

BCIP..... Duchefa

Bis-acrylamide..... Melford

Bovine serum albumin..... Sigma

Bromophenol blue..... BDH

Calcium chloride..... BDH

CHAPS..... Duchefa

Cover glass #1.5H 12 mm..... VWR 630-2190

DABCO..... Aldrich

EDTA..... Sigma

EGTA..... Amresco

Expand™ PCR System..... Sigma 4738250001

EZ-Link™ Sulfo-NHS-Biotin..... Thermo Fisher Scientific

Foetal bovine serum..... Sigma

Formaldehyde..... Promega

G418..... Santa Cruz

GeneRuler 1 kb DNA Ladder..... Fisher Scientific

Glucose..... Duchefa

Glutaraldehyde..... Aldrich

Glycerol..... Fisher Scientific

Glycine ..... Duchefa

Hemin ..... Sigma

HEPES ..... Duchefa

HMI-9 culture medium ..... Invitrogen

Hoechst ..... ThermoFisher Scientific

HRP Substrate ..... Immobilon Western Chemiluminescent

Hypoxanthine ..... Sigma

LB agar ..... Sigma

LB broth ..... Fisher Scientific

Magnesium chloride ..... Sigma

Midiprep kit ..... Promega PureYield Plasmid

Milk powder ..... Marvel

Miniprep kit ..... Promega PureYield Plasmid

NBT ..... Fisher Scientific

Nitrocellulose membrane ..... Whatman

***Not*I** ..... ThermoFisher Scientific

Nucleofector ..... Amaxa

p2708 plasmid ..... gift from Dr Mark Carrington, University of Cambridge

p2T7.177 plasmid ..... gift from Dr Bill Wickstead, University of Nottingham

p3074 plasmid ..... gift from Dr Mark Carrington, University of Cambridge

PageRuler Plus Protein Ladder ..... Fisher Scientific 26619

Paraformaldehyde ..... Riedel-de-Haën

Phleomycin ..... Duchefa

Poly-L-lysine ..... Sigma P1524

Ponceau S ..... Sigma

Potassium chloride ..... Duchefa

Potassium dihydrogen phosphate ... Duchefa

di-Potassium hydrogen phosphate.. BDH

Primers ..... Eurofins

ProLong Gold Mountant ..... Fisher Scientific

Protease-inhibitor cocktail..... Sigma P8340

Protein G..... Sigma P3296

pSMOX plasmid ..... gift from Dr Mark Carrington, University of Cambridge

Puromycin..... Santa Cruz

RCA-I lectin beads ..... Vector Laboratories AL-1083

RCA-I lectin biotinylated..... Vector Laboratories B-1085

SDM-79 culture medium..... Invitrogen

SDS ..... Sigma

Silver nitrate..... Sigma

Sodium acetate ..... Sigma

Sodium carbonate ..... Fisher Scientific

Sodium chloride..... Duchefa

Sodium dihydrogen phosphate..... Sigma

Sodium hydrogen carbonate ..... Fluka

di-Sodium hydrogen phosphate ..... Fluka

Sodium thiosulfate..... Duchefa

Streptavidin agarose beads ..... Sigma 85881

Streptavidin AP..... Sigma S2890

Streptomycin sulfate..... Duchefa

Sucrose..... Duchefa

TEMED ..... Sigma

TES..... Melford

Tetracycline..... Duchefa

Tomato lectin biotinylated..... Vector Laboratories B-1175

Transfer apparatus.....Biometra semi-dry

Tris.....Fisher Scientific

Triton X-100.....Sigma

Tween-20.....Duchefa

Urea.....Duchefa

***Xho* I**.....New England Biolabs

2.2 Methods: Cell Culture and Genetic Manipulation

2.2.1 Cell Culture

All bloodstream-form cells cultured were monomorphic *Trypanosoma brucei brucei*, Lister 427 strain, expressing the MITat 1.1 VSG variant, and all procyclic-form cells were derived from these.

Bloodstream-form cells were maintained at $1x10^5$–$2x10^6$ cells ml^{-1} in HMI-9 culture medium with 10% v/v foetal bovine serum, 36 mM $NaHCO_3$, 0.2 mM β-mercaptoethanol, 50 μg ml^{-1} streptomycin and 50 μg ml^{-1} ampicillin (pH 7.5), at 37°C.

Procyclic-form cells were maintained at $1x10^5$–$3x10^6$ cells ml^{-1} in SDM-79 culture medium with 10% v/v foetal bovine serum, 24 mM $NaHCO_3$, 5 μg ml^{-1} hemin, 50 μg ml^{-1} streptomycin and 50 μg ml^{-1} ampicillin (pH 7.5), at 27°C.

2.2.2 Centrifugation

The centrifugation steps of all protocols were carried out at 4°C.

2.2.3 Agarose-Gel Electrophoresis

Agarose gels were made using 1% agarose in TAE buffer. Ethidium bromide was added directly to the gels prior to setting. The DNA ladder used was GeneRuler 1 kb. Gels were viewed under UV light in a gel-documentation system.

2.2.4 Plasmid Amplification

Existing plasmid stocks were diluted 1 in 50 in sterile nuclease-free dH_2O, to achieve concentrations of 2–20 ng μl^{-1}. DH5α *E. coli* bacteria stock (50 μl) was thawed from -80°C and added to 1 μl diluted plasmid. This mixture was incubated on ice for 20 min, then heat-shocked at 42°C for 2 min before incubation on ice for a further 2 min. LB broth (1 ml) was added under sterile conditions. The mixture was incubated at 37°C for 1 h at 350 rpm. It was centrifuged at 10,000 rcf for 2 min. All but 300 μl of supernatant was discarded, and the pellet was resuspended. 100 μl was lawned onto one LB-agar plate (+ampicillin), and 200 μl was lawned onto another. The plates were incubated at 37°C overnight. A single colony of the transformed bacteria was inoculated into 7 ml LB broth (+ampicillin) in a culture tube and incubated at 37°C overnight at 180 rpm.

Plasmid DNA was purified from a 5-ml sample of this culture using the Promega PureYield miniprep kit, according to the manufacturer's protocol. Correct plasmid amplification was confirmed by running a sample on an agarose gel. The remaining 2 ml transformed culture was added to 200 ml LB broth (+ampicillin) and incubated at 37°C overnight. The culture was then centrifuged at 3200 rcf for 15 min. Plasmid DNA was purified from the pellet using the Promega PureYield midiprep kit, according to the manufacturer's protocol.

2.2.5 Transfection

Method 1 (used for transfection with pSMOX, p2708, p3074 or p2T7.177):

The relevant plasmid (10–20 µg) was linearised by incubation with the appropriate restriction enzyme (1 U per µg DNA) in the appropriate buffer, at 37°C overnight. It was then precipitated in a SpeedVac machine and resuspended in dH_2O (10 µl).

The trypanosome cells to be transfected ($2x10^7$–$3x10^7$ cells) were centrifuged at 1500 rcf for 6 min. The pellet was resuspended in cold Cytomix transfection buffer (100 µl). The 10 µl linearised plasmid was added to an electroporation cuvette. The 100 µl cells-in-Cytomix was added to the same cuvette, and the mixture was immediately electroporated in an Amaxa Nucleofector using the Free Choice Program X-001. The electroporated cells were added to 15 ml culture medium in a flask marked A. The flask was mixed thoroughly and 1 ml was taken into a flask marked B, which contained 13 ml culture medium. The flasks were incubated at 37°C in the case of BSFs or 27°C in the case of PCFs. Selection antibiotic was added after 24 h, and cells were transferred onto a 24-well plate (1 ml per well). Transfected clones grew up after 5–10 days.

Method 2 (used for transfection with pPOTv6 or pPOTv7):

Following purification of the long-primer PCR amplicon (3–5 µg DNA in 80 µl) (see Section 2.2.8), the DNA was precipitated by sequential addition of linear acrylamide (10 µl), sodium acetate (3 M; 10 µl) and molecular-grade ethanol (100%; 300 µl; -20°C). The mixture was incubated at -20°C for 15 min, then centrifuged at 15,000 rcf for 10 min. The pellet was washed once in molecular-grade ethanol (70%; 1 ml; -20°C) and centrifuged at 15,000 rcf for 10 min. The pellet was centrifuged again (dry spin) at 15,000 rcf for 10 min and subsequently left to dry in a sterile biological safety cabinet for ~1 h. The DNA was resuspended in 3–5 µl sterile nuclease-free water.

The trypanosome cells to be transfected ($1x10^7$ cells) were centrifuged at 1500 rcf for 6 min. The pellet was resuspended in Amaxa Nucleofector solution (100 µl) at room temperature. The precipitated PCR product (3–5 µl) was added to an Amaxa electroporation cuvette. The 100 µl cells-in-Nucleofector was added to the same cuvette, and the mixture was immediately electroporated in an Amaxa Nucleofector using the Free Choice Program X-001. The electroporated cells were added to 19 ml culture medium in a flask marked A. The flask was mixed thoroughly and 1 ml was taken into a flask marked B, which contained 17 ml culture medium. The flasks were incubated at 37°C in the case of BSFs or 27°C in the case of PCFs. Selection antibiotic was added after 24 h, and cells were transferred onto a 24-well plate (1 ml per well). Transfected clones grew up after 5–10 days.

2.2.6 Conditional RNAi

The FLA2 RNAi p2T7.177 construct was generated as follows. A 414-bp fragment of the *FLA2* gene (Tb927.8.4110; nucleotides 333–746) was amplified by PCR from genomic DNA using the primers shown in Table 2.1. This fragment was chosen as the region of *FLA2* bearing least similarity to *FLA1*, to avoid cross-reaction with the *FLA1* gene (see Table 3.1 for pairwise alignment of the *FLA2* fragment with the corresponding fragment of *FLA1*; see Appendix C for full alignment of *FLA2* and *FLA1*). The primers introduced ***Xho*I** and ***Bam* HI** restriction sites.

Table 2.1: Primers used for amplification of *FLA2* nucleotides 333–746

Primers	Sequence (5' → 3')
FLA2 FP	**CTCGAG**ACGGCCCCAAGGGAGATGCACTA
FLA2 RP	**GGATCC**TATATCCTTACTTTTTCCAATATCAGACTCCCC

Reaction conditions for the amplification are given in Tables 2.2 and 2.3.

Table 2.2: Reaction mixture used for amplification of *FLA2* nucleotides 333–746

Component	Volume	Final Concentration
$MgCl_2$ (25 mM)	3 µl	1.5 mM
5× Green GoTaq® Flexi Buffer	10 µl	1×
PCR Nucleotide Mix (2.5 mM each)	1 µl	50 µM each
FLA2 FP	1 µl	2 µM
FLA2 RP	1 µl	2 µM
Genomic Template DNA	1 µl	100 ng
GoTaq® DNA Polymerase	1 µl	5 U / 50 µl
dH_2O	32 µl	–

Table 2.3: Thermal cycle used for amplification of *FLA2* nucleotides 333–746

Step	Temperature	Duration (min)	No. of Cycles
Initial Denaturation	94°C	1	1
Denaturation	94°C	0.5	35
Annealing	66°C	1	35
Extension	72°C	1	35
Final Extension	72°C	5	1
Incubation	4°C	indefinite	1

The amplified fragment was purified and subsequently ligated into the *Bam* HI/*Xho* I-digested p2T7.177 RNAi vector (Wickstead et al., 2002) using the LigaFast Rapid DNA Ligation System.

The cloned FLA2 RNAi p2T7.177 vector was sequenced to verify that the correct sequence had been amplified (Appendix D).

The FLA2 RNAi vector was amplified in bacteria (as described in Section 2.2.4) and purified. To generate a FLA2 RNAi cell line, the vector was linearised with *Not*I and transfected (see Section 2.2.5) into a BSF cell line that already carried the pSMOX plasmid. The pSMOX plasmid confers RNAi compatibility because it encodes the T7 RNA polymerase and tetracycline repressor (Poon et al., 2012). Selection was maintained using puromycin (2.5 μg ml^{-1}) and phleomycin (2.5 μg ml^{-1}).

Addition of tetracycline (2 μg ml^{-1}) induced expression of *FLA2* dsRNA, and consequent knockdown of the *FLA2* gene.

A FLA3 RNAi construct had already been generated by insertion of a ~600-bp fragment of the *FLA3* gene (nucleotides 1174–1792) into the p2T7.177 RNAi vector (Woods et al., 2013, Wickstead et al., 2002). The fragment had been chosen to avoid cross-reaction with the similar *FLA1BP* gene.

The FLA3 RNAi p2T7.177 vector was amplified in bacteria (as described in Section 2.2.4) and purified. To generate a FLA3 RNAi cell line, the vector was linearised with *Not*I and transfected (see Section 2.2.5) into a BSF cell line that already carried the pSMOX plasmid. Selection was maintained using puromycin (2.5 μg ml^{-1}) and phleomycin (2.5 μg ml^{-1}).

Addition of tetracycline (2 μg ml^{-1}) induced expression of *FLA3* dsRNA, and consequent knockdown of the *FLA3* gene.

2.2.7 *In-situ* Tagging (with p3074/p2708)

Tagging of FLA3 was achieved using the p3074 plasmid, which adds a Ty1 tag to the C-terminal end of the protein and, separately, using p2708, which adds a HA tag to the C-terminal end of the protein. These vectors are both derived from the pC-PTP plasmid and are used for constitutive expression a tagged transgene (Kelly et al., 2007).

The *FLA3*-Ty1 vector was generated as follows. A 1,195-bp fragment of the *FLA3* gene (residues 1261–2455 of Tb927.5.4570) was amplified by PCR using primers that introduced *Swa* I and *Bam* HI restriction sites at the 5′ and 3′ ends, respectively. The resulting 1,232-bp amplicon was ligated into a pGEM-T plasmid, which was then transformed into bacteria for further amplification. The fragment was subsequently isolated by restriction digest and inserted into the p3074 tagging vector between the *Swa* I site and the *Bam* HI site, in frame with the tag and with a 5-residue linker between the end of the ORF and the start of the tag (Fig. 2.1). This p3074 *FLA3*-Ty1 plasmid was transformed into bacteria, amplified and purified for use in further studies. The *FLA3*-HA p2708 vector was generated in the same manner.

For insertion into the trypanosome, the tagging vector (p3074 or p2708) was cut at the unique *Xho* I restriction site (residue 1781) within the *FLA3* ORF, producing a linear fragment of DNA with regions of homology to the *FLA3* gene at either end. Upon transfection, the plasmid inserted into the endogenous gene location (either of the two available alleles) by homologous recombination (Fig. 2.1). The endogenous 3′ UTR was replaced with that of the *RPA1* gene (replication protein A), which may have altered the stability and/or translational efficiency of the mRNA transcribed from the tagged allele. The homologous ends were long (674 bp at the 5′ end and 521 bp at the 3′ end), which ensured efficient transfection and accurate targeting to the *FLA3* gene locus. The transfected cell line therefore expressed C-terminally-tagged FLA3 constitutively, using the endogenous promoter. Selection was maintained using G418 (2.5 μg ml^{-1}).

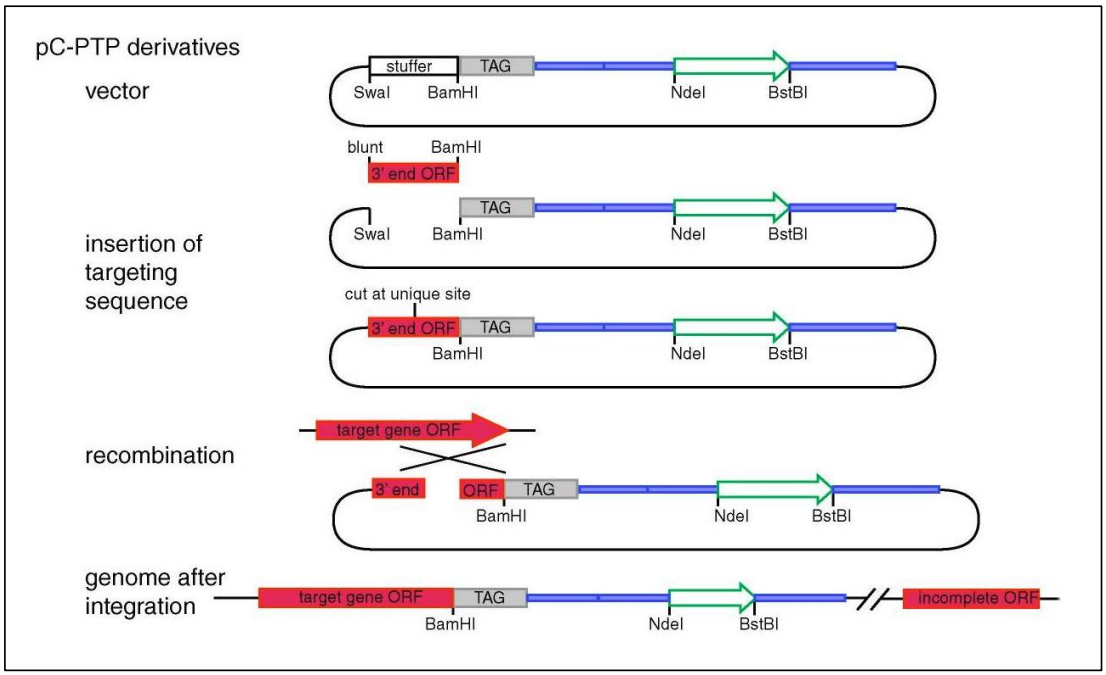


Figure 2.1: *In-situ* tagging by pC-PTP-derived vectors, including p3074 and p2708 (Kelly *et al.* 2007, p. 106, Figure 2, part B). Green arrow, selectable marker gene; red, targeted open reading frame. The 3' end of the target gene is inserted into the plasmid, in frame with the protein tag. The vector is then linearised using a unique restriction site within the ORF. It inserts into the endogenous gene location by homologous recombination. In the cases of p3074 and p2708, the selection marker is G418.

The Ty1 tag is a 10-amino-acid sequence (EVHTNQDPLD) from the major structural protein of the Ty1 virus-like particle, which is a multi-subunit protein structure resulting from the expression of viral structural genes in *Saccharomyces cerevisiae* (Bastin et al., 1996). This sequence was developed as an epitope tag because it is linear and because it can be recognised in different environments. The epitope is recognised by two different monoclonal antibodies that do not cross-react with trypanosomal proteins, and it has been shown not to contain any cryptic addressing signal that would cause it to be mislocalised in the trypanosome. The antibodies that recognise the Ty1 tag are BB2 (which is an IgG) and TYG 5 (which is an IgA). It is quite unusual to have two different antibodies for a specific epitope. This gives the tag potential for double-labelling in

immunofluorescence or immunoelectron microscopy. The p3074 vector adds a 4×Ty1 tandem (4.7 kDa) to its target gene.

The HA tag is a 9-amino-acid sequence (YPYDVPDYA) from the human-influenza haemagglutinin protein. It has been used extensively as an epitope tag because it does not tend to interfere with the activity or distribution of the protein which it tags. The p2708 vector adds a 12×HA tandem (13.2 kDa) to its target gene. Tandem arrays of epitope tags give a stronger signal and allow bivalent binding of the antibody to a single tagged protein, which improves the efficiency of immunoprecipitation. The p2708 vector also includes a TEV-protease site between the protein and tag to facilitate release of the protein if the tag is used for purification (Bastin et al., 1996).

2.2.8 Long-Primer PCR Tagging (with pPOTv6/pPOTv7)

FLA2:

Tagging of FLA2 at the C-terminus was achieved by long-primer PCR tagging. This approach was taken for two reasons: firstly, the method makes use of the 3' UTR of the gene to be tagged, which avoids cross-reaction with *FLA1*; secondly, this method is quicker and easier than traditional restriction-cloning methods because it eliminates the need for time-consuming amplification of the tagging vector in bacteria.

The pPOT series of plasmids was developed specifically for PCR-based endogenous gene tagging in trypanosomatids (Dean et al., 2015). Long-primer PCR tagging takes advantage of the fact that trypanosomes require only 50 nucleotides of identity for targeted homologous recombination.

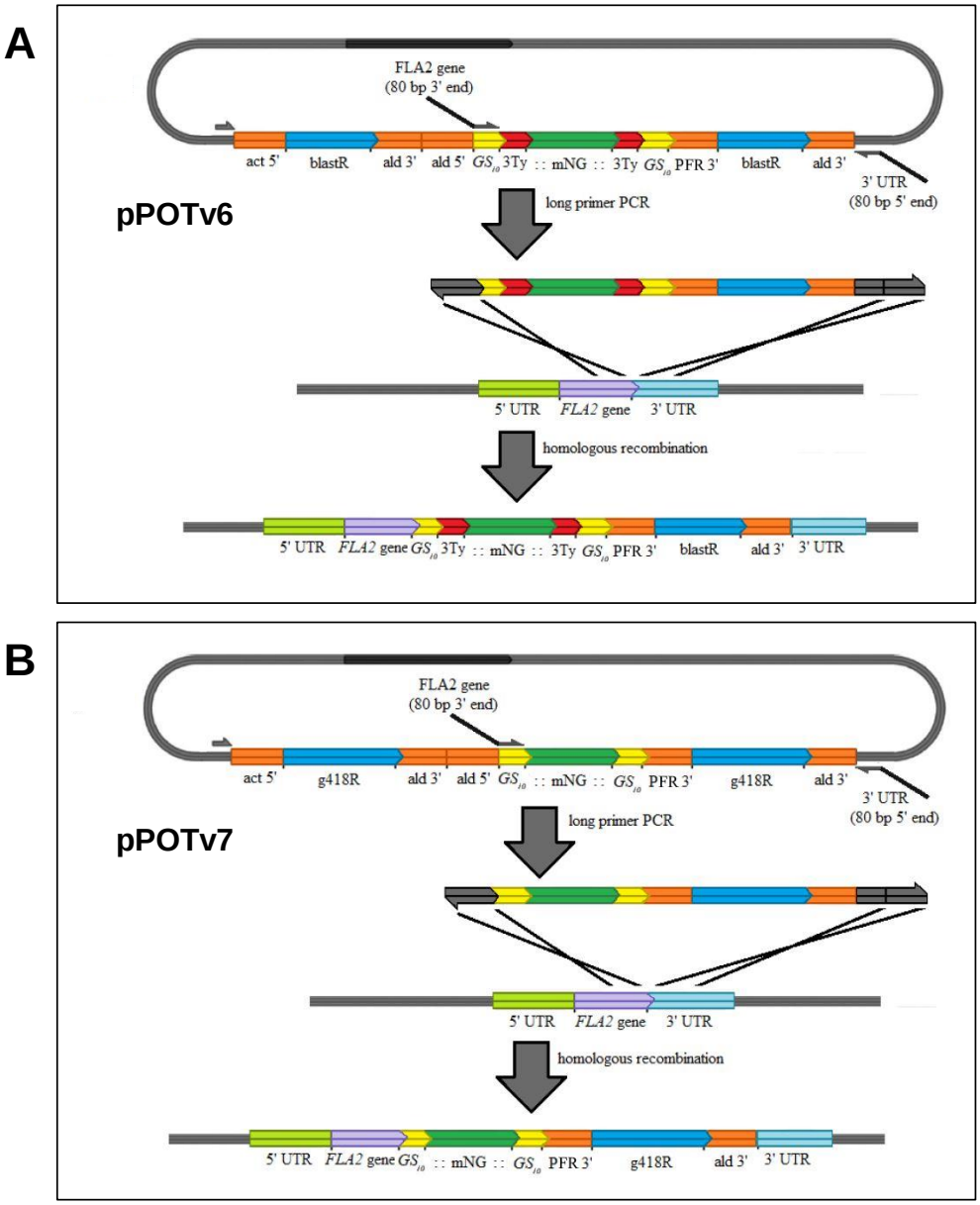


Figure 2.2: Long-primer PCR tagging using the pPOTv6/pPOTv7 plasmids (Dean *et al.*, 2015). Green, mNeonGreen fluorescent tag; lilac, target gene; red, Ty1 tags; yellow, glycine-serine linkers; blue, resistance genes.

For C-terminal tagging, the forward primer includes the last 80 nucleotides of the target ORF (excluding the stop codon) and the reverse primer includes the first 80 nucleotides of the target gene's 3' UTR in reverse complement.

The PCR amplicon inserts into the endogenous locus by homologous recombination.

C-terminal tagging with the pPOTv6 vector is achieved as depicted in Fig. 2.2 (A). The plasmid contains a fluorescent protein (neon green (NG) or scarlet (SR), see below) flanked by three Ty1 residues on each side and followed by the 3′ UTR of the *PFR2* gene, and also a resistance gene (blasticidin). A long forward primer containing the last 80 nucleotides of the gene to be tagged (excluding the stop codon) binds specifically to a spacer region just upstream of the tagging cassette. A long reverse primer containing the first 80 nucleotides of the 3′ UTR of the gene to be tagged (in reverse complement) binds specifically to a spacer region downstream of the resistance gene. The amplicon produced by PCR, therefore, can recombine into the endogenous gene locus by homologous recombination. Thus, the gene is tagged with Ty1-NG-Ty1 (or Ty1-SR-Ty1) at the C-terminus. The 3′ UTR is replaced by the 3′ UTR of *PFR2*, which may alter the stability of the mRNA product of the tagged gene.

The pPOTv7 vector is essentially identical to pPOTv6, except that it lacks the six Ty1 residues (Fig. 2.2 B).

The neon-green protein (NG) conferred by pPOTv6/7 is mNeonGreen (often denoted mNG), a bright 27-kDa monomeric fluorescent protein (Shaner et al., 2013). The scarlet protein (SR) conferred by the variant versions of pPOTv6/7 is mScarlet, a bright 26-kDa monomeric fluorescent protein (Bindels et al., 2017). Inclusive of the six Ty1 residues, the tag conferred by pPOTv6 amounts to ~36 kDa.

The tagging region of the pPOTv6 NG plasmid encodes the following protein sequence:

M G S G S G S G S G S E V H T N Q D P L D G S E V H T N Q D P L D G S
E V H T N Q D P L D K L V S K G E E D N M A S L P A T H E L H I F G S
I N G V D F D M V G Q G T G N P N D G Y E E L N L K S T K G D L Q F S
P W I L V P H I G Y G F H Q Y L P Y P D G M S P F Q A A M V D G S G Y
Q V H R T M Q F E D G A S L T V N Y R Y T Y E G S H I K G E A Q V K G
T G F P A D G P V M T N S L T A A D W C R S K K T Y P N D K T I I S T
F K W S Y T T G N G K R Y R S T A R T T Y T F A K P M A A N Y L K N Q
P M Y V F R K T E L K H S K T E L N F K E W Q K A F T D V M G M D E L
Y K G S E V H T N Q D P L D G S E V H T N Q D P L D G S E V H T N Q D
P L D G S G S G S G S G S G S Stop

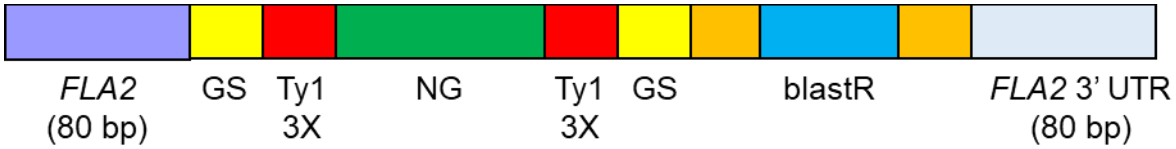
(GS spacers in yellow; Ty1 tags in red; mNeonGreen tag in green)

For the initial FLA2 tagging attempts (described in Sections 5.2.1 and 5.2.2), the forward primer (**FP1**) consisted of the last 80 nt of the *FLA2* gene (excluding the stop codon), followed by 18 nt identical to the region of the pPOT plasmid highlighted in black below.

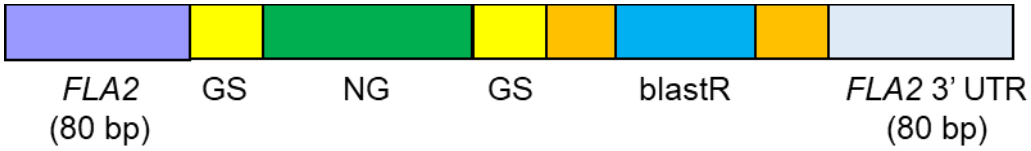
M G S G S G S G S G S E V H T N Q D P L D G S E V H T N Q D P L D G S
E V H T N Q D P L D K L V S K G E E D N M A S L P A T H E L H I F G S
I N G V D F D M V G Q G T G N P N D G Y E E L N L K S T K G D L Q F S
P W I L V P H I G Y G F H Q Y L P Y P D G M S P F Q A A M V D G S G Y
Q V H R T M Q F E D G A S L T V N Y R Y T Y E G S H I K G E A Q V K G
T G F P A D G P V M T N S L T A A D W C R S K K T Y P N D K T I I S T
F K W S Y T T G N G K R Y R S T A R T T Y T F A K P M A A N Y L K N Q
P M Y V F R K T E L K H S K T E L N F K E W Q K A F T D V M G M D E L
Y K G S E V H T N Q D P L D G S E V H T N Q D P L D G S E V H T N Q D
P L D G S G S G S G S G S Stop

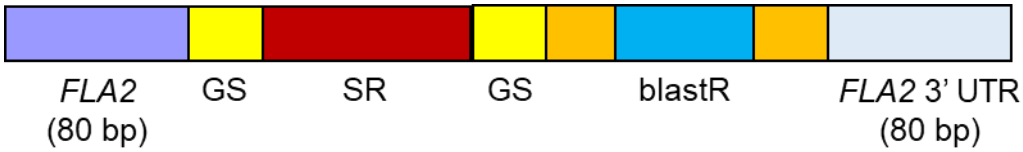
The reverse primer (**RP1**) consisted of the first 80 nt of the *FLA2* 3' UTR (in reverse complement) followed by a 20 nt region identical to a spacer region downstream of the resistance gene on the pPOT plasmid. [All primer sequences used are given in Appendix B.]

Thus, when used in conjunction with pPOTv6, the FP1 and RP1 primers produced the following tagging amplicon.



When used in conjunction with pPOTv7 (SR) or pPOTv7 (NG), respectively, the FP1 and RP1 primers produced the following tagging amplicons.





However, when *FLA2* was tagged using the above amplicons, the tagged protein mislocalised throughout the cell body (Section 5.2.1, 5.2.2). A new forward primer, **FP2**,

was designed to bypass the NeonGreen sequence thereby creating a tagging amplicon that would add just a 3X Ty1 tag to the *FLA2* gene. (Section 5.2.4)

FP2 consisted of the last 80 nt of the *FLA2* gene, followed by 21 nt identical to the black region of the pPOTv6 tagging region below.

M G S G S G S G S G S E V H T N Q D P L D G S E V H T N Q D P L D G S
E V H T N Q D P L D K L V S K G E E D N M A S L P A T H E L H I F G S
I N G V D F D M V G Q G T G N P N D G Y E E L N L K S T K G D L Q F S
P W I L V P H I G Y G F H Q Y L P Y P D G M S P F Q A A M V D G S G Y
Q V H R T M Q F E D G A S L T V N Y R Y T Y E G S H I K G E A Q V K G
T G F P A D G P V M T N S L T A A D W C R S K K T Y P N D K T I I S T
F K W S Y T T G N G K R Y R S T A R T T Y T F A K P M A A N Y L K N Q
P M Y V F R K T E L K H S K T E L N F K E W Q K A F T D V M G M D E L
Y K G S E V H T N Q D P L D G S E V H T N Q D P L D G S E V H T N Q D
P L D G S G S G S G S G S Stop

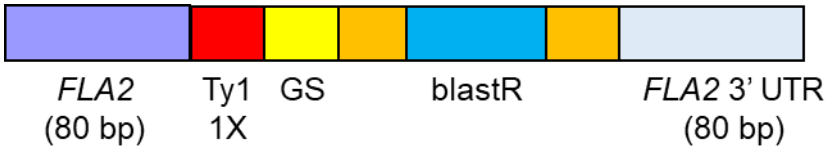
The FP2 and RP1 primers amplified the following tagging amplicon from the pPOTv6 template.

When *FLA2* was tagged with the above amplicon, the tagged protein mislocalised throughout the cell body. A new forward primer, **FP3**, was therefore designed to add a single Ty1 residue to the *FLA2* gene. (Section 5.2.8).

FP3 consisted of the last 80 nt of the *FLA2* gene, followed by 22 nt identical to the last Ty1 residue on the pPOTv6 plasmid, i.e. the black region below:

M G S G S G S G S G S E V H T N Q D P L D G S E V H T N Q D P L D G S
E V H T N Q D P L D K L V S K G E E D N M A S L P A T H E L H I F G S
I N G V D F D M V G Q G T G N P N D G Y E E L N L K S T K G D L Q F S
P W I L V P H I G Y G F H Q Y L P Y P D G M S P F Q A A M V D G S G Y
Q V H R T M Q F E D G A S L T V N Y R Y T Y E G S H I K G E A Q V K G
T G F P A D G P V M T N S L T A A D W C R S K K T Y P N D K T I I S T
F K W S Y T T G N G K R Y R S T A R T T Y T F A K P M A A N Y L K N Q
P M Y V F R K T E L K H S K T E L N F K E W Q K A F T D V M G M D E L
Y K G S E V H T N Q D P L D G S E V H T N Q D P L D G S E V H T N Q D
P L D G S G S G S G S G S Stop

Thus, the FP3 and RP1 primers amplified the following tagging amplicon from the pPOTv6 template.



Again, the tagged protein mislocalised. To make an even smaller tag, a forward primer, **FP4**, was designed to confer a single Ty1 residue without a GS spacer afterwards. (Section 5.2.8)

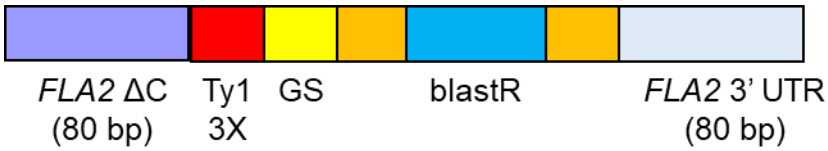
FP4 consisted of the last 80 nt of the *FLA2* gene, followed by 39 nt identical to the last Ty1 residue on the pPOTv6 plasmid, i.e. the black region below, *followed by a stop codon*.

M G S G S G S G S G S E V H T N Q D P L D G S E V H T N Q D P L D G S
E V H T N Q D P L D K L V S K G E E D N M A S L P A T H E L H I F G S
I N G V D F D M V G Q G T G N P N D G Y E E L N L K S T K G D L Q F S
P W I L V P H I G Y G F H Q Y L P Y P D G M S P F Q A A M V D G S G Y
Q V H R T M Q F E D G A S L T V N Y R Y T Y E G S H I K G E A Q V K G
T G F P A D G P V M T N S L T A A D W C R S K K T Y P N D K T I I S T
F K W S Y T T G N G K R Y R S T A R T T Y T F A K P M A A N Y L K N Q
P M Y V F R K T E L K H S K T E L N F K E W Q K A F T D V M G M D E L
Y K G S E V H T N Q D P L D G S E V H T N Q D P L D G S E V H T N Q D
P L D G S G S G S G S G S Stop

Thus, the FP4 and RP1 primers amplified the following tagging amplicon from the pPOTv6 template.

Still, the tagged protein mislocalised. It was decided to remove the C-terminal tail of *FLA2* and add 3X Ty1 tag just after the TM domain. (Section 5.2.9) A new forward primer, **FP5**, was designed to achieve this. FP5 consisted of the 80 nt up to and including the alanine (GCA) at nucleotides 1720–1722 of the *FLA2* gene, followed by the same region of identity with the pPOTv6 plasmid as was used previously for FP2.

Thus, the FP5 and RP1 primers amplified the following tagging amplicon from the pPOTv6 template.



Still, the tagged protein mislocalised. As a final attempt, a forward primer, **FP6**, was designed to replace the C-terminal tail of *FLA2* with a single Ty1 residue. (Section 5.2.10)

FP6 consisted of the 80-nt region of identity to *FLA2* that was used in FP5, followed by the same region of identity to the pPOTv6 plasmid that was used in FP4. Thus, the FP6 and RP1 primers amplified the following tagging amplicon from the pPOTv6 template.

In all of the above tagging experiments, the PCR reaction mixture was as described in Table 2.4, and the PCR cycle was as described in Table 2.5. After amplification of the tagging amplicon, the PCR product was purified using a peqGOLD Cycle-Pure Kit. Precipitation and transfection were performed as described in Section 2.2.5.

Table 2.4: Reaction mixture used for long-primer PCR

Component	Volume	Final Concentration
Expand™ High Fidelity PCR Buffer (10X) with $MgCl_2$	5 µl	1X
DMSO (2.68% in dH_2O)	37.25 µl	–
Forward Primer	2 µl	10 µM
Reverse Primer	2 µl	10 µM
dNTPs (10 mM)	1 µl	200 µM
Template DNA (pPOTv6/7)	2 µl	15 ng
Expand™ High Fidelity Enzyme Mix	0.75 µl	50 µl (2.6 U)

Table 2.5: Thermal cycle used to amplify a tagging amplicon from pPOTv6/pPOTv7

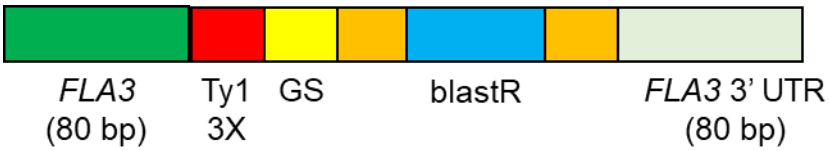
Step	Temperature	Duration	No. of Cycles
Stage 1			**1**
Initial Denaturation	94°C	2 min	
Stage 2			**10**
Denaturation	94°C	15 s	
Annealing	65°C	30 s	
Elongation	72°C	2 min	
Stage 3			**20**
Denaturation	94°C	15 s	
Annealing	65°C	30 s	
Elongation	72°C	2 min + 5 s/cycle	
Stage 4			**1**
Final Elongation	72°C	7 min	
Cooling	4°C	indefinite	

FLA3:

Long-primer PCR tagging was used to add a 3X Ty1 tag to the C-terminus of *FLA3* (Section 4.2.5). The PCR cycle and reaction mixture were the same as described for *FLA2*, above.

The forward primer, **FLA3 FP1**, consisted of the last 80 nt of the *FLA3* gene, followed by the same region of identity to the pPOTv6 plasmid that was used for the FP2 primer above. The reverse primer, **FLA3 RP1**, consisted of the first 80 nt of the *FLA3* 3' UTR (in reverse complement) followed by a 20 nt region identical to a spacer region downstream of the resistance gene on the pPOTv6 plasmid.

Thus, the FLA3 FP1 and FLA3 RP1 primers amplified the following tagging amplicon from the pPOTv6 template.



2.3 Methods: Protein Studies

2.3.1 SDS-PAGE

To prepare cells for SDS-PAGE, they were centrifuged at 1500 rcf for 6 min. The pellet was resuspended in 2 ml PSG in two Eppendorf tubes, and centrifuged at 1500 rcf for 6 min. 500 µl supernatant was discarded from each tube. The pellets were resuspended, combined and centrifuged again at 1500 rcf for 6 min. The resulting pellet was resuspended in equal parts PBS and 2X sample buffer to have a final concentration of $2.5x10^5$ cell-equivalents μl^{-1}. The sample was heated at 100°C for 2 min and then syringe-sheared if excessively viscous.

Samples were loaded in 20-µl aliquots, i.e. the equivalent of $5x10^6$ cells in each well. Gels (10% acrylamide) were run at 100 V for 30 min, then at 120 V for ~1 h 30 min. If not used immediately for western blotting or silver staining, gels were left in the electrophoresis apparatus for up to 24 h.

2.3.2 Western Blotting and Lectin Blotting

SDS-PAGE gels were soaked in transfer buffer and then transferred onto nitrocellulose membrane using semi-dry transfer apparatus. A constant current (3 mA per cm^2 blot) was applied for 2 h.

Blots were stained briefly with Ponceau S solution to reveal all protein bands present. After several quick rinses in deionised water to remove background staining, blots were scanned. A subsequent wash (~2 min on orbital shaker) in TBST removed all Ponceau stain.

Blots were blocked in 10% Marvel in TBST (5 ml) at for 1 h at r.t., on a roller. Primary antibody was then added to the blocking solution, and the blot was incubated on a roller at 4°C overnight. The blot was washed in TBST (5 × 10 min) on an orbital shaker before being incubated in secondary antibody in 10% Marvel in TBST (5 ml) for 1–3 h at room temperature, on a roller. It was washed in TBST again (5 × 10 min) on an orbital shaker.

If the secondary antibody was AP-conjugated, the blot was viewed by colorimetric detection. The blot was placed in AP buffer (10 ml) containing 5-bromo-4-chloro-3-indolyl phosphate (BCIP) (165 µg ml^{-1}) and nitro blue tetrazolium (NBT) (330 µg ml^{-1}). It was incubated on a roller at room temperature for 2–30 min, until bands appeared. The colorimetric reaction was stopped with dH_2O or TBST.

If the secondary antibody was HRP-conjugated, the blot was viewed by chemiluminescence, using Immobilon Western Chemiluminescent HRP Substrate, according to the manufacturer's protocol. It was then viewed under UV light in a gel-documentation system.

In cases where biotinylated lectin was used instead of a primary antibody, the blot was blocked in bovine serum albumin (2% in TBST) instead of Marvel milk. Biotinylated lectin was diluted 1-in-250 in PBS (5 ml) containing BSA (2%) and $MgCl_2$ (5 µl). AP-conjugated streptavidin (1-in-1000 in PBS containing BSA (2%)) was used instead of a secondary antibody.

2.3.3 Cell Fractionation by Detergent

Trypanosome cells ($1x10^8$) were centrifuged at 1500 rcf for 6 min. The pellet was washed once in PSG, then resuspended at $1x10^9$ cells ml^{-1} in lysis buffer containing Sigma P8340 protease-inhibitor cocktail (1 in 100). To this was added an equal volume lysis buffer containing Triton X-100 (2%), to give a final concentration of $5x10^8$ cells ml^{-1} in 1% Triton. The mixture was incubated on ice for 30 min. A sample of this total lysate was taken and prepared for western blotting. The remainder was centrifuged at 18,000 rcf for 15 min. The supernatant was isolated and prepared for western blotting. The pellet was washed once, gently, with lysis buffer (100 µl), without centrifugation. Lysis buffer containing urea (4 M) was added (an equal volume to the volume of the supernatant isolated earlier), and the pellet was resuspended. This pellet fraction was then prepared for western blotting.

2.3.4 Cell Fractionation by Freeze-Thaw

Trypanosome cells ($1x10^8$) were centrifuged at 1500 rcf for 6 min. The pellet was washed once in PSG, then resuspended at $5x10^8$ cells ml^{-1} in lysis buffer containing Sigma P8340 protease-inhibitor cocktail (1 in 200). A hole was pierced in the lid of the Eppendorf, and the tube was dipped into liquid nitrogen for 6 seconds. It was left to thaw to room temperature. This freeze-thaw procedure was repeated twice more. The sample was then vortexed. A sample of this total lysate was prepared for western blotting. The remainder was spun at 18,000 rcf for 15 min. The supernatant was isolated and prepared for western blotting. The pellet was washed once, gently, with lysis buffer (100 µl), without centrifugation. Lysis buffer containing urea (4 M) was added (an equal volume to the volume of the supernatant isolated earlier), and the pellet was resuspended. This pellet fraction was then prepared for western blotting.

2.3.5 Lectin Pulldown

Trypanosomes (1×10^8 cells) were centrifuged at 1500 rcf for 6 min. The pellet was washed once in PSG, then resuspended at $1x10^9$ cells ml^{-1} in lysis buffer containing Sigma P8340 protease-inhibitor cocktail (1 in 100). To this was added an equal volume lysis buffer containing Triton X-100 (2%), to give a final concentration of $5x10^8$ cells ml^{-1} in 1% Triton. The mixture was incubated on ice for 30 min, then centrifuged at 18,000 rcf for 15 min. A sample of the supernatant was prepared for western-blotting ("soluble lysate") and the remainder was put to one side. The pellet was washed once with lysis buffer (100 µl), centrifuged at 18,000 rcf for 10 min and resuspended in lysis buffer containing urea (4 M) at $1x10^9$ cells ml^{-1}. This detergent-insoluble fraction was then prepared for western-blotting.

Ricinus communis agglutinin (RCA-I) lectin (20 µl) was washed by addition of lysis buffer (1 ml) containing Triton X-100 (1%) and centrifugation at 3000 rcf for 10 min. The supernatant was discarded, except for a volume equal to the volume of the pellet, producing a 1:1 slurry. This slurry was added, using a cut-off tip, to the remainder of the detergent lysate that was put aside earlier. The mixture was incubated on a rotator for 1 h at room temperature. It was then centrifuged at 15,000 rcf for 30 seconds. The supernatant ("unbound fraction") was prepared for western blotting. The pellet was washed twice in lysis buffer containing Triton X-100 (1%), with centrifugation at 15,000 rcf

for 30 seconds each time. The resultant pellet ("lectin-bound fraction") was prepared for western-blotting. Before loading the lectin-bound sample on a gel, it was boiled for 3 min to release the protein from the lectin and subsequently centrifuged to pellet the lectin beads. The sample was loaded from the supernatant.

2.3.6 Immunoprecipitation

Method 1: Nondenaturing Conditions

Trypanosome cells ($2x10^8$) were centrifuged at 1500 rcf for 6 min. The pellet was resuspended in PSG (10 ml) to wash and centrifuged again. The washed pellet was resuspended at $1x10^9$ cells ml^{-1} in PSG (-S/G) containing Sigma P8340 protease-inhibitor cocktail (1 in 100). An equal volume PSG (-S/G) containing Triton X-100 (2%) was added, producing a final concentration of $5x10^8$ cells ml^{-1} in 1% Triton. The mixture was incubated on ice for 15 min and centrifuged at 18,000 rcf for 15 min. The pellet was discarded. A sample of the supernatant ("Soluble Lysate") was prepared for western blotting. Primary antibody (BB2 anti-Ty1, 50 µl) was added to the remainder of the supernatant. The mixture was incubated on a rotator at 4°C overnight.

Protein-G agarose slurry (30 µl) was added and the mixture was incubated on a rotator at r.t. for 3 h before centrifugation at 9000 rcf for 30 s. The supernatant ("Unbound" fraction) was prepared for western blotting. The pellet was washed twice in Tris-Triton buffer (1 ml; 9000 rcf; 30 s), once in Tris-NaCl buffer (1 ml; 9000 rcf; 30 s) and twice in Tris buffer (1 ml; 9000 rcf; 30 s and 2 min, respectively). The washed pellet was centrifuged (dry) at 9000 rcf for 2 min, and any remaining supernatant was discarded. Sample buffer (20 µl) was added. The mixture was boiled for 2 min and centrifuged at 9000 rcf for 5 min. The supernatant ("Bound" fraction) was used for western blotting.

Method 2: Denaturing Conditions

Trypanosome cells ($2x10^8$) were centrifuged at 1500 rcf for 6 min. The pellet was resuspended in PSG (10 ml) to wash and centrifuged again. The washed pellet was resuspended at $1x10^9$ cells ml^{-1} in PSG (-S/G) containing Sigma P8340 protease-inhibitor cocktail (1 in 100). An equal volume PSG (-S/G) containing SDS (4%) was added, producing a final concentration of $5x10^8$ cells ml^{-1} in 2% SDS. The mixture was boiled for 3 min. To

the mixture was added two volumes (i.e. 800 µl) PSG (-S/G) containing Triton X-100 (2.5%). The sample was incubated on ice for 1 h, then centrifuged at 9000 rcf for 5 min. The pellet was discarded. A sample of the supernatant was prepared for western blotting ("Soluble Lysate"). Primary antibody (BB2 anti-Ty1, 50 µl) was added to the remainder of the supernatant. The mixture was incubated on a rotator at 4°C overnight.

The protocol was completed exactly as described above in Method 1.

2.3.7 Silver Staining

After SDS-PAGE, a gel was incubated in fixer for at least 1 h. It was washed once in dH2O. It was incubated in incubation solution for 30 min, then washed in dH2O (3 × 5 min). It was incubated in silver solution for 20 min, then washed once in dH2O. It was incubated in developing solution, in the dark, for 5–10 min until bands were visible. It was incubated in stop solution for 30 min, and then washed in dH2O for 30 min. All staining solutions were freshly prepared, and all containers were cleaned with strong acid and ethanol prior to use.

2.3.8 Surface Biotinylation

Trypanosome cells ($2.5x10^8$) were centrifuged at 1500 rcf for 6 min, washed once in PSG (10 ml) and resuspended in PSG (450 µl) in an Eppendorf tube. To this was added a membrane-impermeable biotin solution (2.5 mg EZ-Link Sulfo-NHS-SS-Biotin in 50 µl PSG). The mixture was incubated on ice for 15 min, then quenched by addition of glycine (5 mM in 9.5 ml cold PSG). After a further 10 min incubation on ice, the cells were centrifuged at 1500 rcf for 6 min. The cell pellet was washed twice in PSG (10 ml, then 1 ml), with centrifugation each time. The final pellet was resuspended at $1x10^9$ cells ml^{-1} in 250 µl PSG (-S/G) containing protease-inhibitor cocktail (1 in 100). A sample was taken for western blotting; the remainder was detergent-lysed and immunoprecipitated as described in Section 2.3.6 (Method 1).

2.3.9 Flow Cytometry

Trypanosomes ($2x10^6$ cells) were centrifuged at 1500 rcf for 6 min, resuspended in PSG (1 ml) and centrifuged again. Upon resuspension in PSG at $1x10^7$ cells ml^{-1}, the sample was stored on ice.

Fluorescence data for cells expressing mScarlet-tagged FLA2 were acquired on an LSRFortessa flow cytometer, using the FACSDiva software. Red fluorescence was measured according to an excitation wavelength of 561 nm and an emission range of 610/20.

Fluorescence data for cells expressing mNeonGreen-tagged FLA2 were acquired on an Accuri C6 flow cytometer, using the C6 software. Green fluorescence was measured according to an excitation wavelength of 488 nm and an emission range of 530/30.

Intact cells were gated according to their characteristic forward and side scatter (FSC/SSC). Fluorescence emission was presented as single-parameter histograms.

2.4 Methods: Microscopy

2.4.1 Phase-Contrast Microscopy

Trypanosomes were centrifuged at 1500 rcf for 6 min and resuspended in PSG at $1x10^7$ cells ml^{-1}. An equal volume paraformaldehyde (6% in PBS) was added and the sample was incubated at room temperature for 10 min before centrifugation at 600 rcf for 10 min. The fixed cells were resuspended at $1x10^7$ cells ml^{-1} and added to a polylysine slide (20 µl per well). The slide was allowed to dry completely under sterile air flow for ~1 h, then rinsed in PBS for 2 min with gentle agitation. The attached cells were mounted with ProLong™ Gold, covered with a coverslip and sealed with nail varnish. The slide was viewed on a Zeiss Axiovert 100 TV microscope.

2.4.2 Direct Fluorescence Microscopy

Live cells:

Trypanosomes were centrifuged at 1500 rcf for 6 min, washed once in PBS and resuspended in cold PBS at $1x10^7$ cells ml^{-1}. The sample was viewed in an Ibidi µ-Slide 8-Well Chamber Slide, using a Leica SP8 microscope.

Fixed cells:

Trypanosomes were centrifuged at 1500 rcf for 6 min and resuspended in cold PSG at $2x10^7$ cells ml^{-1}. An equal volume paraformaldehyde (0.6% PFA in PBS) was added, and the mixture was incubated at room temperature for 10 min, centrifuged at 600 rcf for 5 min

and resuspended in cold PSG at $1x10^7$ cells ml^{-1}. The sample was viewed in an Ibidi µ-Slide 8-Well Chamber Slide, using a Leica SP8 microscope.

2.4.3 Immunofluorescence

PFA Method:

Trypanosomes were centrifuged at 1500 rcf for 6 min and resuspended in cold PSG at $2x10^7$ cells ml^{-1}. An equal volume paraformaldehyde (6% PFA in PBS) was added, and the mixture was incubated at room temperature for 10 min, centrifuged at 600 rcf for 8 min and resuspended in cold PSG at $2x10^7$ cells ml^{-1}. This fixed sample was stored at 4°C if not used immediately.

Slides (5-well) were cleaned with ethanol before high-molecular-weight (≥300,000 kDa) poly-L-lysine (20 µl 0.01% solution) was added to each of the three inner wells. The slides were left at room temperature until dry.

The fixed cell samples were added to the polylysine-coated wells (20 µl per well) and allowed to settle for 1 h (humid conditions; r.t.). The slides were lowered gently into PBS containing glycine (1 mg ml^{-1}) for 2 min, with gentle agitation. They were then lowered gently into PBS for 2 min, with gentle agitation. Blocking Buffer A (20µl) was added to each well, and the slides were incubated overnight (humid conditions; 4°C).

The slides were then lowered briefly into PBS to remove the blocking solution. Primary antibody diluted in Blocking Buffer B (20 µl) was added to each well, and the slides were incubated for 2–4 h (humid conditions; r.t.). The slides were washed gently in PBS (3 × 2 min). Secondary antibody diluted in Blocking Buffer C (20 µl) was added to each well, and the slides were incubated for 1–2 h (dark humid conditions; r.t.). From this point onwards, the slides were kept out of direct light. The slides were washed gently in PBS (3 × 2 min). Mounting solution (2 µl) (see Section 2.1.1) or ProLong™ Gold (2 µl) was added to each well. A coverslip was added to each slide and sealed with nail varnish. The slides were viewed by confocal fluorescence microscopy immediately, or else stored at 4°C overnight, in the dark, and viewed the following day.

Methanol Method:

Trypanosomes were centrifuged at 1500 rcf for 6 min and resuspended in cold PBS at $2x10^7$ cells ml^{-1}. They were added to polylysine-coated slides (20 µl per well) and dried under sterile air flow for 1 h, then rinsed gently in cold PBS. Methanol (molecular grade; -20°C) was added (20 µl per well). The slides were incubated at -20°C for 20 min, then rehydrated in PBS for 20 min (r.t.). The fixed samples were blocked in BSA (3% in PBS) for 1 h (r.t.) and subsequently lowered briefly into PBS to remove the blocking solution. Primary antibody diluted in BSA (1% in PBS) (20 µl) was added to each well, and the slides were incubated overnight (humid conditions; 4°C). The slides were washed gently in PBS (3 × 5 min). Secondary antibody diluted in BSA (1% in PBS) (20 µl) was added to each well, and the slides were incubated for 1 h (dark humid conditions; 4°). From this point onwards, the slides were kept out of direct light. The slides were washed gently in PBS (3 × 5 min). Mounting solution (2 µl) was added to each well. A coverslip was added to each slide and sealed with nail varnish. The slides were viewed by confocal fluorescence microscopy immediately.

PFA and Methanol Method:

Trypanosomes were centrifuged at 1500 rcf for 6 min and resuspended in cold PBS at $2x10^7$ cells ml^{-1}. An equal volume paraformaldehyde (6% PFA in PBS) was added, and the mixture was incubated at room temperature for 10 min, centrifuged at 600 rcf for 8 min and resuspended in cold PBS at $2x10^7$ cells ml^{-1}. This fixed sample was added to polylysine-coated slides (20 µl per well) and dried under sterile air flow for 1 h, then rinsed gently in cold PBS. Methanol (molecular grade; -20°C) was added (20 µl per well). The slides were incubated at -20°C for 20 min, then rehydrated in PBS for 20 min (r.t.). The immunofluorescence probing was performed as described in the Methanol Method above.

2.4.4 Confocal Microscopy

Imaging was performed on a Leica SP8 scanning confocal microscope, using the Leica Application Suite X (LAS X) software.

2.4.5 Stimulated Emission Depletion (STED) Microscopy

Poly-L-lysine (0.01%; molecular weight ≥300,000 kDa; 50 µl per coverslip) was added to high-precision round coverslips (#1.5 H) and allowed to settle for 1–2 h. Excess polylysine was subsequently poured off and the coverslips were dried under sterile air flow.

Trypanosomes were centrifuged at 1500 rcf for 6 min, washed once in cold PSG and resuspended in cold PSG at $2x10^7$ cells ml^{-1}. An equal volume paraformaldehyde (6% PFA in PBS) was added, and the mixture was incubated at room temperature for 10 min, centrifuged at 600 rcf for 8 min and resuspended in cold PSG at $2x10^7$ cells ml^{-1}.

The fixed cells were added (20 µl per coverslip) to the polylysine-coated coverslips and allowed to settle for 2 h (humid conditions; r.t.). The coverslips were washed once in glycine (1 mg ml^{-1} in PBS; 2 min) and once in PBS (2 min). All wash steps were performed at 12 rpm.

Blocking Buffer A (20 µl) was added to each coverslip. The coverslips were incubated overnight (humid conditions; 4°C), then rinsed gently in PBS. Primary antibody diluted in Blocking Buffer B (20 µl) was added, and the coverslips were incubated for 3 h (humid conditions; r.t.). The coverslips were washed in PBS (3 x 2 min; 12 rpm). Secondary antibody (green fluorescence only) diluted in Blocking Buffer C (20 µl) was added and the coverslips were incubated for 1 h (dark humid conditions; r.t.). The coverslips were washed in PBS (5 x 2 min; 12 rpm; dark conditions). ProLong™ Gold (1 µl) was added to each coverslip and a plain glass slide was added on top. The samples were left to cure for 24 h (horizontally, in the dark, at r.t.).

The slides were viewed on a Leica SP8 gated STED microscope. Images were deconvoluted using the Huygens Professional Deconvolution Wizard, with a signal-to-noise ratio of 7.

Chapter 3: Function of FLA2 and FLA3; Localisation of FLA3

3.1 Introduction

The first indication that membrane glycoproteins are involved with flagellar attachment in *T. brucei* came with the discovery of FLA1, nearly twenty-five years ago (Nozaki et al., 1996). Evidence that related genes might exist arose through identification of a cross-hybridising band on northern blots of bloodstream-form cells (LaCount et al., 2002). Upon induction of what was supposed to be FLA1-specific RNAi, another gene was knocked down at the same time. This gene, which is 73% similar to *FLA1*, was given the name *FLA2*. Because both genes had been ablated simultaneously, it was not known whether the drastic knockdown phenotype (flagellar detachment, cytokinesis failure, cell death) was attributable to loss of FLA1, loss of FLA2 or loss of both proteins.

Surprisingly, no published study has attempted to resolve this ambiguity, possibly because of the difficulty in producing molecular tools (antibodies/RNAi vectors, etc.) that enable manipulation of FLA2 without cross-reaction with FLA1. However, several high-throughput proteomic studies have suggested that FLA1 is not, in fact, produced at a meaningful level in BSFs (Section 1.4.4) – a finding which strongly implies that the knockdown phenotype observed by LaCount et al. was due to loss of the FLA2 protein.

This thesis investigates the possibility that FLA2 and FLA1 are paralogues: homologous proteins with the same core function, but each adapted for a different life-cycle stage. Prior to characterisation of FLA2 at the protein level (Chapter 5), it was first necessary to prove that FLA2 is indeed required for flagellar attachment in BSFs. This was achieved by generating an RNAi vector specific to FLA2 (Section 3.2.1).

In contrast to FLA2, FLA3 had already undergone preliminary characterisation at the protein level. Woods et al. (2013) used conditional RNAi to show that FLA3 is essential in BSFs because it is required for flagellar attachment. A polyclonal antibody was raised against FLA3 and the protein was observed to localise to the flagellar attachment zone. However, no concrete evidence was given as to how FLA3 might mediate flagellar attachment. What are its binding partners? Does it localise to the flagellar membrane, the cell-body membrane, or both? Is it a component of the junctional complexes that span the intermembrane region of the FAZ?

Woods et al. were hampered by the fact that their anti-FLA3 polyclonal antibody gave considerable background signal when used for immunofluorescence. Furthermore, only a small batch of the antibody was produced. The limited availability and the considerable risk of nonspecific interactions made the antibody unsuitable for high-resolution imaging and co-immunoprecipitation experiments. In this research project, it was decided to overcome these problems by using an *in-situ* tagging approach. FLA3 was tagged at the C-terminus with epitopes that could be recognised by highly specific, commercially available monoclonal antibodies (Section 3.2.3).

This approach revealed that FLA3 is anchored in the flagellar membrane (Section 3.2.4–3.2.7) and demonstrated that FLA3 has a punctate distribution suggestive of localisation to the junctional complexes (Section 3.2.8).

3.2 Results

3.2.1 Conditional Knockdown of FLA2

To assess the function of FLA2, it was first necessary to generate a FLA2 RNAi vector that would not cross-react with FLA1. The sequences of *FLA2* and *FLA1* were compared and a 414-bp region of *FLA2* bearing minimal similarity to *FLA1* (Table 3.1) was amplified from genomic DNA.

The *Bam* HI/*Xho* I-digested *FLA2* gene fragment was purified and ligated into the similarly digested RNAi vector, p2T7.177. The cloned fragment was sequenced to confirm its identity (Appendix D).

This FLA2-specific RNAi vector was subsequently transfected into bloodstream-form cells (see Methods). Three independent clones were generated. The vector inserted stably into the 177-bp repeats of the trypanosome minichromosomes, enabling tetracycline-inducible conditional knockdown of *FLA2* without concomitant knockdown of *FLA1*.

```
  1 AGACGG-CCCCAAGGGAGATGCACTATTCAACAATCCCACCTCAGTTGTA     49
    ||..|| ||...||.| |.|||.|..||.||..||||||||||.||||||.
  1 AGTTGGTCCTTTAGAG-GTTGCTCGTTTTAATCATCCCACCTCTGTTGTT     49

 50 AGCGTCAACGATGATATTTACGTGGCAGACAGGGATAACAACTGTATTAG     99
    .|.||.|||.|||||||||||||||.||||||||||..|.||.||.|.
 50 GGTGTAAACAATGATATTTACGTGGCGGACAGGGATAATGATTGCATGAA     99

100 GCGAATAGATGCTGAGGGCAAT-GTAACCAGATATG-GACCACAAGACCT    147
    |||.||||||||||.|| ||| ||.||.|..|.|| |.||| |.||..|
100 GCGGATAGATGCTGATGG-AATGGTGACGAAGTTTGCGGCCA-ATGAAGT    147

148 TAACAAACCGA--AGGATATTCTCCCATTCA-CATTGA-----ACGGGAC    189
    |.|.|||||||  ||  ||||     |||.| || |||     |||||.|
148 TGATAAACCGAGTAG--TATT-----ATTTATCA-TGAACAAGACGGGGC    189

190 TCAGAATCTATTCATTTCCGACACAGGCAATTCGCAGATCCTTTATGTGC    239
    .|.|..|.||||.|||||.||.||.||.||||||...||    ||||||
190 ACCGGTTTTATTTATTTCGGATACTGGAAATTCAAGGA----TTATGT--    233

240 CACTGGATGCTAATAACACAGTG-----------GTAA---CTACACTAG    275
     |||.|..|.||       ||||           ||||   |.|.|||.|
234 -ACTCGCAGATA-------AGTGTGTCCAACAATGTAACGGCGAAACTTG    275

276 TCGCTGGCTTCCAACCTGGTGTGATGCAAATAAGCAAAGAA-AAGAACAT    324
    |.|..||.||.|||||||||||||||||||||| ||||| ..|||..|
276 TTGGAGGTTTTCAACCTGGTGTGATGCAAATAAGC-AAGAAGCGGAATTT    324

325 GATGTATGTCGTAAAAAACACCTCATGGATTGCAGCGGTTAATTTG---A    371
    .||||||||||.||..|.|||||.||||||||||||..|..||   |
325 CATGTATGTCGTGAAGGAGACCTCGTGGATTGCAGCGGTGGACCTGAATA    374

372 GCAAAACT-GGGGAGTCTGATATTGG-AAAAAGTAA------GGATATAG    413
    ||    || .|.||.||.||   ||| ||.||| ||      |||.||.|
375 GC----CTCAGCGATTCGGA---TGGTAACAAG-AAGAGCTGGGACATCG    416

414 G    414
    .
417 C    417
```

Length:	451
Identity:	292/451 (**64.7%**)
Similarity:	292/451 (**64.7%**)
Gaps:	71/451 (**15.7%**)
Score:	825.5

Table 3.1: Pairwise alignment of the *FLA2* RNAi fragment and the corresponding region of the *FLA1* gene.

Upper row: Nucleotides 333–746 of Tb927.8.4110 (*FLA2*), which constitute the *FLA2* fragment inserted into p2T7.177 to facilitate specific RNAi of *FLA2*.

Lower row: The region of Tb927.8.4010 (*FLA1*) (nucleotides 333–749) bearing the most similarity to the above fragment of *FLA2*.

Pairwise alignment was performed using the EMBOSS Needle tool.

Noninduced clonal cells exhibited normal morphology and grew at just a slightly reduced rate compared to the parental cell line (doubling times: 10.6 h vs 8.1 h). Knockdown of FLA2 had a deleterious effect on growth and morphology. Upon induction of *FLA2* dsRNA by tetracycline, cell division ceased (Fig. 3.1) and the population density gradually declined, indicating cell death. After the 72-h timepoint, recovery of normal growth rate was consistently observed – probably due to loss of expression of *FLA2* dsRNA because of counter selection against the strong negative effect on growth.

