# Lifecycle progression in *Trypanosoma brucei*; genome-wide expression profiling and role of the cell cycle in this process

**Sarah Kabani**

PhD – The University of Edinburgh - 2010

# Abstract

The bloodstream form of *Trypanosoma brucei* differentiates into the stumpy form in the mammalian bloodstream, completing differentiation into the procyclic form on uptake by the tsetse fly. The underlying genetic events occurring during this differentiation process in pleomorphic cell lines were investigated through whole-genome microarray studies of key time points during differentiation from stumpy form cells to the procyclic form found in the insect midgut. The microarray was extensively validated and bioinformatic experiments conducted to detect motifs over represented in stumpy form or slender form cells. A positional-dependent motif was identified that was over represented in stumpy form cells, possibly representing a regulatory domain. The transcripts found to be enriched in stumpy form cells included a chloride channel, although RNAi directed against this gene showed no phenotype, suggesting the protein is redundant, as three other homologous proteins exist in the genome and showed similar mRNA profiles on the microarray.

Stumpy form cells are G0 arrested and two proteins implicated in G0/G1 regulation in other organisms, Target of Rapamycin (Tor) and Cdh1, were investigated in *T. brucei* to determine whether these proteins were involved in differentiation. The result of depletion of either protein was rapid cell death in bloodstream form cells, although treatment with the drug rapamycin did not have any effect on the cells in contrast to other eukaryotes where this drug causes G1 arrest. A method for synchronisation of bloodstream form cells was also designed using a supravital dye and flow cytometry to allow investigation into cell cycle-dependent processes. This method was particularly suitable for harvesting populations enriched in G0/G1 stage cells, however differentiation of the isolated G0/G1 and G2/M populations did not show significantly different differentiation kinetics.

## Declaration

I declare that, other than where stated, all work presented in thesis is my own and has not been previously submitted for examination.

## Acknowledgements

I would like to thank Keith Matthews for his supervision and for giving calming and encouraging advice when I felt things were getting particularly desperate. I am also indebted to Deborah Hall for her assistance and to Katelyn Fenn for help with all things bio-informatical. Also indispensable were Martin Waterfall and Andrew Saunderson for all their assistance with the sorting experiments, and Aziz El Hage for recommendations for the yeast experiments. I would like to thank Alan Ross and Al Ivens for their work on the microarray and Terry Smith for the mass spectrometry work. I am also enormously thankful to Irene Ruberto and her family for hosting me when I was at a conference in Verona.

I am also grateful to the entire Matthews lab for being such wonderful people to work and socialise with, and for being so amenable to all my attempts to make them bring me cakes. I would like to thank my family for being supportive throughout the whole PhD experience and for making the trek up to Edinburgh regularly to visit me. Most especially, I would like to thank Leo Hughes, for having so much faith in me and being the perfect antidote against all the times when science got to be too much.

# Abbreviations

| | |
|---|---|
| APC/C | Anaphase promoting complex/cyclosome |
| APS | Ammonium persulphate |
| BARP | Brucei alanine rich protein |
| BLAST | Basic local alignment search tool |
| bp | Base pair |
| cAMP | Cyclic adenosine monophosphate |
| Cy3 | Cyanine 3 |
| COX | Cytochrome oxidase |
| DAPI | 4,6-diamidine-2, phenylindole |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| EATRO | East African Trypanosomiasis Research Organisation |
| EP | Glutamic acid-proline protein |
| EST | Expressed sequence tag |
| FACS | Fluorescent activated cell sorting |
| FCS | Foetal calf serum |
| FITC | Fluorescein isothiocyanate |
| FKBP12 | FK506 binding protein 12 |
| FRB | Rapamycin binding domain |
| FSC | Forward scatter |
| g | gravities |
| GPEET | Glycine-proline-glutamic acid-glutamic acid-threonine procyclin |
| HAT | Human African Trypanosomiasis |
| HU | Hydroxyurea |
| JCVI | J. Craig Venter institute |
| kb | kilobase |
| LB | Luria broth |
| mRNA | Messenger RNA |
| MVSG | Metacyclic variant surface glycoprotein |
| ncRNA | Non-coding RNA |
| OD | Optical density |

| | |
|---|---|
| Loess | Locally weighted polynomial regression |
| PAD | Proteins associated with differentiation |
| PBS | Phosphate buffered saline |
| PCF | Procyclic form |
| PCR | Polymerase chain reaction |
| PDA | p-phenylenediamine |
| PGK | Phosphoglycerate kinase |
| PIKK | Phosphatidylinositol kinase-like kinase |
| qPCR | Quantitative RNA |
| RNA | Ribonucleic acid |
| RNAi | RNA interference |
| rRNA | Ribosomal RNA |
| S6K | S6 kinase |
| SIF | Stumpy induction factor |
| SL | Spliced leader |
| SM1 | Stumpy motif 1 |
| SSC | Side scatter |
| TAO | Trypanosome alternative oxidase |
| TEMED | N,N,N',N'-tetramethylenediamine |
| TOR | Target of rapamycin |
| TORC | Target of rapamycin complex |
| TOS | Target of rapamycin signalling |
| TRITC | Tetramethyl rhodamine isothiocyanate |
| UTR | Untranslated region |
| UV | Ultra violet |
| VSG | Variant surface glycoprotein |
| ZPFMG | Zimmerman postfusion buffer with glucose |
| YPD | Yeast extract-peptone-dextrose |

# Table of Contents

# Table of Figures

# 1 - Introduction

## *1.1 African Sleeping Sickness*

The parasitic kinetoplastid *Trypanosoma brucei* is the causative agent of both human African sleeping sickness (HAT) and the animal wasting disease Nagana in sub-Saharan Africa. The human disease can be caused by one of two subspecies, *T.b.gambiense* and *T.b.rhodesiense*, resulting in a chronic or acute infection respectively. The parasite carries out part of its life cycle in the tsetse fly and is spread into the mammalian host when the fly takes a bloodmeal. The vector is found across a third of the African continent (Kennedy, 2008), however instances of African sleeping sickness occur in discrete foci (Hide and Tait, 2006). Most cases of HAT are caused by the chronic *T.b.gambiense* form (Fevre et al., 2008), and measures aimed at controlling the number of human cases have been highly successful in reducing the spread of the disease as an animal reservoir of this subspecies is considered relatively insignificant for the transmission of the parasite, in contrast to the *T.b.rhodesiense* form (Simarro et al., 2008). A recently established database aims to generate an atlas of HAT cases, as has been initiated for malaria (http://www.map.ox.ac.uk/), to better understand the epidemiology of the disease. Figure 1.1 shows some of the preliminary data from this project including, importantly, those areas screened that were negative for cases of sleeping sickness (Cecchi et al., 2009).

A parasite infection is detrimental to the fly because exposure to the parasites triggers the immune response, which is an energy-requiring process and necessitates a reduction in fecundity (Hu et al., 2008). Thus, most instances of trypanosome uptake with a bloodmeal do not result in infection, as the parasites are destroyed in the tsetse midgut (Krafsur, 2009). Indeed, tsetse flies have established relationships with maternally-inherited endosymbionts such as *Wigglesworthia* through expression of a tsetse peptidoglycan recognition protein that prevents the fly's immune system from destroying these cells (Wang et al., 2009). The presence of several species of symbionts prevents establishment of a trypanosome infection, possibly through the

release by the symbionts of lectins in response the bloodmeal, which effectively destroy the parasites (Welburn and Maudlin, 1999). However, once a trypanosome population is established, the fly will remain infective for the duration of its lifespan (Kennedy, 2008).

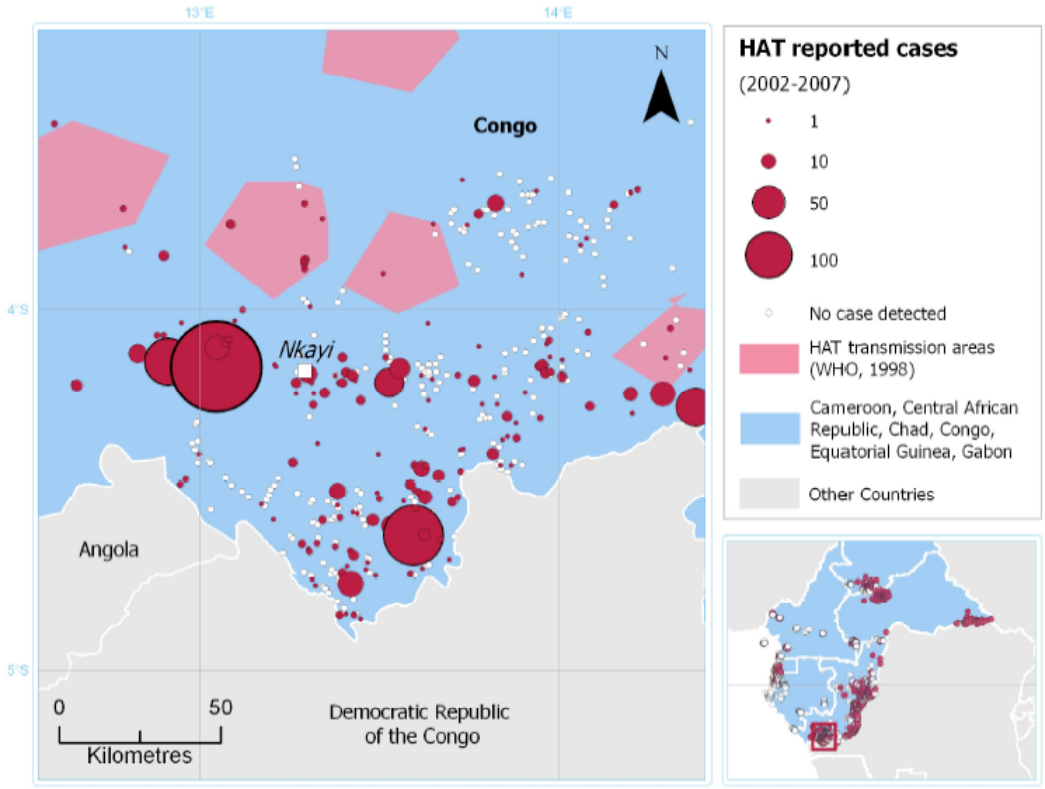



**Figure 1.1. Local level mapping of HAT cases compiled as part of the HAT atlas project**. Incidences of sleeping sickness and screening events leading to the identification of no cases can be mapped accurately enabling better understanding of the epidemiology of HAT (Cecchi et al., 2009).

The disease in humans has two distinct phases; in the early stage the symptoms are general and can be confused with those for other diseases, with typical symptoms including fever and possibly a skin rash. However, it is the symptoms arising during the second stage that are the hallmark of the disease and to which the disease owes its name; once the parasites have crossed the blood-brain barrier and infected the central nervous system, the patient experiences disrupted sleeping patterns which eventually lead to coma and death (Kennedy, 2008). At present only four drugs are prescribed for the treatment of HAT, two of which are available for the early disease, prior to the parasite crossing the blood brain barrier. Treatment for the late stage of

the disease is by either eflornithine (for only *T.b.gambiense*) or melarsoprol (for either subspecies), which has a range of side effects including death in 5% of patients who go on to develop neurological problems. Recently the use of nifurtimox, traditionally used to treat South American trypanosomiasis, has been partially successful in clearing parasites that prove resistant to other drug treatments (Rodgers, 2009). All these drugs require lengthy courses of administration, often intravenously. A better treatment regime for the stage two *T.b.gambiense* infection that has recently passed phase III clinical trials, uses a combination of nifurtimox with eflornithine, allowing the treatment to be completed sooner than with eflornithine alone, without such an intensive regime of infusions, resulting in fewer adverse side effects (Priotto et al., 2009). The use of a combination of drugs is always preferred as the chance of the parasite developing resistance to the drugs is much reduced.

## *1.2 Life cycle*

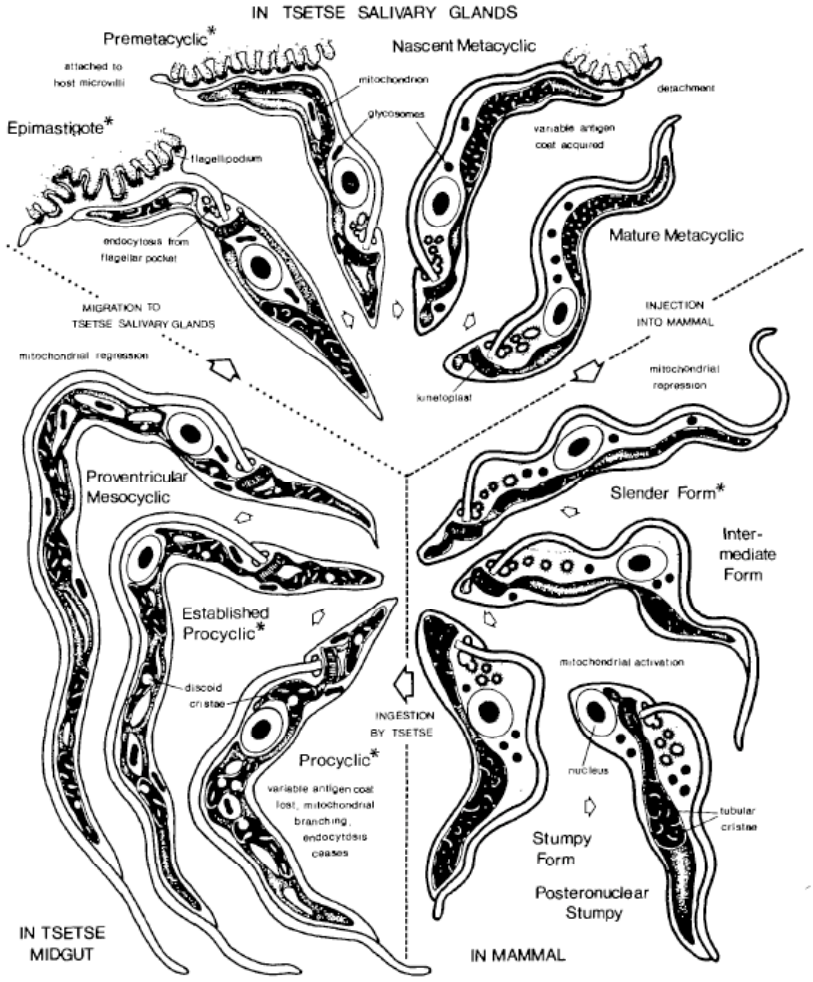



**Figure 1.2. Life cycle diagram of *T. brucei*.** Drawings showing the morphological changes of *T. brucei* as it differentiates into the different forms for survival in the tsetse salivary glands and midgut and the mammalian bloodstream (Vickerman, 1985). The epimastigotes show modification of the flagellum for attachment to the salivary gland, which is lost as the cells differentiate into the metacyclic form which is injected into the mammalian bloodstream when the fly feeds. The metacyclic form differentiates into the slender form in the bloodstream, which in turn differentiates into the stumpy form via an intermediate form in response to density- dependent signalling. The stumpy form is preadapted for uptake into the tsetse fly, and differentiates rapidly into the procyclic form, which crosses into the salivary glands through differentiation in the mesocyclic form.

### 1.2.1. Cell structure

The procyclic form trypanosome cell is around 15µm by 3µm in size (Sherwin and Gull, 1989), however the morphology of the cells changes dramatically as cells progress through the different stages of the life cycle (Figure 1.2) (Vickerman, 1985). All life stages have a flagellum, which exits the cell at the flagella pocket towards the posterior of the cell; the only site of endo- and exocytosis (Sherwin and Gull, 1989). Internally, the organelles of *T. brucei* are kept strictly controlled in terms of number and positioning during the different life cycle stages, including the kinetoplast. The presence of a kinetoplast is a unique feature of all kinetoplastid organisms, where this disc shaped organelle contains the DNA from the single mitochondrion (Liu et al., 2005).

### 1.2.2 Procyclic forms

*T. brucei* stumpy form cells are transmitted between mammalian hosts using, as a vector, the tsetse fly. The number of parasites present in the midgut of the fly three days post infection is a fraction of the number that are uptaken. This is due to the high levels of parasite death that occur as parasites adapt to the new environment, but those surviving after this time will then resume proliferation as a trypomastigote procyclic population (Van Den Abbeele et al., 1999). As stumpy forms complete differentiation to the procyclic form in the midgut, the mitochondrion of the cell rapidly becomes more elaborated, with the formation of a branched network of cristae, as the metabolic requirements of the cell change with the new environment (Vickerman et al., 1988). Other morphological changes accompanying differentiation include the repositioning of the kinetoplast from the extreme posterior of the cell, towards the nucleus, facilitated through the elongation of the posterior of the cell through the extension of cytoskeletal microtubules, although the reason for this is not known (Matthews et al., 1995). Once the differentiation process to procyclic form is completed the cells begin proliferation and the kinetoplasts will segregate. However, kinetoplast segregation can be uncoupled in cells that are unable to reposition their kinetoplasts, showing that differentiation events, although occurring in a strictly ordered manner, are not completely interdependent (Hendriks and Matthews, 2005). The variant surface glycoproteins (VSG) coat covering the

bloodstream form cell is not necessary in procyclic form cells, which instead are coated with a more sparse coat of procyclins, named after the repetitive amino acids found within them, with up to 27 'EP' repeats found in the EP procyclins, and five 'GPEET' repeats in GPEET procyclins. In contrast to the vast array of VSG proteins in bloodstream form cells, only three procyclin proteins are expressed on the surface of procyclic form cells in vivo with all three of these expressed during the early stages of a midgut infection; GPEET, EP1 and EP3, although with time, GPEET expression is lost (Acosta-Serrano et al., 2001). A fourth protein, EP2, can also be expressed but has not been detected in tsetse midgut infections (Acosta-Serrano et al., 2001). The function of the procyclin coat is a mystery; complete depletion of all EP and GPEET genes is not overly detrimental to cells and they are still able to infect tsetse flies, although the procyclin depleted cell lines are easily out-competed by wild type infections with intact procyclin coats (Vassella et al., 2009). However, potential functions for the EP surface coat include protection from proteases found in the insect midgut, or as signalling proteins generated by the removal of their N glycan (Acosta-Serrano et al., 2001). Other cellular features that are less active in this life stage are also down graded, such as endocytosis, which plays an important role in bloodstream form cells, possibly as a method to rapidly remove immune system proteins from the surface of cells (Engstler et al., 2007). Such a process is not required in procyclic form cells and, consequently, the proteins required for endocytosis are seen to rapidly fall as cells differentiate from the stumpy form into the procyclic form (Natesan et al., 2007). Procyclic form cells must migrate to the salivary glands in order to allow future transmission, so first the cells breach the peritrophic membrane to invade the endoperitrophic space and differentiate into mesocyclic form cells (Vickerman, 1985).

### 1.2.3 Epimastigote form

Mesocyclic trypomastigotes are arrested in G0/G1, presumably induced through a factor present in the midgut or nutrient limitation, as removal of these cells into culture triggers them to resume proliferation immediately (Van Den Abbeele et al., 1999). Once the mesocyclic trypomastigotes are released from their cell cycle arrest, they differentiate into a long epimastigote form. The epimastigote form is unusual in that it divides asymmetrically, such that one daughter cell is long like the parental

cell, whereas the other is much shorter. Only the shorter form is capable of attaching to the salivary gland due to the specialised nature of its flagellum, but the long form is required for migration to the salivary gland as the short flagellum of the short form impairs its motility (Van Den Abbeele et al., 1999). Epimastigote form parasites express a BARP surface coat (brucei alanine-rich proteins), although the function of this surface coat is unknown (Urwyler et al., 2007). These cells also start to show an increase in endocytosis again, possibly as a pre-adaptation for life in the mammalian host until, once the cells differentiate into the metacyclic form, the levels of endocytosis almost reach the same level as bloodstream form trypanosomes (Natesan et al., 2007).

### 1.2.4. Metacyclic form

The epimastigote form differentiates into the metacyclic form whilst still attached to the salivary glands, but once formed, the metacyclic cell is released from the epithelial wall to allow transmission the next time the insect takes up a bloodmeal (Vickerman et al., 1988). The attachment to the salivary gland is required for the differentiation to the metacyclic form, as shaking cultures of *T. congolense* (which can complete the life cycle in vitro) prevents differentiation, although the epimastigote forms are still capable of proliferation (Vickerman et al., 1988). Release from the epithelial cells is possible due to the modification of the flagellum in metacyclic forms cells, which becomes smaller and unbranched and so terminating attachment (Vickerman et al., 1988). The differentiation of epimastigotes to metacyclics is not synchronous within the population, with parasites at all stages of development found in close proximity in the salivary glands (Tetley et al., 1987). Prior to injection into a mammalian host, the parasites must again express a VSG coat, although metacyclic cells use a different repertoire of VSGs to those used in slender cells. These metacyclic VSGs (MVSGs), in contrast to bloodstream form VSGs, are expressed monocistronically, although in common with bloodstream VSGs, they are transcribed using RNA polymerase I (Barry et al., 1998). These genes are located at telomeric regions, which allows them to be effectively silenced in the other life stages of the parasite when MVSGs are not required (Ginger et al., 2002). The repertoire of MVSGs is much more limited than the bloodstream form VSGs, with only around 27 different genes to chose from, and interestingly the

nascent metacyclics do not express the same surface antigen as each other (Tetley et al., 1987). It has been suggested that the heterogeneity of MVSGs expressed by the trypanosome population is a defence against the immune system of the mammalian host, which may have previously experienced a *T. brucei* infection and thus raised antibodies against some of the MVSGs (Barry et al., 1998).

### 1.2.5 Bloodstream slender form

Trypanosomes are released into the dermal connective tissue of the host when the tsetse fly feeds, and from here the parasites migrate into the bloodstream via the lymphatic drainage system (Vickerman, 1985). The slender form cells present in the mammalian host are free living in the bloodstream and must evade clearance by the immune system and so have evolved a mechanism of expressing a dense surface coat of proteins solely to shield the essential immunogenic receptors. This surface coat is comprised of VSGs are routinely switched, so that whilst the majority of cells will be cleared through an antibody attack, a few cells will express a different surface coat and will not be recognised by the immune system, meaning the parasite infection is never entirely eliminated from the bloodstream (Seed and Wenck, 2003). Bloodstream form parasites exclusively use glycolysis to generate energy, exploiting the plentiful glucose available in the blood, and so the mitochondrion of these cells is relatively undeveloped, lacking cristae and consequently inactive with respect to ATP generation (Vickerman, 1965).

### 1.2.6 Stumpy form

A universal agreement cannot be reached upon whether stumpy forms are the pre-adapted cells required to survive uptake by a tsetse fly or whether they represent dying cells, which merely act to limit parasite numbers in the bloodstream. The switch from slender form to stumpy forms, which are unable to vary the VSG coat, means that these are rapidly cleared from the bloodstream by the host immune system unless they are taken up by a tsetse fly. The fluctuating levels of parasitaemia that are generated from this periodic clearing of the majority of the trypanosomes allows the infection to continue in a host much longer than in a monomorphic infection, where a single parasite can cause death within five days (Seed and Wenck, 2003). Certainly stumpy cells are growth arrested and

differentiation from the slender form is irreversible, leading to the hypothesis that this differentiation event is used to lower parasitaemia and thus prevent death of the host (Seed and Wenck, 2003, Tyler, 2003). The process of slender forms differentiating into stumpy form cells is clearly a major determinant of the progression of an infection and has been used as one of the main variables in a mathematical model developed to predict antigenic variation during a parasitaemia (Lythgoe et al., 2007).

A number of features inherent to stumpy forms make them good candidates for uptake into the tsetse fly. In response to acidic environments, slender form cells show an increase in cytoplasmic pH, lose their VSG coat and rapidly die, whereas the cytoplasmic pH of stumpy form cells remains unaltered and they show no sensitivity to acid treatment (Nolan et al., 2000). Stumpy cells also show much higher resistance to proteases than slender cells, which the cells would be expected to encounter on uptake into the fly midgut (Sbicego et al., 1999, Nolan et al., 2000). Interestingly, stumpy cells are also more resistant to antibody-mediated killing whilst in the host bloodstream than slender cells (McLintock et al., 1993), a useful attribute considering these cells are incapable of switching their VSG coat. The mechanism for this immune-system resistance appears to stem from the increased rate of endocytosis (Engstler et al., 2007). The motion of the swimming parasite allows the VSG surface coat, and any immunoglobulins attached to it, to be forced towards to posterior of the cells where the flagella pocket is situated where both are exported into the cells, with the VSG recycled and the immunoglobulin destroyed in the lysosomes (Engstler et al., 2007).

Significantly, the mitochondrion of stumpy form cells is larger than slender cells, and a small number of cristae are formed, reflecting the ability of the cells to initiate mitochondrial metabolism upon uptake into the fly where the source of energy is no longer glucose, but proline (Vickerman, 1965). The activation of the mitochondria in differentiating cells appears to initiate in ‘intermediate’ cells; these are cells still slender in appearance, but having committed to differentiate into stumpy forms. The

pre-emptive activation of a mitochondria is a useful attribute for stumpy cells, as once they are ingested into the tsetse fly midgut, the residual glucose from the bloodstream disappears in as little as 15 minutes, thus a rapid change in metabolism is required (Milne et al., 1998). Indeed, studies examining the activity of the mitochondria in pleomorphic bloodstream form populations have shown that both intermediate and stumpy form cells accumulate rhodamine 123, a dye that is confined to mitochondria only when a membrane potential is maintained (Bienen et al., 1991). In slender cells, in contrast, the dye is dispersed throughout the cytoplasm showing that the mitochondria in this life form does not maintain the membrane potential and thus cannot be active (Bienen et al., 1991). Therefore, even prior to the morphological events that allow discrimination of stumpy form cells, the cells are already switching over their metabolic functions in preparation for transmission. Despite this, slender form cells devoid of kinetoplasts (diskinetoplastid cells) created through treatment with intercalating dyes, have been shown to successfully complete differentiation into the stumpy form, although these parasites are unable to enter the cell cycle to become proliferative procyclic form cells (Timms et al., 2002).

One other adaptation that stumpy form cells undergo is to prime themselves to receive the differentiation trigger they will be exposed to in the midgut. It is known that cold shock enhances the differentiation of stumpy form cells to procyclic forms when exposed to cis aconitate or citrate, and recently the genes encoding the transporter for these molecules was characterised (Dean et al., 2009). Stumpy form cells initiate expression of PAD1 and PAD2 which, under cold shock conditions, are exported to the cell surface thus allowing active importation of cis aconitate or citrate resulting in differentiation to the procyclic form (Dean et al., 2009).

### *1.3 Differentiation of slender to procyclic form*

The appearance of stumpy form cells in the bloodstream has been found to occur in a density dependent manner, such that once a critical threshold has been reached, the cells will differentiate from the proliferating slender form into the growth arrested stumpy form (Reuner et al., 1997). In has been hypothesised that a substance of low

molecular weight, termed Stumpy Induction Factor (SIF), is secreted from the parasites and gradually accumulates, which crosses the cell membrane and triggers differentiation (Vassella et al., 1997). Culture adapted cells lose the ability to differentiate in response to SIF and form monomorphic populations of homogenous slender forms, which are unable to progress into the stumpy life cycle stage (Fenn and Matthews, 2007, Vassella et al., 1997). Whilst these cell lines will not differentiate when stimulated with SIF-containing medium, addition of cis-aconitate coupled with temperature reduction to 27˚C induces the monomorphic cells to undergo a non-synchronous differentiation from bloodstream form to procyclic form (Ziegelbauer et al., 1990). This in vitro differentiation can be induced through addition of cis-aconitate or citrate, but not any other of the Kreb cycle intermediates (Sbicego et al., 1999, Brun and Schonenberger, 1981).

Cyclic AMP (cAMP) was known to show differential levels throughout the parasite development cycle, and further study demonstrated that pCPTcAMP, a derivative of cAMP which is capable of crossing the cell membrane, caused differentiation of slender cells into stumpy form cells (Vassella et al., 1997, Breidbach et al., 2002). This differentiation event was even possible in monomorphic cell lines that were unable to differentiate in response to SIF, suggesting that cAMP was a downstream messenger in the SIF pathway (Vassella et al., 1997, Breidbach et al., 2002). A more recent study however, has shown that the products formed from hydrolysis of cAMP analogues are even more potent at inducing differentiation from slender form to stumpy forms, and that analogues resistant to hydrolysis were unable to trigger differentiation, effectively demonstrating that it is the hydrolysis products of cAMP that are responsible for differentiation (Laxman et al., 2006). Other chemical triggers will also induce differentiation from slender form to stumpy form, including the diabetic drug troglitazone which causes pleomorphic cells to take up a ‘stumpy-like’ morphology, although in monomorphic cells treatment with this drug does not result in cell cycle arrest (Denninger et al., 2007). Other differentiation triggers which are more likely to be representative of the events in the fly midgut include pronase,

trypsin and thermolysin, but addition of any of these into the medium of cells also caused high levels of cell death, at least among slender cells (Sbicego et al., 1999).

Cold shock of more than 15˚C has been shown to be sufficient to cause slender form cells to express EP, however the EP protein is retained within the cells (Engstler and Boshart, 2004). Interestingly, stumpy form cells subjected to cold shock also express EP but direct it to the cell surface (Engstler and Boshart, 2004). The cold shock of stumpy form cells also made them hypersensitive to cis aconitate, such that the concentration required to trigger differentiation was reduced to within the micromolar range (Engstler and Boshart, 2004). Subsequently, PAD2 was found to have increased expression in cells subjected to cold shock and the temperature change also caused this transporter protein to be routed onto the surface of the cells, in which position the protein allows the uptake of cis aconitate, thus triggering differentiation (Dean et al., 2009). The role of the temperature reduction in differentiation is intriguing; differentiation of stumpy form cells can be triggered through addition of cis-aconitate without an accompanying temperature reduction to 27˚C, but these cells are unable to re-enter the cell cycle to become proliferative procyclic form cells (Matthews and Gull, 1997). The result of this, and previous studies (Matthews and Gull, 1994), has shown that differentiation from the stumpy form and cell cycle re-entry can be separated, which is interesting as the two events are required to act in tandem for successful establishment of an infection in the tsetse fly.

### *1.4 Cell cycle*

The cell cycle of *T. brucei* varies significantly from the highly conserved events that are seen in most eukaryotic cells. Major cell cycle checkpoints are absent (McKean, 2003) and the mitochondrial DNA present in the kinetoplast is replicated in an ordered fashion in conjunction with the nuclear DNA replication, and has been implicated as being responsible for a novel checkpoint (Ploubidou et al., 1999). Using the microtubule inhibitor rhizoxin, it was discovered that a point existed towards to final stages of nuclear S phases when the drug had no effect on cells, whereas addition of rhizoxin at earlier points, when the kinetoplast was also in S phase, caused mitotic defects resulting in divisions where one daughter had two

nuclei and the other one none (Ploubidou et al., 1999). This generated anucleate cytoplasts, called zoids.

Under normal conditions, the events for the cell cycle are highly predictable, with the length of the cell cycle in procyclic forms having been variously shown to range from 8.5 hours (Woodward and Gull, 1990) to over 12 hours (Siegel et al., 2008a, Siegel et al., 2008b). Detailed analysis of the bloodstream form cell cycle has not been conducted, but from even the limited data we have so far, it is predicted that it would vary from procyclic form cells significantly. Differences between the cell cycle of bloodstream form cells and procyclic form cells known so far include the length of the cell cycle, which, at six hours, is considerably shorter in bloodstream form cells (Vickerman, 1985), to the consequences of depletion of cell cycle genes resulting in different phenotypes between bloodstream form cells and procyclic form cells (Hammarton et al., 2003, Hammarton et al., 2004, Kumar and Wang, 2005, Janzen et al., 2006, Li and Wang, 2006, Gourguechon and Wang, 2009).

In the procyclic form cell, the basal body is formed from the immature pro-basal body early on in G1, prior to replication of any other organelle, and from this the daughter flagellum starts to grow (McKean, 2003) (Figure 1.3A). The nascent flagellum grows along the length of the parental flagellum, attaching via the flagella connector (Figure 1.3B), a feature that is not present in the bloodstream form cells (Briggs et al., 2004). After this, the Golgi apparatus duplicates and then the basal bodies and Golgi separate, in a manner that appears to be coupled (Hammarton, 2007). The migration of the basal bodies initiates before the daughter flagellum has reached its full length (Sherwin and Gull, 1989). The basal bodies are physically associated with the kinetoplasts, thus the separation of the basal bodies also results in separation of the kinetoplasts (McKean, 2003). Prior to cytokinesis, two new pro-basal bodies are generated, which are inherited along with the mature basal body, preparing the daughter cells for further proliferation once cytokinesis is completed (Sherwin and Gull, 1989).

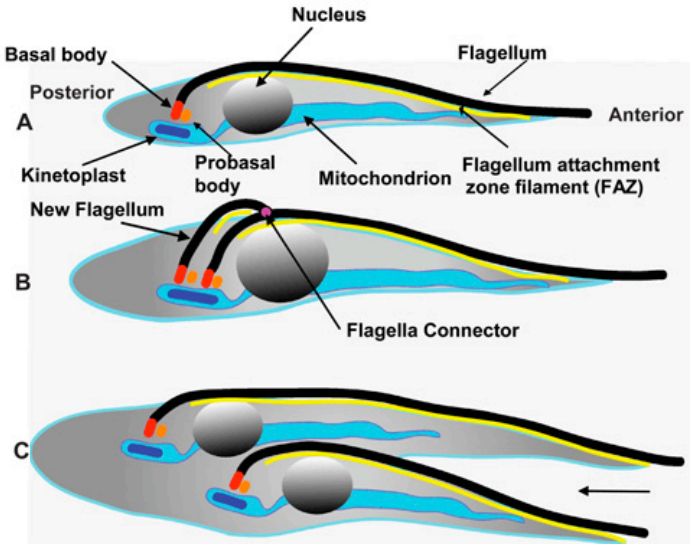



**Figure 1.3. Diagram of procyclic form cell cycle progression**. (A) A cell entering the cell cycle has a single copy of the nucleus, basal body, nucleus and kinetoplast. (B) In early S phase a new flagellum grows from the replicated basal body. (C) Cytokinesis progresses from the anterior to the posterior of the cell (arrow). (Vaughan and Gull, 2008)

### 1.4.1 DNA replication

Kinetoplast DNA is arranged into approximately 50 maxicircles over 20kb in length, mostly comprising genes coding for rRNA and proteins required for respiration, as is the case for the mitochondrial genomes of most organisms (Liu et al., 2005). Separate to the protein coding sequences found on the maxicircles, thousands of minicircles contain the sequences for guide RNAs, these being used to allow editing of the transcripts generated from the maxicircles (Liu et al., 2005). These minicircles are all interlocked, and so must be released and move towards to mitochondrial membrane, where the minicircle replication machinery is situated in order to replicate. This contrasts with the maxicircles which are not released from the network during replication (Liu et al., 2005).

The nuclear DNA is arranged into 11 megabase chromosomes and a large number of minichromosomes, the telomeric regions of which appear to harbour VSGs with the remainder of the sequence consisting of tandem repeats (El-Sayed et al., 2000). A

final DNA element in *T. brucei* is the intermediate chromosomes which do not appear to have any necessary genes, but which appear to harbour VSG genes. The number of these chromosomes does not seem to be strictly important, with different stocks of cells showing varying numbers of intermediate chromosomes (El-Sayed et al., 2000). Replication of all chromosomes initiates shortly after that of the kinetoplast, but takes longer, not being completed until half way through kinetoplast G2 phase (Figure 1.4).

Once all the cellular contents are duplicated and arranged appropriately, a cleavage furrow appears between the two nuclei and basal bodies (Sherwin and Gull, 1989). Cytokinesis is initiated from the anterior of the cell (Figure 1.3C), following the cleavage furrow and takes approximately 0.1 unit of the cell cycle to complete in procyclic form cells (Figure 1.4), but is much faster in bloodstream form cells (Hammarton, 2007).

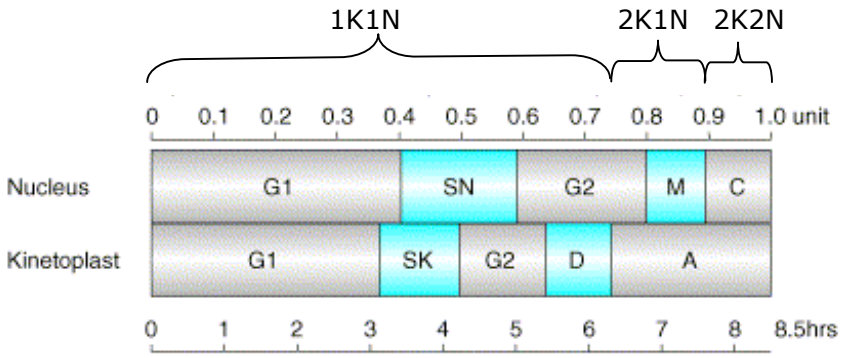



**Figure 1.4. Schematic diagram of sequential events of the cell cycle in procyclic *T. brucei*.** The diagram shows the time intervals for G1, S, and G2 phases of the cell cycle for both the nuclear and kinetoplast DNA. The 'D' phase of kinetoplast replication is the point at which the daughter kinetoplast segregates and is followed by the 'A' phase; the apportioning phase, which encompasses the time that the basal bodies continue to move apart. Mitosis of the nuclear DNA does not occur until part way through the apportioning phase of the kinetoplast division and is followed by cytokinesis (C). Due to the staggering of the duplication of the two types of DNA, visualising of the DNA of the cell during the cell cycle allows identification of cells with different karyotypes. The 1K1N arrangement is seen for the majority of the cell cycle with the 2K1N karyotype present for 0.15 units of replication and 2K2N for the remaining 0.1 unit (Modified from McKean, 2003).

As can be seen, the duplication of the organelles and DNA is tightly linked, however treatment with a number of drugs, and mutations of a range of cell cycle related genes have been found to uncouple cytokinesis from mitosis (Das et al., 1994). This is not possible in other eukaryotes due to the presence of a checkpoint prior to cytokinesis, which is absent in *T. brucei* (Ploubidou et al., 1999).