The morphology of induced cells was observed by fluorescent labelling of the VSG surface coat (Fig. 3.2). By the 16-h timepoint, nearly all cells had at least one flagellum that had no lateral attachment to the cell body. This phenotype was not an artefact of fixation, as it could be seen easily in live cells. The detached flagella continued to beat rapidly, but directional motility was lost [video footage not shown].

After 18 hours of induction of FLA2 RNAi, a multinucleate phenotype was apparent. Cells were stained with DAPI and observed by phase contrast microscopy, then scored according to the number of nuclei and kinetoplasts they possessed (Fig. 3.3). Cells with more than two nuclei or more than two kinetoplasts were considered aberrant and were annotated as >2N. Less than 5% of noninduced cells were >2N, compared to 45% of induced cells. Knockdown of FLA2 appears to cause defects in cytokinesis.

After 24 hours of induction of FLA2 RNAi, normal morphology was lost in most cells (Fig. 3.4) Induced cells exhibited an aberrant rounded morphology with indeterminate polarity and with multiple detached flagella, indicative of cytokinesis failure. Again, this phenotype could be observed easily in live cells.

To verify that the knockdown phenotype was caused directly by loss of FLA2, qRT-PCR was performed in induced and noninduced cells (Fig. 3.5). Induced cells showed a specific reduction of *FLA2* mRNA, with no observable change in the level of control mRNA (actin, not shown). Importantly, there was no reduction of *FLA1* mRNA, confirming that the FLA2 RNAi vector was indeed specific to FLA2. Indeed, after 72 hours of induction of FLA2 RNAi, *FLA1* mRNA levels increased to nearly twice the wildtype expression level. This result suggests that *FLA1* mRNA is upregulated upon knockdown of FLA2. [The qRT-PCR measurements were performed by Derek Nolan.]

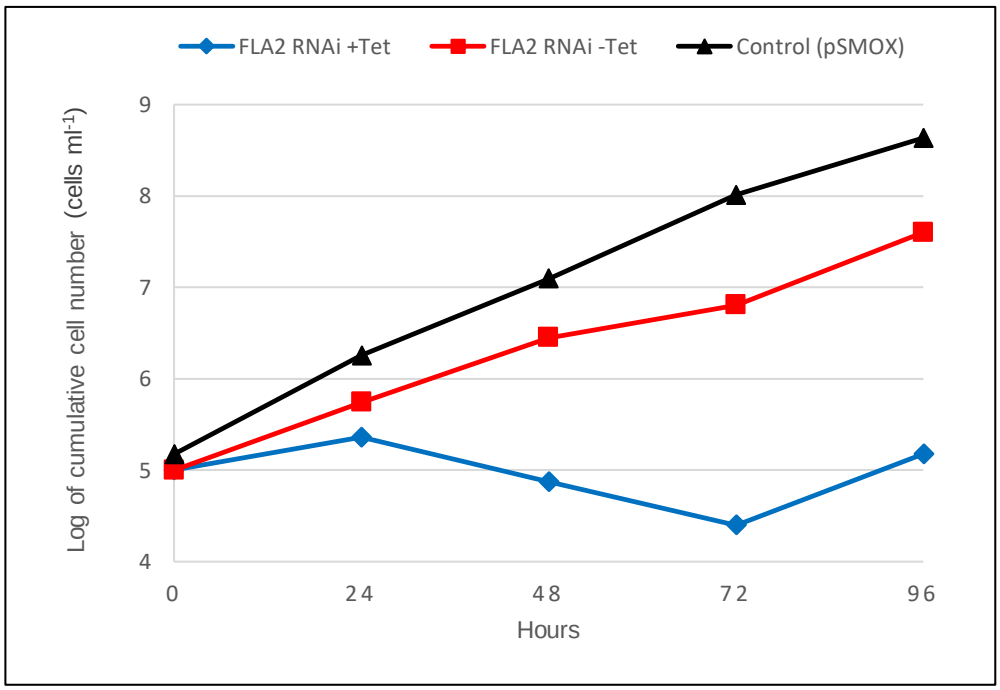


Figure 3.1: FLA2 is essential for growth in bloodstream-form cells. Noninduced FLA2-RNAi-competent BSF cells [red] grew at a similar rate (doubling time: 10.6 h) to the parental cell line [black] (doubling time: 8.1 h). Cell growth ceased upon induction of *FLA2* dsRNA by tetracycline [blue]. The data shown are the averages of two counts of a representative clone.

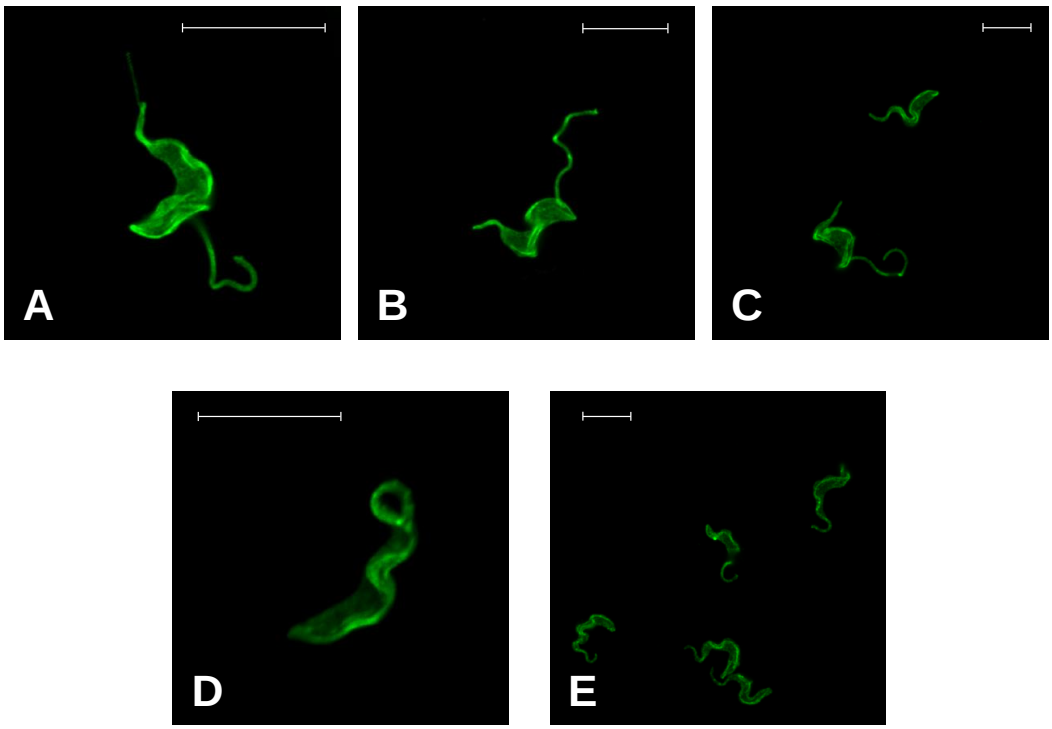


Figure 3.2: Knockdown of FLA2 causes flagellar detachment.

(A–C): Clonal FLA2-RNAi-competent bloodstream-form trypanosomes exhibit flagellar detachment after 16 h exposure to tetracycline. **(D):** The same clone has normal morphology in the absence of tetracycline. **(E):** Wildtype bloodstream-form trypanosomes have normal morphology after 16 h exposure to tetracycline. All cells probed for VSG (green) to obtain a surface view of the whole cell and flagellum. Scale bars are 10 µm.

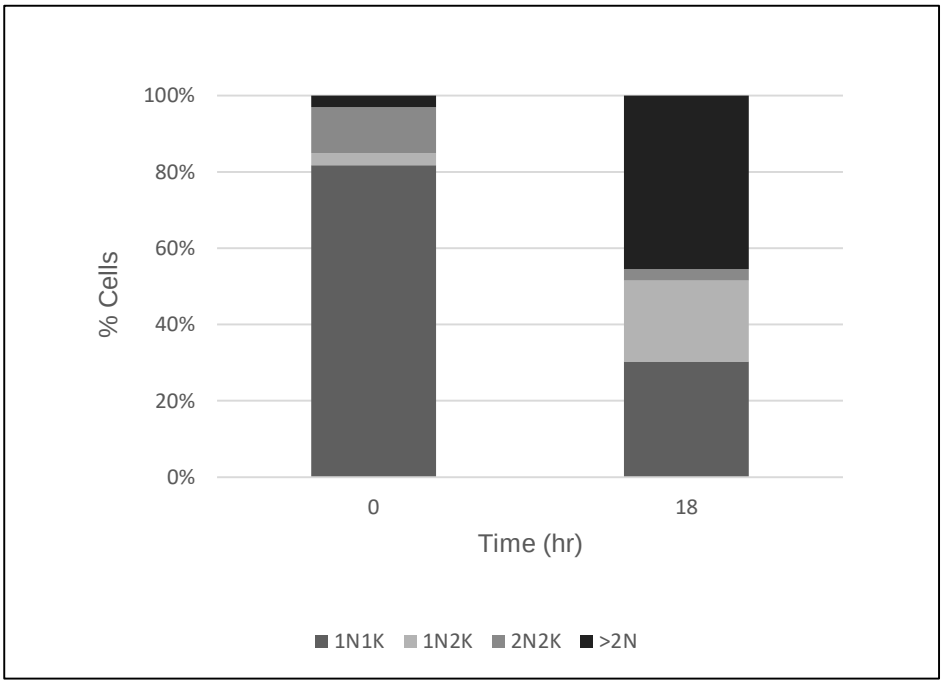


Figure 3.3: A multinucleate phenotype is observed upon knockdown of FLA2.

Analysis of the number of nuclei (N) and kinetoplasts (K) in a representative clone of FLA2-RNAi-competent BSF cells. Cells were assessed for the presence of nuclei and kinetoplasts and scored as 1N1K, 1N2K, 2N2K and cells with clearly more than two nuclei or more than two kinetoplasts (>2N). Noninduced cells (0 h) were mostly 1N1K (>80%); few cells (<5%) had an aberrant nuclear phenotype. Upon induction of FLA2 RNAi with tetracycline (18 h), nearly half the cell population (45%) had more than two nuclei or kinetoplasts, indicating aberrant cell division.

(n=33 at each timepoint)

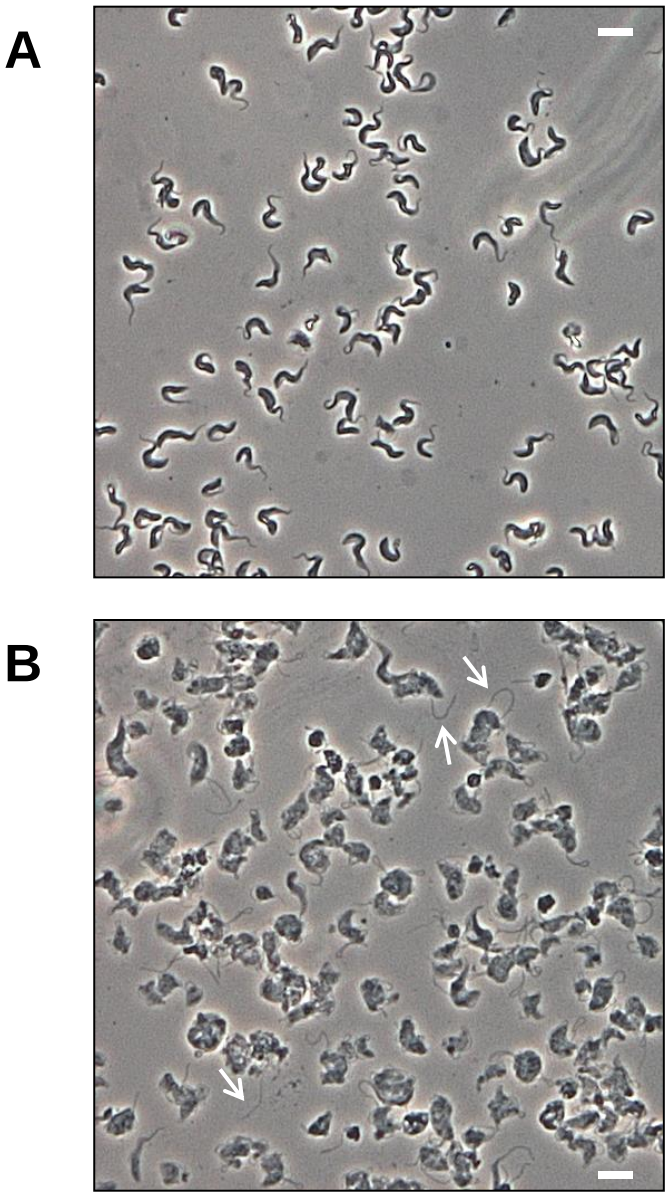


Figure 3.4: Knockdown of FLA2 causes flagellar detachment and loss of morphology.

Clonal FLA2-RNAi-competent bloodstream-form trypanosomes **(A)** exhibit classical vermiform shape with attached flagellum until knockdown of *FLA2* mRNA is induced by tetracycline **(B)**. After 24-h incubation in tetracycline, as shown in image (B), nearly all cells have at least one detached flagellum and most exhibit total loss of cell morphology.

White arrows indicate some of the detached flagella. Scale bars are 10 μm. Images obtained by phase-contrast microscopy.

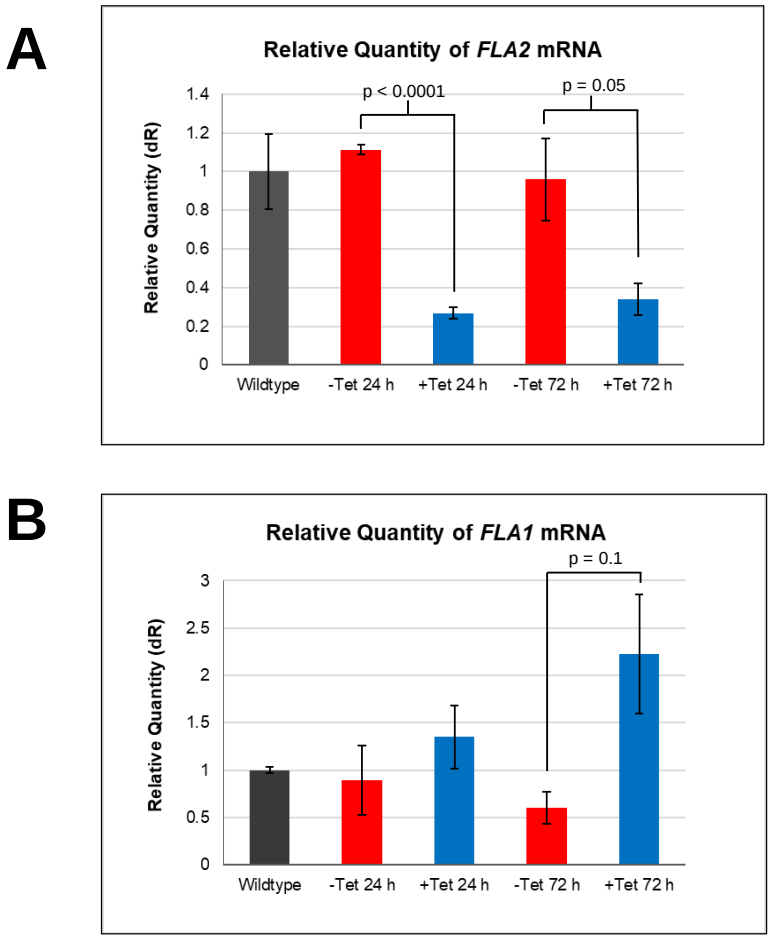


Figure 3.5: Specific reduction of *FLA2* mRNA upon induction of FLA2 RNAi.

(A) Relative levels of *FLA2* mRNA were measured by qRT-PCR in induced and noninduced FLA2-RNAi cells. After 24 h of induction with tetracycline (+Tet), *FLA2* mRNA levels were reduced to approximately one quarter of the wildtype expression level. Noninduced cells (-Tet) expressed *FLA2* mRNA at a similar level to wildtype cells.

(B) Relative levels of *FLA1* mRNA were measured by qRT-PCR in induced and noninduced FLA2-RNAi cells. After 24 h of induction with tetracycline, *FLA1* mRNA levels were slightly increased; by the 72-h timepoint, *FLA1* mRNA levels were more than twice the wildtype expression level. Noninduced cells expressed *FLA1* mRNA at a similar level to wildtype cells.

Data shown are the means of triplicates from a single experiment. Error bars show the standard error of the mean. The p-values were obtained using a two-tailed t-test.

qRT-PCR experiments were performed by Derek Nolan.

These results demonstrate that FLA2 is an essential protein in the bloodstream-form trypanosome. Given the rapid onset of the flagellar-detachment phenotype, it appears that the immediate effect of loss of FLA2 is the prevention of attachment of the newly synthesised flagellum. This flagellar detachment causes subsequent failure of cytokinesis.

3.2.2 Conditional Knockdown of FLA3

To likewise confirm the essentiality of FLA3 in the bloodstream form, the protein was knocked down by conditional RNAi using the FLA3-p2T7.177 construct in a bloodstream-form pSMOX cell line (see Methods). Three independent clones were generated.

Noninduced cells exhibited normal morphology. They grew at a slightly reduced rate compared to the parental cell line (doubling times: 13.8 h vs 9.8 h). Knockdown of FLA3 had a deleterious effect on growth and morphology. Upon induction of *FLA3* dsRNA by tetracycline, cell division ceased (Fig. 3.6) and the population density declined slightly over the first 72 hours. Between the 72-h and 96-h timepoints, recovery of normal growth rate was consistently observed – probably due to loss of expression of *FLA3* dsRNA because of counter selection against the strong negative effect on growth.

The morphology of induced cells was observed by phase-contrast microscopy (Fig. 3.7). It was essentially identical to that of induced FLA2-RNAi cells. By the 18-h timepoint, nearly all cells had at least one flagellum that had no lateral attachment to the cell body. This phenotype could be seen even in live cells. The detached flagella continued to beat rapidly, but directional motility was lost.

After 24–48 hours of induction of FLA3 RNAi, normal morphology was lost in most cells. Induced cells exhibited an aberrant rounded morphology with indeterminate polarity and with multiple detached flagella, indicative of cytokinesis failure. Again, this phenotype could be observed easily in live cells.

(Loss of FLA3 at the protein level upon induction of RNAi was confirmed later by western blotting, see Section 3.2.6.)

These results demonstrate that FLA3 is necessary for flagellar attachment in the bloodstream-form trypanosome and, probably by consequence, is essential for cytokinesis.

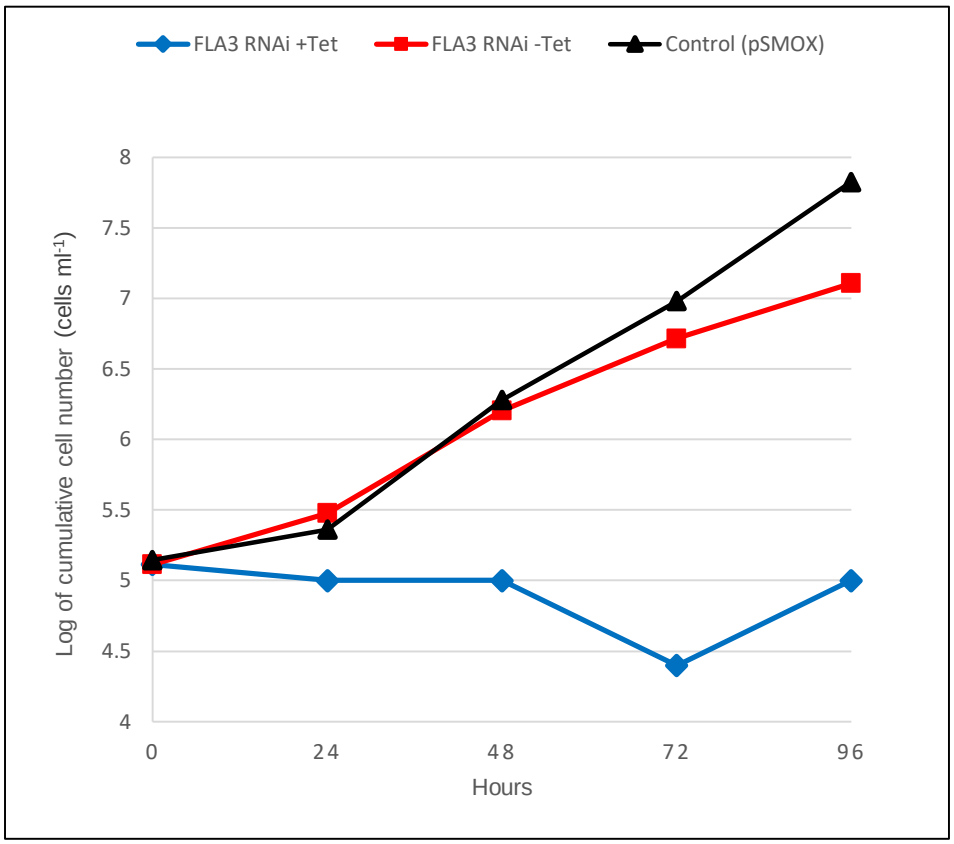


Figure 3.6: FLA3 is essential for growth in bloodstream-form cells. Noninduced FLA3-RNAi-competent BSF cells [red] grew at a slightly reduced rate (doubling time: 13.8 hr) to the parental cell line [black] (doubling time: 9.8 hr). Cell growth ceased upon induction of *FLA3* dsRNA by tetracycline [blue]. The data shown are the averages of two counts of a representative clone.

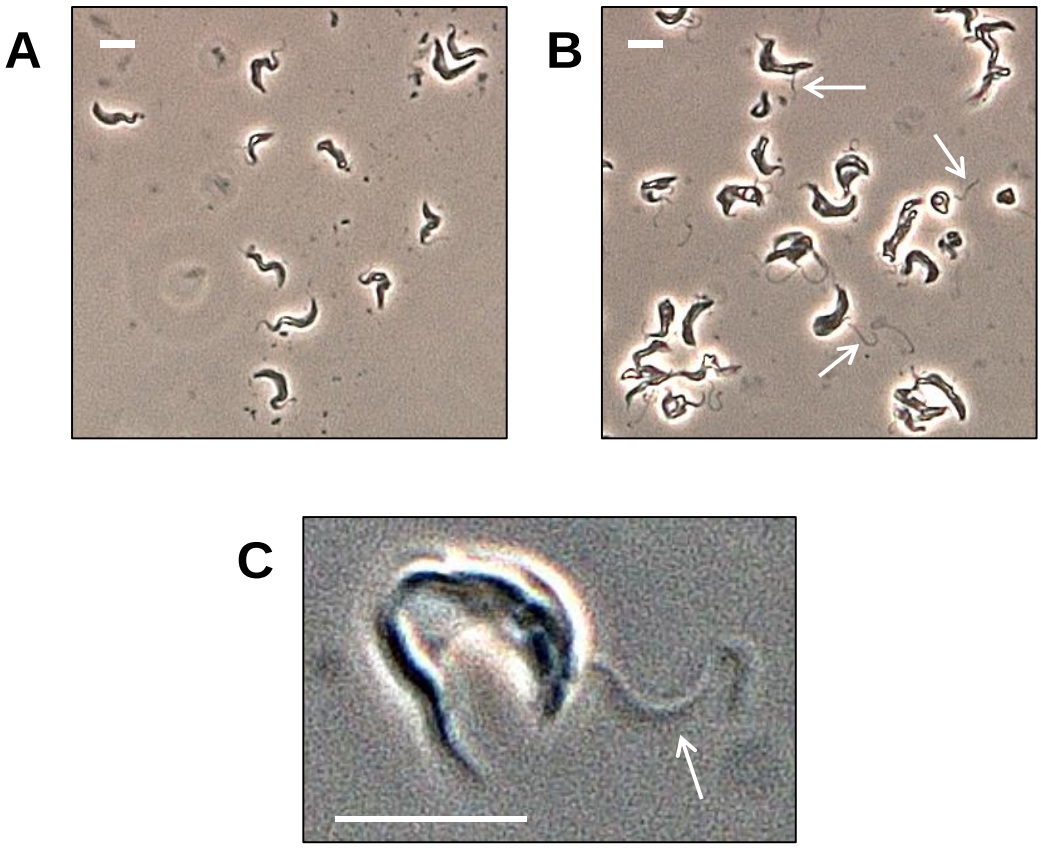


Figure 3.7: Knockdown of FLA3 causes flagellar detachment and loss of morphology.

Clonal FLA3-RNAi-competent bloodstream-form trypanosomes **(A)** exhibit classical vermiform shape with attached flagellum until knockdown of *FLA3* mRNA is induced by tetracycline **(B, C)**. After 18-hr incubation in tetracycline, as shown in images B and C, nearly all cells have at least one detached flagellum and some exhibit total loss of cell morphology.

White arrows indicate some of the detached flagella. Images obtained by phase-contrast microscopy. Scale bars are 10 µm.

3.2.3 *In-situ* tagging of FLA3 with Ty1 and HA tags

To enable analysis of FLA3 at the protein level, it was necessary to add an epitope tag to the protein so that it could be recognised specifically on blots and on fixed cell samples. The *FLA3* gene was subjected to *in-situ* tagging to allow for expression of a FLA3 protein tagged at the C-terminus with a Ty1 or HA tag. This approach inserted the epitope tag directly into the target gene at the endogenous locus by homologous recombination (see Section 2.2.7), thus avoiding the need for ectopic overexpression of the tagged protein. Following transfection with the p3074 (Ty1) or p2708 (HA) tagging vector, transformants were selected using G418, and individual clones were expanded. Two clones expressing FLA3-Ty1 were generated; they performed identically when tested. Only one clone expressing FLA3-HA was generated.

Cells expressing FLA3-Ty1 grew at the same rate as wildtype cells (doubling time: ~7.5 h) whereas cells expressing FLA3-HA grew at a slightly reduced rate (doubling time: ~9 h) (Fig. 3.8). There were no observable differences in morphology, motility or flagellar attachment in tagged cells compared to wildtype cells.

Samples of these cells were probed by western blot, using monoclonal antibodies against Ty1 or HA (Fig. 3.9), to assess expression of tagged FLA3 at the protein level. In each case, a single band was observed at ~160 kDa.

Crucially, the apparent size of the tagged FLA3 (~160 kDa) was substantially higher than the molecular weight predicted by the primary sequence (89 kDa). This result was consistent with extensive *N*-glycosylation of the protein, as observed for wildtype FLA3.

Taken together, these results show that C-terminal tagging of FLA3 did not have any deleterious effect on the post-translational modification of the protein or on cell growth.

The FLA3-Ty1 band seemed more intense than the corresponding FLA3-HA band. Although this difference in intensity could be explained by a difference in expression between the two tagged forms, it seems more likely to be related to the relative detection sensitivities of the Ty1 and HA antibodies.

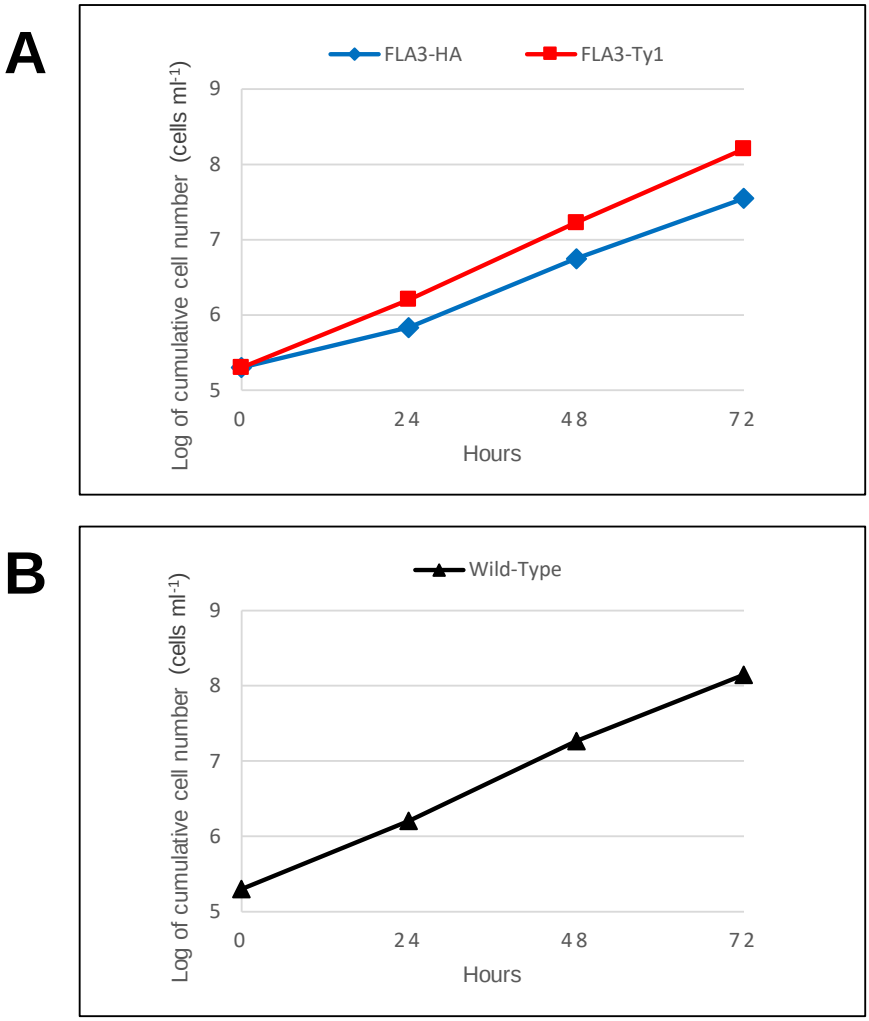


Figure 3.8: Epitope tagging of endogenous FLA3 has little impact on cell growth.

Bloodstream-form cells expressing Ty1-tagged FLA3 [red] **(A)** proliferate at the same rate (doubling time: 7.4 h) as untransfected wild-type cells [black] **(B)** (doubling time: 7.6 h). Cells expressing HA-tagged FLA3 [blue] **(A)** proliferate at a slightly reduced rate (doubling time: 9.3 h).

The data shown are the averages of two counts of representative clones.

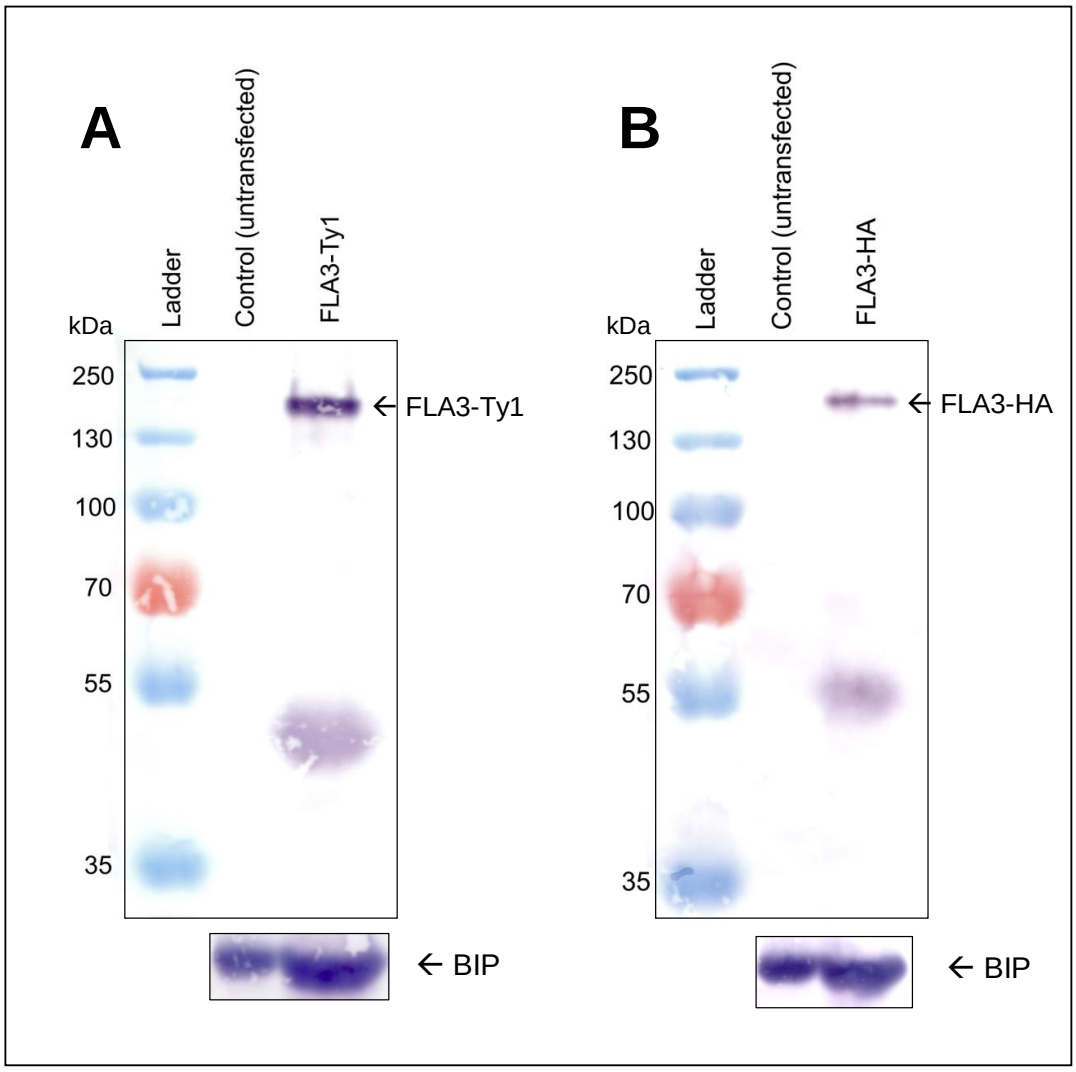


Figure 3.9: Expression of tagged FLA3 in bloodstream-form cells.

(A) Western blot probed for Ty1 confirms that the bloodstream-form cell line that has undergone *in-situ* tagging with p3074 expresses FLA3-Ty1 at the protein level. **(B)** Western blot probed for HA confirms that the bloodstream-form cell line that has undergone in-situ tagging with p2708 expresses FLA3-HA at the protein level. Probing for a chaperone protein, BIP, indicates that the control and test lanes had comparable protein levels.

Interestingly, a diffuse band at ~50 kDa on the FLA3-Ty1 blot (Fig. 3.9 A) and at ~55 kDa on the FLA3-HA blot (Fig. 3.9 B) was reproducibly detected in all tagged cell lines and neither band was detected in the parental cell line employed for transfection. These bands could represent breakdown products of FLA3 – in fact, it is highly probable that they each represent the same fragment (discussed in Section 3.3.9). The apparent size difference between the two bands (~5 kDa) can be attributed simply to the corresponding size difference between the HA and Ty1 tags (8.5 kDa). Since this putative breakdown product retains the epitope tag, it must be derived from the C-terminal end of FLA3.

3.2.4 Localisation of tagged FLA3 to the membrane system

The subcellular distribution of tagged FLA3 was investigated in cell-fractionation experiments (Fig. 3.10).

Upon solubilisation of the cell membrane with Triton X-100 detergent, tagged FLA3 was found almost exclusively in the supernatant fraction. Upon freeze-thaw fractionation, tagged FLA3 was found almost exclusively in the pellet fraction. Taken together, these results indicated that the C-terminal epitope tags did not affect correct localisation of FLA3 to the membrane system.

Thus, FLA3-Ty1 and FLA3-HA both appear to act as integral membrane proteins, in the same manner as wildtype FLA3. Curiously, the putative breakdown product at ~50 kDa appears to be distributed in the same manner, presumably because it retains the C-terminal transmembrane domain.

3.2.5 Immunolocalisation of tagged FLA3 to the FAZ

The localisation of FLA3-Ty1 was investigated by immunofluorescence in fixed cells. Different fixation conditions were tested to optimise visualisation of FLA3-Ty1 (Fig. 3.11). Fixation with 3% paraformaldehyde (PFA) was the most successful of the methods attempted, so it was used for all subsequent localisation studies. FLA3-Ty1 signal was also clearly visible upon sequential fixation with PFA and methanol (see Section 2.4.3), but cell morphology was damaged. Fixation with methanol only (see Section 2.4.3) caused excessive damage to cell morphology under the conditions tested, and only background signal was observed (data not shown).

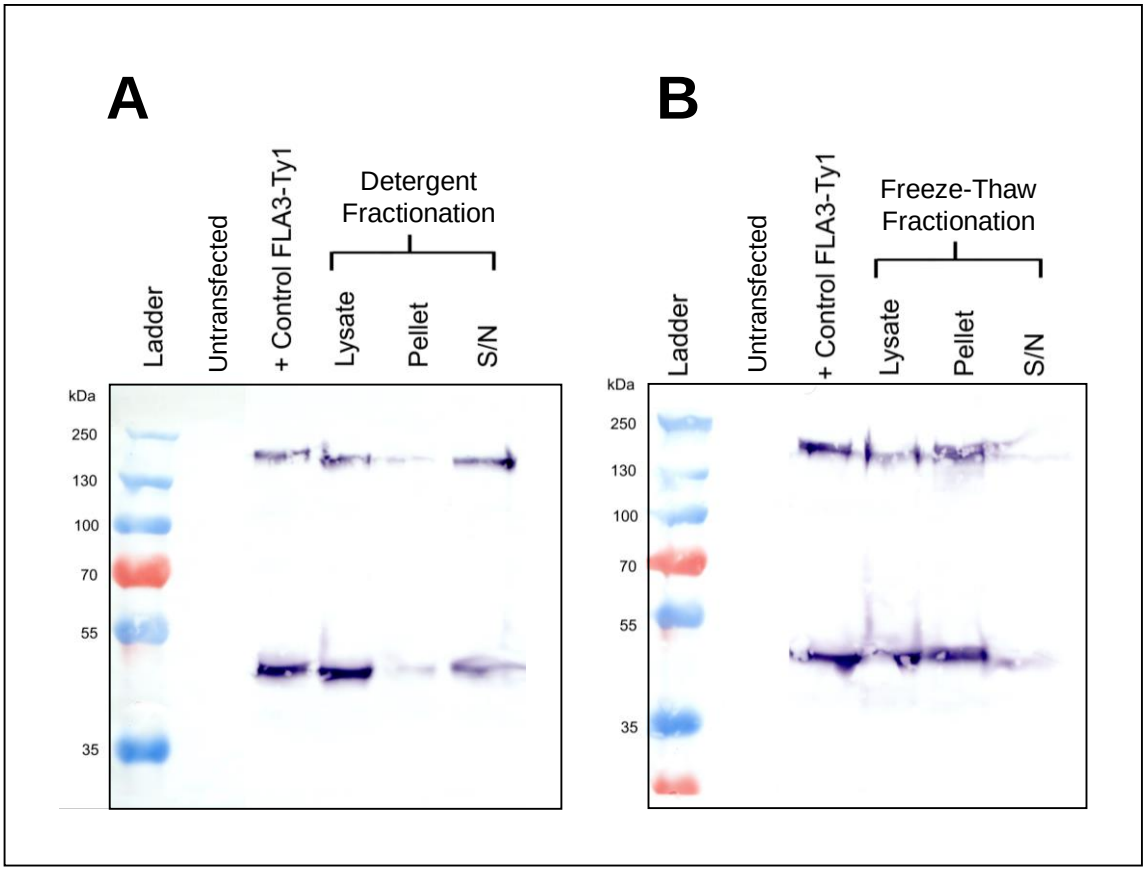


Figure 3.10: Distribution of FLA3-Ty1 after cell fractionation.

(A) Western blot probed for Ty1 confirms location of FLA3-Ty1 in the supernatant (S/N) fraction after detergent fractionation, showing that the protein is solubilised. **(B)** Upon freeze-thaw fractionation, the protein is found in the pellet fraction, because it remains anchored in the membrane system. Similar results were obtained for FLA3-HA [not shown].

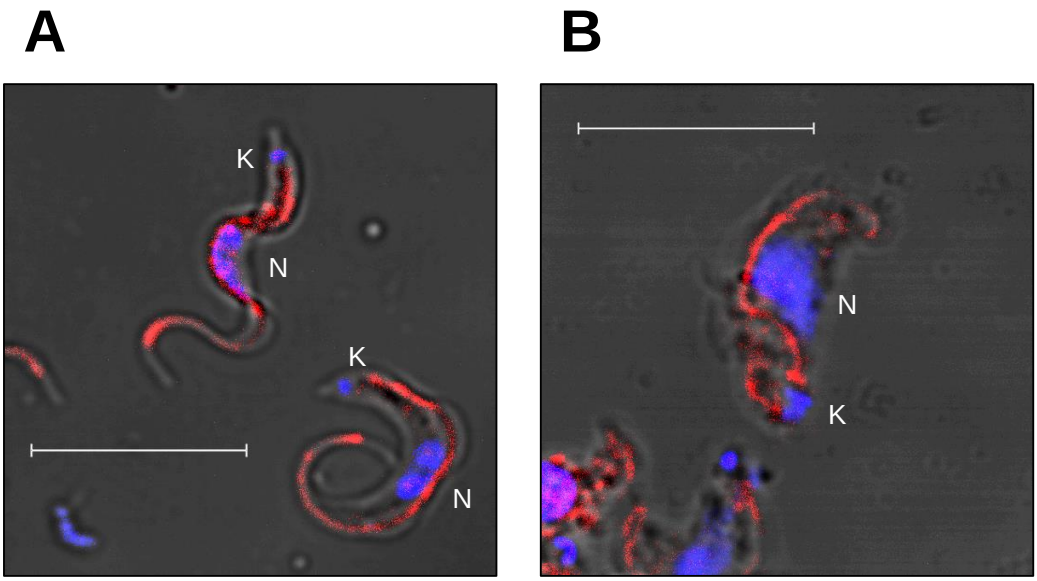


Figure 3.11: Fixation with PFA produces optimal conditions for imaging FLA3-Ty1.

Cells retain their morphology better upon fixation with PFA only **(A)** than upon fixation with PFA and methanol **(B)** or with methanol only (not shown).

FLA3-Ty1 signal is shown in red. Nucleus (N) and kinetoplast (K) are stained in blue. Scale bars are 10 µm.

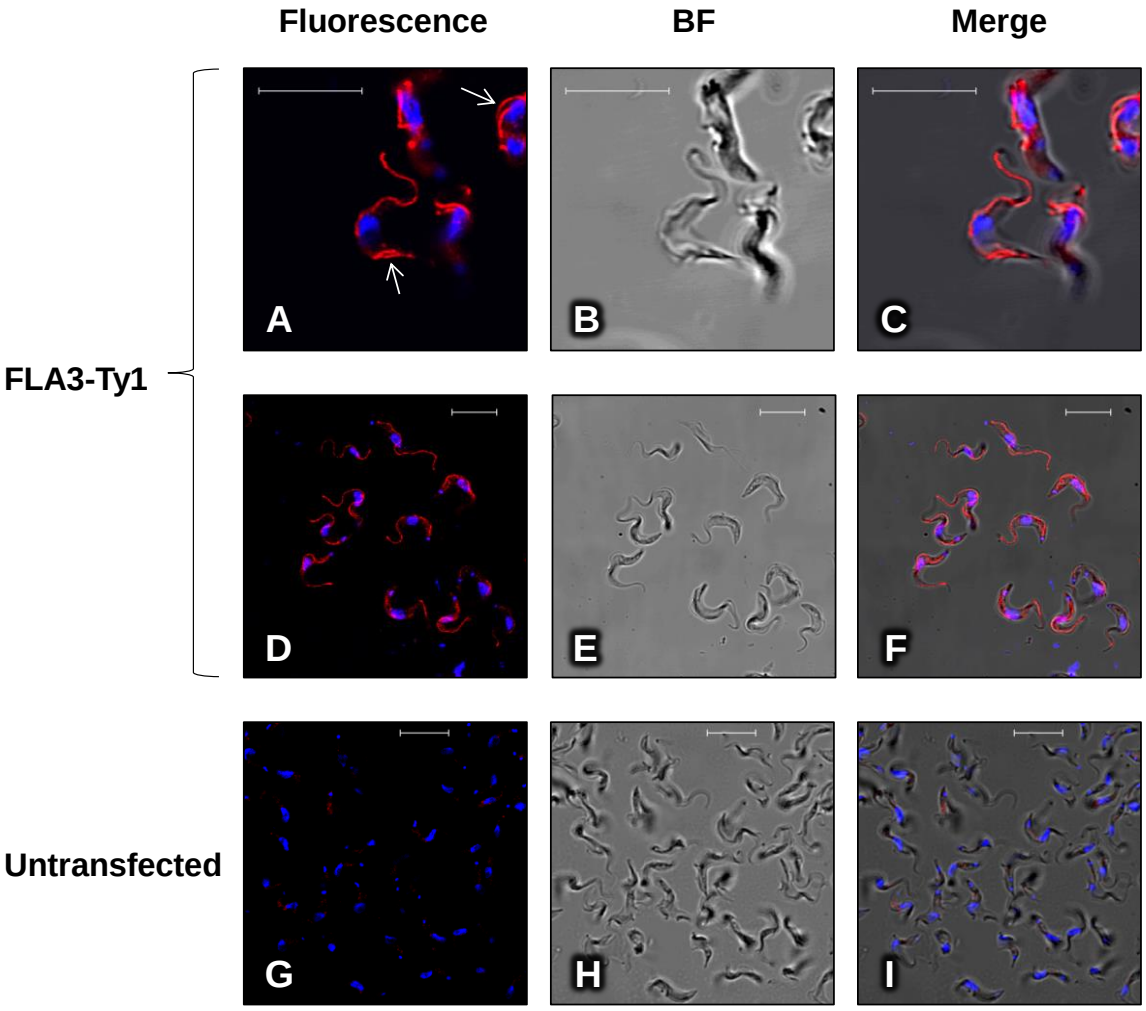


Figure 3.12: L**ocalisation of Ty1-tagged FLA3 to the flagellar attachment zone**.

(A–F) Indirect immunofluorescence shows that clonal bloodstream-form cells transfected with p3074 express FLA3-Ty1 (red), and that the tagged protein localises to the flagellar attachment zone. It is possible to see the growth of new flagella alongside the old flagella in dividing cells (white arrows).

(G–I) Untransfected control cells do not show any fluorescence signal when probed for Ty1.

Nucleus and kinetoplast are stained in blue. All scale bars are 10 μm.

When immunofluorescence was performed in cells expressing FLA3-Ty1, the anti-Ty1 antibody gave a strong signal along the FAZ, with a very low background signal throughout the cell (Fig. 3.12). This result suggests that Ty1-tagged FLA3 behaves in the same manner as the wildtype protein. Indeed, the images obtained here were much clearer than the published images of wildtype FLA3 (Woods et al., 2013), which were obtained using a polyclonal antibody that produced considerable background signal due to nonspecific binding.

The demonstration that Ty1-tagged FLA3 localised correctly indicated that the FLA3-Ty1 cell line was a suitable system in which to carry out more extensive investigations into the localisation and function of the protein.

In cells that were actively dividing, FLA3-Ty1 signal was visible along both the old and the new FAZ (Fig. 3.12 A). There often appeared to be increased intensity of signal at either extremity of the FAZ. Furthermore, it was evident that the FLA3-Ty1 signal was totally absent from the stretch of free flagellum at the anterior end of the trypanosome cell (Fig. 3.13).

To address the question of whether FLA3 distribution along the FAZ is continuous or discontinuous, the images of FLA3-Ty1 immunofluorescence in PFA-fixed cells were examined closely. It was seen that in some images, regions of the FLA3-Ty1 signal appeared as regularly spaced punctate spots (Fig. 3.14), rather than an unbroken line along the FAZ. These findings indicate that the true distribution of FLA3 might be punctate, suggestive of an association with the junctional complexes that appear to act as molecular rivets bridging the intermembrane space (Vickerman, 1969, Sunter and Gull, 2016). This distribution pattern was postulated by Woods et al. but could not be demonstrated clearly because of the high background signal from the anti-FLA3 antibody.

It was not possible to form an unequivocal conclusion as to whether FLA3-Ty1 distribution was punctate or continuous owing to the diffraction limit imposed by ordinary confocal fluorescence microscopy. This issue was subsequently overcome using super-resolution microscopy (Section 3.2.8).

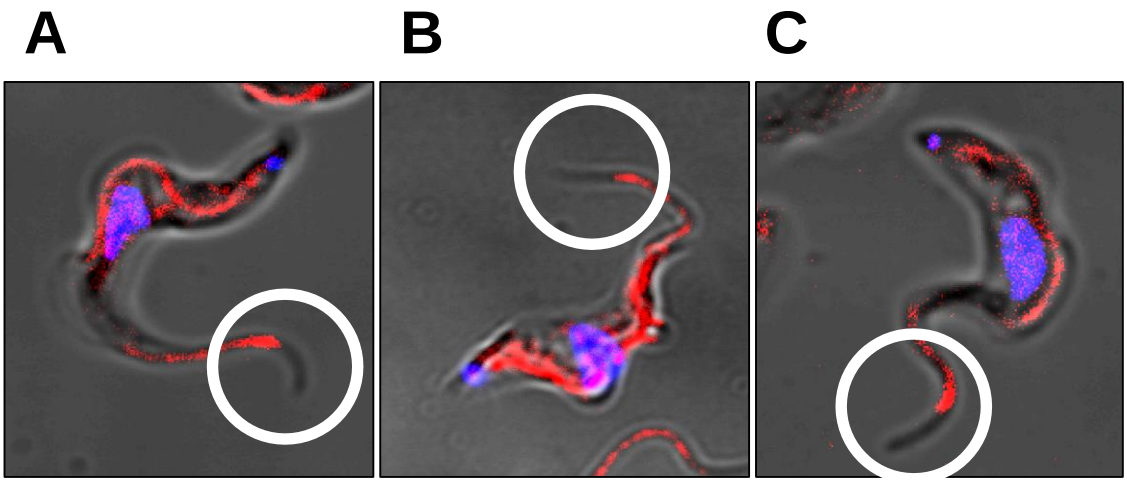


Figure 3.13: FLA3 is absent from the free flagellum.

(A, B, C) Close-up confocal images indicate that the FLA3-Ty1 signal stops abruptly at the distal end of the flagellar-attachment zone, and does not continue on into the stretch of free flagellum at the anterior tip of the trypanosome.

FLA3-Ty1 signal is shown in red. Nucleus and kinetoplast are stained in blue. The free flagellum visible as dark grey line extending beyond the FAZ.

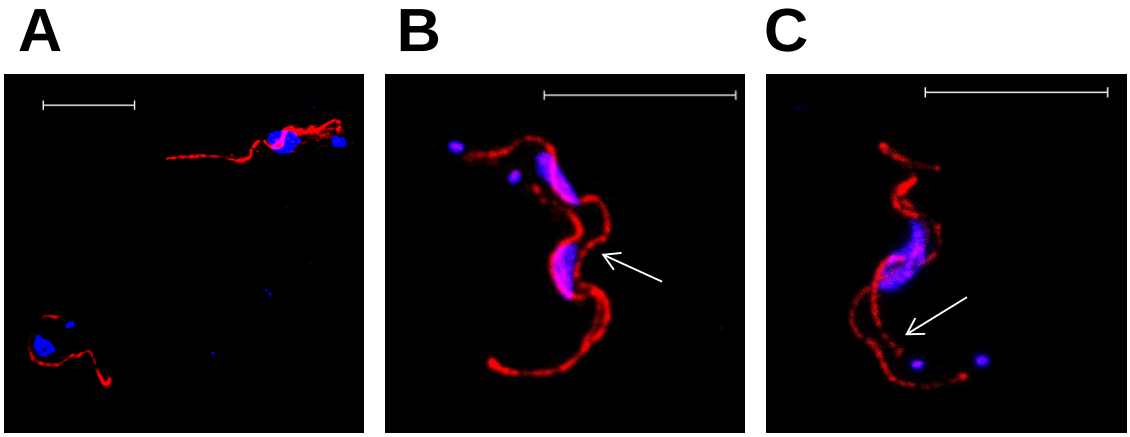


Figure 3.14: FLA3-Ty1 may have a punctate distribution along the FAZ.

(A, B, C) When imaged by confocal microscopy, FLA3-Ty1 signal usually seems to form a continuous line along the FAZ. However, in some images, FLA3-Ty1 signal seems punctate [white arrows]. A higher-resolution imaging technique is necessary to determine the true distribution of FLA3-Ty1.

FLA3-Ty1 signal is shown in red. Nucleus and kinetoplast are stained in blue. Scale bars are 10 μm.

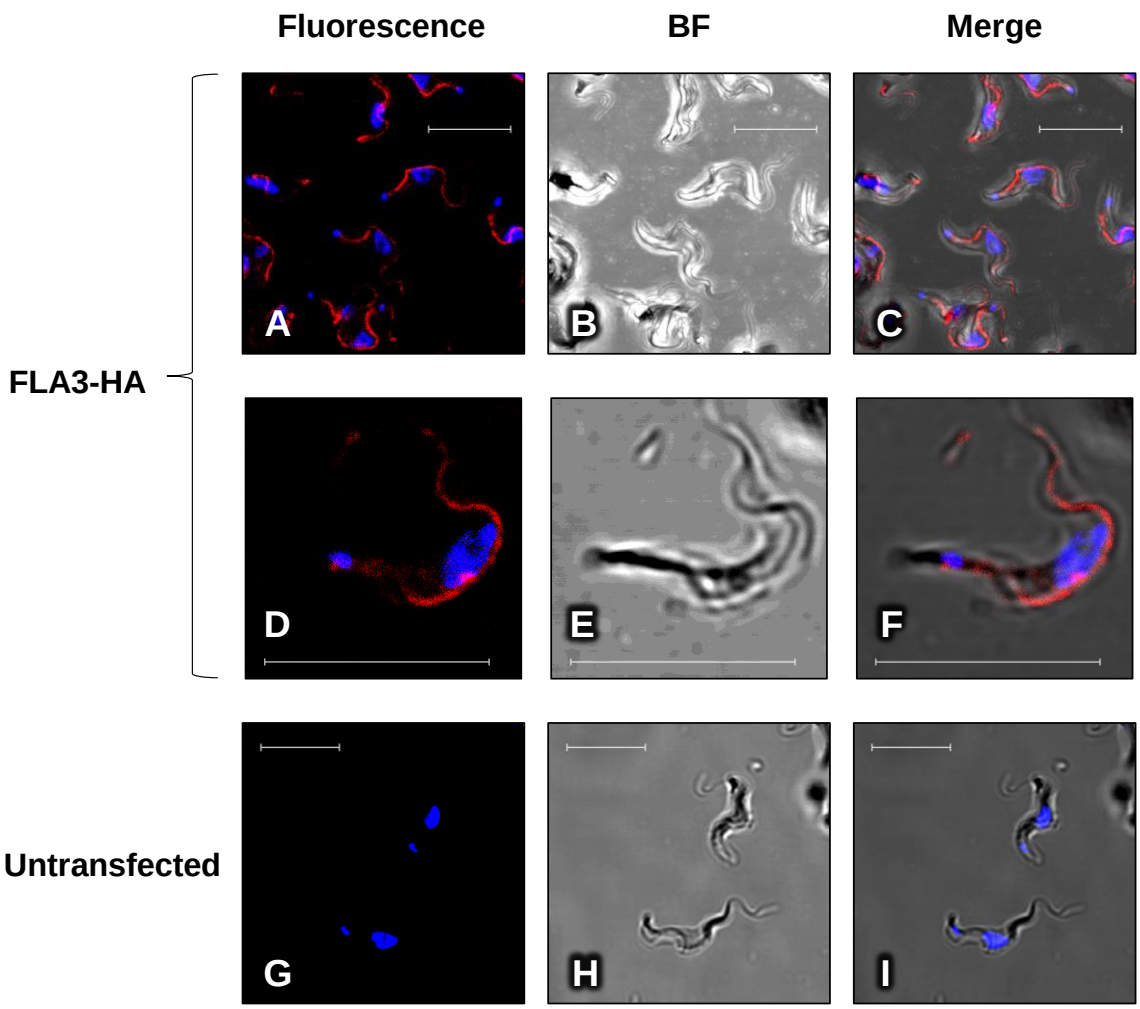


Figure 3.15: Localisation of HA-tagged FLA3 in bloodstream-form cells.

(A–F) Indirect immunofluorescence shows that clonal bloodstream-form cells transfected with p2708 express FLA3-HA (red), and that the tagged protein localises to the flagellar attachment zone.

(G–I) Untransfected control cells do not show any fluorescence when probed for HA.

Nucleus and kinetoplast are stained in blue. All scale bars are 10 μm.

The localisation of FLA3-HA was likewise investigated in PFA-fixed cells by immunofluorescence. The anti-HA antibody gave a strong signal along the FAZ, with a very low background signal throughout the cell (Fig. 3.15), demonstrating that FLA3-HA had the same localisation as FLA3-Ty1. The images obtained were slightly lower in quality than those of FLA3-Ty1, so all further localisation studies were performed using a FLA3-Ty1 clone. However, it is probable that if probing conditions were optimised, the FLA3-HA cell line would also be suitable for further studies, as it was clear that the HA tag did not interfere with normal localisation of FLA3.

3.2.6 Conditional RNAi of FLA3 in cells expressing epitope-tagged FLA3

To confirm beyond question that the correct protein had been epitope tagged for the results shown in Fig. 3.8–3.15, the pSMOX vector and FLA3 RNAi vector were transfected into the FLA3-Ty1 and FLA3-HA cell lines. Non-induced cells had doubling times of 7.2 h and 6.5 h, respectively. Induction of RNAi in these cell lines produced the same cessation of cell growth as in the untagged line (Fig. 3.16), and the morphological defects were the same. Even after just 24 h, most of the cells had detached flagella, and there were many monster cells that had failed to complete cytokinesis. The proportion of abnormal or dead cells increased over the following days, although cell growth again showed signs of recovery by the 96-h timepoint, with outgrowth of cells of normal morphology.

Samples of these tagged-FLA3 RNAi cells were probed for Ty1 and HA, respectively, by western blot (Fig. 3.17). In each case, there was a clear decrease in FLA3 signal in the knockdown cells upon addition of tetracycline, which demonstrated that ablation of the tagged FLA3 protein had occurred successfully. This result confirmed that the ~160-kDa band seen in all clones represents FLA3.

Interestingly, the level of expression of the ~50 kDa FLA3 degradation product remained constant during the knockdown.

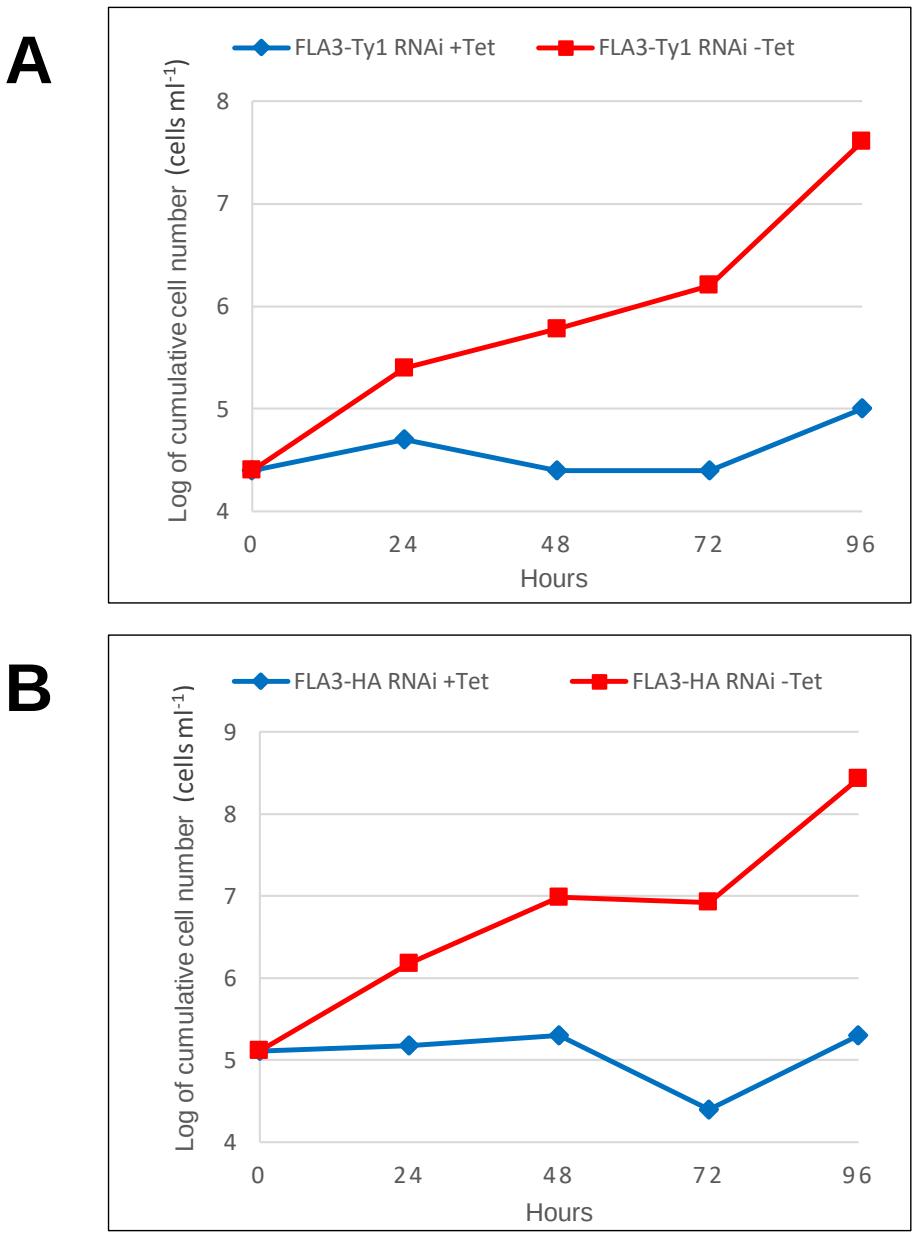


Figure 3.16: Knockdown of FLA3 in epitope-tagged cell lines.

Growth curves showing cell death following tetracycline-induced knockdown of *FLA3* mRNA in clonal RNAi-competent bloodstream-form trypanosomes in which FLA3 was tagged at the C-terminus with Ty1 **(A)** or with HA **(B)**.

The data shown are the averages of two counts of representative clones.

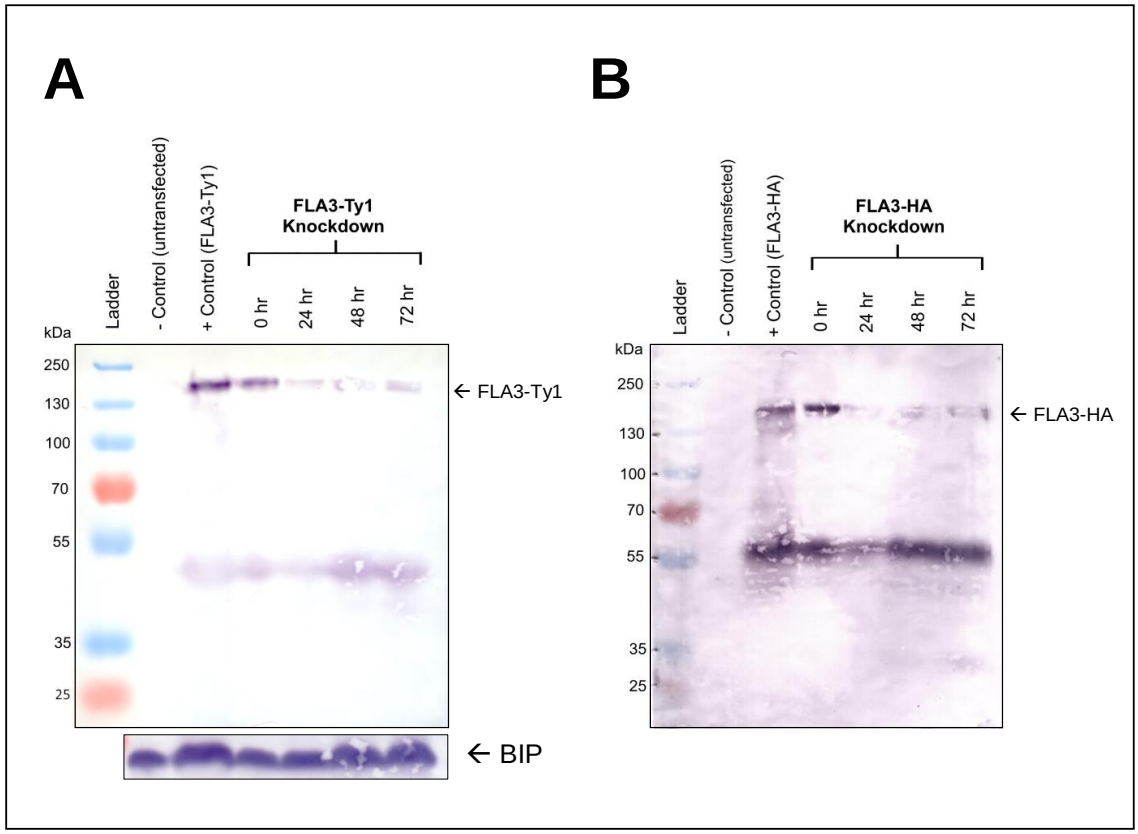


Figure 3.17: Knockdown of tagged FLA3.

(A): Western blot probed for Ty1 shows knockdown of FLA3-Ty1 in RNAi-competent bloodstream-form trypanosomes upon addition of tetracycline. Loading control (probing for BIP) shows that loading was consistent (as do the degradation-product bands at ~50 kDa).

(B): Western blot probed for HA shows knockdown of FLA3-HA in RNAi-competent bloodstream-form trypanosomes upon addition of tetracycline. FLA3 degradation-product bands at ~55 kDa show that loading was consistent.

The positive controls are the parental cell lines, which are not RNAi-competent.

3.2.7 Conditional knockdown of FLA2 in cells expressing Ty1-tagged FLA3

A key question of this thesis was whether FLA3 locates to the flagellar or cell-body side of the FAZ, as this has never been shown conclusively (discussed in Section 3.3.6).

This issue was investigated by localisation of FLA3 in cells where the flagellum was detached from the cell body. Knockdown of FLA2 is an effective method of inducing flagellar detachment (see Section 3.2.1). The pSMOX vector and the FLA2-RNAi vector (p2T7.177) were therefore transfected into a cloned FLA3-Ty1 cell line to produce a clone that expressed tagged FLA3 and was capable of FLA2 RNAi.

Induction of *FLA2* dsRNA by tetracycline produced the same phenotype observed previously. Cell growth ceased (Fig. 3.18) and cells developed an abnormal morphology, with detached flagella. Western blotting confirmed that these cells expressed FLA3-Ty1 at the protein level (Fig. 3.19). Expression of *FLA2* dsRNA did not reduce the expression level of FLA3-Ty1 (indeed, it seemed that FLA3 may even have been upregulated slightly when FLA2 was knocked down) (Fig. 3.19). Taken together, these results suggested that this cell line was a suitable system for investigation of the location of FLA3 upon detachment on the flagellum.

Immunofluorescence was performed to probe for FLA3-Ty1 in these cells. Noninduced cells had normal morphology (Fig. 3.20 A–C). As before, FLA3-Ty1 signal was seen clearly along the FAZ, with very little background signal throughout the cell body.

After 18-h induction of FLA2 RNAi, flagellar detachment was observed in most cells (Fig. 3.20 D–I). Crucially, the FLA3-Ty1 signal was visible along the detached new flagellum. This is a critically important finding. It supports the hypothesis that FLA3 is anchored in the flagellar membrane rather than in the cell-body membrane.

3.2.8 Super-Resolution Localisation of FLA3-Ty1

Another important question regarding localisation of FLA3 was whether the protein is distributed in a continuous line along the FAZ or is rather a component of the punctate junctional complexes. To address this question, cells expressing FLA3-Ty1 were imaged using stimulated emission depletion (STED) microscopy. This technique bypasses the diffraction limit of ordinary light microscopy, facilitating super-resolution imaging.

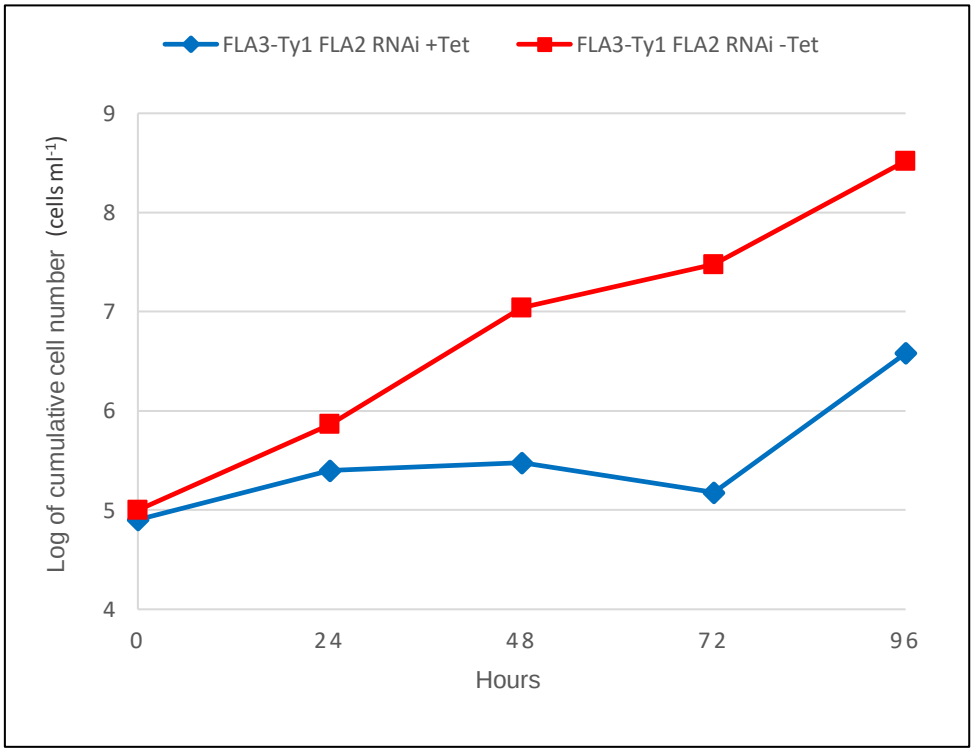


Figure 3.18: Knockdown of FLA2 in cells expressing FLA3-Ty1.

Growth curve showing cell death following tetracycline-induced knockdown of *FLA2* mRNA in RNAi-competent bloodstream-form trypanosomes that express Ty1-tagged FLA3. This confirms that expression of FLA3-Ty1 did not change the phenotypic response to FLA2 RNAi.

The data shown are the averages of two counts of a representative clone.

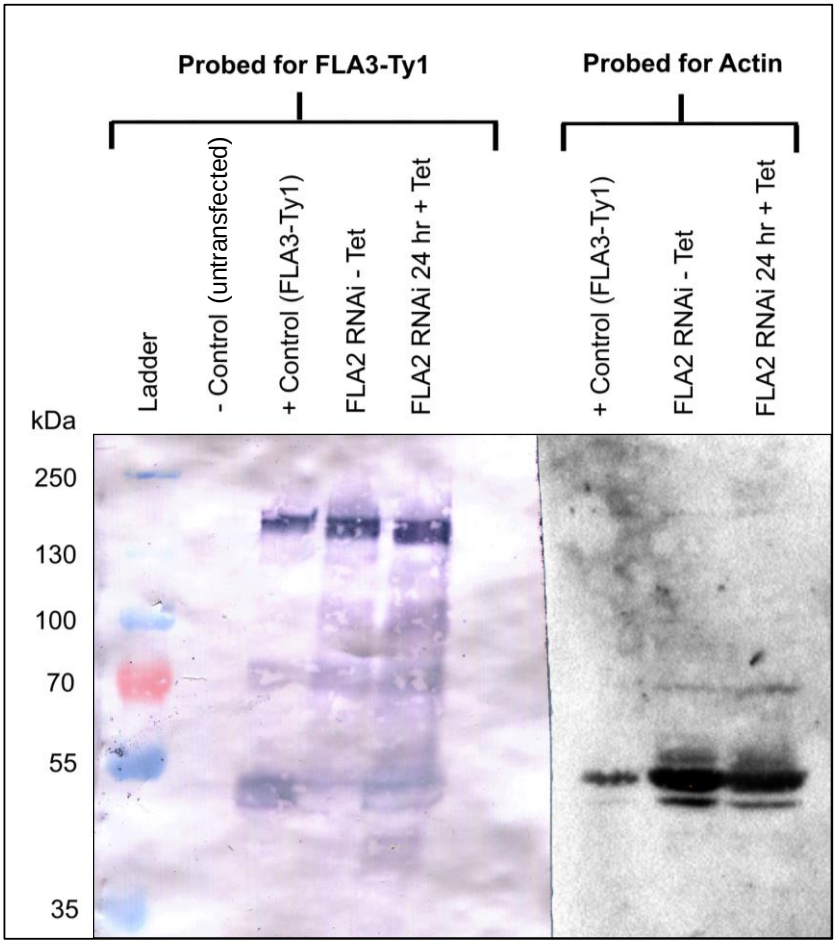


Figure 3.19: Expression of tagged FLA3 in FLA2 RNAi cells.

Western blot probed for Ty1 [left] confirms that FLA2 RNAi-competent bloodstream-form trypanosomes transfected with the p3074 vector express FLA3-Ty1 at the expected molecular weight (~160 kDa). The FLA3-Ty1 bands seem stronger than in the positive control (FLA3-Ty1 cells that are not RNAi-competent), but actin probing [right] shows that this is due merely to inconsistent loading.

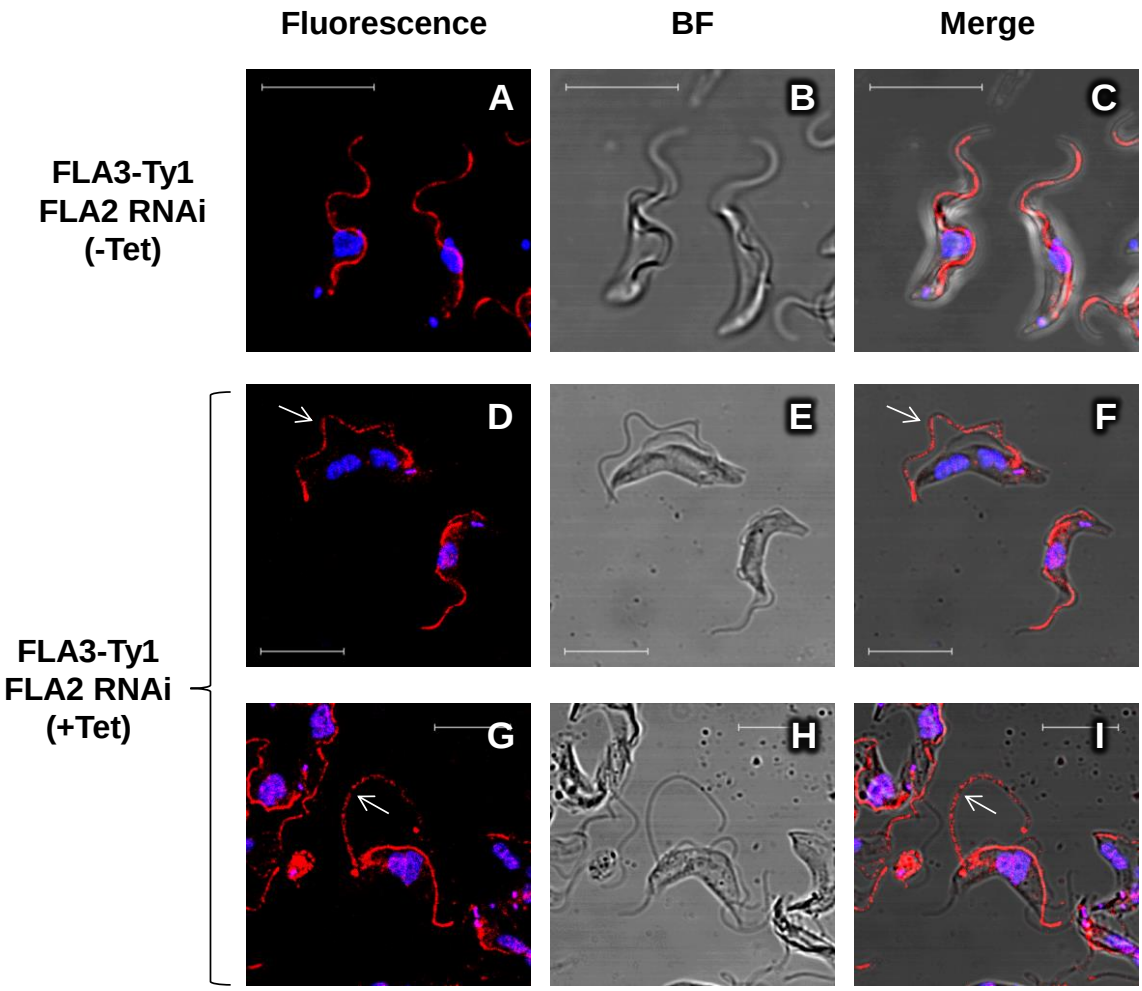


Figure 3.20: Localisation of FLA3-Ty1 in FLA2-RNAi-competent cells.

(A–C): FLA2-RNAi-competent bloodstream-form cells transfected with p3074 express FLA3-Ty1 (red). The tagged protein localises to the flagellar attachment zone (FAZ).

(D–I): Upon induction of FLA2 RNAi (18 h tetracycline) in the same clone, the flagellum detaches and some loss of morphology is evident. Again, FLA3-Ty1 (red) seems to localise to the FAZ. White arrows show localisation of FLA3-Ty1 along the detached flagellum.

Nucleus and kinetoplast are stained in blue. All scale bars are 10 µm.

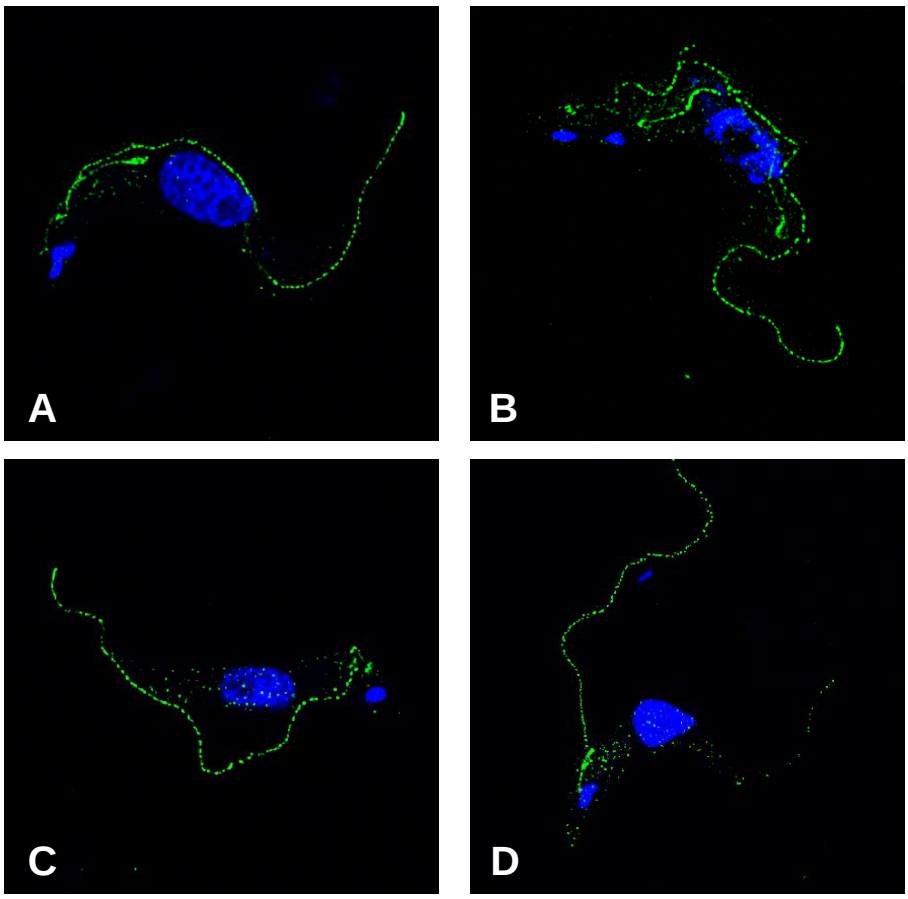


Figure 3.21: Super-resolution imaging demonstrates that FLA3-Ty1 has a punctate distribution along the FAZ.

(A–D) Upon deconvolution, STED super-resolution images of cells expressing FLA3-Ty1 (green) show that the tagged protein localises in a distinctly punctate line along the flagellar attachment zone.

(C, D) In cases where flagella became completely or partially detached because of “leaky” expression of *FLA2* dsRNA by the p2T7.177 vector, the punctate FLA3-Ty1 signal is found on the detached flagellum rather than on the cell body. Nucleus and kinetoplast are stained in blue.

The cells were fixed with PFA and probed for Ty1 by indirect immunofluorescence, then viewed by STED (see Methods). Upon deconvolution, the STED images show that FLA3-Ty1 localises in a distinctly punctate line along the flagellar attachment zone (Fig. 3.21).

In cases where flagella became completely or partially detached because of leaky expression of *FLA2* dsRNA by the p2T7.177 vector, the punctate FLA3-Ty1 signal is found on the detached flagellum rather than on the cell body. This supports the earlier finding (Section 3.2.7) that FLA3 appears to be anchored on the flagellar side of the FAZ only.

Attempts were made to view FLA3-Ty1 by STED upon induction of FLA2 RNAi. However, the FLA2-knockdown cells are extremely fragile and prone to damage during fixation. Consequently, the samples were not of adequately high quality for successful STED imaging.

The distances between adjacent FLA3-Ty1 punctae were measured in nine representative deconvoluted images. Regions of poor resolution were discounted, as were regions with obvious artefactual gaps. In total, 388 distance measurements were made. The mean centre-to-centre distance between adjacent punctae was 176.8 nm, with a standard deviation of 50.4.

3.2.9 Expression and localisation of tagged FLA3 in procyclic cells; conditional RNAi of FLA2 in procyclic cells

Although both FLA2 and FLA3 are bloodstream-stage specific proteins, two preliminary experiments were carried out in procyclic cells.

To investigate whether FLA3 would localise correctly if expressed in procyclic cells, *in-situ* tagging with p3074 was used to add a 4×Ty1 tag to the C-terminus of endogenous FLA3, exactly as performed in bloodstream-form cells in Section 3.2.3. Transfected cells grew at the normal rate and had normal morphology. Western blotting (Fig. 3.22) confirmed that FLA3-Ty1 was expressed at the protein level in a transfected clone. The p3074 tagging vector replaces the 3' UTR of the tagged gene with the 3' UTR of *RPA1* gene, which has implications for the expression of the tagged gene. In *T. brucei*, the 3' UTR of mRNA is important in post-transcriptional regulation, determining the half-life of the mRNA and thus the steady-state level of the encoded protein. Since *RPA1* is a constitutively expressed gene (it encodes a subunit of the DNA-replication protein RPA), its 3' UTR does

not confer stage-specificity on the tagged gene. This may explain why Ty1-tagged FLA3 was expressed at the protein level in procyclic cells even though wildtype FLA3 is bloodstream-stage specific.

Interestingly, the predominant FLA3-Ty1 band on the western blot was ~160 kDa – the same size as in BSFs. This suggests that post-translational modification had occurred correctly. However, there were also bands at ~70, 50, 40 and 25 kDa. These bands might represent breakdown products of FLA3, suggesting that FLA3 is targeted for degradation when expressed in PCFs. The bands are too small to represent complete unmodified FLA3 protein, which should be 89 kDa according to the primary sequence.

The procyclic cells expressing FLA3-Ty1 were probed for Ty1 by indirect immunofluorescence (Fig. 3.23). In marked contrast to the BSF results, FLA3-Ty1 signal was observed only in the posterior region of the cell body, mostly between the nucleus and kinetoplast. This distribution suggests that in procyclic cells, FLA3-Ty1 is confined to the lysosome or to endosomal compartments. The protein may enter the flagellar pocket, but it is not inserted into the flagellar attachment zone.

Having thus reaffirmed the bloodstream-stage specificity of FLA3, a brief investigation was made of the role of FLA2 in procyclic-form cells. The pSMOX vector and FLA2 RNAi vector (p2T7.177) were transfected into wildtype PCF cells, exactly as had been performed in BSF cells in Section 3.2.1. Noninduced FLA2-RNAi-competent procyclic cells grew at the same rate (doubling time: 11.0 h) as the parental cell line (doubling time: 10.6 h) (Fig. 3.24). Upon induction of *FLA2* dsRNA by tetracycline, no effect was observed on cell proliferation or morphology (doubling time: 11.4 h). The fact that FLA2 RNAi has no effect on procyclic cells confirms the published findings that FLA2 is not expressed in PCFs (LaCount et al., 2002). More importantly, it demonstrates that the FLA2 RNAi vector does not cross-react with FLA1, which validates the specificity of the FLA2 knockdown in bloodstream-form cells described in Section 3.2.1.

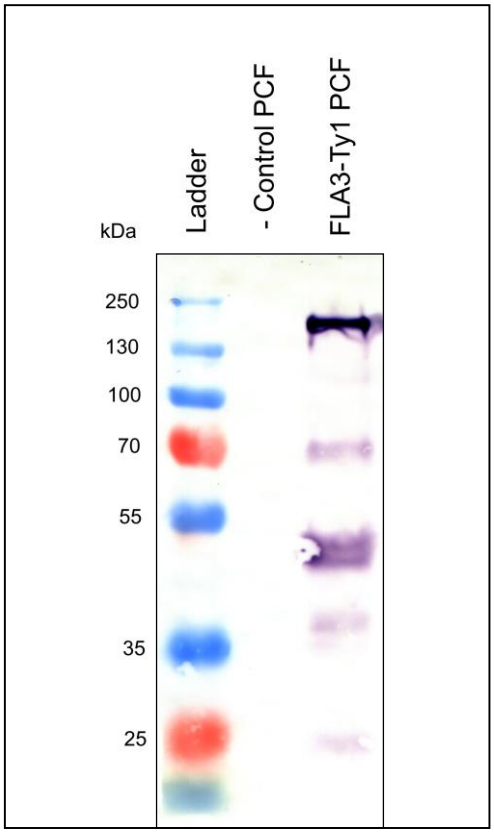


Figure 3.22: Expression of tagged FLA3 in procyclic cells.

Western blot probed for Ty1 confirms that the clonal procyclic-form cell line that has undergone *in-situ* tagging with p3074 expresses FLA3-Ty1 at the protein level. The negative control sample was untransfected procyclic cells.

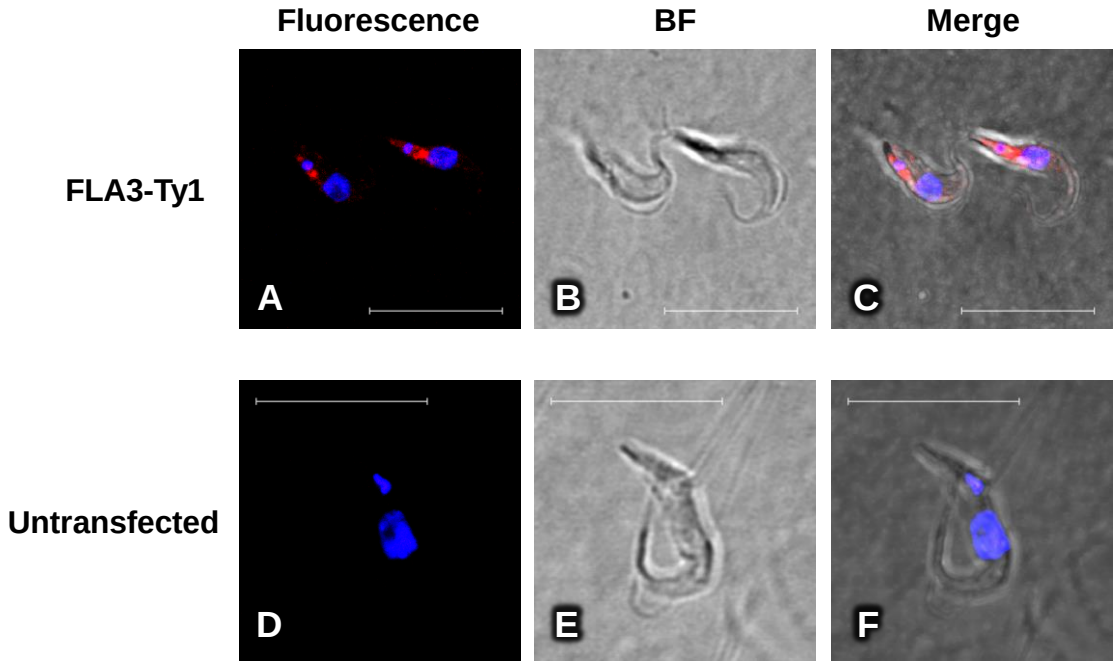


Figure 3.23: Immunofluorescence microscopy shows mislocalisation of Ty1-tagged FLA3 in procyclic cells.

(A–C) Confocal microscopy shows that a procyclic-form clone transfected with p3074 expresses FLA3-Ty1 (red), but that the tagged protein is confined to the region of the cell between the nucleus and kinetoplast.

(D–F) Untransfected control cells do not show any fluorescence signal when probed for Ty1.

Nucleus and kinetoplast are stained in blue. All scale bars are 10 μm.

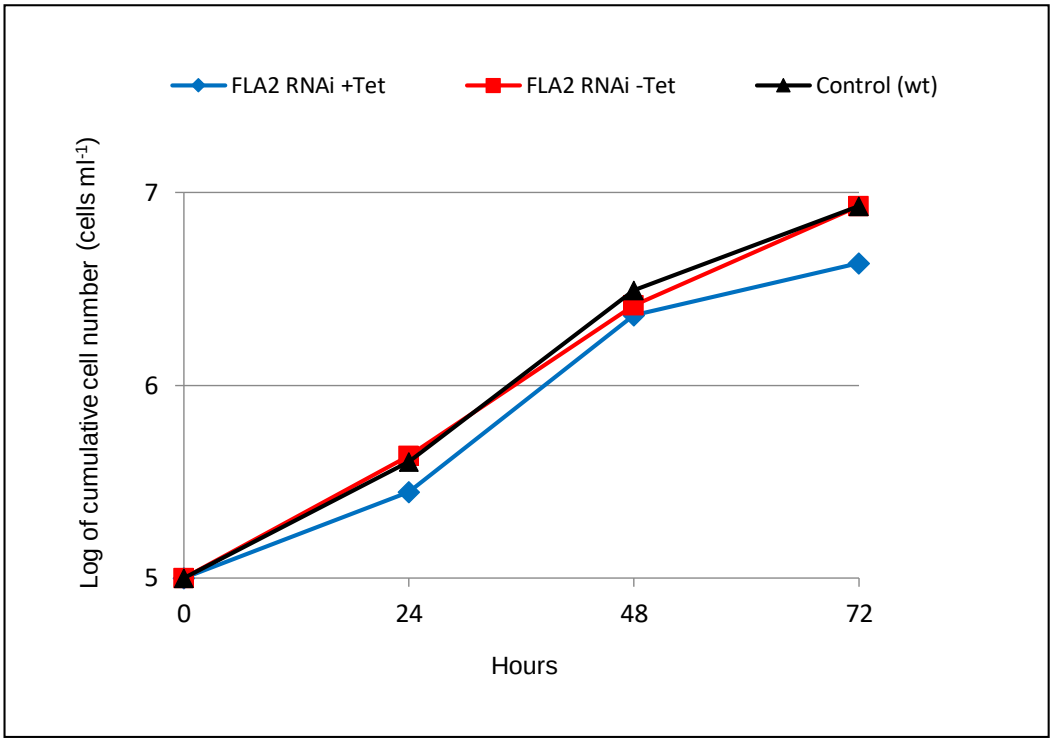


Figure 3.24: Knockdown of FLA2 has negligible effect on growth in procyclic cells.

Noninduced FLA2-RNAi-competent PCF cells [red] grew at the same rate (doubling time: 11.0 hr) as the parental cell line [black] (doubling time: 10.6 hr). The growth rate was not significantly affected upon induction of *FLA2* dsRNA by tetracycline [blue] (doubling time: 11.4 hr).

3.3 Discussion

3.3.1 Generation of a FLA2-specific RNAi vector

The generation of a FLA2-specific RNAi vector was an important step towards understanding the role of this little-characterised protein.

The challenge in producing a FLA2-specific RNAi vector was the sequence similarity between *FLA2* and *FLA1*. As described in Section 1.3.2, the initial attempt to knock down *FLA1* by RNAi (LaCount et al., 2002) had the unintended effect of ablating *FLA2* at the same time. In fact, detection of the *FLA2* transcript on northern blots, probed under conditions of high stringency, was the first evidence for expression of a gene related to *FLA1.* The *FLA1* and *FLA2* genes are nearly identical for the first 300 nucleotides and their sequences exhibit a high level of similarity throughout their length (Appendix C). There is a 44-a.a. region of insertion in FLA2 (amino acids 351–394) that is lacking in FLA1. However, this insertion is a low-complexity polyproline region that is unsuitable for targeting by RNAi because of the high risk of off-target effects. A comparison of the *FLA1* and *FLA2* sequences showed that the region of lowest similarity (64.7% identity) was the fragment (nucleotides 333–746 of *FLA2*) depicted in Table 3.1. This fragment was therefore amplified, digested with *Bam* HI and *Xho* I and inserted between double-headed T7 promoters in the p2T7.177 plasmid, as described in Section 3.2.1, to produce a FLA2-specific RNAi vector.

Off-target effects ("cross-RNAi") are a potential issue in any RNAi experiment (Qiu et al., 2005). However, the siRNAs of *T. brucei* are quite long (24–26 bp) relative to those of most higher eukaryotes (typically 21–23 bp), which somewhat reduces the risk of cross-RNAi (Djikeng et al., 2001). For example, in a study of paraflagellar rod proteins, it was found that knockdown of *PAR2* had no detectable effect on expression of *PFRA* (at the level of northern blotting and western blotting), even though the *PAR2* dsRNA shared 83% identity with *PFRA* (Durand-Dubief et al., 2003). This result was all the more remarkable because when the two sequences were aligned, there were nine blocks of perfect identity >20 bp in length. However, there were only three blocks of identity that were 26 bp or more, so the number of siRNAs sharing complete identity with the *PFRA* mRNA would have been low. Similarly, a 75% knockdown of *TbSIR2RP1*, a histone deacetylase, had no observable effect (by northern blot or western blot) on the expression of *TbSIR2RP2* or

TbSIR2RP3, each of which has 47% identity to *TbSIR2RP1* over the region used in the RNAi construct (García-Salcedo et al., 2003). The efficiency of silencing depends greatly on the amount of siRNA present: there are different thresholds for silencing an RNA of identical sequence and silencing an RNA with related but non-identical sequence, which ensures that off-target effects are minimised (Durand-Dubief et al., 2003).

In view of these results, it is not surprising that the *FLA2* dsRNA employed in this thesis did not affect the *FLA1* transcript. As shown in Table 3.1, the longest region of identity between the selected fragment of *FLA2* and the corresponding region of *FLA1* is a single 24-bp block. In contrast, the 1000-bp fragment of *FLA1* employed by LaCount et al. (2002) had 81% overall identity with the *FLA2* gene – and, crucially, shared 100% identity over the first 300 nucleotides. This long region of identity explains the observed effect on the *FLA2* transcript.

Two important results demonstrated the specificity of the FLA2 RNAi vector generated here.

Firstly, qRT-PCR data (Fig. 3.5 A) showed that upon induction of *FLA2* dsRNA, *FLA2* mRNA levels decreased fourfold, confirming knockdown of *FLA2*. Critically, *FLA1* mRNA levels did not decrease (Fig. 3.5 B), confirming that, as intended, the *FLA2* dsRNA did not ablate *FLA1* mRNA. For the first time, therefore, the function of the FLA2 protein could be investigated independently of FLA1.

Interestingly, *FLA1* mRNA levels increased twofold after 72 hours of induction of FLA2 RNAi (Fig. 3.5 B). It is not known whether this increase was due to a specific upregulation of FLA1 upon loss of FLA2, or to a more general stress response in the cells subjected to FLA2 RNAi; in either case, it confirms that the knockdown effects (described in Section 3.3.2 below) were not due to loss of FLA1.

The qRT-PCR measurements reported here (Appendix E) indicated that expression of the *FLA1* transcript is normally twice as high in PCFs as in BSFs (a BSF:PCF ratio of 0.58), which is in line with the published high-throughput data (Table 1.4.2). Therefore, the observed twofold increase in *FLA1* mRNA upon knockdown of *FLA2* means that in these BSF cells, the *FLA1* transcript was approximately as abundant as it would ordinarily be in wildtype

PCFs. However, the observed increase in expression of the *FLA1* transcript does not necessarily mean that there was an increase in the production of the FLA1 protein.

The second demonstration of the specificity of the FLA2 RNAi vector came when the vector was transfected into PCF cells. Upon induction of *FLA2* dsRNA by tetracycline, no growth effect was observed over a 72-h period (Fig. 3.24), and no morphological effects were observed. (However, northern blotting was not performed to confirm that induction of *FLA2* dsRNA had occurred successfully.) The fact that the procyclic cells continued to proliferate as normal and did not exhibit flagellar detachment implied that *FLA1* was not knocked down by the vector. Knockdown of *FLA1* in PCFs leads to flagellar detachment, growth arrest, multinucleation and eventual cell death (LaCount et al., 2002, Rotureau et al., 2014).

These results indicated that the FLA2 RNAi vector could be used to elucidate the specific function of FLA2 in bloodstream-form cells.

3.3.2 FLA2 is essential for cytokinesis in the bloodstream-form cell

The FLA2 RNAi results presented in Figs. 3.1–3.4 confirm for the first time that FLA2 is required for flagellar attachment and cytokinesis in BSFs.

Taken together, the RNAi data show that ablation of FLA2 causes a rapid defect in the attachment of the new flagellum to the cell body but has no effect on attachment of the existing flagellum, at least during the first 24 hours of induction of *FLA2* dsRNA. This selective effect on the new flagellum supports the finding that, once assembled, the constituents of the FAZ are stable and cannot be exchanged between the old FAZ and the new (Sunter et al., 2015b, Woods et al., 2013).

Detached flagella continued to beat, but directional cell motility was lost, as seen upon knockdown of FLA1, FLAM3, FLA3 and FLA1BP (Rotureau et al., 2014, Woods et al., 2013, Sun et al., 2018). Once detached, flagella did not reattach to the cell body.

Cells with a detached new flagellum failed to undergo cytokinesis. However, mitosis and kinetoplast replication continued unabated, leading to multinucleation and loss of normal cell shape, i.e. "monster" morphology (Broadhead et al., 2006). After 72 h of induction of *FLA2* dsRNA, cell death occurred for the majority of the population.

Noninduced cells grew at a slightly reduced rate compared to the parental cell line (doubling times: 11 h vs 8 h), indicating that low levels of *FLA2* dsRNA were probably produced in the absence of tetracycline.

The effects of knockdown of FLA2, as presented here, are very similar to the effects of concomitant FLA1/FLA2 knockdown in BSFs (LaCount et al., 2002), with two minor exceptions.

Firstly, LaCount et al. observed that cells ceased division almost immediately after addition of tetracycline, with cell death apparent even by the 12-h timepoint. Here, however, the population was consistently observed to double once in the first 24 hours of induction. This discrepancy seems to be a matter of different response times between different cloned cell lines, rather than a qualitative difference in response to *FLA2* dsRNA. The growth rate of the cells, the speed of response to tetracycline and the efficiency of FLA2 knockdown (i.e. the percentage loss of FLA2 at the protein level) will differ between strains and between clones and will affect the onset of the growth phenotype.

Secondly, after 24 hours of induction with tetracycline, LaCount et al.'s FLA2 RNAi cells exhibited flagellar detachment but still retained normal trypomastigote shape; the enlarged monster-cell morphology was not observed until the 30-h timepoint. Here, in contrast, most of the FLA2 RNAi cells had apparently lost normal morphology by the 24-h timepoint (Fig. 3.4). The major cell damage depicted in Fig. 3.4 was somewhat exacerbated by the fixation conditions, however. It is more representative of the appearance of live populations at the 48-h timepoint. The results presented here are therefore in line with those of LaCount et al. with regards to onset of morphological defects.

It is evident, therefore, that specific loss of the FLA2 protein could account for all the phenotypic effects observed by LaCount et al. (2002) in BSF cells. This observation lends considerable weight to the hypothesis that FLA2 and FLA1 are stage-specific paralogues, and that FLA1 is not expressed in BSFs (as suggested by proteomics data, Section 1.4.4). The implications of the stage-specificity of both FLA1 and FLA2 are discussed further in Chapter 5. The ideal way to confirm that FLA1 does not play a role in flagellar attachment in the bloodstream form would be to generate a FLA1-specific RNAi vector. It seems highly probable that expression of specific *FLA1* dsRNA in BSFs would have no effect on growth or morphology.

An interesting feature of the FLA2-RNAi growth curve in Fig. 3.1 is the apparent return to normal growth rate between the 72-h and 96-h timepoints. This phenomenon is consistent with the view that the cells have lost responsiveness to tetracycline, which can occur via mutations in T7 promoters or tetracycline operators, loss of RNAi inserts or defects in the RNAi machinery (Motyka and Englund, 2004). Outgrowth of cells that do not produce dsRNA upon induction with tetracycline ("RNAi revertants") is a common phenomenon: the frequency of RNAi revertants in the bloodstream form has been estimated as 1 in 5000 cells (Chen et al., 2003). However, it most often occurs after a long period of induction (14 days or more), or after long-term passage of noninduced cells (Chen et al., 2003, Motyka and Englund, 2004). For example, when a tetracycline-inducible conditional null mutant was used to knock down the UDP-glucose 4' epimerase (Roper et al., 2002), BSF cells spontaneously recovered from the Tet-induced phenotype after 16 days by deleting the tetracycline-repressor gene (*TETR*). The hypothesis that FLA2-RNAi cells stop producing *FLA2* dsRNA after a few days of induction could be tested by northern blotting. Based on the growth curve data, it is expected that between the 72-h and 96-h timepoints, a decrease in *FLA2* dsRNA would be observed.

LaCount et al. (2002) stated that their FLA1/FLA2 RNAi BSF cells reverted to a tetracycline non-responsive phenotype more quickly than their FLA1/FLA2 RNAi PCF cells – however, it is not clear in what timeframe this reversion occurred. It is mentioned elsewhere in the publication that tubulin RNAi cells lose responsiveness to tetracycline over a period of several months, so they may have been referring to a long-term effect. However, the FLA2 RNAi growth curves presented by LaCount et al. only show data for 56 hours of induction in the case of PCFs and 48 hours of induction in the case of BSFs. It is possible, therefore, that if these growth curves had been continued for 96 h, a reversion to normal growth rate may have been observed.

Another possible explanation for the return to normal growth rate observed in Fig. 3.1 is that the cells responded to the lack of FLA2 by producing FLA1 protein, which was able to mediate flagellar attachment in place of FLA2. This hypothesis is supported by the qRT-PCR measurements (Fig. 3.5, discussed in Section 3.3.1), which show twofold upregulation of *FLA1* mRNA upon knockdown of FLA2. However, the increased *FLA1* mRNA levels do not necessarily mean that FLA1 protein is produced, because the translational control

mechanisms may intervene. Furthermore, it is not known whether FLA1 can insert into the bloodstream-form FAZ, so this explanation remains a speculative one. It is more likely that, in the absence of its binding partner, FLA1BP, FLA1 would fail to locate normally, and would subsequently be degraded. Note that even if FLA1 could compensate for loss of FLA2 *in vitro*, it probably could not do so *in vivo*. It is likely that the exposed regions of FLA1 and FLA2 are adapted to different host environments, in accordance with their stage specificity, as discussed in Chapter 5.

3.3.3 FLA3 is essential for flagellar attachment and cytokinesis in the bloodstream form.

Knockdown of FLA3 in BSF cells produces an almost identical phenotypic response (Figs. 3.6, 3.7, 3.16) to that observed for FLA2 knockdown. Detachment of the newly synthesised flagellum is the first effect observed upon induction of *FLA3* dsRNA. Cells with a detached flagellum fail to undergo cytokinesis, but mitosis, kinetoplast replication and synthesis of new flagella continue. Consequently, this growth arrest is accompanied by progressive loss of normal cell morphology. After two days of induction of FLA3 RNAi, the population declines due to death of unviable monster cells.

This characteristic growth defect (i.e. cytokinesis failure but continued progression of the cell cycle) may occur sooner upon induction of FLA3 RNAi than upon induction of FLA2 RNAi (cf. Fig. 3.6, Fig. 3.1). Whereas induced FLA2-RNAi cells doubled once in the first 24 hours, induced FLA3-RNAi cells failed to divide at all. However, the speed of onset of the growth defect seems to vary slightly from clone to clone. In Fig. 3.16 A, for example, induced FLA3-RNAi cells are seen to double in the first 24 hours. As suggested earlier for FLA2-RNAi cells, the ability of cells to divide after induction of dsRNA will depend on how far they had progressed along the cell cycle at the time of induction.

Noninduced cells grew at a slightly reduced rate compared to the parental cell line (doubling times: 14 h vs 10 h), indicating that low levels of *FLA3* dsRNA were probably produced in the absence of tetracycline.

In terms of morphology, induced FLA3-RNAi cells were indistinguishable from induced FLA2-RNAi cells at the same timepoint. The similarity of the knockdown phenotypes supports the hypothesis that FLA2 and FLA3 perform similar roles in the FAZ. However, it is

obvious that FLA2 cannot compensate for loss of FLA3, and vice versa, suggesting that each protein has a distinct role in the function of flagellar attachment. One obvious explanation is that they may be binding partners (discussed in Chapter 6), which would explain why they cannot function in each other's absence.

Between the 48-h and 72-h timepoints after induction of *FLA3* dsRNA, the cell population approximately halved. This decline in cell number indicates widespread death of cells that have failed to divide but have re-entered the cell cycle several times.

A return to normal growth rate and morphology was consistently observed between the 72-h and 96-h timepoints after induction of *FLA3* dsRNA, as was observed previously for FLA2-RNAi cells. Again, this phenomenon probably indicates the outgrowth of cells that are refractory to the induction of dsRNA upon addition of tetracycline (see Section 3.3.2) and which therefore express normal levels of FLA3.

The possibility that knockdown of FLA3 in BSFs could cause upregulation of FLA1BP, the PCF paralogue of FLA3, and that FLA1BP could compensate for loss of FLA3 is unlikely since FLA3 RNAi has no effect on the transcription of *FLA1BP* mRNA, at least at the 18-h timepoint (Woods et al., 2013).

The effects of FLA3 RNAi presented here are in close agreement with those described in the published literature (Woods et al., 2013), which used the same FLA3 RNAi vector in a different RNAi cell line (the 328.114 single-marker *T. brucei* strain (Wirtz et al., 1999)). The only difference is that the loss of responsiveness to tetracycline, observed here after 96 hours of induction, was not documented before. However, the FLA3-RNAi growth curve presented by Woods et al. describes the first 60 hours of induction only, so it remains possible that a return to normal growth rate would have been seen if that growth curve had been continued for longer.

The parental RNAi-competent cell line used here (Lister 427 MITat 1.1 BSFs transfected with pSMOX) should be superior than the 328.114 single-marker strain in terms of both inducibility of dsRNA and tight repression of dsRNA expression in non-induced cells (Poon et al., 2012). The 328.114 single-marker cell line (also known as SMB or S16) was generated in 1999 to enable expression of T7 RNA polymerase (T7RNAP) and the tetracycline repressor (TETR) under a single selection marker in Lister 427 MITat 1.2 BSF

cells (Wirtz et al., 1999). The *T7RNAP* and *TETR* genes were inserted into the tubulin locus, with the former transcribed by endogenous RNA Pol II and the latter by T7RNAP.

The pSMOX plasmid, on the other hand, was designed in 2012 in view of improved understanding of the trypanosome genome (Poon et al., 2012). This single-marker vector inserts into the tubulin locus; *T7RNAP* and *TETR* are both transcribed by endogenous RNA Pol II read-through. Crucially, however, codon optimisation was performed so that these transgenes would use codons that are used frequently by trypanosomatids. Intergenic sequences from the PFR2 protein were used as UTRs for the transgenes. Consequently, *T7RNAP* and *TETR* were produced to a higher level than previously. Furthermore, a nuclear-localisation signal was added to both the *T7RNAP* and *TETR* genes, improving the efficiency of both encoded proteins. Therefore, when the pSMOX system and the 328.114 system were each tested in conjunction with the p2T7.177 RNAi vector, the pSMOX cell lines performed substantially better and exhibited lower leakiness in the noninduced state (Poon et al., 2012).

It is probable, therefore, that the FLA3-RNAi clones generated here produce a more tightly regulated knockdown effect than those generated previously using an outdated system.

3.3.4 *In-situ* tagging of FLA3

A major impediment to the characterisation of FAZ proteins is the lack of specific tools/reagents such as antibodies, to allow analysis at the protein level. As described in Section 3.1, a polyclonal antibody was raised against FLA3 (Woods et al., 2013), but it was prone to nonspecific interactions and therefore produced excessive background signal during localisation studies. Furthermore, it had to be affinity purified from immune serum. These problems were overcome in this thesis using a plasmid-based *in-situ* tagging approach. To enable precise localisation of FLA3 and to facilitate investigation of the putative interaction partners of FLA3, two cell lines were established in which endogenous FLA3 was tagged at the C-terminus with Ty1 or HA, respectively (Section 3.2.3).

Epitope tagging did not affect the normal processing, localisation and function of FLA3. Firstly, based on electrophoretic mobility, the tagged protein migrated at approximately the same molecular weight (~160 kDa; Fig. 3.9) as the native wildtype protein, which migrates at almost double the size (~89 kDa) predicted by the amino acid sequence

(Woods et al., 2013). This size difference has been shown to be due largely to extensive post-translational modification by addition of *N*-glycans (Woods et al., 2013). Therefore, the presence of a C-terminal Ty1 or HA tag does not seem to affect the normal post-translational modification of the protein.

Indeed, the C-terminal tag should not be in a position to interfere with the glycosylation of FLA3, which occurs solely in the extracellular domain. The C-terminal transmembrane domain (residues 750–773) is followed by a strong stop-transfer signal (KRRR, residues 775–778). According to the positive-inside rule (von Heijne, 1992), the C-terminal tail of FLA3 must therefore be on the cytoplasmic side of the membrane throughout the processing and trafficking of the protein. The extracellular domain, in contrast, must be on the luminal side of the ER/Golgi membranes. Therefore, the C-terminal epitope tag is separated by a membrane from the region of the protein that is glycosylated.

Tagged FLA3 appears as the same size as wildtype FLA3 on a 10% acrylamide gel because the molecular weights of the respective tags (4.7 kDa for (4×) Ty1; 13.2 kDa for (12×) HA) are negligible compared to that of the protein.

Only two bands were observed in western-blot samples of tagged cells (Fig. 3.9), suggesting that no nonspecific tagging had occurred. (The 50/55-kDa band observed on blots of tagged cells is discussed in Section 3.3.9.) Specific loss of the tagged protein was observed upon FLA3 RNAi (Fig. 3.17), confirming the ~160-kDa band to be FLA3.

Furthermore, addition of either tag did not affect targeting of FLA3 to its normal membrane location, since the tagged protein, like the wildtype form, was solubilised by detergent extraction but not by mechanical lysis (Fig. 3.10). The C-terminus was chosen as the site of tag addition to improve the likelihood of correct localisation. FLA3 may have an atypical N-terminal signal sequence (see Section 1.3.3); consequently, insertion of an epitope tag at the N-terminus could have affected the processing/localisation of the protein, or the tag could have been lost when the leader sequence was cleaved.

Crucially, FLA3-Ty1 and FLA3-HA were shown to localise alongside the attached flagellum in intact whole cells (Figs. 3.12, 3.15). The localisation pattern of tagged FLA3 matched that observed when wildtype FLA3 was probed using an affinity-purified polyclonal antibody (Woods et al., 2013) but was much clearer in terms of signal intensity and lack of

background (discussed further in Section 3.3.5). It was concluded that epitope tagging had no apparent effect on positioning of the protein and was therefore suitable for high-resolution localisation studies.

Finally, the tagged cell lines grew at the same rate as wildtype cells (Fig. 3.8) and possessed no observable morphological defects. Knockdown of FLA3 in the tagged cells produced the same loss-of-function effects as in control cells: flagellar detachment and cell death. This supported the hypothesis that the presence of the epitope tag did not alter the functionality of the protein, which is consistent with the view that FLA3's role in attachment is mediated by the extracellular domain.

When the evidence is considered as a whole, it seems evident that both FLA3-Ty1 and FLA3-HA are accurate representations of wildtype FLA3. The immunofluorescence signal was stronger in the FLA3-Ty1 cells than in the FLA3-HA cells, so FLA3-Ty1 was used for all subsequent localisation studies. The difference in signal was probably a matter of the sensitivity of the primary antibodies used, rather than inherent superiority of the FLA3-Ty1 tagging.

The reason for employing two different epitope tags was twofold: firstly, it enabled a greater range of possible experiments (e.g. investigation of the possibility of self-binding of FLA3, see Chapter 4) and secondly, it increased the likelihood of discovering the optimal tagging system for precise localisation of FLA3. Both Ty1 and HA tags have previously proven effective for investigating FAZ proteins. For example, Ty1-tagging has been used to reveal the localisation of FAZ10 (Morriswood et al., 2013) and CC2D (Zhou et al., 2011), while 3×HA tags have been used to localise surface membrane proteins such as FS179 (Oberholzer et al., 2011) and FAZ filament proteins including FAZ2, FAZ8 and KMP11 (Zhou et al., 2015), both endogenously and ectopically. However, each protein-tag interaction is unique and, as such, the success of a tag in a specific situation cannot be predicted in advance. It is therefore important to evaluate multiple tags, as was performed here, and to optimise the experimental conditions according to the circumstances. For example, it was decided to tag FLA3 at the C-terminus to avoid interference with the putative signal sequence at the N-terminus. *In-situ* tagging using the endogenous promoter was chosen so that expression of the tagged protein would match the wildtype expression level as closely as possible, which is important for the accuracy of localisation.

Ectopic overexpression is a useful technique for analysing the real-time dynamics of a protein e.g. its pattern of integration to the FAZ, but it is not suitable for precise localisation studies because overexpression can cause a protein to mislocate.

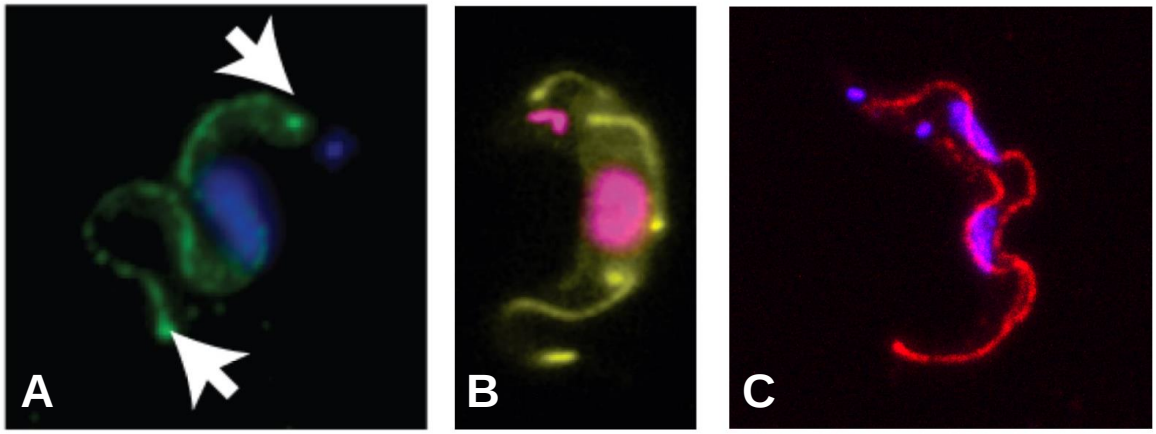


Figure 3.25: Comparison of FLA3 immunofluorescence.

(A) is taken from (Woods *et al.*, 2013). FLA3 (green) probed with anti-FLA3 affinity-purified antibody. Nucleus/kinetoplast stained in blue. The white arrows indicate the apparent increase in FLA3 signal at the extremities of the FAZ. **(B)** is taken from (Gadelha *et al.*, 2015). FLA3-GFP (yellow) probed with anti-GFP antibody. Nucleus/kinetoplast stained in pink. **(C)** is taken from Figure 3.14, above. FLA3-Ty1 (red) probed with anti-Ty1 antibody. Nucleus/kinetoplast stained in blue.

Note that (B) seems to show a continuous signal along the length of the FAZ, whereas the FAZ signal in (A) and (C) seems to be more punctate in certain regions.

MCFIFGVEMS NLAKRPMSLR K**LPQLLLLIM IGIAFVAVEC IG**APVKLPRR VDTVAGQFGV
EGETNGYPNT TRLTEPYALC RGRTNDEILV GSSNSFRNYS RKTKETGTYL RYNVGDSVIS
GSSTINKPRS CVRRGSGNHT IIYFVDDQKD IKYIVGDDVS SFSVPTSGSL NAVAVHEGTL
YVTDQNNKSV WKCGLGGAGK PQSCEEKKFT SVTLDAKPEG IAVTSKGIFV TARDSSNKGA
LLWLDMSGGN RKGNVSGGFV DVFSTESGVL YAATEKELYT VTATDTSLSV TSFAGKNTSQ
CYFPTNGEDI VLCDNSRLLV IEEYEMYVTS KAKHTMRALT LPPVNLTAIF RGRPAPVGYP
NTTIMEQFVA SLTEDVNKAL GTNDSYVDPD SVRVDPDTWE TNFTVFVQQT RFDNTTEEKL
RSLTYTQTDK TVDEYYGLTD EYVYIDTVLV PFCDDASLVT IQRALAREAG RALNFSLVYA
DKPITFGSDV AENVTAVKLL MPHSFKNATT PKQLSAANLT DFAHNLVKDL RASDTRVDIT
FPDPPFNFSA VVPEREQEVR WFVHGKVMKQ LEICERLGSQ GDAAVIAAAA DATARGKANV
TLNTSGVKAN DTGVGPNTTN TAGGANTTAN VAANGTANVI VNPSTNATPT GTTNASVTNT
TERAVPVVAP TQPSNGYAEC RSAITNRTET QNMEPPYDRK HRYEVFLPKK YDFNVSWCVD
IIDWRDLDEM LNNRTDEVVE KSLSWCGHGC **IIAFAVVGSL IAACLVVLAV VLT**SKRRRLA
AVVAPPRPKF VSTVEDDDED RVSNIGVPLT DGKGTTAP

Figure 3.26: Protein sequence of FLA3 (Tb927.5.4570)

Putative *N*-glycosylation sites are highlighted in red. Predicted transmembrane regions are underlined. A hypothesized alternative translation start site is highlighted in turquoise (residue 17). C-terminal proline residues are highlighted in purple.

The other FLA3 gene, Tb927.5.4580, encodes a protein that is 92% identical to the one shown here, and the first 59 residues are the same.

Sequence and predicted TM regions were obtained from the TriTryp database (http://tritrypdb.org/tritrypdb/app/record/gene/Tb927.5.4570).

3.3.5 Localisation of tagged FLA3

Having established that FLA3-Ty1 and FLA3-HA seemed to be accurate representatives of wildtype FLA3, the localisation of the tagged protein was examined closely and compared to the results obtained previously, when FLA3 was probed directly with antibody (Woods et al., 2013) or tagged at the C-terminus with GFP (Gadelha et al., 2015). (The GFP-tagging of FLA3 occurred independently as part of an investigation into the *T. brucei* surface proteome, before FLA3 had been characterised, and was published during the current project.) The visualisations of FLA3 obtained in each of these three instances (Fig. 3.25) show that the protein locates along the entire length of the attached flagellum, but does not extend into the free flagellum, which suggests that FLA3 belongs to the FAZ. It is useful to contrast the results obtained using different methods; the fact that they all give the same localisation attests to their accuracy. The polyclonal-antibody approach has the advantage of probing the wildtype protein, so there is no chance of mislocalisation; however, there is the distinct possibility of unwanted cross-reaction with other antigens. The tagging approach uses only monoclonal antibodies that have been demonstrated to produce little or no background signal in the trypanosome, and it therefore gives a more sharply defined image. Furthermore, the antigen is repeated in the epitope tags (4× Ty1 for p3074; 12× HA for p2708), so the signal is amplified when probed.

Most of the immunofluorescence images of FLA3-Ty1 obtained here seemed to indicate a continuous signal along the FAZ, although certain images seemed to exhibit a more punctate signal in certain regions (Fig. 3.14). Woods et al. (2013) described FLA3 as exhibiting "a strong, punctate fluorescence signal associated with the attachment zone of the flagellum". Gadelha et al. (2015) stated that FLA3-GFP (which they termed ESP8) located to the flagellar pocket and FAZ. In the fluorescence image they showed (Fig. 3.25 B), the FLA3 signal seemed to form a continuous line along the FAZ. There was uncertainty, therefore, about the true distribution of FLA3 within the FAZ. Super-resolution microscopy was required to resolve this issue (discussed in Section 3.3.7).

The fluorescence images presented here sometimes seem to show an increased intensity of FLA3 signal at the proximal (Fig. 3.14 C) and/or distal (Fig. 3.14 B) end of the FAZ. Woods et al. (2013) and Gadelha et al. (2015) also observed an increased FLA3 signal at both extremities of the FAZ (Fig. 3.25 A, B). These points of increased fluorescence may be

due to a localised increase in abundance of the protein in these regions; alternatively, the tag may simply be more accessible to primary antibodies at these locations, which therefore show a disproportionate signal when probed.

Several other FAZ proteins (CC2D, FAZ9, FAZ10, FAZ11) have been shown to exhibit increased signal at the proximal end of the FAZ (Zhou et al., 2011, Sunter et al., 2015b, Morriswood et al., 2013). In the cases of CC2D, FAZ10 and FAZ11, this point of increased signal has been described to overlap with the hook complex (bilobe), at least in procyclic cells. The hook complex is a constitutive cytoskeletal structure; since it is intracellular, however, it seems unlikely that membrane proteins such as FLA3 or FAZ9 could have any significant interaction with it. Besides, FLA3 is postulated here to locate to the flagellar side of the FAZ only (Section 3.3.6), which rules out any direct interaction with the hook complex.

Similarly, several proteins have been shown to have increased signal at the anterior tip of the FAZ in procyclic cells but have not yet been localised in BSFs. These include FAZ4, 6, 7, and 11; furthermore, both FAZ10 and CC2D have sometimes exhibited increased signal at the anterior FAZ tip when probed in PCFs (Sunter et al., 2015b, Morriswood et al., 2013, Moreira et al., 2017, Zhou et al., 2011). As yet, no FAZ protein has yet been demonstrated to give increased signal at the anterior tip of the mature bloodstream-form FAZ. However, if these proteins locate in the same manner in BSFs as they do in PCFs, it is possible that there is a specialised protein complex at the distal tip of the FAZ, which would explain the apparent increase in intensity of FLA3 at that point.

It must be noted, however, that it can be difficult to image the entire length of the FAZ with accuracy. Because of its three-dimensional nature, the FAZ is not always aligned precisely along the focal plane. There can, therefore, be gaps or discrepancies in the FAZ signal that are due to inconsistent imaging. This issue has been tackled here by taking z-stack images that are composed of sets of superimposed images from different focal planes. However, care must still be taken when interpreting differences in intensity to real differences in distribution of a protein.

A further caveat in interpreting the FLA3-Ty1 and FLA3-HA immunofluorescence data must also be emphasised: C-terminal tagging with p3074/p2708 changes the endogenous 3' end of the target gene to the 3' UTR of the *RPA1* gene, which could potentially increase

or decrease expression of the transgene relative to the wild-type gene (Bastin et al., 1996). As described in Section 1.4.2, post-transcriptional regulation of trypanosome gene expression occurs largely through the binding of regulatory proteins to the 3' UTR of mRNA. These regulatory proteins determine the half-life of the mRNA and thereby influence the steady-state level of the encoded protein (Clayton, 2019).

Furthermore, there was a risk that the epitope tag could interfere with any localisation signals or protein-binding motifs that may be present in the C-terminal tail of the protein. No such motifs have yet been identified, except the positively charged KRRR stop-transfer sequence just after the TM domain. Interestingly, the C-terminal tail of FLA3 has five proline residues (Fig. 3.26). Proline acts as a flagellar-targeting signal for the peripheral membrane protein, glycosylphosphatidylinositol phospholipase C (GPI-PLC) (Sunter et al., 2013). A specific proline residue, ~70 amino acids downstream of the acylated region, is necessary for localisation of GPI-PLC to the inner face of the flagellar membrane. When the proline residue was mutated, the protein mislocalised throughout the cell-body membrane. It was therefore postulated that a putative peptidyl prolyl isomerase could regulate entry to the flagellar membrane. It is possible that FLA3 could avail of such a targeting system, if it exists.

However, the close agreement between the images presented here of FLA3-Ty1 and the published images of antibody-probed wildtype FLA3 (Fig. 3.25) demonstrates that C-terminal tagging did not interfere with the localisation of FLA3.

3.3.6 FLA3 locates to the flagellar membrane.

A key question concerning the localisation of FLA3 is whether its TM domain(s) are found in the flagellar membrane or in the cell-body membrane. The FLA1BP protein was seen along the detached flagellum when probed by immunofluorescence in PCFs upon knockdown of *FLA1* (Sun et al., 2013). It was concluded that the protein located to the flagellar membrane. As the putative BSF paralogue of FLA1BP, FLA3 is therefore likely to locate to the flagellar membrane in bloodstream-form cells. Indeed, high-throughput proteomics found FLA3 to be a component of the flagellar-surface proteome when the flagellum was isolated after knockdown of FLA1/FLA2 (Oberholzer et al., 2011), which suggests that FLA3 belongs in the flagellar membrane. However, this proteomics study did not analyse the cell-body membrane, so the possibility that FLA3 was also present in the

cell-body surface proteome was not investigated. Indeed, Woods et al. (2013) provided evidence suggesting that, in fact, FLA3 locates to the cell-body membrane only. In wildtype cells where the flagellum became partially detached during the fixation process, FLA3 was found only on the cell-body side (Woods et al., 2013).

To address this issue, Ty1-tagged FLA3 was expressed in cells capable of FLA2 knockdown, to enable visualisation of FLA3 in cells with detached flagella. If the FLA3 signal localised along the detached flagellum, it would suggest that the protein's TM domain(s) are in the flagellar membrane. Conversely, if the FLA3 signal localised along the cell body, it would suggest that the protein's TM domain(s) are in the cell-body membrane. There was a certain possibility that knocking down a FAZ protein would alter the normal localisation pattern of FLA3, but such an occurrence would be noticeable in a small-scale fluorescence study, and adjustments could be made accordingly.

The immunofluorescence results (Fig. 3.20) gave a strong indication that the TM domain(s) of FLA3 are in the flagellar membrane. Upon induction of FLA2 RNAi, a distinct line of FLA3-Ty1 signal was observed along the detached (new) flagellum. The FLA3-Ty1 signal was also observed along the FAZ of the attached (old) flagellum.

It is well known that in the absence of an attached flagellum, the intracellular FAZ filament does not extend beyond a short stub (Sunter and Gull, 2016). Therefore, when the new flagellum is detached by knockdown of a FAZ protein, the proteins of the cell-body side of the FAZ are often seen to accumulate at the posterior end of the cell. For example, FAZ1 is seen to accumulate at the posterior end of the cell upon FLA3 RNAi (Woods et al., 2013). Here, however, no such accumulation of FLA3-Ty1 is observed upon knockdown of FLA2. This finding supports the view that FLA3-Ty1 is not a component of the cell-body membrane.

Taken together, the data show that FLA3-Ty1 locates to the flagellar membrane rather than the cell-body membrane of the flagellar attachment zone.

The location of FLA3 when the newly synthesised flagellum fails to attach should be an accurate indication of the wild-type positioning of the protein. As the new FAZ forms, FLA3 must first enter the flagellar-pocket membrane, where it must be sorted to the appropriate FAZ membrane (i.e. the flagellar membrane, cell-body membrane, or both). It

is thought that the transmembrane FAZ proteins are trafficked to their respective membrane domains (possibly as pre-assembled hemidesmosome-like structures) before they can be assembled into the growing FAZ structure at the neck of the flagellar pocket (Sunter et al., 2015b). Therefore, trafficking of FLA3 to the appropriate membrane domain should occur independently of interaction with the proteins of the opposing membrane. Just as FLA1 is not required for localisation of FLA1BP (Sun et al., 2013), FLA2 should not be required for localisation of FLA3 even if the two proteins are binding partners.

In contrast, the location of FLA3 in a scenario where an existing flagellum were induced by mechanical force to detach from the cell body would not be an accurate indication of where the protein is usually located. Woods et al (2013) concluded incorrectly that FLA3 located to the cell-body membrane based on the location of the protein after unintentional detachment of the flagellum during the fixation process. It is highly probable that the whole FAZ acted as a single cohesive unit and remained intact on the cell-body side while the rest of the flagellum (PFR and axoneme) detached. (As discussed in Section 6.1, the FAZ seems to be anchored more firmly in the cell body than in the flagellum.) Therefore, the localisation of FLA3 following mechanical shearing of the flagellum was misleading.

3.3.7 FLA3-Ty1 has punctate distribution along the FAZ.

As described in Section 3.3.5, it was not clear in the immunofluorescence images of FLA3-Ty1 and FLA3-HA whether the tagged protein had a continuous or discontinuous (punctate) distribution along the FAZ. To determine the true distribution of FLA3-Ty1, stimulated emission depletion (STED) microscopy was used to produce super-resolution images of the tagged protein (Fig. 3.21). This was the first application of the STED technique to a protein of the flagellar attachment zone. The results were clear: FLA3-Ty1 localised in a distinctly punctate line along the FAZ. The distances between adjacent FLA3-Ty1 punctae were measured in nine representative deconvoluted images; in total, 388 distance measurements were made. The mean centre-to-centre distance between adjacent punctae was 177 nm, with a standard deviation of 50.