### *1.5 Differentiation and cell cycle position*

Cell cycle position appears to be intimately linked with differentiation of the parasites; in both the mammalian and insect host, cells preparing for transmission into the next host are arrested in G0/G1 and only initiate differentiation and cell cycle re-entry upon successful passage into the next environment. Stumpy forms are arrested in G0/G1, allowing them to proceed through differentiation in a synchronous manner, although the exact role of the cell cycle in differentiation is unclear (Li et

al., 2003, Matthews and Gull, 1994, Mutomba and Wang, 1996). Some doubts have been cast on the requirement of G0 arrest for differentiation, as cells stimulated to arrest in different stages of the cell cycle are seen to differentiate with equal efficiency (Li et al., 2003, Mutomba and Wang, 1996). However the methods used to arrest the cells, namely treatment with toxic doses of aphidicolin or hydroxyurea (Mutomba and Wang, 1996) and depletion of subunits of the proteosome (Li et al., 2003), were invariably lethal. Hence the results must be interpreted with care, as dying cells are known to undergo some differentiation events (eg mitochondrial activation, surface antigen exchange) as a default pathway. For example, Mutomba and Wang (1996) used aphidicolin or high doses of hydroxyurea to attempt to synchronise cells before mitosis or before cytokinesis respectively, before inducing the cells to differentiate with cis aconitate addition and temperature reduction. The authors saw no difference between cells in different cell cycle stages and their ability to differentiate. However, the cells had been incubated with the drugs for 12 hours and so were severely stressed, and hence clear conclusions cannot be reached from these experiments (Mutomba and Wang, 1996).

Bass and Wang (1991) also conducted differentiation experiments to demonstrate that slender, intermediate and stumpy forms were all able to differentiate in vitro to procyclic form cells with the same kinetics, concluding that stumpy forms were not required for successful differentiation. However, the classification in this study that a cell was differentiated relied solely on morphological traits rather than using other markers such as surface proteins or cell cycle position, and thus accurate differentiation kinetics could not be determined from this experiment (Bass and Wang, 1991). In contrast to these studies, evidence also exists to show that stumpy cells initiated to differentiate proceed through the cell cycle synchronously, suggesting that the cells must first be primed at G1 (Matthews and Gull, 1994). Supporting this, tracking the re-entry of stumpy form cells into the cell cycle by uptake of BrdU during differentiation revealed that most cells initiate S phase between 6-8 hours after initiation of differentiation (Matthews and Gull, 1994). Slender form cells, in contrast, differentiate in an inefficient, asynchronous fashion,

and tracking of re-entry into the cell cycle of these cells showed that their BrdU incorporation occurred with the same kinetics as differentiation, suggesting that those cells that initiate differentiation, do so in a cell cycle specific fashion (Matthews and Gull, 1994). The authors designed a model to demonstrate the role of cell cycle position with capacity to differentiate (Figure 1.5). This shows the possible outcomes for bloodstream form cells upon reception of the differentiation signal; stumpy cells (A) are prepared for differentiation, already arrested in G0/G1 and in response to the differentiation signal, they differentiate readily. For the slender cells in the population, three remaining options remain, (B) those cells at the 'receptive' point in the cell cycle, probably G1, are able to differentiate, however those not at the receptive window (C) can either continue progressing through the cell cycle until reaching the receptive window, where upon they may differentiate. Alternatively the cells will arrest in response to the unfavourable conditions and die (Matthews and Gull, 1994). This theory supports previous observations from differentiation experiments conducted on monomorphic cell lines, which were shown to differentiate into procyclic forms in an asynchronous manner (Ziegelbauer et al., 1990), suggesting that these cells must first reach the optimum point in the cell cycle before they are able to initiate differentiation.

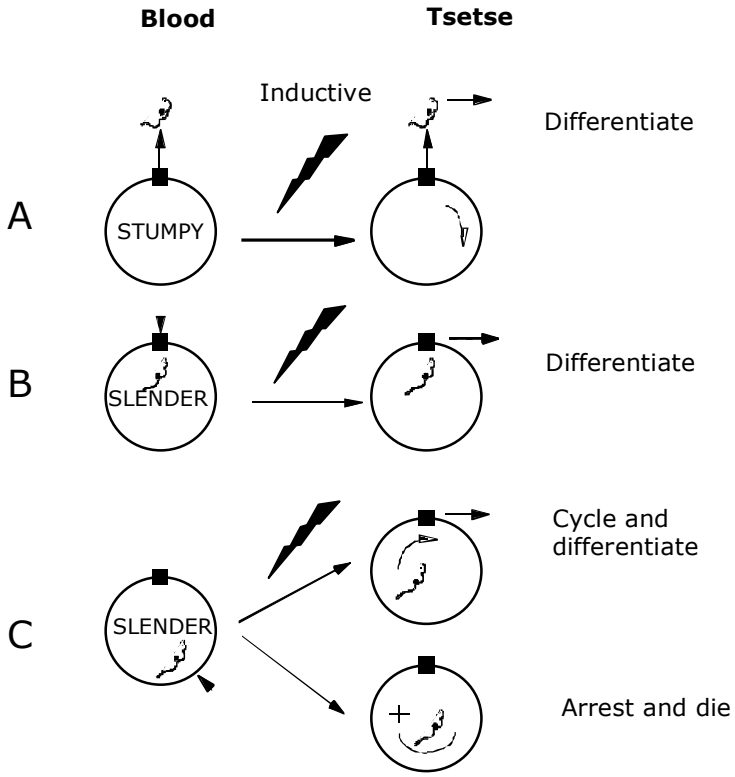



**Figure 1.5. Model for initiation of differentiation**. Arrowheads denote position of a trypanosome in the cell cycle in relation to the window in which cells are receptive to differentiation (filled box). (A) Stumpy cells are G0/G1 arrested and committed to differentiation and thus in the receptive window. (B) Slender cells in the receptive window will also differentiate. (C) Slender cells not in the receptive window may either progress through the remainder of the cell cycle and differentiate, or arrest and die (Matthews and Gull, 1994).

However, Matthews and Gull (1994) then went on to show that returning to a proliferative status does not need to occur in tandem with other features of differentiation. Thus, treating stumpy form cells with aphidicolin in order to block DNA synthesis had no effect on differentiation, despite all cells having arrested before S phase (Matthews and Gull, 1994). Similarly, inhibition of cell cycle progression during differentiation of monomorphic bloodstream form cells also confirmed that growth arrested cells were able to differentiate into morphological procyclic form but were unable to resume proliferation (Mutomba and Wang, 1998). Further evidence came from studies inducing stumpy cells to differentiate at 37˚C,

which caused cells to exchange antigenic coats as expected but not to re-enter the cell cycle (Matthews and Gull, 1997).

Further evidence concerning the role of G0 arrest in differentiation was provided from a study looking at a protein tyrosine phosphatase (TbPTP1), which was required to maintain G0 arrest and whose removal resulted in the spontaneous differentiation of stumpy form cells (Szoor et al., 2006). Strikingly, a subset of slender form cells could also be triggered to differentiate into the procyclic form upon depletion of TbPTP1, and it was proposed that these cells had already undergone cell cycle arrest, and were thus primed for differentiation (Szoor et al., 2006). Taking into account the evidence from all published experiments, it appears as though, whilst G0/G1 arrest appears to be beneficial for stumpy form cells in order to progress through differentiation in a synchronous manner, it is not certain that a specific cell cycle position is absolutely required for differentiation.

## *1.6 Cell cycle synchronisation*

If a method for synchronisation of *T. brucei* populations existed, the cell cycle of the parasite would be able to be studied in much greater detail, and debates over the role of cell cycle position in differentiation could be resolved. However trypanosomes have proved difficult to synchronise using standard techniques to arrest cells such as chemicals or serum starvation.

### 1.6.1 Block-release induced synchronisation

Chemicals used to synchronise populations in other organisms have been trialled on *T. brucei* cultures, including hydroxyurea (HU) and aphidicolin (APH), which cause cell cycle arrest through preventing synthesis of dNTPs and inhibiting the action of DNA polymerase $\alpha$ respectively. Historically these methods were not routinely utilised for synchronising trypanosomes as the drugs proved toxic to the cells (Mutomba and Wang, 1996), however recent work has optimised the dosage and length of treatment with these drugs to find concentrations low enough to arrest cell growth without lethality (Chowdhury et al., 2008, Forsythe et al., 2008). The optimum dose to arrest bloodstream form cells was 10$\mu$M HU, with cells completely

arrested after nine hours, although incubation for longer than this was lethal (Forsythe et al., 2008). If cells were removed from the drug after six hours it was possible to get a population comprising 75% cells with the 2K1N configuration, which, if the size and shape of the kinetoplasts of the 1K1N cells were examined, could be classed as S phase cells for <90% of the cells (Forsythe et al., 2008). When cells were released from the drug, 35% of cells were found to be in the 2K2N configuration three hours after release. This relative synchrony was not maintained however, and so this method was not suitable for isolating pure populations of G1 stage cells.

Serum starvation has also been attempted with varying success; whilst partial synchrony as been recorded previously (Gale et al., 1994a, Gale et al., 1994b), this method proved hard to reproduce, at least until recently optimised (Archer et al., 2009). In this approach, procyclic form cells were grown for two days until the density of the population was between 2-4x$10^7$cells/ml. At this point the media was depleted of nutrients, such that the cells were effectively starved with population growth having reached a plateau with the majority of the population stalled in G1 (Archer et al., 2009). Dilution of these cells into fresh media then allowed the cells to resume proliferation in a semi-synchronous manner. The limitation of this method of synchronisation, however, is that it is confined to procyclic cells alone, as bloodstream forms are extremely sensitive to serum deprivation.

All of the above methodologies are intrinsically flawed since artificially inducing cell cycle arrest does not necessarily inhibit cell growth (Cooper, 2003). Hence, cells with incomplete DNA replication will continue expanding cellular mass during their cell cycle inhibition and upon release, all cells will progress forward into the next cell cycle size regardless of cell size. Abundance of cell cycle regulated transcripts was shown to be uncoupled from DNA replication in human cells arrested with a variety of cell cycle inhibitors, representing a perturbation of the normal relationship between cell growth and cell division (Urbani et al., 1995). Considering also the toxicity of the chemical inhibitors of cell cycle, even using low doses renders it

difficult to interpret results yielded from these cells, as obstruction of the cell cycle will invariably stress the cells, having downstream effects on a variety of cellular activities including gene expression. Numerous microarrays have been launched to find cell cycle regulated genes in yeast cells synchronised in different ways, yet scrutiny of the results shows that the cells were not truly synchronous, and few genes demonstrating cell cycle differential expression overlapped between the different synchronisation methods (Cooper and Shedden, 2003).

### 1.6.2 Sorting cells based upon cell cycle position

Preferably cells should be sorted according to the cell cycle stage they are naturally in, using a 'selective' method such as elutriation where cells are separated according to size, rather than a batch 'block and release' method (Cooper, 2004) such as HU treatment. Better still is a situation where newly separated daughter cells can be isolated after cytokinesis when all cells are the same age. This process is possible in bacterial and eukaryotic cell cultures through the development of a 'baby machine' where cells are adhered to a membrane with a constant stream of media pumped from above so that nascent cells are washed off as they separate from the parent cell, thus allowing in-depth analysis of the unperturbed cell cycle (Helmstetter et al., 2003). Such a set-up for bloodstream form trypanosomes is unlikely to work however, as cells must be motile for cytokinesis to occur (Ralston et al., 2006, Monnerat et al., 2009); thus adherence to a membrane would be likely to prevent cell division.

Sorting trypanosomes based upon DNA content has been used to generate particularly pure populations of G1, S phase and G2/M cells through staining with propidium iodide. However, this is toxic to cells and so can only be used for the examination of fixed cells (Gale et al., 1994a). More recently, supravital dyes have been used to stain the DNA of living cells to allow sorting of viable cells based upon their DNA content (Bradford et al., 2006). Such techniques have been used to sort live procyclic form trypanosomes, although no attempts have been reported to culture the sorted populations or to use this method on bloodstream form cells (Siegel et al., 2008a, Michael Boshart, personal communication). Another method

for sorting cells in specific stages of the cell cycle is to fluorescently tag a known cell cycle regulated gene, making it possible to look at cells with minimal interference with the cell (Coquelle et al., 2006). A number of cell cycle regulated genes are characterised in *T. brucei*, so such a method could be utilised (Archer et al., 2009) (Siegel et al., 2009), (Gale et al., 1994a). The drawback to this method is that a tagged cell line must be produced and the tag must be demonstrated to not detrimentally affect the cell.

## *1.7 Proteins involved in G0/G1 progression*

All stumpy forms are G0 arrested, thus tight regulation must exist to direct slender forms into this quiescent phase and also to maintain this until such time as the cells are taken into the tsetse midgut and can complete differentiation. Surprisingly, no receptor-linked kinases or G protein coupled receptors have been reported in *T. brucei*, so unlike in other organisms where these pathways are major players in G1 progression, other mechanisms must exist for G1 control in trypanosomes (Hammarton, 2007). Previously, a protein homologous to the human PTP1B protein tyrosine phosphatase, which is involved in quiescence in mammalian cells, was shown to be required for inhibition of differentiation in *T. brucei* (Szoor et al., 2006). However, the mechanism causing the activation of the protein to arrest stumpy forms is unknown. Nonetheless, many different proteins are known to function in regulation of G0/G1 in other organisms, two of which include the Target of Rapamycin (Tor) and the APC/C activator Cdh1.

### 1.7.1 Target of Rapamycin

*S. cerevisiae* cells are sensitive to the antifungal agent rapamycin (Heitman et al., 1991). However, spontaneously arising cells were found that were resistant to rapamycin, and further investigations of these resistant cell lines identified mutations in either of two genes that were subsequently named the Target of Rapamycin genes: Tor1 and Tor2 (Koltin et al., 1991). Further characterisation of these two genes showed that they were highly homologous, and investigations in other organisms showed that Tor genes were evolutionarily highly conserved and have subsequently been identified in every organism studied to date with the notable exception of *Plasmodium* ssp. However, yeast were the only organisms reported to date to have

two Tor genes which form distinct complexes (TORC1 and TORC2) and function in different pathways to control a vast array of cellular events. Interestingly, TORC2 only ever contains Tor2, but TORC1 can be composed of either Tor1 or Tor2 (Loewith et al., 2002). Apart from fission and budding yeast, all other organisms have a single Tor gene coding for a protein that is able to function in both TORC1 and TORC2 (Figure 1.6A). The mammalian homologue to Tor, mTor, which shows most identity to yeast Tor2 (Sabatini et al., 1994), is ubiquitously expressed throughout all cells of the body, but is particularly high in skeletal muscle (Loewith et al., 2002, Jacinto et al., 2004). Raptor, the binding partner of mTor, is also highly conserved throughout the eukaryotic lineage and is hypothesised to act as a scaffold protein to facilitate the association of mTor with its substrates (Inoki et al., 2005). The interaction with mTor is dependent on a number of HEAT repeats and WD40 domains in raptor, and deletion series of the protein showed that the majority of the protein was required for binding (Kim et al., 2002). However, there is also evidence showing that raptor can increase the association with mTor under conditions of high nutrients, with the higher association resulting in an inhibition of mTor kinase activity (Kim et al., 2002). Thus, it is possible that raptor may act as a regulator of Tor activity. The LST8 protein is required for both The TORC1 and TORC2 complexes, and is a remarkably simple protein, consisting almost entirely of seven WD40 repeats (Loewith et al., 2002).

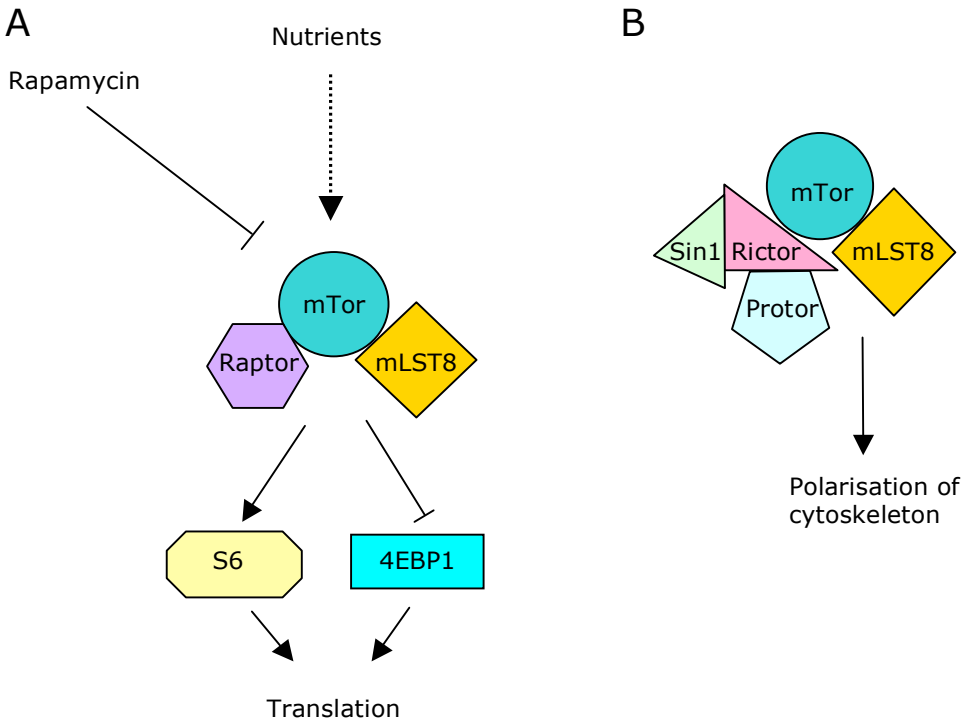



**Figure 1.6. Diagram of the binding partners and pathways of mTor in the TORC1 and TORC2 complexes**. Mammalian Tor and LST8 are components of both of the Tor complexes, but the other binding partners differ. (A) TORC1 also consists of a potential regulatory protein raptor, and this complex signals to regulate translation in response to nutrient availability and is sensitive to rapamycin. (B) TORC2 also contains Sin1, Rictor and Protor, is involved in polarisation of the actin cytoskeleton and is insensitive to rapamycin inhibition.

Rapamycin treatment or Tor depletion in budding yeast causes cells to undergo highly similar morphological and cytological changes as are seen under starvation conditions (Koltin et al., 1991). When investigated in more depth, it became clear that Tor was required for the regulation of entry in quiescence; rapamycin treated cells arrested division in early G1 and gene expression of known G0-regulated transcripts matched that for quiescent cells (Barbet et al., 1996). Tor also appears to be involved in the regulation of autophagy, via a number of different mechanisms (Diaz-Troya et al., 2008). In higher eukaryotes, TORC1 is known to directly interact with members of the ATG family of autophagy genes, and also regulates transcription factors responsible for the initiation of transcription of genes required for autophagy (Chan et al., 2001, Chang and Neufeld, 2009).

### 1.7.1.1 Structure of Tor proteins

Tor proteins are identifiable through the presence of a rapamycin binding domain (FRB), but also possess several more general domains of PI3 kinases such as a PI3 PI4 kinase domain, a FAT domain and, at the extreme C terminus, a FATC domain (Figure 1.7). Despite the PI3 kinase domain in the protein, the presence of the FAT and FATC domains and the serine/threonine activity of Tor, place this protein into the phosphatidlyinositol kinase-like kinase family (PIKK) (Hay and Sonenberg, 2004). This C terminal domain appears to be the most conserved, with overall conservation gradually increasing from the N terminus, which harbours the HEAT repeats, down to the FATC domain (Sabatini et al., 1994). FATC domains are unique to PIKKs, and the final 31 amino acids of mTor are highly conserved with the other human PIKKs (Takahashi et al., 2000). The FATC domain appears to be involved in regulation of mTor activity, as deleting the three cysteine residues within the domain, thus abolishing the disulphide bonds required for the structural integrity of the FATC domain, results in loss of kinase activity of mTor (Takahashi et al., 2000). HEAT repeats are involved in protein-protein interactions, so would be expected to vary between organisms, and these domains appear to be required for the interaction with the components of the Torc1 and 2 complexes (Kim et al., 2002, Sabatini et al., 1994).

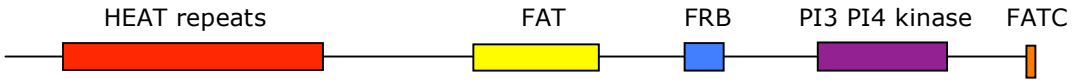



**Figure 1.7. Schematic diagram showing the architecture of a TOR protein.** The HEAT repeats are required for Tor association with the remaining proteins on the TORC complexes. The FRB domain is required for binding rapamycin. The FATC domain may be involved in regulation of the protein through folding to interact with the FAT domain and the PI3/PI4 kinase domain is the catalytic part of the protein.

New roles for the Tor proteins are discovered constantly, such as the evidence showing that TORC1 may directly interact with ribosomal DNA promoters in *S. cerevisiae* (Li et al., 2006). This function for Tor requires the protein to be localised to the nucleus, and a nuclear localisation sequence was discovered towards the N

terminus. The nuclear localisation of Tor was lost during starvation or rapamycin treatment, when the protein was directed to the cytoplasm via a nuclear exporting sequence, thus preventing activation of ribosomal gene transcription (Li et al., 2006). Treatment with rapamycin or nutrient depletion are also known to affect rRNA transcription, which is conducted by polymerase III, but has little effect of tRNA transcription despite these gene groups both being transcribed by the same polymerase (Wei et al., 2009). The mechanism for this is via TORC1 phosphorylation of the polymerase III negative regulator, Maf1, which localises to the cytoplasm in its phosphorylated form but, upon depletion of TORC1, will move to the nucleolus and inhibit transcription of rDNA (Wei et al., 2009).

Tor appears to control one of the pathways required for the regulation of translation. Two major substrates of Tor are the 4E-BP1 and S6 kinase (S6K), and through these proteins, a variety of translational activities can be controlled (Hay and Sonenberg, 2004). Both of these target proteins contain TOS (target of rapamycin signalling) motifs, without which they are incapable of phosphorylation by Tor (Inoki et al., 2005). The role of 4E-BP1 is to inhibit the action of the eukaryotic translation initiation factor 4E (eIF4E), which forms a complex with several other initiation factors to form eIF4F. 4E-BP competitively binds to one of these proteins, preventing the formation of the active complex, thus inhibiting the activity of the translation initiation factors and so inhibiting translation (Inoki et al., 2005). TORC1 acts to phosphorylate 4E-BP1 in response to nutrient levels, and upon phosphorylation of the substrate, 4E-BP1 is no longer able to bind to eIF4E and so it can instead form the active eIF4F complex which is then functional to recruit the translational machinery to mRNAs (Raught et al., 2001). Certain transcripts have been seen to be particularly sensitive to rapamycin, and analysis of the 5' UTRs of these transcripts has shown them to be unusually long and having complicated structures, a feature that obliges a higher requirement for eIF4F (Gingras et al., 2001). The second major substrate of mTor, S6K, is required for the translation of mRNAs containing a 5' Tract Of Pyrimidine (TOP) structure; these primarily comprising transcripts for the translational machinery and ribosomes (Raught et al.,

2001). Tor phosphorylation of S6K results in its activation, thereby increasing the translation of proteins required for translation and thus increasing the capacity of the cell for translation (Inoki et al., 2005). Both S6K and 4E-BP are also targeted by PI3K in a separate regulatory pathway, with PI3K and Tor phosphorylating different residues in the substrates (Inoki et al., 2005).

Mounting evidence has shown a role for Tor in G1 progression in a variety of organisms. Originally characterised in *S. cerevisiae*, treatment with rapamycin caused cells to arrest as unbudded cells, presumably in G1 arrest (Heitman et al., 1991). In *S. pombe*, nutritional depletion causes cells to arrest in G1 and then to undergo sexual differentiation; a mechanism which appeared to be Tor-regulated, as Tor2 knock down resulted in a similar effect, with cells becoming smaller and entering quiescence, which was fully reversible (Alvarez and Moreno, 2006). Further characterisation of Tor2 in fission yeast demonstrated that Tor2 mutants showed overexpression of genes required in nitrogen starvation, but no difference was seen in genes activated in response to glucose starvation, thus Tor2 is specifically regulated by nitrogen availability (Matsuo et al., 2007). Closer examination of the amino acid regulation of mTor in human cell lines was able to narrow down the substrates for regulation and found that mTor was specifically activated in response to glutamine availability (Nicklin et al., 2009). It has also been shown that microinjection of the FRB domain of mTor into human cells causes a G1 specific cell cycle arrest (Vilella-Bach et al., 1999). Disruption of Tor in *C. elegans* causes cells to arrest in division, although these mutant nematodes did not differentiate to the duaer form, as might be expected considering that this is the natural arrested form that spontaneously arises under conditions of nutrient limitation (Long et al., 2002). Therefore, CeTor is involved in cell cycle progression, but causes arrest in a mode distinct from that occurring under nitrogen starvation.

### 1.7.1.2 TORC2

The function of the TORC2 complex was discovered to be in the regulation of the polarisation of the actin cytoskeleton, and is under different regulation to TORC1 as

the kinase activity of TORC2 showed no effect in response to serum deprivation or amino acid starvation (Jacinto et al., 2004). The Tor binding partner in the TORC2 complex was named rictor in mammals (rapamycin insensitive companion of Tor) due to the fact that this complex is not disrupted in response to rapamycin (Sarbassov et al., 2004). Another component of TORC2 is LST8, as in TORC1, but two additional proteins are also required; Sin1 and Protor (Figure 1.6B) (Pearce et al., 2007). Interestingly, the *S. cerevisiae* and *S. pombe* orthologues of rictor do not show a high level of conservation with mammalian rictor, in contrast to all the other members of the Tor pathways, with the region of highest homology confined to a 200 residue region within which the identity reaches 8% (Sarbassov et al., 2004). Depletion of the TORC2 complex in human cell lines gave rise to misshapen cells due to the aberrant formation of thick actin fibres (Sarbassov et al., 2004). Signalling to the actin cytoskeleton occurs via PKA, however other organisms using the PKA signalling pathway for other functions have evolved TORC2 functions specialised for different pathways. For example, the slime mould *Dictyostelium* uses cAMP signalling to regulate chemotaxis and analysis on cell lines ablated for components of the TORC2 complex has shown that this complex is directly required for this signalling pathway (Lee et al., 2005).

### 1.7.1.3 Rapamycin

Rapamycin is produced by the soil-dwelling bacterium *Streptomyces hygroscopicus*, (Arndt et al., 1999) and was originally used medically as an anti-fungal agent. Upon detection of its anti-proliferation functions, however, rapamycin is now more commonly used as an immunosuppressive drug for transplant patients, and is being increasingly tested for treating a range of cancers (Ballou and Lin, 2008). Three rapamycin analogues have since been developed for human use, one of which has been modified to be water-soluble in order to aid administration to patients (Ballou and Lin, 2008). A recent application of rapamycin has been in extending the life span of mice, where treatment with the drug late on in life increased life span by up to 14% (Harrison et al., 2009). The method of action of the rapamycin appeared to be from delaying onset of cell deterioration related to ageing, such as cellular stress (Harrison et al., 2009).

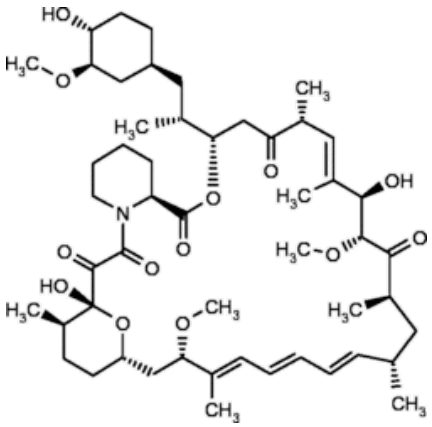


**Figure 1.8. Structure of rapamycin**. The anti-fungal drug rapamycin binds to the FRB domain of Tor only when bound to FKBP12.

Rapamycin is unable to bind Tor directly; and mutational studies in *S. cerevisiae* identified the ligand FKBP12 as a protein that, when mutated, conferred rapamycin resistance (Koltin et al., 1991). FKBP12 is a peptidyl-prolyl cis-trans isomerase protein and the usual function of this protein the cell is in the formation of a calcium channel (Gothel and Marahiel, 1999). However, this protein is also responsible for rapamycin binding, and deletion of FKBP12 from a rapamycin sensitive cell is sufficient to confer complete rapamycin resistance (Kunz et al., 1993, Crespo et al., 2005). FKBP12 is extremely well conserved, showing 100% identity between human and rabbit orthologues and, in all organisms showing sensitivity to rapamycin, FKBP12 has been shown to bind to rapamycin and in this conformation interact with Tor at the rapamycin binding domain (FRB), thus inhibiting the activity of TORC1. Extensive characterisation of the FRB domain has been carried out in yeast to identify the residues essential for rapamycin binding, and three highly conserved residues (serine 1975, tryptophan 2042 and phenylalanine 2049) were found, that when mutated abolished rapamycin sensitivity (Stan et al., 1994, Lorenz and Heitman, 1995). The crystal structure of the interaction between rapamycin-FKBP12-mTor has been solved and shows that the residues identified as essential for rapamycin binding all lie at the FRB-FKBP12-rapamycin interface (Choi et al., 1996). The same residues required for rapamycin binding to yeast Tors are conserved for mTor, and substitution assays of the corresponding serine residue of

mTor (2035) (Figure 1.9) showed that only replacement with an alanine residue prevented rapamycin resistance (Chen et al., 1995).

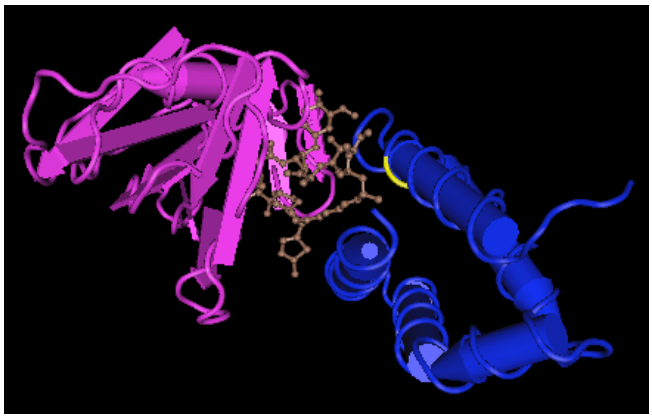


**Figure 1.9. The structure of human FKBP12-rapamycin complex interacting with mTor**. The blue structure is the rapamycin binding domain of mTor, the pink structure is human FKBP12 and the brown structure rapamycin. Highlighted in yellow is the conserved serine residue of mTor (serine 2035, correlating to serine 1975 of *S. cerevisiae* Tor2) at the mTor-rapamycin interface (Choi et al., 1996).

Rapamycin has been shown to specifically inhibit the activity of the TORC1 complex in a wide variety of organisms, from mammals, to slime mould, plants, and yeast (Loewith et al., 2002, Cutler et al., 2001, Kim et al., 2002, Crespo et al., 2005, Heitman et al., 1991), but sensitivity to rapamycin is not universal across all life forms. Indeed *C. elegans* shows no effects upon treatment with the drug, despite having a highly conserved FKBP12 receptor and the FRB domain of CeTor showing high levels of identity to human and yeast Tor, importantly possessing all the essential residues including the conserved serine (Long et al., 2002). Thus, other mechanisms apart from mutations in the FRB domain exist to prevent sensitivity to rapamycin. The authors speculate that the nematodes may have evolved a means to prevent uptake of rapamycin, possibly in response to the presence of this chemical in the soil – the natural habitat of these organisms (Long et al., 2002). The helminth parasite *Schistosoma mansoni* has also been found to be resistant to rapamycin, although the reason for resistance was not investigated (Rossi et al., 2002).

### 1.7.1.4 TbTOR

During the course of this PhD, the Tor pathway in *T. brucei* was independently studied by a number other groups and extensively characterised (Barquilla et al., 2008, Barquilla and Navarro, 2009b, Barquilla and Navarro, 2009a, Monnerat et al., 2009). The results of these studies were surprising and some findings were in direct contrast to preliminary data from our lab, so certain aspects of the Barquilla et al study (2008) were re-examined in this thesis. Interestingly, the authors described four Tor orthologues in *T. brucei,* comprising clear Tor1 and Tor2 homologues with an additional two Tor-like proteins, with TbTor1 exclusively present in TORC1 and TbTor2 uniformly present in TORC2 (Barquilla et al., 2008). Closer scrutiny however, suggests that the Tor-like 2 protein is not a member of the Tor family but instead a member of the more general PIKK family as it lacks the rapamycin binding domain (FRB) found in all Tor proteins so far identified, which is present even in the gene coding for the rapamycin insensitive Tor homologue in *C. elegans*.

Incubation of bloodstream form *T. brucei* cells with rapamycin results in specific inhibition of the TORC2 pathway (Barquilla et al., 2008), in contrast to all other organisms where TORC2 is invariably insensitive to rapamycin action. Interestingly, the rapamycin-TbFKBP12complex is unable to bind to TbTor2 whilst it is associated with the TbTORC2 complex, but does associate with free TbTor2, suggesting that the phenotype seen is the result of depletion of available TbTor2 (Barquilla et al., 2008). This suggestion appears validated by the equivalent phenotype of cells when RNAi is directed against TbTor, both regimes resulting in cells with multiple nuclei and flagella, with hugely enlarged flagella pockets (Barquilla et al., 2008). This phenotype stems from the role of TbTORC2 in polarisation of the actin cytoskeleton, in keeping with the function in yeast. The function of TbTor1 also mirrors that for in other organisms, with depletion in *T. brucei* resulting in G1 arrest and a decrease in cell size and protein synthesis (Barquilla et al., 2008). TbTor1 depletion in procyclic form cells has been shown to result in deregulation of the cell cycle, with disappearance of post mitotic cells and accumulation of cells with abnormal numbers of nuclei and kinetoplasts, ultimately resulting in death (Monnerat et al., 2009).

Results generated in our lab contradict some of these findings, and will be described in Chapter 5.

Rapamycin treatment of *T. brucei* bloodstream form cells was shown to cause a reduction in cell number; the dose causing a 50% decrease in cell number was calculated as 152nM. A homologue of Tor in the closely related protozoan parasite *Leishmania* species has not been investigated, however treatment of *Leishmania major* with rapamycin has been shown to kill these parasites with an IC50 of 4.9μM (Madeira da Silva et al., 2009), which is surprisingly different to the concentration eliciting the same response in *T. brucei*, considering how similar the Tor proteins are between the species. Interestingly, at lower doses of rapamycin, the treated *Leishmania major* cells appear to have better survival rates in comparison to untreated cells (Madeira da Silva et al., 2009). This phenotype of Tor inhibition resulting in extended lifespan has similarities to the depletion of TbTor1 which, prior to cell cycle arrest and death, increases the cells resistance to stress caused from growing to high density (Barquilla and Navarro, 2009a). In addition, visualisation of TbTORC1 depleted cells shows certain characteristics resembling autophagy, such as the appearance of double-membraned vacuoles, suggesting that one of the roles of TbTORC1 is in the regulation of autophagy, a role that has been suggested in higher eukaryotes (Diaz-Troya et al., 2008, Barquilla et al., 2008, Barquilla and Navarro, 2009a).

### 1.7.2 Cdh1

The Anaphase Promoting Complex/Cyclosome (APC/C) plays a major role in cell cycle progression through controlling the degradation of cell cycle regulatory proteins and is highly conserved (Li and Zhang, 2009). The part played by the APC/C in degradation is through catalysing the addition of ubiquitin molecules onto targets and thus labelling them for proteolysis (Peters, 2006). The human APC/C is composed of 12 different subunits, although the exact arrangement of these subunits is unknown; all that is certain is that the combined molecular weight of the APC/C is too high for each subunit to be present only once (Dube et al., 2005). Further, the multiple appearance of the largest subunits can be rejected based on the high weights

of these proteins, meaning that there must be more than a single copy of at least one of the smaller subunits (Dube et al., 2005). Nevertheless, these 12 subunits only represent the core of the complex, with different proteins recruited to confer different substrate recognition at transition points of the cell cycle (Peters, 2006). These recruited proteins are termed activator proteins and these are required to direct the APC/C to destroy only the specific targets at the different stages of the cell cycle, with the Cdc20 substrate required for the metaphase-anaphase transition and Cdh1 for regulation of G1 (Skaar and Pagano, 2008). A third APC/C activator, mfr1, has also been identified in yeast, which is only required in meiosis (Blanco et al., 2001), and two further meiosis specific co-activators of the APC/C exist in *Drosophila melanogaster* (Peters, 2006). In addition, four distinct homologues of Cdh1 were discovered in chicken, with each having a specific set of target substrates (Wan and Kirschner, 2001). A common feature of all these coactivators (but only the Cdh1-C homologue in chicken) is the presence of a C box and an IR tail at the extreme C terminus, required for interaction with the APC/C (Peters, 2006). Cdh1 was originally discovered in *Saccharomyces cerevisiae* in a screen for orthologues of Cdc20, and over expression of this gene was found to result in cell cycle arrest leading to lethality (Visintin et al., 1997) (Schwab et al., 1997). Shortly afterwards the homologue was discovered in *S. pombe* (Yamaguchi et al., 1997), with over expression resulting in growth arrest and cell lines depleted of Cdh1 were unable to arrest in G1 in response to nitrogen starvation as occurs in wildtype cells (Kitamura et al., 1998).