What does the punctate nature of the FLA3-Ty1 signal inform us about the localisation of FLA3 within the complex architecture of the FAZ? This question is discussed in detail in Chapter 6. It is tempting to speculate that FLA3 might be a component of the regularly spaced junctional complexes ("*macula adherens*") that are seen to traverse the intermembrane space when the FAZ is imaged by electron microscopy (Vickerman, 1969, Höög et al., 2012, Sunter and Gull, 2016). The centre-to-centre distances between these junctional complexes have been measured as 95 nm in BSFs (Vickerman, 1969) and 122 nm ±25 in PCFs (Höög et al., 2012). These structures are thought to connect the flagellum to the cell body by acting as molecular rivets, but their architecture is not known.

Attempts were made to image FLA3-Ty1 by STED upon induction of FLA2 RNAi. These attempts were unsuccessful, for technical reasons. However, in cases where the flagellum became partly or completely detached because of "leaky" expression of *FLA2* dsRNA by the p2T7.177 vector in the absence of tetracycline, it was clear that the FLA3-Ty1 signal along the detached flagellum was punctate (Fig. 3.21 C, D). The centre-to-centre distances between adjacent punctae were the same in the detached regions as in the attached regions. This observation suggests that the even spacing of the junctional complexes is retained when the flagellum detaches. It may even indicate that the flagellar "hemidesmosomes" of the junctional complexes can be assembled and spaced correctly in the absence of flagellar attachment, but this is not certain because it is not clear whether the detached regions in Fig. 3.21 (C, D) were initially attached to the cell body.

FLA1BP has likewise been shown to have a punctate distribution along the flagellar membrane (in procyclic-form cells) (Sun et al., 2013). Sun et al. did not measure the distances between punctae, probably because the YFP-FLA1BP signal was not very even, and the resolution of the images was low. Judging from the size of the scale bars on the images shown, it seems that the average distance between adjacent FLA1BP punctae (when the flagellum is detached by FLA1 RNAi) is approximately 400–500 nm. However, the resolution and signal intensity are insufficient to give an accurate measurement.

The apparent similarity between the distributions of FLA3 and FLA1BP further supports the hypothesis that the two proteins are stage-specific paralogues that each play a similar role in flagellar attachment. It thus seems extremely likely that FLA3 must interact with FLA2, just as FLA1BP interacts with FLA1.

3.3.8 Topological organisation of FLA3 in the flagellar membrane

FLA3 was originally identified as a protein present in the tomato-lectin-binding fraction isolated from bloodstream forms of *T. brucei* (Woods et al., 2013). Among this cohort of putative TL-binding proteins, FLA3 attracted particular attention because the primary sequence, as reported in the genome database, indicated the possible presence of two transmembrane domains, one at either terminus. The region of the protein between the two putative TM domains (residues 43–749) is hydrophilic and contains all the possible sites of *N*-glycosylation; it seems, therefore, to be externally disposed. The predicted C-terminal transmembrane region (residues 751–773) is followed by four basic residues, so according to the positive-inside rule (von Heijne, 1992), the C-terminal tail of the protein (45 a.a.) is almost certainly intracellular. However, the topology of FLA3 in the membrane remains unclear, as it has never been established whether the hydrophobic region near the N-terminus really constitutes a transmembrane domain or rather indicates an atypical signal sequence.

Transmembrane domains normally consist of a 15–30 a.a. region with an overrepresentation of hydrophobic residues (Käll et al., 2004). Residues 22–42 of FLA3 certainly fit this criterion (Fig. 3.26). FLA3 may therefore have a large extracellular domain that is flanked by its two TM domains, which are each embedded in the same membrane, as depicted in Fig. 3.27 (Woods et al., 2013).

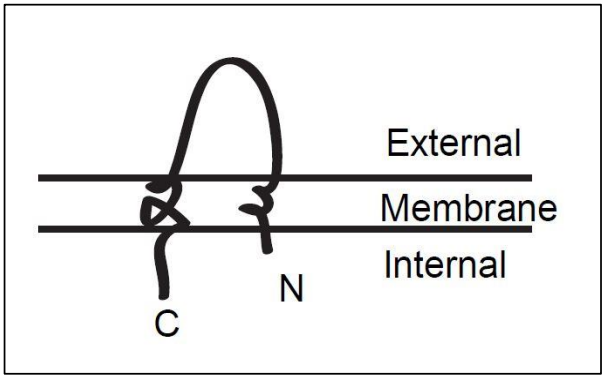


Figure 3.27: A possible topological organisation of FLA3 within the flagellar membrane of the FAZ.

Taken from Woods et al. (2013) Supplementary Figures.

However, the hydrophobic region in question could, alternatively, be a long cleavable signal sequence. The similarity between TM domains and signal peptides is a major problem for topology-prediction software (Käll et al., 2004, Tsirigos et al., 2018). A typical cleavable signal sequence has a hydrophobic region ("h-region") 10–15 residues in length, preceded by a positively-charged region ("n-region") 1–12 residues in length and followed by a somewhat polar region ("c-region") containing small neutral residues that form a cleavage site for the signal peptidase complex (Käll et al., 2004, Guo et al., 2018). Broadly speaking, the N-terminus of FLA3 matches this description, except that the putative h-region is a little long – and, more significantly, the putative n-region is exceedingly long (21 residues).

The length of the putative n-region does not rule out the possibility that FLA3 has a cleavable signal sequence, though. Long ER-targeting signal peptides are known to exist – in fact, an estimated 2% of eukaryotic signal peptides are over 40 residues in length (Hiss and Schneider, 2009), and some are more than 50 residues (Martoglio and Dobberstein, 1998). For example, a human IL-15 precursor has an isoform with a 48-residue signal sequence and murine C4b-binding protein has a 56-residue signal peptide (Hiss and Schneider, 2009). Viral signal peptides can be 100 amino acids in length. Since long signal peptides are extremely difficult to predict *in silico*, they may be more abundant than is currently known. In some cases, long signal peptides have been shown to have signalling functions after cleavage from the mature protein, but this is not necessarily true of all long signal peptides (Hiss and Schneider, 2009). The n-region contributes most to the variability in length of signal peptides (Martoglio and Dobberstein, 1998). It is quite possible, therefore, that FLA3 has a long signal peptide (cleavage would probably occur after the proline at residue 44) rather than an N-terminal TM domain, meaning that the mature protein has single-pass topology.

If FLA3 has indeed a long atypical cleaved signal peptide, there is an intriguing possibility that the signal peptide serves to prevent premature binding of FLA3 to its putative interaction partners prior to the transport of the protein to the FAZ. Since flagellar membrane FAZ proteins and pellicular-membrane FAZ proteins must enter the FAZ on opposite sides of the intermembrane gap, it would seem essential that they do not bind to one another during their passage through the exocytic pathway. If FLA3 were to form a

tight interaction with, for example, FLA2, immediately upon translation in the endoplasmic reticulum, it is difficult to see how the two proteins could be subsequently separated and trafficked to their respective locations correctly. A long signal peptide could conceivably block any premature interactions between FLA3 and its binding partners, possibly by acting as a transient transmembrane domain. The peptide would then have to be cleaved off once FLA3 was separated from the cell-body cohort, perhaps upon entry to the flagellum. Prevention of premature binding is not among the usual functions of signal peptides (Owji et al., 2018), but it would not be entirely without precedent. A recent investigation of the mammalian PAR2 protein (protease-activated receptor 2) has shown that the protein's signal peptide seems to prevent premature interaction with intracellular proteases in the ER or Golgi apparatus (Liu et al., 2020). It is perhaps worth considering that the putative atypical signal peptide of FLA3 could have an atypical function of this sort.

The signal sequences of a given species have a characteristic set of motifs ("h-motifs") that are recognised by the Sec61 translocon, facilitating entry to the ER (Duffy et al., 2010). The putative signal peptide of FLA3 contains eight of the known *T. brucei* h-motifs (Fig. 3.28), confirming that this region of the protein acts an ER-targeting signal. However, this observation does not resolve the signal peptide vs transmembrane domain debate. If the N-terminal hydrophobic region of FLA3 is a TM domain, it must be a Type II signal anchor. As such, it would be expected to possess h-motifs in the same way as a cleavable Type I signal peptide.

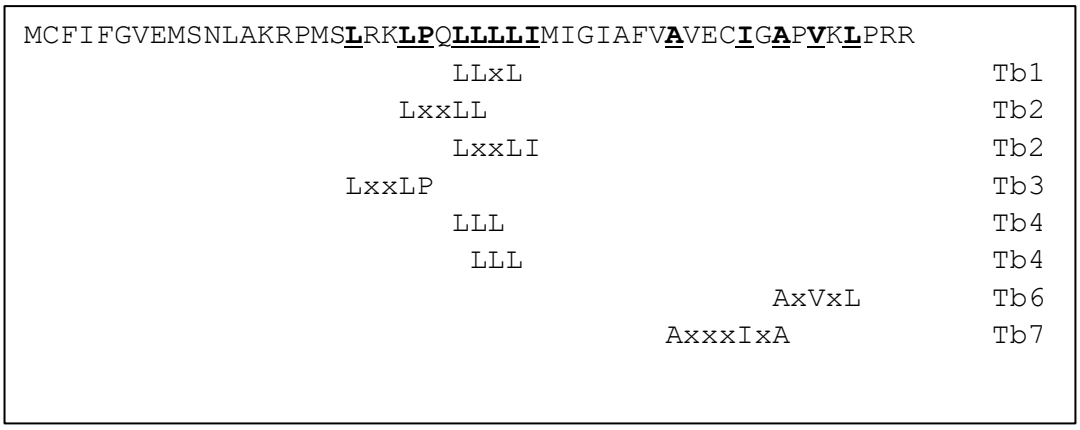


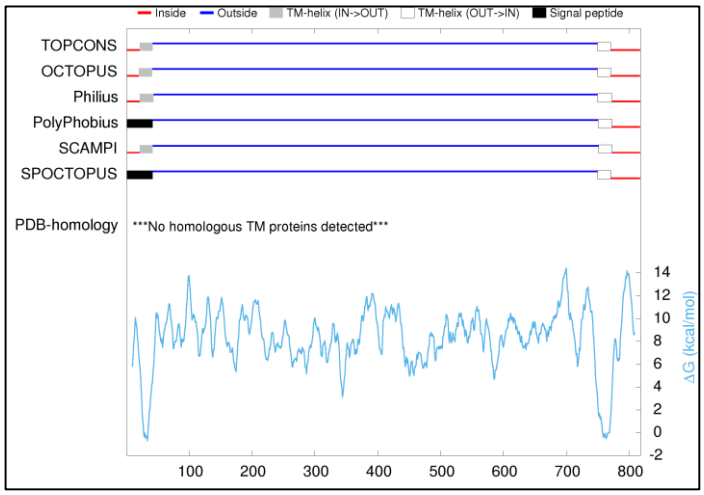
Figure 3.28: The putative signal peptide of FLA3 contains eight h-motifs that can be recognised by the Sec61 translocon in *T. brucei*, facilitating entry to the endoplasmic reticulum.

Upper line: residues 1–50 of FLA3.

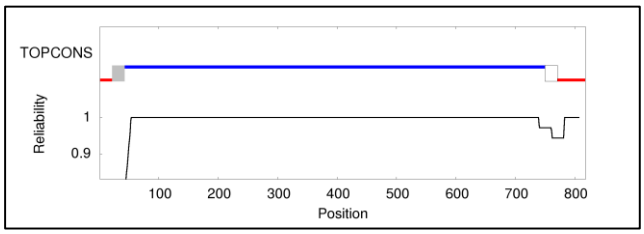
Annotations on the right-hand side refer to the names of the *T. brucei* h-motifs as described by Duffy et al. (2010).

The PCF homologue of FLA3, FLA1BP, likewise has an N-terminal hydrophobic region that is annotated in the TriTryp database as being a TM domain. However, this hydrophobic region has been interpreted to be a cleavable signal sequence (Sun et al., 2013). Sun et al. added a YFP tag to the N-terminal region of FLA1BP, immediately after this putative signal sequence. The tag was subsequently confirmed to localise on the extracellular surface of the cell, but this would be the case whether the hydrophobic region is a signal sequence or a TM domain. It is very probable that FLA3 and FLA1BP possess the same membrane topology, although there are rare cases in which paralogous proteins differ in this respect. In plants, for example, there are two closely related cytochrome P450 enzymes that have different topology in the ER (one TM vs two TM), and which therefore have distinct roles despite possessing the same enzymatic activity (Tsirigos et al., 2018).

A



B



C

Predicted signal peptide and TM-helix positions (position starting from 1):

TOPCONS	**TM1:** 22-42,	**TM2:** 750-770
OCTOPUS	**TM1:** 21-41,	**TM2:** 750-770
Philius	**TM1:** 22-43,	**TM2:** 750-772
PolyPhobius	**SP:** 1-43,	**TM1:** 751-772
SCAMPI	**TM1:** 22-42,	**TM2:** 752-772
SPOCTOPUS	**SP:** 1-43,	**TM1:** 750-770
PDB-homology	***No homologous TM proteins detected***	

Sequence and predicted topologies: (i: inside the membrane, o: outside of the membrane, M: membrane region)

```
                1                                           41
Seq.            MCFIFGVEMS NLAKRPMSLR KLPQLLLLIM IGIAFVAVEC IGAPVKLPRR
TOPCONS         iiiiiiiiii iiiiiiiiii iMMMMMMMMM MMMMMMMMM MMooooooooo
OCTOPUS         iiiiiiiiii iiiiiiiiii MMMMMMMMM MMMMMMMMM Mooooooooo
Philius         iiiiiiiiii iiiiiiiiii iMMMMMMMMM MMMMMMMMM MMMooooooo
PolyPhobius     SSSSSSSSSS SSSSSSSSSS SSSSSSSSSS SSSSSSSSSS SSooooooooo
SCAMPI          iiiiiiiiii iiiiiiiiii iMMMMMMMMM MMMMMMMMM MMooooooooo
SPOCTOPUS       SSSSSSSSSS SSSSSSSSSS SSSSSSSSSS SSSSSSSSSS SSooooooooo
PDB-homology
```

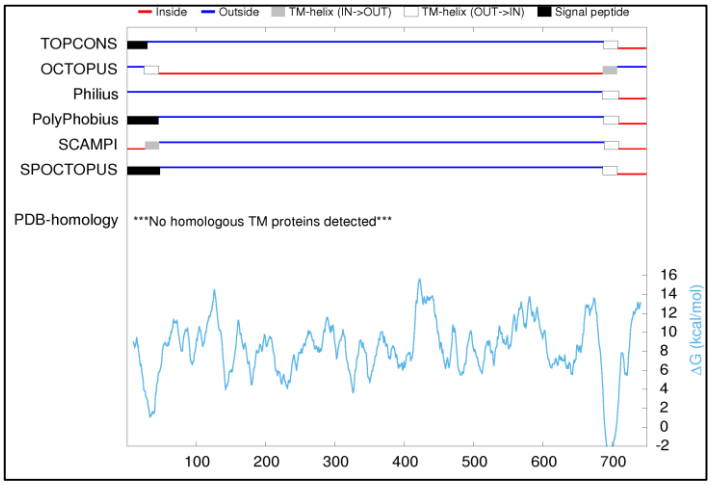
Figure 3.29: TOPCONS topology predictions for FLA3

(A) Topology predictions by the prediction programs on the TOPCONS web server, and a hydrophobicity plot of FLA3.

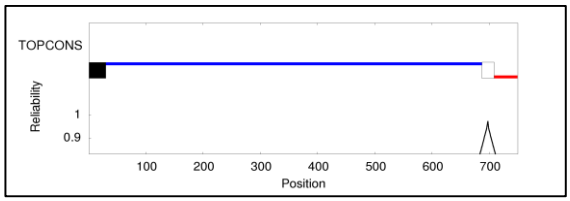
(B) The TOPCONS prediction program does not assign high confidence to the prediction that the N-terminal region of FLA3 has a TM domain.

(C) The topology predictions made by different prediction programs are not in complete consensus regarding the N-terminus of FLA3.

D



E



F

Predicted signal peptide and TM-helix positions (position starting from 1):

TOPCONS	**SP:** 1-31,	**TM1:** 688-708
OCTOPUS	**TM1:** 26-46,	**TM2:** 687-707
Philius	**TM1:** 686-709	
PolyPhobius	**SP:** 1-47,	**TM1:** 688-709
SCAMPI	**TM1:** 27-47,	**TM2:** 689-709
SPOCTOPUS	**SP:** 1-49,	**TM1:** 687-707
PDB-homology	***No homologous TM proteins detected***	

Sequence and predicted topologies: (i: inside the membrane, o: outside of the membrane, M: membrane region)

```
                1                                                   41
Seq.            MPLWKQTNCE VETMNVREVV GTVHLGYVSQ MLLLVATVAT IVVRSGAAPI
TOPCONS         SSSSSSSSSS SSSSSSSSSS SSSSSSSSSS oooooooooo oooooooooo
OCTOPUS         oooooooooo oooooooooo oooooMMMMM MMMMMMMMMM MMMMMMiiii
Philius         oooooooooo oooooooooo oooooooooo oooooooooo oooooooooo
PolyPhobius     SSSSSSSSSS SSSSSSSSSS SSSSSSSSSS SSSSSSSSSS SSSSSSoooo
SCAMPI          iiiiiiiiii iiiiiiiiii iiiiiiMMMM MMMMMMMMMM MMMMMMMooo
SPOCTOPUS       SSSSSSSSSS SSSSSSSSSS SSSSSSSSSS SSSSSSSSSS SSSSSSSoo
PDB-homology
```

Figure 3.29 cont'd: TOPCONS topology predictions for FLA1BP

(A) Topology predictions by the prediction programs on the TOPCONS web server, and a hydrophobicity plot of FLA1BP.

(B) The TOPCONS prediction program does not assign high confidence to the prediction of the topology of FLA1BP.

(C) The topology predictions made by different prediction programs are not in consensus regarding the N-terminus of FLA1BP.

Up-to-date membrane topology prediction algorithms do not resolve the ambiguity (Fig. 3.29). Some programs describe the N-terminal hydrophobic region of FLA3/FLA1BP as a signal peptide; others as a TM domain. The TOPCONS algorithm is designed to distinguish signal peptides from transmembrane domains (Tsirigos et al., 2015); it predicts that the hydrophobic region in question is a TM domain. However, even the best prediction methods have limited accuracy (Tsirigos et al., 2018), so it cannot be stated conclusively whether FLA3 or FLA1BP have signal sequences or TM domains at their respective N-termini until specific experimental data are obtained.

It is interesting that if the methionine at residue 17 of the FLA3 protein sequence were used as a translation start site, the resultant protein would have a classical short N-terminal signal peptide (Fig. 3.26). The hydrophobic region at residues 22–42 would be preceded by a short (6-residue) hydrophilic sequence – a typical n-region. However, three transcriptomics studies have used RNA-Seq or spliced leader trapping to identify genes that have different translation start sites to those annotated in the published genome, and none of these studies have suggested any alteration to the predicted open reading frame of FLA3 (Kolev et al., 2010, Siegel et al., 2010, Nilsson et al., 2010). The only way to determine the actual start site with certainty would be to clone and sequence the *FLA3* cDNA. Although *FLA3* was originally identified by expression cloning, the longest cDNA obtained was a ~1.4-kb cDNA fragment encoding the last 344 amino acids of the protein, followed by the first 329 bp of the 3′ UTR (Woods et al., 2013).

An elegant way to determine whether the N-terminus of FLA3 constitutes a signal peptide or TM domain would be to clone the putative signal sequence (residues 1–42), attach it to the N-terminus of a reporter gene such as GFP and place the fusion gene under inducible control in bloodstream-form *T. brucei*. If the peptide acts as a signal sequence, the GFP fluorescence will be observable in the ER and Golgi apparatus only, as the fusion protein will be secreted upon reaching the cell surface (Ho et al., 2006). If, on the other hand, the peptide acts as a transmembrane signal anchor, the GFP should locate to the surface of the cell.

3.3.9 Processing of tagged FLA3

A striking feature of all western-blot analyses of epitope-tagged FLA3 is the specific and highly reproducible detection of a band at ~50 kDa in FLA3-Ty1 samples and ~55 kDa in FLA3-HA samples (Fig. 3.9). The band, which is diffuse and relatively weak, probably represents a breakdown product of FLA3. The fact that it is present in both FLA3-Ty1 and FLA3-HA cell lines seems to suggest that it is not merely the product of an off-target gene that was tagged unintentionally. Indeed, the length of the homologous ends used to insert the tagging vector specifically to the *FLA3* gene (674 bp at the 5′ end and 521 bp at the 3′ end, see Section 2.2.7) makes it extremely unlikely that off-target tagging could have occurred. (Furthermore, the band was also seen when FLA3 was tagged using an entirely different tagging method, discussed in Section 4.3.7.)

The fact that this protein band retains the epitope tag means that it must be derived from the C-terminal end of FLA3. The band is approximately one third of the size of the mature protein. However, it is not possible to determine the possible cleavage site in the mature protein based on this size information because the positions and sizes of the various *N*-glycan side chains (of which there may be up to 29) are not known.

During the initial studies of the wildtype FLA3 protein (Woods et al., 2013), immune serum was raised against a peptide consisting of the C-terminal 18 amino acids of the FLA3 protein and subsequently used to probe whole-cell lysates. Extensive nonspecific reaction was observed but, interestingly, when the 18-mer peptide antigen was included with the immune serum, two protein bands were lost (one at ~160 kDa and one at ~45 kDa, in addition to a weak smear of proteins between 45 and 55 kDa). It seems possible that the 50/55-kDa bands observed in samples of FLA3-Ty1 and FLA3-HA could represent the same fragment as that ~45-kDa band. The differences in molecular weight would be due to the presence of the respective epitope tags.

Curiously, the 50/55-kDa band is not lost upon conditional knockdown of FLA3, even 72 h after induction of *FLA3* dsRNA by tetracycline (Fig. 3.17). While the absence of a detectable loss of this band might suggest that it is an off-target protein, a more likely possibility is that the 50/55-kDa band represents a relatively stable intermediate in the breakdown of FLA3. During FLA3 RNAi, gradual loss of FLA3 protein occurs due to normal turnover of the protein. However, FLA3 must be a very stable protein (as demonstrated by

the fact that the old flagellum remains attached during FLA3 RNAi). If there is a rate-limiting step in the degradation of FLA3, it is plausible that the level of a stable intermediate would remain constant even several days after induction of FLA3 RNAi. Furthermore, it is possible that the C-terminal epitope tag artificially stabilises this degradation product.

A formal but unlikely possibility is that the 50/55-kDa band represents a truncated isoform of FLA3, produced by alternative *trans*-splicing of the *FLA3* mRNA within the coding sequence, that escapes RNAi degradation because it lacks the region of the *FLA3* gene (residues 1174–1792) employed in the RNAi experiment. Such a putative short transcript would be expressed at its normal rate while the full-size FLA3 protein was knocked down. This scenario would, formally, explain the observed results. However, a recent study has shown that alternative *trans*-splicing is not used in the trypanosome to produce different protein isoforms (Soulette et al., 2019). Alternative splice acceptor sites (SAS) are certainly used to produce transcripts with different 5′ UTRs (Siegel et al., 2011), but to date the only example of functional alternative splicing in the coding sequence of a gene is the case of isoleucyl-tRNA synthetase (Rettig et al., 2012). An SAS within the isoleucyl-tRNA synthetase gene in *T. brucei* is used to produce a functional isoform of the protein that lacks a mitochondrial-targeting signal peptide. The discovery that there are many SAS within coding sequences in the trypanosome genome (Siegel et al., 2010) led to the hypothesis that alternative *trans*-splicing could be used to diversify the proteome (Siegel et al., 2011). Ribosomal profiling suggests that this is not the case – however, alternative *trans*-splicing may lead to the production of pervasive unwanted truncated transcripts that must be degraded if translated (Soulette et al., 2019).

Could the 50/55-kDa band represent such a truncated isoform, stabilised perhaps by the presence of the epitope tag? This question hinges on whether the *FLA3* gene has any SAS within the CDS. Nilsson (2010) found two such SAS (although Siegel (2010) and Kolev (2010) did not). If translation were to occur from the next ATG downstream of these SAS (residue 693 of the FLA3 protein sequence), the resulting peptide would be ~14-kDa, with just two putative *N*-glycosylation sites. This hypothetical peptide would be much too small to be responsible for the 50/55-kDa band observed on western blots here.

On balance, then, it seems most likely that the ~50/55-kDa band observed here must represent a stable intermediate produced during normal turnover of the mature FLA3 protein, but possibly stabilised by the epitope tag. The reason it is not lost upon ablation of *FLA3* mRNA is because there is sufficient extant FLA3 protein in the cell to maintain the normal steady-state level of proteolysis. Since this putative stable breakdown product fractionates as a membrane protein (Fig. 3.10), it probably locates to the endosomal/lysosomal compartments. An interesting experiment would be to block the proteolytic pathway using protease inhibitors. Subsequent loss of the ~50/55-kDa band would confirm that the band represents a degradation product.

Pulse-chase labelling would be an ideal way to demonstrate that the 50/55-kDa band represents a degradation product. In this method, the cells would be incubated with ^{35}S methionine/cysteine for 5–10 min and then transferred to "cold" medium. By immunoprecipitating FLA3-Ty1 at progressive time intervals and subsequently probing for the radioactive label, the processing of FLA3 over time could be visualised. At first, the protein would be seen maturing to the 160-kDa form. The mature protein would remain stable before being gradually degraded. The 50-kDa band would be observed at these later timepoints only, if it is indeed a FLA3 degradation product.

Further insights into the 50/55-kDa band are described in Section 4.3.7.

3.3.10 Expressing FLA3-Ty1 in procyclic cells

The FLA3 protein is bloodstream-stage specific (Woods et al., 2013). Although trace amounts of *FLA3* mRNA occur in procyclic-form cells, expression of *FLA3* dsRNA has no effect on growth or morphology of the cells (Woods, 2007). (Conversely, expression of *FLA1BP* dsRNA has no effect on growth or morphology of BSFs.)

To investigate whether FLA3 would localise to the FAZ if expressed in procyclic cells, *FLA3* was endogenously tagged with Ty1 at the C-terminus in wildtype PCFs. Western blotting confirmed the FLA3-Ty1 protein was expressed (Fig. 3.22), which was interesting because FLA3 is not ordinarily expressed in PCFs. The tagging vector changed the 3′ UTR to that of the constitutive *RPA1* gene. This change of UTR was evidently sufficient to completely alter the expression profile of FLA3. It must be the case that regulatory motifs in the *FLA3* 3′ UTR determine the stage-specificity of the FLA3 protein. This is as expected because the 3'

UTR is known to bind the regulatory proteins that mediate the regulated pathway of mRNA degradation and that control the translational efficiency of a transcript (Section 1.4.2).

The main FLA3-Ty1 band in PCFs was ~160 kDa, the same size as in BSFs, which suggests that post-translational modification occurred to the same extent as in BSFs. Interestingly, several other bands were also seen (Fig. 3.22), at 70, 50, 40 and 25 kDa, respectively. These probably represent C-terminal breakdown products of FLA3. The 50-kDa breakdown-product band appears as a double band in PCFs, whereas it was a single band in BSF samples (Section 3.3.9). The presence of extra breakdown-product bands indicates that when expressed in procyclic cells, FLA3 is not stable and is degraded.

When these cells were probed by immunofluorescence (Fig. 3.23), the FLA3-Ty1 signal was confined to the posterior end of the cell, particularly to a specific region between the nucleus and kinetoplast. This localisation could indicate that FLA3-Ty1 is targeted to the endosomal/lysosomal compartments, which would be consistent with the earlier observation of multiple breakdown bands. Alternatively, FLA3-Ty1 may enter the flagellar pocket, a possibility that could be tested by biotinylating the surface of the cell (including the pocket), isolating FLA3-Ty1 by immunoprecipitation and subsequently probing with streptavidin.

The immunolocalization data indicate that FLA3-Ty1 does not enter the flagellar attachment zone when expressed in procyclic cells, possibly because of the absence of stage-specific interacting proteins or because of differences in the trafficking pathway. The apparent degradation of FLA3-Ty1 may also be because of the absence of interacting proteins. It is often the case that when one member of a protein complex is missing, the other members of the complex also get degraded – for example, FAZ1 and FAZ8 get degraded when FAZ2 is knocked down by RNAi (Zhou et al., 2015).

3.3.11 Summary of Chapter 3

The important novel findings presented in this chapter are the confirmation that FLA2 is essential for flagellar attachment and cytokinesis in bloodstream-form cells, the discovery that FLA3 is located in the flagellar membrane of the FAZ rather than the cell-body membrane, and the demonstration that tagged FLA3 exhibits a regularly-spaced punctate distribution along the FAZ.

Chapter 4: Interaction Studies and Glycosylation Studies of FLA3

4.1 Introduction

As a flagellar-adhesion protein, FLA3 must bind to other proteins in the flagellar attachment zone in order to function. Immunoprecipitation is a useful way to identify the binding partners of a protein of interest. For example, FLA1BP was discovered by co-immunoprecipitation (Co-IP) with FLA1 (Sun et al., 2013) and the FAZ proteins FA2–7 were discovered through Co-IP with FAZ1 and with each other (Sunter et al., 2015b). As a preliminary step towards identifying the putative binding partners of FLA3, a Ty1-tagged version of the protein was therefore immunoprecipitated (Section 4.2.1).

To see what proteins, if any, had co-precipitated with FLA3, three methods were used to probe the IP samples: silver staining (Section 4.2.1), lectin blotting (Section 4.2.4) and biotinylation followed by streptavidin probing (Section 4.2.2). Each method gave a different insight into the possible binding partners of FLA3.

Furthermore, the biotinylation experiment demonstrated that FLA3 is expressed on the surface of the trypanosome cell and is accessible to small molecules (Section 4.2.2), which is consistent with the finding in Chapter 3 that FLA3 locates to the punctate junctional complexes that span the intermembrane region of the FAZ. As a vital component of the junctional complexes, FLA3 could plausibly act as a dimer or as an oligomer. Expression of two tagged versions of FLA3 simultaneously enabled investigation of the hypothesised ability of FLA3 to bind to itself (Section 4.2.5).

Perhaps the most striking attribute of FLA3 is its extensive glycosylation. The carbohydrate side chains of FLA3, which comprise nearly half the molecular weight of the mature protein, are likely to play an immune-evasive role by concealing the protein from the adaptive immune response of the host. Given the extent of glycosylation, however, it also seems reasonable to suppose that these glycans are important to the function of FLA3. It has even been suggested that carbohydrate-carbohydrate interactions between FAZ glycoproteins may be fundamental to the maintenance of flagellar adhesion (Woods et al., 2013). Nonetheless, the glycans of FLA3 have never been characterised.

As a preliminary step towards identifying the glycans of FLA3, the protein was tested for binding to two lectins: tomato lectin and RCA-I. Tomato lectin binds to poly-*N*-acetyllactosamine (poly-LacNAc) when this disaccharide is present in three or more linear

repeats (Merkle and Cummings, 1987); RCA-I binds to galactose (Wu et al., 2006). Both lectins were found to recognise FLA3 (Sections 4.2.3, 4.2.4). This would seem to be consistent with the hypothesis that FLA3 belongs to a group of large poly-LacNAc glycoproteins associated with endocytosis and the flagellar pocket in BSF cells (Atrih et al., 2005), discussed in Section 4.3.4.

4.2 Results

4.2.1 Immunoprecipitation of tagged FLA3

Immunoprecipitation is a widely used method to investigate protein-protein interactions. The results presented in Chapter 3 demonstrated that *in-situ* tagged FLA3-Ty1 locates to the FAZ and appears to be subject to normal post-translational processing. Therefore, experiments were performed to determine whether Ty1-tagged FLA3 could be immunoprecipitated, as a first step towards identifying the putative binding partners of FLA3.

Immunoprecipitation experiments were performed using whole-cell lysates prepared under denaturing (SDS) and nondenaturing (Triton X-100) conditions. These lysates, along with the bound and nonbinding fractions obtained following immunoprecipitation with mouse anti-Ty1, were subject to western blot analysis using anti-Ty1. The results presented in Fig. 4.1 demonstrate successful immunoprecipitation of FLA3-Ty1 from both lysates.

In both cases, a band of ~160 kDa was obtained in the bound fraction that co-migrated with the FLA3-Ty1 band in the total lysate. This band was not detectable in the unbound/nonbinding fraction, indicating that all the tagged protein had been immunoprecipitated.

It was not possible to determine whether the ~50-kDa FLA3-Ty1 degradation product was immunoprecipitated. It was present in each total-lysate sample, but the corresponding region of the bound-fraction lane was obscured by the intense band of the anti-Ty1 heavy chain.

Under denaturing conditions, a ~110-kDa band was observed in the bound fraction. This may represent a minor degradation product of FLA3-Ty1.

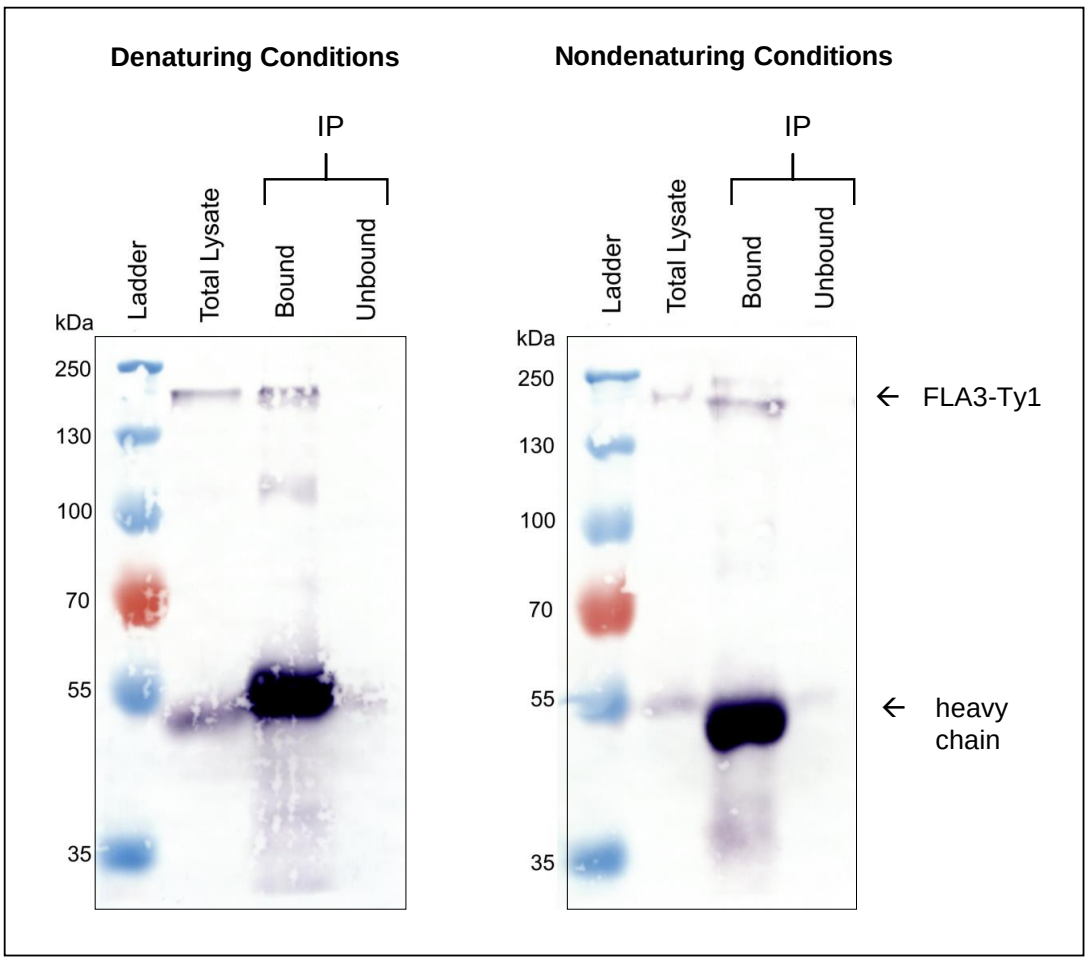


Figure 4.1: Immunoprecipitation of FLA3-Ty1 from detergent lysate.

Detergent-lysed cells expressing FLA3-Ty1 were subjected to immunoprecipitation (IP) using anti-Ty1 and Protein G agarose, as described in the Methods section. The immunoprecipitate samples were probed with anti-Ty1. The blots show that FLA3-Ty1 was successfully immunoprecipitated.

(Note that the bound samples each represent $1x10^7$ cells, while the lysate and unbound samples each represent $2x10^6$ cells under denaturing conditions and $7x10^6$ cells under nondenaturing conditions.)

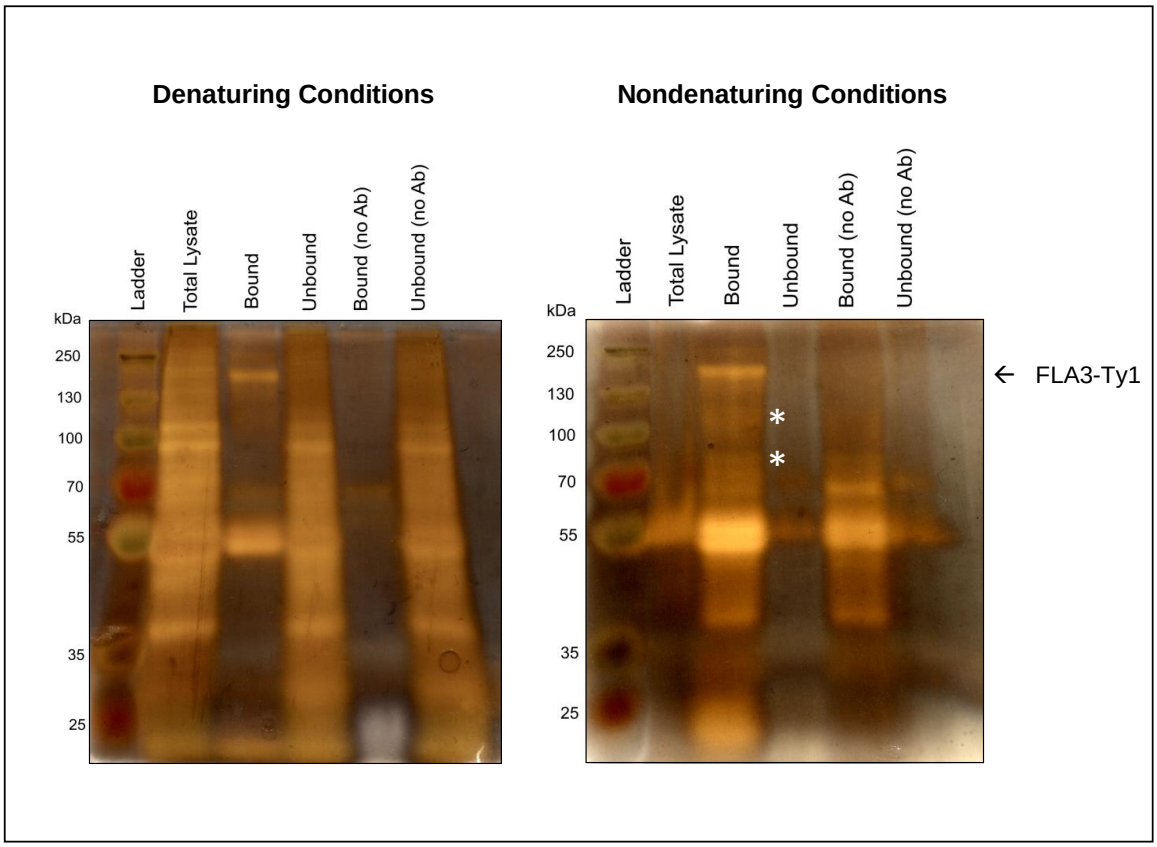


Figure 4.2: Silver Staining confirms immunoprecipitation of FLA3-Ty1.

The samples from the immunoprecipitation described in Fig. 1 were silver-stained to reveal all the proteins present. The "no Ab" samples are controls in which no antibody was added to the Protein G. A ~160-kDa band indicates the presence of FLA3-Ty1 in the “Bound” fractions under denaturing and nondenaturing conditions.

(The bound samples each represent $1x10^7$ cells, while the unbound samples represent $2x10^6$ cells under denaturing conditions and $7x10^6$ cells under nondenaturing conditions.)

Having demonstrated that FLA3-Ty1 could be immunoprecipitated, the search for putative binding partners of FLA3 could commence. To determine whether any proteins had co-precipitated with FLA3-Ty1, the samples from the immunoprecipitation were silver-stained to reveal all proteins present (Fig. 4.2). Under denaturing conditions, FLA3-Ty1 should immunoprecipitate alone, because all noncovalent interactions are broken; under nondenaturing conditions, its binding partners should be pulled down also. Therefore, comparison of the bands obtained under the two lysis conditions should reveal which ones correspond to binding partners.

Under denaturing conditions, three major bands were observed in the bound fraction on the silver stain: one at ~160 kDa, which probably corresponds to FLA3-Ty1; one at ~55 kDa, probably representing the anti-Ty1 heavy chain; and one at ~20 kDa, probably representing the anti-Ty1 light chain. As expected, there were no other significant bands in the bound fraction. There were some very weak bands, but these had evidently precipitated due to non-specific interaction with the Protein G beads, as they were also present in a control sample ("Bound (no Ab)") in which no anti-Ty1 was present.

Under nondenaturing conditions, the same three major bands were present in the bound fraction on the silver stain: FLA3-Ty1 (~160 kDa), anti-Ty1 heavy chain (~55 kDa) and anti-Ty1 light chain (~20 kDa). Again, there were also some weak nonspecific bands that were also present in the control sample ("Bound (no Ab)") in which no anti-Ty1 was present.

Critically, however, there appeared to be several weak bands in the bound fraction that were not present in the control sample. The bands in question (~115 kDa, ~110 kDa and ~80 kDa, respectively) are highlighted with asterisks on Fig. 4.2. These proteins may have co-precipitated with FLA3-Ty1 because they have specific noncovalent interactions with FLA3. In other words, these weak bands may represent binding partners of FLA3.

4.2.2 Surface Biotinylation of FLA3-Ty1

A crucial question concerning FLA3 is whether the protein is accessible to small molecules. To test this, intact live cells expressing FLA3-Ty1 were surface-labelled with membrane-impermeant NHS-biotin (see Methods 2.3.8). The cells were subsequently lysed with non-denaturing detergent. FLA3-Ty1 was immunoprecipitated with anti-Ty1 antibody and

subjected to SDS-PAGE. The samples were blotted onto nitrocellulose membrane and probed with streptavidin to reveal all biotinylated proteins (Fig. 4.3 A).

Two biotinylated proteins were seen in the anti-Ty1-bound fraction: a ~160-kDa band corresponding to FLA3-Ty1, and a ~60-kDa band probably corresponding to VSG. The identity of the putative FLA3-Ty1 band was confirmed by reprobing the blot for Ty1: the anti-Ty1 antibody recognised the same ~160-kDa band (Fig. 4.3 B). The fact that FLA3-Ty1 was biotinylated demonstrates that the protein is found on the surface of the cell and is accessible to small molecules, at least *in vitro*.

To verify that only surface proteins had been biotinylated, the blot was reprobed for a known intracellular endoplasmic-reticulum protein – the chaperone, BIP (Fig. 4.3 C). Trace amounts of BIP were present in the immunoprecipitated fraction (possibly because of nonspecific binding, but possibly due to an interaction with nascent FLA3-Ty1), at the expected size of 80 kDa. Significantly, this band was not recognised by streptavidin. This shows that the biotin probing was indeed specific to surface proteins.

The ~60-kDa biotinylated protein in the immunoprecipitated sample is likely to be VSG, although this cannot be stated for certain. VSG (isotype MITat 1.1) usually appears on SDS-PAGE gels as a diffuse ~50-kDa band, but it could easily appear larger when extensively biotinylated. It is not surprising that VSG can be biotinylated, given that it is a surface protein. However, its presence in the anti-Ty1-bound fraction is interesting because it could indicate an interaction between FLA3-Ty1 and VSG. It is also possible, though, that presence of VSG in this fraction is due merely to nonspecific binding to the Protein-G beads. Given the abundance of VSG in the trypanosome, the protein is liable to nonspecific interactions.

Interestingly, the ~50-kDa FLA3-Ty1 degradation product was not recognised by streptavidin (although it was present in the immunoprecipitated fraction (data not shown)). This indicates that the degradation product is not present on the surface of the cell.

The critical finding of the surface-biotinylation experiment was that FLA3-Ty1 is a surface-exposed protein, as was predicted by analysis of its primary structure.

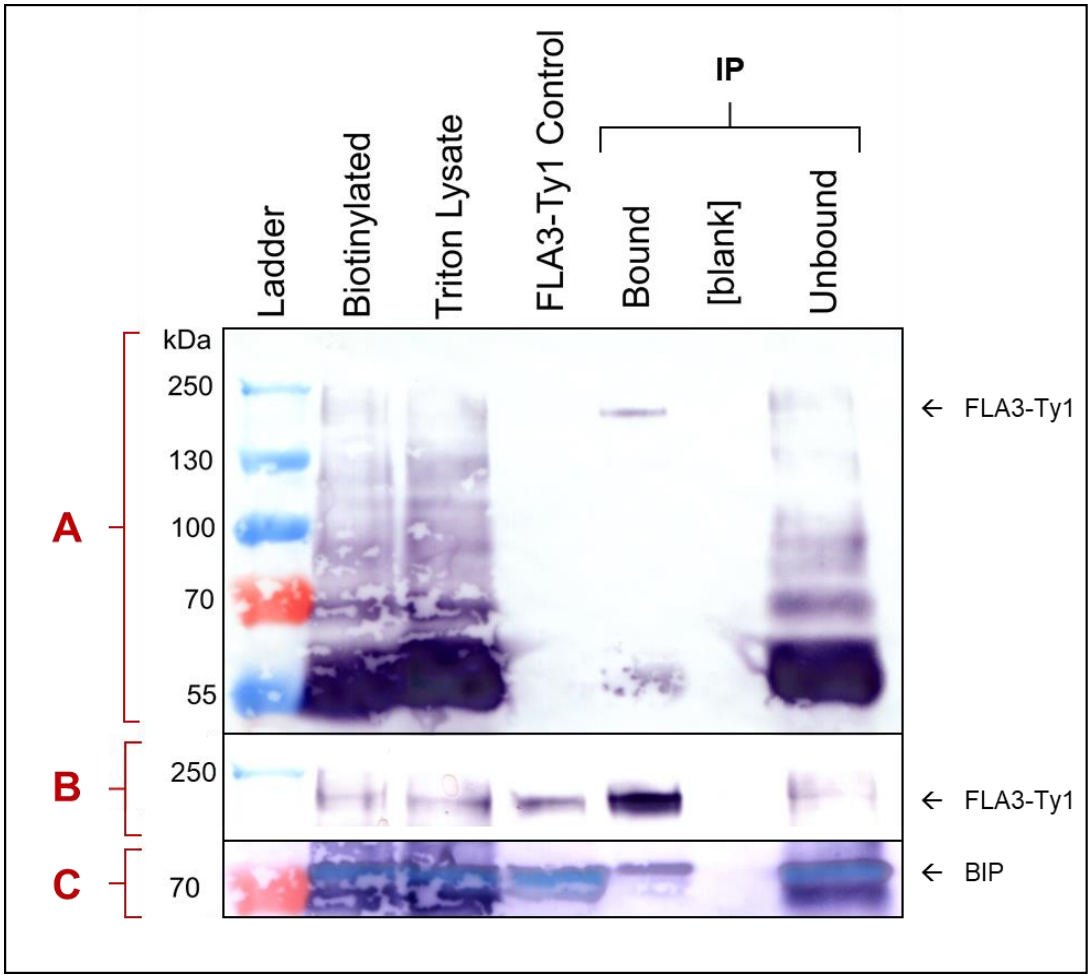


Figure 4.3: Surface biotinylation of FLA3-Ty1.

Intact bloodstream-form cells expressing FLA3-Ty1 were surface-biotinylated and subsequently lysed with Triton X-100. Immunoprecipitation (IP) was performed, using anti-Ty1 and Protein-G agarose, to concentrate the tagged FLA3. The biotinylation and IP samples were probed with streptavidin **(A)** to reveal all biotinylated proteins. The presence of a ~160 kDa band in the bound fraction indicates that FLA3-Ty1 was successfully biotinylated.

The blot was subsequently probed for Ty1 **(B)**, to confirm the identity of the aforementioned band by showing that it bound the anti-Ty1 antibody.

Finally, the blot was probed for an intracellular protein, BIP **(C)**. As expected, BIP was present in all lanes, although it had not been not visible by streptavidin probing. This confirms that the biotinylation was specific to surface proteins.

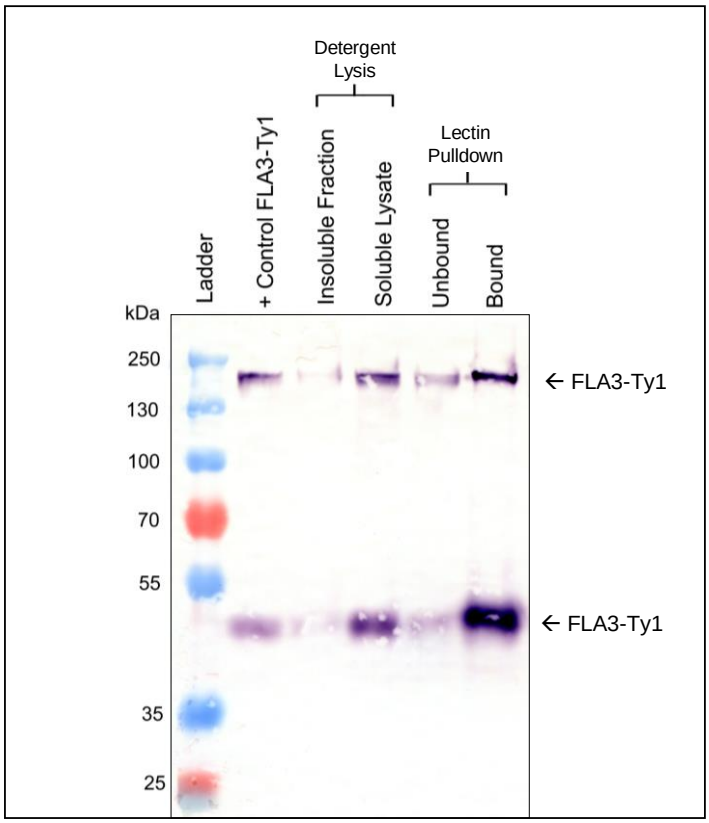


Figure 4.4: Lectin pulldown of FLA3-Ty1.

Western blot probed for Ty1 confirms pulldown of FLA3-Ty1 by RCA-I lectin. Bloodstream-form cells expressing FLA3-Ty1 (p3074; Clone 1) were incubated in Triton X-100 detergent and divided into detergent-insoluble and detergent-soluble fractions. The Soluble Lysate sample was subsequently incubated with lectin beads and fractionated into supernatant (Unbound) and pellet (Bound) fractions. The blot shows that most of the FLA3-Ty1 was found in the bound fraction.

(Note that the bound sample represents $8x10^7$ cells, the lysate and unbound samples represent $5x10^6$ cells and the insoluble fraction represents $1x10^7$ cells.)

4.2.3 Lectin Pulldown of FLA3

As discussed in Section 4.1, FLA3 may belong to a class of proteins modified by *N*-linked linear poly-*N*-acetyllactosamine (poly-LacNAc) side chains. To investigate this hypothesis, FLA3 was tested for binding to RCA-I lectin, a protein that binds specifically to galactose residues.

Bloodstream-form cells expressing FLA3-Ty1 were solubilised with Triton X-100 detergent. The lysate was incubated with RCA-I beads, and fractionated (see Methods). A western blot probed for Ty1 revealed that most of the FLA3-Ty1 was found in the lectin-bound fraction (Fig. 4.4). The experiment was repeated in cells expressing FLA3-HA. Most of the solubilised FLA3-HA was found in the lectin-bound fraction (data not shown).

These results are a strong indication that mature FLA3 contains galactose residues, supporting the hypothesis that the protein has poly-LacNAc side chains.

4.2.4 Investigation of Potential Glycoprotein Binding Partners of FLA3

To investigate the possibility that FLA3 interacts with other glycoproteins in the flagellar attachment zone, FLA3-Ty1 was immunoprecipitated under nondenaturing conditions, as described in Section 4.2.1. The immunoprecipitate samples were lectin blotted with biotinylated RCA-I lectin (Fig. 4.5 A) to see if any glycoproteins with galactose residues had co-precipitated with FLA3-Ty1.

In the immunoprecipitate sample, five bands can be seen, labelled 1–5 (Fig. 4.5 A). Band 2 (~160 kDa) almost certainly represents FLA3-Ty1, indicating that immunoprecipitation was successful. Band 5 (~50–55 kDa) probably represents the stable FLA3-Ty1 breakdown product, which binds strongly to RCA-I, as seen in Fig. 4.4. Band 5 may also contain some signal corresponding to the anti-Ty1 heavy chain, because mouse IgG1 antibodies are known to contain galactose residues (de Haan et al., 2017).

Interestingly, Bands 1, 3 and 4 in the immunoprecipitate sample may represent glycoproteins that co-precipitated with FLA3-Ty1. These three bands (~180 kDa, ~130 kDa and ~60 kDa) were barely visible when the immunoprecipitation was repeated under denaturing conditions (Fig. 4.5 B), which seems to indicate that their presence in the IP

sample was due to a reversible noncovalent interaction with FLA3-Ty1. The identities of these three bands are unknown. None are of the correct size to represent VSG.

To confirm that these three proteins had not been precipitated merely because of non-specific interaction with the Protein-G-agarose beads, a control "immunoprecipitation" was performed without any primary antibody. Under these conditions, no proteins precipitated (Fig. 4.5 C), which implies that the three unidentified bands in Fig. 4.5 (A) precipitated due to a specific interaction with FLA3-Ty1 (or with the anti-Ty1 antibody).

A complementary investigation of the potential glycoprotein binding partners of FLA3 was performed using a different lectin – tomato lectin, which binds to linear *N*-acetyllactosamine repeats. As before, FLA3-Ty1 was immunoprecipitated under nondenaturing conditions. The immunoprecipitate samples were lectin blotted with biotinylated tomato lectin (Fig. 4.6 A), to see if any glycoproteins with *N*-acetyllactosamine residues had co-precipitated with FLA3-Ty1.

A faint band at ~160 kDa indicated the presence of FLA3-Ty1. This suggests that FLA3 binds to tomato lectin, but not as strongly as it binds to RCA-I. This is the first indication that FLA3 interacts directly with tomato lectin. No other clear bands were revealed by tomato lectin in the immunoprecipitate sample, suggesting that FLA3 does not interact with other glycoproteins that have linear *N*-acetyllactosamine residues. There appeared to be an extremely faint band at ~180 kDa, but it was too indistinct to be certain.

To verify the identity of the ~160-kDa band, the blot was reprobed with anti-Ty1 (Fig. 4.6 B). A strong Ty1 band superimposed exactly over the previous faint band, confirming it to be FLA3-Ty1. Two other bands were also seen: one at ~55 kDa, representing the anti-Ty1 heavy chain, and one at ~50 kDa, representing the stable FLA3-Ty1 breakdown product. As the latter band was not observed when the blot was initially probed with tomato lectin, it seems that the FLA3-Ty1 breakdown product lacks the glycans that are recognised by tomato lectin. It is probable, therefore, that these glycans are attached to the large N-terminal region of FLA3 that is missing from the 50-kDa degradation product.

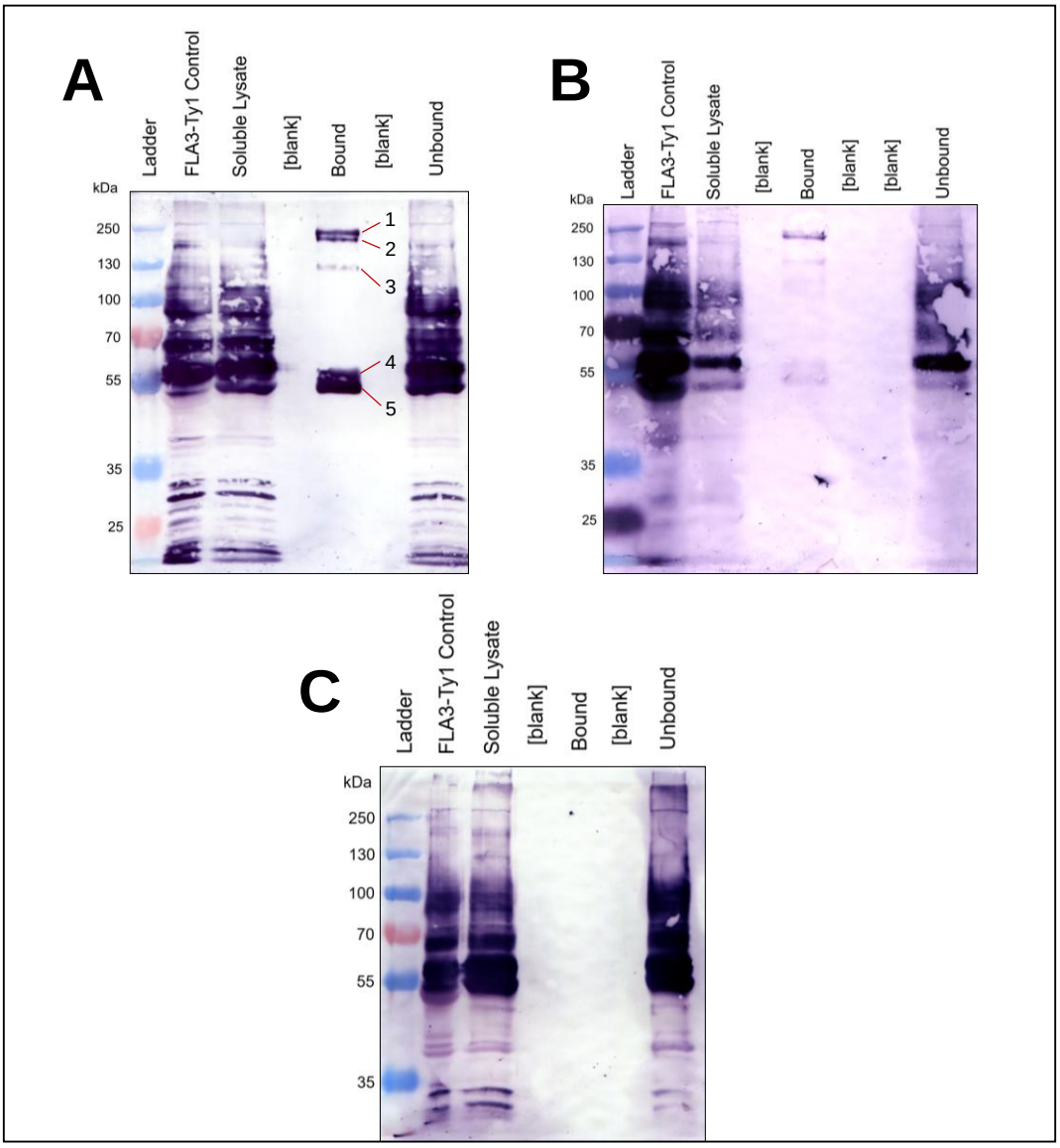


Figure 4.5: Investigation of potential glycoprotein binding-partners of FLA3. Detergent-lysed cells expressing FLA3-Ty1 (p3074; Clone 1) were subjected to immunoprecipitation using anti-Ty1 and Protein G agarose. The immunoprecipitated samples were probed with RCA-I lectin to reveal their glycoprotein content.

(A) In nondenatured lysates (see Methods), FLA3-Ty1 was found in the α-Ty1 antibody-bound fraction. Also present were two unidentified bands, which were greatly depleted **(B)** when the experiment was repeated in denatured lysates.

(C) A control experiment, in which no α-Ty1 antibody was used, showed no nonspecific bands in the Protein-G-bound fraction.

(In all blots, the bound sample represents $2x10^8$ cells.)

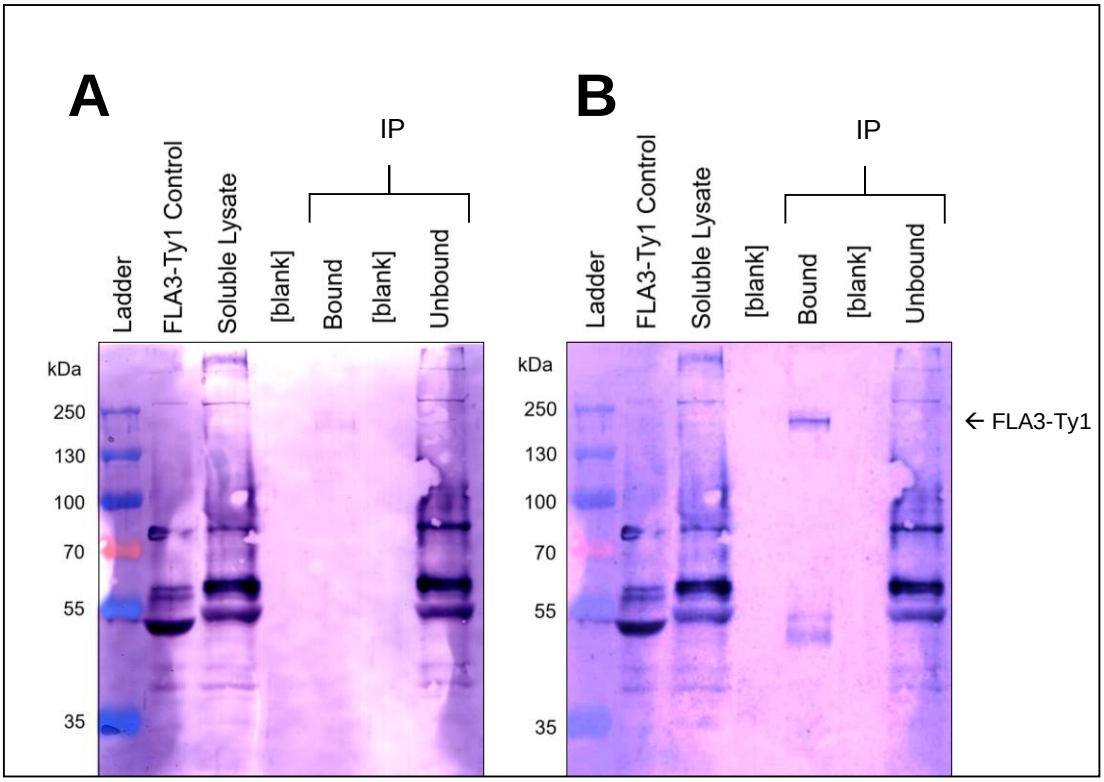


Figure 4.6: FLA3-Ty1 can bind to tomato lectin.

Detergent-lysed cells expressing FLA3-Ty1 (p3074; Clone 1) were subjected to immunoprecipitation using anti-Ty1 and Protein G agarose. The immunoprecipitated samples were probed with biotinylated tomato lectin and, subsequently, with streptavidin.

(A) Upon streptavidin probing, a weak band at the expected size of FLA3-Ty1 was observed in the α-Ty1 antibody-bound fraction. **(B)** The blot was reprobed for Ty1. The FLA3-Ty1 band colocalised with the existing weak tomato-lectin band, suggesting that it was indeed FLA3-Ty1 that had bound the tomato lectin.

(Note that the bound sample represents $2x10^8$ cells, while all other samples represent $5x10^6$ cells.)

4.2.5 Investigating the Possible Oligomerisation of FLA3

To test for possible dimerisation or oligomerisation of FLA3, long-primer PCR tagging with the pPOTv6 vector (see Methods) was used to add a 3×Ty1 tag to the C-terminal end of FLA3, at the endogenous locus, in cells already expressing FLA3-HA (p2708). As there are two endogenous *FLA3* genes, each with two alleles, there was no difficulty in expressing two tagged forms of the protein simultaneously in one cell line (using different selection markers). Western blots confirmed that the transfected clones expressed both FLA3-Ty1 and FLA3-HA (Fig. 4.7).

Indirect immunofluorescence was performed to confirm that this new version of FLA3-Ty1 localised correctly. Upon probing, Ty1 signal was observed in a distinct line along the flagellar attachment zone, with little or no background signal in the cell body (Fig. 4.8 A). This was the same distribution pattern that had been observed previously when FLA3 was tagged with Ty1 using a more traditional tagging method (Fig. 3.9).

Indirect immunofluorescence also confirmed that FLA3-HA localised to the FAZ in the double-tagged cells (Fig. 4.8 B). The FLA3-Ty1 and FLA3-HA signals colocalised almost exactly (Fig. 4.8 C). This cell line was used to investigate whether there was any interaction between the two tagged forms of the protein. An interaction between FLA3-Ty1 and FLA3-HA would be consistent with the view that FLA3 functions as a dimer or oligomer.

FLA3-Ty1 was immunoprecipitated under non-denaturing conditions as described in Section 4.2.1. The immunoprecipitate samples were blotted onto membrane and probed with an anti-HA antibody (Fig. 4.9). A faint band (~160 kDa) (marked with an asterisk) was seen in the Ty1-antibody-bound fraction, which may represent trace amounts of FLA3-HA. However, a much more intense FLA3-HA band was observed in the nonbinding fraction, demonstrating that most of the FLA3-HA in the cells had not co-precipitated with FLA3-Ty1.

Further investigations are necessary to determine whether FLA3 is monomeric, dimeric or oligomeric.

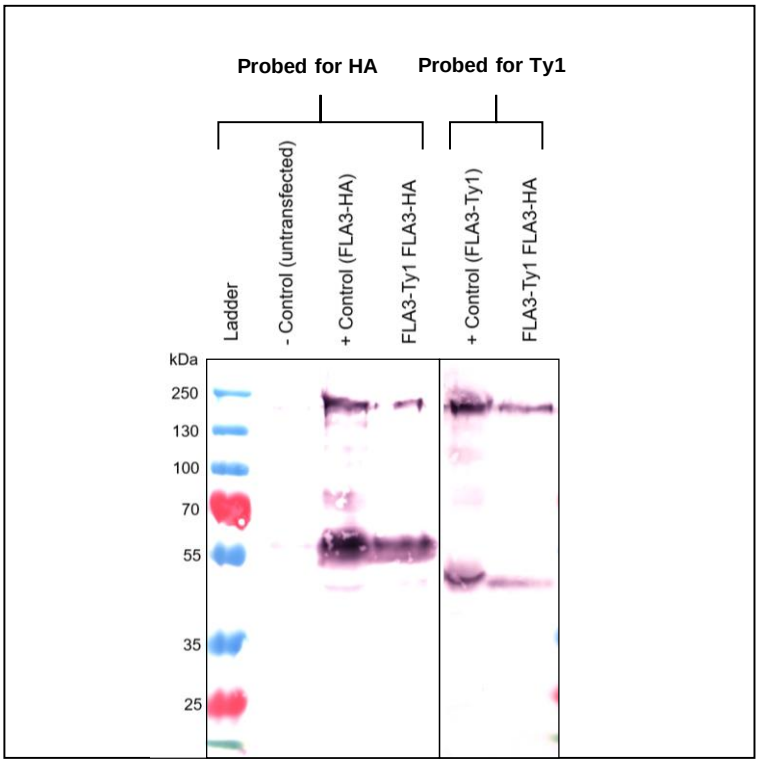


Figure 4.7: Simultaneous expression of FLA3-Ty1 and FLA3-HA.

Long-primer PCR tagging with pPOTv6 was used to add a 3X Ty1 tag to the C-terminal end of FLA3, at the endogenous locus, in bloodstream-form cells already expressing FLA3-HA. Western blots probed for HA and Ty1, respectively, confirm expression of both FLA3-HA and FLA3-Ty1 in a representative clone.

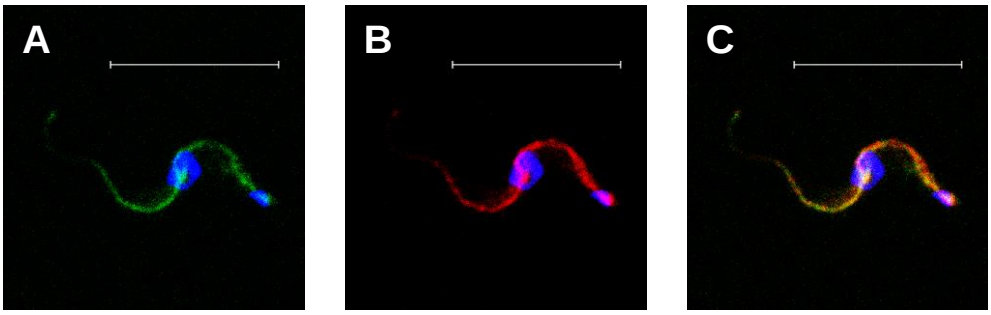


Figure 4.8: Colocalisation of FLA3-Ty1 and FLA3-HA.

Long-primer PCR tagging with pPOTv6 was used to add a 3X Ty1 tag to the C-terminal end of FLA3, at the endogenous locus, in bloodstream-form cells already expressing FLA3-HA.

(A) Upon probing with anti-Ty1 (green), FLA3-Ty1 is observed along the flagellar attachment zone. **(B)** Upon probing with anti-HA (red) in the same clone, FLA3-HA is observed along the flagellar attachment zone. **(C)** Merging the signals shows that FLA3-Ty1 and FLA3-HA colocalise.

Nucleus and kinetoplast are stained in blue. Scale bars are 10 μm.

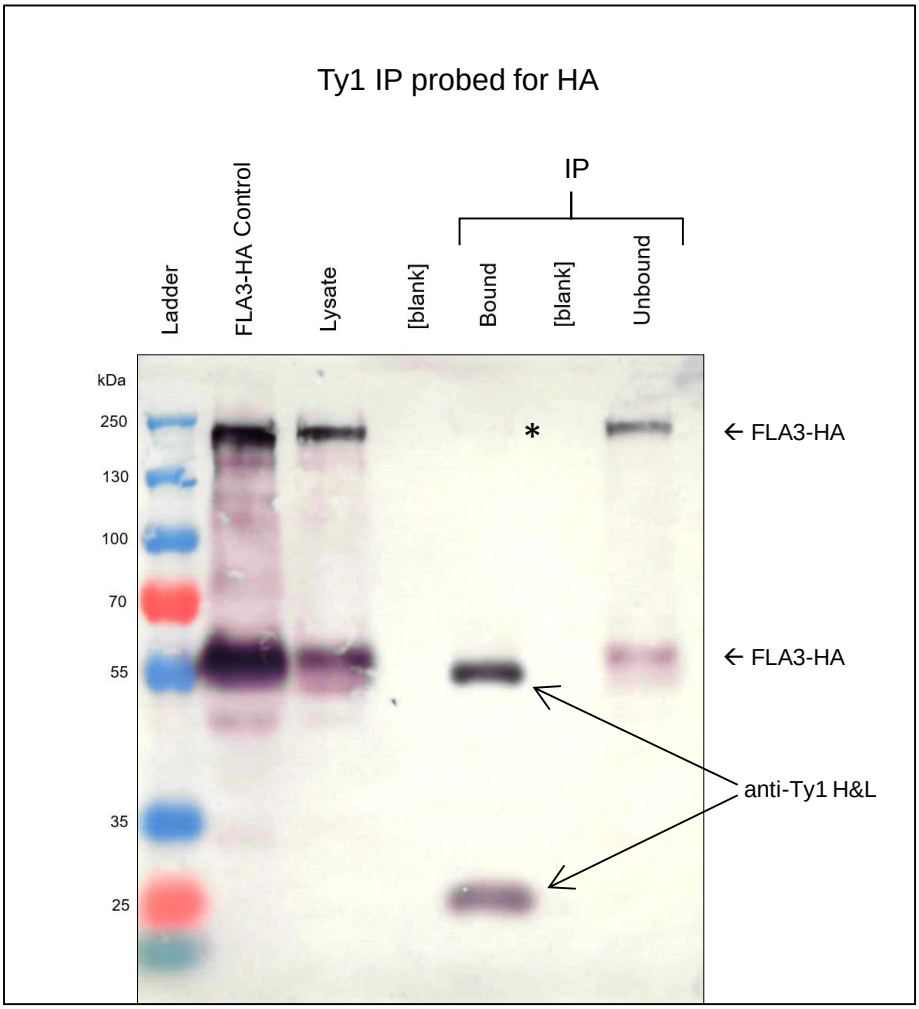


Figure 4.9: Does FLA3-Ty1 interact with FLA3-HA?

Detergent-lysed cells expressing FLA3-Ty1 and FLA3-HA (Clone R3) were subjected to immunoprecipitation using anti-Ty1 and Protein G agarose. The immunoprecipitate samples were probed with anti-HA on a western blot. No FLA3-HA was observed in the antibody-bound fraction. (Note that the bound sample represents $2x10^8$ cells, while the lysate and unbound samples represent $5x10^6$ cells and $4x10^6$ cells, respectively.)

4.3 Discussion

4.3.1 Immunoprecipitation and FLA3

Successful epitope tagging of FLA3 allowed investigation of possible protein-protein interactions involving FLA3. As yet, no binding partners of FLA3 have been identified. However, investigations of FLA1BP, the procyclic-form paralogue of FLA3, provide clues to their possible identities.

Based on the finding in Section 3.2.7 that FLA3 is anchored in the flagellar membrane, the current model of FLA3 orientation is that the short C-terminal tail of FLA3 is within the flagellum while the large N-terminal region is in the extracellular intermembrane region of the FAZ. It seems, therefore, that FLA3 resembles FLA1BP in localisation and topology, as well as in sequence. FLA1BP is known to bind, via its N-terminal region, to FLA1 and, via its C-terminal region, to FLAM3 (although the latter interaction may be indirect) (Sun et al., 2013, Rotureau et al., 2014). It seems highly likely, then, that FLA3 could bind, via its N-terminus, to FLA2 (the BSF paralogue of FLA1), and, via its C-terminal region, to the constitutive protein, FLAM3. Other obvious candidates for binding to FLA3 are FS179, FAZ5 and VSG. The putative calcium channel, FS179, seems to locate to the flagellar membrane of the FAZ (Oberholzer et al., 2011). FAZ5 is a protein that interacts with the intracellular FAZ filament but has several transmembrane domains (Sunter et al., 2015b). It could therefore be another component of the junctional complexes, although this has not yet been shown directly. In bloodstream-form cells, VSG dimers form a continuous surface coat over both the cell body and the flagellum (Bartossek et al., 2017). It therefore seems probable that FLA3 would be in close proximity to VSG.

However, the FAZ is an extremely complex membrane-membrane junction. There may be many more proteins in the intermembrane region than have been discovered thus far. It is impossible, therefore, to describe the architecture of the FAZ with any confidence until much more empirical data has been generated regarding the interactions mediated by the known FAZ proteins, such as FLA3.

Interaction studies were not performed when FLA3 was originally discovered (Woods et al., 2013). The immunofluorescence images produced using the antibody show considerable background signal, suggestive of nonspecific binding. Furthermore, the antibody was only

produced in a small-scale batch (personal communication), meaning that it could be used only for a limited number of experiments. In contrast, the stable Ty1-tagging of FLA3 performed in Chapter 3 is ideal for Co-IP because it facilitates highly specific recognition of FLA3 by the monoclonal BB2 anti-Ty1 antibody. Indeed, one of the main reasons for tagging FLA3 was to enable efficient immunoprecipitation of the protein.

There is excellent precedent for using Co-IP of tagged proteins to identify unknown binding partners. For example, Sunter et al. (2015) discovered FAZ2, FAZ3 and FAZ4 by co-immunoprecipitation with YFP-tagged FAZ1, and subsequently discovered FAZ5, FAZ6 and FAZ7 by co-immunoprecipitation with YFP-tagged FAZ2 and FAZ4 (Sunter et al., 2015b). Even more significantly, Sun et al. (2013) discovered FLA1BP through Co-IP with a YFP-tagged FLA1 mutant and were able to show reverse Co-IP whereby FLA1 co-precipitated with YFP-tagged FLA1BP.

The YFP tags on FLA1 and FLA1BP were added to the N-termini of the proteins, immediately after the respective signal sequences – i.e. to the extracellular region (Sun et al., 2013). For the current work, FLA3 was tagged at the C-terminus, to avoid interference with the putative N-terminal signal sequence. The Ty1 tag is thought therefore to locate to the intraflagellar region. However, the exact position of the tag on the membrane protein in question should not greatly affect which binding partners are pulled down during Co-IP. Prior to immunoprecipitation, the lipid membranes of the cell are solubilised with nondenaturing detergent, so the membrane protein in question (e.g. FLA3) is released into solution in such a way that, in theory, all its noncovalent interactions with proteins from either side of the membrane remain intact.

The domain structure of FLA3, however, means that Co-IP is probably more likely to reveal the proteins that bind to the N-terminal region, rather than those that bind to the C-terminal tail. The short C-terminal domain of FLA3 is hypothesised to bind to large structural proteins such as FLAM3 and ClpGM6, as had been proposed in the case of FLA1BP (Rotureau et al., 2014, Sunter and Gull, 2016). These large structural proteins are probably attached firmly to the flagellar cytoskeleton (Sunter et al., 2015a); it seems, therefore, that they are relatively unlikely to be solubilised by detergent.

In any case, it is the interactions with the N-terminal region of FLA3 that are of most interest. The intermembrane zone is the least well characterised part of the FAZ, and it could be the region most accessible to therapeutic drugs.

As soon as it was confirmed that FLA3-Ty1 could be immunoprecipitated (Fig. 4.1), silver-staining was used to search for co-precipitated proteins. Silver-staining is a useful method to reveal all the proteins present in a sample. It is far more sensitive than other nonspecific protein stains such as Ponceau S or Coomassie blue. Silver-staining was used to great effect for the discovery of FLA1BP (Sun et al., 2013).

The silver stains of the FLA3-Ty1 IP (Fig. 4.2) were successful: the FLA3-Ty1 band could be observed clearly in the antibody-bound fraction, and it was clear that there was little nonspecific binding to the Protein-G-agarose beads. Most importantly, there seemed to be some protein bands (~115 kDa, ~110 kDa and ~80 kDa, respectively) that co-precipitated with FLA3-Ty1 under nondenaturing conditions. These bands could represent binding partners of FLA3-Ty1. Their identities are not known, but a possible candidate for the ~80-kDa band is discussed in Section 4.3.3.

One of the drawbacks of silver-staining is the difficulty of identifying bands of interest. This problem is sometimes overcome using mass spectrometry, but this technique requires the band of interest to contain a substantial amount of protein. Sun et al. used mass spectrometry of their FLA1 immunoprecipitation to identify the FLA1BP band (Sun et al., 2013). However, this IP was performed in procyclic cells, with ~$4x10^9$ cell equivalents in the bound fraction. Obtaining a similarly large cohort of BSF cells is relatively difficult. Bloodstream-form cells remain viable only up to a maximal concentration of ~$2.5x10^6$ cells ml^{-1}, so the IP experiments performed here have only ~$2x10^8$ cell equivalents in the respective antibody-bound fractions. There is therefore insufficient protein to facilitate mass spectrometry identification of weak bands.

A further disadvantage of the silver-stain approach is that silver-staining does not amplify the signal from a given band in the same way as a western blot or a lectin blot. Therefore, there may be other proteins that co-precipitate with FLA3-Ty1 but are not visible on the silver stain. It must be noted that even in the most efficient immunoprecipitations, proteins do not always co-precipitate stoichiometrically. Co-precipitating proteins will not

pull down as efficiently as the target protein and will therefore be present in relatively small amounts in the bound fraction.

Given the lack of antibodies for most FAZ proteins, the best solution for identification of the binding partners of FLA3 is epitope-tagging of likely candidates. Attempts to tag FLA2 are detailed in Chapter 5. However, more general identification methods, including biotinylation and lectin probing, can also be informative about the attributes of co-precipitating proteins. These approaches are discussed in the forthcoming sections.

4.3.2 Surface biotinylation of FLA3

Surface biotinylation is a convenient way to label the proteins that are on the surface of a cell. *N*-Hydroxysulfosuccinimide (NHS) esters of biotin react with exposed primary amine groups (usually lysine residues and polypeptide N-termini) to form stable amide bonds (Shimogawa et al., 2015). Since these esters are membrane-impermeant, only surface proteins are labelled.

Surface biotinylation was performed here for two reasons: firstly, to assess whether FLA3 was accessible on the surface of the trypanosome and secondly, to see if any other surface proteins interact with FLA3-Ty1.

The finding that FLA3-Ty1 can indeed be biotinylated using this technique (Fig. 4.3; Section 4.2.2) confirms that the protein is surface-accessible – at least to small molecules. This is consistent with the prediction, based on the FLA3 sequence, that most of the protein is extracellular. Furthermore, it has profound implications for the consideration of FLA3 as a potential drug target. The two fundamental characteristics of a druggable protein are its essentiality for the viability of the pathogen and its accessibility to therapeutic agents. It seems that FLA3 has both these attributes.

Accessibility to NHS-biotin does not necessarily imply accessibility to circulating proteins such as antibodies. It seems likely, looking at cross-sections of the FAZ (e.g. Fig. 1.5), that the VSG surface coat may form a substantial barrier between the intermembrane FAZ and the extracellular medium. This would explain why the host adaptive immune response cannot clear the trypanosome population by targeting FLA3. However, it is plausible that

small-molecule therapeutic drugs could, like NHS-biotin, penetrate the VSG and bind to FLA3.

Furthermore, exciting new inroads in the field of nanobody (Nb) research have shown that nanobodies can be engineered to target trypanosome surface epitopes that are inaccessible to conventional antibodies – including invariant regions of the VSG coat (Stijlemans et al., 2017, Stijlemans et al., 2004). Nanobodies are small (~15 kDa), single-domain antibody fragments derived from the heavy-chain antibodies of camelid species (Stijlemans et al., 2011). They recognise their target epitopes with high specificity, and they bind with high affinity. It has been proposed that nanobodies could be used therapeutically as a targeted toxin-delivery system (Stijlemans et al., 2017). Perhaps a nanobody directed against FLA3 could deliver a specific inhibitor of flagellar attachment to the intermembrane region of the FAZ, producing a lethal effect on the parasite without damage to host tissue.

Having confirmed that FLA3-Ty1 could be biotinylated, attention was turned once more to the protein's putative binding partners. Following surface biotinylation, FLA3-Ty1 was immunoprecipitated to see whether other surface proteins would co-precipitate.

Apart from FLA3-Ty1, the only biotinylated band observed in the IP fraction was a diffuse ~60-kDa band (Fig. 4.3). This band is the right size to represent biotinylated VSG (discussed further in Section 4.3.5), but it is not certain that it co-precipitated specifically with FLA3-Ty1. Silver staining has shown (Fig. 4.2) a diffuse but intense ~55-kDa band to precipitate under nondenaturing conditions even when no anti-Ty1 was present. This suggests that VSG may interact nonspecifically with the Protein-G-agarose beads.

A repeat of this experiment might benefit from the inclusion of a VSG-depletion step. VSG is an overwhelmingly abundant protein in the BSF trypanosome ($5x10^6$ dimers per cell; 10% of total protein content) (Tiengwe et al., 2016). Detection of low-abundance surface proteins can be improved greatly by depletion of VSG, e.g. using GPI-phospholipase C (GPI-PLC) (Shimogawa et al., 2015).

The fact that no other biotinylated bands were observed upon IP of FLA3-Ty1 does not necessarily imply that FLA3 does not interact with any other surface proteins. Firstly, proteins that lack an exposed primary amine will not react with NHS-biotin. Secondly,

solubilisation issues or unintended disruption of noncovalent interactions could have preventing the binding partners of FLA3 from precipitating in sufficient quantities to be visible by streptavidin probing. Again, performing the immunoprecipitation using a larger cohort of cells might produce more data.

A fresh approach towards discovering the putative binding partners of FLA3 was taken by means of lectin blotting, discussed in Sections 4.3.4 and 4.3.5.

4.3.3 FLA3 and the Endoplasmic-Reticulum Chaperone Protein, BIP

As mentioned in Section 4.2.2, the surface-biotinylation blot was reprobed for the intracellular chaperone protein, BIP, to confirm that biotinylation had been specific to surface proteins (Fig. 4.3 C). A clear BIP band was observed that had not been recognised by streptavidin, confirming that the biotin probing had indeed been specific to surface proteins.

Unexpectedly, however, BIP was observed in the antibody-bound fraction, which should only contain proteins that interact with FLA3-Ty1. This finding raised the possibility of an interaction between FLA3-Ty1 and BIP.

BIP (usually styled BiP) is a soluble protein that locates to the ER lumen (Bangs et al., 1993). It is an isoform of the heat-shock protein, Hsp70, and is highly conserved across the eukaryotic domain (Bentley et al., 2019). In the trypanosome, BIP (also known as Grp78) has been shown to act as a chaperone that assists in the correct folding of VSG (Bangs et al., 1996). BIP also forms transient associations with a discrete set of large polypeptides in BSFs and PCFs. Since BIP has been shown to associate with secretory proteins in other systems, it is thought that this set of large proteins must be secretory proteins whose folding is assisted by BIP (Bangs et al., 1996).

The fortuitous finding, here, that BIP is present in a FLA3-Ty1 IP sample suggests that FLA3 may be one of the large molecular-weight extracellular proteins that bind transiently to BIP in the ER. There is significantly less BIP in the antibody-bound fraction (Fig. 4.3 C) relative to the nonbinding fraction, but this is easily explained by the fact that most of the BIP in the ER is probably associated with nascent VSG.

Further investigation is required to confirm that the putative interaction between FLA3 and BIP is specific. It is noteworthy that one of the weak bands observed when a FLA3-Ty1 IP sample was silver-stained (marked with an asterisk on Fig. 4.2, discussed in Section 4.3.1) is the correct size (~80 kDa) to represent BIP. This weak band was only seen under nondenaturing conditions and was postulated to represent a binding partner of FLA3-Ty1. This is consistent with the suggestion that FLA3 interacts specifically with BIP.

4.3.4 The Glycosylation of FLA3

Having investigated the putative binding partners of FLA3 by means of silver-staining and streptavidin probing, a third approach – lectin blotting – was performed. Lectin blotting allowed investigation of both the glycan side chains and the potential glycoprotein interaction partners of FLA3.

The lectin results presented in this chapter demonstrate that (a) FLA3-Ty1 can be precipitated by RCA-I lectin, at least under nondenaturing conditions (Fig. 4.4); (b) FLA3-Ty1 can be detected by RCA-I lectin blotting (Fig. 4.5); (c) FLA3-Ty1 can also be detected by tomato lectin blotting (Fig. 4.6).

Tomato lectin (TL) (*Lycopersicon esculentum* agglutinin (LEA)) interacts with high affinity with glycans containing three or more linear units of the repeating disaccharide poly-*N*-acetyllactosamine (poly-LacNAc) (β-Gal-(1→4)-β-GlcNAc-(1→3))n (Merkle and Cummings, 1987, Kawashima et al., 1990). It also binds to the core and stem regions of high-mannose complex-type *N*-glycans (Kaltner et al., 2018, Oguri, 2005) and the Man-(1→4)-β-GlcNAc-(1→4)-β-GlcNAc trisaccharide core of paucimannose glycans (Schwartz et al., 2013). These interactions, however, are probably weaker than the interactions with poly-LacNAc (Brosson et al., 2016).

RCA-I lectin (a non-toxic ricin) is less specific than TL in its interactions: it binds to galactose residues. It binds galactose much more strongly when in the form of Gal-GlcNAc oligosaccharides (Wu et al., 2006).

FLA3 is known to be extensively *N*-glycosylated. Digestion with PNGase F reduces its apparent molecular weight from ~160 kDa to ~100 kDa (Woods et al., 2013). The glycan side chains of FLA3 have never been characterised. They have been hypothesised to include poly-LacNAc, because FLA3 was originally isolated from a TL-binding fraction

(Woods et al., 2013, Rubotham et al., 2005). However, it is also possible that the presence of FLA3 in this fraction was due to noncovalent interaction with a TL-binding protein, rather than a direct interaction with TL (Rubotham et al., 2005).

The TL-binding fraction of *T. brucei* includes a group of proteins that possess unusual poly-LacNAc glycans and which are postulated to be associated with endocytosis (Nolan et al., 1999). The suggested role for these poly-LacNAc glycans in endocytosis remains unconfirmed, as there is evidence for and against the hypothesis. One of the crucial questions is whether the *T. brucei* transferrin receptor (TfR) has poly-LacNAc side chains; a recent investigation suggests that it does, although this point has been disputed in the past (Koeller et al., 2020, Mehlert et al., 2012).

It was further suggested that poly-LacNAc could possibly serve as a signal to prevent diffusion out of the pocket into the surface plasma membrane (Nolan et al., 1999). Since FLA3 locates to the flagellar membrane, it may require such a signal to prevent it from entering the wrong membrane domain.

The glycans of the TL-binding fraction were characterised as follows: BSF cells were solubilised and incubated with ricin beads (Atrih et al., 2005). It was found that 85% of the ricin-binding fraction also bound TL. The TL-binding glycans were isolated and characterised. A fraction of unusual poly-LacNAc glycans were observed. They were considered unusual because (a) they were very large (~54 LacNAc residues per glycan); (b) they contained mostly β-GlcNAc-(1→**6**)-β-Gal-(1→4) linkages, rather than β-GlcNAc-(1→**3**)-β-Gal-(1→4) linkages, producing a distinct compact shape; and (c) they were of very simple composition, with no sialylation, sulfation or fucosylation (Atrih et al., 2005).

Could FLA3 be one of the proteins bearing the giant poly-LacNAc glycans described by Atrih et al.? The results presented in Section 4.2 certainly suggest that it could contain poly-LacNAc among its side chains. The finding that FLA3-Ty1 could be precipitated by RCA-I lectin beads (Fig. 4.4) strongly suggested that RCA-I could recognise the glycans of FLA3. Most, although not all, of the FLA3-Ty1 was found in the lectin-bound fraction. However, this lectin pulldown was tested only under nondenaturing conditions, so the possibility remained that FLA3-Ty1 was precipitated due to noncovalent interaction with a lectin-binding protein, rather than direct interaction with RCA-I.

Subsequently, FLA3-Ty1 was immunoprecipitated with anti-Ty1 and then lectin-blotted with biotinylated RCA-I (Fig. 4.5). A clear band of ~160 kDa was observed under denaturing and nondenaturing conditions, confirming direct recognition of FLA3-Ty1 by RCA-I. It is very probable, therefore, that FLA3 has galactose residues in its *N*-glycans.

The specific test for poly-LacNAc is interaction with tomato lectin (TL). FLA3-Ty1 was immunoprecipitated as before and then lectin-blotted with biotinylated TL (Fig. 4.6). The result was interesting: a ~160-kDa corresponding to FLA3-Ty1 was observed, but it was very weak. There are several possible explanations for this weak interaction. FLA3-Ty1 may contain linear poly-LacNAc repeats among its side chains, but in a relatively small quantity. The high amount of FLA3-Ty1 in the IP sample ($2x10^8$ cell equivalents) may result in nonspecific binding to the lectin. Alternatively, TL may recognise the mannose core region of the glycans in question, resulting in an interaction that is specific but considerably weaker than an interaction with poly-LacNAc. Brosson (2016) noted that weak interactions are sometimes observed between TL and glycoproteins on blots that are not observed under the more stringent conditions of glycoprotein pulldown with TL-agarose. Therefore, to ascertain whether the interaction between FLA3-Ty1 and TL is truly an indicator of poly-LacNAc repeats, it might be best to perform a lectin-pulldown experiment similar to the one performed here using RCA-I.

The glycans of FLA3 are certainly large enough to include the giant poly-LacNAc chains described by Atrih et al. (2005). A 54-residue repeat of the LacNAc disaccharide (383.35 Da) would add approximately 21 kDa to the molecular weight of a glycoprotein. In comparison, the glycans of FLA3 contribute at least 60 kDa to the molecular weight of the mature protein. However, the apparent weakness of the interaction between FLA3-Ty1 and tomato lectin makes it seem unlikely that FLA3 has such large poly-LacNAc repeats. Comparison with a known poly-LacNAc glycoprotein would perhaps give an indication of the relative amount of poly-LacNAc present.

Taken together, the results presented here suggest that FLA3 has many galactosyl residues, and possibly a low to moderate amount of poly-LacNAc.

The question remains as to the function of the glycans of FLA3. In addition to the tentative hypothesis that poly-LacNAc could act as a sorting signal that prevents membrane glycoproteins diffusing from the FP membrane into the exposed pellicular membrane

(Nolan et al., 1999), there is a distinct possibility that poly-LacNAc could play a role in immune evasion. Poly-LacNAc is used in the mammalian host for immune cell signalling and for maintenance of membrane stability in lysosome and erythrocyte membranes (Zhou, 2003). It is a component of the i/I antigens of erythrocytes (Zhou, 2003) and it acts as a scaffold for the ABO blood-group antigens (Stanley and Cummings, 2017). It is tempting to speculate, therefore, that the presence of LacNAc on invariant trypanosome surface proteins helps to conceal the parasite from the host immune system by mimicking host glycans.

In the case of FLA3, the intriguing possibility remains that the protein's glycans could have a direct function in flagellar attachment.

Galactose synthesis is essential in BSFs (Roper et al., 2002). Trypanosomes cannot take up galactose from their surroundings; they use an epimerase to convert glucose to galactose. Knockout of this epimerase was shown to cause cessation of cell division, and subsequent cell death. However, Roper et al. did not state whether flagellar detachment was observed in these knockout cells, so it cannot be stated whether galactose is required for flagellar attachment. Besides, the loss of galactose would produce pleiotropic downstream effects, so the observed cell death cannot be ascribed to a single cause.

Fascinatingly, a connection has been noted between fucosylation and flagellar attachment – at least in procyclic-form cells (Turnock et al., 2007). Conditional knockout of fucose synthesis caused flagellar detachment in PCFs. It was hypothesised that the PCF-specific flagellar adhesion glycoprotein, FLA1, could be fucosylated and that its fucose residues could contribute directly to flagellar attachment. This hypothesis was strengthened by the observation that the *T. cruzi* homologue of FLA1, GP72, is known to have fucose residues.

In contrast, conditional knockout of fucose synthesis in bloodstream-form cells was not seen to cause flagellar detachment, although cell division ceased and the cells exhibited aberrant morphology (Turnock et al., 2007). It would seem, therefore, that the bloodstream-form flagellar adhesion glycoproteins (FLA2 and FLA3) probably do not require fucose residues to mediate flagellar attachment.

Still, the apparent reliance of the PCF FAZ on specific glycan components suggests that the BSF FAZ might have similar requirements. Further investigations of the structure and function of the glycosyl side chains of FLA3 are necessary.

4.3.5 Potential Glycoprotein Binding Partners of FLA3

The lectin blots shown in Figs. 4.5 and 4.6 were performed not just to investigate the glycans of FLA3, but also to see whether other glycoproteins had co-precipitated with FLA3-Ty1.

As described in Section 4.2.4, biotinylated RCA-I lectin recognised three unidentified protein bands that appeared to co-precipitate specifically with FLA3-Ty1. These bands (labelled 1, 3 and 4 on Fig. 4.5 A) were clearly visible in the antibody-bound fraction when FLA3-Ty1 was immunoprecipitated under nondenaturing conditions. They were also apparent in the antibody-bound fraction when FLA3-Ty1 was immunoprecipitated under denaturing conditions (Fig. 4.5 B) but they were much weaker under these conditions. These bands could potentially represent glycosylated binding partners of FLA3.

It might be expected that if these three bands represent proteins that interact noncovalently with FLA3-Ty1, they would not be observed at all in the IP sample prepared under denaturing conditions. However, it is possible that the conditions employed (2% SDS) failed to disrupt all protein-protein interactions completely. Including β-mercaptoethanol and performing a longer boiling step (5 min instead of 3 min) might have allowed more efficient denaturation.

The fact that no proteins precipitated when a mock immunoprecipitation was performed without anti-Ty1 antibody (Fig. 4.5 C) confirms that the bands were not precipitated due to a nonspecific interaction with the Protein G beads. However, the lectin blot depicted in Fig. 4.5 A would have to be replicated before it could be stated with certainty that the three unknown protein bands are indeed interaction partners of FLA3, rather than accidental contaminants.

Considered as potential binding partners of FLA3, the three unidentified bands in question (~180 kDa, ~130 kDa and ~60 kDa, respectively) present certain features of interest. It is certain that the proteins in question must be galactosylated, since they were recognised by RCA-I. The relatively low intensity of the 130-kDa band ("Band 3") could indicate a

relatively low quantity of protein; alternatively, it could indicate an abundant protein with relatively few galactose residues.

An obvious question is whether any of these bands is likely to represent FLA2, in accordance with the hypothesis that FLA3 may mediate flagellar attachment by binding to FLA2. However, this question is difficult to answer because the molecular weight of mature wildtype FLA2 is unknown. The predicted molecular weight of FLA2 is 64 kDa, but the protein is postulated to be extensively glycosylated. Tagging experiments were performed (see Chapter 5) to enable better characterisation of FLA2. Several tagged versions of FLA2 were visualised by western blotting. If the post-translational modifications of these tagged versions are representative of the post-translational modification of the wildtype protein, it is predicted here that mature wildtype FLA2 has a molecular weight of ~100 kDa. If this prediction is accurate, none of the bands observed in Fig. 4.5 represent FLA2.

What of the other potential binding partners of FLA3 that were proposed in Section 4.3.1? FS179 (304 kDa), FLAM3 (468 kDa) and ClpGM6 (~1500 kDa) are too large to resolve on the 10% acrylamide SDS-PAGE gels used here; FLAM3 and ClpGM6 are unlikely to be glycosylated in any case, as these proteins are intracellular. The transmembrane protein FAZ5 is perhaps a possible candidate for the ~60-kDa band ("Band 4"). The predicted molecular weight of FAZ5 is 67 kDa, but the protein has been shown to migrate at ~80 kDa with a ~27-kDa tag (Sunter et al., 2015b), so the mature wildtype protein may be in the region of 50–60 kDa. However, FAZ5 must not be extensively glycosylated if it migrates close to its predicted size, whereas the intensity of Band 4 suggests a glycoprotein that is heavily galactosylated.

Band 4 seems too large to represent VSG. The MITat 1.1 isotype of VSG usually migrates as a 50-kDa band. It is possible that VSG contributes to Band 5 (~50–55 kDa), in addition to the FLA3-Ty1 degradation product and the anti-Ty1 heavy chain.

There is an interesting possibility that Band 4 represents the same protein that was observed as a diffuse 60-kDa band on the surface biotinylation blot (Fig. 4.3; discussed in Section 4.3.2). The latter band was originally hypothesised to represent biotinylated VSG, but it may, alternatively, represent a distinct surface glycoprotein that co-precipitates with FLA3-Ty1 and interacts with RCA-I lectin.

In summary, then: the RCA-I lectin blot (Fig. 4.5 A) reveals that at least three galactosylated proteins appear to co-precipitate specifically with FLA3-Ty1, but the identities of these putative binding partners remain unknown.

Little can be concluded from the biotinylated tomato lectin blot (Fig. 4.6) regarding the putative glycoprotein binding partners of FLA3. The fact that the tomato lectin did not recognise any bands that co-precipitated with FLA3-Ty1 would seem to suggest that FLA3-Ty1 does not interact with any glycoproteins that have linear poly-LacNAc repeats. However, the possibility remains that the proteins that co-precipitated with FLA3-Ty1 did so in quantities that were too low to facilitate recognition by TL. It can merely be stated that no evidence for poly-LacNAc interaction partners was found under the conditions tested.

4.3.6 Does FLA3-Ty1 Interact with FLA3-HA?

In considering the possible architecture of the junctional complexes in which FLA3 performs an essential role, it is possible that FLA3 behaves as an oligomer. In maintaining lateral attachment of the flagellum along the cell body, the junctional complexes must withstand the vigorous force of the flagellum as it beats continuously. They must, therefore, form an extremely strong, rivet-like connection between the flagellar and subpellicular cytoskeletons. The dramatic knockdown phenotype of FLA3 suggests that the protein performs a fundamental structural role in flagellar adhesion. It seems likely that interactions between FLA3 and its putative binding partners would be far more efficient if multiple adjacent copies of the FLA3 protein operated in tandem. Certainly, the strong immunofluorescence signal when tagged FLA3 is viewed by super-resolution microscopy (Fig. 3.20) would seem to suggest that multiple copies of FLA3 must locate to each observed punctum. It seems plausible, therefore, that FLA3 might interact with itself, forming an oligomeric structure.

This hypothesis was tested by expressing two different tagged versions of FLA3 (FLA3-Ty1 and FLA3-HA) in the same cell line. Having first verified that both tagged forms localised correctly (Fig. 4.8), FLA3-Ty1 was immunoprecipitated under nondenaturing conditions, and the IP sample was probed for HA (Fig. 4.9). A faint FLA3-HA band was indeed visible, but it was clear that almost all the FLA3-HA complement had remained in the nonbinding fraction.

It is not certain, from this result, whether any significant interaction occurs between tagged FLA3 molecules. If Co-IP conditions were optimised, a more definite result might be produced. Judgement as to whether wildtype FLA3 behaves as an oligomer is reserved until further investigations are made.

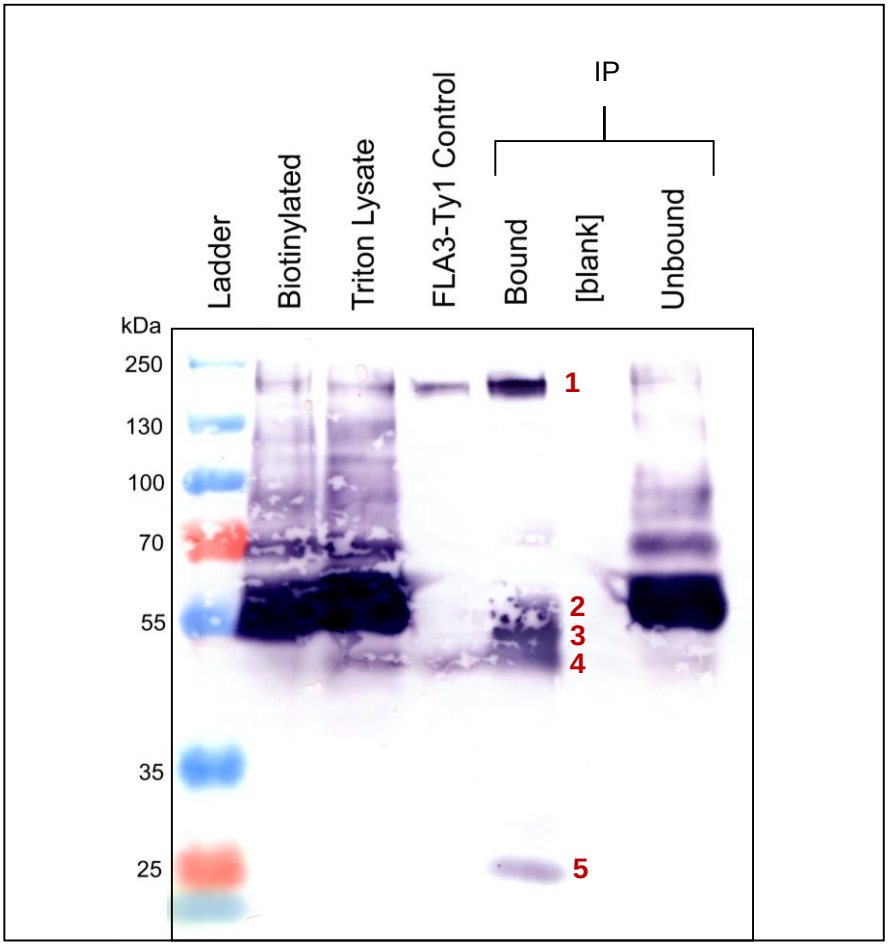


Figure 4.10: Surface biotinylation of FLA3-Ty1, probed for biotin and Ty1.

The streptavidin-probed blot shown in Fig. 4.3 was reprobed for Ty1. The bands in the antibody-bound fraction are as follows:

1) FLA3-Ty1
2) Unidentified biotinylated protein
3) Anti-Ty1 heavy chain
4) Stable FLA3-Ty1 degradation product
5) Anti-Ty1 light chain.

Bands 3, 4 and 5 were not visible by streptavidin probing alone.

4.3.7 Further Characterisation of the Stable FLA3 Degradation Product

The immunoprecipitation experiments presented in this chapter shed some further light on the stable ~50-kDa FLA3 breakdown product discussed in Section 3.3.9.

The reproducibility of this ~50-kDa band was further verified when it was observed in samples of cells in which FLA3 was tagged with Ty1 by means of long-primer PCR tagging with the pPOTv6 tagging vector (Fig. 4.7). This was confirmation that the putative stable FLA3 degradation product was not an artefact of the traditional *in-situ* tagging performed previously with the p3074 and p2708 vectors.

In the initial immunoprecipitations of FLA3 (Fig. 4.1), it was not possible to discern whether the FLA3 degradation product had been immunoprecipitated because, if present, it was obscured by the intense anti-Ty1 heavy-chain band. Subsequent immunoprecipitations were better optimised: they were performed using significantly more cells and significantly less primary antibody, according to the protocol described in Chapter 2. It therefore became possible to resolve the ~55-kDa anti-Ty1 heavy chain and the ~50-kDa FLA3 degradation product, as seen clearly in Fig. 4.6 B and Fig. 4.10. The FLA3 breakdown product can be reliably immunoprecipitated in cell lines in which FLA3 is tagged with Ty1.

Interestingly, Fig. 4.10 also indicates that the FLA3-Ty1 degradation product is not found on the surface of the cell. The ~50-kDa band is clearly visible in the antibody-bound fraction upon probing with anti-Ty1 but was not visible previously when the same blot was probed with streptavidin (Fig 4.3 A). This means that the protein fragment in question had not been biotinylated, indicating that it was not present on the cell surface. (Since there are many lysine residues throughout the FLA3 sequence, it is very unlikely that there are simply no accessible primary amines for the biotin to bind.) It is not surprising that the FLA3 degradation product is intracellular, as it probably locates to the lysosome.

When FLA3-Ty1 was solubilised and incubated with RCA-I lectin beads (Fig. 4.4), the ~50-kDa FLA3 degradation product was seen as an intense band in the lectin-bound fraction, which shows that it is galactosylated. In contrast, the FLA3 breakdown product was not observed upon probing with tomato lectin (Fig. 4.6), which suggests that unlike full-length FLA3, it lacks poly-LacNAc.

The FLA3 breakdown product must correspond to the C-terminal end of the FLA3 protein because it retains the Ty1/HA tag. It is approximately one-third of the molecular weight of the full-size protein. It would seem, therefore, that the TL-binding glycans must be attached somewhere along the N-terminal two-thirds of the mature protein. At least some of the RCA-I-binding glycans must attach to the C-terminal third of the protein.

4.3.8 Summary of Chapter 4

The crucial findings presented in this chapter are the confirmation that FLA3 is surface exposed, the demonstration that several as yet unidentified proteins appear to co-precipitate with FLA3 (including some surface glycoproteins), and the evidence that the glycans of FLA3 are recognised both by RCA-I lectin and by tomato lectin.

Chapter 5: Localisation of FLA2

5.1 Introduction

The confirmation that FLA2 is essential for flagellar attachment and cytokinesis in the bloodstream-form cell (Section 3.2.1) indicated that this protein merited closer investigation. Since FLA2 had a high level of sequence similarity with FLA1, and since the two proteins had a similar role in flagellar attachment, it seemed reasonable to assume that FLA2 would have the same location and structure as FLA1 (LaCount et al., 2002, Sun et al., 2013, Sunter and Gull, 2016). However, FLA2 has never been studied at the protein level.

To investigate the expression, location and putative interaction partners of FLA2, it was first necessary to epitope-tag the protein, as had been performed for FLA3 (Chapter 3). However, tagging FLA2 was a more challenging prospect because of the high sequence similarity between the *FLA2* and *FLA1* genes (Appendix C). To ensure specific tagging of *FLA2*, long-primer PCR tagging was used (see Section 2.2.8). This approach took advantage of the fact that the 3′ UTRs of the *FLA2* and *FLA1* genes are dissimilar (Table 5.1). Although the forward primer could not distinguish between *FLA2* and *FLA1*, the reverse primer should specifically recognise the 3′ UTR of *FLA2*. Using this method, it was possible to insert a tag (a Ty1 epitope tag, fluorescent protein, or both) at the C-terminus of FLA2 to allow detection by immunofluorescence or western blot.

When FLA2 was tagged in this manner with a fluorescent protein (mScarlet or mNeonGreen), the tagged protein was observed to localise throughout the cell body (Sections 5.2.1–5.2.3). This result was unexpected as, given its role in flagellar attachment, FLA2 would be expected to localise to the FAZ; it suggested that the tagged protein mislocalised because of the presence of a large tag (~36 kDa in total) at the C-terminus. However, upon successive tagging attempts with increasingly small tags (Sections 5.2.4; 5.2.8–5.2.10), each tagged version of FLA2 mislocalised. It is probable, therefore, that the short (16-a.a.) cytoplasmic C-terminal tail of FLA2 may play a role in localisation of FLA2 to the FAZ, and that addition of a tag to this region interferes with normal trafficking or interaction with other proteins.

Since the tagged versions of FLA2 generated here were thought to mislocalise, interaction studies were not performed. However, fractionation experiments demonstrated that Ty1-

tagged FLA2 behaved as a membrane protein (Section 5.2.4), and lectin pulldown with RCA-I showed that FLA2-Ty1 was glycosylated (Section 5.2.5).

Interestingly, the apparent molecular weights of the tagged versions of FLA2 are consistent with the hypothesis that the "FLA1" protein observed in bloodstream-form lysates by Nozaki et al. (1996) was most likely FLA2. This hypothesis is discussed in Section 5.3.

```
1       ttggagctgcaagtgacttcggagtcactcatggatacgtaatggaggtg     50
        .|...|||.|.||||.|.|.|..|||||.|..|.||||||.||||.|.|.
1       ctaccgctacgagtggcattgatgtcacaccagaatacgtgatggtgctc     50

51      a-cagc--cggtggggaatagtgagaggcaacctctgcgac--tacttcg     95
        | ||.|  | ||||.||| |..||.||..|.|.||    ||  .|||||
51      accatctac-gtggagaa-acagaaagatatcatc----acaggacttc-     93

96      catgaggttagagttgggtgggtgt---gcgtt-tctggtcgt---catg    138
                ||.|||||.||   ||.|| |.| .||||   ||
94      --------------tgtgtgggggtttagcattgtat-ttcgtttgca--    126

139     tgatggctacaccttgctcattttgggtgaggaaacaaatctttttgtgc    188
              |||.|||  .|.|||.||       ||||.||..||.||..||.
127     ------ctatacc--acacatgtt-------gaaagaacgctgttgttgt    161

189     agaaaatacgtatttatgacagtaagaaaaagcgt---gtgatt------    229
        |...||.|||.||.|||   ||.||..||..||.|   .|.|||
162     atgtaacacgcatgtat---agcaattaagggcattacatcattaatttg    208

230     ----acgt-----------aagtgatagaataggta------    250
            ||||           |.||..|.|||||..||
209     gcagacgttccgttttcccatgtttttgaataattaagattg    250
```

Length: 292
Identity: 143/292 (**49.0%**)
Similarity: 143/292 (**49.0%**)
Gaps: 84/292 (**28.8%**)
Score: 245.0

Table 5.1: Pairwise alignment of the *FLA2* 3′ UTR and the *FLA1* 3′ UTR.

Upper row: The first 250 nucleotides of the *FLA2* 3′ UTR (same for both Tb927.8.4060 and Tb927.8.4110)

Lower row: The first 250 nucleotides of the *FLA1* 3′ UTR

UTR sequences from Siegel (2010). Pairwise alignment was performed using the EMBOSS Needle tool.

5.2 Results

5.2.1 Long-Primer PCR Tagging of FLA2 with a Fluorescent-Protein Tag (FLA2-SR / FLA2-NG)

To investigate the localisation of FLA2, the protein was tagged at the C-terminus with a fluorescent-protein tag, mScarlet. This was achieved by long-primer PCR tagging of the endogenous *FLA2* gene using the pPOTv7 tagging construct as the template (Section 2.2.8). Briefly, the mScarlet tag and the downstream selectable marker (blasticidin resistance) were amplified from the pPOTv7 vector by PCR, using primers homologous to the *FLA2* gene and to the *FLA2* 3' UTR, respectively. Wildtype bloodstream-form cells were transfected with the resulting amplicon and subjected to selection with blasticidin. Several clones were subsequently isolated.

Expression of mScarlet-tagged FLA2 (FLA2-SR) at the protein level was confirmed in these clones by flow cytometry (Fig. 5.1). Fluorescence data were acquired from live clonal cells on an LSRFortessa flow cytometer. Of the 10,445 intact tagged cells analysed, 98.4% showed positive for mScarlet fluorescence, confirming that the clone expressed FLA2-SR. There was very little background signal in parental wildtype cells (Fig. 5.1 D–F).

The localisation of FLA2-SR was investigated by direct fluorescence microscopy (Fig. 5.2). Live transfected cells could be seen to emit a weak red fluorescence throughout the cell body; there was no background signal from wildtype cells. This localisation pattern was unexpected as FLA2 is hypothesised to be a component of the flagellar attachment zone because of its role in flagellar adhesion. It seemed likely that the mScarlet tag prevented normal localisation of FLA2. However, direct fluorescence is a relatively insensitive method to determine the localisation of a low-abundance protein such as FLA2. It was decided to tag FLA2 with a Ty1 epitope tag, which would facilitate immunodetection with a sensitive monoclonal antibody and would also enable western blotting.

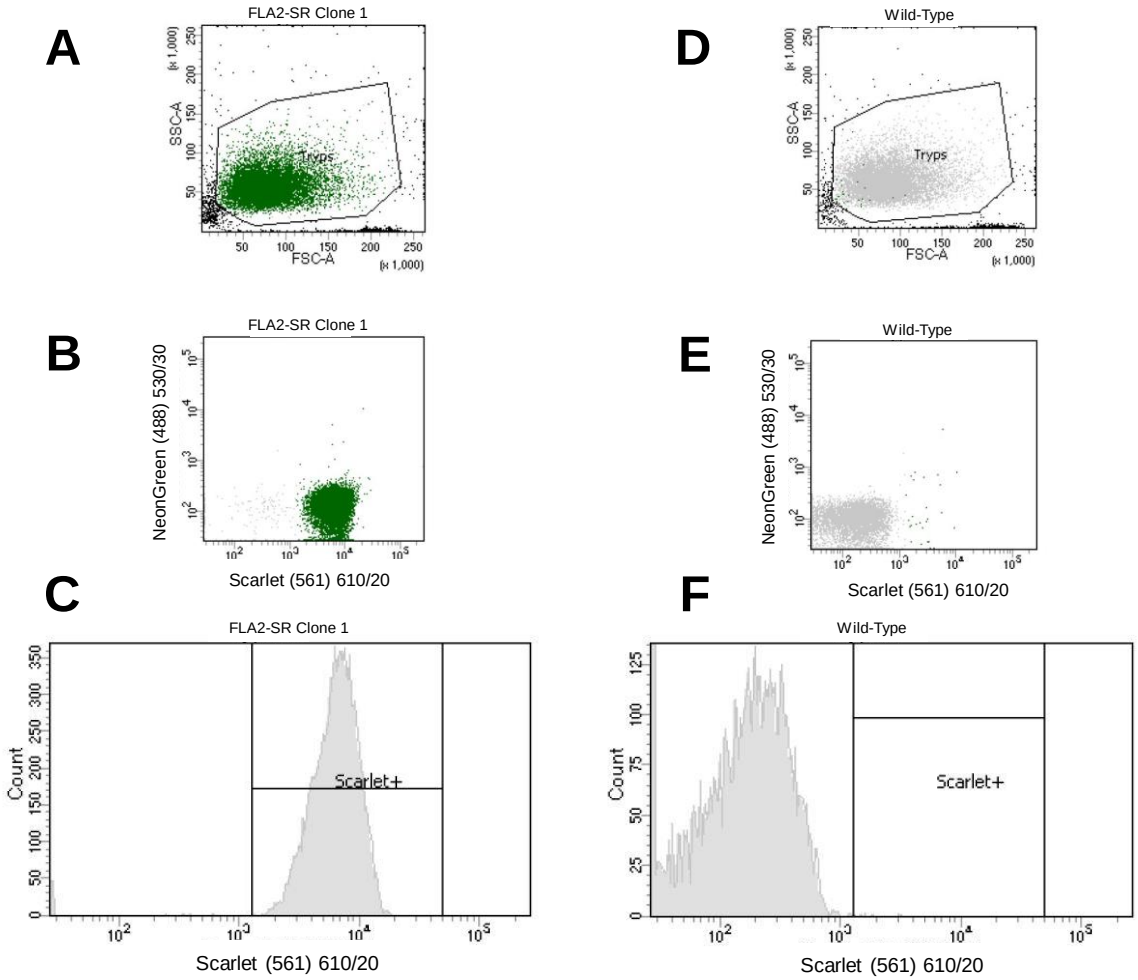


Figure 5.1: Expression of mScarlet-tagged FLA2 in bloodstream-form cells. Long-primer PCR tagging with pPOTv7 was used to add an mScarlet tag to the C-terminus of the *FLA2* gene at the endogenous locus. **(A–C)** Expression of tagged FLA2 ("FLA2-SR") at the protein level was confirmed by flow cytometry of a representative clone (see Methods). **(A)** Intact cells were gated according to size. **(B)** Red fluorescence was measured according to an excitation wavelength of 561 nm and an emission range of 610/20. (Green fluorescence was also measured as a negative control). **(C)** A histogram of fluorescence intensity shows that the cells emitted red fluorescence. Of the 10,445 intact cells measured, 98.4% showed positive for mScarlet fluorescence.

(D–F) Wildtype cells were analysed by flow cytometry in the same manner. Of 10,103 intact cells measured, only 0.3% were within the mScarlet threshold.

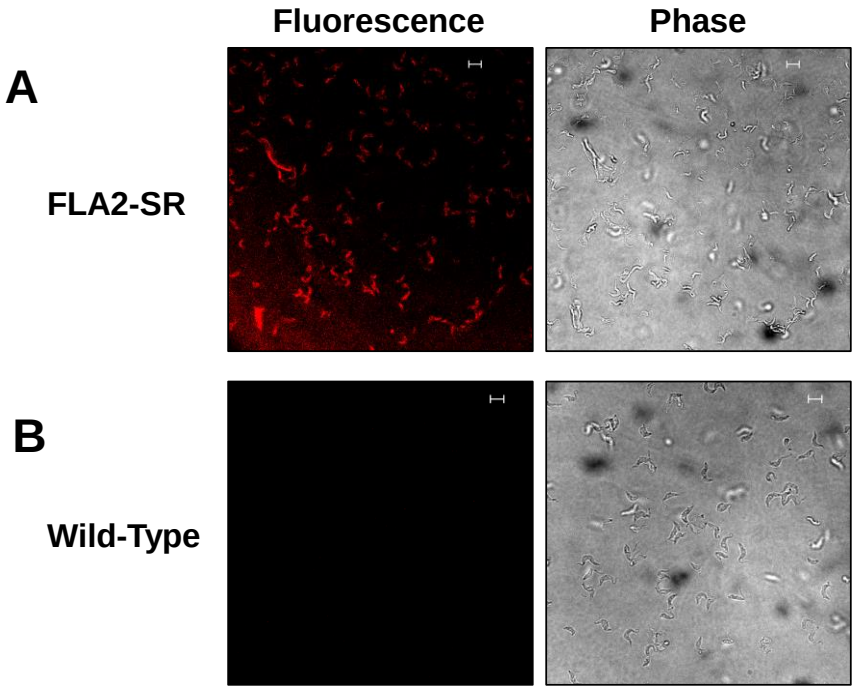


Figure 5.2: Localisation of mScarlet-tagged FLA2 in bloodstream-form cells.

(A) Direct fluorescence further confirmed that cells expressing FLA2-SR (see Fig. 1) emit red fluorescence, whereas untagged control cells **(B)** do not. However, the signal appears to be mislocalised throughout the cell, suggesting that the mScarlet tag prevents normal localisation of FLA2.

Scale bars are 10 µm.

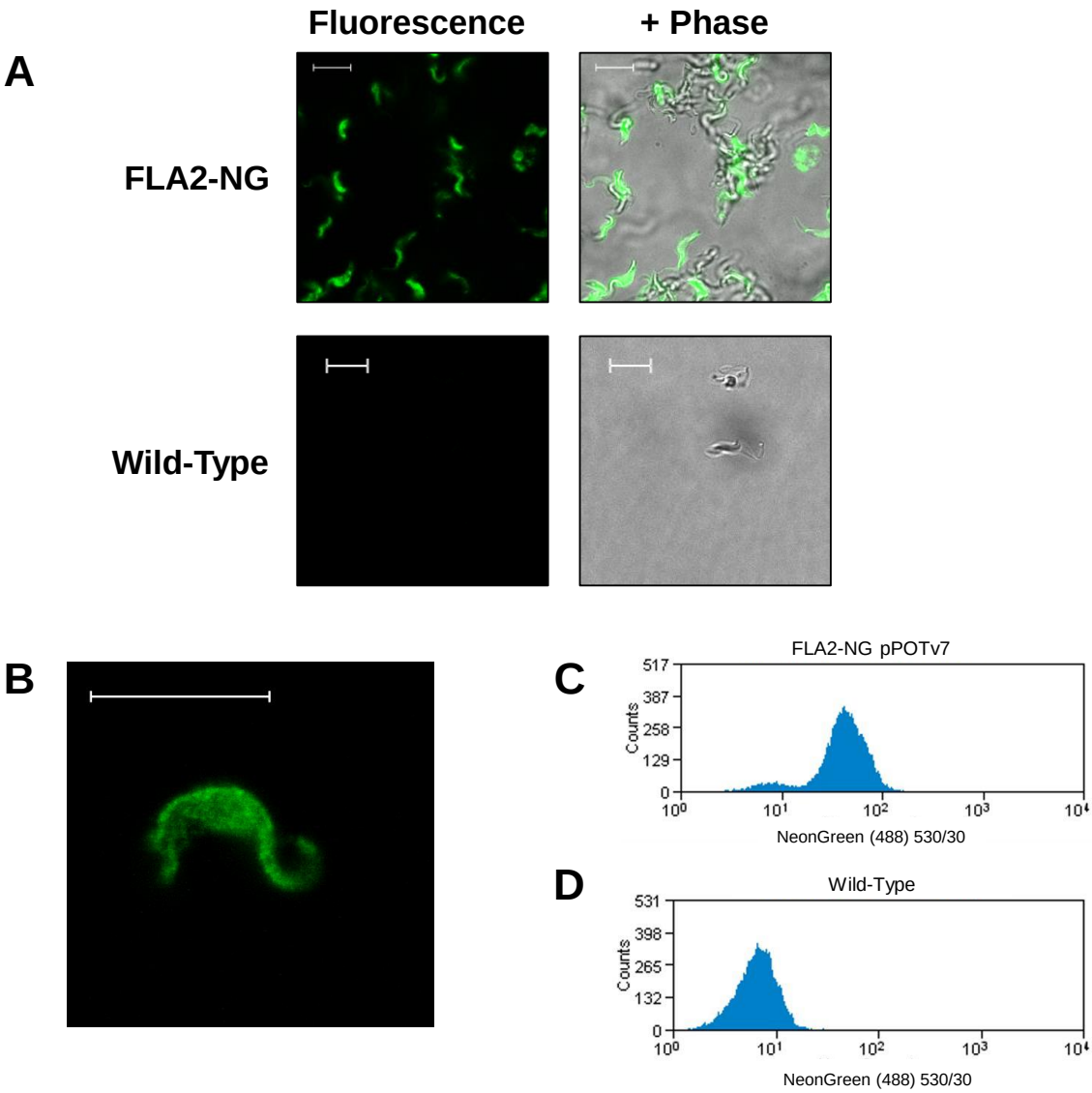


Figure 5.3: Localisation of mNeonGreen-tagged FLA2 in bloodstream-form cells.

Long-primer PCR tagging with pPOTv7 was used to add an mNeonGreen (NG) tag to the C-terminus of the *FLA2* gene at the endogenous locus. **(A)** Direct fluorescence microscopy of clonal live cells expressing FLA2-NG shows a weak green fluorescence throughout the cell body. Wildtype cells do not fluoresce.

(B) Upon fixation, the cells emit the same mislocalised green signal by direct fluorescence. Scale bars are 10 µm.

(C) Flow cytometry of the same clone confirms that the cells emit green fluorescence (median fluorescence intensity: 42.79) while untagged wildtype cells **(D)** do not fluoresce (median fluorescence intensity: 6.54).

Firstly, however, FLA2 was tagged with a different fluorescent-protein tag, to see if this would have a more satisfactory result.

Long-primer PCR tagging was carried out exactly as before, using a variant of the pPOTv7 vector that added the mNeonGreen fluorescent protein sequence to the 3′ end of the endogenous *FLA2* gene. Again, live cells were investigated by direct fluorescence microscopy (Fig. 5.3 A). There was no background signal from parental wildtype cells. Transfected cells expressing mNeonGreen-tagged FLA2 (FLA2-NG) emitted green fluorescence. As before, the signal was observed throughout the cell body. To demonstrate that the diffuseness of the mNeonGreen signal was not due to the motility of the cells, the cells were fixed with PFA and viewed by direct fluorescence (Fig. 5.3 B). The signal was the same as before: a uniform green signal throughout the cell body. Flow cytometry data (Fig. 5.3 C, D) were acquired at an excitation wavelength of 488 nm and an emission range of 530/30. The median fluorescence intensity (MFI) of the transfected clone was 42.79, whereas that of wildtype cells was only 6.54. It seemed, therefore, that FLA2-NG was expressed, but that the mNeonGreen tag probably prevented normal localisation of the protein.

5.2.2 Long-Primer PCR tagging of FLA2 with a Combined Epitope Tag and Fluorescent-Protein Tag (FLA2-Ty1-NG)

To enable localisation of FLA2 by indirect immunofluorescence, a fresh tagging attempt was made. Long-primer PCR tagging with the pPOTv6 tagging vector (see Section 2.2.8) was used to add a Ty1-mNeonGreen-Ty1 tag to the C-terminus of endogenous *FLA2* in bloodstream-form cells that were already FLA2-RNAi-competent (p2T7.177 Clone 1, see Section 3.2). The repeated Ty1 epitopes would allow amplification of the signal during each antibody-probing step, producing a stronger signal than with direct fluorescence and perhaps enabling more precise localisation of the tagged protein.

Expression of tagged FLA2 (FLA2-Ty1-NG) was confirmed by western blotting (Fig. 5.4). The tagged protein appeared as a single band of ~140 kDa. Taking into account the size of the Ty1-NG-Ty1 tag (~36 kDa), FLA2 appears to be considerably larger (by about 40 kDa) than the size (64 kDa) predicted from the primary sequence of the protein based on

the annotated open reading frames (Tb927.8.4060/4110). This discrepancy between the expected and observed size suggests that the mature protein is extensively post-translationally modified, as is the case for FLA1 and FLA3 – or, alternatively, that nonspecific off-target tagging had occurred.

Since tagging had been performed in FLA2 RNAi-competent cells, it was possible to verify that FLA2 had been tagged. Knockdown of *FLA2* mRNA was induced by addition of tetracycline. Specific loss of the FLA2-Ty1-NG band was observed (Fig. 5.5), confirming the identity of the band.

The localisation of FLA2-Ty1-NG was investigated by indirect immunofluorescence (Fig. 5.6). Upon probing for Ty1, the clonal cells expressing FLA2-Ty1-NG showed strong red signal throughout the cell body, while the parental cell line showed no signal. It seems highly unlikely that this distribution is representative of the true localisation pattern of wildtype FLA2. The similarity of FLA2 to the known FAZ-specific protein, FLA1, coupled to the fact that FLA2 mediates flagellar attachment, suggests strongly that FLA2 is a FAZ protein or is associated with the flagellum. It is probable, therefore, that the addition of the Ty1-NG-Ty1 tag to the C-terminal end of FLA2 caused the protein to mislocalise.

It is possible that mislocalisation of tagged FLA2 occurs because the tag obscures a putative localisation signal or blocks interactions with other FAZ proteins. Alternatively, mislocalisation may be due to unintentional overexpression of the tagged protein. The pPOTv6 tagging vector changes the 3' UTR of the tagged gene to the 3' UTR of the *PFR2* paraflagellar-rod gene. As described in Section 1.4, the 3' UTR determines the stability of an mRNA in *T. brucei*, and thereby influences the expression level of the encoded protein. It could be the case, therefore, that the tagging system used here causes overexpression of tagged FLA2, leading to mislocalisation of the surplus protein. In that case, however, some FAZ signal should be apparent, in addition to signal throughout the cell body. Furthermore, when FLA3 was tagged in the same way, using the pPOTv6 tagging construct as the template (Section 4.2.5), the tagged protein localised correctly to the FAZ, with no apparent overexpression.

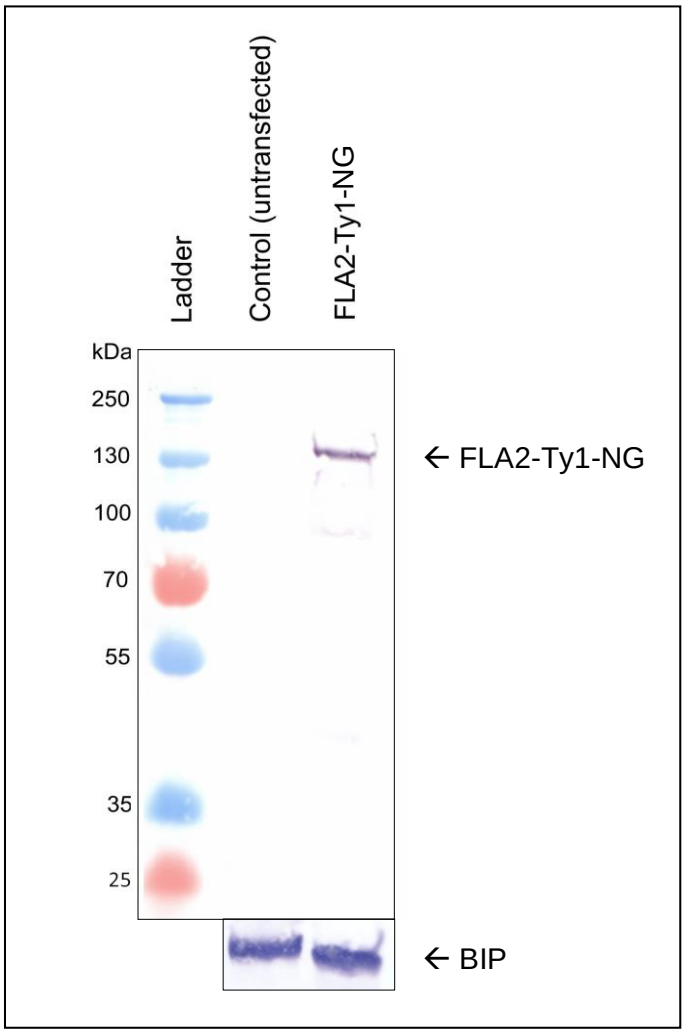


Figure 5.4: Expression of FLA2 with C-terminal Ty1 and NG tags.