### 1.7.2.1 Structure of Cdh1

The Cdh1 protein has seven WD40 repeats that make up the large majority of the C terminus which form a propeller structure (Figure 1.10 B), with the 'outside' of the propeller containing the most evolutionarily conserved residues which are shown to be involved in protein interactions with the substrates (Kraft et al., 2005). Interestingly, Cdh1 also has two destruction boxes located towards the N terminus (Figure 1.10A) that on other proteins targets them as substrates for Cdh1-mediated degradation (Listovsky et al., 2004). In keeping with the other Cdh1 substrates, the function of the destruction boxes in Cdh1 was indeed found to be required for

ubiquitination of the protein, allowing Cdh1 levels to be regulated in a negative feedback loop such that high levels of the protein triggers degradation of free Cdh1 via the APC/C associated with Cdh1 (Listovsky et al., 2004). When active, Cdh1 associates with the APC/C within the nucleus, and a well-conserved nuclear localisation sequence has been identified (Zhou et al., 2003a). Cdh1 and Cdc20 also have a C box towards the N terminus that is not found in other proteins without WD40 domains, consisting of seven amino acids (Schwab et al., 2001). The final two residues of the protein form an IR motif, again present in both Cdc20 and Cdh1, which together with the C box, allows binding of Cdh1 to the APC/C (Burton et al., 2005).

Many of the motifs seen in Cdh1 are shared with those of Cdc20, however a region downstream of the C box was identified uniquely in Cdh1 homologues, and so named the Cdh1 specific motif (CSM) (Tarayre et al., 2004).

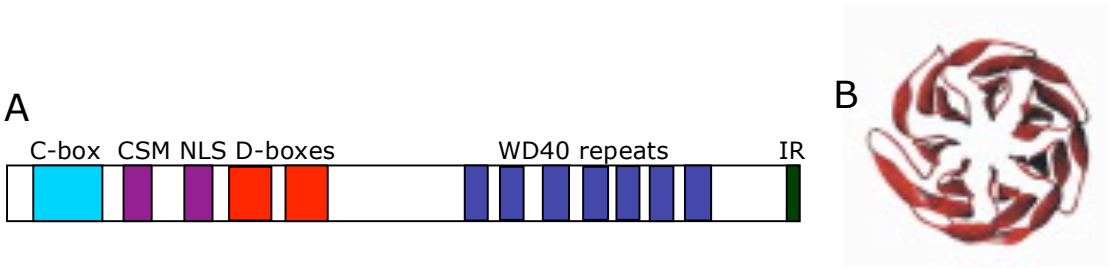



**Figure 1.10. Structure of Cdh1**. (A) Schematic showing the domains of Cdh1 with the N terminal C-box required for substrate recognition and C terminal IR motif for APC/C interaction. The NLS (nuclear localisation sequence) directs the protein into the nucleus where it is able to interact with the APC/C. Two destruction boxes are required for self-mediated degradation (Listovsky et al., 2004). The seven WD40 repeats are involved in protein interactions with the substrate and form the propeller structure shown in (B) Propeller structure of Cdh1 (Dube et al., 2005)

Regulation of Cdh1 is multifaceted as is often seen in proteins at the top of control pathways. The Cdh1 activity in higher eukaryotes, as determined by degradation of its substrates, is highest in G1 and in cells arrested in G0 through serum starvation or naturally quiescent in differentiated tissues (Brandeis and Hunt, 1996). The mRNA of both Cdh1 and Cdc20 show similar profiles, peaking during G2 and maintained until the loss of transcript early on in G1 phase (Fang et al., 1998). The protein levels of Cdc20 mirror that of the mRNA, but interestingly despite the fluctuations seen in the mRNA levels of Cdh1, the protein levels remain fairly stable (Fang et al., 1998), leading to the hypothesis that only newly synthesised protein is active due to phosphorylational inactivation of older protein (Inbal et al., 1999). However another, more plausible, possibility for protein regulation is the physical separation of the protein from its substrates. In human cells the action of Cdh1 has been shown to be repressed via sequestration to the cytoplasm through phosphorylation of the nuclear localisation signal (Zhou et al., 2003a, Huynh et al., 2009) resulting in Cdh1 localisation to the cytoplasm during the earlier parts of mitosis and during G2/M (Jaquenoud et al., 2002). Phosphorylation of Cdh1 also causes it to disassociate from APC/C thus resulting in its inactivation (Huynh et al., 2009). Additionally Cdh1 can autoubiquitinate and thus trigger its own destruction (Huynh et al., 2009). The phosphorylation also causes Cdh1 to bind to Msn5, which as it binds at the C box,

may help regulate activity of Cdh1 as this protein will be competitively binding to the same site as the substrates (Jaquenoud et al., 2002).

Other proteins are also involved in the regulation of Cdh1 activity, such as Emi1, a zinc-finger-containing protein, which appears to bind to Cdh1 and thus prevent the association with the APC/C (Harper et al., 2002). Interestingly, this protein is able to bind to and regulate both Cdh1 and Cdc20 in some organisms, but is specific for Cdh1 only in *Drosophila* (Pesin and Orr-Weaver, 2008).

Cdh1 associates with the APC/C via interactions with the Cdc27 subunit when in a low phosphorylation state (Huynh et al., 2009). More recently it has been shown that the activators Cdh1 and Cdc20 both bind, via different motifs, to the subunits Cdc27 and Cdc23, with the substrate binding to both the APC/C and the activator (Matyskiela and Morgan, 2009) (Figure 1.11). In contrast, the means by which Cdc20 activation of the APC/C is regulated appears to be via phosphorylation of the APC/C itself, specifically the Cdc27 subunit, although there is conflicting evidence relating to this (Blanco et al., 2000). The various targets of Cdh1 in the different cell cycle phases are not all degraded at the same time; certain proteins must be removed sooner than others, and the mechanism for this has been shown to be the mode of polyubiquitination of the substrate (Rape et al., 2006). Those proteins that are removed early on during activation of Cdh1 were shown to have the entire length of ubiquitin molecules attached immediately after the substrate was associated with the APC/C-Cdh1 complex, however other proteins removed later on would still form an interaction with the APC/C-Cdh1 complex, but only a single ubiquitin molecule was attached and the protein released, thus requiring the protein to return to the complex and gradually accumulate enough ubiquitins to ensure recognition by the proteosome (Rape et al., 2006).

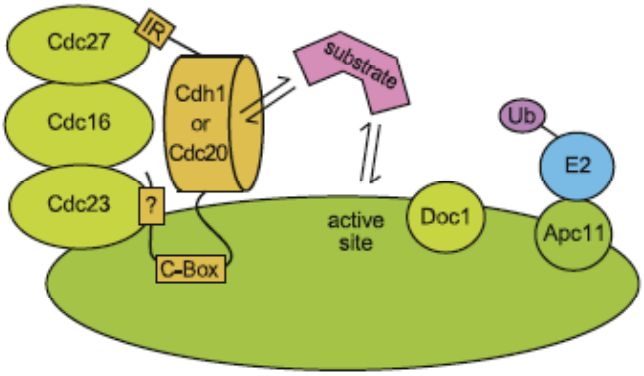



**Figure 1.11. Mode of Cdh1 binding to APC/C**. The activators Cdh1 and Cdc20 bind to the same region of APC/C via binding to the subunits Cdc23 and Cdc27 (from (Matyskiela and Morgan, 2009)).

New targets of the APC/C in eukaryotes are being found with impressive regularity, as are new regulatory mechanisms, so there is still much as yet unknown of the action of Cdh1, even in yeast.

### 1.7.2.2 APC/C in *T. brucei*

Homologues to seven of the APC/C subunits have been identified in *T. brucei*, using the protein sequences from *S. cerevisiae* to run pBLAST searches (Kumar and Wang, 2005). A Cdc20 homologue in *T. brucei* was also identified in this study, but a putative homologue of Cdh1 was not discovered, however closer inspection of the protein makes it difficult to necessarily designate it as Cdc20 rather than Cdh1; this will be discussed further in Chapter 5. The six subunits of the APC/C in *T. brucei* that were known to be essential in yeast were depleted in both procyclic form and monomorphic bloodstream form cells, but only Cdc27 and APC1 depletion resulted in growth defects, although the phenotypes seen differed between the two life cycle stages (Kumar and Wang, 2005). Whilst the effects of Cdc27 or APC1 depletion in procyclic form cells resulted in cells arrested in metaphase in common with the phenotype in *S. cerevisiae*, the same depletions in bloodstream form cells appeared to arrest the cells in late anaphase (Kumar and Wang, 2005). This result is maybe not as surprising as it first appears, considering the different roles of other cell cycle proteins between procyclic form and bloodstream form cells, the TbAPC/C could have different targets in the different life cycle stages and thus disruption of the

TbAPC/C would result in the accumulation of different cell cycle proteins. The similarities between the Cdc27 and APC1 of *T. brucei* with the yeast homologues were comparatively low considering the high level of conservation in the proteins involved in the APC/C pathway in other organisms, only 20% and 17% identity respectively, and neither were able to rescue null yeast cell lines (Kumar and Wang, 2005). Taken together it is clear that the APC/C in *T. brucei* is quite different from that in higher eukaryotes, but of particular interest is the presence of only one co-activator; Cdc20, so it will be interesting to discover how the substrates of TbAPC/C are phase specifically targeted for destruction.

## *1.8 Gene expression regulation*

A plethora of mechanisms exist for regulation of gene expression in kinetoplastids, which differ in many ways from that of higher eukaryotes in part because of the unusual method of transcription in these organisms.

### 1.8.1 Transcription in *T. brucei*

The genomes of the kinetoplastids are arranged into polycistronic units encompassing genes for both related and unrelated functions and timing of expression (Vanhamme and Pays, 1995). Hence, gene expression control in these organisms is largely post-transcriptional, with a variety of mechanisms in place to ensure tight regulation (Figure 1.12). Few promoters have been identified in the genome, although the ribosomal DNA promoter has been found and the promoters for the surface antigens VSG and EP are known, as are those for the metacyclic VSGs which are arranged into monocistronic units (Clayton, 2002). All these genes are transcribed by RNA PolI, which in higher eukaryotes is used exclusively to transcribe ribosomal RNA. It has been speculated that the reason for the novel use of PolI for transcription of the surface antigens is due to its higher processivity, making it able to transcribe these genes which are required in much higher abundance than other proteins in the cell, and yet have limited gene copy number (Clayton, 2002). The putative promoter sequences identified for PolI show a high level of homology to those in higher eukaryotes but have little sequence homology to each other (Palenchar and Bellofatto, 2006). The RNA polymerase II promoters by contrast, have proved more elusive, although it has been shown that the promoters for the SL

RNA genes are recognised by PolII (Gilinger and Bellofatto, 2001). Recent work has suggested that, rather than the method of transcription regulation seen in organisms where transcription is monocistronic via conserved sequences in the 5'UTR of genes that recruit transcription factors, in trypanosomes, DNA conformation may play a role in determining sites of transcription initiation. Histone variants showing poor levels of stability are enriched at regions of transcription start sites, effectively causing the chromatin to open up and thus making this area accessible to transcription factors (Siegel et al., 2009). The lack of regulation at the level of transcription explains the observation that the *T. brucei* genome has such a limited repertoire of predicted transcription factors (Campbell et al., 2003).

Once the polycistronic unit has been transcribed the message must be processed, which in trypanosomes requires that the individual genes are trans-spliced and polyadenylated (Figure 1.12B). A Spliced Leader (SL) is annealed to the 5' end of the nascent mRNA at a conserved splice acceptor site, stabilising the transcript and making it amenable for translation (Palenchar and Bellofatto, 2006). The SL RNA genes are arranged as multiple repeats throughout the genome and account for a large proportion of the RNA synthesis in the cell in order to have enough available for each transcript generated (Campbell et al., 2003).

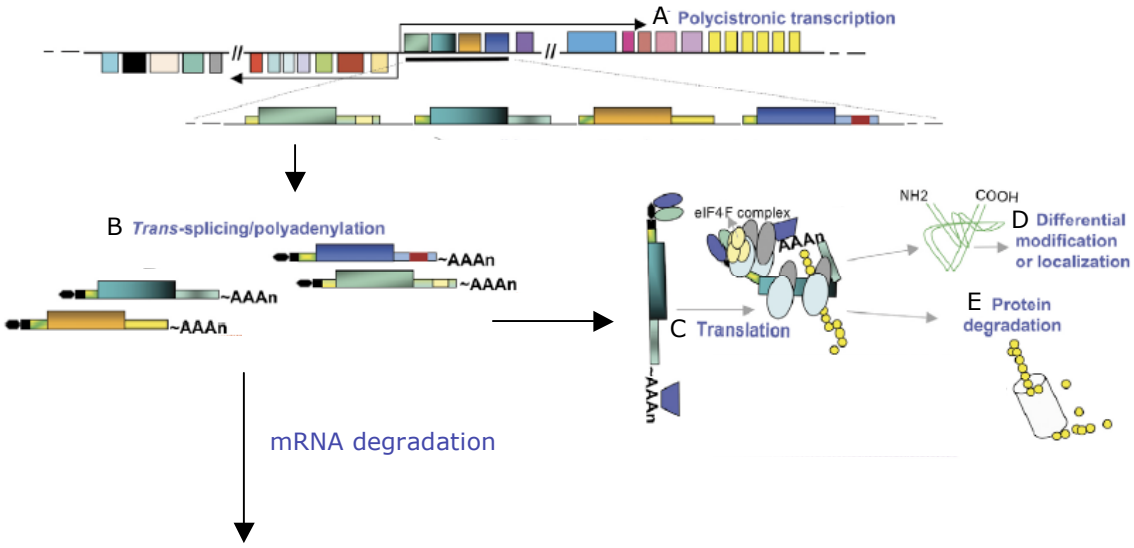



**Figure 1.12. Overview of the processes involved in transcription in *T. brucei*.** (A) Polycistronic transcription occurs from ill-defined PolII promoters. (B) Each mRNA is then prepared for translation through the addition of a 5' SL fragment and 3' polyadenylation. (C) The fate of these mRNAs is either degradation or translation into proteins. (D) Those transcripts that are not degraded can then be translated with the final proteins either sent to their required localisation or regulated through localisation to a place where it will be inactive, (E) alternatively the protein may be degraded (Modified from (Haile and Papadopoulou, 2007)).

As there is little or no regulation of genes at the transcriptional level, other methods exist to affect the levels of transcripts. Degradation of either the transcript or the mature protein is one such method, as is the sequestering of a protein from its site of action (Figure 1.12D, E).

### 1.8.2 Regulatory elements

A number of regulatory elements have been found, particularly within the untranslated regions (UTR) of genes, conferring stage specific expression to them. Most of the regulatory elements identified in kinetoplastid genes have been within the 3'UTRs. Typically in higher eukaryotes, regulatory sequences in the 3'UTR are AU rich, and such AU rich regions have been discovered to be required for the procyclic specific expression of EP1, phosphoglycerate kinase B (PGKB), pyruvate phosphate dikinase and many of the COX subunits (Haile and Papadopoulou, 2007). The regulatory region for these genes is a loosely conserved 26mer (Hotz et al.,

1995). More than one regulatory element has been identified in the 3'UTR of EP1, two of which positively regulate gene expression and a Loop II region, which acts as a negative regulator. The Loop II region is targeted by negative regulators to prevent expression of this gene in the bloodstream form, but in the procyclic form a zinc finger protein, TbZFP3, competes for binding to the Loop II domain, likely preventing the negative regulators from binding and causing destabilisation of the mRNA (Walrad et al., 2009).

Not all regulatory elements found in kinetoplastid 3'UTRs are AU rich however. For example, a 450 bp region in the 3'UTR of genes up regulated during the amastigote stage in *Leishmania infantum,* showed no enrichment of AU (Boucher et al., 2002). Interestingly this region appeared to regulate expression levels through enhancing translational efficiency.

### 1.8.3 RNA binding proteins

As is seen from the action of TbZFP3, many of the gene regulatory mechanisms rely on RNA binding proteins, with depletion of these genes leading to severe phenotypes, (Estevez, 2008, Walrad et al., 2009, Hartmann et al., 2007, Hartmann and Clayton, 2008). Kinetoplastids have significantly more of the proteins involved in RNA binding than other organisms that conduct expression regulation at the transcriptional level (Archer et al., 2009). For example, the homologues of the UBP1 and UBP2 proteins that were originally found in *T. cruzi,* and which contain a RNA recognition motif, have been shown in *T. brucei* to regulate mRNA levels. Thus, over expression of TbUBP2 leads to an increased half life in one of its substrates, CFB1 – a protein potentially involved in cell division, and both over expression or depletion of the proteins leads to growth inhibition of both bloodstream and procyclic form cells (Hartmann et al., 2007). Similarly, the RNA recognition motif (RRM) containing protein DRBD3 was identified through its ability to associate with the 3'UTR of PGKB and was found to be involved in the regulation of a range of mRNAs, all of which had a conserved sequence in their 3'UTR, known to be over represented in the 3'UTR of genes specifically expressed in the procyclic form (Estevez, 2008). More recently, the targets of one protein of the Pumillo family of

RNA binding proteins, Puf9, were identified, with Puf9 found to regulate the timing of expression of these transcripts via a conserved motif in their 3'UTR. Puf9 appeared to bind this element and specifically stabilised these mRNAs during S phase (Archer et al., 2009).

### 1.8.4 G-quadruplexes

Recently, Guanine-rich sequences forming a tertiary structure (G quadruplex) in promoter regions of *Plasmodium falciparum* have been implicated in transcription regulation (Smargiasso et al., 2009). G-quadruplexes cause conformational alterations in the DNA structure and can recruit binding proteins which will then prevent access by transcriptional machinery, or alternatively, require specific helicases in order to unwind these regions. Depletion of the helicase required to unwind the G-quadruplex leads to down regulation of the genes containing these sequences in their promoters. Poly-G sequences have been identified at transcription strand switch regions in *T. brucei*, potentially indicating a role in regulation, although it is not clear whether this would be a positive or negative regulation or whether, instead, the G-quadruplex structure could act as a barrier to prevent the RNA polymerase from proceeding in the wrong direction (Siegel et al., 2009).

### 1.8.5 Base J

A further novel mechanism of gene regulation is through the modification of certain T residues into ß-D-glucosyl-hydroxymethyuracil, otherwise known as base J. This phenomenon only occurs in the bloodstream form of *T. brucei* and the major role appears to be in the silencing of the inactive telomeric VSG genes (Borst and Sabatini, 2008). Interestingly, base J also shows enrichment at both divergent and convergent polycistronic switching sites, potentially suggesting these modifications also play a role in transcription initiation at the level of chromosome structure (Cliffe et al., 2009).

## *1.9 Microarray*

Despite the largely post-transcriptional regulation of gene expression in *T. brucei*, the use of mRNA stability as a method of regulation means that the amount of transcript can give an indication of the importance of a particular protein in certain situations.

Microarrays have become an increasingly common method for comparing differences in gene expression under various conditions in *T. brucei,* such as the effects of drug treatments and gene knock down and over expression experiments (Estevez, 2008, Denninger et al., 2007, Hartmann et al., 2007) or to identify mRNAs bound by RNA binding proteins (Archer et al., 2009). Numerous studies have also been conducted to examine the expression profiles of genes in the different life cycle forms of the kinetoplastids. These early experiments were conducted prior to the publication of the complete genome sequence, often on custom made glass slides, and thus are difficult to interpret. Moreover, all have been carried out on cultured monomorphic bloodstream form cells (Diehl et al., 2002, Brems et al., 2005). Other microarray studies of the life cycle have only considered a subset of the genome (El-Sayed et al., 2000, Koumandou et al., 2008). The first attempt at a microarray in *T. brucei*, looked at a *T.b.rhodesiense* strain and used 400 expressed sequence tags (ESTs) as probes and found that 13% of these showed differential regulation between bloodstream and procyclic form cells (El-Sayed et al., 2000). Subsequent studies using greater number of ESTs in order to gain almost total genome coverage have found that this level of regulated expression is probably an artefact of using highly expressed mRNA, which is more likely to show differential expression. Taking this into account, the study calculates the percentage of differentially expressed transcripts between these life stages as closer to 2% (Brems et al., 2005, Diehl et al., 2002). These early microarray studies gave little insight into the global regulation of gene expression but are good proof-of-principle studies demonstrate that *T. brucei* will provide interesting information regarding gene expression regulation between different life cycle stages in future experiments (Duncan, 2004).

Koumandou et al (2008) designed microarray slides for the analysis of genes involved in membrane transport, comprising around 8% of the genome alongside a set of over 100 genes of known expression profile used for reference (Koumandou et al., 2008). Using these slides the difference between monomorphic bloodstream forms grown in culture were compared with another monomorphic cell line harvested from a rodent infection, with 35 genes found to be up-regulated in the in vitro

cultured cells with only three up-regulated in the in vivo derived cell lines (Koumandou et al., 2008). Considering this microarray only investigated 796 genes, extrapolation based on this experiment would predict that almost 5% of the genome is differentially regulated between in vivo and in vitro cultured bloodstream form cells. This suggested that experiments on parasites derived from long term culture may give a very different profile to in vivo derived cells. For example, the percentage of differentially regulated genes found between the in vivo and in vitro cultured cells is comparable with percentage difference between the bloodstream form and procyclic form cells, which are known to be vastly different in many facets of their cell physiology (Koumandou et al., 2008). Similar conclusions have been reached in studying *Leishmania* cells where comparisons of axenically cultured *Leishmania infantum* amastigote form cells with those harvested from an infection showed that 40% of the genome is up-regulated in the axenic cells (Rochette et al., 2009). A microarray study has been used to look at the differences between the bloodstream form and procyclic form in a monomorphic cells line and compare these with the differences seen between these two life stages in a pleomorphic cell line (Brems et al., 2005). Interestingly the majority of the regulated transcripts were the same in both cell lines, with only two additional transcripts found to show differential regulation in the pleomorphic cell line (Brems et al., 2005). Life cycle stage microarrays have also been conducted on *T. cruzi*, showing a greater proportion of genes that are stage regulated than either *T. brucei* or the *Leishmania* ssp (Minning et al., 2009).

However, to date none of the microarrays conducted to investigate differences in expression profiles in different life cycle stages have looked at the transient changes associated with differentiation from one form to the other. The lack of investigation of gene expression throughout a time course of differentiation means little is known of the genetic events required for differentiation, and a microarray study on *Leishmania donovani* looking at gene expression over a time course during differentiation, found that many genes are transiently regulated during differentiation of promastigotes to amastigotes (Saxena et al., 2007). It is therefore likely that

similar events will occur in *T. brucei* and investigation of this will allow a better understanding of the process of differentiation.

### *1.10 The Aims of this PhD*

- Investigate orthologues of proteins known to be involved in G0/G1 control in other organisms and determine whether they play a role in differentiation in *T. brucei.*
- Develop a method for isolation of cells in different cell cycle stages and assess the differentiation capacity of cells in these different stages.
- Examine the global gene expression profile of differentiating cells to identify genes involved in key stages.

# 2 – Materials and Methods

## *2.1 Trypanosomes*

### 2.1.1 Strains

The culture adapted stains used throughout this study are all derived from the non-human infective subspecies *T. b. brucei* Lister 427 strain. These cells are all monomorphic; they are unable to take on the stumpy morphology and are unresponsive the stumpy induction factor (SIF). The details relating to the culture history of this cell line are available on Professor George Cross' website (http://tryps.rockefeller.edu). RNA interference (RNAi) was carried out in the single marker cell line S16. This bloodstream form cell line expresses the T7 RNA polymerase and tetracycline repressor and contains a neomycin resistance marker (Wirtz et al., 1999). These cultures were maintained in 2.5µg/ml G418 to prevent selective loss of the toxic T7 polymerase.

The over-expression cell line 427-449 were used for differentiation experiments as this cell line had been demonstrated to have retained the ability to differentiate in response to cis-aconitate at high levels. The 427-449 cell line was derived from 427 cells transfected with the plasmid pHD449, incorporating a phleomycin resistance marker and a tetracycline repressor gene that allows inducible ectopic expression. A benefit of this cell line is that the presence of the pHD449 plasmid is not detrimental to the cell and thus this construct is stably maintained without the need to select with antibiotics.

Two pleomorphic cell lines were also used in this study, which naturally differentiate from the slender to stumpy form. The pleomorphic cell line AnTat1.1 (Van Meirvenne et al., 1975) was used to generate stumpy form cells and for use in the differentiation experiments for the microarray as this cell line is still able to complete both the mammalian and fly stages of the life cycle and the kinetics of differentiation of this cell line are well characterised. A second pleomorphic cell line, EATRO

(East African Trypanosomiasis Research Organisation) 2340 *T. brucei brucei* strain (Cunningham and Vickerman, 1962), were also used to generate RNA to compare expression of stumpy-specific transcripts in different strains.

### 2.1.2 Culturing

Cells were cultured in HMI-9 medium (Hirumi and Hirumi, 1989) supplemented with 20% foetal calf serum (FCS) (Gibco) and 10µg/ml penicillin/streptomycin, at 37˚C in 5% $CO_2$, in vented flasks (Greiner). Cell density was kept between $2x10^5$ cells/ml to $4x10^6$ cells/ml. Cell counts were conducted on a Beckman Coulter Counter Z2, calibrated to 2.5-7µm to exclude cell debris. Differentiation of stumpy form cells was carried out through transferring cells into SDM-79 (Brun and Schonenberger, 1979, Cross and Manning, 1973), supplemented with 10% FCS and 2.5µg/ml haemin, with 10mM glycerol and 6mM cis aconitate. For differentiation, cells were transferred to flasks with non-vented lids and incubated at 27˚C.

Cryopresevation of cells was carried out through centrifuging 10ml of a logarithmically growing culture at 600xg for 10 minutes. All but 0.5ml of the supernatant was removed and the cells resuspended in the remaining media, then a further 0.5ml of HMI-9 medium supplemented with 14% glycerol was added dropwise. The mixture was allowed to stand and then transferred to -80˚C in a cryovial.

### 2.1.3 Isolation of parasites from blood

Mice were immune-suppressed through treatment with cyclophosphamide 24 hours prior to injection with parasites. After three days for isolation of slender form cells, or six days for stumpy form cells, the blood was harvested through cardiac puncture and collection into pre-warmed syringes by a licensed animal worker. The trypanosomes were then purified out from the blood in PSG at 37$^o$C via DEAE cellulose columns (Lanham and Godfrey, 1970). Harvested parasites were then centrifuged at 800xg for 10 minutes in a clinical centrifuge, resuspended in HMI-9 medium pre-warmed to 37˚C, at approximately $1x10^8$/ml and settled for 1 hour.

### 2.1.4 Differentiation

Monomorphic bloodstream form cells were induced to differentiate in HMI-9 medium at a density of $5x10^5$ cells/ml. Cis aconitate was added to a final concentration of 6mM, vented lids replaced with solid lids, and cultures transferred to a 27˚C incubator. Pleomorphic AnTat 1.1 cells were induced to differentiate from stumpy form cells extracted from blood, through transferring the cells at a density of $2.5x10^6$ cells/ml into SDM79 with 6mM cis aconitate, pre-warmed to 27˚C. The differentiation medium also contained 10mM glycerol to allow continued expression of GPEET (Vassella et al., 2000).

### 2.1.5 Transfection

Between $1\text{-}2x10^7$ bloodstream form trypanosomes were centrifuged at 600xg for 10 minutes, washed once in half volume of ZPFMG and resuspended in 0.5ml of ZPFMG. 15µg plasmid DNA linearised overnight with Not1, purified using a Qiagen PCR Purification Kit (following the manufacturers instructions) and resuspended in 50µl $H_20$ was mixed with the cells in 0.4cm electroporation cuvettes. Electroporation was conducted using a BTX 830 Electo Square Porator at 1700v, with three pulses for 100µs carried out at 200ms intervals, and immediately transferred to 10ml non-selective HMI-9 medium and incubated for 20 hours at $37^oC$ in 5% $CO_2$. Cells were then diluted 1 in 4 in selective media (hygromycin at 1.5µg/ml and neomycin at 2.5µg/ml) in 24 well plates for seven days and successfully transfected cell lines were diluted 1 in 5 into 6 well plates before being propagated in 10ml flasks.

## *2.2 Viable isolation of cells in specific cell cycle stages*

Bloodstream form cells were harvested during exponential growth by centrifugation for 10 minutes in a clinical centrifuge (600x*g*) at room temperature. The parasite pellet was then resuspended to a density of $1x10^6$ cells/ml in HMI-9 medium supplemented with 2% FCS and 10µg/ml penicillin/streptomycin. Vybrant® DyeCycle™ Violet (Molecular Probes, Invitrogen V35003) was added to a final concentration of 1µg/ml and the cell suspension incubated for 30 minutes at 37˚C, the tube being protected from light by wrapping in aluminium foil. The samples

were then concentrated into a one-quarter volume of staining media prior to sorting on a FACSAria (BD Biosciences). Autoclaved PSG was used as the sheath fluid in order to minimise cell stress, and the machine cooled to 4˚C prior to cell sorting, such that the temperature of the sorting sample would have been below 20˚C, in order to limit cell metabolic activity. Intact cells were gated based on FSC/SSC profiles, combined with pulse processing for doublet discrimination, in order to select single cells. The dye was excited using a 407nm Violet laser and emission detected via a 450/40 bandpass filter. Gates were set up to collect only the 2C fraction (G0/G1 cells) and 4C fraction (G2, mitotic and post-mitotic cells), with the 2C gate set to include 1 quartile either side of the G0/G1 peak and the 4C gate set to exclude the first quartile of the G2/M peak, to ensure tight discrimination and selection of these cell cycle stages. In all cases, cells were sorted under conditions of minimal ambient light to prevent photoinactivation of the Vybrant® DyeCycle™ Violet. Sort time did not exceed 90 minutes, with cells collected into filter-sterilised 100% FCS and maintained below 20˚C until the overall sort for all cells was completed. A sample of the recovered cells was re-analysed on the FACSAria to confirm purity of the populations. The cells diluted in PSG were then centrifuged at 600xg for 10 minutes and resuspended in HMI-9 medium supplemented with 40% FCS and cultured at 37˚C.

### *2.3 Viability assay using alamarBlue®*

Viability of cells treated with Vybrant® DyeCycle™ Violet stain and exposed to temperature reduction was investigated by subjecting cells to the same staining regime as for the sort and then analysing cellular metabolism via the redox indicator alamarBlue® (Serotec). Simulating the sorting protocol, 10ml cells were stained at $1x10^6$ in 2% FCS supplemented HMI-9 with 1µg/ml Vybrant® DyeCycle™ Violet at 37˚C for 30 minutes, with a control set of cells not treated with the DNA stain incubated in parallel. Both control and stained cells were then split into two falcon tubes and 20ml of PSG added to each tube to simulate the dilution in sheath fluid experience by the cells in the FACS sorter and incubated for 90 minutes to replicate the average length of a sort, with one of the tubes placed at 4˚C to mimic the FACS machine temperature and the other remaining at 37˚C to establish whether the

cooling step had a detrimental effect on the cells. The cells were then counted and centrifuged for 10 minutes at 600xg and resuspended at $2x10^6$cells/ml in 40% FCS supplemented HMI-9. 200µl of each sample was loaded in triplicate into a 96 well plate (Fisher Scientific) with 10% v/v alamarBlue®, with a further control of unmanipulated, unstained control cells grown in normal 20% FCS supplemented HMI-9, and wells containing 20% FCS and 40% FCS supplemented media alone and with 10% v/v alamarBlue®. Excess cells were transferred into a 6 well plate, sealed with Parafilm and incubated at 37˚C. The 96 plate was then incubated for 1 hour at 37˚C in 5% $CO_2$ with the lid placed on but not sealed, after which a MicroAmp™ optical adhesive film (Applied Biosystems) was adhered over the wells. The plate was loaded into a BioTek® Elx808 plate reader at a temperature of 37˚C and the change in absorbance of the alamarBlue® was then recorded at 540 and 595nm hourly over a 24 hour time course using Gen5 software. The data were adjusted through application of a correction factor to adjust for the use of a non-preferable wavelength of 595 instead of 600nm using the equation:

$$\frac{\text{Absorbance at 540 of media + alamarBlue® – Absorbance at 540 of media alone}}{\text{Absorbance at 595 of media + alamarBlue® – Absorbance at 595 of media alone}}$$

calculated for each time point. The percentage reduction at each time point of the alamarBlue® could then be calculated by the equation:

% reduction = Absorbance at 540 –(Absorbance at 595 x correction factor) x100.

## *2.4 Yeast experiments*

### 2.4.1 Strain and culturing

The yeast strain used was *Saccharomyces cerevisiae* BY4741, with the genotype MAT**a**, his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 (kindly supplied by Aziz El Hage). This strain has a GFP tag fused onto the Net1 gene to allow visualisation of the rDNA repeats on chromosome XII. Frozen stocks of cells in YPD + 15% glycerol were streaked onto YPD plates and incubated at 30˚C, and single colonies used to inoculate into YPD liquid media. Liquid cultures were grown in autoclaved conical flasks with a capacity at least ten times larger than the volume of liquid, and incubated at 30˚C with constant shaking.

### 2.4.2 Growth assays

#### 2.4.2.1 Growth in liquid culture

The optical density of an overnight culture was assessed using the 600nm wavelength on an Eppendorf BioPhotometer and diluted to an OD600 of 0.005 into prewarmed YPD media. The cells were then incubated with shaking until the OD600 of the culture reached 0.01, at which point the cells would be in logarithmic growth. The culture was then split into autoclaved glass vials with a 35ml capacity and 3ml of the culture was placed into each vial with various doses of rapamycin (LC Laboratories), the solvent (100% ethanol) or left untreated. The cultures were returned to the incubator and grown for a further eight hours with the OD600 recorded every two hours. Alternatively, 100µl samples of the drugged cultures were loaded in triplicate into 96 well plates, sealed with a MicroAmp™ optical adhesive film (Applied Biosystems) and incubated in a BioTek® Elx808 plate reader at 30˚C with continuous shaking set to 'medium'. This machine does not have a 600nm setting, thus the OD595 was recorded every hour for 12 hours using Gen5 software.

#### 2.4.2.2 Spot assays

Yeast cells grown in liquid culture to an OD600 of 0.01 were used to make 10-fold serial dilutions with sterile water and 10µl spotted onto YPD plates containing either ethanol, 100nM, 500nM or 1µM rapamycin. The plates were allowed to dry and then incubated at 30˚C for 48 hours.

### *2.5 Molecular biology*

### 2.5.1 Small scale DNA preparation of plasmids

Overnight bacterial cultures were prepared through inoculation of a single colony into 2ml of LB with 100µg/ml ampicillin and incubation in a shaking incubator at 37˚C overnight. The DNA was extracted from these cells using the rapid alkaline extraction procedure (Birnboim and Doly, 1979). 1.5ml of the overnight culture was centrifuged in an Eppendorf at maximum speed for 5 minutes and the supernatant removed. The pellet was resuspended into 100µl ice cold Solution I through

pipetting and incubated on ice for 5 minutes. 200µl Solution II was then added, mixed through inverting the tube and incubated on ice for 5 minutes. 150µl of cold Solution III was added and the neutralisation reaction allowed to continue for 5 minutes through incubation on ice. The cell lysate was centrifuged at maximum speed for 10 minutes after which time the supernatent was transferred into an Eppendorf containing 900µl 100% ethanol, and incubated on ice for 15 minutes. Another centrifugation at maximum speed for 10 minutes was conducted to pellet the DNA, with the supernatant discarded. The DNA was washed in 200µl 70% ethanol, air-dried and resuspended in 40µl sterile water containing 20µg/ml RNase.

### 2.5.2 Large scale DNA preparation of plasmids

Large scale preparations of DNA plasmids were carried out using a Qiagen midi-prep kit following the manufacturer's instructions from a 200ml overnight culture.

### 2.5.3 Separation of DNA on an agarose gel

Agarose gels of either 30ml of 100ml were cast in a tank (Bio Rad), through melting agars in TAE, with Safe View (NBS Biologicals) added to a final concentration of 3.5%. Routinely a 1% gel was used for resolution of sequences, although for separation of DNA sequences less than 400bp, 2% gels were used. DNA loading buffer was added to samples and these were run alongside 5µl of the size marker SmartLadder (Eurogentec). Gels were immersed in TAE buffer and subjected to electrophoresis at 100V until the DNA was resolved. DNA was visualised on a UV light course and photographed using GeneSnap software (Syngene).

### 2.5.4 Cloning

Chemically competent *Esherichia coli* XL1-Blue (Stratagene) cells were transformed with plasmid constructs. Cells, stored at -80˚C, were thawed on ice and 100µl was gently added to transforming DNA then incubated on ice for 20 minutes. The cells were heat-shocked at 42˚C for 45 seconds, briefly cooled on ice, then recovered in 100µl LB broth by incubation at 37˚C with shaking for 1 hour. The entire reaction was then plated onto LB agar with 100µ/ml ampicillin and incubated overnight at 37˚C.

### 2.5.5 Polymerase Chain Reaction (PCR)

PCRs were carried out using as a template genomic DNA from monomorphic 427 cells. Genomic DNA was prepared using an animal tissue extraction kit (Qiagen) following the manufacturer's instructions. Standard PCR reactions consisted of a 50µl reaction comprising: 20% 5x Green Buffer, 0.2mM dNTPs, 1.5mM $MgCl_2$, 1µM each of forward and reverse primers and 0.0375 units/µl GoTaq. The 5x buffer used contains loading dyes, thus allowing the final product to be loaded directly into an agarose gel. Both the GoTaq and buffers were purchased from Promega and primers ordered from Sigma. Reactions were conducted in a ThermoCycler (Thermo Electron Corporation), with an initial denaturation step of 5 minutes at 95˚C, followed by 30 cycles of 30 seconds at 95˚C, a 30 second annealing step at 60˚C and an elongation step of 2 minutes at 72˚C, with a single final elongation at 72˚C for 5 minutes.

### 2.5.6 Restriction enzyme digests

Digestion of DNA using restriction enzymes was typically carried out using: up to 15µg DNA, 10% 10x recommended buffer (NEB), 10 units enzyme and 10% 10xBSA as required. Reactions were incubated at 37˚C for 90 minutes, or overnight for large amounts of DNA.