Long-primer PCR tagging with pPOTv6 was used to add a Ty1-NG-Ty1 epitope tag to the C-terminus of the endogenous *FLA2* gene. Western blot probed for Ty1 confirms expression of FLA2-Ty1-NG at the protein level. FLA2 is a 64-kDa protein that appears here as ~140 kDa due to the ~34 kDa tag and, furthermore, to extensive glycosylation.

Probing for a chaperone protein, BIP, indicates that loading was consistent.

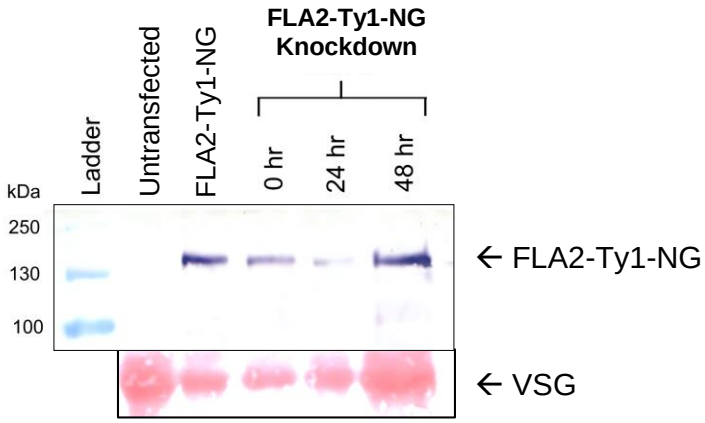


Figure 5.5: Knockdown of tagged FLA2 in bloodstream-form cells.

The FLA2-Ty1-NG clone described in Fig. 4 had been generated in a FLA2-RNAi-competent parental cell line. Western blot probed for Ty1 shows knockdown of FLA2-Ty1-NG in these cells upon addition of tetracycline. Significant loss of FLA2-Ty1 is observed in the first 24 hr, although the protein is not ablated completely.

Ponceau-stained VSG bands shown as loading control.

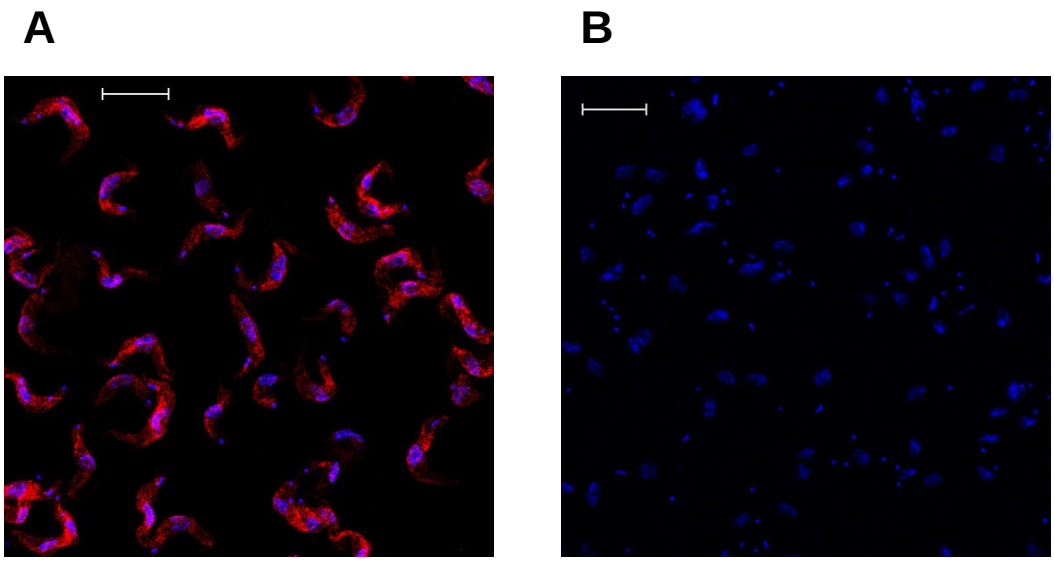


Figure 5.6: Localisation of Ty1-NG-tagged FLA2 in bloodstream-form cells.
Localisation of FLA2-Ty1-NG in the clone described in Fig. 4 was investigated by immunofluorescence (see Methods). **(A)** Upon probing for Ty1 (red), the cells show a strong signal throughout the cell body. Addition of the bulky tag may cause FLA2 to mislocalise. **(B)** In the absence of anti-Ty1 antibody, the transfected cells show no signal.
Nucleus and kinetoplast stained in blue. Scale bars are 10 μm.

5.2.3 Simultaneous Expression of Tagged FLA2 (FLA2-Ty1-NG) and Tagged FLA3 (FLA3-HA)

One of the main reasons for tagging FLA2 was to facilitate an investigation of the possible interaction between FLA2 and FLA3. To further this end, FLA2 was tagged with a Ty1-NG-Ty1 tag, exactly as described in Section 5.3, in trypanosomes expressing FLA3-HA (p2708, see Section 3.4). The cells were transfected with the pPOTv6 amplicon and subjected to selection with blasticidin and G418. Several clones were generated. This tagging attempt was performed at the same time as the previous attempt (Section 5.2.2), before it became clear that FLA2-Ty1-NG does not localise to the flagellar attachment zone.

Expression of FLA2-Ty1-NG at the protein level in these cells was confirmed by flow cytometry (Fig. 5.7). Fluorescence data for live clonal cells (Clone C2) were acquired on an Accuri C6 flow cytometer. Of the 17,299 intact tagged cells measured, 98.0% showed positive for mNeonGreen fluorescence, confirming that the clone expressed FLA2-Ty1-NG.

There was very little background in parental wildtype cells (Fig. 5.7 C, D). A comparison of the histograms (Fig. 5.7 B vs D) shows that the median fluorescence intensity of the tagged cells was higher than that of the control cells by nearly two orders of magnitude.

Western blotting showed that FLA2-Ty1-NG migrated as a ~140-kDa band (Fig. 5.8 A). Weaker bands at ~100 kDa and ~50 kDa seemed to indicate breakdown products of FLA2-Ty1-NG. Reprobing the blot for HA showed that the cells still expressed FLA3-HA (Fig. 5.8 B). It seemed, therefore, that this cell line would be useful for colocalisation and interaction studies. However, it was first necessary to ascertain the localisation of FLA2-Ty1-NG in these cells.

The cells were fixed with PFA and probed for Ty1 by indirect immunofluorescence (Fig. 5.9 A). Strong red signal throughout the cell body suggested that FLA2-Ty1-NG had mislocalised, just as it had in the previous tagging attempt (Section 5.2.2). Altering the fixation method (Fig. 5.9 B, C) had no effect on the localisation of FLA2-Ty1-NG.

It was reasoned that if FLA2-Ty1-NG did not localise to the flagellar attachment zone, it was not an accurate representation of the wildtype protein and was therefore unsuitable for further studies.

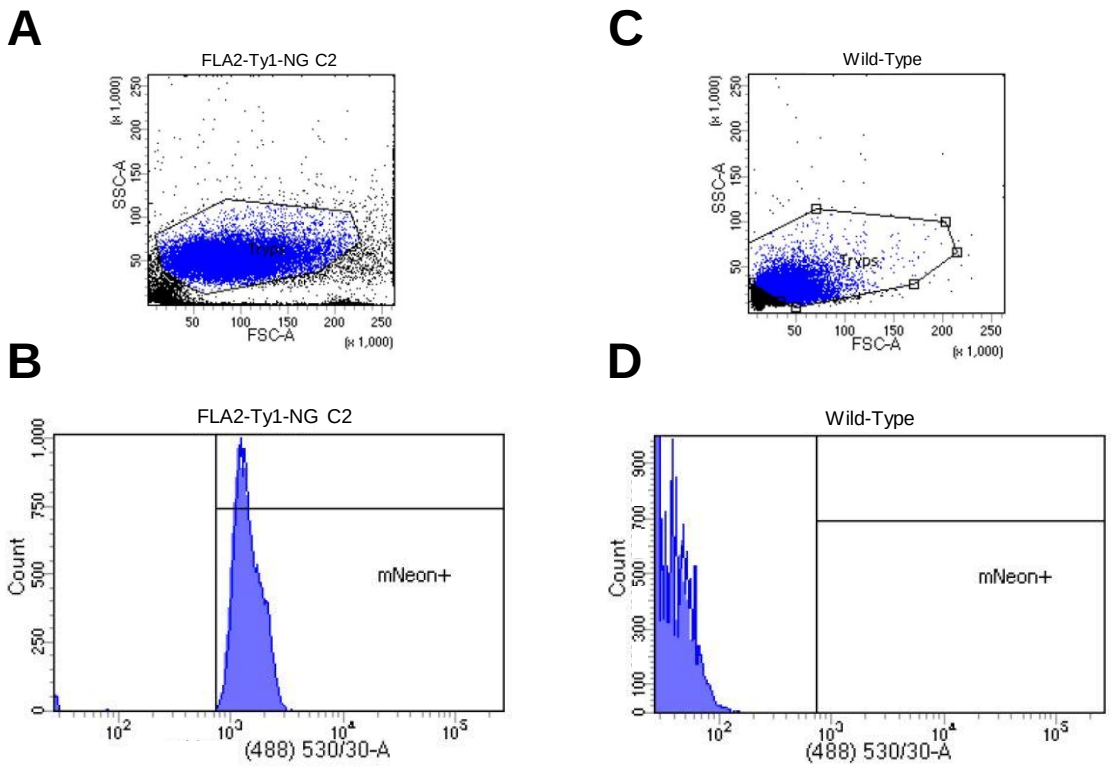


Figure 5.7: Expression of FLA2-Ty1-NG in a cell line expressing FLA3-HA.

To facilitate interaction studies, FLA2 was tagged in cells that expressed tagged FLA3. Long-primer PCR tagging with pPOTv6 was used to add a Ty1-NG-Ty1 tag to the C-terminal end of *FLA2* in bloodstream-form cells already expressing FLA3-HA (p2708). Expression of FLA2-Ty1-NG at the protein level was confirmed by flow cytometry of a representative clone (C2) (see Methods). **(A)** Intact cells were gated according to size. **(B)** Green fluorescence was measured according to an excitation wavelength of 488 nm and an emission range of 530/30. A histogram of fluorescence intensity shows that the cells emitted green fluorescence. Of the 17,299 intact cells measured, 98.0% showed positive for mNeonGreen fluorescence.

(C, D) Wildtype cells were analysed by flow cytometry in the same manner. Of 24,891 intact cells measured, only 0.4% were within the mNeonGreen threshold.

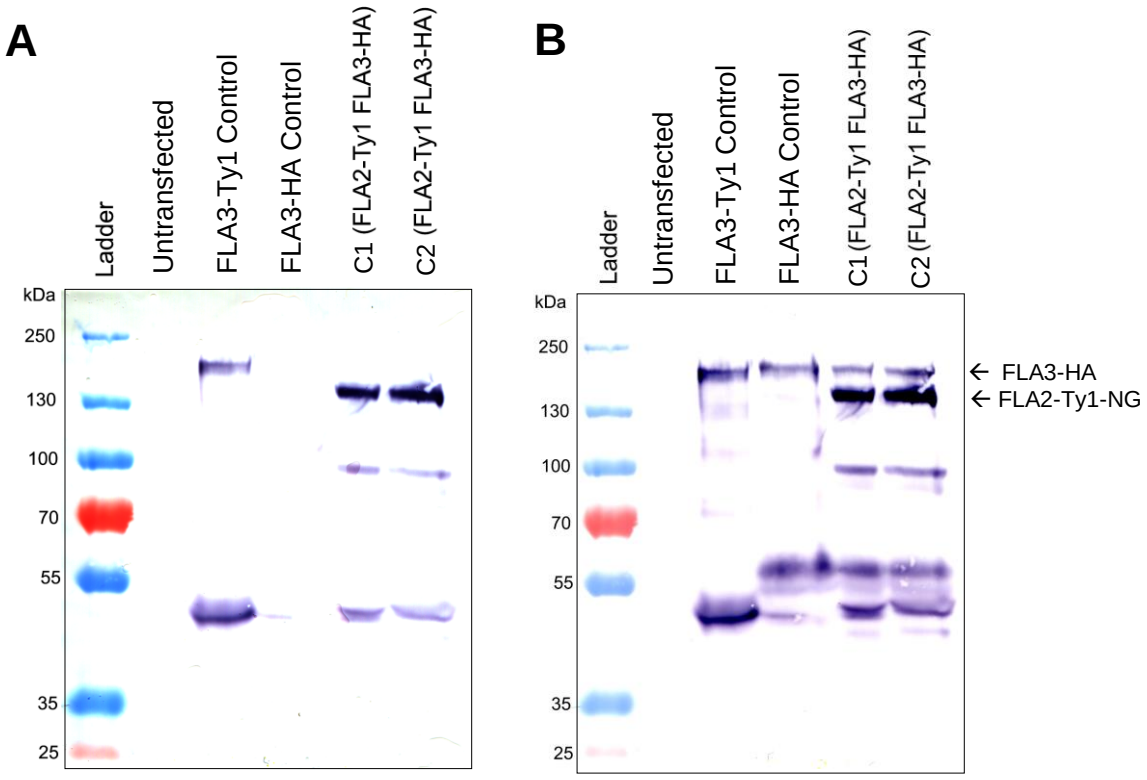


Figure 5.8: Simultaneous expression of FLA2-Ty1-NG and FLA3-HA.

To confirm expression of both FLA2-Ty1-NG and FLA3-HA in the cell line described in Fig. 7, samples of two clones (C1, C2) were blotted onto membrane. **(A)** Probing for Ty1 confirms expression of FLA2-Ty1-NG in both clones (C1, C2). **(B)** Reprobe of the same western blot confirms simultaneous expression of FLA3-HA.

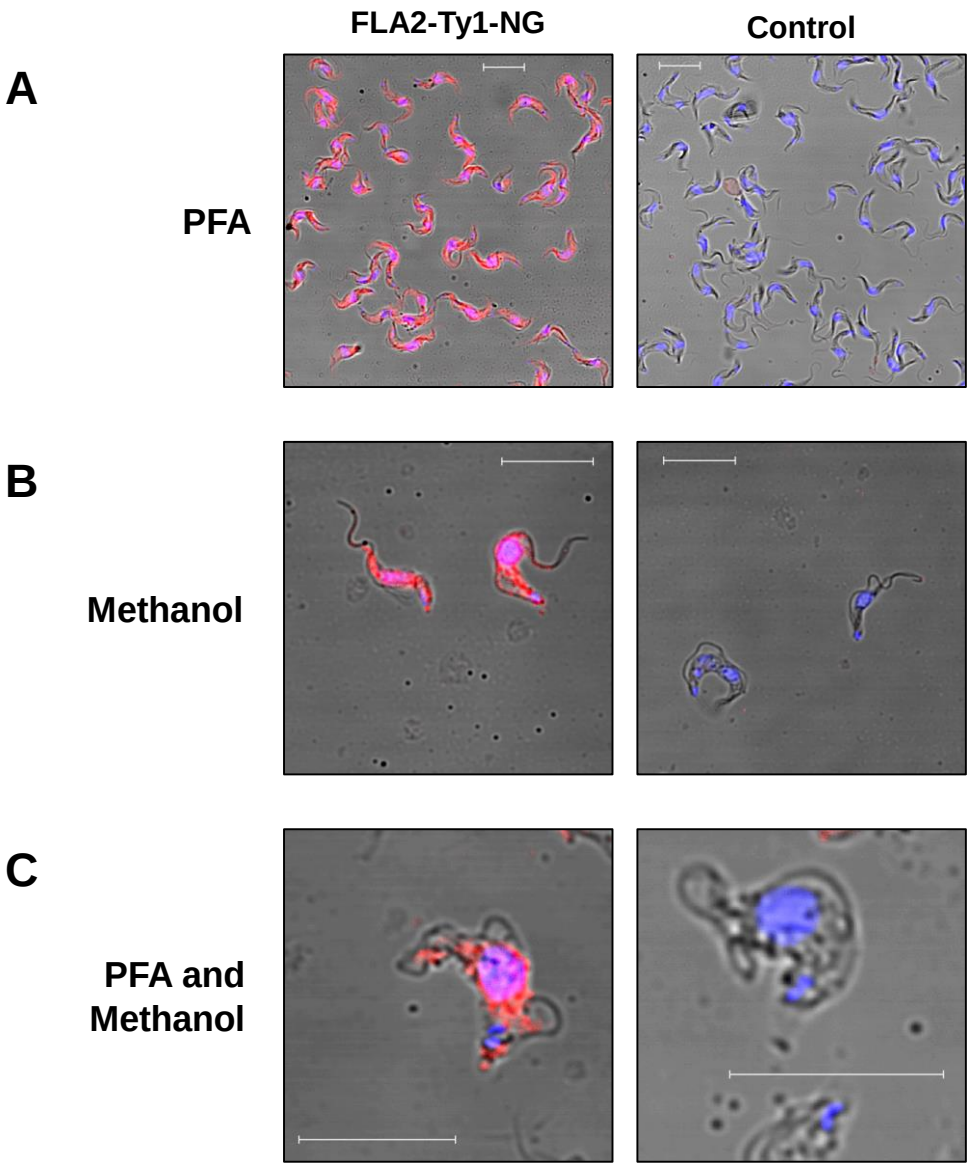


Figure 5.9: Localisation of FLA2-Ty1-NG is the same in different cell lines and under different fixation conditions.

(A) Bloodstream-form cells expressing FLA2-Ty1-NG and FLA3-HA (Clone 2, described in Fig. 7) were fixed with PFA. Immunofluorescence showed strong Ty1 signal (red) throughout the cell body. This localisation pattern was the same as had been seen for FLA2-Ty1-NG in a FLA2-RNAi cell line (Fig. 6). Fixation with methanol **(B)** or with PFA and methanol **(C)** did not alter the localisation pattern. Control cells, from the parental cell line expressing FLA3-HA, did not show red fluorescence when probed.

Nucleus and kinetoplast are stained in blue. All scale bars are 10 µm.

5.2.4 Long-Primer PCR Tagging of FLA2 with a (3×Ty1) Epitope Tag

Thus far, each attempt to tag FLA2 had been unsuccessful because the tag appeared to interfere with localisation of the protein. The tags used (mScarlet / mNeonGreen / Ty1-mNeonGreen-Ty1) each contained a monomeric fluorescent protein and added ~30 kDa to FLA2. It was hypothesised that these large tags interfered with entry of FLA2 into the FAZ, and that a smaller tag might therefore be more successful.

A modified long-primer PCR tagging approach using the same pPOTv6 vector was used to add a 3×Ty1 tag (5.5 kDa) to the C-terminus of the endogenous *FLA2* gene in wildtype cells. This approach employed a forward primer that generated an amplicon that lacked the mNeonGreen but retained the 3×Ty1 tag (see Section 2.2.8).

Western blotting of transfected bloodstream-form cells (Fig. 5.10 A) demonstrated that FLA2-Ty1 was expressed at the protein level in three clones (denoted A1, A2, A3). In each clone, FLA2-Ty1 was observed as a single band of ~110 kDa. Taking into account the size of the tag (~5.5 kDa), FLA2-Ty1 appears to be ~40 kDa larger than the size (64 kDa) predicted from the primary sequence. This size difference is consistent with extensive post-translational modification.

As a preliminary investigation of the subcellular distribution of FLA2-Ty1, cell fractionation experiments were performed (Fig. 5.10 B, C). Upon solubilisation of the cell membrane with Triton X-100 detergent, tagged FLA2 was found exclusively in the supernatant fraction. Upon freeze-thaw fractionation, tagged FLA2 was found exclusively in the pellet fraction. Taken together, these results indicated that FLA2-Ty1 behaved as an integral membrane protein, consistent with the presence of the predicted transmembrane domain.

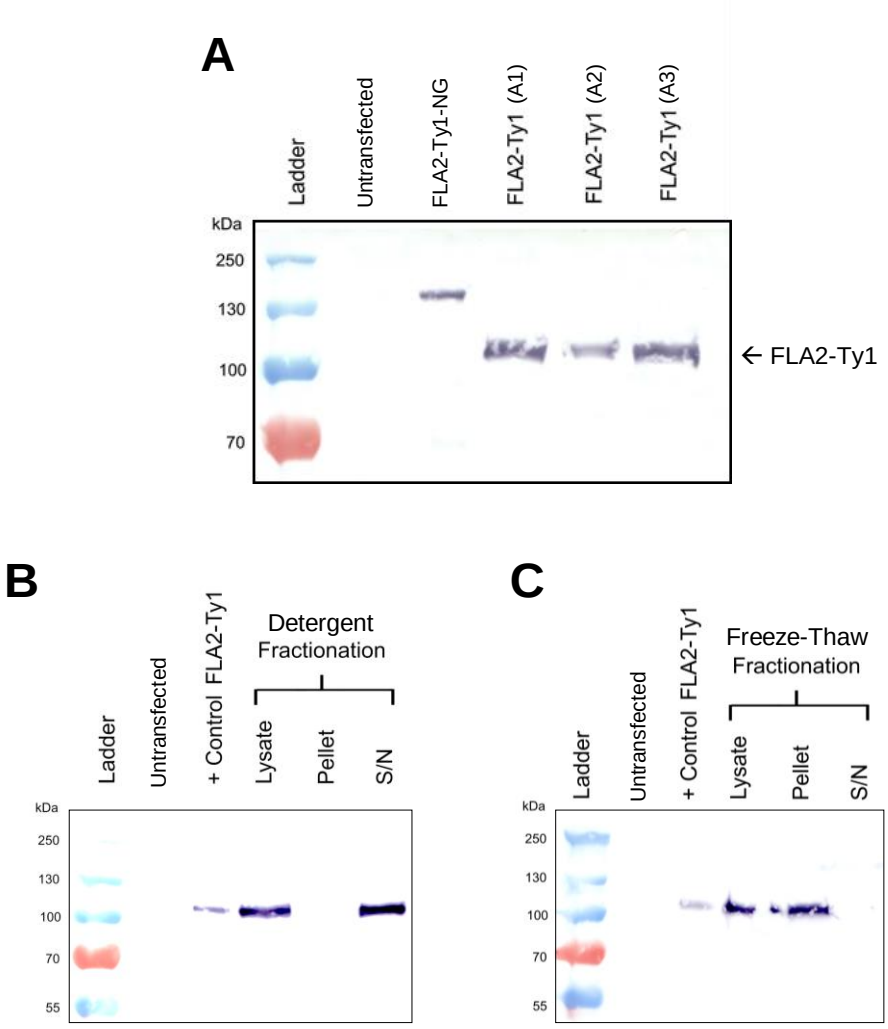


Figure 5.10: Expression of Ty1-tagged FLA2 without a fluorescence tag; distribution of the tagged protein after cell fractionation.

Long-primer PCR tagging with pPOTv6 was used to add a 3×Ty1 tag only (i.e. without a fluorescent protein) to the C-terminus of the *FLA2* gene at the endogenous locus in wildtype BSFs. **(A)** Western blot probed for Ty1 confirms expression of FLA2-Ty1 in three clones (A1, A2, A3). FLA2-Ty1 appears at ~110 kDa. **(B)** Western blot probed for Ty1 confirms location of FLA2-Ty1 in the supernatant (S/N) fraction after detergent fractionation of Clone A1, showing that the protein is solubilised. **(C)** Upon freeze-thaw fractionation of Clone A1, the protein is found in the pellet fraction, because it remains anchored in the membrane system.

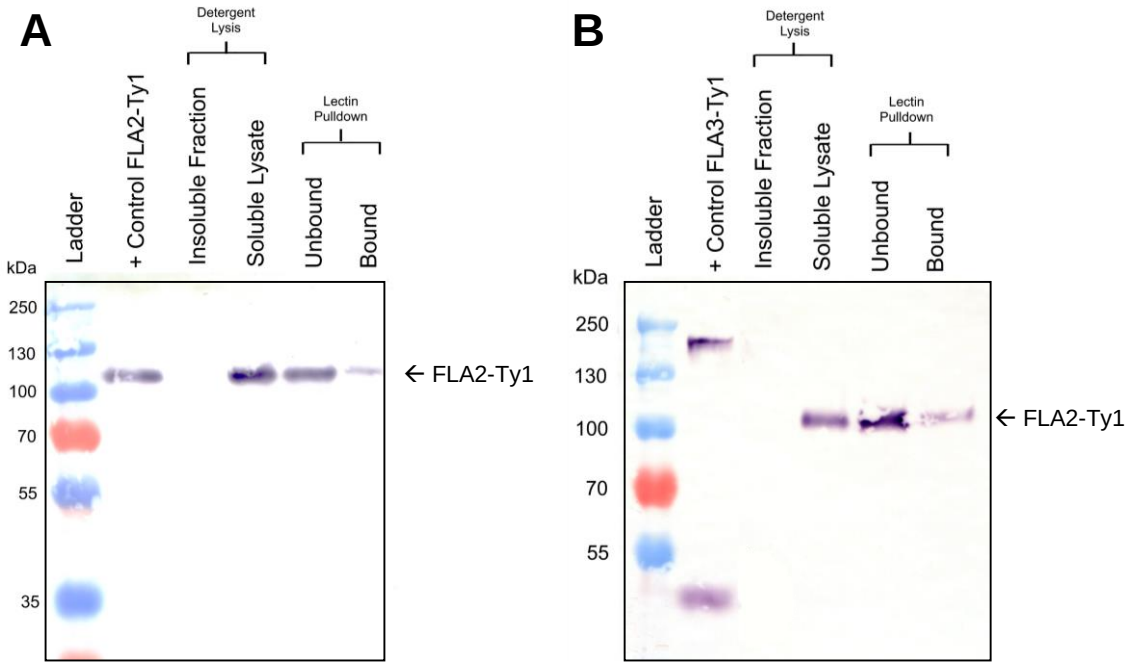


Figure 5.11: Lectin pulldown of FLA2-Ty1.

(A, B) Western blot replicates probed for Ty1 confirm pulldown of FLA2-Ty1 by RCA-I lectin. Bloodstream-form cells (Clone A3) expressing FLA2-Ty1 (3×Ty1) were incubated in Triton X-100 detergent and divided into detergent-insoluble and detergent-soluble fractions. The Soluble Lysate sample was subsequently incubated with lectin beads and fractionated into supernatant (Unbound) and pellet (Bound) fractions. The blots shows that some of the FLA2-Ty1 was found in the bound fraction. (Note that the bound samples represent $1.4x10^8$ cells (A) and $7.5x10^7$ cells (B), respectively. The lysate and unbound samples represent $5x10^6$ cells and the insoluble fractions represent $1x10^7$ cells.)

5.2.5 Lectin Pulldown of FLA2-Ty1 (3×Ty1)

As described above, tagged FLA2 migrated as a single band ~40 kDa larger than the size predicted by the primary sequence, suggesting that the protein was post-translationally modified. This evidence of extensive modification suggested that trafficking of the nascent protein through the ER and Golgi apparatus occurred correctly. Indeed, since the epitope tag was added to the cytoplasmic domain of FLA2 and the predicted glycosylated region is in the extracellular domain of the protein, there was no reason for tagging to interfere with normal processing of the protein in the secretory pathway. It was likely, therefore, that tagged FLA2 underwent the same post-translational modifications as wildtype FLA2 – most likely *N*-glycosylation and *O*-glycosylation, as occur on FLA1 (Nozaki et al., 1996). It was decided, therefore, to investigate the glycosylation of FLA2-Ty1 by means of a lectin pulldown, as had been performed for the FLA3 protein in Section 4.4.

One of the clones expressing FLA2-Ty1 (3×Ty1; Clone A3) was solubilised with Triton X-100 detergent. The lysate was incubated with RCA-I lectin beads, and fractionated (see Methods). Western blotting revealed that some of the FLA2-Ty1 was found in the lectin-bound fraction, although most remained in the unbound fraction. To confirm that this distribution was representative, the experiment was repeated. Both replicates gave the same result (Fig. 5.11 A, B). It seems that FLA2 binds to RCA-I lectin, but more weakly than FLA3 (*cf.* Fig. 4.6).

This result suggests that FLA2 contains galactose and/or *N*-acetylgalactosamine residues in its glycan side chains. Alternatively, FLA2 may be found in the lectin-bound fraction solely because it interacts with a lectin-binding protein (such as FLA3). This hypothesis would explain why the interaction between FLA2 and RCA-I is comparatively weak.

5.2.6 Localisation of FLA2-Ty1 (3×Ty1)

The primary function of the cell line expressing FLA2-Ty1 (described in Section 5.5) was to facilitate localisation of FLA2. The cells were therefore fixed with PFA and probed for Ty1 by indirect immunofluorescence. A representative clone (Clone A3) is shown in Fig. 5.12. Strong Ty1 signal was observed throughout the cell body (Fig. 5.12), just as had been observed previously when FLA2 was tagged with Ty1-mNeonGreen-Ty1.

It seemed, therefore, that the 3×Ty1 tag impaired localisation of FLA2 to the flagellar attachment zone, despite its small size.

5.2.7 Putative Carbohydrate-Deficient Clone of FLA2-Ty1 (3×Ty1)

Unexpectedly, one clone of the BSF cell line expressing FLA2-Ty1 (3×Ty1) (described in Sections 5.5–5.7) behaved differently to the other clones on a western blot.

When the five positive transfected clones (denoted A1–A5) were probed for Ty1 (Fig. 5.10 A; 5.13 A), FLA2-Ty1 appeared as a single band in each case. In four of these clones, the FLA2-Ty1 band was ~110 kDa, as stated earlier. However, in Clone A4, the FLA2-Ty1 band was only ~90 kDa. This band was still considerably larger than the predicted molecular weight (64 kDa), even allowing for the weight of the Ty1 tag (~5.5 kDa), indicating that significant post-translational modification had occurred. It was hypothesised that in Clone A4, FLA2-Ty1 was expressed but underwent incomplete processing and was therefore only partially glycosylated.

It was postulated that if the depleted size of the FLA2-Ty1 band in Clone A4 was indeed due to deficient glycosylation, the protein might lack the ability to bind to RCA-I lectin. A lectin pulldown was therefore performed, exactly as in Section 5.6. In Clone A4, no FLA2-Ty1 was observed in the RCA-I lectin-bound fraction; in Clone A5, much of the FLA2-Ty1 was found in the lectin-bound fraction (Fig. 5.13 D). This was a clear demonstration that the smaller version of FLA2-Ty1 lacked the ability to bind to RCA-I. It seems, therefore, that Clone A4 expresses a (presumably artefactual) version of FLA2-Ty1 that is inadequately glycosylated.

To see whether the differential modification of FLA2-Ty1 affected localisation of the protein, Clones A4 and A5 were probed for Ty1 by indirect immunofluorescence. Both clones showed strong red signal throughout the cell body (Fig. 5.13 B, C), as had been observed previously in Clones A1–A3. It seemed that the differential modification of FLA2-Ty1 had no effect on the distribution of the protein within the cell.

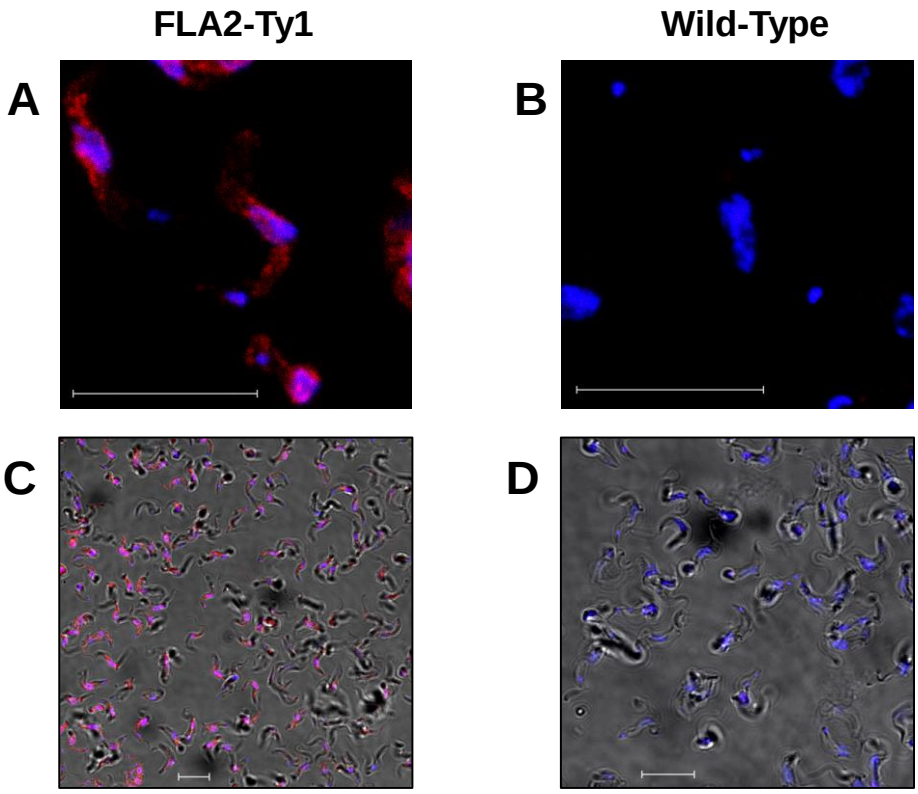


Figure 5.12: Localisation of Ty1-tagged FLA2 (3×Ty1) in bloodstream-form cells.

Bloodstream-form cells (Clone A3; described in Fig. 10) expressing FLA2-Ty1 (3×Ty1) without a fluorescent protein tag were probed by immunofluorescence to determine the localisation of FLA2-Ty1.

(A, C) Upon probing for Ty1 (red), the cells show a strong signal throughout the cell body. **(B, D)** Wildtype cells show no signal when probed for Ty1.

Nucleus and kinetoplast stained in blue. Scale bars are 10 μm.

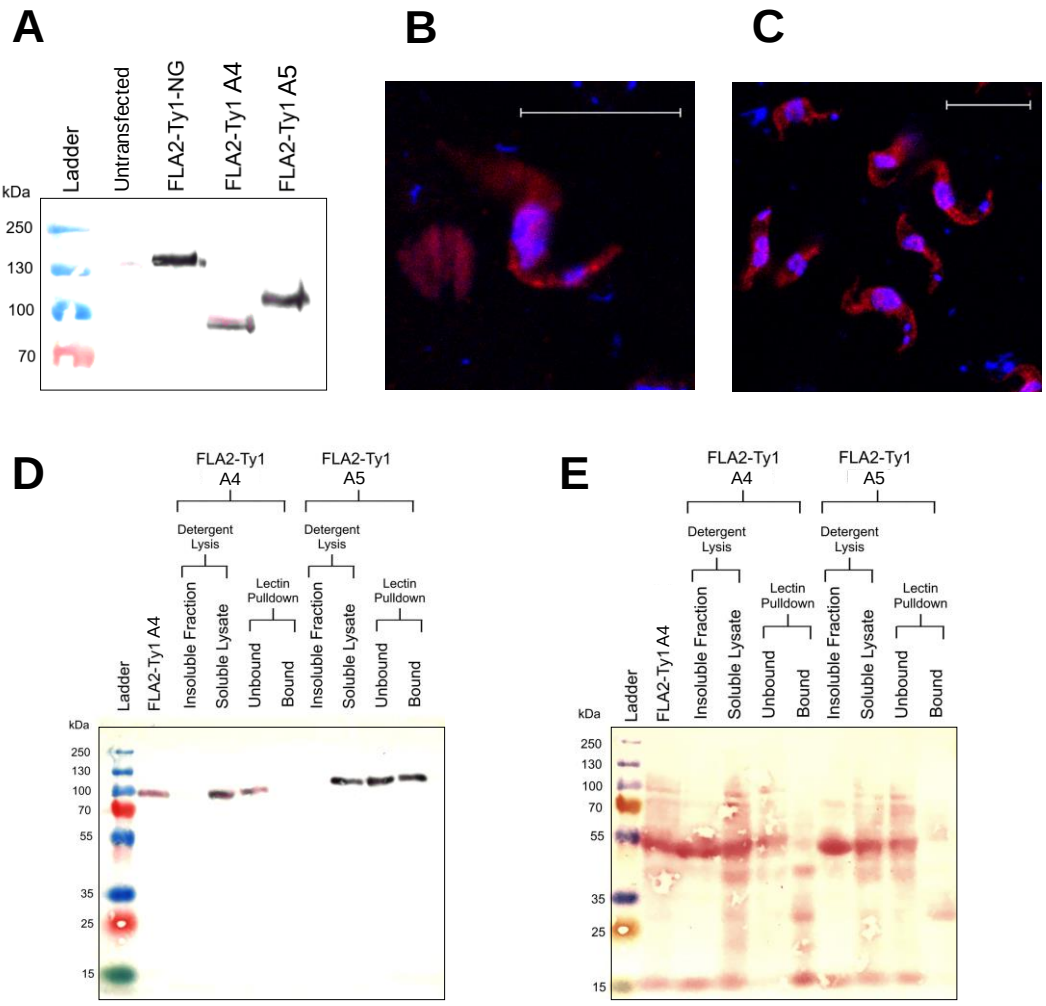


Figure 5.13: Putative carbohydrate-deficient clone of FLA2-Ty1 (3×Ty1) fails to bind to RCA-I lectin.

(A) Western blot of the cell line described in Fig. 10 demonstrates that one clone (Clone A4) expresses FLA2-Ty1 at a lower apparent molecular weight (~90 kDa) than the other transfected clones (represented here by Clone A5), which express FLA2-Ty1 at ~110 kDa. Immunofluorescence microscopy of Clones A4 **(B)** and A5 **(C)** does not show an observable difference between the two clones when probed for Ty1. **(D)** A lectin-pulldown experiment demonstrates that the FLA2-Ty1 expressed by Clone A4 does not bind to RCA-I lectin, whereas that expressed by Clone A5 binds strongly to the lectin. (Note that the bound samples represent $8.4x10^7$ cells (A4) and $1.2x10^8$ cells (A5), respectively.) **(E)** Ponceau stain of the same blot is shown as loading control.

5.2.8 Long-Primer PCR Tagging of FLA2 with a (1×Ty1) Epitope Tag

In an attempt to overcome the problem of mislocalisation of tagged FLA2, it was decided to tag FLA2 with the smallest possible tag – a single Ty1 epitope (1.2 kDa). It was hypothesised that this small tag would cause little steric hindrance and might therefore allow FLA2-Ty1 to localise in the same manner as the wildtype protein.

Long-primer PCR tagging with the pPOTv6 vector was used to add a 1×Ty1 tag to the C-terminus of the endogenous *FLA2* gene in wildtype cells. As before, a new forward primer was used that ensured that the tagging amplicon would confer the desired Ty1 tag only (see Methods). In fact, two tagging attempts were made, each with a slightly different forward primer. The tagging amplicon produced using primer "FP3" conferred a 1×Ty1 tag with a GS spacer between the *FLA2* gene and the tag. The tagging amplicon produced using primer "FP4" conferred a 1×Ty1 tag without a GS spacer. Each tagging amplicon was transfected, separately, into wildtype bloodstream-form cells; each transfection produced one respective clone. The clones were named FP3 and FP4 to indicate which forward primer had been used in each case.

A western blot probed for Ty1 confirmed that FLA2-Ty1 (1×Ty1) was expressed at the protein level in both clones (Fig. 5.14). In each case, FLA2-Ty1 appeared as a single band of ~100 kDa. The size of the band was suggestive of correct post-translational modification. The GS spacer is of negligible size (0.2 kDa), so there was no apparent difference between the respective FLA2-Ty1 bands in Clones FP3 and FP4. In both clones, the FLA2-Ty1 (1×Ty1) band was noticeably weaker than the FLA2-Ty1 (3×Ty1) band or FLA2-Ty1-NG band in the control lanes. This fact is attributable to the decreased antibody avidity when the epitope is monomeric.

Localisation of FLA2-Ty1 (1×Ty1) in clones FP3 and FP4 was investigated by indirect immunofluorescence (Fig. 5.15). In both cases, discrete patches of Ty1 signal were observed within the cell body. It seemed that FLA2-Ty1 (1×Ty1) was probably restricted to the endosomal compartments of the cell, which may indicate that the tagged protein is targeted for degradation.

As the localisation of FLA2-Ty1 (1×Ty1) appeared to be aberrant, no further studies were performed in these cell lines.

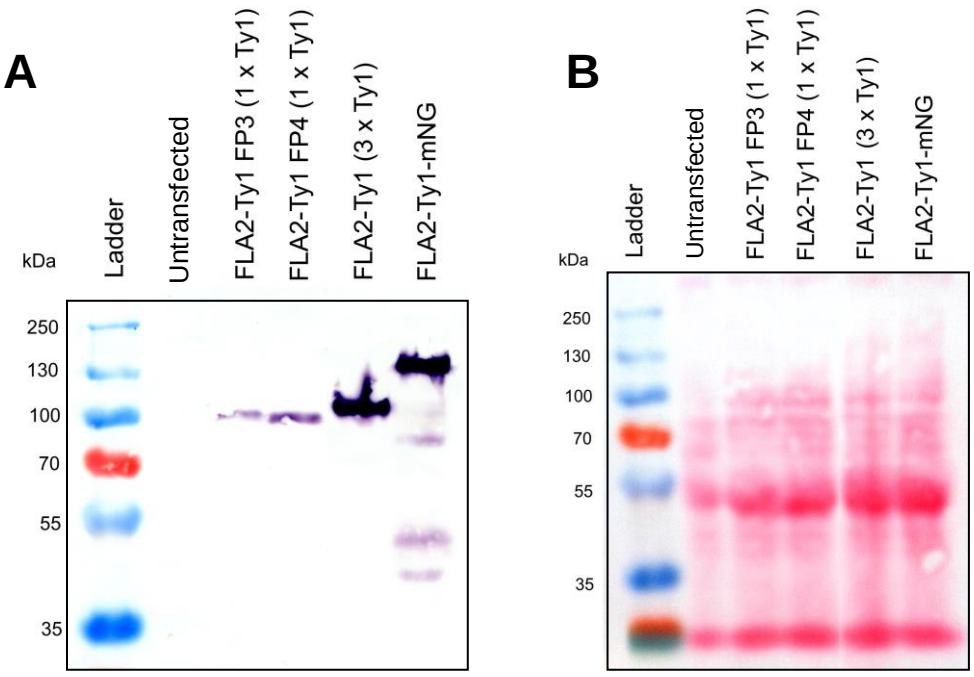


Figure 5.14: Expression of FLA2 with a single Ty1 residue in bloodstream-form cells.

Since epitope tagging at the C-terminus appeared to interfere with localisation of FLA2, an attempt was made to add the smallest tag possible. Long-primer PCR tagging was used to add a 1×Ty1 tag to the C-terminus of the *FLA2* gene at the endogenous locus. For the transfected clone denoted “FP3”, a GS spacer was included between the *FLA2* gene and the Ty1 residue; for clone “FP4”, the GS spacer was omitted.

(A) Western blot probed for Ty1 confirms expression of FLA2-Ty1 in both clones. FLA2-Ty1 (1×Ty1) appears as a ~100-kDa band. **(B)** Ponceau-stain of the same blot is shown as loading control.

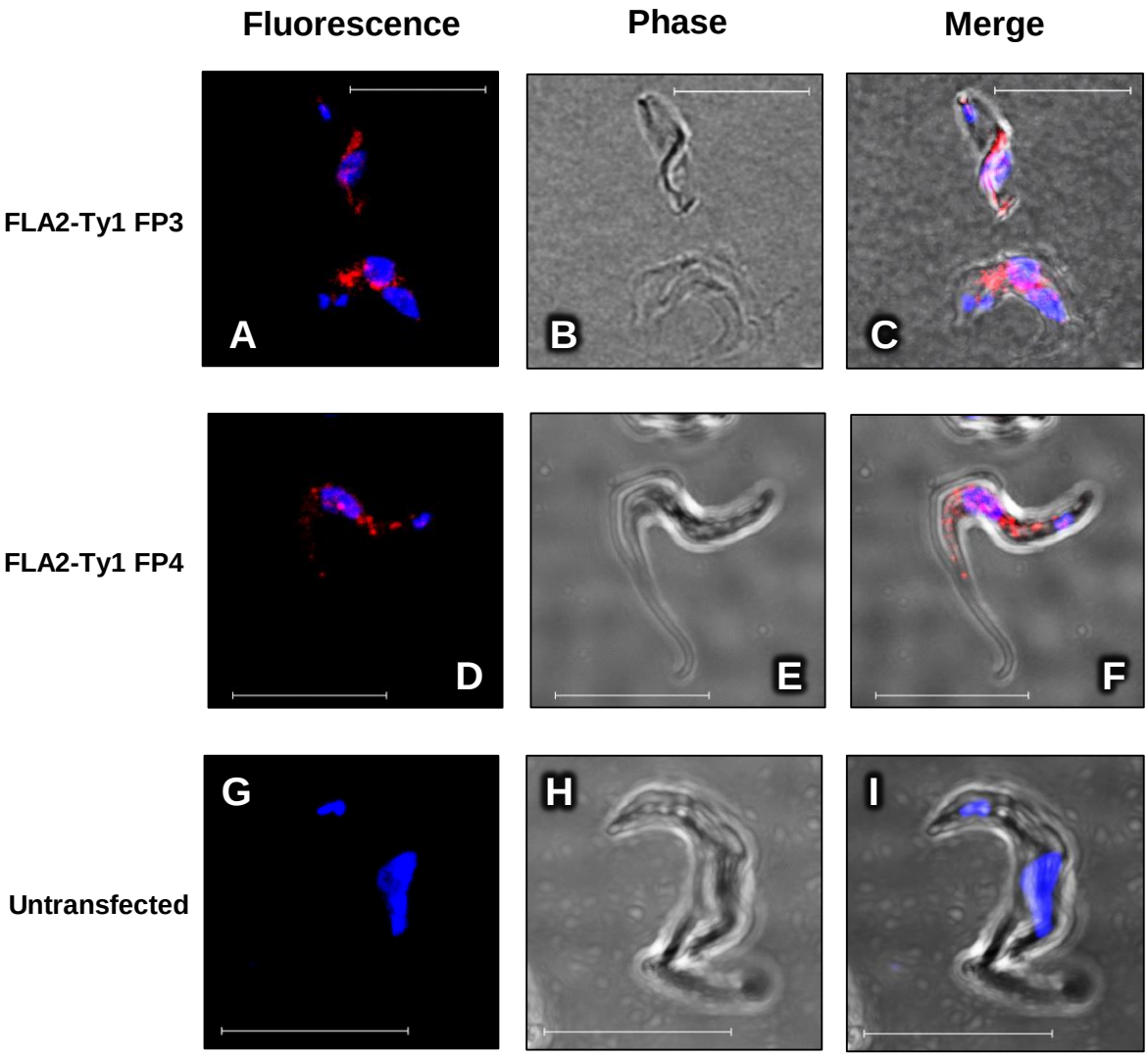


Figure 5.15: Localisation of FLA2 with a single Ty1 residue in bloodstream-form cells.

The clonal cell lines FP3 and FP4 (described in Fig. 14), which express FLA2-Ty1 (1×Ty1), were fixed with PFA.

(A–C) Upon probing for Ty1 (red), FP3 cells expressing FLA2-Ty1 (1×Ty1, with GS spacer) show a patchy, mostly post-nuclear, signal that may be indicative of endosomal/lysosomal localisation. **(D–F)** Upon probing for Ty1 (red), FP4 cells expressing FLA2-Ty1 (1×Ty1, without GS spacer) show similar signal.

(G–I) Wildtype cells show no signal when probed for Ty1.

Nucleus and kinetoplast stained in blue. Scale bars are 10 µm.

5.2.9 Replacement of the C-terminal Domain of FLA2 with an Epitope Tag (3×Ty1)

It was postulated that steric hindrance was the reason why the addition of various tags at the C-terminus of FLA2 has caused the protein to mislocalise: in other words, there was insufficient space to accommodate the tag in the narrow confines of the FAZ region. To circumvent this problem, the C-terminal tail of FLA2 (16 a.a.) was deleted and replaced by a Ty1 tag (3×Ty1). There was precedent for this unconventional approach: it had been shown previously that FLA1, the homologue of FLA2, can localise correctly to the FAZ even if its C-terminal tail is deleted (Sun et al., 2013).

As before, long-primer PCR tagging with the pPOTv6 vector was used to achieve the desired tagging of FLA2 ("FLA2 Δ C-Ty1"). This time, a new primer ("FP5") was used, which was homologous to the region of *FLA2* just upstream of the C-terminal tail. The tagging amplicon therefore replaced the C-terminal tail of FLA2 with a Ty1 tag (3×Ty1) in a single step (see Methods). This was performed in wildtype bloodstream-form cells.

Expression of FLA2 Δ C-Ty1 at the protein level was confirmed by western blotting of a transfected clone (Fig. 5.16). The single FLA2 Δ C-Ty1 (3×Ty1) band was ~110 kDa, the same size as FLA2-Ty1 (3×Ty1), because the loss of the C-terminal tail of the protein (1.8 kDa) does not significantly alter the molecular weight.

Localisation of FLA2 Δ C-Ty1 was investigated by indirect immunofluorescence. Upon probing for Ty1, the tagged protein appeared to be mislocalised (Fig. 5.17). The cells showed either a strong red signal throughout the cell body or a patchy, mostly post-nuclear, signal suggestive of lysosomal/endosomal localisation.

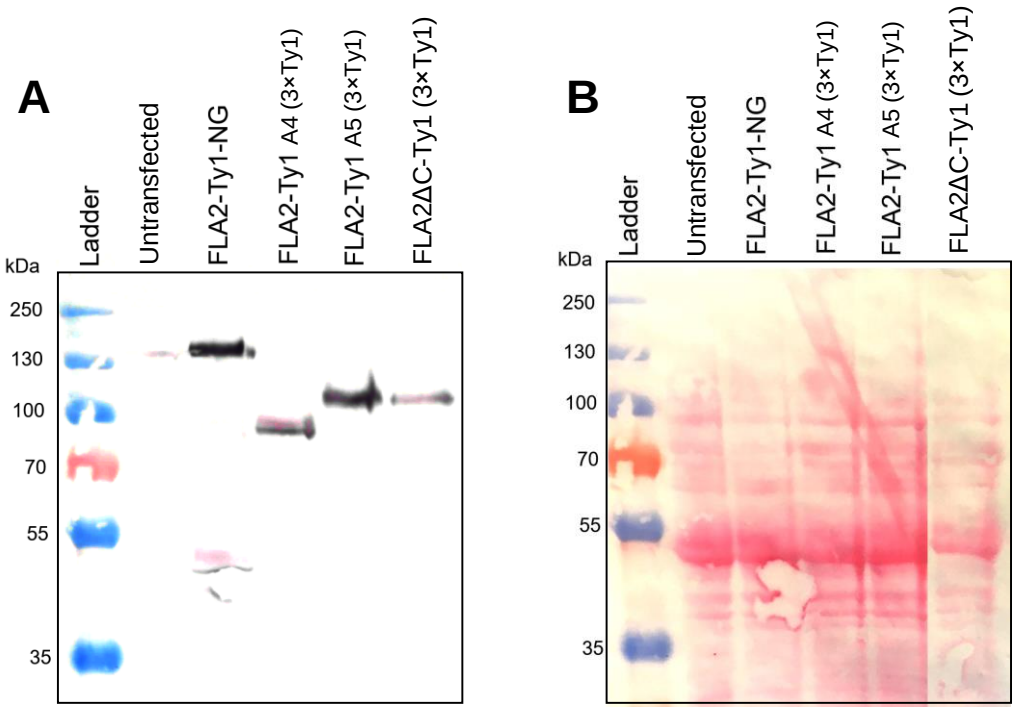


Figure 5.16: Replacement of the C-terminal tail of FLA2 with a (3×) Ty1 tag.
Long-primer PCR tagging was used to replace the C-terminal end of *FLA2* (i.e. the 16 amino acids downstream of the transmembrane domain) with a (3×) Ty1 tag in wildtype bloodstream-form cells (see Methods).

(A) Western blot probed for Ty1 shows expression of FLA2ΔC-Ty1 (3×Ty1) in one transfected clone (rightmost lane). FLA2ΔC-Ty1 (3×) appears as a ~110-kDa band, the same size as FLA2-Ty1 (3×), since the loss of the C-terminal tail of the protein does not significantly alter the molecular weight.

(B) Ponceau stain of the same blot is shown as loading control.

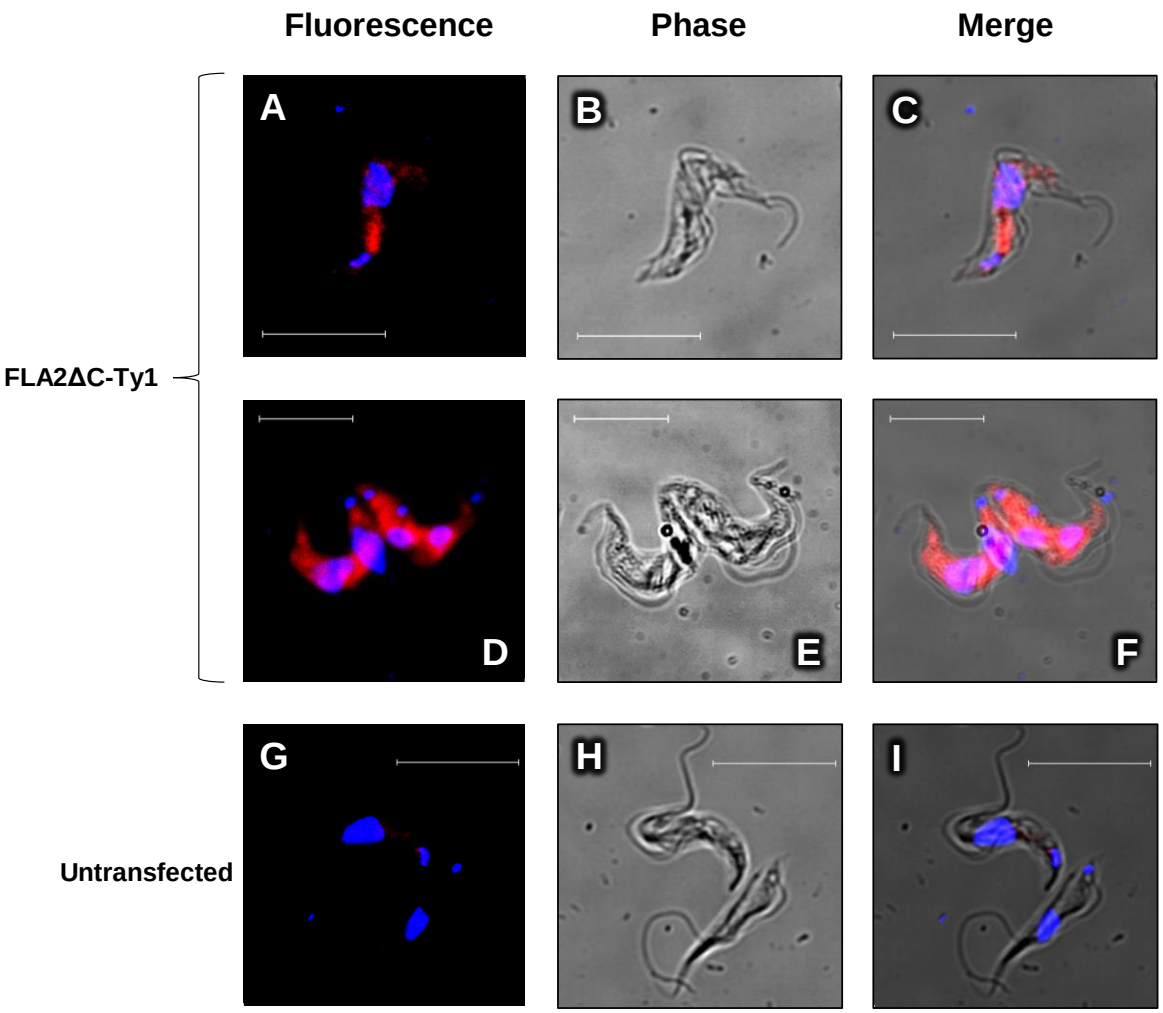


Figure 5.17: Localisation of a (3×Ty1)-tagged FLA2 mutant that lacks a C-terminal tail.

The clone (described in Fig. 16) expressing FLA2ΔC-Ty1 (3×Ty1) was fixed with PFA and examined by immunofluorescence. Upon probing for Ty1 (red), the cells sometimes show a patchy, mostly post-nuclear, signal suggestive of endosomal/lysosomal localisation **(A–C**), and sometimes a mislocalised signal throughout the cell body **(D–F)**.

(G–I) Wildtype cells show no signal when probed for Ty1.

Nucleus and kinetoplast stained in blue. Scale bars are 10 μm.

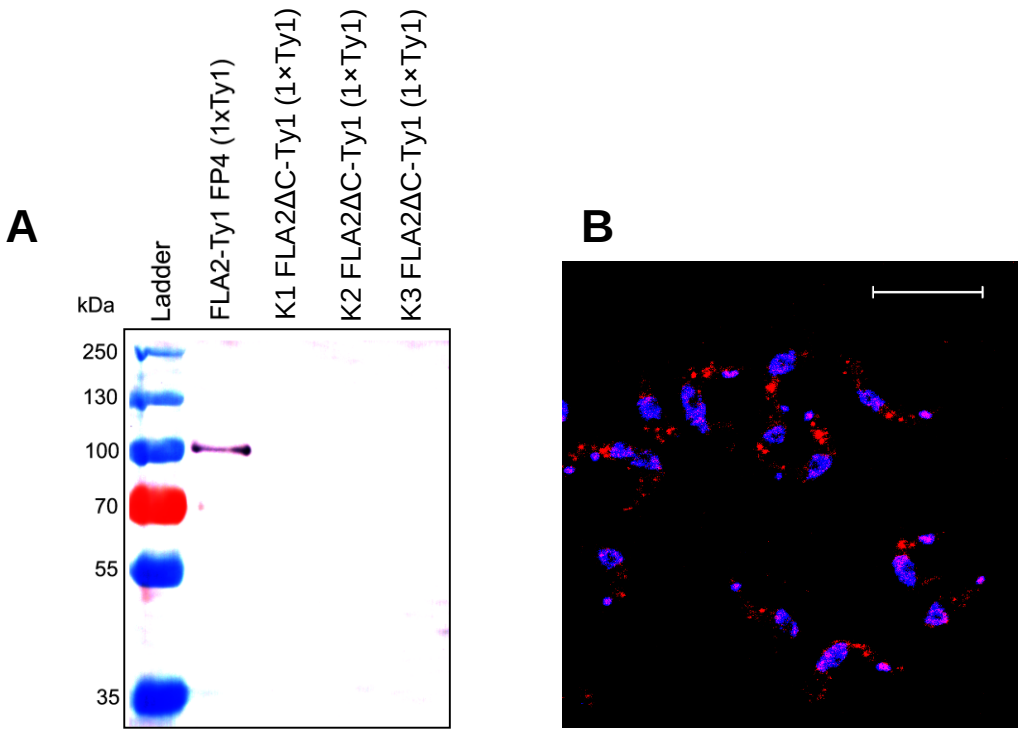


Figure 5.18: Replacement of the C-terminal tail of FLA2 with a single Ty1 residue.

Long-primer PCR tagging was used to replace the C-terminal end of *FLA2* (i.e. the 16 amino acids downstream of the transmembrane domain) with a single Ty1 residue in wildtype bloodstream-form cells (see Methods).

(A) Western blot probed for Ty1 fails to show expression of FLA2ΔC-Ty1 in any of three clones (K1, K2, K3).

(B) However, upon probing for Ty1 (red), fixed cells (Clone K2) show small patches of signal in the posterior end of the cell.

Nucleus and kinetoplast stained in blue. Scale bar is 10 μm.

5.2.10 Replacement of the C-terminal Domain of FLA2 with an Epitope Tag (1×Ty1)

As a final attempt to tag and localise FLA2, the C-terminal tail of the protein was replaced with a single Ty1 residue. It was reasoned that this small tag would cause the least steric hindrance possible. Tagging was achieved exactly as described in Section 5.10 but using a forward primer (FP6) that ensured that the epitope tag consisted only of a single Ty1 residue.

The tagging amplicon was transfected into wildtype bloodstream-form cells. Several clones were generated that were resistant to the selection antibiotic (Clones K1, K2, K3). However, a western blot probed for Ty1 failed to reveal any FLA2ΔC-Ty1 (1×Ty1) bands (Fig. 5.18 A).

Indirect immunofluorescence was performed to ascertain whether FLA2ΔC-Ty1 (1×Ty1) was expressed at the protein level (Fig. 5.18 B). Upon probing for Ty1, the transfected cells (Clone K2) showed discrete patches of signal in the posterior end of the cell. It appeared that FLA2ΔC-Ty1 (1×Ty1) was expressed, but the protein was mislocalised, possibly to the lysosome and endosomal compartments.

It was concluded that C-terminal tagging is not a suitable method of localising FLA2.

5.3 Discussion

5.3.1 In bloodstream-form cells, FLA2 performs the role previously ascribed to FLA1.

The FLA2 tagging experiments performed here did not determine the localisation pattern of FLA2. Collectively, however, the evidence generated in this thesis lends strong support to the hypothesis that FLA2 is the BSF paralogue of FLA1 and, in fact, that FLA1 is not produced at the protein level in the bloodstream stage of the *T. brucei* life cycle.

When FLA1 was first discovered, it seemed evident that the protein was present in both PCF and BSF life-cycle stages (Nozaki et al., 1996). Western blots probed with an anti-FLA1 antibody showed a single band in both stages; fixed PCFs and BSFs probed with the same antibody showed specific signal along the flagellar attachment zone. Furthermore, repeated attempts to produce a *FLA1* double knockout were unsuccessful, suggesting that the protein is essential in both life-cycle stages.

The subsequent finding that a related gene, *FLA2*, was transcribed in BSFs and was knocked down concomitantly with *FLA1* during RNAi experiments (LaCount et al., 2002) raised the possibility that FLA1 and FLA2 might both play a role of flagellar attachment in BSFs. (In fact, there are two *FLA2* genes, 98% identical to each other. Both genes are close neighbours of *FLA1* and *FLA1BP* on chromosome 8.) LaCount et al. (2002) suggested that FLA1 and FLA2 might be functionally redundant, or else that both proteins were required in the bloodstream-form cell. They did not consider the possibility that FLA1 might not be present in BSFs, probably because (a) Nozaki et al (1996) had reported apparently convincing evidence that FLA1 is present in BSFs and (b) there was still no evidence that FLA2 was actually produced at the protein level.

More recently, however, high-throughput proteomic studies have suggested that expression of the FLA1 protein is stage-specific. As summarised in Table 1.7, proteomic data show that the FLA1 protein is significantly more abundant in PCFs than in BSFs (Urbaniak et al., 2012, Butter et al., 2013). Indeed, FLA1 has essentially the same PCF:BSF ratio (11.77) as that of mitochondrial aspartate aminotransferase (mASAT) (11.62) (Urbaniak et al., 2012). Empirical results show that mASAT is virtually undetectable by western blot in bloodstream-form cells (Marciano et al., 2008). In contrast, FLA2 has a PCF:BSF ratio of 0.06 (Urbaniak et al., 2012), which is the same as that of GPI-specific phospholipase C (GPI-PLC), which is generally accepted to be BSF stage-specific (Carrington et al., 1998). It is therefore likely that in the bloodstream form, FLA2 performs the role in flagellar attachment that is performed by FLA1 in the procyclic form.

In the intervening eighteen years since LaCount et al. (2002) detected the *FLA2* transcript, no further studies of either FLA1 or FLA2 have been performed in the bloodstream form. Consequently, there is no direct, published evidence for the proposed functionality of FLA2, and there has been no explicit discussion of the data of Nozaki et al. (1996) in the context of the sequenced *T. brucei* genome.

The results presented in Chapter 3 demonstrated clearly that FLA2 is essential for flagellar attachment in BSFs (discussed in Section 3.3.2), while the results presented in Chapter 5 indicate indirectly that the protein observed by Nozaki et al (1996) in bloodstream-form cells, using the anti-FLA1 antibody, was actually FLA2, not FLA1. This hypothesis is consistent with all the published data. It implies that FLA1 and FLA2 are stage-specific

paralogues, which is more plausible than the earlier hypothesis (LaCount et al., 2002) that FLA1 and FLA2 are both required in BSFs, but only FLA1 is required in PCFs.

The tagging experiments presented in this chapter suggest that the molecular weight of mature FLA2 is probably ~100 kDa, which is the same size as the "FLA1" band seen by Nozaki et al. (1996) in BSF samples probed with anti-FLA1. Until now, the molecular weight of FLA2 was unknown.