### 2.5.7 Ligation

Ligation into pGem T-Easy was carried out directly from PCR products via the T overhangs. The reaction comprised: 50ng pGemT-Easy vector (Promega), 5% supplied 2x Buffer (Promega), approximately 150ng insert DNA and 1 Weiss Unit T4 Ligase (Promega), and were incubated at room temperature for 1 hour. For ligation into other vectors, restriction digests were first carried out in order to generate complementary stick ends and both the insert DNA and the vector were separated on an agarose gel and excised using a clean scalpel to remove any uncut fragments, and purified using a gel purification column (Machery Nagel). These ligation reactions comprised: 10% 10x ligase buffer (Promega), 0.15U/µl T4 DNA ligase and a ratio of 6:1 of insert:vector DNA, and were incubated at 4˚C overnight.

### 2.5.8 Sequencing

Prior to transfection into cells, RNAi constructs were sequenced to ensure the insert was correct either from the pGem-T-Easy plasmid using the universal M13 sequencing primers, or from the RNAi plasmid using the primers designed to amplify the insert. All sequencing was performed by the Gene Pool (Edinburgh University) using an ABI 3730 capillary sequencing machine. 10-40ng of plasmid DNA was required in addition to a primer, either the M13 sequencing or a custom primer. Results were analysed using 4Peaks version 1.7.1.

### 2.5.9 RNA interference

*T. brucei* has the required cellular machinery for RNA interference, and this was exploited for the inducible depletion of genes of interest. Primers were designed against a unique sequences of the gene using RNAit (Redmond et al., 2003) to generate sequences of approximately 500bp, incorporating XhoI and BamHI restriction sites onto the ends of the sequences to allow subcloning. The amplified product was subcloned into pGem-T-Easy (Promega) using AT cloning, and the construct was then transformed into bacterial cells. The construct was released from the pGem vector through digestion with XhoI and BamHI and ligated into the p2T7Tablue vector (http://trypanofan.path.cam.ac.uk/trypanofan/vector/), cut with the same enzymes to produce complementary sticky ends. Once stably transfected into S16 cells, the expression of this construct could be inducibly expressed through the addition of 1µg/ml tetracycline. Clones successfully transfected with the construct were then maintained in hygromycin to prevent loss of the construct through selective pressure.

### 2.5.10 Immunofluorescence

1 million cells were harvested per slide through centrifugation if cells for 5 minutes at 700xg, washing once in PBS and resuspended in 100µl PBS. The cell suspension was gently spread over the slide using the side of the tip and allowed to air dry. The slide was then immersed in cold methanol for at least 30 minutes at -20˚C. The cells were rehydrated in PBS for 5 minutes then blocked in 5% BSA/PBS for 10 minutes, and stained with mouse anti-EP antibody (Cederlane Laboratories) diluted 1:500, and rabbit anti-AnTat 1.1 VSG antibody diluted 1:10,000, in 5% BSA/PBS for 45

minutes in a humidity chamber. Unbound primary antibody was removed through washing three times for 5 minutes in PBS, then slides were probed with anti-mouse conjugated TRITC (Sigma) diluted 1:50 in 5% BSA/PBS and anti-rabbit conjugated FITC (Sigma), also diluted 1:50, for 45 minutes in a humidity chamber. The slides were washed three times for 5 minutes and then stained with 10µg/ml 4', 6-diamnidino-2-phenylindole (DAPI) for 2 minutes and washed a final time. The slides were mounted in 40µl MOWIOL with 10% of the anti-fade agent p-Phenylenediamine (PDA) and a coverslip added. Slides were allowed to set overnight at 4˚C before visualisation on a Zeiss Axioskop 2 plus microscope. Images were captured using QCapture and measurements made using ImageJ 64., with at least 100 cells measured per slide.

### 2.5.11 qRT-PCR

#### 2.5.11.1 Primer design

qPCR primers were designed to amplify a 143-217 bp amplicon using Primer3 (Rozen and Skaletsky, 2000), with constraints in place to avoid picking primers with more than two consecutive repeated bases. Primers also had to end in a CG clamp and have a melting temperature of 60˚C. All primer pairs were screened using Amplify version 3.1 (http://engels.genetics.wisc.edu/amplify/) to assess for formation of primer dimers and finally the primers were used as templates in BLAST searches to ensure they would uniquely amplify regions from the single target gene. All primer pairs were tested in PCR reactions against genomic DNA using an annealing temperature of 60˚C to ensure they were capable of amplification at this temperature and did not form primer dimers.

#### 2.5.11.2 qRT-PCR reaction

cDNA was made from the same RNA as was used for the microarray using Oligo $(dT)_{15}$ primer to specifically amplify mRNA, in a 20µl total volume using the Reverse Transcription System (Promega). Briefly, 1µg sample RNA was incubated at 70˚C for 10 minutes then cooled on ice. The reaction components were then added to the RNA, comprising: 5mM $MgCl_2$, 1x reverse transcription buffer, 1mM dNTP mixture, 25 units RNasin ribonuclease inhibitor, 15 units AMV reverse

transcriptase, 0.5µg oligo$(dT)_{15}$ primer. The reaction was incubated at 42˚C for 15 minutes, 95˚C for 5 minutes and then placed on ice for 5 minutes and finally diluted 1 in 5 with nuclease-free water. cDNA was stored at -20˚C until required. Alongside each preparation of cDNA from RNA, a No-RT control was also generated, where the same reaction was prepared in the absence of reverse transcriptase.

For qRT-PCR, 20µl reactions were set up containing 5µl cDNA in 50% SYBR Green Master Mix (Roche) with 0.5µM forward and reverse primers on ice. In order to avoid the possibility of contamination, a master mix solution was prepared of the SYBR Green and primers and added to the wells, with the cDNA added afterwards and mixed through gentle pipetting. One row of wells was prepared at a time and all other wells were covered with Parafilm. Plates were sealed with sealing foils (Roche) and centrifuged at 1500xg for 2 minutes prior to loading into the qRT-PCR machine. Amplification reactions were carried out in white 96 well plates (Roche), in duplicate, in a LightCycler 480 machine (Roche) comprising a 10 minute pre-incubation step at 95˚C followed by 40 cycles of 95˚C for 10s, 55˚C for 20s, 72˚C for 10s with a single acquisition read taken at 82˚C. A meltcurve analysis was performed from 65˚C to 97˚C with two acquisitions taken per ˚C, to test for primer dimer contamination. No-RT controls were included for each 96 well plate, whereby the qRT-PCR reaction was conducted using as a template, cDNA which had been generated alongside the sample cDNA, but which had not had reverse transcriptase added. In addition, no-template controls were also run in each experiment for which the cDNA was replaced with $dH_2O$.

Prior to using the primers to amplify sample cDNA, the primers were first tested against serially diluted cDNA taken from *T. b. brucei* EATRO 3420 stumpy form cells, to assess the PCR efficiency of the primers. The cDNA was diluted 1:1000, 1:100, 1:10 or left undiluted and the Cp of the samples plotted on a graph from which the steepness of the curve could be calculated. From the curve slope it was possible to calculate the PCR efficiency of the primer set, and those not demonstrating an

efficiency between 90-110% were rejected. qPCR primer sequences and PCR efficiencies are listed in Appendix I. Relative quantification was calculated using the Pfaffl equation (Pfaffl, 2001) using as a standard Tb10.389.0540, a gene encoding a hypothetical protein, which was identified as being a stably expressed gene across all time points by array analysis.

## *2.6 Northern Blotting*

### 2.6.1 Preparation of riboprobe

A riboprobe was generated from the same unique sequence used for the RNAi, using the Roche DIG RNA labelling kit. The unique sequence was ligated into pGEM and orientation discovered through amplification of the insert using a combination of M13 and gene-specific forward and reverse primers. The insert was amplified using the M13 primers and, depending on the orientation of the insert, either the S6P or T7 RNA polymerase used to generate an antisense RNA probe. The probe was stored at -80˚C.

### 2.6.2 Separation of mRNA on a denaturing gel

RNA was prepared from trypanosomes during the slender and stumpy lifecycle stages taken from a rodent, and procyclic stage grown in culture, or from cultured monomorphic bloodstream form cells using a Qiagen RNA-easy kit according to the manufacturer's instructions, completed with optional centrifugation step. The RNA samples that were to be used for the microarray were also subjected to the optional DNase treatment. All RNA samples were resuspended into 50µl RNase-free $H_2O$ and the concentration of the RNA was assessed either through loading 1.5µl of the sample onto a nanodrop (Thermo Scientific) or through dilution of 1µl sample into 99µl of $H_2O$ and assessment of RNA concentration through absorbance on an Eppendorf BioPhotometer.

A 150ml denaturing gel (3% formaldehyde, 1.2% agarose in 1xMOPS) was prepared in a fume cupboard. 2µg of RNA was denatured in 9µl formamide, 3µl 37% formaldehyde, 2µl 10xMOPS and 2µl RNA loading buffer (Appendix I) at 65$^{o}$C for 10 minutes and immediately loaded into the gel. The RNA was separated in a

running buffer of 1x MOPS at 150 volts for approximately 90 minutes at which point the RNA would be adequately resolved. The gel was stained with ethidium bromide for 15 minutes, destained twice for 30 minutes in water and photographed on a UV transilluminator in order to view the rRNA.

### 2.6.3 Capillary blotting of mRNA

MME paper was soaked in 10x SSC, and used as a wick in a reservoir of 10x SSC. The gel was flipped upside down onto the wick so that once the mRNA was transferred into the membrane it would be in the same orientation as the gel had been loaded. A positively-charged nylon membrane was soaked in 10x SSC and carefully placed onto the gel so as to avoid trapping air bubbles. Two layers of MME filter paper (Whatman) were soaked in 2x SSC and laid on top of the membrane, with approximately 10cm of paper towels stacked on top and weighted down using ~500g weight. SaranWrap was used to cover to reservoir to prevent evaporation of the liquid. The RNA was allowed to transfer overnight via capillary action and fixed onto the membrane using a UV crosslinker (UVI tech Ltd, UK) using a setting of 0.12 joules for 120 seconds.

### 2.6.4 Hybridisation of riboprobe

The membrane was probed overnight after pre-hybridising for 1 hour at 68$^{o}$C in hybridisation buffer (5xSSC. 50% formamide, 0.02% SDS, 2% blocking solution). The pre-hybridisation solution was discarded and 2µl of riboprobe boiled for 10 minutes in 100µl hybridisation buffer was added to 7ml of prewarmed hybridisation buffer and added directly to the membrane and left to hybridise in rolling tubes in a Hybaid oven at 68˚C overnight. The following day the membrane was washed two times for 30 minutes at 68$^{o}$C in 2xSSC/0.1%SDS, then once more for 30 minutes at 68$^{o}$C in 0.5xSSC/0.1%SDS. Washing and blocking steps conducted at room temperature were carried out in plastic lunch boxes which were washed with detergent and 70% ethanol between each step. The membrane was rinsed in Wash Buffer (0.3% Tween 20 in Maleic Acid) for 1 minute at room temperature and blocked with constant agitation for one hour in Maleic Acid with 1% Blocking solution (Roche). After this, the membrane was incubated in Maleic Acid with 1% Blocking solution and Anti-DIG (1:25000) for 30 minutes and then washed three

times in Wash Buffer. The membrane was then soaked in Detection Buffer (100mM Tris-HCl, 100mM NaCl, pH9.5) for 2 minutes and placed in Bag 'W'. The excess Detection Buffer was removed and the membrane was labelled with 1:100 CDP-Star (Boehringer) in Detection Buffer through incubation in Bag 'W' for 2 minutes. Excess Detection Buffer was removed and the Bag 'W' was sealed in a heat-sealer and incubated at 37$^{o}$C for 15 minutes to activate the CDP-Star then exposed on X-ray film (Kodak). Routinely, the riboprobes used gave a good signal after 30 minutes of exposure.

## *2.7 Western Blot*

### 2.7.1 Protein preparation

$2x10^7$ trypanosomes were pelleted for 10 minutes at 600xg , washed in PBS, and pelleted again. The pellets was resuspended in 100µl Laemmli sample buffer and boiled for 5 minutes. Protein samples were stored at -20˚C until needed.

### 2.7.2 Protein resolution on acrylamide gels

SDS acrylamide gels were made up at 6% in order to resolve the 270kDa Tor protein as follows: 6% acrylamide-bisacrylimide stock (Severn Biotech Ltd), 375 mM Tris pH8.8, 0.1 % ammonium persulphate (APS) (w/v), 0.1% SDS, 0.05 % tetramethylethylenediamine (TEMED). The gel was poured into Bio-Rad PROTEAN II casting apparatus, covered with a layer of 70% ethanol and, once set, the ethanol was rinsed off with $dH_2O$ and the stacking gel poured on top. The stacking gel comprised 125 mM Tris pH6.8, 5% acrylamide-bisacrylimide stock, 0.1% SDS, 0.1% ammonium persulphate and 0.1% TEMED. 15µl of protein samples were loaded per lane, equivalent of $3x10^6$ cells and run alongside All Blue prestained protein marker (BioRad). The protein were separated through applying a voltage of 150V for ~60 minutes, or until the dye front had reached the end of the gel.

### 2.7.3 Transfer

The proteins were transferred onto nylon membrane (Whatman) using a wet-transfer system (Bio-Rad). The gel was sandwiched with the membrane between four layers of filter paper (Whatman) and finally two sponges, each layer placed carefully to

avoid trapping air bubbles, and soaked in transfer buffer. Transfer buffer was optimised for transfer of large proteins (20% methanol, 25mM Tris, 150 glycine) and the proteins were blotted at 0.30A with amperage constant for 90 minutes, prevented from overheating using an ice insert, and checked for efficiency of transfer through visualising the transfer of the prestained ladder. The membrane was rinsed briefly in $dH_2O$ and stained with ponceau (0.4% Ponceau S, 3% trichloroacetic acid (TCA)) to confirm transfer of the proteins, and the ponceau removed through repeated washes with $dH_2O$.

### 2.7.4 Detection of proteins

The membrane was blocked in 5% skimmed powdered milk (Marvel) in PBS/0.1% Tween 20 for 1 hour with constant agitation, and then stained with 1:500 anti-TbTor polyclonal antibody overnight at 4˚C. The membrane was then washed three times for 10 minutes in PBS/0.1% Tween 20, before incubation with 1:5000 anti-rabbit-HRP conjugate secondary antibody (Sigma) in 5% milk in PBS/0.01% Tween 20 for 45 minutes at room temperature with constant agitation. The membrane was washed a further three times for 10 minutes in PBS/0.1% Tween 20. The membrane was placed into Bag 'W' and Enhanced chemiluminescence (ECL) substrate (GE Healthcare) was added for exactly one minute. The Bag 'W' was sealed and X-ray film was exposed to the membrane for up to 5 minutes.

## *2.8 Microarray*

The labelling of the cDNA for microarray, quality assessment of RNA samples and microarray hybridisation were all carried out by Alan Ross (Division of Pathway Medicine, University of Edinburgh).

### 2.8.1 Preparation of labeled cDNA

The RNA from cells at different points of differentiation were assessed for quality using a Thermo Scientific Nano-Drop™ 1000 Spectrophotometer to assess the quantity and then quality analysed on an Agilent RNA 6000 Nanochip (lab-on-a-chip), using a 2100 Agilent Bioanalyser. The RNA was used as a template to generate labelled cDNA using the Agilent low input amplification protocol. Briefly, a primer containing poly dT and the T7 polymerase promoter was annealed to 500 ng of polyA+ RNA and reverse transcriptase used to synthesize the first and second

strands of cDNA. Next, cRNA was synthesized from the double stranded cDNA using T7 RNA polymerase as specified by the manufacturer, with cyanine 3 (Cy3) labelled CTP included in the nucleotide pool and incorporated into the cDNA.

### 2.8.2 Microarray hybridization and quality control

Microarray hybridisations were carried out by Alan Ross. 5ug of the Cy3 labelled cRNAs were hybridised to the JCVI *Trypanosoma brucei* microarrays (version 3) for 18 hours at 60°C, washed, and scanned using an excitation wavelength of 532nm and Cy3 detection filter in an Agilent microarray scanner (G2505B). Microarray images were quantified using QuantArray software version 3 (Genomic Solutions).

### 2.8.3 Statistical analysis

All statistical analysis was conducted by Al Ivens (Fios Genomics).

#### 2.8.3.1 Normalisation

The spiked *Arabidopsis thaliana* standards were used to assess the background hybridisation of each microarray and subtracted from the signal intensities for each gene. An average fluorescent intensity was calculated for each gene from the two spots on the slides. The mean intensity of the entire slide was calculated and the slides showing poor fluorescent signals were removed from further analysis. Initially the slides were normalised within arrays to correct for variation in hybridisation over the entire surface of the slide, using print tip loess. The normalised slides were compared to each other using Pearson correlation and the slides showing poor signal intensities were removed from further analysis. The slides were then normalised between arrays using quantile normalisation.

#### 2.8.3.2 Detection of significantly differentially expressed genes

The signal intensity of the individual genes was then compared to that of the genes in each different time point and the difference in intensity was $\log_2$ transformed. The fold change of each gene was then calculated and subjected to a moderated t-test to identify those genes showing statistically significant differential expression. An adjusted p value was calculated for each data point to remove errors arising multiple testing using the Benjamini and Hochberg method (Benjamini and Hochberg, 1995).

The genes that were found to show different levels of expression at a significance level of at least $p<0.05$ were then sorted into 'profiles', through implementing a trinary code on the data when compared to the expression level in stumpy form cells. Those time points showing an increase in expression level above a $\log_2$ of 1 when compared to the level in stumpy form cells were given the value +1, whilst those with a decreased expression level in contrast to stumpy forms of $\log_2$ of 1 were given a value of -1, with changes in expression between these two points assigned a value of 0. The profiles were sorted according to the trinary codes of a gene expression level throughout the time course, with a possibility for 243 different profiles.

# 3 – Microarray Analysis Through Differentiation

Whole genome microarray studies examining differences in life cycle expression patterns in *Trypanosoma brucei* have previously only focussed on cultured monomorphic bloodstream form and procyclic form cells, ignoring the differences that must have arisen in cells that have been laboratory adapted versus wild type field isolates and also genes expressed specifically in the stumpy form that bridges the gap between slender and procyclic forms (Diehl et al., 2002, El-Sayed et al., 2000, Koumandou et al., 2008, Brems et al., 2005). However, an RNA fingerprinting experiment conducted on pleomorphic slender, stumpy and procyclic form cells showed that numerous genes do show a different expression profile between slender and stumpy forms, with two genes seen to have an increased level of expression in stumpy forms (Mathieu-Daude et al., 1998). Similar life cycle microarrays have been carried out on parasites of the *Leishmania* species, one of which explored the transient differences occurring during the differentiation of the promastigote form found in the sandfly to the amastigote form that survives in the mammalian macrophage (Saxena et al., 2007). From this study, a number of genes were found to show temporary up or down regulation, with a small number of genes fluctuating between the promastigote and amastigote form, which had never before been detected due to the omission of the transition phase between the life stages in previous experiments. Thus we were interested in dissecting the expression differences between pleomorphic slender and stumpy forms, as well as changes occurring during the differentiation of the bloodstream form stumpy stage to the procyclic form to identify genes involved in the differentiation process.

Very little is so far known about the genetic control of events involved in differentiation, so our experiment was designed to look at gene expression at key points during differentiation that coincide with known morphological or cytological events; 1 hour at which point the proteins comprising the procyclic surface coat are expressed, 6 hours which marks the loss of the bloodstream form VSG surface coat and should also correspond with the beginning of the re-entry to the cell cycle of the

parasites; 18 hours and 48 hours to distinguish between early and more established procyclic forms. In order to generate data that could be meaningfully subjected to statistical analyses, we used multiple biological replicates, rather than technical repeats, which are not considered to provide useful additional information (Allison et al., 2006). Thus, five stumpy form populations and their subsequent differentiation to 48 hours, and four slender populations were generated for microarray analysis.

### *3.1 Experimental design*

Figure 3.1 shows, in schematic form, the experimental regime used in the assays. Ten mice were immune suppressed through treatment with cyclophosphamide one day prior to injection with *T. brucei brucei* AnTat1.1 parasites. After three days, blood was harvested from four of the mice by cardiac puncture and the slender trypanosomes were purified over DE52 columns using PSG. It is desirable to have a homogeneous population of cells when conducting a microarray experiment (Ness, 2007), therefore to ensure that the slender populations had no stumpy cell contamination, the parasites were isolated when the parasitaemia was below $2x10^7$cells/ml, well beneath the density at which stumpy forms appear. Between 4.6-$8.1x10^7$ trypanosomes were harvested in total from each mouse from the slender infections. The remaining mice were left until day six by which time the parasite infection consisted predominantly of stumpy form cells, and these were similarly purified. Each mouse used to generate stumpy form cells yielded approximately $1x10^8$ parasites. Unfortunately, two of the mouse infections for the stumpy populations resulted in a poor recovery of parasites, and therefore these two samples were pooled following purification; all other mouse infections were analysed separately.

In order to limit any 'cold shock' induced gene expression (Engstler and Boshart, 2004), syringes and PSG were kept at 37˚C and, once purified, the trypanosomes were rapidly placed into pre-warmed HMI-9 medium for one hour at 37˚C to allow them to recover from the purification procedure. Both the slender and stumpy form cells were resuspended into the HMI-9 medium to a density of $1x10^7$ cells/ml. The

entire population was then used to make slides and extract RNA from the slender samples, whilst for the stumpy form populations, only 2ml of each sample was removed to make RNA and slides. The remaining 8ml was diluted 1:4 so that the cell density was $2.5x10^6$ cells/ml into SDM-79 media with 10mM glycerol pre-warmed to 27˚C, and induced to differentiate through addition of 6mM cis-aconitate. 10ml samples were then taken at each time point to generate RNA.

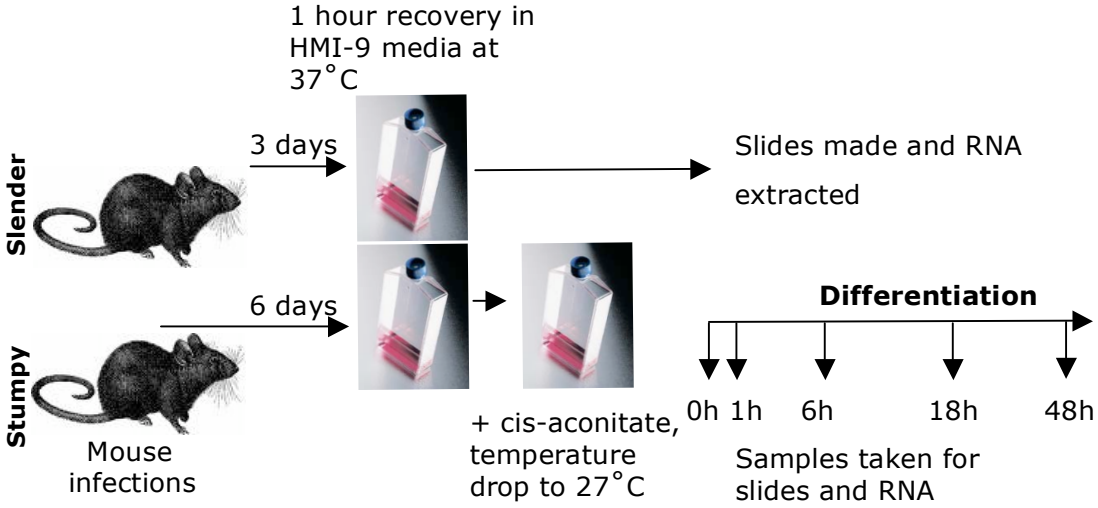



**Figure 3.1. Schematic diagram showing sample generation strategy used for the microarray analysis**. The experimental set up shows how slender form parasites were isolated after three days of infection in a mouse followed by a brief 1 hour recovery in HMI-9 media. Thereafter, cells were fixed and spread onto slides or lysed for RNA extraction. Stumpy cells were isolated after six days of infection, recovered in HMI-9 media then differentiation initiated by temperature reduction to 27˚C and addition of 6mM cis-aconitate. Slides and RNA samples were prepared at 0, 1, 6, 18 and 48 hours after the initiation of differentiation.

Samples from the cells undergoing differentiation were taken at certain time points and cell counts conducted, to ensure that cells had successfully resumed proliferation following differentiation. From Figure 3.2 it is possible that see that bio-replicate #5 shows the poorest growth, and also had the lowest number of cells initially. This is the pooled sample from the combination of the parasites from two mouse infections where neither returned a high yield of trypanosomes. Nonetheless, all of the differentiating samples showed an increase in cell number following differentiation.

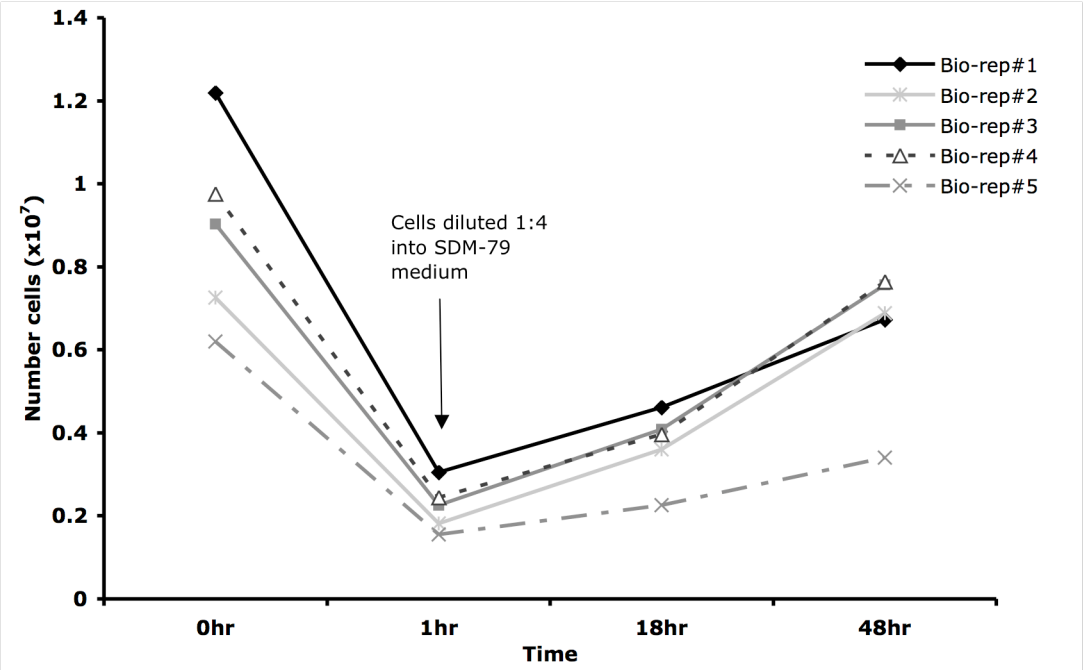



**Figure 3.2. Growth of differentiating cells**. Samples from the cells were taken through differentiation to conduct cell counts to ensure cells were still viable. Bio-replicate#5 is derived from pooled samples from two mice and has the poorest growth, yet there is an increase in cell numbers over the course of differentiation.

As the cells had been demonstrated to be viable and showed proliferation following differentiation, the samples taken for RNA were processed for the microarray.

### *3.2 Quality control of RNA samples*

The quality of RNA used in a microarray hybridisation experiment is the determining factor for whether meaningful results are achieved (Morey et al., 2006). To obtain the highest quality RNA possible, RNA was generated using the Qiagen RNeasy kit, and analysed by gel electrophoresis to ensure it was intact. For our analysis it was also essential that the cells were progressing through differentiation in the expected manner. To this end, extensive cytological analysis of the cells during differentiation was carried out to ensure the expected timecourse of differentiation events was observed.

RNA was prepared from cells in such a manner so as to prevent degradation; cells were pelleted and washed in PSG before being lysed and RNA harvested using an RNeasy kit (Qiagen), with on-column DNase step being used to remove any contaminating gDNA. 5µl aliquots of each of the RNA samples were then loaded onto denaturating formaldehyde gels and the RNA profile visualised after electrophoresis using ethidium bromide to determine the RNA integrity. Figure 3.3 shows that the RNA was intact and gives a measure of relative concentration during the time series. For example, the increased intensity of the RNA at 18 hours after the initiation of differentiation correlates with the cell division of the parasites at some point between 6 and 18 hours as the same volume of RNA was loaded without adjustment made for concentration or derived cell number. This is in keeping with the cell counts taken at 18 and 48 hours through differentiation.

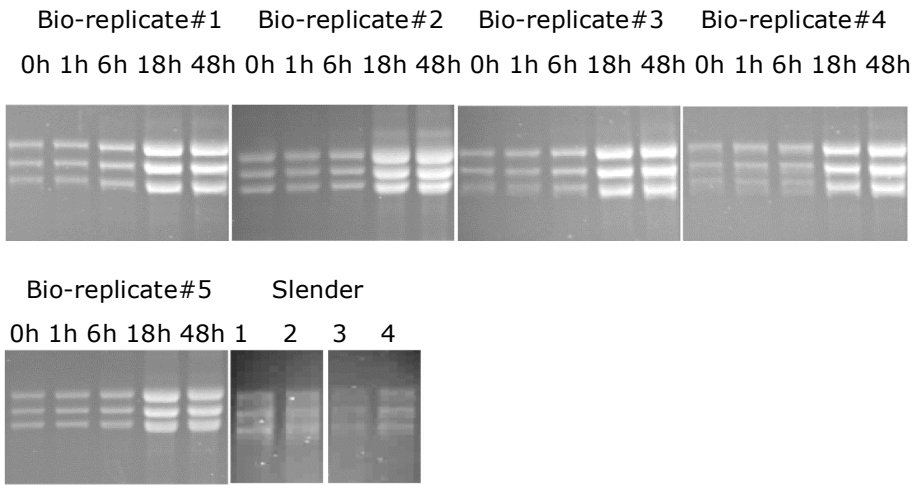



**Figure 3.3. RNA profiles of the samples used for microarray hybridisation visualised by ethidium bromide staining after resolution on denaturating gels**. RNA was generated using a Qiagen RNeasy kit and samples visualised on formaldehyde gels to assess the quality of the RNA. The upper panel shows the rRNA from RNA samples generated from the first four biological replicates of stumpy cells and during their subsequent differentiation; the lower panel shows the fifth stumpy sample and the RNA generated from the slender samples.

In addition to visualisation of the RNA by formaldehyde gel electrophoresis, all samples were also quality controlled using a Thermo Scientific NanoDrop™ 1000 Spectrophotometer to assess the RNA quantity and calculate their 260/280 ratio. It

was found that all samples showed a 260/280 ratio of greater than 2.1; a good indicator that the RNA did not have protein contamination. The limitation of this measurement is that it is unable to distinguish between RNA and any contaminating DNA. However, the possibility of DNA contamination was excluded in later analyses when the same RNA samples were used for cDNA generation for qPCR analysis of array fidelity (Section 4.3). In these experiments, RNA samples without reverse transcriptase failed to generate any template nucleic acid for amplification in the qPCR reaction, demonstrating an absence of contaminating DNA. Hence, the RNA generated was considered suitable for microarray hybridisation.

## *3.3 Confirmation that cells were undergoing differentiation*

Prior to using the RNA generated for microarray experiments, it was first necessary to analyse the cells the RNA had been derived from to ensure the cells had differentiated as expected. Few molecular markers of stumpy form cells have been established and thus cells are usually designated as stumpy based upon morphological traits. Fortunately, the kinetics of stumpy accumulation of AnTat1.1 cells in vivo is well established and predictable, hence we estimated that harvesting bloodstream form cells from mouse infections at $>2x10^8$cells/ml would result in populations comprising at least 80% stumpy form cells; a density that was reached after six days of infection. Morphological observation of the harvested cells confirmed this hypothesis. Evidence was then required that the stumpy form cells had differentiated and this was generated using analysis of the surface antigens on the cells, cell cycle re-entry and kinetoplast repositioning.

### 3.3.1 Surface coat exchange

Over the course of differentiation, the parasite must switch surface coats from the variant surface antigen (VSG) coat expressed during the bloodstream stage to allow protection from the host immune system, to the procyclin surface coat present on procyclic form parasites, thought to be protective against proteases in the insect midgut (Acosta-Serrano et al., 2001). In *T. brucei* AnTat1.1 cells, the major VSG expressed is VSG AnTat1.1 and antibodies are available against both this antigen and EP proteins. Therefore, slides of cells throughout differentiation were co-labelled

with antibodies against VSG AnTat1.1 and EP to track antigen coat switching (Figure 3.4).

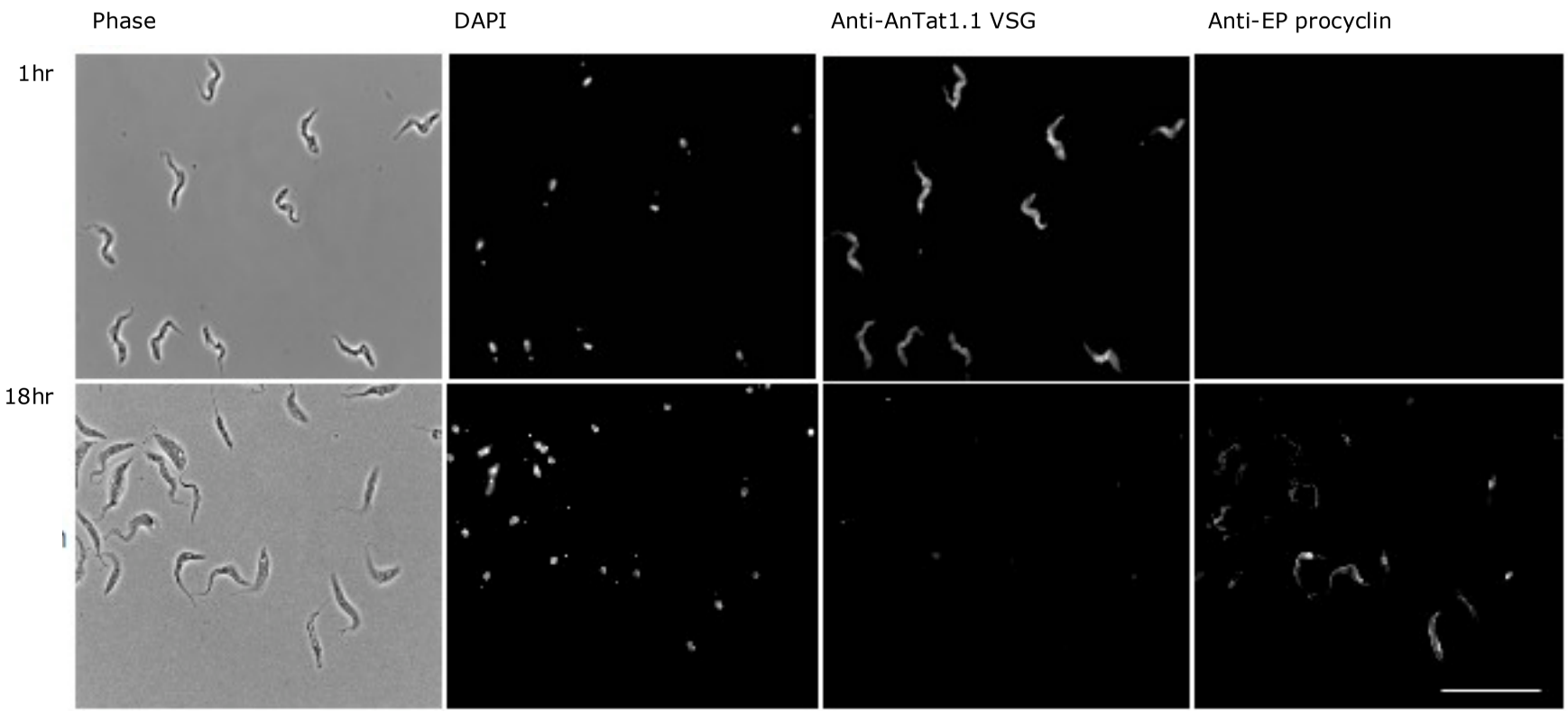



**Figure 3.4. Immunofluorescent analysis of surface antigens during differentiation**. (A) Cells at 1 hour through differentiation express solely the VSG surface coat. (B) At 18 hours through differentiation, cells have re-entered the cells cycle and are expressing predominantly EP procyclin. Line represents 50μm.

Morphological analysis of the slides alone was sufficient to confirm the cells had differentiated, as stumpy form and procyclic form cells are easily distinguishable purely by shape and structure (Compare the upper and lower panels of the phase images in Figure 3.4). This was verified through in-depth analysis of the surface staining of the slides for each biological replicate at each time point, which demonstrated that antigen switching had clearly taken place. Therefore, to track the kinetics of the switching the presence of VSG and EP on the cell surface was scored for each biological replicate over the time course. Coat exchange is an early process during differentiation, and in *T. brucei* AnTat1.1, EP has been detected on the surface of cells within the first three hours of differentiation (Sbicego et al., 1999). Figure 3.5A shows the gain of EP is rapid, with an average of 87% (range 79-94%) of cells expressing this protein on the surface by 6 hours into differentiation, with 100% of cells expressing exclusively EP at 48 hours. VSG loss is initiated less quickly (Figure 3.5B), resulting in a window in which cells express both EP and VSG on their surfaces, with over 90% of cells still expressing the VSG antigen at 6 hours, although nearly all cells have successfully shed their VSG coat by 18 hours.

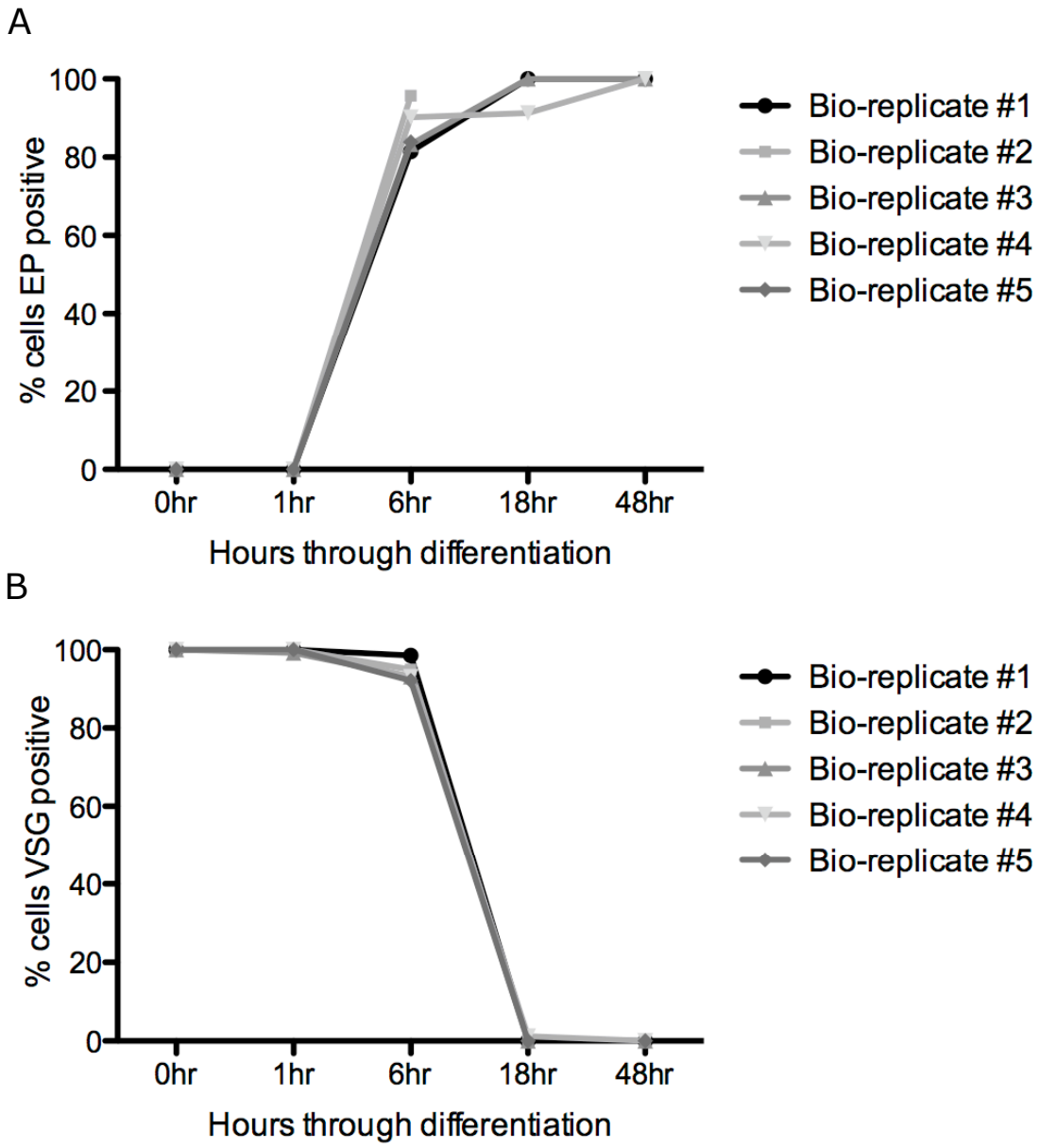



**Figure 3.5. Cell surface protein of differentiating cells**. Methanol fixed cells were dual stained with antibodies against the cell surface proteins EP and VSG AnTat1.1 and scored as either positive or negative for each. At least 250 cells were counted for each slide, and all five biological replicates were scored for time points 0, 1 and 6 hours, but only replicates #1, 3 and 4 were examined at 18 and 48 hours due to slide quality. (A) EP is expressed on the cell surface almost completely by 6 hours whilst (B) VSG loss is completed by 18 hours post differentiation.

EP procyclin is introduced into differentiating cells through incorporation into the VSG surface coat already present (Roditi et al., 1989), and at 6 hours into

differentiation, cells were seen to express both antigens equally. The antigen switch from VSG to EP over the time course is a good indicator of differentiation. Whilst EP surface expression can be triggered through coldshock in the absence of differentiation (Engstler and Boshart, 2004), when coupled with the corresponding loss of VSG, this data further confirmed that cells had differentiated.

### 3.3.2 Kinetoplast repositioning

Once harvested, it was also possible to monitor differentiation using the distance between the kinetoplast and the nucleus as a marker. Over the course of differentiation, through a combination of migration of the kinetoplast towards the nucleus and outgrowth of microtubules at the posterior end of the cell, the kinetoplast is seen to move from the extreme posterior of the cell in bloodstream form slender and stumpy forms, closer towards the nucleus. The extent of the kinetoplast repositioning has been measured previously and showed that the kinetoplast-posterior distance increased from approximately 1μm in bloodstream form cells, to 4μm in procyclic populations, although the reason for this is unclear (Matthews et al., 1995).

To monitor this parameter, kinetoplast-posterior distance was measured for the entire time course of one biological replicate and then the 0 hour and 6 hour time points of all replicates, to establish whether all biological replicates were behaving in a similar manner. Figure 3.6A shows schematically how the measurements were made, with a line drawn from the centre of the kinetoplast to very posterior and measured using ImageJ 64. For each time point, a minimum of 100 cells were measured, and in later time points, once cells had resumed division and thus had two kinetoplasts, only cells with one kinetoplast were examined.

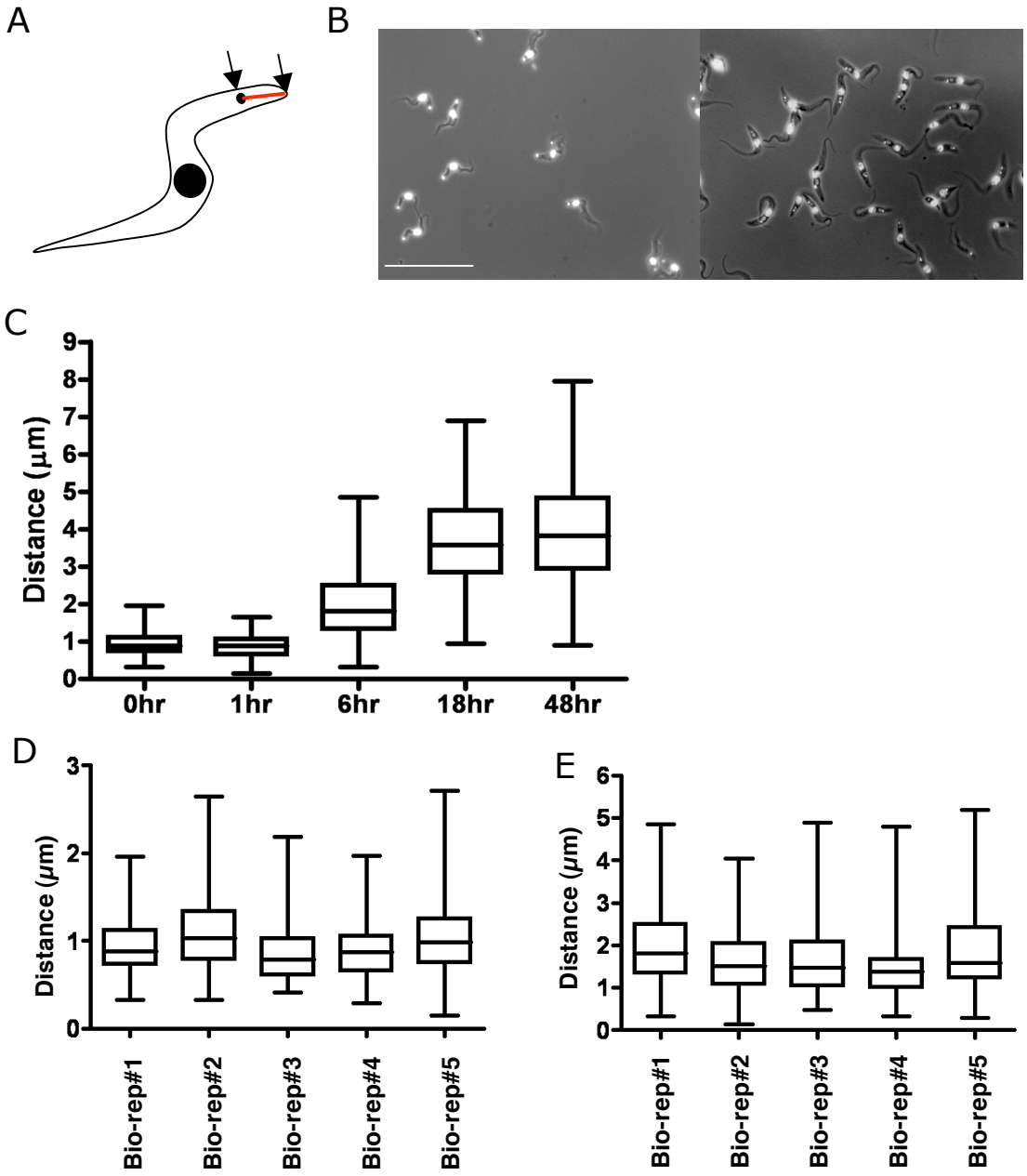



**Figure 3.6. Analysis of the kinetoplast-posterior distance during differentiation.** (A) Schematic diagram illustrating the measurement made. A line was drawn from the centre of the kinetoplast to the extreme posterior of the cell and distance calculated using ImageJ 64 software. (B) DAPI-stained parasites from bio-replicate #1 at 0 hour (left panel) and 48 hours (right panel). Line represents 50µm. Box and whiskers graphs plotting the mean and upper and lower quartiles of kinetoplast-posterior distance over the differentiation time course for (C) bio-replicate #1, and (D,E) time points 0 and 6 hours plotted for each bio-replicate. For each sample at least 100 cells were counted.

Figure 3.6C shows that there was a clear increase in distance between the kinetoplast and the posterior of the cells throughout the time course, initiating between 1 and 6 hours into differentiation, whereby the distance in stumpy form cells was around 1µm, increasing to 1.5µm at 6 hours, up to 4µm at 18 and 48 hours, which proved highly reproducible between biological replicates (Figure 3.6D and E). Larger variations in size were seen for the later time points, as can been seen by the extent of the whiskers at the 18 and 48 hour time points, but this is in keeping with previous work (Matthews et al., 1995). From Figure 3.6B it is possible to see that stumpy form cells position the kinetoplast at the extreme posterior of the cells, whilst at 48 hours there was a clear movement of the kinetoplast towards the nucleus. Previous work has also examined the distance between the nucleus and the kinetoplast (Matthews et al., 1995) which was found to decrease throughout differentiation, as can been seen through comparison of the two panels in Figure 3.6B. The changes in morphology and organelle positioning from microscopy confirms that cells differentiated in the expected manner.

### 3.3.3 Cell cycle analysis of differentiating cells

Due to the quiescent nature of stumpy cells, which are G0 arrested, cell cycle analysis can be used to plot the re-entry into the cell cycle, and as such can be wielded as a tool to confirm that cells are progressing through differentiation. Therefore, to investigate cell cycle stage, parasites were methanol fixed and the cellular DNA was visualised using DAPI staining (see Figure 3.7) and analysis carried out of the number of kinetoplasts and nuclei in each cell, with at least 250 cells scored from each slide. DNA replication of the nucleus and kinetoplast of trypanosomes occurs in a tightly co-ordinated manner such that cells with 1 kinetoplast and 1 nucleus (1K1N) must be either in G0/G1 or S phase (See Figure 1.4). Similarly, cells with a 2K1N configuration can be identified as in G2 or early mitosis, whereas 2K2N cells are post-mitotic. During normal logarithmic growth *in vivo*, the majority of cells display the 1K1N configuration as demonstrated in all slender populations, with only approximately 15% (10-19%) cells 2K1N and around 10% (7-13%) 2K2N (Figure 3.7). Stumpy form cells, by comparison were almost entirely in the 1K1N configuration (>99%), confirming that these populations were growth arrested. Over the course of differentiation it was possible to see the cells in

each biological replicate gradually resumed cell cycle progression, at 18 hours into differentiation between 3-18% of cells were in the 2K1N configuration and a further 0.5-4% were 2K2N. The 48 hour time point showed similar proportions of cells. Thus the cells in each biological replicate re-entered the cell cycle, with approximately the same kinetics between the samples, despite being derived from separate mouse infections.

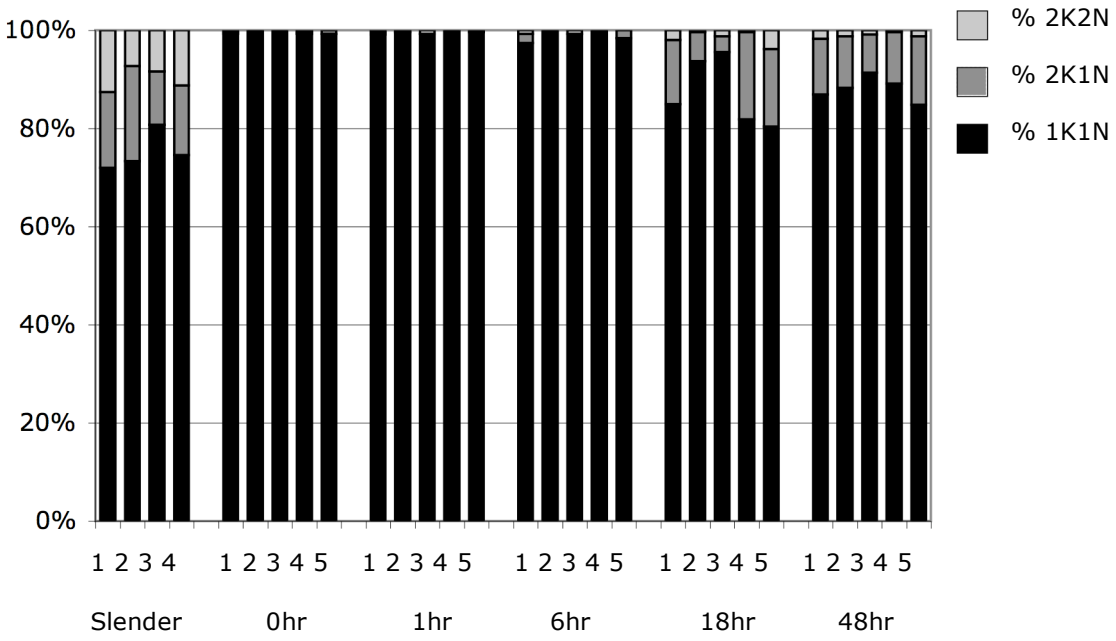



**Figure 3.7. Karyotype analysis of cells throughout differentiation**. Nuclear and kinetoplast DNA was visualised through DAPI staining of cells on each bio-replicate over the course of differentiation. At least 250 cells were counted for each sample. Proliferating slender cells had a largely 1K1N karyotype, whilst stumpy cells all showed the 1K1N configuration, consistent with G0 arrest. After 6 hours the cells re-entered the cell cycle with 2K1N and 2K2N cells present at 18 and 48 hours.

Previous tracking of cell configurations through differentiation has shown predominantly 1K1N cells with a sudden appearance of 2K1N cells at 10 hours into differentiation, proceeded shortly afterwards by a lesser increase in 2K2N cells (Matthews and Gull, 1994). Therefore, it appears that we have missed the point of cell cycle re-entry in our choice of time points, and somewhere between the 6-18 hour timepoint an abrupt increase in 2K1N and 2K2N cells would have been seen.

Thus the karyotypes at the 18 and 48 hour timepoints are more representative of the proportions seen in normal proliferative cells.

All together, these results demonstrated that the RNA was of a good enough quality to be used for the microarray hybridisations, and analysis of the cells from which the RNA was derived assured us that the cells had progressed through differentiation in the expected manner. The RNA samples could therefore be hybridised to the microarrays.

### *3.4 Hybridisation of RNA to microarray slides*

Hybridisations were carried out using JCVI *Trypanosoma brucei* v3 glass arrays, which comprise 8594 *T. brucei* genes, with a single 70mer oligonucleotide target for each gene spotted in duplicate within the same grid. Additionally, the slides contain 500 control 70mers derived from genes from *Arabidopsis thaliana*, with known amounts of labelled cDNA probes from this organism 'spiked' into each test sample in order to calculate a reference signal intensity that can be compared against all arrays in an experiment. Further "obsolete" 70mers are also present to calculate background hybridisation; these sequences do not have 100% homology to any *T. brucei* genes, thus any fluorescence signal from these loci are the result of non-specific hybridisation and can thus be used as a measure of background signal intensity. The calculated background signal is then subtracted from the overall intensity of each array.

The microarray hybridisations were carried out by Dr Alan Ross in the group of Professor Peter Ghazals at the Division of Pathway Medicine, University of Edinburgh, with quality control conducted by Fios Genomics. Single colour arrays were used in this analysis rather than the more widely used competitive, two-colour arrays. Such one-colour arrays have been shown to be as sensitive and reproducible as two-colour arrays (Patterson et al., 2006) but without the issue of dye bias. Therefore technical replicates involving dye swap experiments are not required for these arrays, allowing more biological repeats to be conducted instead. The

comparisons between the different samples were then carried out after all slides had been hybridised and normalised.

As the microarray examines differences in gene expression level, only mRNA was required, which was uniquely amplified through exploitation of the polyA+ tail, which is present on mRNA but not rRNA. Using the T7 promoter for generation of antisense RNA results in linear amplification, rather than exponential amplification as is the case with PCR, thus the relative proportion of the different mRNAs is maintained (Phillips and Eberwine, 1996). Hence, cyanine 3 (Cy3) labelled cDNA was generated from the total RNA samples as is shown in Figure 3.8. A primer containing poly dT and the T7 polymerase promoter was annealed to the poly A+ RNA, which was then amplified using reverse transcriptase. The Cy3 label was incorporated into the samples via the presence of labelled CTP in the reaction mix when synthesising cRNA from the double stranded cDNA. The labelled antisense cRNA samples were then hybridised to individual slides and scanned.

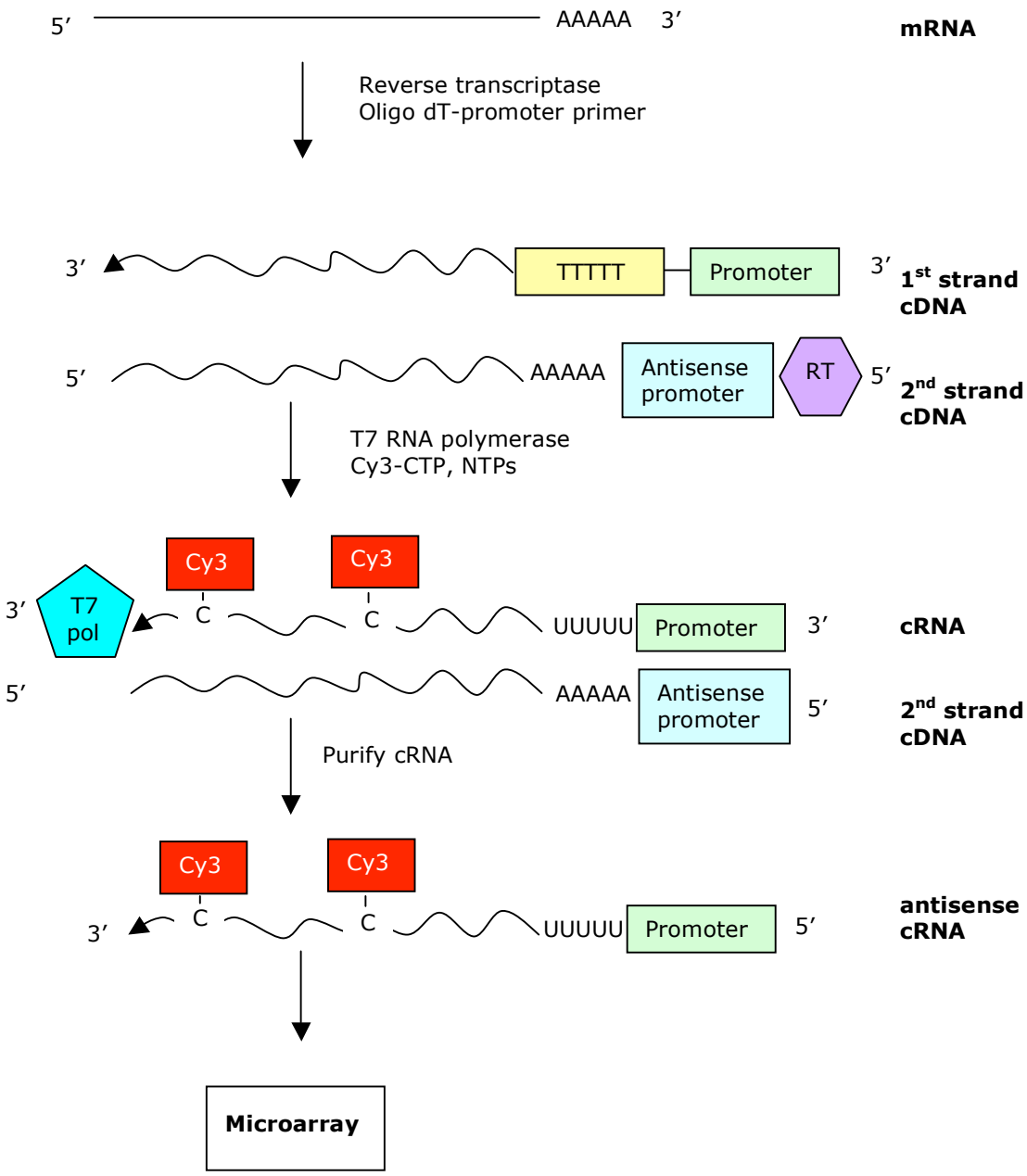



**Figure 3.8. Schematic diagram showing Cy3 incorporation into cRNA for hybridisation onto microarray slides**. In order to amplify the mRNA, a T7 polymerase promoter and oligo dT sequence were annealed onto the poly A+RNA and reverse transcriptase added to produce double stranded cDNA. This was then amplified using the T7 RNA polymerase to make the cRNA using Cy3-labeled CTP, which was incorporated into the cRNA product. The cRNA was then purified prior to hybridisation. (From the Agilent low RNA input fluorescent linear amplification kit protocol).

The hybridised arrays were then scanned using an excitation wavelength of 532 nm and Cy3 detection filter on an Agilent microarray scanner and images quantified using QuantArray software version 3.

### *3.5 Quality control of hybridisations*

The method of extracting genuine data from the raw intensity values generated from the microarray hybridisations is summarised in Figure 3.9. The background fluorescence generated from non-specific hybridisation was calculated from the obsolete and blank spots and subtracted from the intensity values for each transcript. The spiked-in *Arabodopsis* standards were used to adjust the overall intensities from each slide such that the standards were equal across all slides. Each array was then subjected to within-slide normalisation, which accounted for the errors introduced from the printing procedure and for spatial effects. Finally the slides were normalised between arrays to allow comparisons with the other time points and biological replicates. All data processing steps were carried out by Fios Genomics.

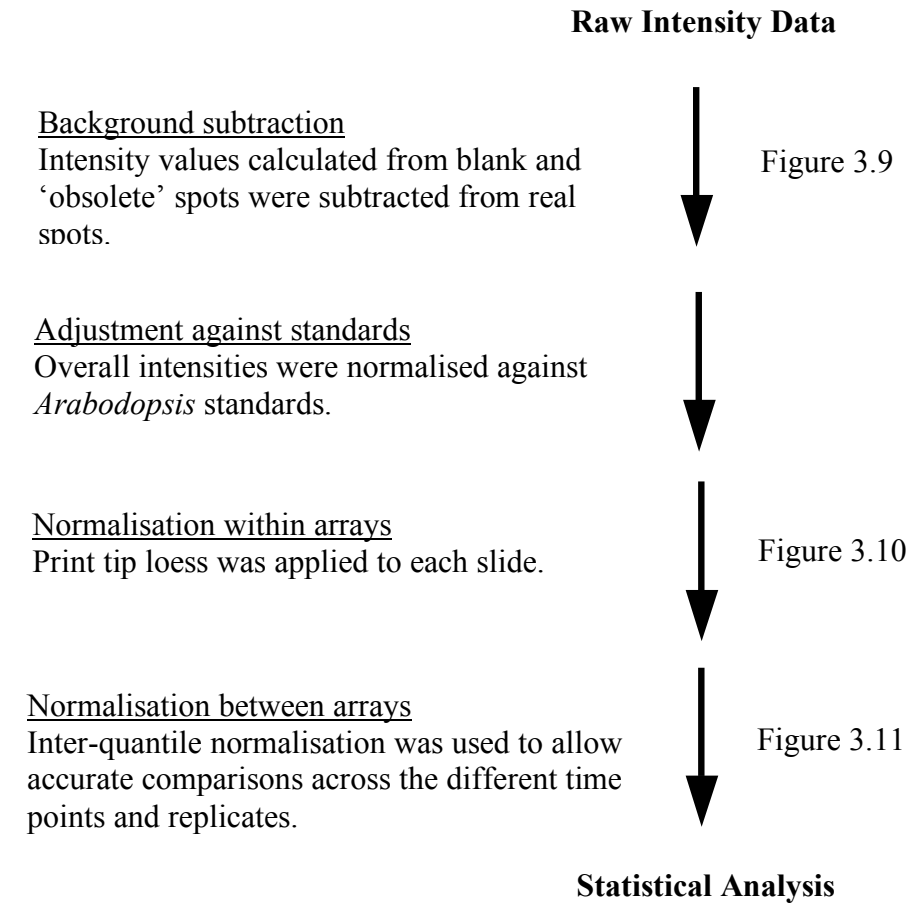



**Figure 3.9. Flow diagram illustrating the data processing steps applied to the raw microarray data.** The sequence of background adjustment and normalisation of the microarray are presented alongside the figures to which they relate.

Before any data analysis was initiated, the raw data taken from the microarrays was first 'cleaned' through background subtraction, and transformed into logarithms to base 2. Transforming the data in this manner allows a better spread of the data than is achieved using raw values, hence making it more suitable for analysis (Stekel, 2003). Thus, each integer on the log scale represents a doubling of the raw data such that a $\log_2$ value of 8 corresponds to a raw intensity of 256 and a $\log_2$ value of 9 corresponds to a raw intensity value of 512 (Stekel, 2003).

In order to gauge the average fluorescent intensity of each slide, the background intensity was subtracted from the overall slide intensity. Background signal occurs through non-specific hybridisation and was calculated using a combination of empty spots and probes without 100% homology to any transcripts on the array (obsolete spots). The 4048 empty spots on the arrays have no oligo probes, so any fluorescence detected from these spots is entirely due to background. However, the use of empty spots alone for background calculation has been criticised because the absence of DNA at a spot is not reflective of the situation for the real spots (Draghici, 2003), therefore the 'obsolete' probes are also included, which are more representative of the real spots. Following background subtraction, control probes from the *Arabidopsis thaliana* standards were used as a quality control method to assess the efficiency of the hybridisations. These probes are designed against 500 targets known to give a range of fluorescent intensities entirely independently of the experimental procedure (PFGR), thus the success of the hybridisation can be inferred from the deviation of the signal recorded from these spots from the expected fluorescence, and the individual arrays can all be normalised against these standards. The mean fluorescent intensity of the entire slide was then calculated for each of the arrays and drawn as a box plot along with the median intensity of the fluorescence for the entire set of slides (Figure 3.10). The box plots show the $\log_2$ intensities of each slide in order to determine whether the fluorescent intensity of any of the slides strayed greatly from the median. These graphs show the median $\log_2$ intensity of all spots for each slide at the centre of the box, with the edges showing the upper and lower quartiles, these encompassing the 50% of spots with a fluorescent intensity

surrounding the median intensity. The whiskers represent the spots with maximum and minimum intensities from each slide.

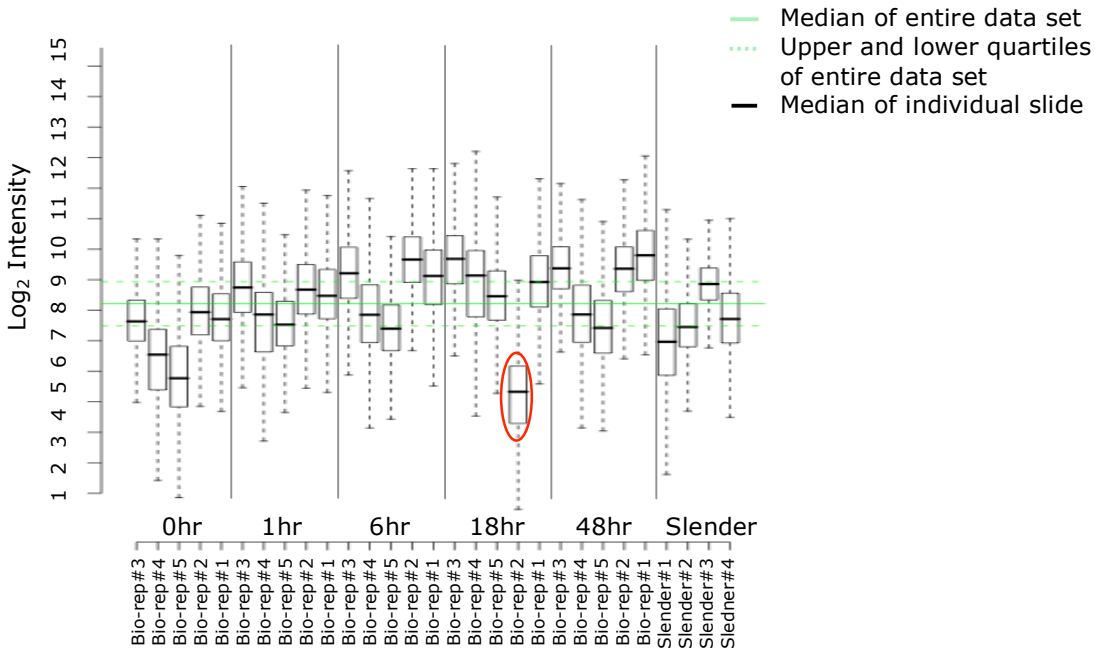



**Figure 3.10. Box plots of overall intensities of individual arrays**. The box plots were generated from un-normalised data with background subtracted. The solid green line represents the median of the entire data set, and black lines the median of the individual sample. Bio-replicate#2 18h was withdrawn from the later comparisons due to its very low intensity.

In Figure 3.10 it is possible to see that the slide for biological-replicate#2 at 18 hours through differentiation (circled) has a much lower median fluorescence intensity compared to all other slides and was thus deemed to be of too poor a quality to retain in the data analysis.

### *3.6 Normalisation within arrays*

Variations seen between different probes on the arrays might arise either due to biological or technical variations. In order to eliminate the variations occurring due to technical discrepancies, slides were normalised within, and between arrays. Within-array normalisation was accomplished using a version of loess (locally weighted polynomial regression), print tip loess. This is the preferred method to adjust for variations arising from the physical printing of the oligos onto the arrays (Smyth and Speed, 2003). A flaw of printed microarrays is that the tips used to

physically spot the probes onto the slides can become worn and so oligos situated in areas of the array spotted from tips that are slightly worn or damaged in some manner will have bound to the slide with varying efficiency. Print tip loess will correct for inaccuracies arising for these mechanical abnormalities and also works as a method to normalise locational effects due to spatial blocks on the slides being printed as a block. Similarities between the background-subtracted data, normalised within arrays, for the individual slides were calculated using Pearson correlation. The results were then represented as heatmaps in order to better visualise the distribution of intensity and to identify any other substandard slides (Figure 3.11).

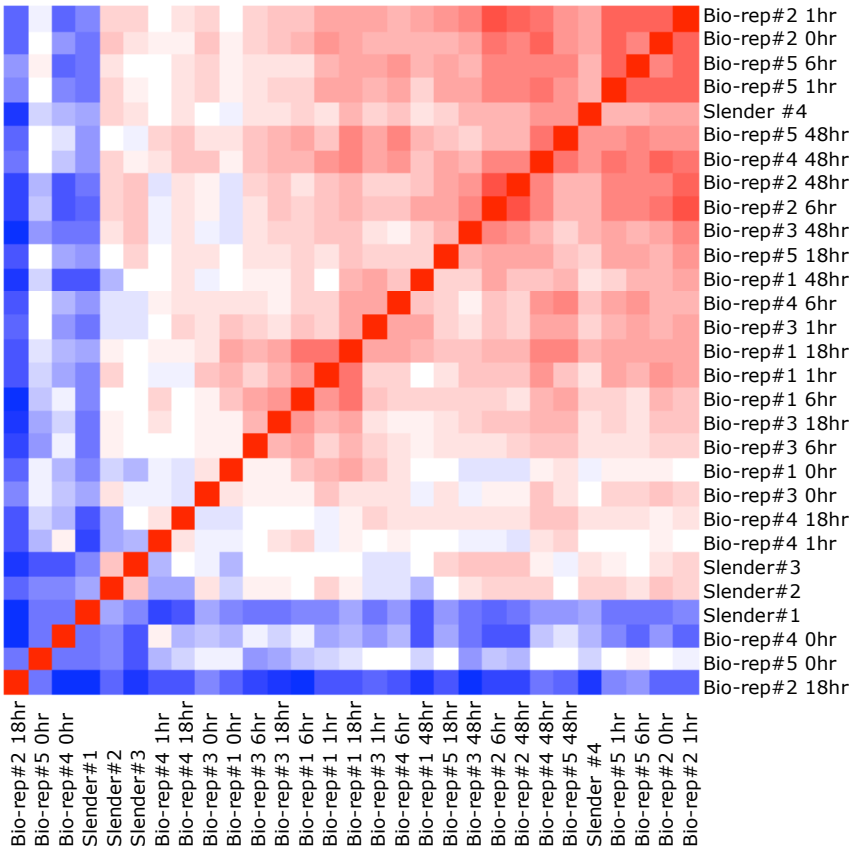



**Figure 3.11. Correlation plot, normalised within arrays**. Pearson correlations were calculated of the average intensity for each array with respect to each other array and the results are presented as a heatmap, such that blue shows a negative correlation and red a positive correlation. Four of the arrays; bio-replicate#4 0 hour, bio-replicate#5 0 hour, bio-replicate#2 18 hour and slender#1, show little correlation with the other arrays so were removed from further analysis.

Pearson correlations were calculated for all arrays to judge the degree of similarity in average intensity between each of the arrays. A value of +1 (red) is representative of a perfect correlation, thus each slide when compared against itself is red, whilst -1 (blue) is a negative correlation. Looking at the heatmap it was clear that, in addition to bio-replicate#2 18 hour which had been shown to be of unsatisfactory quality from the box plot, three other arrays were also unsuitable. Bio-replicate#4 0 hour, bio-replicate#5 0 hour and slender#1 all showed largely negative correlation with the other arrays in the analysis. All four of the unsuitable arrays were removed from further analysis, as the poor hybridisation efficiency of these slides could skew results gained from normalisation.

### *3.7 Normalisation between arrays*

The remaining slides were then normalised between arrays to remove variations arising from non-biological differences (Smyth and Speed, 2003). As the individual arrays were all identical it was then possible to use quantile normalisation to smooth the distribution of intensities between arrays to allow between-array normalisation. Following all the normalisation, large differences occurring for technical reasons ought to have been removed and thus all subsequent data should represent true biological differences.

Figure 3.12 shows density plots of three of the time points before and after normalisation between arrays. The individual slides were all normalised within array prior to global normalisation between arrays. The 6 hour time point is a good example of a set of arrays with large variations in magnitude of intensity, whilst the three stumpy form samples that remained after quality control all show quite similar profiles even before normalisation between arrays. Following normalisation, all biological replicates within and between time points showed the same profile, allowing meaningful comparisons to then be done without bias of data from slides with higher average intensities.

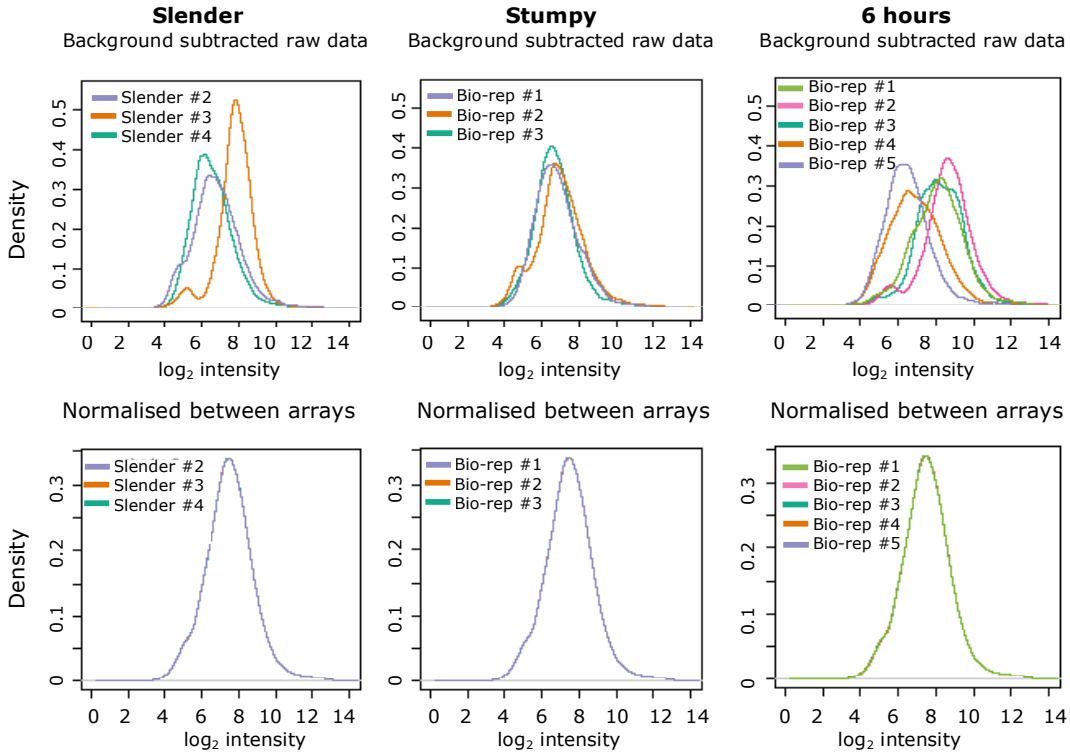



**Figure 3.12. Density plots of slides normalised between arrays**. As an example, density plots of three of the time points are shown after (top panel) background subtracted raw data, and (lower panel) after slides had been normalised between arrays. Only the arrays that had been shown to have reasonable intensities were considered, which is why only three biological replicates were included in between array normalisation of the slender and stumpy form arrays.