The molecular weights of the various tagged versions of FLA2 generated here are summarised in Table 5.2.

Table 5.2:

Tagged FLA2	Apparent Size (approx.)	Size of Tag	Estimated Size Without Tag
FLA2-Ty1-NG	140 kDa	36 kDa	104 kDa
FLA2-Ty1 (3× Ty1)	110 kDa	5.5 kDa	105 kDa
FLA2-Ty1 (3× Ty1)*	90 kDa	5.5 kDa	85 kDa*
FLA2-Ty1 (1× Ty1)	100 kDa	2 kDa	98 kDa
FLA2ΔC-Ty1 (3× Ty1)	105 kDa	-4.2 kDa; +5.5 kDa	104 kDa

*putative carbohydrate-deficient clone

In each case, when the size of the tag is subtracted from the apparent size of the mature protein on a western blot, the molecular weight of the FLA2 is consistently ~100 kDa. As seen for FLA1, FLA3 and FLA1BP, the FLA2 protein migrates at a much larger size than is predicted from the amino acid sequence (61 kDa after cleavage of the putative signal sequence), probably due to extensive post-translational modification, most likely glycosylation. It seems probable that the ~100-kDa size observed here is representative of the mature wildtype protein, since the presence of a C-terminal tag on the intracellular part of the protein should not affect the processing of the extracellular/luminal domain, which is the location of all predicted *N*- and *O*-glycosylation sites. The data presented in Fig. 5.10 (B, C) confirm that FLA2-Ty1 (3× Ty1) behaves as a membrane protein, while Figs. 5.11 and 5.13 show that FLA2-Ty1 (3× Ty1) possesses glycan side chains that can bind

RCA-I lectin, so it certainly seems likely that tagged FLA2 is processed normally in the ER and Golgi. (The putative carbohydrate-deficient clone described in Section 5.2.7 is a marked exception to this statement.) It is therefore postulated here that mature wildtype FLA2 has a molecular weight of ~100 kDa.

As described in Section 1.3.1, a polyclonal antibody raised against a recombinant protein corresponding to part of the FLA1 open reading frame recognised a single diffuse 100-kDa band in bloodstream-form cells (Nozaki et al., 1996). It now seems likely that this band was FLA2, not FLA1, because it is very likely that the antibody in question would have cross-reacted with FLA2, given the high degree of sequence similarity between the two proteins. In fairness, this problem could not have been anticipated by Nozaki et al. (1996) because FLA2 was unknown at the time. The antibody was raised against residues 81–312 of the FLA1 protein sequence. As shown in Table 5.3, this region of FLA1 is 58% identical, and 75% similar, to the corresponding region of FLA2. No other antibody has since been raised against either FLA1 or FLA2.

```
Query 81      CDVGGGGSTVRLVNKTGIYTVAGNLAKRGDEVGPLEVARFNHPTSVVGVNNDIYVADRDN 140
              CDVGGGGSTVRLVNKTGIYT+AG+L +RG++ GP   A FN+PTSVV VN+DIYVADRDN
Sbjct 81      CDVGGGGSTVRLVNKTGIYTIAGSLTQRGNKDGPKGDALFNNPTSVVSVNDDIYVADRDN 140

Query 141     DCMKRIDADGMVTKFAANEVDKPSSIIYHEQDGAPVLFISDTGNSRIMYSQISVSNNVTA 200
              +C++RIDA+G VT++   +++KP  I+    +G   LFISDTGNS+I+Y  +  +N V
Sbjct 141     NCIRRIDAEGNVTRYGPQDLNKPKDILPFTLNGTQNLFISDTGNSQILYVPLDANNTVVT 200

Query 201     KLVGGFQPGVMQISKKRNFMYVVKETSWIAAVDLNSLSDSDGNKKSWDIANMSCLDYVDA 260
               LV GFQPGVMQISK++N MYVVK TSWIAAV+L+   +SD   KS DI +++CL Y  A
Sbjct 201     TLVAGFQPGVMQISKEKNMMYVVKNTSWIAAVNLSKTGESDIG-KSKDIGDVTCLHYKSA 259

Query 261     LMLTEDENELYYYGEPNGESHIMSLEL-NSSETGLLCPKKVMEWMHGPIVSLV 312
              LMLT+ E++LYYYGEPNGE +IMSLE+   S     C + +++W +  IVSL+
Sbjct 260     LMLTQAEDKLYYYGEPNGEGYIMSLEVERPSPASQWCAQSMLKWDYDRIVSLL 312
```

Identities:	135/233 **(58%)**
Positives:	176/233 **(75%)**
Gaps:	2/233 **(0%)**
Expect:	7e-92

Table 5.3: BLASTP alignment of residues 81–312 of *FLA1* against the *FLA2* gene.
Query = Residues 81–312 of *FLA1* (Tb927.8.4010)
Subject = The *FLA2* gene (Tb927.8.4110)
BLASTP performed using NCBI Blast.

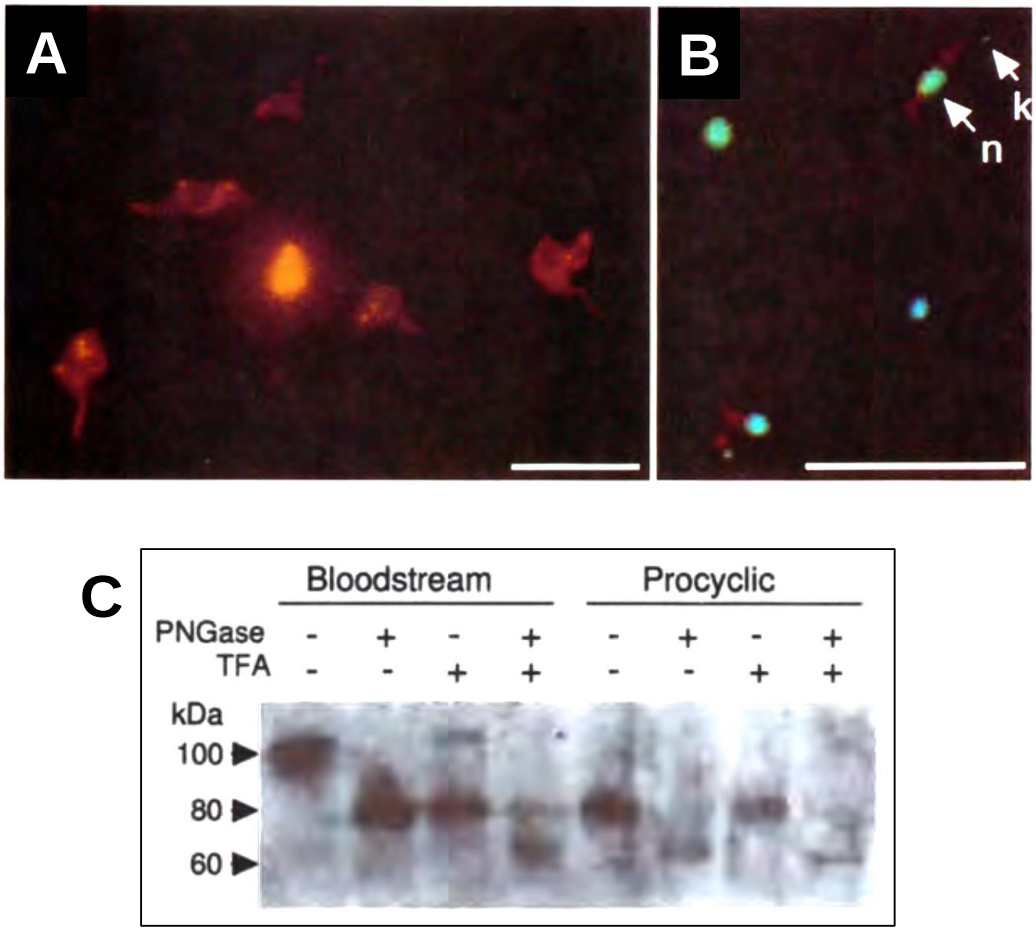


Fig. 5.19: **Investigations of FLA1 in bloodstream-form cells (reproduced from Nozaki et al. 1996).**

Fixed BSFs were probed with polyclonal anti-FLA1 **(A)** or with anti-FLA1 and Hoechst stain **(B)**. Kinetoplast = k; nucleus = n.

(C) Cell lysates were treated with PNGase F (to remove *N*-glycans), TFA (to remove *O*-glycans) or both. Samples were electrophoresed, blotted onto membrane and probed with polyclonal anti-FLA1 antibody.

The fact that the polyclonal antibody generated by Nozaki et al. was not specific for FLA1 necessitates re-interpretation of the data presented by Nozaki et al. (1996). Firstly, the immunofluorescence images in BSFs (reproduced in Fig. 5.19 A, B), which were presented in the study as FLA1, actually portray FLA2. Nozaki et al. describe the anti-FLA1 signal as being "concentrated along the flagellum"; however, the images are of low resolution and have high background signal. Therefore, FLA2 still needs to be tagged and localised to determine whether the distribution of the protein is continuous or punctate along the FAZ, and to enable investigation of the interaction partners of the protein (discussed further in Section 5.3.2).

Secondly, the suggestion that FLA1 is differentially glycosylated in different life-cycle stages (Nozaki et al., 1996) is probably incorrect. Nozaki et al. (1996) investigated the glycosylation of FLA1 by treating cell lysates with PNGase F (to remove *N*-linked glycans) or trifluoroacetic acid (TFA) (to remove *O*-linked glycans); their results are reproduced in Fig. 5.19 (C). In PCFs, it was observed that the 80-kDa FLA1 band was reduced to 60 kDa upon treatment with PNGase F; it was concluded that FLA1 has *N*-glycans. Treatment with TFA had little effect; it was concluded that FLA1 does not possess *O*-glycans in procyclic cells. In BSFs, it was observed that the 100-kDa "FLA1" band (postulated here to represent FLA2) was reduced to 80 kDa upon treatment with either PNGase F or TFA, or to 60 kDa upon treatment with both. It was concluded that FLA1 has *N*-glycans and *O*-glycans in bloodstream-form cells.

If, as postulated here, the 100-kDa band recognised by the anti-FLA1 antibody in BSFs actually represented FLA2, it is clear from Fig. 5.19 (C) that FLA2 possesses both *N*-glycans and *O*-glycans. This is a highly plausible suggestion because FLA2 has a 44-a.a. insertion, absent from FLA1, that is eminently suitable for *O*-linked glycosylation:

VPP**S**P**S**P**S**P**S**A**S**P**S**P**S**P**S**VQPPPPPPPPPPPPPPPPI**T**PNPDPP

This sequence contains nine serine/threonine residues (underlined) that could act as sites for *O*-linked glycosylation. *O*-glycosylation usually occurs on residues that are proximate to proline residues (Hema Thanka Christlet and Veluraja, 2001), and this insertion contains many proline residues.

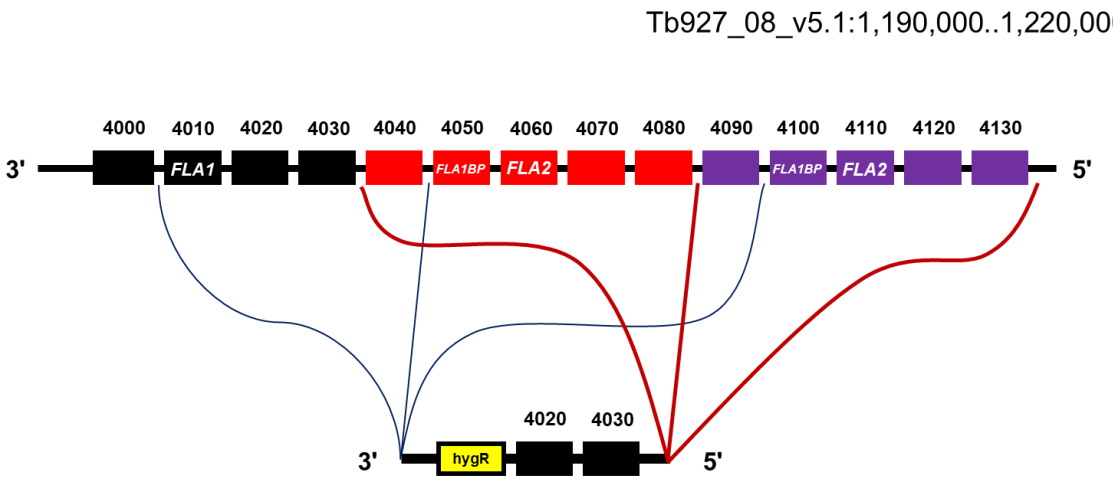


Fig. 5.20: *FLA1* knockout construct (Nozaki et al. 1996) could knock out *FLA2* also.

The *FLA1* knockout construct generated by Nozaki et al (1996), which contained a hygromycin resistance selection marker (hygR) could insert into multiple loci within the genomic location shown here. The 5' end of the *Bam*HI / *Bss*HII-digested knockout construct was identical to three distinct regions on chromosome 8, flanking the 5' ends of ORFs Tb927.8.4030, 4080 and 4130, respectively. The 3' end of the construct was likewise identical to three distinct regions on chromosome 8, flanking the 3' ends of ORFs Tb927.8.4010, 4050 and 4100, respectively.

Consequently, when the construct inserted into the genome by homologous recombination, it could knock out *FLA1* or either of the *FLA2* genes. There was also a possibility that the construct could recombine so as to knock out two of those genes simultaneously, or even to knock out all three genes at once.

One counterpoint to the hypothesis that FLA1 is not produced in BSFs is the fact that, despite repeated attempts, Nozaki et al. (1996) could not generate a (-/-) double knockout of *FLA1* in bloodstream-form cells (or in PCFs). Why should this be the case, if *FLA1* is not an essential gene in BSFs? Firstly, a negative result (failure to generate a double knockout) cannot constitute proof of a positive conclusion (essentiality of *FLA1*). Even more importantly, it now transpires that the *FLA1* knockout construct would have knocked out *FLA2* also (Fig. 5.20).

At the time of the discovery of FLA1, the *T. brucei* genome had not yet been sequenced, so Nozaki et al. were unaware that *FLA1* belongs to a section of chromosome 8 that has been duplicated twice, forming a tandem repeat, as depicted in Fig. 1.7. Briefly, the *FLA1* knockout construct was generated using a plasmid that contained a 4.8-kbp *Bam*HI–*Bgl* II fragment of chromosome 8, which consisted of the full *FLA1* CDS flanked by 2.2 kb on the 5' side and 0.9 kb on the 3' side. Outward-facing primers were used to amplify the whole plasmid except the *FLA1* CDS. An antibiotic-resistance gene was ligated into the resulting amplicon, creating a knockout vector. Prior to transfection into trypanosomes, the vector was restriction-digested with *Bam*HI and *Bss*HII to release a linear stretch of DNA that would insert into the trypanosome genome by homologous recombination.

Using the annotation of the TREU927 genome (version 5.1), the linear knockout construct consisted of the region Tb927_08_v5.1:1195816..1200453. A TriTryp BLAST search of the 300 bp at either extremity of this fragment shows 99–100% identity to corresponding regions on either side of each of the *FLA2* genes. In other words, this knockout construct could insert into any of three distinct loci on chromosome 8 (depicted schematically in Fig. 5.20). The construct could therefore knock out *FLA1* or *FLA2*, or even both at once (depending on where recombination happened to occur). Since there are two *FLA2* genes, the construct would in fact be more likely to knock out a *FLA2* allele than a *FLA1* allele. Indeed, given the length of the homologous ends, it is even possible that the construct could insert in such a way as to knock out the entire locus at once (i.e. Tb927.8.4010–4130). Furthermore, as shown in Fig. 5.20, if *FLA2* were knocked out, the genes Tb927.8.4070/4080 or Tb927.8.4120/4130 would be replaced by the related genes Tb927.8.4020/4030, with unknown consequences. The *FLA1BP* gene would also be knocked out.

Therefore, the fact that Nozaki et al. (1996) could not produce a double knockout using this vector is not an indication that FLA1 is required in the bloodstream form. Collectively, all the evidence suggests that FLA1 is PCF-specific and that FLA2 is the BSF-specific paralogue of FLA1. Since the level of *FLA1* mRNA in PCFs is only 2–4 times higher than in BSFs (Table 1.4), it seems that most control of FLA1 production in the bloodstream form must occur at the translational or post-translational levels.

The finding that the FLA1 protein is not utilised in the bloodstream-form cell emphasises the importance of FLA2. As an essential surface protein in the clinically relevant life-cycle stage of the parasite, FLA2 has been much overlooked and requires more thorough investigation.

5.3.2 Outlook for further investigations of FLA2

Since FLA1 and FLA2 are homologues, and since knockdown of FLA2 in BSFs (Section 3.2.1) produces the same phenotypic effect (flagellar detachment and cytokinesis failure) as knockdown of FLA1 in PCFs (Sun et al., 2013), it is extremely probable that the two proteins perform essentially the same function and that the differences between the proteins are adaptations to their respective life-cycle stages. In PCFs, FLA1 has been shown to locate to the cell-body membrane of the FAZ and to mediate flagellar attachment by binding to FLA1BP, a glycoprotein that has an apparently punctate distribution along the flagellar membrane of the FAZ (Sun et al., 2013). It therefore seems plausible that in BSFs, FLA2 likewise locates to the cell-body membrane of the FAZ and mediates flagellar attachment by binding to FLA3, the BSF homologue of FLA1BP, which has been shown conclusively in this thesis to have punctate distribution along the flagellar membrane of the FAZ. If FLA2 and FLA3 are indeed interaction partners, it is probable that FLA2 has the same punctate spacing as FLA3 (177 nm ±50, Section 3.3.7).

However, FLA2 has yet to be located satisfactorily, and no interaction has yet been shown between FLA2 and FLA3 (or between FLA2 and any other FAZ protein, for that matter). To enable these questions to be resolved, FLA2 will have to be tagged in a way that does not interfere with the normal processing, localisation and behaviour of the protein.

The results presented here (Section 5.2) show that C-terminal tagging is not a suitable method to determine the true localisation of FLA2. Whether FLA2 was tagged at the C-terminus with a simple epitope tag or a large fluorescent protein – and even if the intracellular tail of the protein was removed to make more room for the tag – the tagged version of the protein appeared to mislocalise.

In the cases of FLA2-SR, FLA2-NG, FLA2-Ty1-NG, FLA2-Ty1 (3×Ty1) and FLA2ΔC-Ty1 (3×Ty1), the tagged protein was distributed throughout the cell body. It is not clear whether these tagged versions of FLA2 ever reach the plasma membrane or if, rather, they remain in the secretory system. A surface biotinylation experiment would have resolved this question. The surface of the cell could be labelled with membrane-impermeable NHS-biotin (as performed in Section 4.2.2); subsequent cell lysis and immunoprecipitation with anti-Ty1 would reveal whether FLA2-Ty1 had been biotinylated, thereby demonstrating whether the protein was accessible on the surface of the cell.

In the cases of FLA2-Ty1 (1×Ty1) and FLA2ΔC-Ty1 (1×Ty1), the tagged protein appeared to be confined to small discrete loci within the cell, mostly posterior to the nucleus and anterior to the kinetoplast (Figs. 5.15, 5.18). It is probable that one of these discrete patches of signal represents the lysosome, which would indicate that the mislocalised protein is targeted for degradation. The other small patches may correspond to endosomal compartments or vesicles.

It is not clear why addition of a C-terminal tag to FLA2 causes the protein to mislocalise. In Chapter 3, C-terminal tagging did not affect normal localisation of the FLA3 protein, even upon addition of a 13.2-kDa 12×HA tag. The domain architectures of FLA2 and FLA3 are essentially the same, so there is no immediately obvious reason why the observed effects of C-terminal tagging were so different.

One possible reason why a C-terminal tag might interfere with entry of FLA2 to the FAZ is steric hindrance: the tag might be too big to allow the protein to access the narrow confines of the cell-body side of the flagellar attachment zone. This was a plausible explanation in the case of the bulky fluorescent-protein tags, mNeonGreen and mScarlet. However, it seems very unlikely to be relevant in the case of the small Ty1 tags, which were only 2–6 kDa (Table 5.2). Furthermore, deletion of the 16-a.a. intracellular tail of FLA2

(-4.2 kDa) and addition of a single Ty1 tag (+2 kDa) resulted in a net loss to the size of the C-terminal region, and yet the tagged protein still mislocalised.

Another possible reason for mislocalisation is that the tag could block a putative localisation signal in the intracellular tail of FLA2 that is required for trafficking of the protein to the flagellar pocket or to the FAZ. Alternatively, the tag could prevent a putative interaction between FLA2 and another FAZ protein (intracellular or transmembrane) that needs to occur before either protein can be targeted to the FAZ.

A final possibility is that tagging FLA2 caused unintentional overexpression of the protein. The pPOTv6 tagging construct changes the 3' UTR of the tagged gene to the 3' UTR of the *PFR2* paraflagellar-rod gene, which may increase the expression level of the encoded protein. According to this hypothesis, in the instances where tagged FLA2 was observed throughout the cell body, a certain proportion of the tagged protein may have localised correctly to the FAZ, while the surplus protein mislocalised. Such was the case with FLA1 when a YFP-tagged version of the FLA1 protein was deliberately overexpressed using a procyclin promoter (Sun et al., 2013). When fixed, non-permeabilised cells were viewed by direct fluorescence, mislocalised YFP signal was observed throughout the interior of the cell. In contrast, when the cells were probed with anti-GFP (which could not penetrate into the non-permeabilised cells), tagged FLA1 was observed along the FAZ only. Sun et al. were able to visualise YFP-FLA1 in this manner because the tag had been added to the extracellular domain of the protein, just after the putative signal sequence. In the tagged versions of FLA2 produced here, the tag must remain inside the cell, so surface probing for the tag is not possible.

However, given that tagging FLA3 at the C-terminus using the pPOTv6 tagging construct did not seem to cause overexpression or mislocalisation of the tagged protein, it seems most likely that the mislocalisation of FLA2 is in some way related to its position on the cell-body side of the FAZ. Environmental factors on the tightly packed cell-body region must be less accommodating to change than those on the flagellar side of the FAZ.

It seems, therefore, that the best strategy to localise FLA2 would be to add an epitope tag just after the putative signal-sequence region of the *FLA2* gene. This approach would be a form of internal tagging, which is more challenging than simply inserting a tag at the N- or C-terminus, because multiple cloning steps are required. Furthermore, the regions

immediately upstream of the respective start codons of *FLA1* and *FLA2* are identical, so it would be difficult to ensure specific tagging of FLA2 only. A possible method is outlined in Fig. 5.21. Firstly, the whole *FLA2* gene downstream of the putative signal cleavage site would have to be amplified using at least one primer that was FLA2-specific and would not amplify *FLA1*. This cloned fragment (~1.7 kb) would then have to be inserted into a tagging vector such as pPOTv7 or pPOTv6 (Dean et al., 2015), immediately downstream of the tag sequence. Long-primer PCR could then be performed. The forward primer would contain 80 bp homologous to the putative signal sequence of *FLA2*; the reverse primer would contain 80 bp homologous to the 3' UTR of *FLA2*. Utilising the 3' UTR in this manner should ensure specific targeting to the *FLA2* gene. In this way, a tagging amplicon would be produced that would insert into the endogenous *FLA2* locus, adding an epitope tag immediately after the N-terminal signal sequence.

It is essential for FLA2 to be tagged and localised before any studies can performed to determine the interaction partners of the protein.

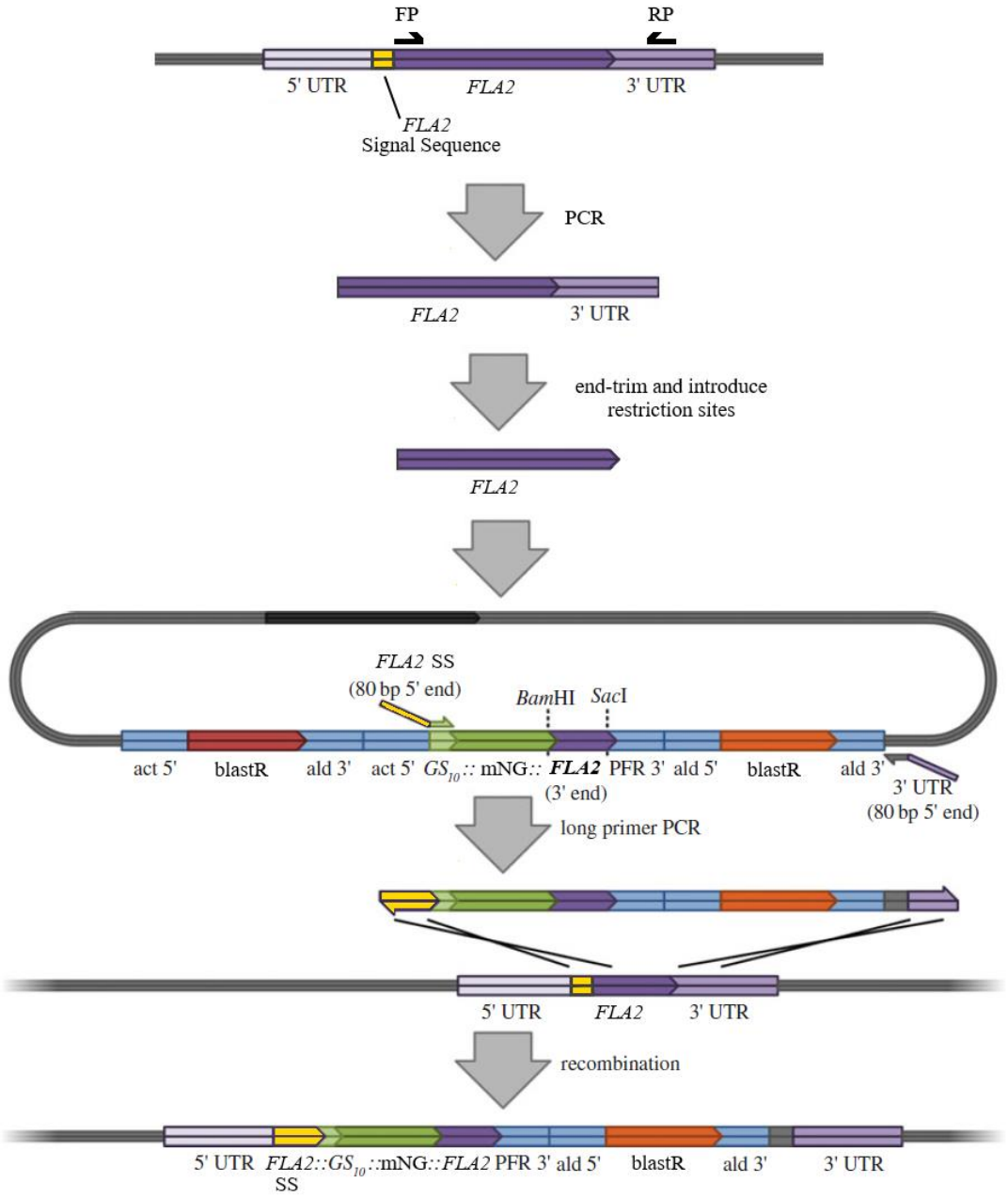


Fig. 5.21: A possible method for adding an epitope tag immediately after the signal sequence (SS) of *FLA2*.

The entire *FLA2* gene except the signal sequence is amplified from genomic DNA using a forward primer (FP) that recognises the region after the signal sequence, and a reverse primer (RP) that recognises the 3' UTR. The amplified fragment is end-trimmed to remove the UTR and to insert *Bam*HI and *Sac*I restriction sites. The fragment is inserted into the pPOTv7 tagging vector. Long-primer PCR with an FP homologous to the *FLA2* signal sequence and an RP homologous to the 3' UTR generates a tagging amplicon that inserts into the endogenous *FLA2* locus.

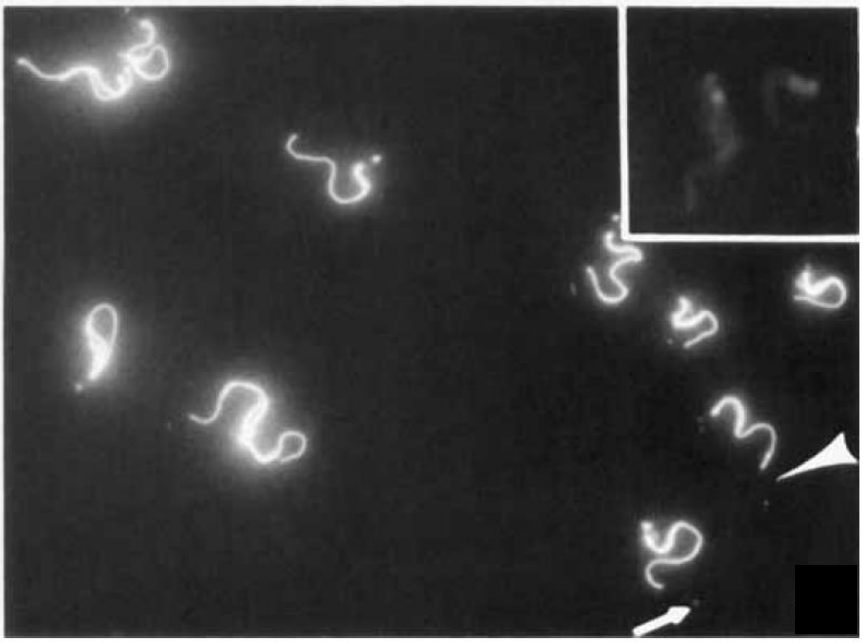


Fig. 5.22: RCA-I lectin specifically recognises the FAZ in fixed BSFs (reproduced from Brickman and Balber (1990)).

Non-permeabilised bloodstream-form *T. b. rhodesiense* were fixed with formaldehyde to denature the VSG surface coat, then subsequently probed with fluorescein-conjugated RCA-I lectin. The lectin bound specifically to the flagellar attachment zone (and to a single spot at the tip of the free flagellum [white arrows]).

The insert shows that the lectin does not bind the cells in the presence of excess galactose, confirming specificity of binding.

5.3.3 Glycosylation, Stage-Specificity and the FLA Glycoproteins

Thirty years ago, an investigation of lectin binding in *T. brucei rhodesiense* showed that there are galactosylated glycoproteins in what is now termed the flagellar attachment zone (Brickman and Balber, 1990). It is postulated here that these glycoproteins are FLA2 and FLA3.

Live bloodstream-form cells were probed with fluorescein-conjugated RCA-I lectin, which binds specifically to galactose (Brickman and Balber, 1990). The fluorescence signal was restricted to the anterior side of the flagellar pocket. The cells were subsequently fixed with formaldehyde, to disrupt/denature the VSG coat. Binding of RCA-I lectin was greatly increased upon fixation and, interestingly, the fluorescence signal was now observed as a strong, bright line alongside the attached flagellum (reproduced in Fig. 5.22). The signal did not extend to the free flagellum, indicating that the lectin-binding glycoprotein(s) were confined to the FAZ. Brickman and Balber further stated that the glycoproteins in question comprised a small group of Triton X-100 soluble glycoproteins carrying both ConA-binding and complex ricin-binding oligosaccharides, although they did not present these data.

Both FLA2-Ty1 and FLA3-Ty1 are shown in this thesis to bind RCA-I lectin (Figs. 5.11, 5.13 D, 4.4) and to be soluble in Triton X-100 (Figs. 5.10 B, 3.10 A). It seems probable, therefore, that FLA2 and FLA3 are among the proteins that were observed by Brickman and Balber (1990) along the intermembrane region of the FAZ.

(In Section 5.2.5, it was suggested that the comparative weakness of the apparent interaction between FLA2 and RCA-I raised the possibility that FLA2 was found in the lectin-bound fraction only because it co-precipitated with a lectin-binding protein such as FLA3. However, the apparent efficiency of the RCA-I pulldown of FLA2-Ty1 in Clone A5 (Fig. 5.13 D) makes it seem more likely that RCA-I can bind directly to FLA2, though not as strongly as to FLA3.)

The most interesting aspect of the observations made by Brickman and Balber (1990) was the fact that the lectin (which was ≥120 kDa) could not access the glycoproteins of the FAZ in live cells. This evidence demonstrates that when the VSG surface coat is in its native configuration, the extracellular intermembrane region of the FAZ is efficiently shielded

from large macromolecules in the extracellular medium – as expected, given that invariant surface proteins such as FLA2 and FLA3 would otherwise be targeted by host antibodies. It has been shown elsewhere that other invariant surface proteins are likewise inaccessible to lectins and antibodies (which are typically ~150 kDa) in live BSFs, but are accessible upon fixation (Balber and Frommel, 1988, Jackson et al., 1993).

It seems probable that FLA2 and FLA3 are situated in close proximity to the VSG dimers. It is possible that the extensive glycosylation of these two proteins acts as an insulating barrier, preventing excess attraction between the VSG and the FAZ and thereby ensuring that the lateral mobility of VSG in the plasma membrane is not impeded. A recent study has shown that the *N*-glycans on VSG have a buffering function of this sort (Hartel et al., 2016).

The FLA glycoproteins of the bloodstream form have a larger glycan component than their respective procyclic-form counterparts, although all four proteins appear to be extensively glycosylated (Table 5.4). As well as, potentially, forming a buffer against the nearby VSG, the glycans of the BSF-specific FLA glycoproteins may also help to shield the intermembrane region of the FAZ from the host adaptive immune response. As discussed in Section 4.3.4, there is a distinct possibility that the glycans of FLA2/3 may mimic those of the i/I antigens of erythrocytes and the ABO blood-group antigens, so that the host cannot raise antibodies against them.

The glycans of PCF-specific FLA1/1BP are exposed to the harsh environment of the tsetse-fly midgut. Rather than protecting from host immune factors, these glycans may protect FLA1/1BP from digestive proteases. *N*-glycosylation is known to confer resistance to proteolysis in many proteins (Niu et al., 2016). It seems very probable that the procyclic-form FLA glycoproteins have different sets of carbohydrates to their bloodstream-form counterparts, because the roles of the glycan side chains differ between life-cycle stages.

Table 5.4:

	Predicted Size	Predicted Size (without signal peptide)	Size on SDS-PAGE (approx.)	Estimated Glycan Component
FLA1	59 kDa	56 kDa	80 kDa	25 kDa
FLA2	64 kDa	61 kDa	100 kDa	40 kDa
FLA3	89 kDa	84 kDa	160 kDa	75 kDa
FLA1BP	83 kDa	78 kDa	80/120 kDa	40 kDa

The stage-specificity of the FLA proteins themselves probably stems from the need to integrate with the respective surface coats of the different life-cycle stages. It may be the case that all trypanosome surface proteins are stage specific for this reason. All invariant surface glycoproteins (ISGs) that have been investigated thus far are BSF-specific and they are all Type I transmembrane proteins with some predicted structural similarity to VSG (Quintana et al., 2018). The transferrin receptor (TfR) is likewise BSF-specific and has a VSG-like structure (Salmon et al., 1997). It is plausible that all BSF surface proteins must be able to fit alongside the VSG dimers in such a way that they do not protrude above the VSG layer or interfere with the lateral mobility of the VSG. Similarly, PCF surface proteins may need to be adapted to suit the needs of the procyclin coat proteins. This model would explain why the surface-exposed FLA glycoproteins are developmentally regulated while all known intracellular and intraflagellar FAZ proteins are expressed constitutively throughout the life cycle.

Even *in vitro*, it seems that FLA glycoproteins cannot function in the "wrong" life-cycle form. In Section 3.2.9, it was seen that when FLA3-Ty1 was induced in procyclic cells, the tagged protein mislocalised to the lysosome/endosomal compartments and appeared to be targeted for degradation. Interestingly, the TrypTag initiative, which aims to tag and localise all *T. brucei* proteins (Dean et al., 2017), has seen a similar result in the case of

FLA2. Dean et al. have recently tagged the *FLA2* gene (Tb927.8.4060) at the C-terminus with mNeonGreen (images available on TrypTag.org) and have investigated the localisation of the tagged protein in procyclic cells (all TrypTag work is performed in PCFs). They observed that the tagged protein is produced, but it mislocalises in a reticulate pattern throughout the ER and cell body, and on the nuclear membrane. It seems, therefore, that the BSF glycoproteins cannot enter the procyclic-form FAZ. It could be the case that in order to access the flagellar attachment zone, FLA2 and FLA3 must first recognise and bind to each other in the neck region of the flagellum, forming a bridge across the intermembrane space, before they can be assembled into the nascent FAZ. It would be interesting to see what would happen if FLA2 and FLA3 were induced simultaneously in procyclic cells.

5.3.4 Summary of Chapter 5

The main novel findings presented in this chapter are the discovery of the molecular weight of FLA2 after post-translational modification, the evidence that FLA2 is glycosylated with glycans that can bind RCA-I lectin, and the demonstration that FLA2 does not localise to the flagellar attachment zone in bloodstream-form cells when the protein is tagged at the C-terminus.

Chapter 6: Final Discussion

6.1 The Structure of the FAZ

The flagellar attachment zone of *T. brucei* is an enigmatic structure. The regularly spaced punctate structures of the FAZ have been viewed and investigated using many different techniques, but redundant terminology and contradictory reports have confused the description of the attachment region.

In recent years, the structure of the FAZ in transverse section has been well described (Sunter and Gull, 2016). However, there has been no detailed description or representation of the FAZ in longitudinal section in the published literature, for several reasons. Firstly, the relationship between the distinct punctate structures seen along the intracellular side of the FAZ in TEM images and the filamentous structures seen traversing the intermembrane region in SEM images was initially unclear. (The consensus now is that there is a 1:1 relationship between the intracellular punctae and the extracellular *macula adherens* – together, they form a junctional complex, see below.) Secondly, most known FAZ proteins locate to the intracellular region, usually termed the "FAZ filament" but it remains unclear whether these proteins have punctate or continuous distribution along the FAZ and how they are positioned in relation to one another. Thirdly, as Sunter et al. (2016) stated explicitly, there is known to be heterogeneity of structure along the length of the FAZ, but this heterogeneity has yet to be clearly defined.

Perhaps the most informative micrograph of the FAZ in longitudinal section is the one shown in Fig. 6.1 A. This TEM image, from Sherwin and Gull (1989), shows Zones 1–6 of the FAZ quite clearly (compare with Fig. 6.1 B). The insert in Fig. 6.1 A gives a schematic representation of a junctional complex.

It seems evident from Fig. 6.1 that the junctional complex (shown in yellow) reaches from the PFR to the diffuse intracellular structure that has been termed Zone 6 (shown in lilac). Each junctional complex includes one punctum of the FAZ filament (marked in Fig. 6.1 with a red X), which is the structure that is easily visible as a dark spot in many TEM images, e.g. Fig. 1.5 C. Each junctional complex also includes one *macula adherens*/staple (marked in Fig. 6.1 with a blue cross), an extracellular structure linking the flagellar and cell-body membranes (see Fig. 1.5 E).

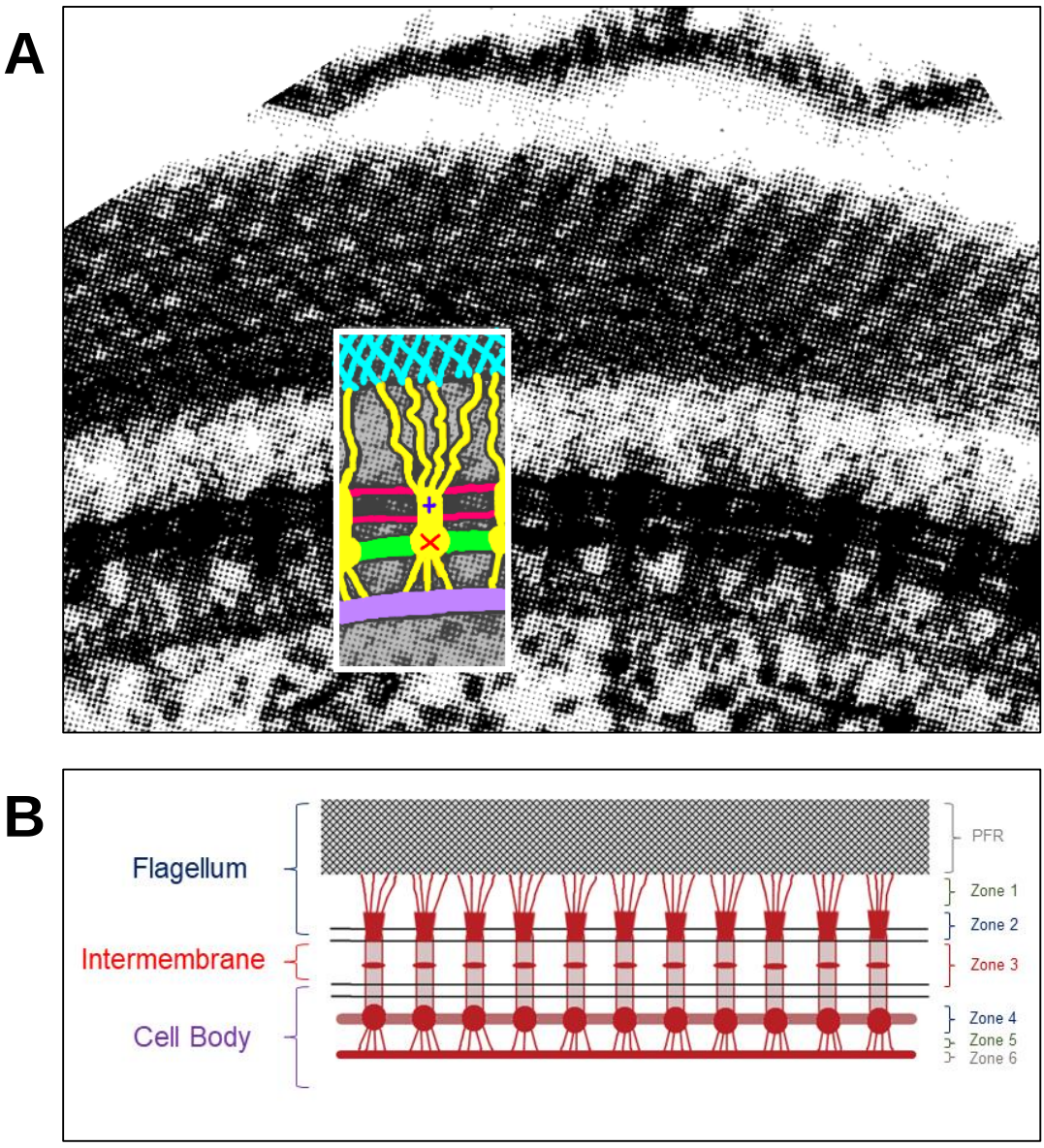


Figure 6.1: Longitudinal section of the FAZ (adapted from Sherwin and Gull (1989)

(A) Main image: transmission electron micrograph of the FAZ.

Insert: schematic overlay of a junctional complex (yellow), with the intermembrane staple region marked with a blue cross, and a punctum of the FAZ filament marked with a red X. The PFR is in light blue; the flagellar membrane and cell-body membrane are in pink; the Zone 4 filament is in green and the diffuse Zone 6 is in lilac.

(B) Schematic diagram of the FAZ in longitudinal section, reproduced from Fig. 1.7.

Critically, it is clear that most zones of the FAZ are discontinuous/punctate along the length of the attachment region. The only parts of the FAZ that seem to be continuous are (a) the filamentous structure connecting the punctae of Zone 4, (b) Zone 6 and (c) Zone 7, the microtubule quartet (not visible in Fig. 6.1). It is not known whether Zone 8, which connects Zones 4 and 7, is continuous or discontinuous.

Strictly speaking, it cannot be stated definitively whether the intermembrane region of the FAZ (Zone 3) is continuous or discontinuous. However, there is no evidence of any continuous membrane-membrane adhesion along the FAZ. It seems very likely that the junctional complexes are the only connection between the flagellum and the cell body, as implied by Sunter et al. (2016) when they described the junctional complexes as acting like desmosomes. The punctate localisation of FLA3-Ty1 in this thesis supports the hypothesis that all extracellular components of the FAZ locate to the junctional complexes. Further data are needed regarding the localisations of FAZ5, FAZ9, FS179 and, of course, FLA2.

It is possible that the discontinuity of the FAZ is necessary to create the immense flexibility that is required for the characteristic screw-like motion of the trypanosome cell.

Several studies have measured the apparent spacing (periodicity) of the punctate structures of the FAZ, and have reached widely differing values, as summarised in Table 6.1.

Table 6.1: Centre-to-centre distances between regularly spaced FAZ structures

In cases where distance values were not explicitly given in the text, distance values have been estimated based on the published images.

Source	Structures	Distance (nm)	Measured / Estimated	Life-Cycle Stage	Viewing Method
(Trépout, 2020)	FAZ "sticks"	**37.0 ±8.5**	measured	BSF	cryo-STEM
(Sherwin and Gull, 1989)	punctate connections	**~60**	estimated	PCF	TEM
(Höög et al., 2012)	"staples"	**71 ±20**	measured	PCF	cryo-ET
(Vickerman, 1969)	*macula adherens*	**95**	measured	BSF	TEM
(Höög et al., 2012)	*macula adherens*	**122 ±25**	measured	PCF	cryo-EM
(Moreira et al., 2017)	FAZ10	**~145**	estimated	PCF	immunogold-EM
(this thesis)	FLA3	**177 ±50**	measured	BSF	STED
(Sun et al., 2013)	FLA1BP	**~400–500**	estimated	PCF	confocal

An obvious reason for the disparity among the distance values in Table 6.1 is that these measurements were made using different fixation and visualisation conditions. Techniques such as high-pressure freezing and dehydration will affect the apparent distances between cellular structures. There is no reason to conclude that the FAZ "sticks", *macula adherens*, "staples" and punctate connections are different structures; rather, they are all parts of the junctional complexes portrayed in Fig. 6.1. Höög et al. (2012) supposed that the "staples"

and *macula adherens* were separate structures. However, the difference in the apparent distances between adjacent "staples" and between adjacent *macula adherens* was merely an artefact due to the use of different visualisation techniques. Subsequent publications by the same authors make it clear that the term "staple" refers to Zone 3 of the *macula adherens*/junctional complex (Sunter and Gull, 2016).

The distance between FLA3-Ty1 punctae measured in this work (Table 6.1) is consistent with the hypothesis that FLA3 is a component of the junctional complexes. The relatively large apparent distance between FLA1BP punctae (Sun et al., 2013) is probably not representative of the true spacing of FLA1BP; the YFP-FLA1BP signal was relatively patchy and would benefit from super-resolution imaging.

It would be reasonable to suppose that there might be a difference in the distances between adjacent junctional complexes in BSFs and PCFs, given that there are perceptible differences in the motility of the two life-cycle stages (BSF motion is distinctly more rapid and more convoluted than PCF motion). However, no such difference is apparent in Table 6.1. Most studies of the FAZ are performed in either PCFs or BSFs. Until an investigation is performed in which distance measurements are made using the same technique in both life-cycle stages, it is impossible to state whether such a difference exists.

It is also possible that the distance between the flagellar membrane and the cell-body membrane differs between life-cycle stages. Vickerman (1969) stated that the intermembrane gap is 35–40 nm in BSFs and ~20 nm in PCFs. Höög et al. (2012) found the intermembrane gap to be 26 ±2 nm in plunge-frozen PCFs. Different fixation/freezing conditions can have a significant impact on apparent distances. Again, until a study is performed using the same visualisation technique in both life-cycle stages, it is difficult to make reliable comparisons.

It is unclear whether the surface coat proteins can pass between the junctional complexes. It seems probable that they can. Vickerman (1969) states that, since VSG is 12–15 nm in height, there is a space of at least 8 nm between the flagellar VSG and the cell-body VSG in the intermembrane region of the FAZ, which would suggest that lateral movement of VSG in both membranes is probably not impeded by the apposition of the membranes. Recent work has shown that the height of the VSG dimer varies because of the flexible configuration of the protein (Bartossek et al., 2017). The MITat 1.1 VSG isoform, for

example, fluctuates from 14–16 nm in height. This flexibility would presumably enhance the freedom of movement of VSG in the FAZ region. Molecular modelling has suggested that the procyclin in the PCF surface coat is 14–18 nm in height (Roditi et al., 1989), so it may not be able to pass easily between the junctional complexes in the limited confines of the PCF FAZ.

It is tempting to suppose that the intermembrane gap in the BSF FAZ is wider than its PCF counterpart specifically to allow free passage of VSG molecules. As discussed in Chapter 5, the BSF-specific FLA proteins are larger than their PCF counterparts by 20–40 kDa. It is possible that this size distances reflects a need for longer junctional complexes in the bloodstream-form cell. However, it remains uncertain whether the intermembrane distance in live cells is really as wide as it appears to be in cells fixed for electron microscopy.

As the FAZ is assembled, the regular spacing of the junctional complexes must be maintained. It seems from electron micrographs of the FAZ, such as Fig. 6.1, that the FAZ is more elaborate and substantial on the cell-body side than on the flagellar side. The junctional complexes seem to be firmly embedded in the Zone 4 filament that runs the whole length of the attachment zone, and they are further anchored by Zones 5 and 6. It seems that during FAZ assembly, the Zone 4 filament must hold the junctional complexes in position relative to one another while the whole nascent FAZ structure proceeds in the anterior direction. As proposed by Sunter et al. (2016), the connection between the flagellar side of the FAZ (Zone 1/2) and the flagellum cytoskeleton (PFR) must be weak and transient during FAZ outgrowth, to allow the flagellum and FAZ to slide against one another. Indeed, it seems that Zone 1/2 may remain the weakest part of the mature FAZ. In this thesis, it is shown that FLA3 locates to the flagellar membrane, with its extracellular domain in Zone 3 and its C-terminal tail in Zone 2. Nonetheless, when Woods et al. (2013) viewed FLA3 in cells in which the flagellum was (unintentionally) mechanically detached during fixation, the protein was seen only on the cell-body side of the FAZ. It seems that when excessive force is applied to the BSF FAZ, the connections in Zone 1 or 2 are more likely to break than the connections in Zone 3. It is likely that while the junctional complexes themselves act as molecular rivets or bolts, certain components of Zone 1/2 act

as a nut, securing each "bolt" in position so that it can withstand the force of flagellar beating.

Like FLA3, FLA1BP has its extracellular domain in Zone 3 and its C-terminal tail in Zone 2. The extracellular domain of FLA1 is also in Zone 3, but its C-terminal tail is attached to Zone 4. Although FLA2 has not yet been localised, it seems overwhelmingly likely to have the same localisation and orientation as its homologue, FLA1. As discussed above, it is postulated here that all the membrane proteins of the FAZ are associated with the discrete junctional complexes, and that there is no continuous membrane-membrane join along the adhesion zone. Therefore, it is hypothesised here that when FLA2 is localised, it will be found to have a punctate distribution.

To enable better comprehension of the FAZ structure, more localisation studies of the known FAZ proteins are required. It will be possible to build up a much more detailed picture of the region if experiments are performed in which two or three FAZ proteins are tagged and imaged simultaneously, to allow comparison of their respective localisations. Immunogold electron microscopy may be useful in this regard, but immunogold images are sometimes quite patchy and difficult to interpret. The work performed on FLA3-Ty1 in this thesis has shown that STED microscopy is another potential way to image FAZ proteins in high resolution. With the advent of multicolour STED (Winter et al., 2017), it should be possible to compare the localisations of several tagged FAZ proteins at once.

If there are proteins other than the FLA glycoproteins in the intermembrane region of the junctional complexes, they are probably membrane proteins too. The most likely candidates, based on current information, are FAZ5 and FS179. Super-resolution imaging of these two proteins might be very informative.

Höög et al. (2012) observed an electron-dense midplate in the intermembrane region of the junctional complex in PCF cells. It is not yet known what this midplate signifies, but it appears to occupy the expected position of the interface between FLA1 and FLA1BP.

Trépout (2020) has stated that the centre-to-centre distance between adjacent BSF junctional complexes is reduced towards the distal end of the FAZ (32.4 nm ±9.3) compared to that towards the proximal end (44.0 nm ±9.0). This change in spacing has not been observed in other studies and was not observed here in the case of FLA3-Ty1.

6.2 Conclusions and Future Work

Few investigations of the bloodstream-form flagellar attachment zone have been performed, even though the BSF is the clinically relevant form of the parasite. This work has sought to remedy this deficit by investigating the membrane glycoproteins of the bloodstream-form FAZ.

Here, it is shown that FLA3 localises to the flagellar membrane of the FAZ, rather than to the cell-body membrane as was initially thought. It is demonstrated that the FLA3 protein is surface-exposed and has a regularly spaced punctate distribution suggestive of localisation to the intermembrane region of the junctional complexes. It is shown that FLA3 can be tagged at the C-terminus and subsequently used for immunoprecipitation experiments. It is demonstrated that FLA3 can bind to RCA-I lectin and tomato lectin, which reveals information about the glycan side chains of the protein.

For the first time, it is demonstrated that FLA2 is essential for flagellar attachment and cytokinesis in the bloodstream form. The argument is presented that FLA2 is the BSF paralogue of FLA1, and the FLA1 is probably not relevant in the bloodstream form, contrary to the initial findings. It is also shown that FLA2 can bind to RCA-I lectin.

Collectively, the data presented here strongly support the hypothesis that FLA2 and FLA3 perform equivalent roles in the BSF to their respective homologues, FLA1 and FLA1BP, in the PCF. To confirm this hypothesis, it will be necessary to show interaction between FLA2 and FLA3. As discussed in Section 5.3.2, FLA2 must be epitope tagged at the N-terminus. When tagged versions of FLA2 and FLA3 are expressed and localised in the same cell line, it will be possible to perform co-immunoprecipitation experiments to investigate the putative interaction between the two proteins.

Furthermore, tagging FLA2 at the N-terminus should facilitate super-resolution localisation by STED, a technique which has been shown here to be an informative way to image FAZ proteins.

Proximity-dependent biotinylation (BioID) (Morriswood et al., 2013) could be a useful method to investigate the composition of the intermembrane region of the junctional complexes.

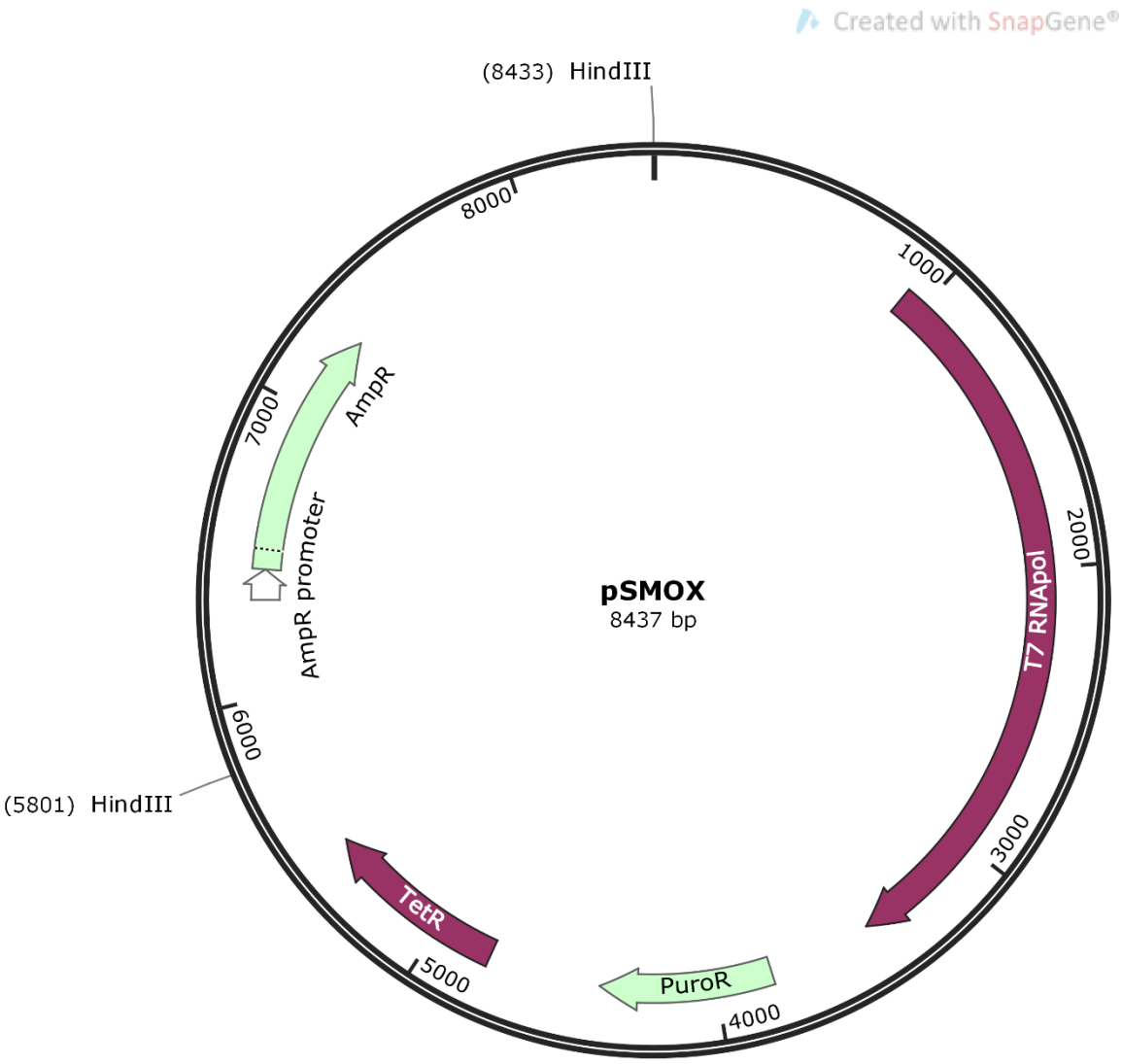
Generation of a FLA1-specific RNAi vector would resolve once and for all the question of whether FLA1 has any role in the bloodstream form. However, owing to the high sequence similarity between FLA1 and FLA2, it remains doubtful whether it is possible to generate an RNAi construct that will target FLA1 only.

Since it has been shown here that FLA2 and FLA3 are both galactosylated, it would be interesting to see whether inhibition of the galactosylation pathway would have any effect on flagellar attachment. It is not known whether the glycans of FLA2 and FLA3 are involved directly with attachment.

Finally, it is advised that future investigations into the flagellar attachment zone of the trypanosome should use only the term "junctional complex" to refer to the molecular rivets that connect the flagellum to the cell body. The different parts of the junctional complexes can be specified using the "zone" terminology established by Sunter et al. (2016). The other names that have been given to the junctional complexes in the past are ambiguous, which hampers the description of this fascinating structure.

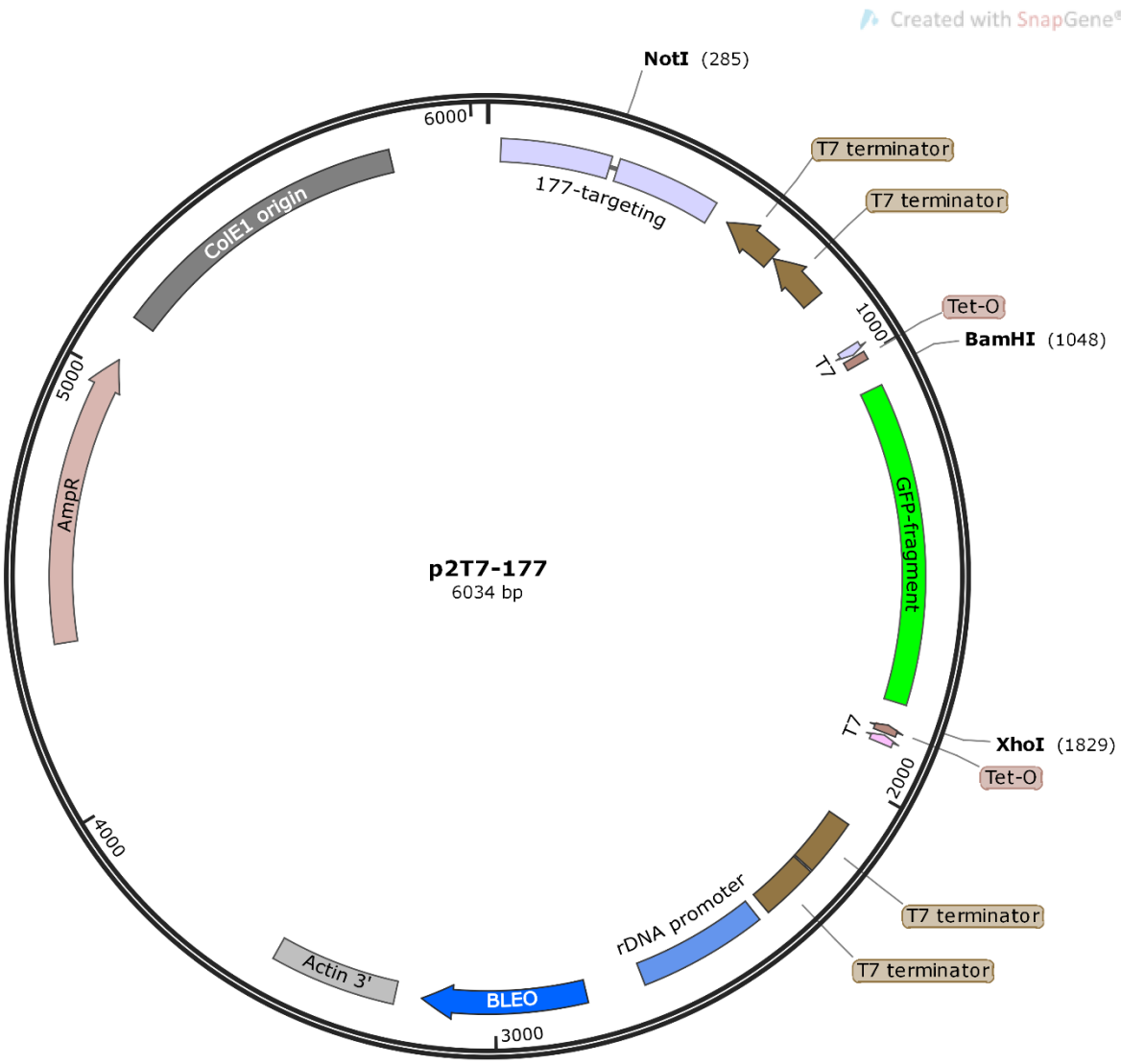
Appendices

Appendix A: Plasmids



The **pSMOX vector** (Poon et al., 2012) was used to generate trypanosome cell lines expressing T7 RNApol and TetR (see Section 3.3.3).

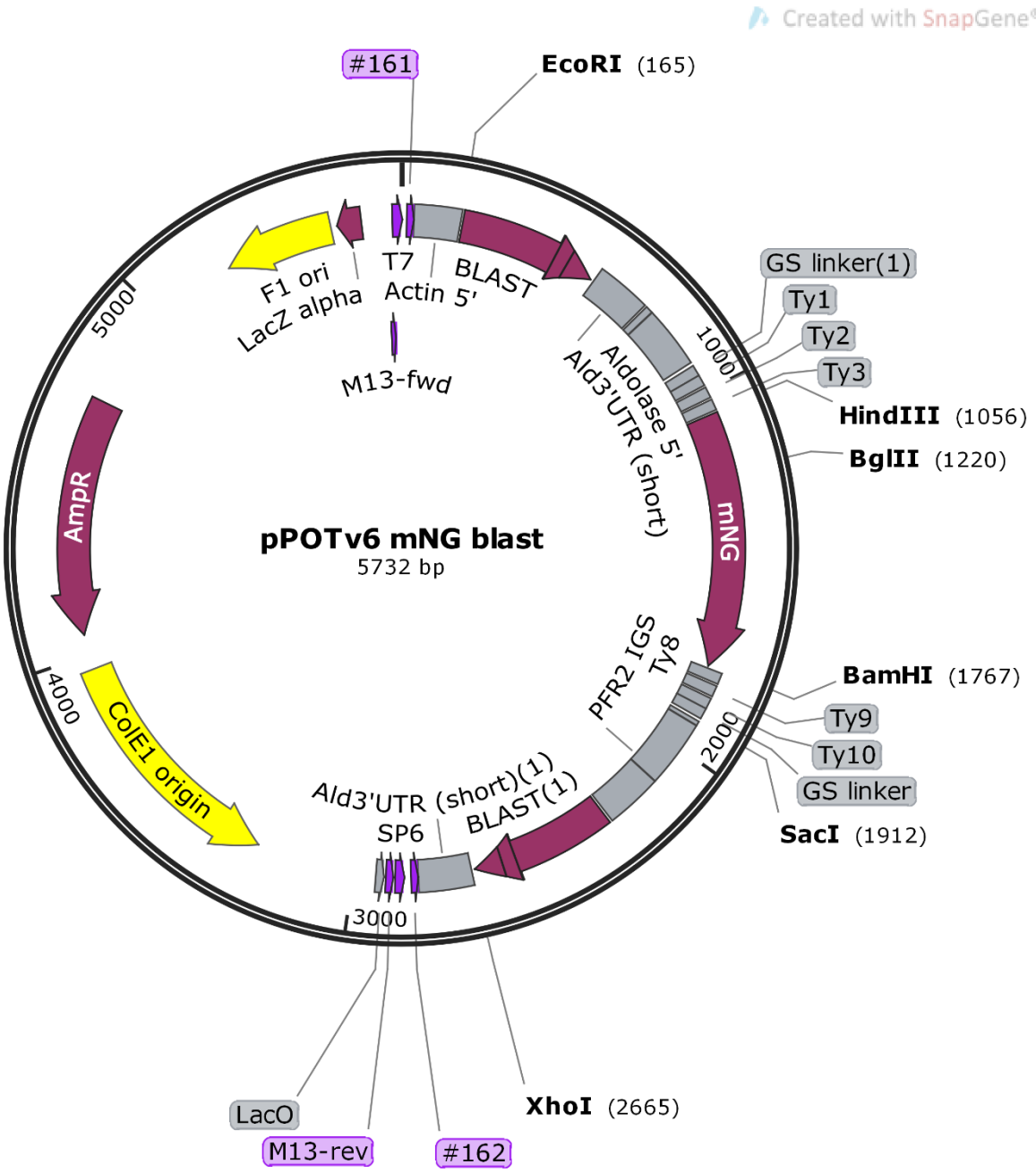
Linearisation was achieved by *Hin*dIII digest. The vector inserted into the tubulin locus.