Once the data were all normalised so that each array showed the same distribution of intensity, it was then possible to conduct the comparisons between the different time points.

### *3.8 Summary*

Based on the quality control performed on the RNA prior to hybridisation onto microarray slides and then subsequent analysis resulting in the removal of four arrays from further analysis, we felt we could have confidence in interpreting the microarray as relevant biological changes. The extensive analysis conducted to ensure that the cells used to generate the RNA for the microarray were differentiating as expected was also reassuring that the microarray data was certain to be representative of genuinely differentiating cells. The next step was to carry out a

statistical analysis to examine differentially regulated genes, although prior to this the average was taken of the duplicate spots on the arrays and then the averages were calculated for each oligo across all arrays of the same time point.

# 4 - Microarray Validation

The previous chapter demonstrated that the cells used to generate RNA for the microarray analysis were differentiating efficiently and on the expected timescales and showed that the hybridisations had been carried out satisfactorily. The next step required was to validate the results from the microarray. Therefore, stringent statistical analyses were carried out to identify differentially regulated genes and these were compared against reference genes of known expression profile and through alternative methods of detecting expression level. This chapter presents the results of this analysis.

## *4.1 Statistical analysis*

Once the arrays for the individual samples had been normalised (see Sections 3.8-3.9), all samples were then compared to each other and analysed for significance using a moderated t-test. The moderated t-test is an extension of the original t-test, tailored to make it more appropriate for microarray analysis; the standard error for each gene is moderated based on the standard error for every other gene analysed, resulting in a higher level of reliability due to the smoothing of the standard errors (Smyth, 2005). The data were log2 transformed to make interpretation easier through conferring equal weighting to all values; thus a log2 value of 1 represents a doubling and a value of -1 means the signal is halved. The fold change of each gene was then calculated for each time point and the moderated t-test was used to assess the probability that each gene was significantly differentially expressed. At p value <0.001 (0.1%), 407 non-redundant genes were found to be differentially expressed. Over 40% of the genes in the list were annotated as 'hypothetical conserved' genes, a further 6% were 'hypothetical' and 9% were 'hypothetical unlikely' transcripts. In-depth analysis of these genes found to be most significantly differentially expressed was encouraging. Many of the transcripts present in this group had had their expression profiles previously investigated and the microarray data showed a strong level of agreement with the published literature (Table 4.1).

| Gene | Expression profile | Reference |
|---|---|---|
| PSSA2 | Increased expression in procyclic form cells | (Jackson et al., 1993) |
| Hexose transporter 1 | Increased expression in bloodstream form cells | (Bringaud and Baltz, 1993) |
| ESAG6 | Increased expression in bloodstream form cells | (Hobbs and Boothroyd, 1990) |
| META1 | Increased expression in metacyclic promastigotes form *Leishmania*<br>Confirmed increased expression in procyclic form *T. brucei* | (Ramos et al., 2004)<br><br>(Diehl et al., 2002)<br>(Queiroz et al., 2009) |

**Table 4.1. Confirmation of the expression profiles of previously published developmentally regulated genes with the microarray data.** The published information relating to these genes refers to expression profile detected between monomorphic bloodstream form cells with cultured procyclic form cells.

Table 4.1 lists a few genes with experimentally determined differential expression that correlated well with the microarray data. Other genes with well-characterised expression profiles are analysed in more detail below.

Figure 4.1 shows a heatmap of all the genes differentially expressed at a significance level of 0.1%, colour coded to illustrate their expression levels at each time point when compared against expression in stumpy form cells. From this heatmap it is possible to see that the vast majority of differences were between the slender group and all others, with a large group of genes strongly up-regulated in slender forms only (dark red lines), whereas a smaller group comprised genes down-regulated only in the slender forms (dark blue lines). These observations were anticipated, because the slender cells were derived from independent mouse infections to the time-course samples. Furthermore, the different biological slender replicates also showed a great deal of variation between themselves. In contrast, each of the separate biological replicate samples used in the time course series were derived from a single infection, and so would be expected to show more similarity to each other. The processes involved in cell differentiation from the slender form into the stumpy form are not known; however based upon the growth arrest and altered morphology of stumpy forms, a correspondingly large alteration in expression profile is not unexpected. Many of the genes at this level of significance were known to be required for specific

life cycle stages and further validation has been conducted largely on the genes falling into this list.

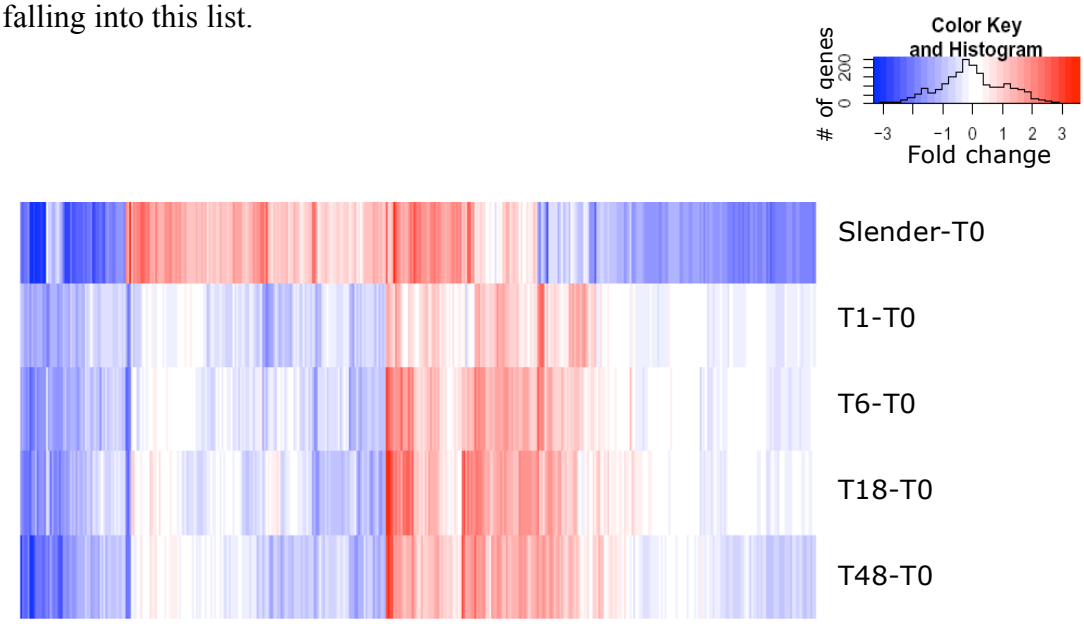



**Figure 4.1. Heatmap of the genes differentially expressed at the 0.1% level**. The heatmap shows the expression profiles of all the genes that were differentially expressed significantly to a p value of 0.001 with a minimum 2-fold change between the level in stumpy form cells with the level in at least one other time point. Each vertical line represents one gene, with the expression colour coded such that blue represents down-regulation in comparison to stumpy, red is up-regulated and white is unchanged. The key contains a histogram showing the number of genes falling into each profile.

These 407 highly significant genes represent less than 5% of the entire genome, so whilst all genes in this list can confidently be considered differentially expressed, numerous other potentially regulated genes will have been eliminated using such a stringent cut-off. At a p value <0.05 (5% significance), 3396 non-redundant genes are identified as being differentially expressed between all comparisons. However when the comparisons were narrowed down to only genes showing a minimum 2-fold change between the level in stumpy form cells and at least one other time point, 2166 genes remained differentially expressed. In order to better dissect the differential expression of these genes in the microarray, a trinary code was assigned to each gene such that for each time point relative to stumpy form, the expression was deemed to be either unchanged (0), increased (1) or decreased (-1). For a gene

to be classed as significantly differentially expressed, the difference in expression in at least one of the comparisons within the time course had to be significant at 5% and the cut off for a gene to be considered increased or decreased was set to a greater than 2-fold change. Therefore a transcript with the profile -1-1-1-1-1 would be down-regulated at least 2-fold at all time points in comparison with stumpy forms and is thus enriched in stumpy cells, whilst the 00000 profile would represent genes that remain unchanged over the course of differentiation. For each of the five comparisons a total of 243 distinct profiles were possible using the trinary code. However, only 74 different profiles were observed with 15 of these profiles only being represented by a single gene. The observation that most genes followed a limited set of expression profiles is a good indicator of the quality of the microarray data as, had we seen multiple genes with wildly fluctuating profiles, it would be unlikely that these represented genuine biological events.

Files can be downloaded comprising spreadsheets listing the 407 transcripts differentially expressed at the 0.1% level. Lists of all genes differentially expressed between slender and stumpy forms are also available at both the 5% and 10% significance level. In addition, tables for each pairwise combination are also available, including results of the statistical analysis. All information can be accessed from http://www.biomedcentral.com/1471-2164/10/427/additional/. A more user-friendly database is due to be launched at TriTrypDB, where it is possible for the researcher to choose the time points to compare, the cut-off level for fold change and the option to select only transcripts up or down regulated, or both.

## *4.2 Validation of microarray against published data*

### 4.2.1 Phosphoglycerate kinase genes

Validation of the microarray was first conducted through comparison of the microarray data against genes shown to be differentially expressed in the literature, although almost all of the published work concentrates on the differences between monomorphic cultured bloodstream form and procyclic forms only, with very few including expression analysis of stumpy forms. In *T. brucei*, the three

phosphoglycerate kinase genes have previously been shown to be differentially regulated between bloodstream form and procyclic forms (Gibson et al., 1988). The different genes code for the bloodstream form-specific PGKC that localises to the glycosome, the procyclic form-specific protein PCKB, present in the cytosol and PGKA, which is expressed at similar levels in both life stages, albeit at low levels, and also localises to the glycosome (Alexander and Parsons, 1991).

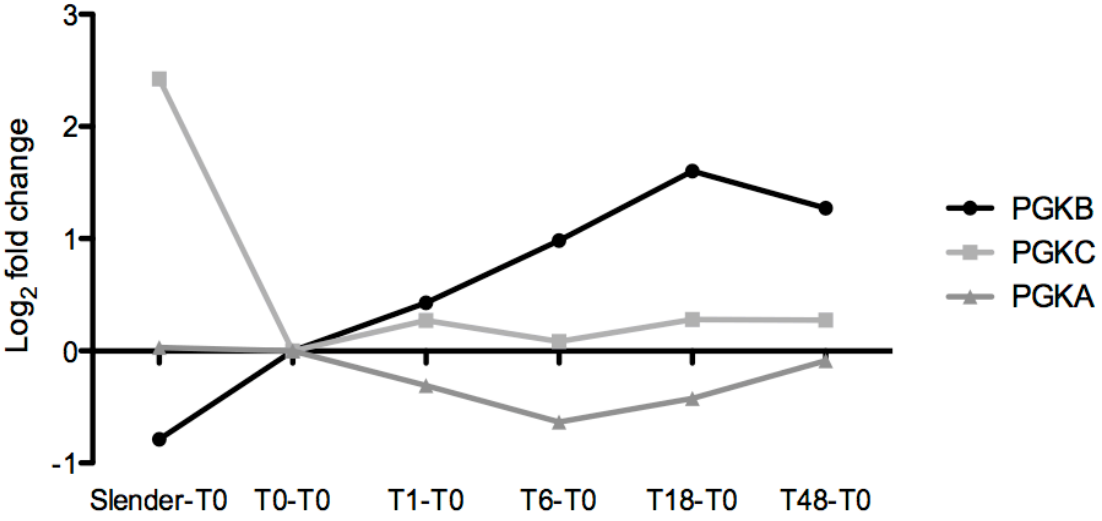



**Figure 4.2. Expression profiles of the three phosphoglycerate kinase determined by microarray**. The bloodstream form specific transcript PGKC is expressed at high levels in slender form cells but not in stumpy form cells or the rest of the differentiation time course. PGKB shows increased expression as cells go through differentiation, as would be expected from a procyclic form specific transcript. PGKA has a largely stable expression profile throughout all time points of differentiation.

Figure 4.2 shows the expression profiles of PGKA, PGKB and PGKC, confirming previous observations (Gibson et al., 1988). All genes show the expected profile in the microarray analysis, with PGKC being expressed predominantly in bloodstream form cells, with a clear drop in expression in stumpy form cells in comparison to the slender forms, whilst PGKB shows a more gradual increase across differentiation, to peak expression at 48 hours, as is expected of a gene that functions specifically in procyclic form cells. Finally, also as expected, PGKA has a non-fluctuating expression profile across the time course.

### 4.2.2 Genes involved in proliferation

As stumpy forms are growth arrested, it was possible to validate the microarray through analysis of the expression of genes required for active proliferation, which would be expected to be lowest in stumpy form cells. The expression profiles of the histone array were examined as a marker of cellular proliferation. Figure 4.3 shows that there is a rapid drop in expression of the histone H2A and H2B array and histone 4 in stumpy forms, which gradually recover to peak between 6-18 hours, as cells are actively dividing once more, and then maintained at a slightly lower level at 48 hours. This is consistent with previous work examining differentially expressed transcripts during differentiation where levels of H2B were seen to peak between 9-12 hours into synchronous differentiation (Matthews and Gull, 1998, Garcia-Salcedo et al., 1999). Histone H3 follows a similar profile but at much lower expression levels with expression not resuming after the drop in stumpy forms until 6 hours into differentiation. Less histone H3 transcript may be expected in comparison with the other histone genes, as histone H3 is represented by only seven copies in the genome, in comparison with histones H2A, H2B and H4, which have between 11-14 copies each.

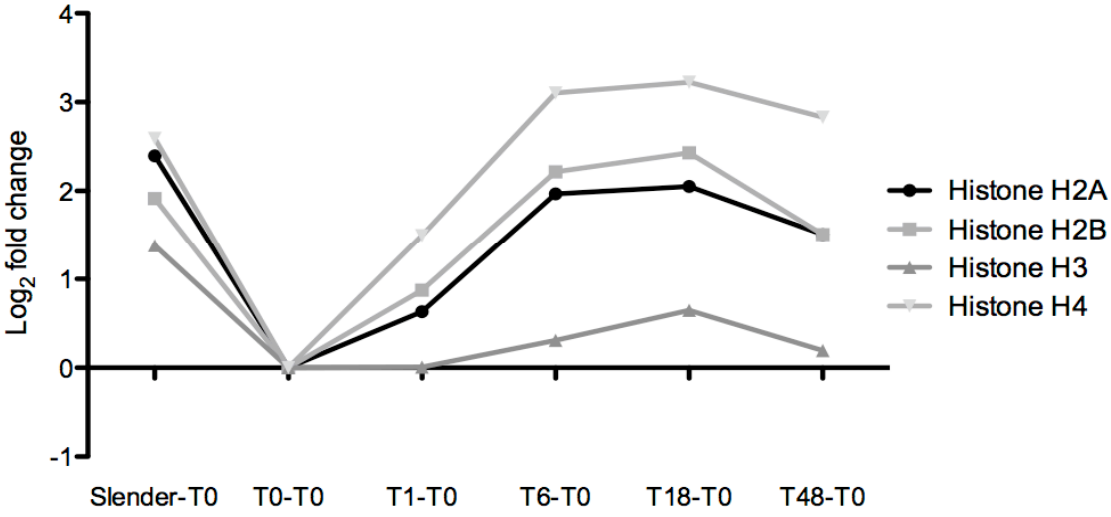



**Figure 4.3. Expression profile of histone transcripts**. The histones all demonstrate a decrease in mRNA abundance in stumpy form cells. This is followed by elevation of mRNA levels at 1 hour into differentiation, which is rapidly increased over the time course, peaking between 6-18 hours for histone H2A, H2B and H4, with H3 resuming expression after 6 hours at lower levels to the others.

The peak of the histone transcripts at between 6-18 hours through differentiation is in keeping with the anticipated point at which cells would be in S phase. Studies of DNA synthesis in differentiating cells has shown that cells undergo S phase between 6-8 hours into differentiation (Matthews and Gull, 1994), and the mRNA levels for histones H2A, H2B and H4 have been shown to peak in S phase (Ersfeld et al., 1996), therefore the profile is in line with the published knowledge on histone transcript regulation. The histones were among the most significantly differentially expressed genes and all showed the expected profile, thus providing good validation of the microarray data.

### 4.2.3 Genes up-regulated in procyclic forms

Major differences in the environment between the mammalian bloodstream and tsetse fly midgut require trypanosomes to alter a variety of systems including their surface coat and their metabolism. The procyclin surface coat is composed of both EP and GPEET procyclins, with GPEET found on earlier procyclic forms, this waning as the cells become more established (Roditi et al., 1998). The rapidly increasing expression of E2 and EP3-2 in Figure 4.4A was expected, as the procyclins are not expressed in bloodstream forms, however the peak at 1 hour is a little sooner than might be expected; the same cell line has previously been shown to have maximal EP mRNA between 4-6 hours (Pays et al., 1993, Rolin et al., 1993). This difference could be contributed to by the reduced temperature experienced during cell purification, which is known to elevate procyclin mRNA levels (Engstler and Boshart, 2004), although syringes and media were pre-warmed to reduce any temperature fluctuations. Alternatively, as previous work used Northern blots to analyse expression levels, these may have been less sensitive than the microarray, such that EP expression was detected later than seen here. Additionally, in previous studies the total EP levels were examined, whilst we were able to distinguish between the different EP procyclin transcripts so will have been more able to dissect subtle changes. A peak in expression of the procyclin mRNAs through differentiation was expected because cells have to rapidly cover the entire cell surface with the new surface coat early in differentiation and, once fully differentiated, the levels can be reduced (Pays et al., 1993). Whilst EP2 is actively transcribed in procyclic cells it is not translated in parasites in culture; however the

expression profile for EP3-2, which is translated, was very similar. The two EP3-2 transcripts shown on the graphs are derived from different loci: Tb927.6.520 and Tb927.6.480, which are 88% identical. The microarray slides did not contain a probe against EP1, however the results from the other EP and GPEET transcripts are a good indication that the microarray data accurately distinguished procyclic form specific regulation.

The profile for GPEET is also in keeping with the literature (Figure 4.4A). The relative proportions of GPEET and EP in the surface coat vary between individual cultures, however, the order of appearance of the different proteins is similar between stocks, with GPEET only ever expressed during the early stages of an infection (Vassella et al., 2000). Both GPEET and EP can be detected on the surface of the cell from three hours into differentiation (Vassella et al., 2001), with GPEET peaking twelve hours into differentiation, around six hours after the peak in EP protein levels (Vassella et al., 2000). The microarray profile of the mRNA levels of GPEET showing a peak in expression shortly after the peak in EP expression is in keeping with these earlier observations of the protein levels.

The cytochrome oxidase complex is not required in bloodstream form parasites as they are able to survive on the abundant glucose present in the host bloodstream. However, once the parasites are taken up into the tsetse fly they must derive energy from available proline via mitochondrial respiration, and hence the COX subunits are up-regulated in procyclic form cells (Mayho et al., 2006). In our analysis, the COX subunits V, VIII and IX were amongst the 407 most statistically significant differentially regulated genes ($p< 0.001$), whilst subunit VI was significantly differentially expressed at $p<0.05$. Figure 4.4B shows the increasing expression of COX V, VI and IX through differentiation, of which the early expression of COX V and IX show pre-adaptation for tsetse uptake through up-regulation of these subunits in the stumpy form. COX VIII also has lower expression in slender forms versus procyclic forms, but is lowest in stumpy forms. These data are in accordance with

published work which studied the differences between bloodstream and procyclic form expression levels (Mayho et al., 2006).

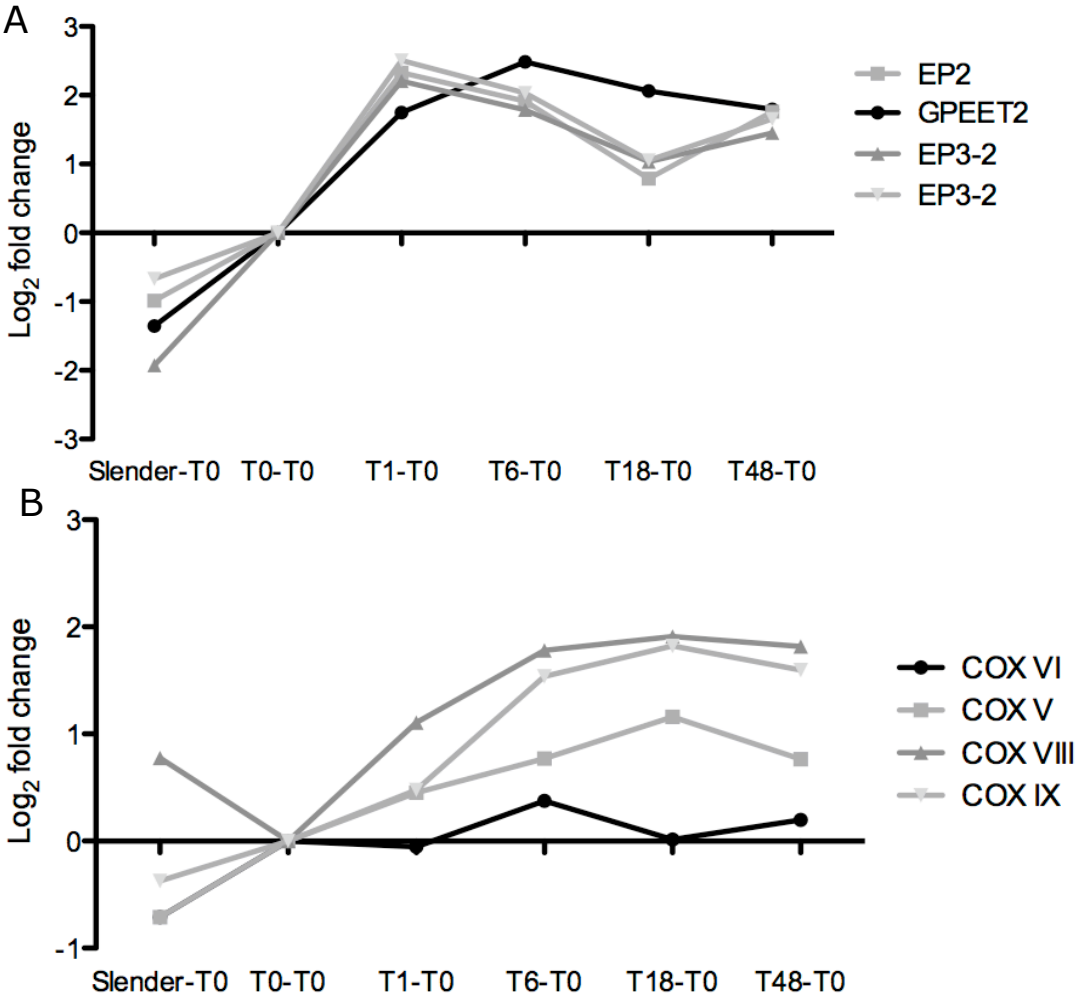



**Figure 4.4. Expression profile of procyclic specific transcripts by microarray**. Expression profile of (A) the EP and GPEET procyclins. The expression profiles of EP2 and EP3-2 are virtually identical, with both transcripts seen to have already increased in stumpy form cells, peaking at 1 hour through differentiation and then remaining expressed at a high level throughout differentiation, as expected. A similar profile was seen for GPEET, which peaks slightly later at 6 hours through differentiation. (B) COX subunits expression profiles showed that, as expected, all subunits demonstrated lower expression in bloodstream forms than at 48 hours through differentiation. COX V, IX and, to a lesser extent, VI showed an increase in expression in stumpy forms compared to slender forms, whereas COX VIII showed a drop in expression in stumpy form, which was rapidly reversed by 1 hour through differentiation.

Hence, the expression levels of genes required specifically in the procyclic life stage are rapidly increased during differentiation as expected, further validating the microarray.

### 4.2.3.1 Unusual expression profiles of procyclin-associated genes

The 0.1% list is so statistically significant that no miss-calls should feature. Thus it was surprising that two procyclin-associated-genes (PAGs), a member of the PAG1 family Tb10.70.1310 and of the PAG2 family Tb10.70.1300, showed a decrease in expression over the course of differentiation, contrary to the expected profile (Figure 4.5). Conventionally, PAGs form part of the procyclin polycistronic unit, are transcribed by RNA polymerase I and show similar expression profiles to the procyclins, with higher expression in procyclic form cells than bloodstream form cells (Berberof et al., 1996). In contrast, these PAGs significantly differentially expressed at the 0.1% level, were found to be chromosome internal copies. Nonetheless, although the expression of these genes had not before been tested, we expected them to follow similar profiles as the sub-telomeric copies in the procyclin gene locus. However, preliminary results from a deep sequencing project have shown the same trend for these two genes (Isabel Roditi, personal communication), implying that for this cell line at least, the observed expression trend was genuine. This provided more credence to the microarray data and highlighted an interesting trend. Subsequent analysis of the procyclin-locus associated PAG2 (Tb10.6k15.0060) and PAG4 (Tb10.6k15.0070), which were significantly expressed at the 5% level, showed that these too followed similar profiles, with decreasing mRNA levels over the differentiation time course. However, PAG3 (Tb927.6.490), also significant at the 5% level, did show the expected profile, with transient up-regulation in expression over differentiation, similar to the profiles of the procyclin transcripts (Figure 4.5A).

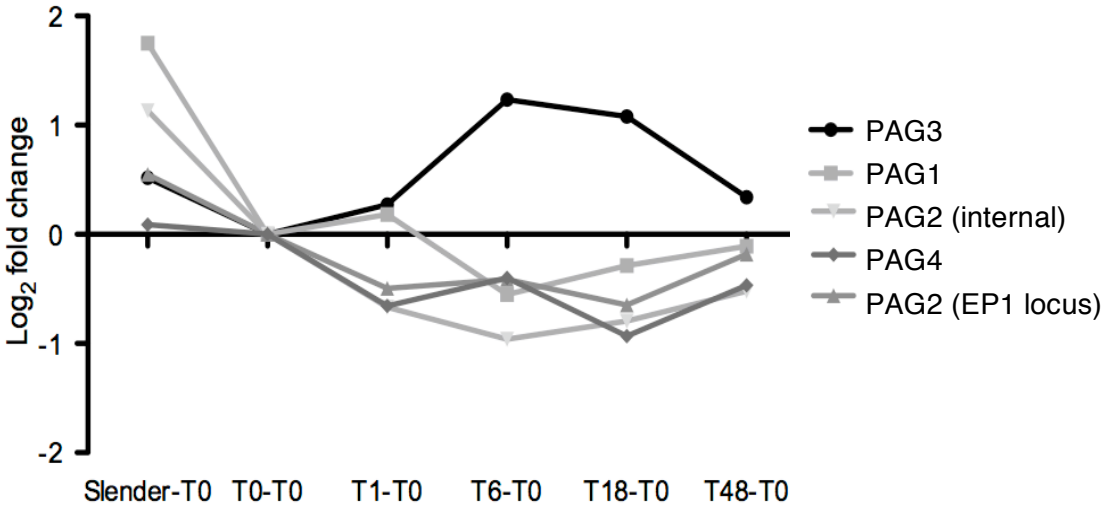



**Figure 4.5. Expression profiles of members of the PAG family as determined by microarray**. The chromosome internal PAG genes Tb10.70.1300 and Tb10.70.1310 that were significant at the 0.1% level show a sharp decrease in expression from slender to stumpy form and continued to decrease throughout the time course. Tb10.6k15.0060 (PAG2) and Tb10.6k15.0070 (PAG4) in the EP procyclin locus are significantly differentiation expressed at $p<0.05$ and show a similar profile. Tb927.6.490 (PAG3) shows an increase in expression during differentiation, peaking at 6 hours.

The expression profile of the PAG genes found in the 0.1% significance list was surprising, yet was confirmed through subsequent deep-sequencing technology. The similar profiles seen for the procyclin locus-associated PAG 2 and 4 genes was also surprising. Investigation of the microarray probes used for the Tb10.6k15.0060/70 locus and Tb10.70.1300/10 locus showed that the probes used for the microarray were specific, eliminating the potential for cross hybridisation.

The unusual expression profile of some of these PAGs warrants follow-up work. In particular; it is interesting that there is no correlation between the location of the genes or the type of RNA polymerase used for transcription and the expression profiles, with PAG 2 and 4 in the EP1/2 locus and PAG3 in the EP3 locus showing entirely different profiles despite both being transcribed by RNA polymerase I. The mRNA levels of PAG3 have been analysed previously in monomorphic bloodstream form and procyclic form cells where negligible levels were found in bloodstream

form (Berberof et al., 1996). As we do not know if the levels seen at 48 hours through differentiation are reflective of the steady-state levels of PAG3 in established procyclic form cells, and as monomorphic cultures were used for the previous study, it is not possible to use these results as a benchmark for validation of the microarray data.

### 4.2.4 Genes up-regulated in slender forms

The JCVI version 3 microarray slides used in these experiments were designed using the *T. brucei* TREU927/4 genome reference strain, and hence lack a probe for the AnTat1.1 VSG expressed by the parasites used. Hence the most obvious marker of bloodstream form cells could not be examined. However, the invariant surface glycoproteins (ISGs), which are shielded by the VSG coat, were represented on the microarray and these are also known to show developmentally regulated expression (Ziegelbauer and Overath, 1992). Figure 4.6A shows the expression profiles of the five ISG genes showing differential expression at the 0.1% significance level. All have increased expression in slender form cells and have immediate down regulation of expression in stumpy form cells. The expression profile of the ISGs therefore is in keeping with the protein expression, which is seen to disappear entirely in procyclic form cells (Ziegelbauer and Overath, 1992).

The metabolic differences between bloodstream form cells and procyclic forms also provide good expression markers of life stage to validate the microarray profiles. Figure 4.6B shows the expression profile for fructose-biphosphate aldolase, a protein localised to the glycosomes and required in glycolysis, showing, as expected, up-regulation in slender forms. Previous analysis of the mRNA levels of fructose-biphosphate aldolase using Northern blots have demonstrated a 6-fold increase between procyclic form and bloodstream form cells (Clayton, 1985). The result here of a 4-fold decrease between slender forms and cells 48 hours through differentiation is slightly less, however not unexpected, as microarrays tend to underestimate differences (Wurmbach et al., 2003). Interestingly however, the activity of the enzyme shows elevation of up to 30-fold in the bloodstream form with respect to

procyclic form (Hart et al., 1984), possibly due to the extremely short half life of the protein in procyclic form cells (Clayton, 1987).

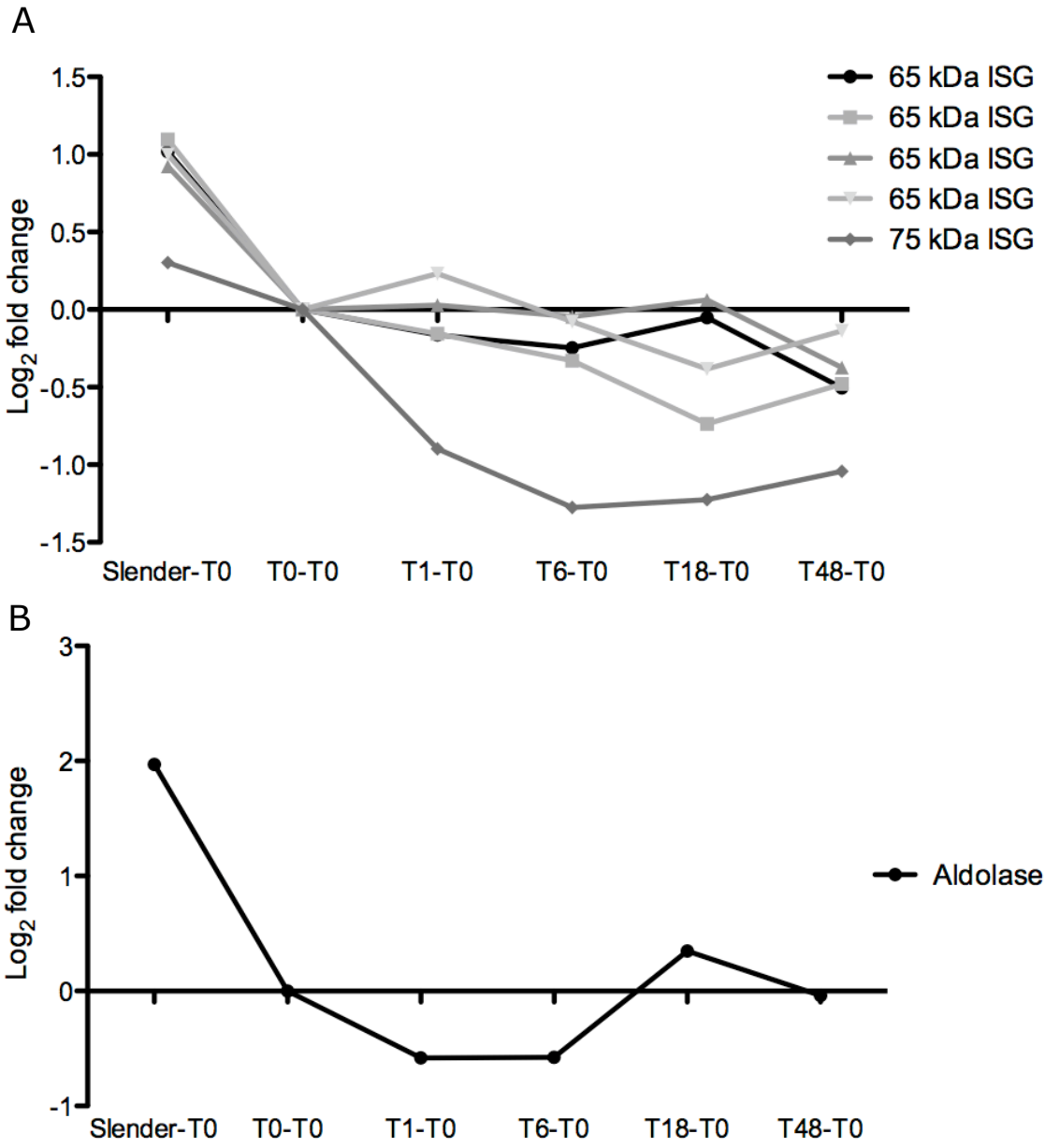



**Figure 4.6. Expression profiles of bloodstream specific transcripts**. (A) All ISG transcripts showed increased expression in slender form cells and were down-regulated in stumpy form cells, staying at low levels for the remainder of the time course. (B) The expression profile for aldolase, differentially regulated at a significance of 0.1%, also shows up-regulation in slender form cells as expected.

To conclude, the slender specific transcripts followed expected expression profiles.

### 4.2.5 Genes up-regulated in stumpy form cells.

Few genes are expected to be expressed in stumpy form cells due to the associated growth arrest and fewer still are known to be up-regulated in stumpy cells. However, recently a family of carboxylate transporters have been characterised, which are required for citrate/cis-aconitate dependent differentiation (Dean et al., 2009). These PAD genes lie in an array of eight gene members, of which five were shown to be significantly differentially expressed on the microarray, PAD6 at the 0.1% level, and PAD1, 3, 4 and 8 at the 5% level. Figure 4.7 shows the expression profile of each gene, demonstrating that PAD1, 6 and 8 are all up-regulated in stumpy forms, with PAD8 in particular being expressed predominantly in stumpy form cells. PAD 3 and 4 appear to be more highly expressed during the later events of differentiation, demonstrating a similar profile to PAD2, which has been shown to be up-regulated in procyclic form cells (Dean et al., 2009). Whilst the coding regions of the genes within the array are all highly similar, they do possess largely different 3'UTRs, so it is not surprising that they show different expression profiles, although all are seen to be down-regulated in slender forms, matching the analysis of Dean et al (2009).

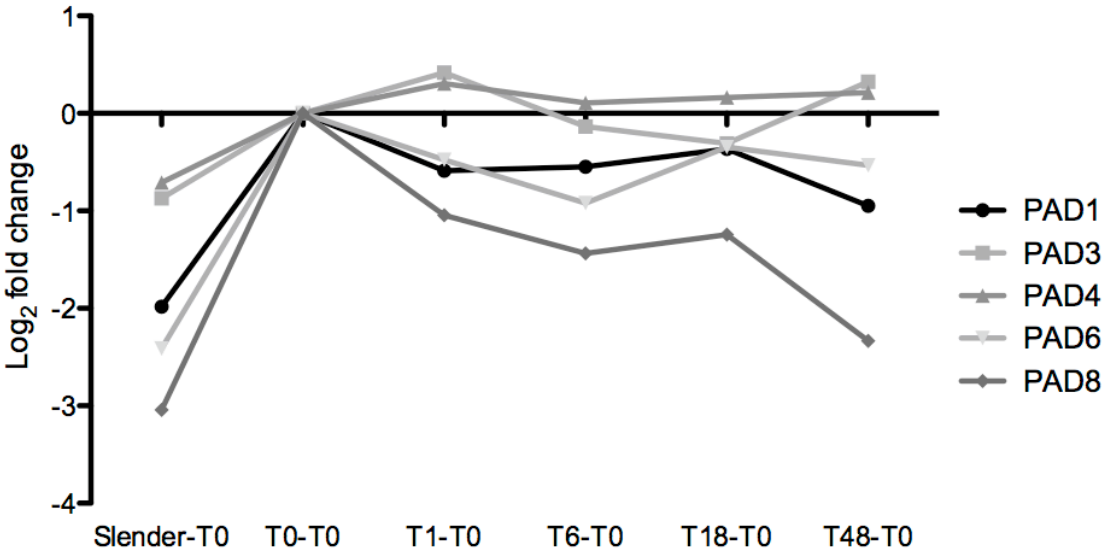



**Figure 4.7. Expression profiles of members of the PAD array**. Five of the eight members of the PAD array showed significant differential expression, with PAD1, 6 and 8 being highly stumpy-enriched, and PAD 3 and 4 proving to be down-regulated in slender form cells.