The **p2T7.177 RNAi vector** (Wickstead et al., 2002) was used to generate FLA2/3 knockdown constructs (see Section 2.2.6).

The fragments of the target genes (*FLA2/3*) were cloned into the vector using *Bam* HI and *Xho*I, replacing the stuffer region of the vector.

Linearisation was achieved by *Not* I digest. The vector inserted into the 177-bp repeats of the minichromosomes.



The **pPOTv6 tagging vector** (Dean et al., 2015) was used for the long-primer PCR tagging experiments described in Section 2.2.8. The pPOTv7 tagging vector is identical except that it lacks the three Ty1 residues ("Ty") on either side of the fluorescent protein sequence ("mNG").

Appendix B: Primers

Regions in red are identical to regions of the *FLA2* or *FLA3* CDS or UTR, as described in detail in Chapter 2. Regions in black are identical to the primer-binding regions of the pPOTv6/pPOTv7 tagging vectors. The total number of nucleotides in each primer is given in brackets.

FLA2.pPOT.FP1

TTGCCATCATGCTGCTGTGGTTGCTCATAGCAAGTCCACCGAATGCAAGATCTACTTTTGTGCGACTGCCGGCGCTGGAGGGTTCTGGTAGTGGTTCC **(98)**

FLA2.pPOT.RP1

GGTTGCCTCTCACTATTCCCCACCGGCTGTCACCTCCATTACGTATCCATGAGTGACTCCGAAGTCACTTGCACGTCCAACCAATTTGAGAGACCTGTGC **(100)**

FLA2.pPOT.FP2

TTGCCATCATGCTGCTGTGGTTGCTCATAGCAAGTCCACCGAATGCAAGATCTACTTTTGTGCGACTGCCGGCGCTGGAGGGGATGGACGAATTGTATAAG **(101)**

FLA2.pPOT.FP3

TTGCCATCATGCTGCTGTGGTTGCTCATAGCAAGTCCACCGAATGCAAGATCTACTTTTGTGCGACTGCCGGCGCTGGAGGGCTCAGAGGTTCATACGAACC **(102)**

FLA2.pPOT.FP4

TTGCCATCATGCTGCTGTGGTTGCTCATAGCAAGTCCACCGAATGCAAGATCTACTTTTGTGCGACTGCCGGCGCTGGAGGGCTCAGAGGTTCATACGAACCAAGATCCTCTTGATGGATAA **(122)**

FLA2.pPOT.FP5

AGGACCTGCGTGCTGCAGGAATCGCCTCCTCAGTGCTCTTGGGCCTTCTTGCCATCATGCTGCTGTGGTTGCTCATAGCAGGGATGGACGAATTGTATAAG **(101)**

FLA2.pPOT.FP6

AGGACCTGCGTGCTGCAGGAATCGCCTCCTCAGTGCTCTTGGGCCTTCTTGCCATCATGCTGCTGTGGTTGCTCATAGCAGGCTCAGAGGTTCATACGAACCAAGATCCTCTTGATGGATAA **(122)**

FLA3.pPOT.FP1

CAACAGTGGAAGATGATGAAGAAGATCGTGTATCAAACATCGGTGTGCCACTGACCGATGGGAAGGGAACCACCGCACCGGGGATGGACGAATTGTATAAG **(101)**

FLA3.pPOT.RP1

CGATTATATTCTATGCATGCAAACACACCACACAAACACAAGGAATGTGATGTTACACATACAGCTACAACGCACGTGACCCAATTTGAGAGACCTGTGC **(100)**

Appendix C: Pairwise alignment of the *FLA2* and *FLA1* genes.

Upper row: *FLA2* (Tb927.8.4110)

Lower row: *FLA1* (Tb927.8.4010)

Pairwise alignment was performed using the EMBOSS Needle tool.

```
  1 ATGGGTGGCAGGACTGAATCACGGGAAGCATTGGCCGCACTCGTGGCAGC     50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
  1 ATGGGTGGCAGGACTGAATCACGGGAAGCATTGGCCGCACTCGTGGCAGC     50

 51 GTTGTTGCTGCTGCTGGGTTTTGTGAGTCCCGTTATAGGAGACCAAACCG    100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 51 GTTGTTGCTGCTGCTGGGTTTTGTGAGTCCCGTTATAGGAGACCAAACCG    100

101 TGGGCACCAAGGTGATTGTGAACCTCTTCAACAGTTGTATCAACTGTAGT    150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 TGGGCACCAAGGTGATTGTGAACCTCTTCAACAGTTGTATCAACTGTAGT    150

151 AGTGGGCAGGCGGATGGCATAAACGGTACTGGTCGCTTATTCAAGGCGGT    200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 AGTGGGCAGGCGGATGGCATAAACGGTACTGGTCGCTTATTCAAGGCGGT    200

201 GGGTGGGTTCTTTAAGAACAAAAGCTTTCCTTTACTTTTATGTGATGTTG    250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 GGGTGGGTTCTTTAAGAACAAAAGCTTTCCTTTACTTTTATGTGATGTTG    250

251 GAGGTGGTGGATCCACCGTGCGACTCGTAAACAAAACTGGCATATATACA    300
    ||||||||||||||||||||||||||||||||||||||||||||||||||.
251 GAGGTGGTGGATCCACCGTGCGACTCGTAAACAAAACTGGCATATATACC    300

301 ATTGCAGGCAGTCTGACGCAA--CGTGGTAATAAAGACGGCCCCAAG--G    346
    .||||.||.|.|||  .||||  ||||||      ||||     |||  |
301 GTTGCGGGAAATCT--TGCAAAGCGTGGT------GACG-----AAGTTG    337

347 G------AGA---TGCACTATTCAACAATCCCACCTCAGTTGTAAGCGTC    387
    |      |||   |||.|..||.||..||||||||||.||||..|.||.
338 GTCCTTTAGAGGTTGCTCGTTTTAATCATCCCACCTCTGTTGTTGGTGTA    387

388 AACGATGATATTTACGTGGCAGACAGGGATAACAACTGTATTAGGCGAAT    437
    |||.||||||||||||||||.||||||||||..|.||.||.|.|||.||
388 AACAATGATATTTACGTGGCGGACAGGGATAATGATTGCATGAAGCGGAT    437

438 AGATGCTGAGGGCAAT-GTAACCAGATATG-GACCACAAGACCTTAACAA    485
    |||||||||.|| ||| ||.||.|..|.|| |.||| |.||..||.|.||
438 AGATGCTGATGG-AATGGTGACGAAGTTTGCGGCCA-ATGAAGTTGATAA    485

486 ACCGA--AGGATATTCTCCCATTCA-CATTGA-----ACGGGACTCAGAA    527
    |||||  ||  ||||     |||.| || |||     |||||.|.|.|..
486 ACCGAGTAG--TATT-----ATTTATCA-TGAACAAGACGGGGCACCGGT    527

528 TCTATTCATTTCCGACACAGGCAATTCGCAGATCCTTTATGTGCCACTGG    577
    |.||||.|||||.||.||.||.||||||...||    ||||||   |||.|
528 TTTATTTATTTCGGATACTGGAAATTCAAGGA----TTATGT---ACTCG    570

578 ATGCTAATAACACAGTG-----------GTAA---CTACACTAGTCGCTG    613
    ..|.||       ||||           ||||   |.|.|||.||.|..|
571 CAGATA-------AGTGTGTCCAACAATGTAACGGCGAAACTTGTTGGAG    613

614 GCTTCCAACCTGGTGTGATGCAAATAAGCAAAGAA-AAGAACATGATGTA    662
    |.||.|||||||||||||||||||||| ||||| ..||||.|.|||||
614 GTTTTCAACCTGGTGTGATGCAAATAAGC-AAGAAGCGGAATTTCATGTA    662
```

```
 663 TGTCGTAAAAAACACCTCATGGATTGCAGCGGTTAATTTG---AGCAAAA    709
     ||||||.||..|.|||||.|||||||||||||..|..||   |||
 663 TGTCGTGAAGGAGACCTCGTGGATTGCAGCGGTGGACCTGAATAGC----    708

 710 CT-GGGGAGTCTGATATTGG-AAAAAGTAA------GGATATAGGTGATG    751
     || .|.||.||.||   ||| ||.||| ||      |||.||.|...||.
 709 CTCAGCGATTCGGA---TGGTAACAAG-AAGAGCTGGGACATCGCCAATA    754

 752 TGACGTGTCTGCACTA---CAAAAGTGCTCTGATGCTGACTCAAGCAGAG    798
     |||..||||..|.||   |.|   |||..|||||.|.|||.|.|..|||
 755 TGAGTTGTCTCGATTATGTCGA---TGCCTTGATGTTAACTGAGGACGAG    801

 799 GATAAACTATATTATTACGGTGAGCCCAATGGAGAGGGGTACATCATGTC    848
     .||.|.||.|||||||||||||.||.||.||.|||.|..|.||||||||
 802 AATGAGCTCTATTATTACGGTGAACCTAACGGCGAGAGCCATATCATGTC    851

 849 ACTAGAGGTGGAACGCCCATCT----CCAGCATCCCAATGGTGTGCGC--    892
     .||.|| .|.|||..||  |||    ||.|  ||    ||.||||||.|
 852 CCTTGA-ATTGAATTCC--TCTGAAACCGG--TC----TGTTGTGCCCTA    892

 893 AGTCAA--TGTTGAAATGGGACTATG-AT---CGGATTGTCTCATTGTTG    936
     ||  ||  ||.||.|.|||    ||| ||   |.||||||||||.||.|.
 893 AG--AAGGTGATGGAGTGG----ATGCATGGGCCGATTGTCTCACTGGTT    936

 937 AGGGTGAATGACAAAGAGTATTATGCTACAACAGAAACGGCCGTGTATGC    986
     |.||||||||.....||.||||||||.|.|||.|||||.||.||||||.||
 937 AAGGTGAATGGATGTGAATATTATGCCATAACCGAAACAGCTGTGTACGC    986

 987 CATTCGTGATGTAGGCTATACACCGACGCCAACGCCTCCACCAACGCCCC   1036
     |.|.||||||...|.|||||.||||.|         |.||||.|.|||
 987 CGTCCGTGAGAAATGCTATTCACCAA---------CACCACTACCGC---   1024

1037 CTCCCACGCCTTCGGTGCCGCCATCACCTTCTCCCAGCCCATCAGCATCG   1086
                |||||            .||||||
1025 ------------CGGTG------------ACTCCCA--------------   1036

1087 CCGTCACCATCACCATCAGTGCAGCCGCCACCGCCACCGCCACCACCACC   1136
                              |||||.|||
1037 --------------------------CGCCAACGC----------------   1045

1137 ACCGCCGCCGCCGCCGCCGCCGATTACTCCCAATCCGGATCCCCCTTCAC   1186
                                        ||||       .||||
1046 -----------------------------------CCGG-------GTCAC   1054

1187 CTCCGAGACCCACAGCTGTTGCCGCCTTTCGCGCTTCTTCGTTCCCAATT   1236
     ||||||||.||||||||||.||||||||||||||||||.||||||||||.
1055 CTCCGAGATCCACAGCTGTGGCCGCCTTTCGCGCTTCCTCGTTCCCAATG   1104

1237 G---ATGAT--GTAAAACTTATGCATGCTCTGTATTCGTGGATTATGAGG   1281
     |   |||||  |     ||.||||||.||.||||||||.||||||||.|..
1105 GGAAATGATCGG-----CTCATGCACGCCCTGTATTCATGGATTATAAAA   1149

1282 GATGTTGAGATCG-CTTTCGGCACAGATGACTTTTTTGCTGTTTTCCTCC   1330
     |||||.|||.|.| ||||| |.||||||.|||||.|.|||||||.||||||.|
1150 GATGTGGAGGTAGCCTTTC-GTACAGAGGACTTCTATGCTGTGTTCCTTC   1198

1331 CCCCAGGTTACAATGACGTGCCTGGTGGTACAAATGTTTCGACGTGGACA   1380
     |.||.|||.|.|.|.||||||||||||||||||||||||||||||||||
1199 CACCCGGTGAGATTAACGTGCCTGGTGGTACAAATGTTTCGACGTGGACA   1248

1381 AACTTGACAGCTCTAGA-GAACTTCGACAAGAACATCCTTATTACCAACT   1429
     |||.|.||.||..| || ||||.|.||.||.|.|||||||||||||||||
1249 AACCTTACTGCAAT-GATGAACCTTGAGAATACCATCCTTATTACCAACT   1297

1430 ACCAAGGG--CCTCCTGGTAGGAGCAAGCAAGAAC---TACAGAAGCGGT   1474
     ||  ||||  ||.||.||.|.|||||||   |.||   ..|||||||||.
1298 AC--AGGGGCCCCCCCGGCATGAGCAAG---GGACGTGCGCAGAAGCGGC   1342
```

```
1475 TGTTCTCGTCTCCATGGAACTGGACACGGAAGTTTCTTGATGAATTGGTA   1524
     ||||||||||||||||||..|||||.||||.|||.|||||.|||.||...
1343 TGTTCTCGTCTCCATGGACGTGGACCCGGATGTTCCTTGAAGAAGTGTCG   1392

1525 CGACAAGAGGAGTGGGATCACCTATCACCATTCTGCATGGTGAAGTGTGT   1574
     |..||||||.|||||..||||||.|..||..|||||||||||||||||||
1393 CAGCAAGAGCAGTGGAGTCACCTGTTGCCGCTCTGCATGGTGAAGTGTGT   1442

1575 GGATGATTGTCAGTACATTTCCTTCAATGAATCTCTGTGTGTAGACGAGC   1624
     |||||||||.|||.|||||.||||...|||.||..|.||.|.|||.||
1443 GGATGATTGCCAGCACATTACCTTAGCTGAGTCCGTATGCGGAGATGA--   1490

1625 TTCA-GT---TCCAGGTGCATGAGGAC-CTGCGTG-CT-GCAGGAATCG-   1666
     .||| ||   ||.||.|||     ||| .||.||| || |..||.||||
1491 CTCACGTCTATCTAGATGC-----GACGGTGTGTGTCTCGGTGGTATCGT   1535

1667 --CCTCCTCAGTGCTCTTGG-GC-------------CTTCTTGCCATCAT   1700
       |||||.||.|  |||||| ||             |||||||..|.|.||
1536 TTCCTCCGCACT--TCTTGGTGCGGTAGCTATCGTACTTCTTTTCCTGAT   1583

1701 GCTGCTGTGGTTGCTCATAGCAAGTCCACCGAATGCAAGATCTACTTTTG   1750
     |     |||||      .|||..||.||.||.|   |||.|| .|.||||
1584 G-----GTGGT------AAGCCCGTACAACGTA---AAGTTC-GCCTTTG   1618

1751 TGCGACTGCCGGCGCTGGAGTAG-   1773
     |.||..||||||||.|.|||| |
1619 TACGGTTGCCGGCGTTCGAGT-GA   1641
```

Length: 1874
Identity: 1276/1874 (**68.1%**)
Similarity: 1276/1874 (**68.1%**)
Gaps: 334/1874 (**17.8%**)
Score: 4426.0

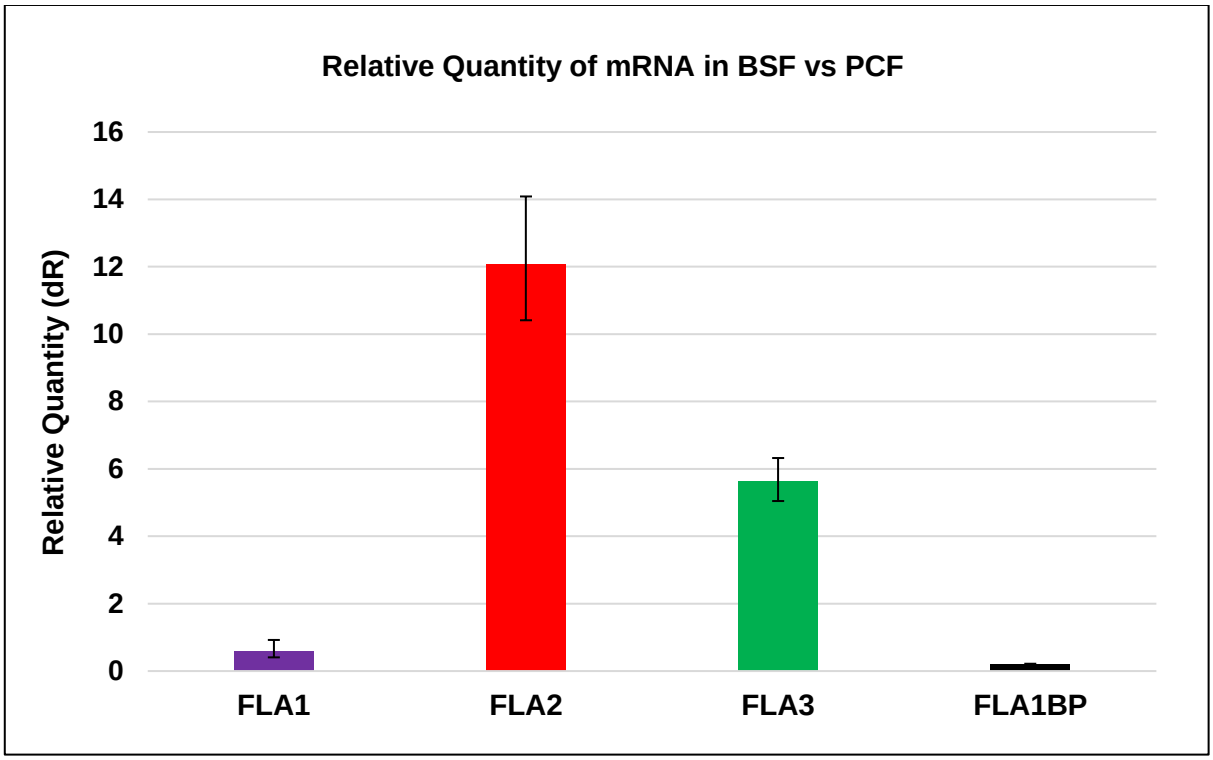
Appendix D: Sequence of the FLA2 RNAi construct

The p2T7-177/FLA2 construct was sequenced by Eurofins MWG Operon using the FLA2 forward and reverse primer (see Table 2.1).

Combining the two sequences gives a sequence with only two nucleotides different (red) compared to the original FLA2 gene sequence.

```
1      AGACGGCCCCAAGGGAGATGCACTATTCAACAATCCCACCTCAGTTGTAAGCGTCAACGA 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
333    AGACGGCCCCAAGGGAGATGCACTATTCAACAATCCCACCTCAGTTGTAAGCGTCAACGA 392

61     TGATATTTACGTGGCAGACAGGGATAACAACTGTATTAGGCGAATAGATGCTGAGGGCAA 120
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
393    TGATATTTACGTGGCAGACAGGGATAACAACTGTATTAGGCGAATAGATGCTGAGGGCAA 452

121    TGTAACCAGATATGGACCACAAGACCTTAACAAACCGAAGGATATTCTCCCATTCACATT 180
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
453    TGTAACCAGATATGGACCACAAGACCTTAACAAACCGAAGGATATTCTCCCATTCACATT 512

181    GAACGGGACTCAGAATCTATTCATTTCCGACACAGGCAATTCGCAGATCCTTTATGTGCC 240
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
513    GAACGGGACTCAGAATCTATTCATTTCCGACACAGGCAATTCGCAGATCCTTTATGTGCC 572

241    ACTGGGTGCTAATAACACAGTGGTAACTACACTAGTCGCTGGTTTCCAACCTGGTGTGAT 300
       |||||  |||||||||||||||||||||||||||||||||  |||||||||||||||||
573    ACTGGATGCTAATAACACAGTGGTAACTACACTAGTCGCTGGCTTCCAACCTGGTGTGAT 632

301    GCAAATAAGCAAAGAAAAGAACATGATGTATGTCGTAAAAAACACCTCATGGATTGCAGC 360
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
633    GCAAATAAGCAAAGAAAAGAACATGATGTATGTCGTAAAAAACACCTCATGGATTGCAGC 692

361    GGTTAATTTGAGCAAAACTGGGGAGTCTGATATTGGAAAAAGTAAGG 407
       |||||||||||||||||||||||||||||||||||||||||||||
693    GGTTAATTTGAGCAAAACTGGGGAGTCTGATATTGGAAAAAGTAAGG 739
```

Appendix E: qRT-PCR data showing the expression levels of *FLA* mRNA in wildtype bloodstream-form cells relative to wildtype procyclic-form cells.



Relative levels of *FLA1, FLA2, FLA3* and *FLA1BP* mRNA were measured by qRT-PCR in wildtype BSF and PCF cells.

The levels of each transcript were normalised against actin mRNA. The relative quantification was calculated by the ΔΔCt method and expressed as a ratio of induced to noninduced or of bloodstream to procyclic form, as described previously (Woods et al., 2013).

The measured BSF:PCF ratios were as follows:

FLA1	0.58
FLA2	12.08
FLA3	5.63
FLA1BP	0.20

qRT-PCR experiments were performed in collaboration with Derek Nolan.

References

AKIYOSHI, B. & GULL, K. 2013. Evolutionary cell biology of chromosome segregation: insights from trypanosomes. *Open Biology,* 3**,** 130023.

ALSFORD, S., TURNER, D. J., OBADO, S. O., SANCHEZ-FLORES, A., GLOVER, L., BERRIMAN, M., HERTZ-FOWLER, C. & HORN, D. 2011. High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Research,* 21**,** 915-924.

AN, T. & LI, Z. 2018. An orphan kinesin controls trypanosome morphology transitions by targeting FLAM3 to the flagellum. *PLoS Pathog,* 14**,** e1007101.

ATRIH, A., RICHARDSON, J. M., PRESCOTT, A. R. & FERGUSON, M. A. J. 2005. Trypanosoma brucei Glycoproteins Contain Novel Giant Poly-N-acetyllactosamine Carbohydrate Chains. *Journal of Biological Chemistry,* 280**,** 865-871.

AUTY, H., TORR, S. J., MICHOEL, T., JAYARAMAN, S. & MORRISON, L. J. 2015. Cattle trypanosomosis: the diversity of trypanosomes and implications for disease epidemiology and control. *Rev Sci Tech,* 34**,** 587-98.

BALBER, A. E. & FROMMEL, T. O. 1988. Trypanosoma brucei gambiense and T. b. rhodesiense: concanavalin A binding to the membrane and flagellar pocket of bloodstream and procyclic forms. *J Protozool,* 35**,** 214-9.

BANGS, J. D., BROUCH, E. M., RANSOM, D. M. & ROGGY, J. L. 1996. A soluble secretory reporter system in Trypanosoma brucei. Studies on endoplasmic reticulum targeting. *J Biol Chem,* 271**,** 18387-93.

BANGS, J. D., UYETAKE, L., BRICKMAN, M. J., BALBER, A. E. & BOOTHROYD, J. C. 1993. Molecular cloning and cellular localization of a BiP homologue in Trypanosoma brucei. Divergent ER retention signals in a lower eukaryote. *J Cell Sci,* 105 (Pt 4)**,** 1101-13.

BARRY, P. 2013. *Investigation of the role of the flagellar adhesion glycoprotein, FLA3, in the T. brucei parasite species.* Senior Sophister Thesis, Trinity College Dublin.

BARRY, P. 2017. *The Roles of Cytoplasmic Aspartate Aminotransferase in Trypanosoma brucei.* PhD thesis, Trinity College Dublin.

BARTOSSEK, T., JONES, N. G., SCHÄFER, C., CVITKOVIĆ, M., GLOGGER, M., MOTT, H. R., KUPER, J., BRENNICH, M., CARRINGTON, M., SMITH, A.-S., FENZ, S., KISKER, C. & ENGSTLER, M. 2017. Structural basis for the shielding function of the dynamic trypanosome variant surface glycoprotein coat. *Nature Microbiology,* 2**,** 1523-1532.

BASTIN, P., BAGHERZADEH, Z., MATTHEWS, K. R. & GULL, K. 1996. A novel epitope tag system to study protein targeting and organelle biogenesis in Trypanosoma brucei. *Mol Biochem Parasitol,* 77**,** 235-9.

BASTIN, P., PULLEN, T. J., MOREIRA-LEITE, F. F. & GULL, K. 2000. Inside and outside of the trypanosome flagellum:a multifunctional organelle. *Microbes Infect,* 2**,** 1865-74.

BENTLEY, S. J., JAMABO, M. & BOSHOFF, A. 2019. The Hsp70/J-protein machinery of the African trypanosome, Trypanosoma brucei. *Cell stress & chaperones,* 24**,** 125-148.

BERRIMAN, M., GHEDIN, E., HERTZ-FOWLER, C., BLANDIN, G., RENAULD, H., BARTHOLOMEU, D. C., LENNARD, N. J., CALER, E., HAMLIN, N. E., HAAS, B., BOHME, U., HANNICK, L., ASLETT, M. A., SHALLOM, J., MARCELLO, L., HOU, L., WICKSTEAD, B., ALSMARK, U. C., ARROWSMITH, C., ATKIN, R. J., BARRON, A. J., BRINGAUD, F., BROOKS, K., CARRINGTON, M., CHEREVACH, I., CHILLINGWORTH, T. J., CHURCHER, C., CLARK, L. N., CORTON, C. H., CRONIN, A., DAVIES, R. M., DOGGETT, J., DJIKENG, A., FELDBLYUM, T., FIELD, M. C., FRASER, A., GOODHEAD, I., HANCE, Z., HARPER, D., HARRIS, B. R., HAUSER, H., HOSTETLER, J., IVENS, A., JAGELS, K., JOHNSON, D.,

JOHNSON, J., JONES, K., KERHORNOU, A. X., KOO, H., LARKE, N., LANDFEAR, S., LARKIN, C., LEECH, V., LINE, A., LORD, A., MACLEOD, A., MOONEY, P. J., MOULE, S., MARTIN, D. M., MORGAN, G. W., MUNGALL, K., NORBERTCZAK, H., ORMOND, D., PAI, G., PEACOCK, C. S., PETERSON, J., QUAIL, M. A., RABBINOWITSCH, E., RAJANDREAM, M. A., REITTER, C., SALZBERG, S. L., SANDERS, M., SCHOBEL, S., SHARP, S., SIMMONDS, M., SIMPSON, A. J., TALLON, L., TURNER, C. M., TAIT, A., TIVEY, A. R., VAN AKEN, S., WALKER, D., WANLESS, D., WANG, S., WHITE, B., WHITE, O., WHITEHEAD, S., WOODWARD, J., WORTMAN, J., ADAMS, M. D., EMBLEY, T. M., GULL, K., ULLU, E., BARRY, J. D., FAIRLAMB, A. H., OPPERDOES, F., BARRELL, B. G., DONELSON, J. E., HALL, N., FRASER, C. M., et al. 2005. The genome of the African trypanosome Trypanosoma brucei. *Science,* 309**,** 416-22.

BERTIAUX, E., MORGA, B., BLISNICK, T., ROTUREAU, B. & BASTIN, P. 2018. A Grow-and-Lock Model for the Control of Flagellum Length in Trypanosomes. *Current Biology*.

BINDELS, D. S., HAARBOSCH, L., VAN WEEREN, L., POSTMA, M., WIESE, K. E., MASTOP, M., AUMONIER, S., GOTTHARD, G., ROYANT, A., HINK, M. A. & GADELLA, T. W., JR. 2017. mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nat Methods,* 14**,** 53-56.

BRICKMAN, M. J. & BALBER, A. E. 1990. Trypanosoma brucei rhodesiense bloodstream forms: surface ricin-binding glycoproteins are localized exclusively in the flagellar pocket and the flagellar adhesion zone. *J Protozool,* 37**,** 219-24.

BROADHEAD, R., DAWE, H. R., FARR, H., GRIFFITHS, S., HART, S. R., PORTMAN, N., SHAW, M. K., GINGER, M. L., GASKELL, S. J., MCKEAN, P. G. & GULL, K. 2006. Flagellar motility is required for the viability of the bloodstream trypanosome. *Nature,* 440**,** 224-7.

BROSSON, S., FONTAINE, F., VERMEERSCH, M., PEREZ-MORGA, D., PAYS, E., BOUSBATA, S. & SALMON, D. 2016. Specific Endocytosis Blockade of Trypanosoma cruzi Exposed to a Poly-LAcNAc Binding Lectin Suggests that Lectin-Sugar Interactions Participate to Receptor-Mediated Endocytosis. *PloS one,* 11**,** e0163302-e0163302.

BRUN, R., BLUM, J., CHAPPUIS, F. & BURRI, C. 2010. Human African trypanosomiasis. *The Lancet,* 375**,** 148-159.

BUTTER, F., BUCERIUS, F., MICHEL, M., CICOVA, Z., MANN, M. & JANZEN, C. J. 2013. Comparative proteomics of two life cycle stages of stable isotope-labeled Trypanosoma brucei reveals novel components of the parasite's host adaptation machinery. *Molecular & cellular proteomics : MCP,* 12**,** 172-179.

CARRINGTON, M., CARNALL, N., CROW, M. S., GAUD, A., REDPATH, M. B., WASUNNA, C. L. & WEBB, H. 1998. The properties and function of the glycosylphosphatidylinositol-phospholipase C in Trypanosoma brucei. *Molecular and Biochemical Parasitology,* 91**,** 153-164.

CHEN, Y., HUNG, C.-H., BURDERER, T. & LEE, G.-S. M. 2003. Development of RNA interference revertants in Trypanosoma brucei cell lines generated with a double stranded RNA expression construct driven by two opposing promoters. *Molecular and Biochemical Parasitology,* 126**,** 275-279.

CLAYTON, C. 2019. Regulation of gene expression in trypanosomatids: living with polycistronic transcription. *Open biology,* 9**,** 190072-190072.

CLAYTON, C. & SHAPIRA, M. 2007. Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Molecular and Biochemical Parasitology,* 156**,** 93-101.

DE HAAN, N., REIDING, K. R., KRIŠTIĆ, J., HIPGRAVE EDERVEEN, A. L., LAUC, G. & WUHRER, M. 2017. The N-Glycosylation of Mouse Immunoglobulin G (IgG)-Fragment Crystallizable Differs Between IgG Subclasses and Strains. *Frontiers in immunology,* 8**,** 608-608.

DE RYCKER, M., BARAGAÑA, B., DUCE, S. L. & GILBERT, I. H. 2018. Challenges and recent progress in drug discovery for tropical diseases. *Nature,* 559**,** 498-506.

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.

DEAN, S., SUNTER, J. D. & WHEELER, R. J. 2017. TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. *Trends in Parasitology,* 33**,** 80-82.

DEAN, S. D. & MATTHEWS, K. R. 2007. Restless gossamers: antibody clearance by hydrodynamic flow forces generated at the surface of motile trypanosome parasites. *Cell host & microbe,* 2**,** 279-281.

DEEKS, E. D. 2019. Fexinidazole: First Global Approval. *Drugs,* 79**,** 215-220.

DEJUNG, M., SUBOTA, I., BUCERIUS, F., DINDAR, G., FREIWALD, A., ENGSTLER, M., BOSHART, M., BUTTER, F. & JANZEN, C. J. 2016. Quantitative Proteomics Uncovers Novel Factors Involved in Developmental Differentiation of Trypanosoma brucei. *PLoS pathogens,* 12**,** e1005439-e1005439.

DESQUESNES, M., HOLZMULLER, P., LAI, D.-H., DARGANTES, A., LUN, Z.-R. & JITTAPLAPONG, S. 2013. Trypanosoma evansi and Surra: A Review and Perspectives on Origin, History, Distribution, Taxonomy, Morphology, Hosts, and Pathogenic Effects. *BioMed Research International,* 2013**,** 22.

DJIKENG, A., SHI, H., TSCHUDI, C. & ULLU, E. 2001. RNA interference in Trypanosoma brucei: cloning of small interfering RNAs provides evidence for retroposon-derived 24-26-nucleotide RNAs. *Rna,* 7**,** 1522-30.

DOCAMPO, R. 2011. Molecular parasitology in the 21st century. *Essays Biochem,* 51**,** 1-13.

DUFFY, J., PATHAM, B. & MENSA-WILMOT, K. 2010. Discovery of functional motifs in h-regions of trypanosome signal sequences. *Biochemical Journal,* 426**,** 135-145.

DURAND-DUBIEF, M., KOHL, L. & BASTIN, P. 2003. Efficiency and specificity of RNA interference generated by intra- and intermolecular double stranded RNA in Trypanosoma brucei. *Mol Biochem Parasitol,* 129**,** 11-21.

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**,** 505-515.

FADDA, A., RYTEN, M., DROLL, D., ROJAS, F., FARBER, V., HAANSTRA, J. R., MERCE, C., BAKKER, B. M., MATTHEWS, K. & CLAYTON, C. 2014. Transcriptome-wide analysis of trypanosome mRNA decay reveals complex degradation kinetics and suggests a role for co-transcriptional degradation in determining mRNA levels. *Mol Microbiol,* 94**,** 307-26.

FORT, C., BONNEFOY, S., KOHL, L. & BASTIN, P. 2016. Intraflagellar transport is required for the maintenance of the trypanosome flagellum composition but not length. *J Cell Sci*.

GADELHA, C., ZHANG, W., CHAMBERLAIN, J. W., CHAIT, B. T., WICKSTEAD, B. & FIELD, M. C. 2015. Architecture of a Host-Parasite Interface: Complex Targeting Mechanisms Revealed Through Proteomics. *Mol Cell Proteomics,* 14**,** 1911-26.

GARCÍA-SALCEDO, J. A., GIJÓN, P., NOLAN, D. P., TEBABI, P. & PAYS, E. 2003. A chromosomal SIR2 homologue with both histone NAD-dependent ADP-

ribosyltransferase and deacetylase activities is involved in DNA repair in Trypanosoma brucei. *The EMBO journal,* 22**,** 5851-5862.
GIBSON, W. 2015. Liaisons dangereuses: sexual recombination among pathogenic trypanosomes. *Research in Microbiology,* 166**,** 459-466.
GOOD, M. C., GREENSTEIN, A. E., YOUNG, T. A., NG, H.-L. & ALBER, T. 2004. Sensor Domain of the Mycobacterium tuberculosis Receptor Ser/Thr Protein Kinase, PknD, forms a Highly Symmetric β Propeller. *Journal of Molecular Biology,* 339**,** 459-469.
GUNASEKERA, K., WUTHRICH, D., BRAGA-LAGACHE, S., HELLER, M. & OCHSENREITER, T. 2012. Proteome remodelling during development from blood to insect-form Trypanosoma brucei quantified by SILAC and mass spectrometry. *BMC Genomics,* 13**,** 556.
GUO, H., SUN, J., LI, X., XIONG, Y., WANG, H., SHU, H., ZHU, R., LIU, Q., HUANG, Y., MADLEY, R., WANG, Y., CUI, J., ARVAN, P. & LIU, M. 2018. Positive charge in the n-region of the signal peptide contributes to efficient post-translational translocation of small secretory preproteins. *The Journal of biological chemistry,* 293**,** 1899-1907.
HAILE, S. & PAPADOPOULOU, B. 2007. Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Current Opinion in Microbiology,* 10**,** 569-577.
HARTEL, A. J. W., GLOGGER, M., JONES, N. G., ABUILLAN, W., BATRAM, C., HERMANN, A., FENZ, S. F., TANAKA, M. & ENGSTLER, M. 2016. N-glycosylation enables high lateral mobility of GPI-anchored proteins at a molecular crowding threshold. *Nature Communications,* 7**,** 12870.
HAYES, P., VARGA, V., OLEGO-FERNANDEZ, S., SUNTER, J., GINGER, M. L. & GULL, K. 2014. Modulation of a cytoskeletal calpain-like protein induces major transitions in trypanosome morphology. *The Journal of Cell Biology,* 206**,** 377.
HEMA THANKA CHRISTLET, T. & VELURAJA, K. 2001. Database Analysis of O-Glycosylation Sites in Proteins. *Biophysical Journal,* 80**,** 952-960.
HISS, J. A. & SCHNEIDER, G. 2009. Architecture, function and prediction of long signal peptides. *Briefings in Bioinformatics,* 10**,** 569-578.
HO, H. H., HE, C. Y., DE GRAFFENRIED, C. L., MURRELLS, L. J. & WARREN, G. 2006. Ordered assembly of the duplicating Golgi in Trypanosoma brucei</em&gt. *Proceedings of the National Academy of Sciences,* 103**,** 7676.
HÖÖG, J. L., BOUCHET-MARQUIS, C., MCINTOSH, J. R., HOENGER, A. & GULL, K. 2012. Cryo-electron tomography and 3-D analysis of the intact flagellum in Trypanosoma brucei. *J Struct Biol,* 178**,** 189-98.
HUGHES, L., TOWERS, K., STARBORG, T., GULL, K. & VAUGHAN, S. 2013. A cell-body groove housing the new flagellum tip suggests an adaptation of cellular morphogenesis for parasitism in the bloodstream form of Trypanosoma brucei. *J Cell Sci,* 126**,** 5748-57.
JACKSON, D. G., WINDLE, H. J. & VOORHEIS, H. P. 1993. The identification, purification, and characterization of two invariant surface glycoproteins located beneath the surface coat barrier of bloodstream forms of Trypanosoma brucei. *Journal of Biological Chemistry,* 268**,** 8085-8095.
JENSEN, B. C., RAMASAMY, G., VASCONCELOS, E. J., INGOLIA, N. T., MYLER, P. J. & PARSONS, M. 2014. Extensive stage-regulation of translation revealed by ribosome profiling of Trypanosoma brucei. *BMC Genomics,* 15**,** 911.
KÄLL, L., KROGH, A. & SONNHAMMER, E. L. L. 2004. A Combined Transmembrane Topology and Signal Peptide Prediction Method. *Journal of Molecular Biology,* 338**,** 1027-1036.

KALTNER, H., GARCÍA CABALLERO, G., LUDWIG, A.-K., MANNING, J. C. & GABIUS, H.-J. 2018. From glycophenotyping by (plant) lectin histochemistry to defining functionality of glycans by pairing with endogenous lectins. *Histochemistry and Cell Biology,* 149**,** 547-568.

KAWASHIMA, H., SUEYOSHI, S., LI, H., YAMAMOTO, K. & OSAWA, T. 1990. Carbohydrate binding specificities of several poly-N-acetyllactosamine-binding lectins. *Glycoconj J,* 7**,** 323-34.

KELLY, S., REED, J., KRAMER, S., ELLIS, L., WEBB, H., SUNTER, J., SALJE, J., MARINSEK, N., GULL, K., WICKSTEAD, B. & CARRINGTON, M. 2007. Functional genomics in Trypanosoma brucei: a collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci. *Mol Biochem Parasitol,* 154**,** 103-9.

KHARE, S., NAGLE, A. S., BIGGART, A., LAI, Y. H., LIANG, F., DAVIS, L. C., BARNES, S. W., MATHISON, C. J. N., MYBURGH, E., GAO, M.-Y., GILLESPIE, J. R., LIU, X., TAN, J. L., STINSON, M., RIVERA, I. C., BALLARD, J., YEH, V., GROESSL, T., FEDERE, G., KOH, H. X. Y., VENABLE, J. D., BURSULAYA, B., SHAPIRO, M., MISHRA, P. K., SPRAGGON, G., BROCK, A., MOTTRAM, J. C., BUCKNER, F. S., RAO, S. P. S., WEN, B. G., WALKER, J. R., TUNTLAND, T., MOLTENI, V., GLYNNE, R. J. & SUPEK, F. 2016. Proteasome inhibition for treatment of leishmaniasis, Chagas disease and sleeping sickness. *Nature,* 537**,** 229-233.

KOELLER, C. M., TIENGWE, C., SCHWARTZ, K. J. & BANGS, J. D. 2020. Steric constraints control processing of glycosylphosphatidylinositol anchors in Trypanosoma brucei. *Journal of Biological Chemistry*.

KOHL, L. & BASTIN, P. 2005. The flagellum of trypanosomes. *Int Rev Cytol,* 244**,** 227-85.

KOHL, L., ROBINSON, D. & BASTIN, P. 2003. Novel roles for the flagellum in cell morphogenesis and cytokinesis of trypanosomes. *Embo j,* 22**,** 5336-46.

KOLEV, N. G., FRANKLIN, J. B., CARMI, S., SHI, H., MICHAELI, S. & TSCHUDI, C. 2010. The Transcriptome of the Human Pathogen Trypanosoma brucei at Single-Nucleotide Resolution. *PLOS Pathogens,* 6**,** e1001090.

KRAMER, S., QUEIROZ, R., ELLIS, L., HOHEISEL, J. D., CLAYTON, C. & CARRINGTON, M. 2010. The RNA helicase DHH1 is central to the correct expression of many developmentally regulated mRNAs in trypanosomes. *Journal of cell science,* 123**,** 699-711.

KUMARI, P., AESCHIMANN, F., GAIDATZIS, D., KEUSCH, J. J., GHOSH, P., NEAGU, A., PACHULSKA-WIECZOREK, K., BUJNICKI, J. M., GUT, H., GROßHANS, H. & CIOSK, R. 2018. Evolutionary plasticity of the NHL domain underlies distinct solutions to RNA recognition. *Nature Communications,* 9**,** 1549.

LACOMBLE, S., VAUGHAN, S., GADELHA, C., MORPHEW, M. K., SHAW, M. K., MCINTOSH, J. R. & GULL, K. 2009. Three-dimensional cellular architecture of the flagellar pocket and associated cytoskeleton in trypanosomes revealed by electron microscope tomography. *J Cell Sci,* 122**,** 1081-90.

LACOUNT, D. J., BARRETT, B. & DONELSON, J. E. 2002. Trypanosoma brucei FLA1 is required for flagellum attachment and cytokinesis. *J Biol Chem,* 277**,** 17580-8.

LACOUNT, D. J., BRUSE, S., HILL, K. L. & DONELSON, J. E. 2000. Double-stranded RNA interference in Trypanosoma brucei using head-to-head promoters. *Molecular and Biochemical Parasitology,* 111**,** 67-76.

LAI, D.-H., HASHIMI, H., LUN, Z.-R., AYALA, F. J. & LUKEŠ, J. 2008. Adaptations of Trypanosoma brucei to gradual loss of kinetoplast DNA: Trypanosoma equiperdum

and Trypanosoma evansi are petite mutants of T. brucei. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 1999-2004.

LIU, B., LEE, G., WU, J., DEMING, J., KUEI, C., HARRINGTON, A., WANG, L., TOWNE, J., LOVENBERG, T., LIU, C. & SUN, S. 2020. The PAR2 signal peptide prevents premature receptor cleavage and activation. *PloS one,* 15**,** e0222685-e0222685.

MACGREGOR, P., SZOOR, B., SAVILL, N. J. & MATTHEWS, K. R. 2012. Trypanosomal immune evasion, chronicity and transmission: an elegant balancing act. *Nat Rev Microbiol,* 10**,** 431-8.

MANFUL, T., FADDA, A. & CLAYTON, C. 2011. The role of the 5'-3' exoribonuclease XRNA in transcriptome-wide mRNA degradation. *RNA (New York, N.Y.),* 17**,** 2039-2047.

MARCIANO, D., LLORENTE, C., MAUGERI, D. A., DE LA FUENTE, C., OPPERDOES, F., CAZZULO, J. J. & NOWICKI, C. 2008. Biochemical characterization of stage-specific isoforms of aspartate aminotransferases from Trypanosoma cruzi and Trypanosoma brucei. *Molecular and Biochemical Parasitology,* 161**,** 12-20.

MARIC, D., EPTING, C. L. & ENGMAN, D. M. 2010. Composition and sensory function of the trypanosome flagellar membrane. *Current opinion in microbiology,* 13**,** 466-472.

MARTOGLIO, B. & DOBBERSTEIN, B. 1998. Signal sequences: more than just greasy peptides. *Trends in Cell Biology,* 8**,** 410-415.

MCALLASTER, M. R., IKEDA, K. N., LOZANO-NUNEZ, A., ANRATHER, D., UNTERWURZACHER, V., GOSSENREITER, T., PERRY, J. A., CRICKLEY, R., MERCADANTE, C. J., VAUGHAN, S. & DE GRAFFENRIED, C. L. 2015. Proteomic identification of novel cytoskeletal proteins associated with TbPLK, an essential regulator of cell morphogenesis in Trypanosoma brucei. *Mol Biol Cell,* 26**,** 3013-29.

MEHLERT, A., WORMALD, M. R. & FERGUSON, M. A. 2012. Modeling of the N-glycosylated transferrin receptor suggests how transferrin binding can occur within the surface coat of Trypanosoma brucei. *PLoS Pathog,* 8**,** e1002618.

MERKLE, R. K. & CUMMINGS, R. D. 1987. Relationship of the terminal sequences to the length of poly-N-acetyllactosamine chains in asparagine-linked oligosaccharides from the mouse lymphoma cell line BW5147. Immobilized tomato lectin interacts with high affinity with glycopeptides containing long poly-N-acetyllactosamine chains. *J Biol Chem,* 262**,** 8179-89.

MOREIRA-LEITE, F. F., SHERWIN, T., KOHL, L. & GULL, K. 2001. A trypanosome structure involved in transmitting cytoplasmic information during cell division. *Science,* 294**,** 610-2.

MOREIRA, B. P., FONSECA, C. K., HAMMARTON, T. C. & BAQUI, M. M. 2017. Giant FAZ10 is required for flagellum attachment zone stabilization and furrow positioning in Trypanosoma brucei. *J Cell Sci,* 130**,** 1179-1193.

MORRISWOOD, B., HAVLICEK, K., DEMMEL, L., YAVUZ, S., SEALEY-CARDONA, M., VIDILASERIS, K., ANRATHER, D., KOSTAN, J., DJINOVIC-CARUGO, K., ROUX, K. J. & WARREN, G. 2013. Novel bilobe components in Trypanosoma brucei identified using proximity-dependent biotinylation. *Eukaryot Cell,* 12**,** 356-67.

MOTYKA, S. A. & ENGLUND, P. T. 2004. RNA interference for analysis of gene function in trypanosomatids. *Current Opinion in Microbiology,* 7**,** 362-368.

NILSSON, D., GUNASEKERA, K., MANI, J., OSTERAS, M., FARINELLI, L., BAERLOCHER, L., RODITI, I. & OCHSENREITER, T. 2010. Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of Trypanosoma brucei. *PLoS Pathog,* 6**,** e1001037.

NIU, C., LUO, H., SHI, P., HUANG, H., WANG, Y., YANG, P. & YAO, B. 2016. N-Glycosylation Improves the Pepsin Resistance of Histidine Acid Phosphatase Phytases by Enhancing Their Stability at Acidic pHs and Reducing Pepsin's Accessibility to Its Cleavage Sites. *Applied and Environmental Microbiology,* 82**,** 1004.

NOLAN, D. P., GEUSKENS, M. & PAYS, E. 1999. N-linked glycans containing linear poly-N-acetyllactosamine as sorting signals in endocytosis in Trypanosoma brucei. *Current Biology,* 9**,** 1169-S1.

NOZAKI, T., HAYNES, P. A. & CROSS, G. A. M. 1996. Characterization of the Trypanosoma brucei homologue of a Trypanosoma cruzi flagellum-adhesion glycoprotein. *Molecular and Biochemical Parasitology,* 82**,** 245-255.

OBERHOLZER, M., LANGOUSIS, G., NGUYEN, H. T., SAADA, E. A., SHIMOGAWA, M. M., JONSSON, Z. O., NGUYEN, S. M., WOHLSCHLEGEL, J. A. & HILL, K. L. 2011. Independent Analysis of the Flagellum Surface and Matrix Proteomes Provides Insight into Flagellum Signaling in Mammalian-infectious Trypanosoma brucei. *Molecular & Cellular Proteomics : MCP,* 10**,** M111.010538.

OBERHOLZER, M., LOPEZ, M. A., RALSTON, K. S. & HILL, K. L. 2009. Approaches for functional analysis of flagellar proteins in African trypanosomes. *Methods Cell Biol,* 93**,** 21-57.

OGURI, S. 2005. Analysis of sugar chain-binding specificity of tomato lectin using lectin blot: recognition of high mannose-type N-glycans produced by plants and yeast. *Glycoconj J,* 22**,** 453-61.

OWJI, H., NEZAFAT, N., NEGAHDARIPOUR, M., HAJIEBRAHIMI, A. & GHASEMI, Y. 2018. A comprehensive review of signal peptides: Structure, roles, and applications. *European Journal of Cell Biology,* 97**,** 422-441.

POON, S. K., PEACOCK, L., GIBSON, W., GULL, K. & KELLY, S. 2012. A modular and optimized single marker system for generating Trypanosoma brucei cell lines expressing T7 RNA polymerase and the tetracycline repressor. *Open Biol,* 2**,** 110037.

QIU, S., ADEMA, C. M. & LANE, T. 2005. A computational study of off-target effects of RNA interference. *Nucleic acids research,* 33**,** 1834-1847.

QUEIROZ, R., BENZ, C., FELLENBERG, K., HOHEISEL, J. D. & CLAYTON, C. 2009. Transcriptome analysis of differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons. *BMC Genomics,* 10**,** 495-495.

QUINTANA, J. F., PINO, R. C. D., YAMADA, K. & ZHANG, N. 2018. Adaptation and Therapeutic Exploitation of the Plasma Membrane of African Trypanosomes. *Genes,* 9**,** 368.

RETTIG, J., WANG, Y., SCHNEIDER, A. & OCHSENREITER, T. 2012. Dual targeting of isoleucyl-tRNA synthetase in Trypanosoma brucei is mediated through alternative trans-splicing. *Nucleic Acids Res,* 40**,** 1299-306.

ROBINSON, D. R., SHERWIN, T., PLOUBIDOU, A., BYARD, E. H. & GULL, K. 1995. Microtubule polarity and dynamics in the control of organelle positioning, segregation, and cytokinesis in the trypanosome cell cycle. *J Cell Biol,* 128**,** 1163-72.

RODITI, I., SCHWARZ, H., PEARSON, T. W., BEECROFT, R. P., LIU, M. K., RICHARDSON, J. P., BÜHRING, H. J., PLEISS, J., BÜLOW, R., WILLIAMS, R. O. & ET AL. 1989. Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of Trypanosoma brucei. *The Journal of cell biology,* 108**,** 737-746.

ROJAS, F., SILVESTER, E., YOUNG, J., MILNE, R., TETTEY, M., HOUSTON, D. R., WALKINSHAW, M. D., PEREZ-PI, I., AUER, M., DENTON, H., SMITH, T. K., THOMPSON, J. & MATTHEWS, K. R. 2019. Oligopeptide Signaling through TbGPR89 Drives Trypanosome Quorum Sensing. *Cell,* 176**,** 306-317.e16.

ROPER, J. R., GÜTHER, M. L. S., MILNE, K. G. & FERGUSON, M. A. J. 2002. Galactose metabolism is essential for the African sleeping sickness parasite Trypanosoma brucei</em&gt. *Proceedings of the National Academy of Sciences,* 99**,** 5884.

ROTUREAU, B., BLISNICK, T., SUBOTA, I., JULKOWSKA, D., CAYET, N., PERROT, S. & BASTIN, P. 2014. Flagellar adhesion in Trypanosoma brucei relies on interactions between different skeletal structures in the flagellum and cell body. *J Cell Sci,* 127**,** 204-15.

RUBOTHAM, J., WOODS, K., GARCIA-SALCEDO, J. A., PAYS, E. & NOLAN, D. P. 2005. Characterization of two protein disulfide isomerases from the endocytic pathway of bloodstream forms of Trypanosoma brucei. *J Biol Chem,* 280**,** 10410-8.

SALMON, D., HANOCQ-QUERTIER, J., PATURIAUX-HANOCQ, F., PAYS, A., TEBABI, P., NOLAN, D. P., MICHEL, A. & PAYS, E. 1997. Characterization of the ligand-binding site of the transferrin receptor in Trypanosoma brucei demonstrates a structural relationship with the N-terminal domain of the variant surface glycoprotein. *The EMBO journal,* 16**,** 7272-7278.

SCHÖTTLER, S. 2010. *Functional characterisation of a putatitive flagellar adhesion protein in Trypanosoma brucei.* Senior Sophister Thesis, Trinity College Dublin.

SCHWARTZ, K. J., PECK, R. F. & BANGS, J. D. 2013. Intracellular trafficking and glycobiology of TbPDI2, a stage-specific protein disulfide isomerase in Trypanosoma brucei. *Eukaryotic cell,* 12**,** 132-141.

SERRICCHIO, M. & BUTIKOFER, P. 2011. Trypanosoma brucei: a model micro-organism to study eukaryotic phospholipid biosynthesis. *Febs j,* 278**,** 1035-46.

SHANER, N. C., LAMBERT, G. G., CHAMMAS, A., NI, Y., CRANFILL, P. J., BAIRD, M. A., SELL, B. R., ALLEN, J. R., DAY, R. N., ISRAELSSON, M., DAVIDSON, M. W. & WANG, J. 2013. A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. *Nat Methods,* 10**,** 407-9.

SHAW, A. P., CECCHI, G., WINT, G. R., MATTIOLI, R. C. & ROBINSON, T. P. 2014. Mapping the economic benefits to livestock keepers from intervening against bovine trypanosomosis in Eastern Africa. *Prev Vet Med,* 113**,** 197-210.

SHAW, S., DEMARCO, S. F., REHMANN, R., WENZLER, T., FLORINI, F., RODITI, I. & HILL, K. L. 2019. Flagellar cAMP signaling controls trypanosome progression through host tissues. *Nature Communications,* 10**,** 803.

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.

SHIMOGAWA, M. M., SAADA, E. A., VASHISHT, A. A., BARSHOP, W. D., WOHLSCHLEGEL, J. A. & HILL, K. L. 2015. Cell Surface Proteomics Provides Insight into Stage-Specific Remodeling of the Host-Parasite Interface in Trypanosoma brucei. *Mol Cell Proteomics,* 14**,** 1977-88.

SIEGEL, T. N., GUNASEKERA, K., CROSS, G. A. M. & OCHSENREITER, T. 2011. Gene expression in Trypanosoma brucei: lessons from high-throughput RNA sequencing. *Trends in Parasitology,* 27**,** 434-441.

SIEGEL, T. N., HEKSTRA, D. R., WANG, X., DEWELL, S. & CROSS, G. A. M. 2010. Genome-wide analysis of mRNA abundance in two life-cycle stages of Trypanosoma brucei

and identification of splicing and polyadenylation sites. *Nucleic Acids Research,* 38**,** 4946-4957.

SLACK, F. J. & RUVKUN, G. 1998. A novel repeat domain that is often associated with RING finger and B-box motifs. *Trends in Biochemical Sciences,* 23**,** 474-475.

SOHAIL, M. 2005. Gene Silencing by RNAi Interference: Technology and Application.

SOULETTE, C. M., OLIVERIO, O. & ROY, S. W. 2019. On the Function of Trans-Splicing: No Evidence for Widespread Proteome Diversification in Trypanosomes. *Genome biology and evolution,* 11**,** 3014-3021.

STANLEY, P. & CUMMINGS, R. D. 2017. Structures Common to Different Glycans. *In:* VARKI A, C. R., ESKO JD, ET AL. (ed.) *Essentials of Glycobiology [Internet], 3rd edition, Chapter 14.* Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.

STIJLEMANS, B., CALJON, G., NATESAN, S. K. A., SAERENS, D., CONRATH, K., PÉREZ-MORGA, D., SKEPPER, J. N., NIKOLAOU, A., BRYS, L., PAYS, E., MAGEZ, S., FIELD, M. C., DE BAETSELIER, P. & MUYLDERMANS, S. 2011. High Affinity Nanobodies against the Trypanosome brucei VSG Are Potent Trypanolytic Agents that Block Endocytosis. *PLoS Pathogens,* 7**,** e1002072.

STIJLEMANS, B., CALJON, G., VAN DEN ABBEELE, J., VAN GINDERACHTER, J. A., MAGEZ, S. & DE TREZ, C. 2016. Immune Evasion Strategies of Trypanosoma brucei within the Mammalian Host: Progression to Pathogenicity. *Frontiers in Immunology,* 7**,** 233.

STIJLEMANS, B., CONRATH, K., CORTEZ-RETAMOZO, V., VAN XONG, H., WYNS, L., SENTER, P., REVETS, H., DE BAETSELIER, P., MUYLDERMANS, S. & MAGEZ, S. 2004. Efficient targeting of conserved cryptic epitopes of infectious agents by single domain antibodies. African trypanosomes as paradigm. *J Biol Chem,* 279**,** 1256-61.

STIJLEMANS, B., DE BAETSELIER, P., CALJON, G., VAN DEN ABBEELE, J., VAN GINDERACHTER, J. A. & MAGEZ, S. 2017. Nanobodies As Tools to Understand, Diagnose, and Treat African Trypanosomiasis. *Frontiers in Immunology,* 8**,** 724.

SUBOTA, I., JULKOWSKA, D., VINCENSINI, L., REEG, N., BUISSON, J., BLISNICK, T., HUET, D., PERROT, S., SANTI-ROCCA, J., DUCHATEAU, M., HOURDEL, V., ROUSSELLE, J. C., CAYET, N., NAMANE, A., CHAMOT-ROOKE, J. & BASTIN, P. 2014. Proteomic analysis of intact flagella of procyclic Trypanosoma brucei cells identifies novel flagellar proteins with unique sub-localization and dynamics. *Mol Cell Proteomics,* 13**,** 1769-86.

SUN, S. Y., KAELBER, J. T., CHEN, M., DONG, X., NEMATBAKHSH, Y., SHI, J., DOUGHERTY, M., LIM, C. T., SCHMID, M. F., CHIU, W. & HE, C. Y. 2018. Flagellum couples cell shape to motility in Trypanosoma brucei. *Proc Natl Acad Sci U S A*.

SUN, S. Y., WANG, C., YUAN, Y. A. & HE, C. Y. 2013. An intracellular membrane junction consisting of flagellum adhesion glycoproteins links flagellum biogenesis to cell morphogenesis in *Trypanosoma brucei*. *J Cell Sci,* 126**,** 520-31.

SUNTER, J., WEBB, H. & CARRINGTON, M. 2013. Determinants of GPI-PLC localisation to the flagellum and access to GPI-anchored substrates in trypanosomes. *PLoS Pathog,* 9**,** e1003566.

SUNTER, J. D., BENZ, C., ANDRE, J., WHIPPLE, S., MCKEAN, P. G., GULL, K., GINGER, M. L. & LUKEŠ, J. 2015a. Modulation of flagellum attachment zone protein FLAM3 and regulation of the cell shape in Trypanosoma brucei life cycle transitions. *Journal of Cell Science,* 128**,** 3117-3130.

SUNTER, J. D. & GULL, K. 2016. The Flagellum Attachment Zone: 'The Cellular Ruler' of Trypanosome Morphology. *Trends Parasitol*.

SUNTER, J. D., VARGA, V., DEAN, S. & GULL, K. 2015b. A dynamic coordination of flagellum and cytoplasmic cytoskeleton assembly specifies cell morphogenesis in trypanosomes. *Journal of Cell Science,* 128**,** 1580-1594.

TIENGWE, C., MURATORE, K. A. & BANGS, J. D. 2016. Surface proteins, ERAD and antigenic variation in Trypanosoma brucei. *Cellular microbiology,* 18**,** 1673-1688.

TRÉPOUT, S. 2020. In situ structural analysis of the flagellum attachment zone in Trypanosoma brucei using cryo-scanning transmission electron tomography. *bioRxiv***,** 2020.02.14.949115.

TSANG, K. K. W. 2012. *Attachment of the flagellum in procyclic forms of Trypanosoma brucei.* Senior Sophister thesis, Trinity College Dublin.

TSIRIGOS, K. D., GOVINDARAJAN, S., BASSOT, C., VÄSTERMARK, Å., LAMB, J., SHU, N. & ELOFSSON, A. 2018. Topology of membrane proteins—predictions, limitations and variations. *Current Opinion in Structural Biology,* 50**,** 9-17.

TSIRIGOS, K. D., PETERS, C., SHU, N., KALL, L. & ELOFSSON, A. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res,* 43**,** W401-7.

TURNOCK, D. C., IZQUIERDO, L. & FERGUSON, M. A. 2007. The de novo synthesis of GDP-fucose is essential for flagellar adhesion and cell growth in Trypanosoma brucei. *J Biol Chem,* 282**,** 28853-63.

URBANIAK, M. D., GUTHER, M. L. & FERGUSON, M. A. 2012. Comparative SILAC proteomic analysis of Trypanosoma brucei bloodstream and procyclic lifecycle stages. *PLoS One,* 7**,** e36619.

VASQUEZ, J.-J., HON, C.-C., VANSELOW, J. T., SCHLOSSER, A. & SIEGEL, T. N. 2014. Comparative ribosome profiling reveals extensive translational complexity in different Trypanosoma brucei life cycle stages. *Nucleic Acids Research,* 42**,** 3623-3637.

VAUGHAN, S. 2010. Assembly of the flagellum and its role in cell morphogenesis in Trypanosoma brucei. *Curr Opin Microbiol,* 13**,** 453-8.

VAUGHAN, S., KOHL, L., NGAI, I., WHEELER, R. J. & GULL, K. 2008. A repetitive protein essential for the flagellum attachment zone filament structure and function in Trypanosoma brucei. *Protist,* 159**,** 127-36.

VICKERMAN, K. 1969. On The Surface Coat and Flagellar Adhesion in Trypanosomes. *Journal of Cell Science,* 5**,** 163.

VICKERMAN, K. 2009. "Not a very nice subject." Changing views of parasites and parasitology in the twentieth century. *Parasitology,* 136**,** 1395-402.

VON HEIJNE, G. 1992. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J Mol Biol,* 225**,** 487-94.

WALL, R. J., RICO, E., LUKAC, I., ZUCCOTTO, F., ELG, S., GILBERT, I. H., FREUND, Y., ALLEY, M. R. K., FIELD, M. C., WYLLIE, S. & HORN, D. 2018. Clinical and veterinary trypanocidal benzoxaboroles target CPSF3. *Proceedings of the National Academy of Sciences,* 115**,** 9616.

WHEELER, R. J. 2017. Use of chiral cell shape to ensure highly directional swimming in trypanosomes. *PLoS Comput Biol,* 13**,** e1005353.

WICKSTEAD, B., ERSFELD, K. & GULL, K. 2002. Targeting of a tetracycline-inducible expression system to the transcriptionally silent minichromosomes of Trypanosoma brucei. *Mol Biochem Parasitol,* 125**,** 211-6.

WINTER, F. R., LOIDOLT, M., WESTPHAL, V., BUTKEVICH, A. N., GREGOR, C., SAHL, S. J. & HELL, S. W. 2017. Multicolour nanoscopy of fixed and living cells with a single STED beam and hyperspectral detection. *Scientific Reports,* 7**,** 46492.

WIRTZ, E., LEAL, S., OCHATT, C. & CROSS, G. A. 1999. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. *Mol Biochem Parasitol,* 99**,** 89-101.

WOODS, K. 2007. *Identification and characterisation of proteins from the tomato lectin binding fraction of Trypanosoma brucei.* PhD thesis, Trinity College Dublin.

WOODS, K., NIC A'BHAIRD, N., DOOLEY, C., PEREZ-MORGA, D. & NOLAN, D. P. 2013. Identification and Characterization of a Stage Specific Membrane Protein Involved in Flagellar Attachment in *Trypanosoma brucei*. *PLoS ONE,* 8**,** e52846.

WU, A. M., WU, J. H., SINGH, T., LAI, L. J., YANG, Z. & HERP, A. 2006. Recognition factors of Ricinus communis agglutinin 1 (RCA(1)). *Mol Immunol,* 43**,** 1700-15.

ZHOU, D. 2003. Why are glycoproteins modified by poly-N-acetyllactosamine glycoconjugates? *Current Protein and Peptide Science,* 4**,** 1-9.

ZHOU, Q., HU, H., HE, C. Y. & LI, Z. 2015. Assembly and maintenance of the flagellum attachment zone filament in Trypanosoma brucei. *Journal of Cell Science*.

ZHOU, Q., HU, H. & LI, Z. 2014. New Insights into the Molecular Mechanisms of Mitosis and Cytokinesis in Trypanosomes. *International review of cell and molecular biology,* 308**,** 127-166.

ZHOU, Q., LIU, B., SUN, Y. & HE, C. Y. 2011. A coiled-coil- and C2-domain-containing protein is required for FAZ assembly and cell morphology in Trypanosoma brucei. *J Cell Sci,* 124**,** 3848-58.