As the initial validation of the microarray against the literature had proven largely accurate, we next used other methodologies to investigate whether the microarray results were biologically accurate.

### *4.3 Validation of the microarray using qPCR*

#### 4.3.1 qPCR control analysis and method evaluation

Quantitative PCR (qPCR) is a standard method to verify microarray data (Morey et al., 2006), and by using SYBR® Green-based methodologies, it is possible to analyse numerous different targets. In this approach, SYBR® Green is added into the qPCR reaction and during the extension phase is incorporated into the newly amplified PCR product, in which configuration it is activated and fluoresces. The increasing level of fluorescence is measured in real time and a crossing-point (Cp, also called Ct) value is assigned to each sample, representing the point at which the level of fluorescence crosses the threshold above background. Once the Cp values were determined for sample genes and a reference gene, relative quantification calculations were carried out using the Pfaffl calculation (Pfaffl, 2001):

$$\text{Ratio} = \frac{(E_{\text{target}})^{\Delta \text{CP}_{\text{target}}(\textit{control-sample})}}{(E_{\text{ref}})^{\Delta \text{CP}_{\text{ref}}\textit{(control-sample)}}}$$

where E is the PCR efficiency of the primers, target refers to the test genes and the reference gene is the normalisation control gene. In this equation, control refers to the stumpy form cDNA so that all other samples are calculated with reference to stumpy form levels of transcript.

#### 4.3.1.1 Primer design

Quantitative PCR was used as a means to validate the microarray data by assaying the same RNA samples that had been used to hybridise to the microarray. The individual RNA sample from each of the biological replicates were not pooled, in order to avoid potential contamination of good samples if one sample should prove to be anomalous (Allison et al., 2006), so all analyses were carried out on single sample

sets. We were particularly interested in confirming that the genes identified as being enriched in stumpy form cells were correct, so primers were designed to amplify a 143-217 base pair fragment from all transcripts in the -1-1-1-1-1 and -10-1-1-1 tables using the primer design program, Primer3 (Rozen and Skaletsky, 2000). The primers were then assessed for their suitability in a qPCR reaction by cross-reactivity screening through running BLAST searches on the primer pairs, and by assessing their ability to form primer dimers using 'Amplify 3.1' (http://engels.genetics.wisc.edu/amplify/). All primers were designed to very stringent criteria in order to ensure they were suitable for qPCR, these criteria included: a melting temperature of 60˚C, the primers must end in a GC clamp and must not contain more than two consecutive repeated residues. Genes for which it was impossible to generate primers to these specifications were not analysed. Prior to using the primers in a qPCR reaction, they were first tested on gDNA in a standard PCR reaction and the products visualised on a 2% agarose gel (Figure 4.8).

| Lane # | Gene |
|---|---|
| 1 | Tb10.6k15.3640 |
| 2 | Tb927.4.1020 |
| 3 | Tb10.70.6180 |
| 4 | Tb09.211.2300 |
| 5 | Tb10.26.0220 |
| 6 | Tb927.3.2710 |
| 7 | Tb927.1.4450 |
| 8 | Tb927.7.7160 |
| 9 | Tb927.5.3400 |
| 10 | Tb10.389.0540 |
| 11 | Tb09.160.2110 |
| 12 | Tb10.6k15.0300 |
| 13 | Tb927.7.3980 |
| 14 | Tb10.61.3040 |
| 15 | Tb927.6.4960 |
| 16 | Tb10.389.0650 |
| 17 | Tb927.8.1270 |
| 18 | Tb927.7.3170 |
| 19 | Tb927.5.2320 |
| 20 | Tb927.6.4180 |
| 21 | Tb927.5.4110 |

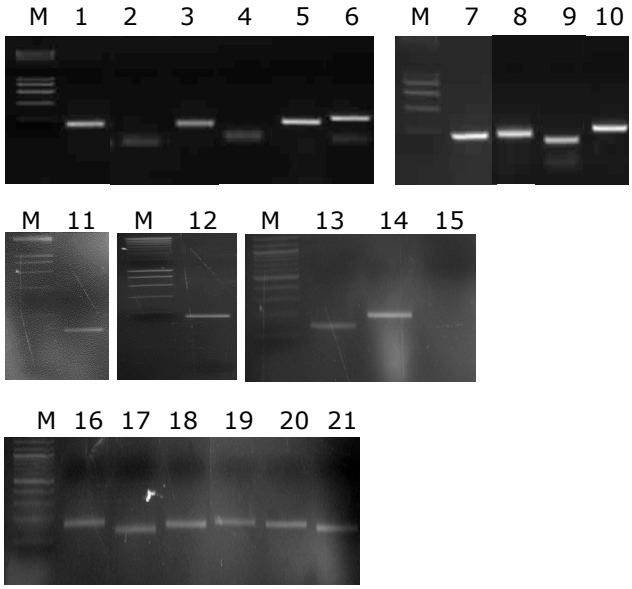



**Figure 4.8. PCR amplification of gDNA using primers designed for qPCR**. qPCR primers were used in a standard PCR reaction using gDNA as a template at an annealing temperature of 60˚C to test specificity of the primers. The primers used in lanes 2, 4 and 15 failed to amplify under these conditions and were therefore not used in qPCR experiments. All other primers amplified only one product corresponding in each case to the expected size and, with the exception of the failed primers, no primer pairs formed primer dimers, so these all underwent further testing for suitability in qPCR reactions. Gene accession numbers relating to row number are listed in the table. M= marker.

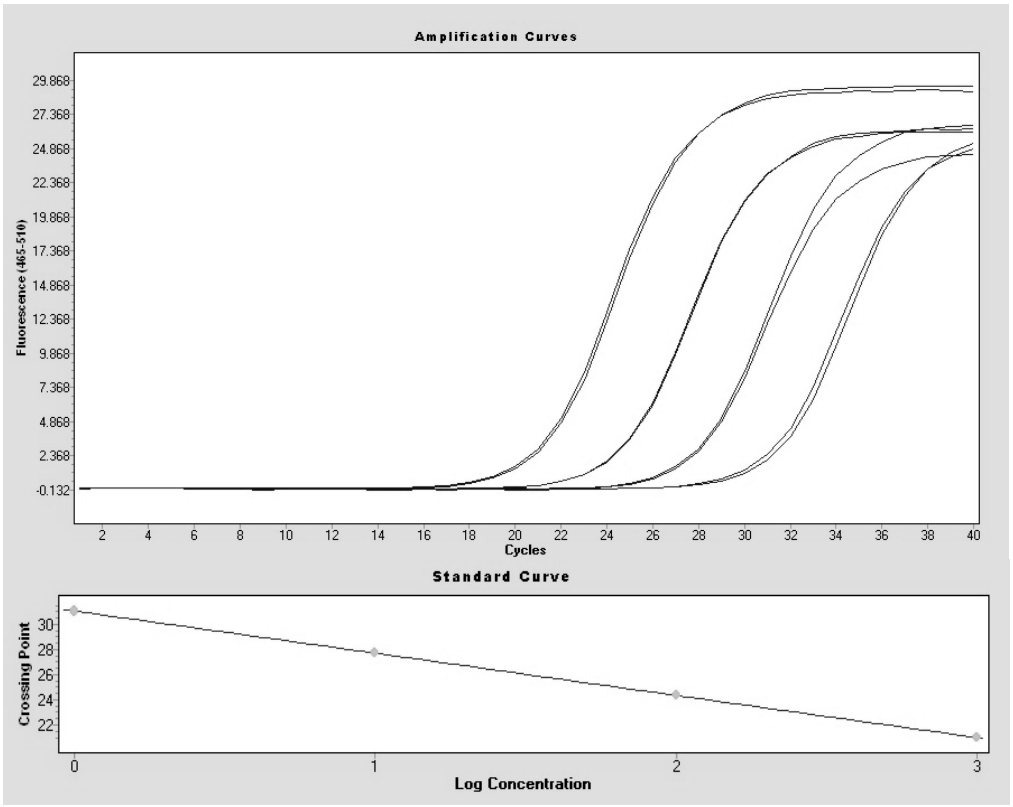
Quantification of a qPCR reaction is conducted through comparing the level of a gene of interest against a reference control gene, however in order for the comparisons to be accurate, the primer pairs for the different genes must have similar PCR efficiencies to each other. The delta delta Ct (ddCT) calculation is frequently used to quantify the difference between the crossing point of the control gene with that of the sample gene, automatically assuming the efficiency of the primers are 100%. The Pfaffl method is essentially the same equation, however the differing PCR efficiency of the primers is taken into account for the calculation (Pfaffl, 2001). In order to establish the PCR efficiency of the primer pairs used here, serial dilutions of stumpy cDNA from *T. brucei brucei* EATRO 2340 cells derived from a rodent

infection were assayed. Ten-fold dilutions were prepared from the cDNA and qPCR reactions carried out in duplicate, from which standard curves were generated. PCR efficiency was calculated from:

$$\text{Efficiency} = 10^{(-1/\text{slope})}-1$$

Only primer pairs providing a PCR efficiency between 90 and 110% were used for the validation experiment. Figure 4.9B shows the standard curve generated from the amplification curves (Figure 4.9A) of Tb10.389.0540, which was one of the two genes used for normalisation as it was discovered in the microarray to be expressed at a constant level in all life stages. This confirmed that the qPCR reaction was giving an accurate qualification of transcript levels in the cDNA dilutions for these primers.

A




B

Slope = -3.359

Efficiency = $10^{(-1/-3.359)}$-1= 98.5

**Figure 4.9. PCR efficiency of qPCR primers calculated from standard curves.** (A) Amplification curves of 10-fold serial dilutions of *T. b. brucei* EATRO 2340 stumpy form cDNA. Primers against the normalisation control, Tb10.389.0540, were used in a qPCR reaction to amplify stumpy form cDNA to test the efficiency of the primers. (B) From the amplification curves, a standard curve was generated by plotting the crossing points for each dilution. The curve of the slope from the standard curve was -3.359, which corresponds with a PCR efficiency of 98.5%.

At the end of each qPCR experiment, a melting curve was conducted to ensure that the fluorescence signal recorded was not the result of primer dimers or non-specific amplification. Melting curves were generated through sequential increase of temperature immediately following the qPCR reaction, from 65˚C to 97˚C, taking a fluorescent reading every 0.5˚C. If only one product is present, the meltcurve should show a sudden drop in fluorescence as the entire product melts simultaneously, becoming single stranded and thus dissociating from the SYBR® Green, which therefore stops fluorescing, as can be seen in Figure 4.10A. The melt peak calculated from the melting curve shows the point at which the product melts. In the case of the

alternative oxidase (TAO) primers the melting point is at approximately 89˚C (Figure 4.10B). A number of primers designed for the qPCR did form primer dimers, which impacted on the efficiency of the amplification rendering them unsuitable for accurate results. Figure 10 C and D show the melting curve and associated melt peak of Tb927.5.4110, with strong primer dimers appearing in the sample at the lower concentrations of template. This rendered these primers unsuitable for qPCR.

In addition, the acquisition temperature was set to 82˚C in order to reduce the possibility of detecting signal from primer dimers, which should have a melting temperature below 80˚C. Figures 4.10A and B clearly show that only one product was amplified by the primers designed against TAO, indicating that the set of primers does not form primer dimers or generate non-specific products. Four of the primer sets caused poor PCR efficiency through to formation of primer dimers so were not tested on experimental samples. The remaining primer sets that did not generate primer dimers were used to validate the microarray data.

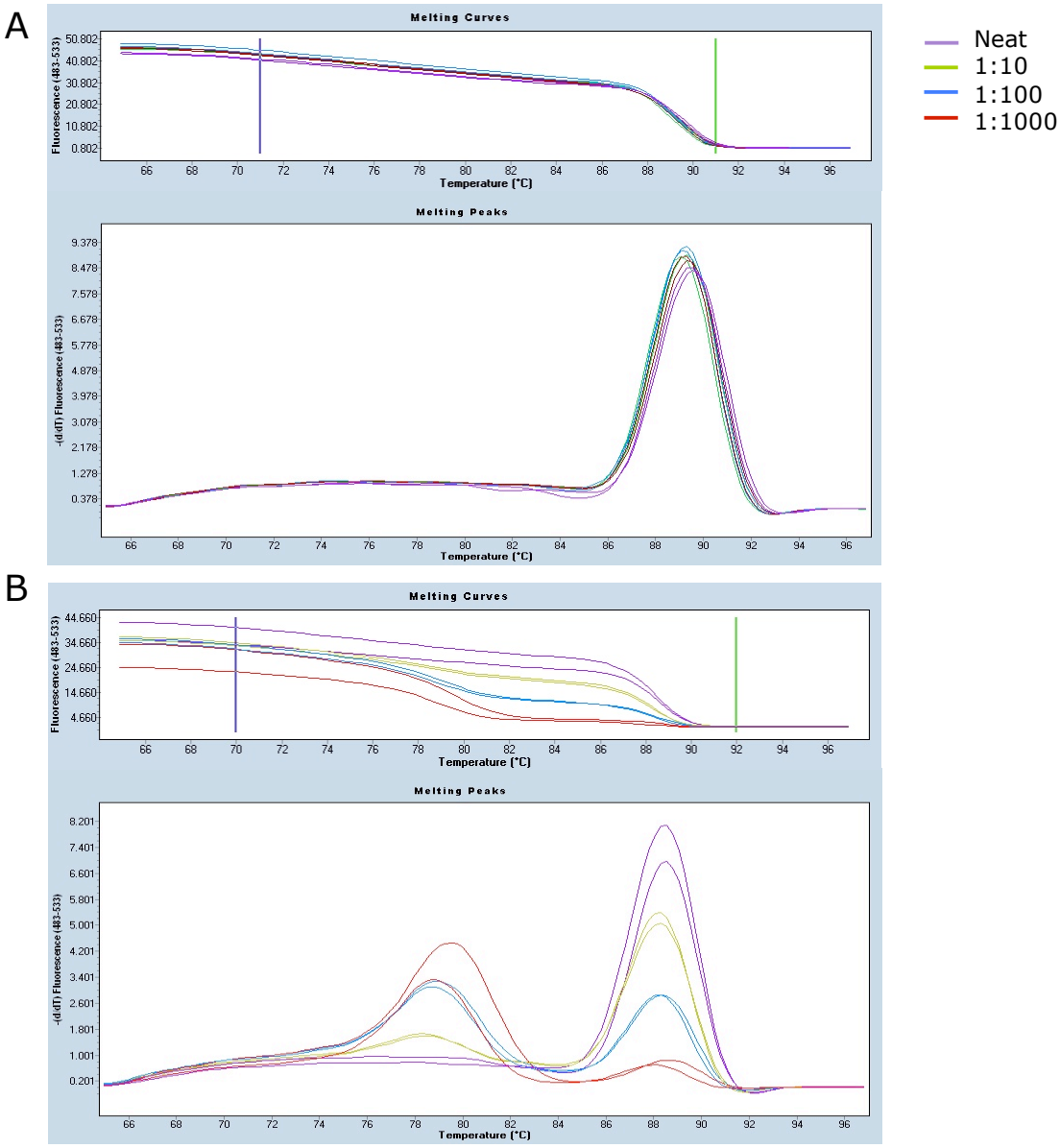



**Figure 4.10. Meltcurve analysis of qPCR products to detect potential contaminating primer dimer products or non-specific amplification**.
Melting curves and melt peaks generated from the qPCR reaction on the serially diluted cDNA to calculate PCR efficiency for (A) TAO, showing a sharp drop on the melting curve representing the single peak at 89˚C seen in the melt peak graph, and (B) Tb927.5.4110, showing two peaks representing the melting of two products. The product with the melting temperature of 88˚C is the desired product, whereas the sequence melting at 79˚C reflects primer dimers. An increase in primer dimers is seen with decreasing starting template as at higher template concentrations, the primers preferentially bind the template.

### 4.3.1.2 Choice of normalisation control

For qPCR to be accurate, a housekeeping gene is required, against which all the data can be normalised. However, the differentiation process in *T. brucei* involves radical changes in gene expression and even conventionally used housekeeping genes, such as tubulin, are not necessarily equally expressed throughout these changes. Using the normalised microarray data, it was possible to search for a gene with the lowest standard deviation across all individual samples tested at each time point, which also had a relatively high level of expression. Such a transcript could act as a suitable control for comparison against other, regulated transcripts. To this end, primers were designed against Tb10.389.0540, which codes for a hypothetical protein and had a standard deviation over all 25 samples of 0.32, and for comparison a second control was designed against Tb09.160.2110, also encoding a hypothetical protein, which displayed a standard deviation of 0.37. The trypanosome alternative oxidase (TAO) gene had been identified previously by microarray as up-regulated in bloodstream form when compared with procyclic form cells (Diehl et al., 2002). In our extended analysis, which looked at key time points during differentiation, TAO was found to be significantly elevated in stumpy form cells, and so was used as a good candidate to trial the two controls. Thus, samples were taken from different biological-replicates and qPCR carried out on different days, this yielding similar general profiles for both normalisation controls. However, Figure 4.11 demonstrates the problems involved in choosing a good normalisation gene for qPCR; whilst both of the genes chosen confirm that TAO is up-regulated in stumpy forms, they show different expression levels across the rest of the time course. This might be a result of using different biological replicates but it also highlights the limitations of using qPCR as a method of validating microarrays, as the normalisation procedure for qPCR cannot be nearly as 'global' as for the normalisation steps used in the microarray analysis, potentially leading to inaccurate or bias normalisation. Nonetheless, for all subsequent qPCRs Tb10.389.0540 was used for normalisation.

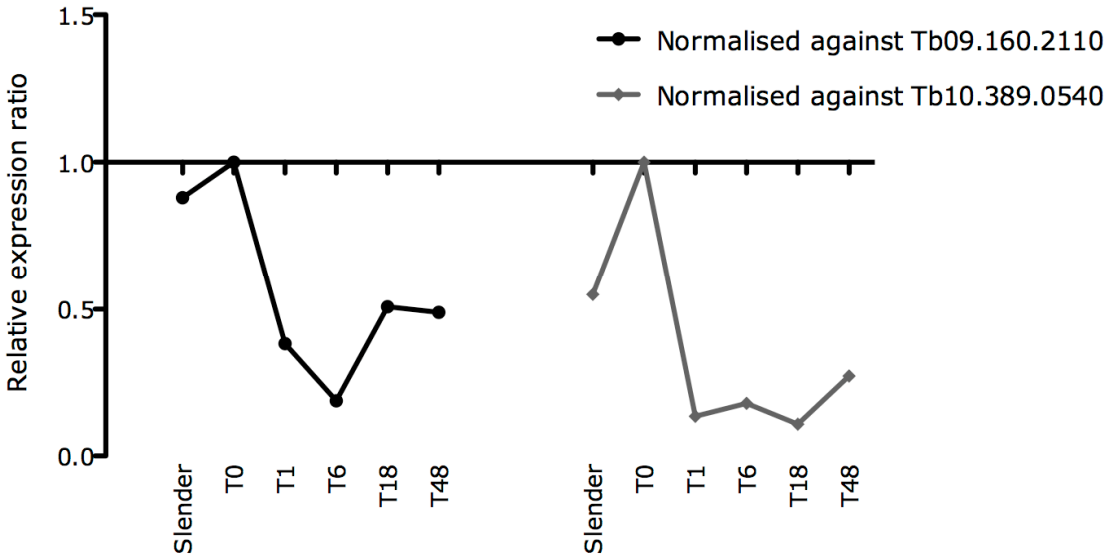



**Figure 4.11. Comparison of two candidate genes to act as the normalisation control in qPCR**. qPCR results using the cDNA generated from differentiating cells used in the microarray from TAO primers using different biological replicates, run on different days and normalised using (A) Tb09.160.2110 and (B) Tb10.389.0540. Samples were expressed relative to their expression levels in stumpy form.

### 4.3.2 Microarray analysis and evaluation by qPCR

To assess the expression levels of transcripts predicted to be up-regulated in stumpy form cells by microarray analysis, cDNA was generated using total RNA from a single bio-replicate time course, a slender sample and established procyclic form 449 cells from culture. Oligo dT primers were used to amplify only the mRNA and not the rRNA, as was used for the microarray, and qPCR carried out. In each experiment a 'no-RT' control of the stumpy form sample was included, whereby the qPCR reaction was carried out on template cDNA, generated from RNA in parallel to the sample cDNA, but which lacked the reverse transcriptase, to ensure amplification was from the cDNA and not contaminating gDNA. Additionally, a 'no-template' control was included, whereby the cDNA was replaced with water in the qPCR reaction, to test for genomic contamination from any other reagent source. The generated results were then presented as relative to their expression in stumpy form.

Although many of the genes tested gave the expected profile by qPCR, a number of qPCR experiments were found to give very different results to those in the microarray. These were reproducible on different biological replicates in different reactions, even with different dilutions of the cDNA to ensure that there was no saturation of the signal. Figure 4.12A shows genes whose expression profile from qPCR matched that of the array analysis, whereas Figure 4.12B shows the qPCR profiles of the genes found to be in disagreement to the microarray. For those that did not agree with the array profile, no pattern was observed in the profiles ruling out a technical error in assigning profiles to these genes.

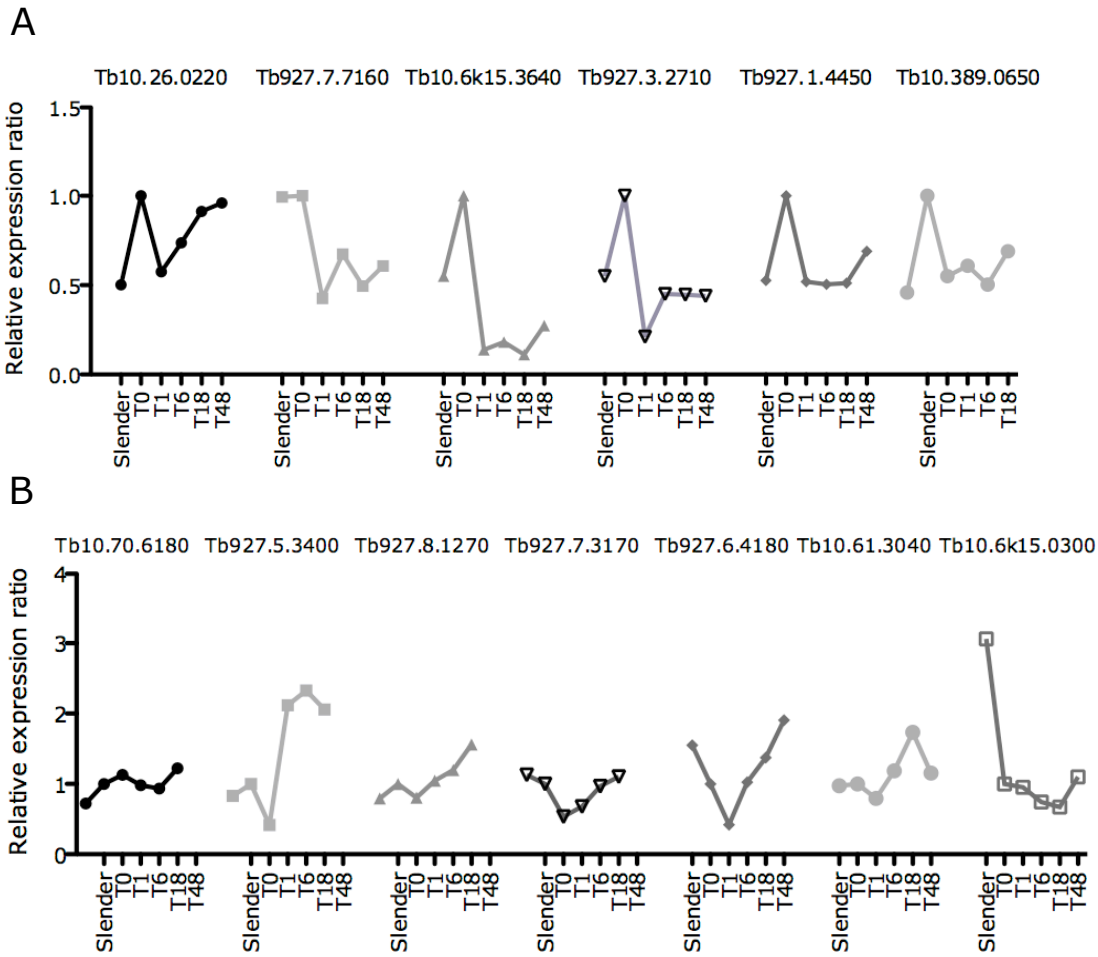



**Figure 4.12. Validation of microarray results using qPCR**. qPCR validation of genes classed as 'stumpy-enriched' by the microarray of (A) the six genes that followed the correct general profile and (B) the seven genes that showed different profiles to the microarray. Even where the qPCR results agreed with the trend determined by the microarray, the level of change often differed considerably.

In conclusion, the qPCR data gave conflicting results to the microarray in some cases, and for some of the genes that did show correct trends, the precise expression profile differed from the array data. Such a discrepancy between qPCR and microarray data is often seen (Holzer et al., 2006).

### *4.4 Northern blot validation*

In order to generate verification data for those stumpy-enriched genes that could not be tested using qPCR, either due to failure to design unique primers, inability of the primers to amplify the product at a 60˚C annealing temperature or failure to show a

good enough PCR efficiency, Northern blots were used as an alternative method of validation. The gene for alternative oxidase (TAO), which had produced successful qPCR data, was also included to gain further confirmation as this gene has a high level of expression, thus making it a good candidate for Northern analysis. The TAO probe had previously been generated by Melanie Buelmann, and Northern probes against two other candidates, Tb09.211.2300 (a putative RNA helicase) and Tb927.1.4450 (a potential calcium channel), were generated through using the programme RNAit (Redmond et al., 2003) to identify unique sequences of the gene. Primers were designed to amplify the unique sequence, which was then inserted into a cloning vector (pGem-T-easy), the orientation of the sequence ascertained and either T7 or SP6 RNA polymerase was added to generate an antisense riboprobe from transcription initiation sites in the vector. Northern Blots were generated using the same mRNA as used for the microarrays, with the addition of mRNA isolated from established procyclic form cells (PCF 449) and then riboprobes were used to detect the levels of the transcripts throughout the time course. Figure 4.13 shows that Tb10.6k15.3640 (TAO) and Tb927.1.4450 had the expected profiles with a higher level of expression for stumpy forms compared with the other time points. However, Tb09.211.2300 showed a greatly divergent result to that seen on the microarray, with the highest levels of expression at 18 and 48 hours through differentiation, and much lower levels seen in stumpy form cells.

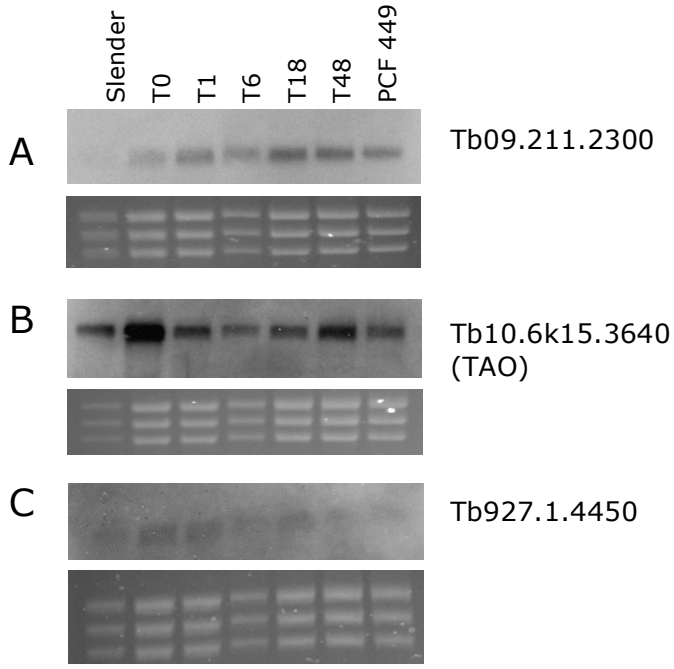



**Figure 4.13. Expression of transcripts identified as up-regulated in stumpy form cells validated using Northern blots**. Northern blots were conducted on the same mRNA as used for the microarray to verify the expression profiles of genes detected in the microarrays as stumpy-enriched. (A) Tb09.211.2300 does not follow the trend suggested in the microarray because expression levels were found to be lower in slender and stumpy forms and increase over the time course. (B) Tb10.6k15.3640 shows the highest expression levels in stumpy forms, with reduced levels in cells 18 hours into differentiation, confirming the microarray and qPCR data. (C) Tb927.1.4450 shows an enrichment in stumpy form cells, in agreement with the microarray data, although its overall level of expression is apparently low in comparison to the other transcripts assayed. The ethidium bromide stained ribosomal RNA is shown below as a loading control.

Again, the lack of correlation between the microarray and Northern analysis for one of the transcripts merely corroborates the limited worth of alternative methods of validation of microarray data. In particular, Northern blotting does not use any internal normalisation controls; rather the normalisation is based on the equal loading of rRNA based on the visual assessment from the ethidium bromide stained gels. This is not effective as a loading control for the stumpy form sample because it has been shown that in this life form, despite hugely reduced levels of mRNA, rRNA levels remain constant (Pays et al., 1993). Thus an equal amount of rRNA will not

correlate with an equal amount of total RNA loaded in each sample. In contrast, the microarray normalisation takes into account the expression levels of thousands of genes simultaneously. With these constraints taken into account it is perhaps more surprising that any of the Northern data showed similar results to the microarray rather than the fact that one gene showed a different profile.

### *4.5 Comparisons against other microarray analyses in* T. brucei

Historically, most microarray analyses carried out on *T. brucei* have looked only at culture-adapted parasites, and compared only monomorphic bloodstream form cells against procyclic cells. Hence, a detailed comparison of our microarray data against these previously published experiments is not likely to be a worthwhile examination. Certainly, the assumption that monomorphic cells are a suitable equivalent for field isolated bloodstream form slender cells is likely to be wrong, and the compounding distinction between in vitro and in vivo grown parasites is liable to generate further differences. Following the publication of our microarray data (Appendix 2), two similar microarray analyses were also published investigating differentiation of *T. brucei* (Jensen et al., 2009, Queiroz et al., 2009). The Queiroz et al (2009) study was conducted on the same microarray slides using the same cell line, however the parasites used for the differentiation experiment were derived from in vitro differentiated cells, which do not develop into true stumpy form cells. The Jensen et al (2009) study was conducted on a custom tile array, such that up to eight probes were designed for each gene. This study also looked at the differences between parasites grown in vitro and those isolated from in vivo infections and concluded that the differences in environment had very little effect of the gene expression of the parasites (Jensen et al., 2009).

In their paper, Queiroz et al (2009) conducted a brief comparison of their data set against our data set and conclude that the majority of genes show similar trends in both studies, although the genes defined in the different studies as showing the most significant differential expression only overlap to a certain degree. The timing of events during differentiation are distorted in the Queiroz et al (2009) study, as the

parasites were not true stumpy forms before differentiation, but had merely been grown to a high density, at which point they are described as "resembling stumpy forms", however the fact that many genes show comparable trends of expression is encouraging.

The Quieroz et al (2009) experimental design differed from ours in that the arrays were competitive two-colour analyses, with samples from each time point hybridised alongside procyclic form cDNA. In order to facilitate a comparative analysis against our data, which was presented relative to expression in stumpy form cells, the authors converted our data to present each time point with respect to expression at 48 hours, this sample being considered equivalent to established procyclic form cells. The limited set of genes chosen for the comparisons show highly similar profiles through differentiation, with Tb11.02.2310, a prostaglandin f synthase, proving almost identical (Queiroz et al., 2009). However, the sets of genes differentially expressed at the 5% significance level did not overlap extensively between the studies, rendering further comparisons unfeasible. The authors conducted extensive bio-informatic analysis of the data derived from the in vitro data set, grouping together genes with similar profiles and subsequently assigning likely functions to the members of these 'clusters' (Queiroz et al., 2009). Summary analysis of our data set did not reveal similar functions for the genes in the different profiles. The Gene Ontology annotations of the genes falling into the different profiles were investigated (Kabani et al., 2009), but in general were not revealing due to the poor quality of the annotations. Thus, whilst it is possible to speculate on the nature of the genes, functional analyses of the genes will be required to determine whether transcripts with the same profile play similar roles or interact with each other.

The Jensen et al (2009) study used a different cell line, and the cells used to harvest the RNA for later time points in differentiation do not come from the same population as was used to harvest RNA for the stumpy form cells. Nevertheless, some of the comparisons conducted were the same in both studies, such as slender from compared with stumpy form. In this comparison, out of the 33 transcripts seen

by Jensen et al (2009) as showing an increase in expression in stumpy form cells when compared to slender form cells, 21 of these demonstrated the same expression profile in our array (Table 4.2). This gives a correlation rate of 64%, where most of the transcripts showing conflicting results were the 'hypothetical conserved' and ESAG9 genes.

| Gene | Accession number | Confirmed in this study |
|---|---|---|
| aquaporin 3 | Tb927.6.1520 | ✔ |
| ESAG4 | Tb11.03.0990 | ✔ |
| ESAG9 | Tb927.1.5220 | ✔ |
| | Tb927.5.120 | ✔ |
| | Tb927.5.4620 | ✔ |
| | Tb927.7.170 | ✔ |
| | Tb09.160.5400 | ✔ |
| | Tb09.160.5430 | x |
| | Tb09.v1.0330 | x |
| ESAG9-like | Tb09.142.0370 | x |
| | Tb09.142.0380 | x |
| glutamate dehydrogenase | Tb09.160.4310 | x |
| GPEET2 procyclin | Tb927.6.510 | ✔ |
| hypothetical conserved | Tb927.2.2140 | ✔ |
| | Tb927.6.4270 | ✔ |
| | Tb927.7.4270 | x |
| | Tb09.160.0465 | x |
| | Tb09.211.1620 | ✔ |
| | Tb09.v1.0490 | x |
| | Tb10.389.1860 | ✔ |
| hypothetical conserved, aminotransferase domain | Tb927.4.2240 | ✔ |
| hypothetical protein | Tb10.70.2840 | x |
| | Tb10.70.2850 | x |
| PAD2 | Tb927.7.5940 | ✔ |
| mitochondrial processing peptidase alpha subunit | Tb11.02.1480 | ✔ |
| 2-oxoglutarate dehydrogenase E2 component | Tb11.01.3550 | ✔ |
| protein phosphatase with EF-Hand domains | Tb927.8.1130 | ✔ |
| purine nucleoside transporter | Tb09.160.5480 | ✔ |
| pyruvate dehydrogenase complex E3 binding protein | Tb10.70.5380 | ✔ |
| RNA-binding protein RBP5 | Tb11.01.3915 | ✔ |
| succinyl-coA:3-ketoacid coA transferase mitochondrial | Tb11.02.0290 | ✔ |
| transketolase | Tb927.8.6170 | x |
| VSG-related VR2.1 | Tb11.01.4560 | x |

**Table 4.2. Comparison of stumpy enriched transcripts identified from two microarray studies**. The genes found by Jensen et al (2009) as showing increased expression in stumpy form cells when compared with slender cells were analysed in our microarray. 64% of the transcripts show the same profile in this comparison between the studies, with the majority of conflicting profiles found in the ESAG9 and 'hypothetical' genes.

The concurrence rate of 64% might appear low, however considering the differences between cell lines, isolation of parasites from blood and differences in microarray technologies used between the studies, such a level of agreement is considerable.

### *4.6 Characterisation of the stumpy enriched genes*

The extensive biological validation of the samples used to generate the RNA for the microarray and the validation of the resulting profiles against the published data, gave us confidence to investigate further some of the 'stumpy-enriched' transcripts identified. These comprised genes up-regulated only in stumpy cells, and those still up-regulated 1 hour into differentiation. Using the cut-off significance of a p value of 0.05, 65 genes met these criteria. Within this cohort, 40% were genes annotated by GeneDB to be open reading frames unlikely to code for functional proteins or sequence orphans, and a further 33.8% were genes coding for proteins conserved with those in another organism. There was no obvious pattern to the genes within these two profiles, with the remaining genes with functional annotations having wide-ranging predicted protein functions. However, three potential ion channels were identified: a chloride channel protein (Tb10.26.0220), a calcium pump protein (Tb927.5.3400), which is a homologue of the endoplasmic reticulum calcium transporters, and a gene putatively coding for a subunit of a potassium channel (Tb927.1.4450). This gene has previously been subjected to RNAi analysis in bloodstream form cells, and no growth, motility or karyotype defects were observed, however the cell line was not tested for differentiation capacity (Subramaniam et al., 2006).

In addition, two ATP-dependent DEAD/H RNA helicase genes were present in the 'up-in-stumpy' group (Tb09.211.2300 and Tb10.70.6180). That two members of this family should be enriched in stumpy form cells was intriguing; these proteins have wide-ranging functions but little is known of their role in *T. brucei*. Previously, a DEAD box protein HEL64, with similarities to the nuclear-localised mammalian p68 protein, was identified in *T. brucei*, however this protein was found to be localised to the cytoplasm so was therefore not a likely homologue of p68 (Missel et al., 1999).

Attempts to create cells null for HEL64 in procyclic form cells were unsuccessful, suggesting that this protein was essential in this life cycle stage (Missel et al., 1999). With the caveat that similarities to homologues in other organisms does not necessarily mean a shared function in *T. brucei*, a BLASTP search was conducted against Tb09.211.2300 which showed high similarities to the *S. cerevisiae* protein Spb4, whilst Tb10.70.6180 showed high sequence conservation to RNA helicase II/Gu. The yeast Spb4 protein was shown to be required for processing the 27SB pre-rRNA subunits (de la Cruz et al., 1998), and RNA helicase II/Gu has been shown in mammalian cells to interact with the ribosomal protein L4, and so aid in the production of 28S rRNA (Yang et al., 2005). Therefore, it may be possible that these transcripts enriched in stumpy form cells, are required for ribosome synthesis. However, whilst the rate of rRNA synthesis in stumpy form cells has not be tested, for mitochondrially-encoded ribosomes at least, rRNA synthesis has been shown to occur at the same rate between bloodstream form and procyclic form cells even though the protein levels differ (Michelotti et al., 1992). It is equally possible that these proteins perform roles unrelated to those in higher eukaryotes, as another DEAD-box protein in *T. brucei* has been shown to be involved in RNA editing (Missel et al., 1997). Interestingly, an unrelated DEAD/H RNA helicase in *Plasmodium* has been shown to be essential in post-translational regulation of mRNAs required for release from quiescence, through sequestering them into the cytoplasm and thus regulating the correct timing of development (Mair et al., 2006).

To begin to address the function of these stumpy enriched genes, an RNAi experiment in cultured monomorphic bloodstream form cells was carried out for the putative chloride channel gene. Firstly, a unique segment within the gene was identified in Tb10.26.0220, using RNAit (Redmond et al., 2003), and then cloned into the p2T7TAblue vector (Alibu et al., 2005), which integrates into the ribosomal spacer region. Figure 4.14A shows a Northern blot of two clones generated after transfection of an RNAi construct directed against Tb10.26.0220. Although some knockdown of the mRNA was achieved for this gene, no growth phenotype was observed in these cell lines even after six days of induction (Figure 4.14B).

Similarly, no morphological phenotype was observed. A number of explanations may account for these observations. Firstly the RNA knockdown may have been insufficient to generate a phenotype. Secondly, it may be that this gene is not required in cultured bloodstream form cells; these experiments were carried out in monomorphic culture-adapted cells, which may not correspond to the situation in pleomorphic slender cells. A third possibility for the absence of phenotype upon knockdown of this gene may stem from the fact that this chloride channel is a member of a conserved family through eukaryotes, and two related putative chloride channel Tb11.01.8420 and Tb09.211.0430, showing similar expression profiles (Figure 4.15), are present in the *T. brucei* genome. This raises the possibility that these proteins may have redundant functions.

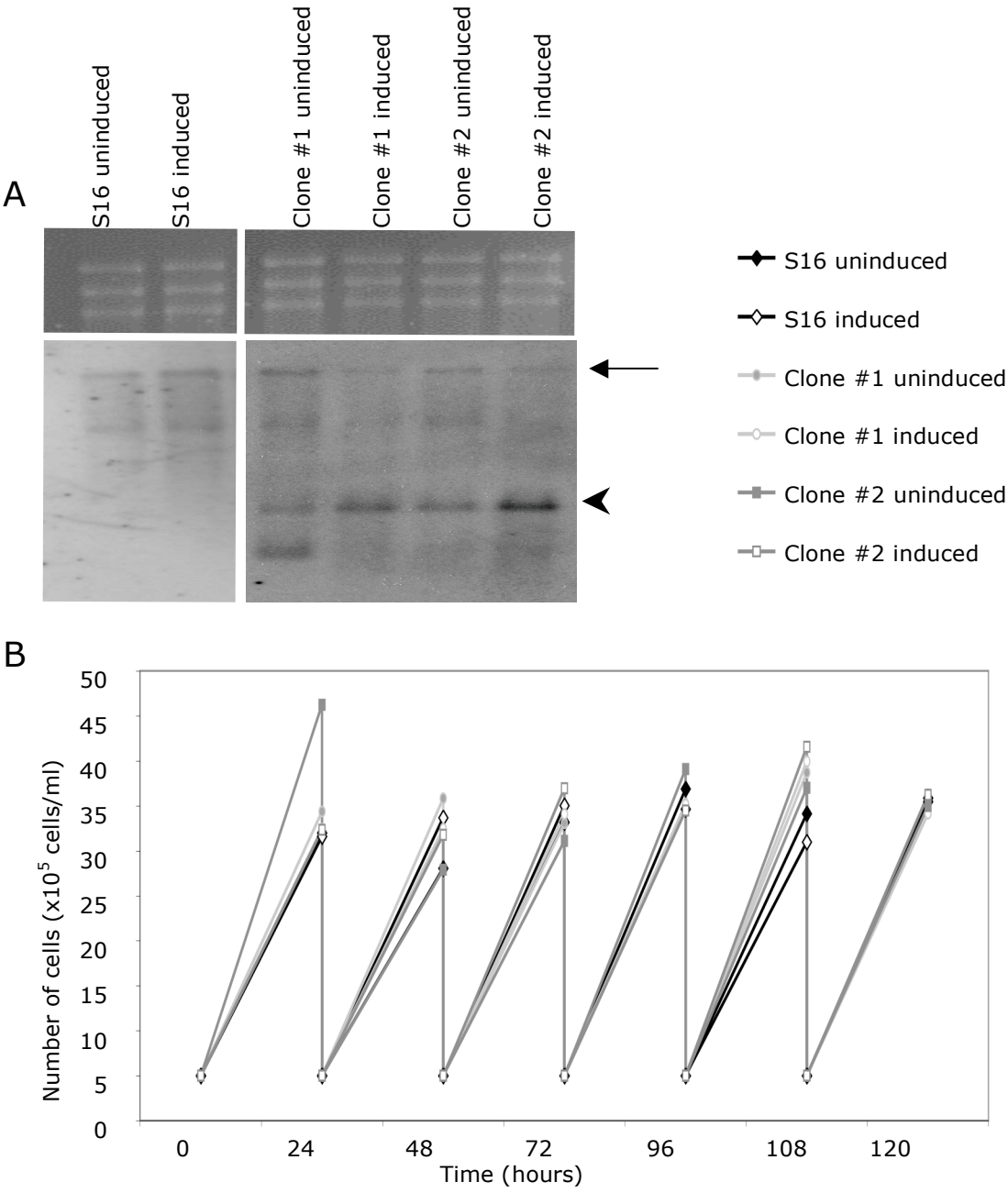



**Figure 4.14. RNA analysis and growth kinetics of Tb10.26.0220 depleted cell lines**. (A) Northern blots of the parental single marker bloodstream form cells with and without tetracycline and two clones showing inducible knockdown of Tb10.26.0220 through RNAi. These cells lines show double stranded RNA (arrowhead) even without induction through tetracycline, suggesting that the expression of the RNAi construct is occurring at low levels in the absence of induction. However, a much higher level of knockdown is seen upon treatment with tetracycline. The transcript is indicated with an arrow. (B) Growth kinetic analysis of the two clones showed no observable growth phenotype.

Using Tb10.26.0220 as a query in a BLASTP search identifies two other similar proteins, which also have conserved domains in keeping with chloride channels. Figure 4.15 shows the expression profiles for each of these putative chloride channel proteins: Tb10.26.0220, Tb11.01.8420 (30% identity to Tb10.26.0220) and Tb09.211.0430 (35% identity to Tb10.26.0220). These related genes show similar expression profiles to Tb10.26.0220 but fall below the cut-off for being designated 'stumpy enriched', that is, there is not a 2-fold difference between the expression level seen in stumpy form cells with those from the other time points. Significantly however, all three transcripts were strongly elevated in stumpy form cells, although only Tb10.26.0220 shows transient expression during differentiation, unlike the other two genes which are retained in late differentiated cells (Figure 4.15). The similarity of these proteins at the amino acid level, and their similar expression profiles, could imply that the proteins are at least partially redundant. These chloride channel proteins show good similarity to the human chloride channel 7 protein (between 22-29% identity), which has been implicated in controlling pH levels in lysosomes through its action as a Cl-/H+ antiporter to maintain a pH of 4 (Graves et al., 2008). In trypanosomes, the pH of lysosomes is kept constant in bloodstream forms at 4.84, with lysosomes playing an important role in the degradation of glycolysis products (McCann et al., 2008). Interestingly, lysosomes have been shown to co-localise to glycosomes during differentiation and glycosomes are entirely recycled during differentiation (Herman et al., 2008). Moreover, lysosomes are particularly enlarged in stumpy forms (Herman et al., 2008), so an up-regulation of proteins required for maintenance of pH in lysosomes in stumpy forms is quite consistent with these functions.

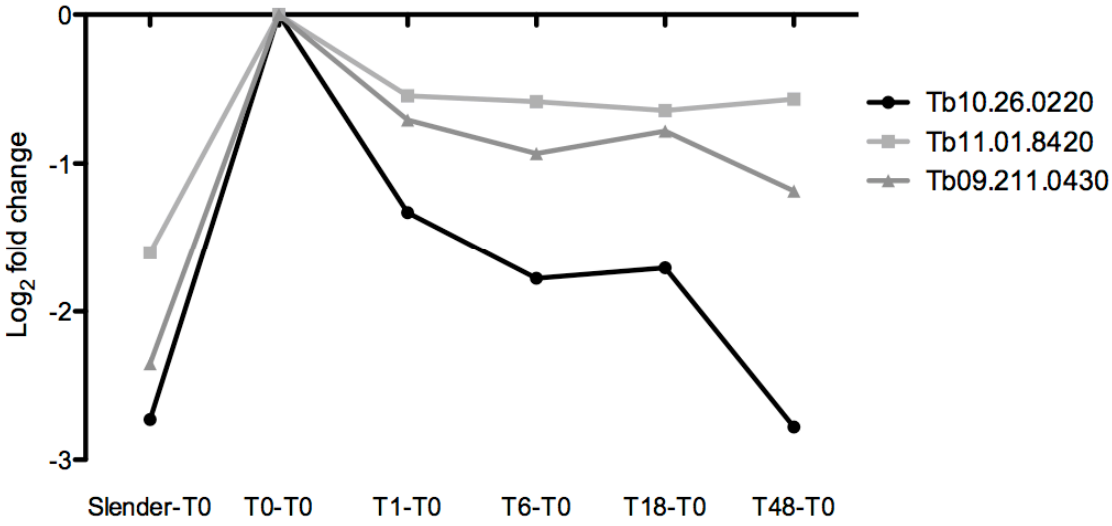



**Figure 4.15. Expression profiles of three putative chloride channels follow a similar pattern**. Tb10.26.0220 shows a highly-stumpy enriched expression profile, and two other genes that were found to be homologous via a BLASTP search show similar profiles to Tb10.26.0220 and even more similar profiles to each other.

Further investigation into these chloride channels and the other genes shown to have increased expression in stumpy form cells will be interesting, but may need to be conducted on pleomorphic cell lines in order to provoke relevant results.

### *4.7 Genes in polycistronic units do not have similar expression profiles*

As transcription in *T. brucei* is polycistronic, we analysed the profiles of adjacent genes to determine whether genes with similar expression profiles were physically clustered together along the chromosome, invoking an operonic organisation. Previous examination of genes arranged in tandem copies, such as the PGK and PAD genes, has shown that these genes show different regulation to each other between bloodstream form and procyclic form life cycle stages (Gibson et al., 1988, Dean et al., 2009). Figure 4.16 shows, as an example, the expression profiles of all genes differentially expressed to a significance of $p<0.5$ levels along chromosome 11, where genes up-regulated in comparison to stumpy form are coloured red, and those down-regulated are blue. Adjacent genes do not show similar profiles to each other, thus our analysis confirms, on a global scale, that no clustering of differentially regulated genes occurs in *T. brucei*. Along chromosome 11, the genes which showed statistically significant differential expression in at least one time point at the $p<0.5$

level show 48 different profiles; over 60% of all the profiles seen across the entire genome. 60 of the 385 genes significantly differentially regulated at $p<0.5$, show an unchanging profile (00000), hence the changes in expression levels, whilst statistically significant, are of a magnitude of less than 2-fold. The same analysis was conducted for all 11 chromosomes and each showed an absence of clustering of similarly expressed genes.

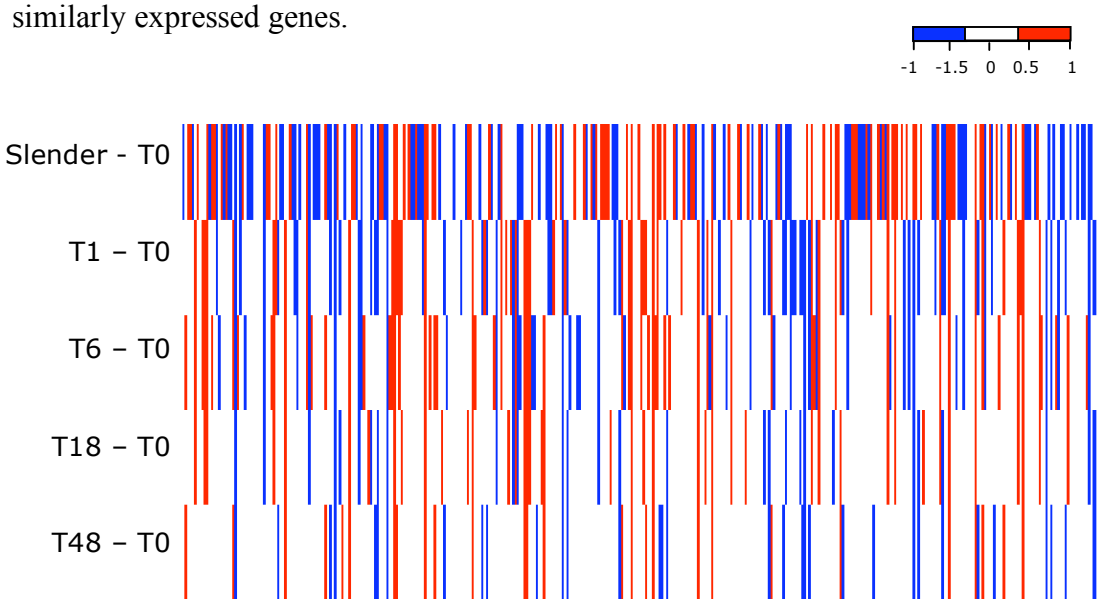



**Figure 4.16. Expression profile of the 385 genes significantly differentially expressed at p<0.5 with respective location on chromosome 11**. Genes have been assigned their trinary code, where red represents genes showing a 2-fold increase compared to stumpy forms, blue represents genes whose signal is halved when compared to stumpy forms. White represents genes that show a significant differential expression in comparison to stumpy form cells, but the change was less than 2-fold. Genes that were not significantly differentially expressed at any time point were removed from the diagram, thus this is not a scale model because spaces are not left for the genes removed.

The lack of physical grouping of genes with similar expression profiles only confirms on a genome-wide scale that which had been previously shown on a smaller scale, despite transcription in *T. brucei* being conducted polycistronically. This must mean that other mechanisms exist to regulate the genes post-transcriptionally. Several mechanisms for post-transcriptional regulation have already been shown to function in trypanosomes including: regulation of the stability of the RNA, its translation into protein, the sequestration of the protein, turnover of the protein and

action of RNA binding proteins (Walrad et al., 2009, Mayho et al., 2006, Hotz et al., 1995, Colasante et al., 2007). Genes in *T. brucei* are transcribed polycistronically, however it is rare for regulatory activity of genes to be determined by their 5'UTR, although DOT1B has two sequences in the 5'UTR which show conservation to a consensus sequence for cell cycle regulation (Janzen et al., 2006). Instead, in this organism, regulatory domains within the 3'UTR of several stage regulated transcripts have been shown to have an influence on the expression of a gene (Mayho et al., 2006, Colasante et al., 2007, Hotz et al., 1995, Berberof et al., 1995).

### *4.8 Regulatory motifs in stumpy-enriched genes*

As gene regulation in trypanosomes is almost exclusively post-transcriptional, the 3'UTRs of genes shown to be up-regulated in stumpy cells with respect to slender cells were analysed for conserved nucleotide motifs. For this, one established procedure is to analyse the first 300nt following the stop codon, this representing the average length of the 3'UTR (Mayho et al., 2006). Motifs were identified using the oligo-analysis (words) programme in RSAT (regulatory sequence analysis tools) (Thomas-Chollier et al., 2008). Using this approach, conserved hexamer and octomer sequences were sought, because this length represents the average extent of sequence conservation within RNA binding regions in other organisms. First a background frequency was established through scoring the frequency of each possible hexamer and octomer sequence present in the 3'UTRs of all genes from chromosomes 1 and 2, these comprising a total of 883 genes. The 3'UTRs of all 40 genes in the -1-1-1-1-1, that is, all genes showing a statistically significant increase in expression levels in stumpy forms, were investigated to find significantly over-represented hexamer and octomer sequences using the background frequency. No octomer sequences were found to be conserved in the 'up-in-stumpy' group, but two hexamer motifs, TCTTAC and TTCTTA were found to be over-represented within this cohort. TCTTAC was found in 17.5% of the sequences, compared to 6.2% in the 3'UTRs of all 11008 genes in the genome, and TTCTTA was present in 30% of the 3'UTRs of stumpy-enriched genes, but in only 11.5% of genes in the whole genome. These two sequences are related and often both would be present in the 3'UTR of the same gene, sometimes overlapping, such that 33% of genes up-

regulated in stumpy form cells had at least one of these motifs in their 3'UTR. The TCTTAC and TTCTTA sequences were thus designated as stumpy motif 1 (SM1) and stumpy motif 2 (SM2) respectively.

The same procedure was also carried out on genes that were down-regulated in slender forms compared with stumpy forms, regardless of the expression levels at any other time point. There were over 3700 genes which qualified as 'down in slender' when compared to stumpy form cells at a p value of <0.05, but in order to only investigate the most significant genes, an adjusted p value was calculated for each gene using the Benjamini and Hochberg method (Benjamini and Hochberg, 1995). The application of this calculation is standard in microarray analysis to reduce the false discovery rate. Briefly, the genes are sorted according to the p values generated from the moderated t-test and ranked, with the gene with the smallest p value assigned a rank of 1 up until n, then the p value of the gene is multiplied by n and divided by its own rank, resulting in an adjusted p value. Only genes that were down-regulated in slender form in comparison to stumpy forms at an adjusted p value of 0.01-0.07 were analysed, this comprising 107 genes. Again the most significant hexamer found was TCTTAC (SM1), present in the 3'UTR of 14.01% of genes. Interestingly, the location of the motif did not appear random. The motif was found in 16 of the genes, with three copies present in Tb10.70.6180, two of which overlapped and were not included in the positional analysis. There appeared to be a positional bias of the motif to between 151-200 nucleotides into the 3'UTR, with seven transcripts preferentially positioning the sequence in this segment (Figure 4.17B). Looking at all genes from the entire genome with this motif, no positional bias was detected (Figure 4.17A) with the motif found uniformly represented in each segment of the 3'UTR.

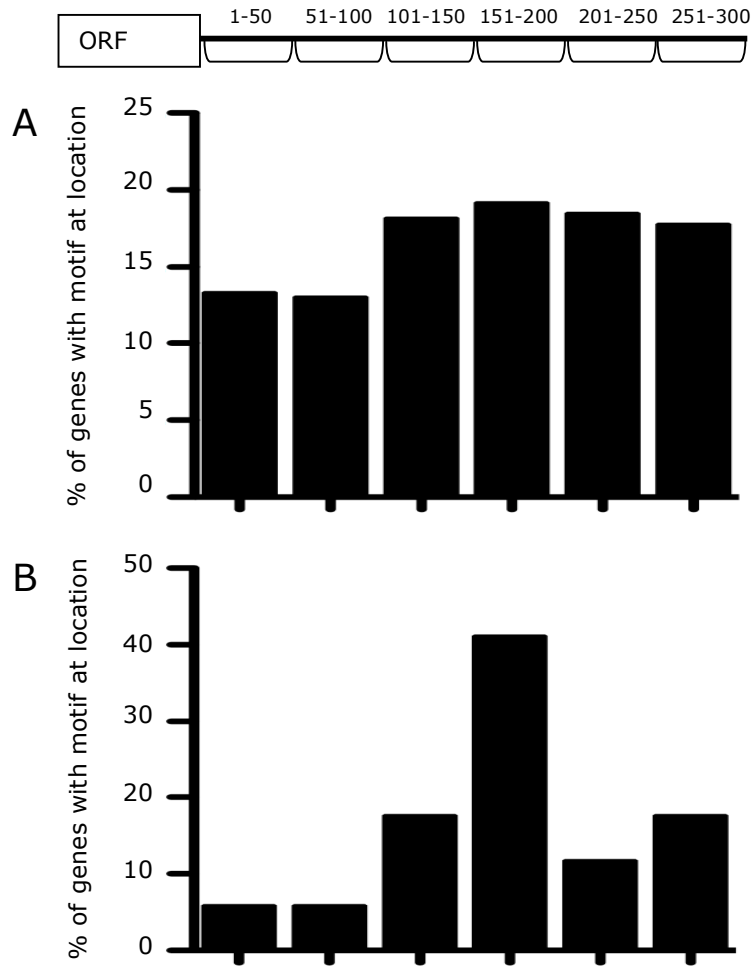



**Figure 4.17. Positional context of the TCTTAC 3' UTR hexamer**. (A) The motif is found apparently randomly distributed across the entire length of the 3'UTRs of all genes within the genome found to harbour this sequence, whilst (B) this hexamer is found to show a positional bias at 151-200 nucleotides downstream of the stop codon in the genes up-regulated in stumpy cells versus slender cells.

The preferential positioning of this hexamer within the 3'UTR adds credibility to the function of this motif as a binding region and thus suggests that it could have a role in the regulation of gene expression of transcripts required for cells once differentiated from the slender form.

### *4.9 Genes annotated as 'hypothetical unlikely'*

The presence of a large number of genes recorded as 'unlikely' to encode proteins in the stumpy-enriched set was a little surprising, but cannot be ignored due to both the

stringent statistics and the overwhelming number of them. These genes have been annotated as 'unlikely' by the gene-calling software, Glimmer, for a variety of reasons, and individually analysing them shows a spectrum of: very short genes, often clustering together or amongst VSG genes; open reading frames with very high AT content; and genes that lie at strand switch regions (Table 4.3). Strand switch regions have been implicated as potential regions of transcription initiation due to the enrichment of histone variants in the regions which open up the DNA structure, allowing easier access for the transcription machinery (Siegel et al., 2009).

In accordance with the genomic position data, 'unlikely' genes clustered together do not share similar expression profiles. Further, the quantity of 'unlikely' genes in the stumpy-enriched list is in keeping with the proportion seen in the other profiles as, out of the 557 'unlikely' genes represented on the microarray (609 genes are annotated as 'hypothetical unlikely' in the entire genome), almost 40% of them were differentially expressed in one of more timepoints at the $p<0.05$ level. Nonetheless, the 'unlikely' genes present in the stumpy-enriched group were amongst the most elevated transcripts. Interestingly, four out of the eleven 'unlikely' genes in the stumpy-enriched group contained at least one of the SM hexamers in their 3'UTR (Table 4.3).

| Accession number | Size (bp) | Other information | SM1 or SM2 motif? |
|---|---|---|---|
| Tb09.160.3000 | 195 | On opposite strand | |
| Tb09.v1.0050 | 111 | At strand switch region, amongst VSGs | SM2 |
| Tb927.1.390 | 285 | At strand switch region | |
| Tb09.160.4070 | 261 | 66% AT | |
| Tb09.211.4050 | 315 | At strand switch region | SM2 |
| Tb09.142.0190 | 270 | On opposite strand, amongst VSGs | |
| Tb09.211.1710 | 300 | On opposite strand | |
| Tb927.1.4510 | 306 | 63% AT | SM1 and SM2 |
| Tb927.1.5310 | 645 | On opposite strand | SM1 and SM2 |
| Tb09.160.2080 | 462 | On opposite strand | |
| Tb09.211.1970 | 546 | On opposite strand | |

**Table 4.3. The genes showing increased expression in stumpy form cells in comparison to all other time points of differentiation which are annotated as 'hypothetical, unlikely'**. Many of these genes are on the opposite strand of the DNA to the other expressed genes and many are less than 400bp long. Two of the genes were GC poor, which may explain their annotation status. Four genes have an SM motif, which may play a role in regulation of their expression.

It is possible that the annotation is incorrect for some of these genes; Gopal et al (2003) found that by applying different conditions to rank ORFs by their likelihood to produce proteins, they were able to reclassify a number of the genes annotated as 'hypothetical, unlikely' on chromosome 1. Using these optimised definitions, 14% of the genes met the conditions to instead be as considered as protein coding (Gopal et al., 2003). However, proteomic studies of the mitochondrion did not identify any proteins annotated as 'hypothetical, unlikely', implying that such ORFs are not translated into proteins (Panigrahi et al., 2009) nonetheless, it cannot be discounted that the transcripts instead have roles as functional RNA.

A recent bio-informatic analysis of the non-coding RNAs in *T. brucei* suggested that non-coding RNAs (ncRNA) do not play a role as microRNAs, and indeed it is unlikely that this pathway exists in trypanosomes (Mao et al., 2009). Other than acting in the RNA interference pathway, another major function of ncRNAs in higher eukaryotes appears to be in the regulation of transcription through binding to RNA binding proteins, or directly to proteins, involved in transcription initiation (Ponting et al., 2009). This is unlikely to be the function in kinetoplastids. However, it is highly possible that they are binding to proteins or RNA binding proteins, which may

allow them to complete for binding sites onto these molecules and thus aid regulation in a post-transcriptional fashion. Non-coding RNA has also been implicated in maintaining the structure of paraspeckles, putative storage vesicles for RNA held within the nucleus, as well as other cytoskeletal structures (Wilusz et al., 2009). Studies reporting functions for ncRNAs are increasing, and it is entirely possible that these 'hypothetical, unlikely' transcripts in *T. brucei* will fulfil some of these roles.

### *4.10 Analysis of lipids*

Inositol phosphorylceramide (IPC) has previously been shown by electrospray mass spectrometry (ESI-MS) to be a major component of procyclic form cells (Guther et al., 2006), but is absent in bloodstream form cells (Martin and Smith, 2006). A tandemly linked array of four genes (SLS1-4) has been implicated in IPC and sphingomyelin synthesis (Sutterwala et al., 2008). Analysis of this array by microarray demonstrated that only SLS1 was up-regulated over 6-fold between slender and stumpy form cells (Figure 4.18C); the remainder of the array showed no significant differential expression. In order to determine whether these changes in gene expression level resulted in differences in the lipid make-up of stumpy form cells, a mass spectrometry analysis of the lipids present in the monomorphic bloodstream form and stumpy forms was conducted by Terry Smith (University of St Andrews). Initial analysis used ESI-MS to detect differences in the type of all phospholipids between bloodstream form cells and stumpy form cells. Total lipids were extracted from the cells and a voltage applied to create an aerosol which could then be ionised, with the resulting ions separated via their mass to charge ratio (m/z). When no difference in lipid content was observed between stumpy and bloodstream form cells, a tandem ESI-MS/MS was conducted using, as parent ions, each phospholipid class. This allowed only one class of phospholipid to be selected at a time, which was then fragmented and analysed. With this more focussed method it was possible to distinguish differences in the phosphatidylinositol (GPIno) phospholipids, with significant peaks at 778.7m/z and 833.4m/z in stumpy form cells. This information correlates to the number of carbons and degrees of saturation of the compound: that is, the number of rings and multiple bonds present. The peaks seen in Figure 4.18B can be identified as IPC, C34:1 and C38:1, which were found to

be specific to stumpy form cells and absent in slender forms (Figure 4.18A).

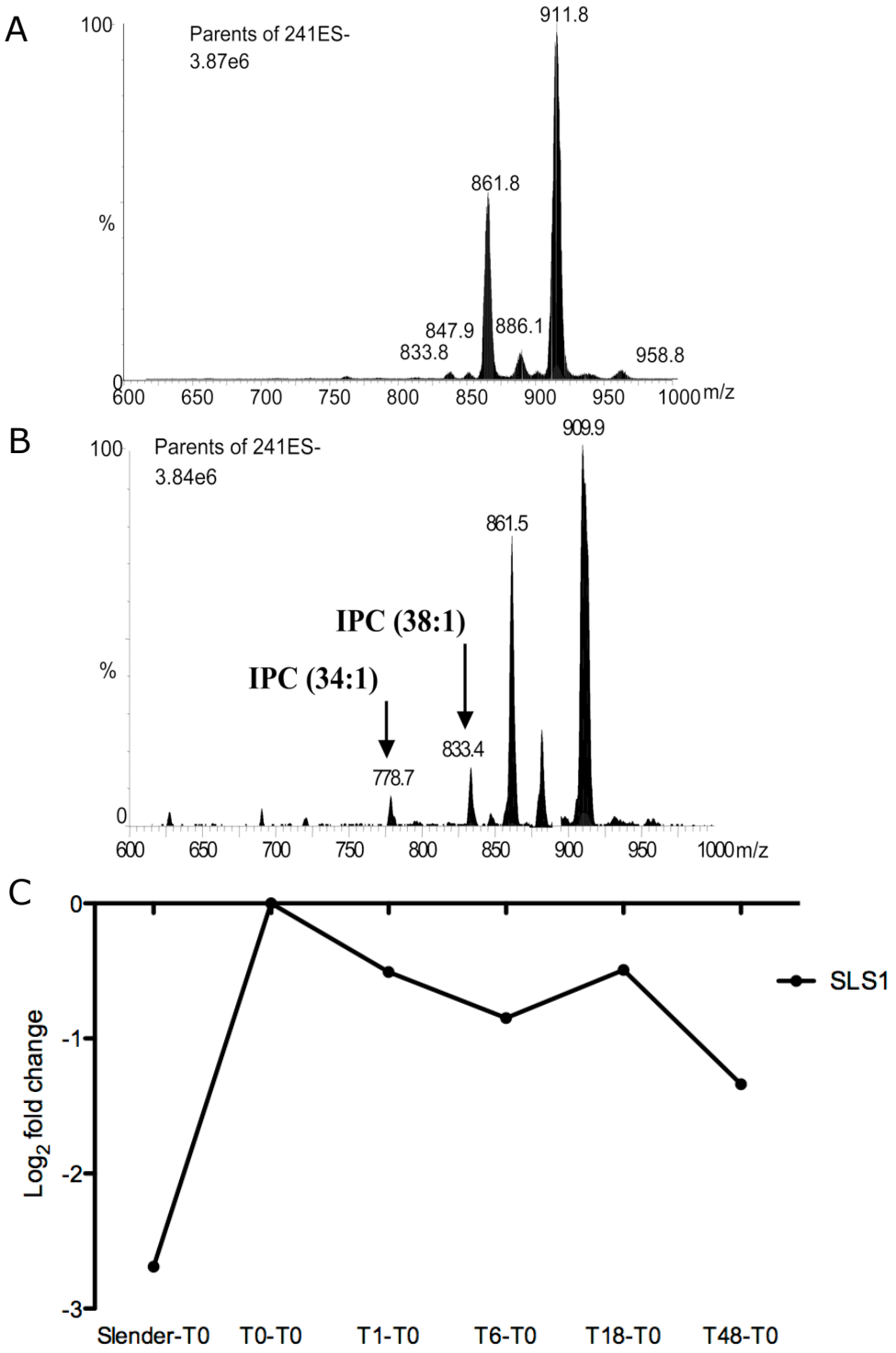



**Figure 4.18. Phospholipid composition of monomorphic slender and stumpy cells**. GPIno make-up of (A) monomorphic slender and (B) stumpy cells using ES-MS/MS. In the stumpy cells there are two significant peaks corresponding to IPC (34:1) and IPC (38:1). (C) The microarray expression profile of TbSLS1 shows up-regulation in stumpy forms.

The presence of IPC in stumpy forms suggests that IPC is being produced to prepare cells for uptake into the tsetse fly, or has a specific function in stumpy forms. Subsequent analysis of the genes in the array has shown that, whilst SLS3 and 4 are responsible for the synthesis of both sphingomyelin and ethanolamine phosphorylceramide, SLS1 is exclusively an IPC synthase (Jay Bangs, personal communication). Hence the up-regulation of this gene in stumpy forms fits well with the coinciding production of IPC, and provides biological validation of a prediction from the microarray analysis that had been previously unknown.

### *4.11 Summary*

The analysis of the microarray profiles against previously known differentially expressed genes supports their overall accuracy. However, the stringency of the statistical cut-off has meant that several genes that are known to be highly expressed in stumpy forms are missing from the list containing the genes significantly differentiation expressed at $p<0.05$. The stringency used however, means that any genes identified as differentially expressed can be confidently believed and thus a number of stumpy-enriched genes are now being pursued.

Similar trends in expression were seen for proliferation-associated and procyclic form specific transcripts, both showing rapid increases in expression immediately after cells begin differentiation. This observation is not surprising; stumpy cells are growth arrested, such that once they resume cell growth they must rapidly synthesise new histones as well as initiating developmental changes. Once this initial synchronous return to proliferation has ended, the population will remain dividing and the levels of histones will drop slightly as these are produced at a steady state level. The same sudden increase in expression can be seen for the procyclic specific transcripts as they must quickly adapt to the new environment and so produce massive amounts of the required RNAs which peak during differentiation and decrease into a stable level. Indeed, there is a global 'burst' in RNA levels immediately following release from G0 arrest, of both the ribosomal and messenger RNA (Pays et al., 1993).

The statistics applied to the microarray analysis were deliberately stringent, to take into account the multiple testing issues and to minimise the frequency of false discoveries from the results. Inevitably this will have caused us to exclude some genuine data through false negatives. However, this approach was deemed preferable to potentially investigating genes that later proved to be false positives. The consequences of this can be seen clearly in the microarray data through examining genes of known expression profiles which did not fulfil the stringent statistical criteria required to place them at an adjusted 5% significance level. One clear example of this are the PAD genes, where PAD1, which has been published previously as showing increased expression in stumpy form cells (Dean et al., 2009), did not show a significant level of differentiation expression at $p<0.05$, but did show the expected profile when the level of significance was reduced to $p<0.1$.

Despite this, we cannot rule out that in our results there will be false positives; it would be unlikely if every single gene in our significant lists acted exactly as our microarray predicts, but the unexpected profiles of the PAG genes with confirmation from an independent lab cautions against distrusting results that differ from those anticipated. A further power of the array analysis was the ability to follow transcript trends over the time course of differentiation; this provided considerable reassurance of transcript profiles over and above projects derived from a single life cycle stage

The use of qPCR and Northern blots for validation proved unreliable, as about half of transcripts correlated well with the microarray data, whilst the remaining genes had varying profiles. The simple explanation for this is that it is impossible to normalise Northern blots or qPCRs in a manner equivalent to the microarray. The normalisation for the microarray was against every single gene in the genome both within and between arrays, whilst the qPCRs were only ever normalised against one gene at a time in one sample and Northerns could only be normalised against total ribosomal RNA levels. In addition, the microarray was averaged over five biological replicates whilst, due to scarcity of samples, Northerns and qPCRs were at most conducted on only two of the replicates. Thus the microarray remains the more

powerful determinant of differential regulation, but it will be necessary to create a more suitable normalisation regime for Northerns and qPCRs or to find other methods to validate the microarrays.

# 5 – Role of G0 in Differentiation

It has been shown that stumpy cells are G0 arrested and proceed into differentiation in a synchronous manner (Matthews and Gull, 1994). However, it is also possible to trigger the differentiation of asynchronous bloodstream form monomorphic slender cells using cis-aconitate and temperature reduction (Ziegelbauer et al., 1990), and the field is still divided as to whether or not G0 is essential for differentiation (see Section 1.5). We wanted to pursue this study in a variety of ways; ie looking for homologues of proteins shown in other organisms to be required for G0 arrest, and through analysing differentiation of cells isolated from specific cell cycle stages. Historically it has not been possible to satisfactorily synchronise trypanosomes, and methods for arresting cell cycle progression using toxic chemicals were rejected on the grounds that stressed cells are not good models for differentiation (discussed in Section 1.6.1). However, using a supravital DNA dye, it was possible to sort cells into populations enriched for certain cell cycle stages.

## *5.1 FACS sorting to isolate populations in different cell cycle phases*

### 5.1.1 Experimental design

The Vybrant® DyeCycle™ stains only fluoresce when bound to double stranded DNA, and this feature can be exploited to stain the cellular DNA of cells. Three stains utilising different lasers are available, and the Violet dye was chosen as it has been shown to be less toxic when staining stem cells, both in terms of the chemical itself and through the use of the violet laser which is less harmful to cells than the shorter laser lines required for the other Vybrant® DyeCycle™ stains (Morrow, 2007). Subsequently, both the DyeCycle™ Orange and Green, despite being used successfully for procyclic form staining (Siegel et al., 2008a), have been shown to be toxic for bloodstream form cells (Michael Boshart, personal communication).

Cells were harvested from logarithmically growing monomorphic cell cultures and resuspended at $1x10^6$ in HMI-9 media supplemented with 2% FCS, since higher percentages of serum interfere with flow cytometry. Vybrant® DyeCycle™ Violet

was added to a final concentration of 1μg/ml and cells were incubated at 37˚C for 30 minutes (Figure 5.1). In order to reduce the duration of the sort, the cells were concentrated after staining into approximately 3ml of the original staining media such that their density would then be ~$1.6x10^7$cells/ml, at which density the sort was generally completed within 90 minutes. Sorts were conducted on a FACSAria (BD Biosciences) machine, with the dye excited using the violet laser and detection emission using the 450/40 bandpass filter. The dye was found to be highly photosensitive, so all experiments were conducted using minimal ambient light. During the course of a sorting experiment it was possible to observe a gradual shift in fluorescence intensity of the stained sample (compare Figure 5.4A with Figure 5.4B), thus for the duration of the sort the gates were required to be constantly adjusted to allow for the shifting position of the cells within a gate. At the end of each sort a sample of the recovered populations were re-examined to assess purity of obtained samples.

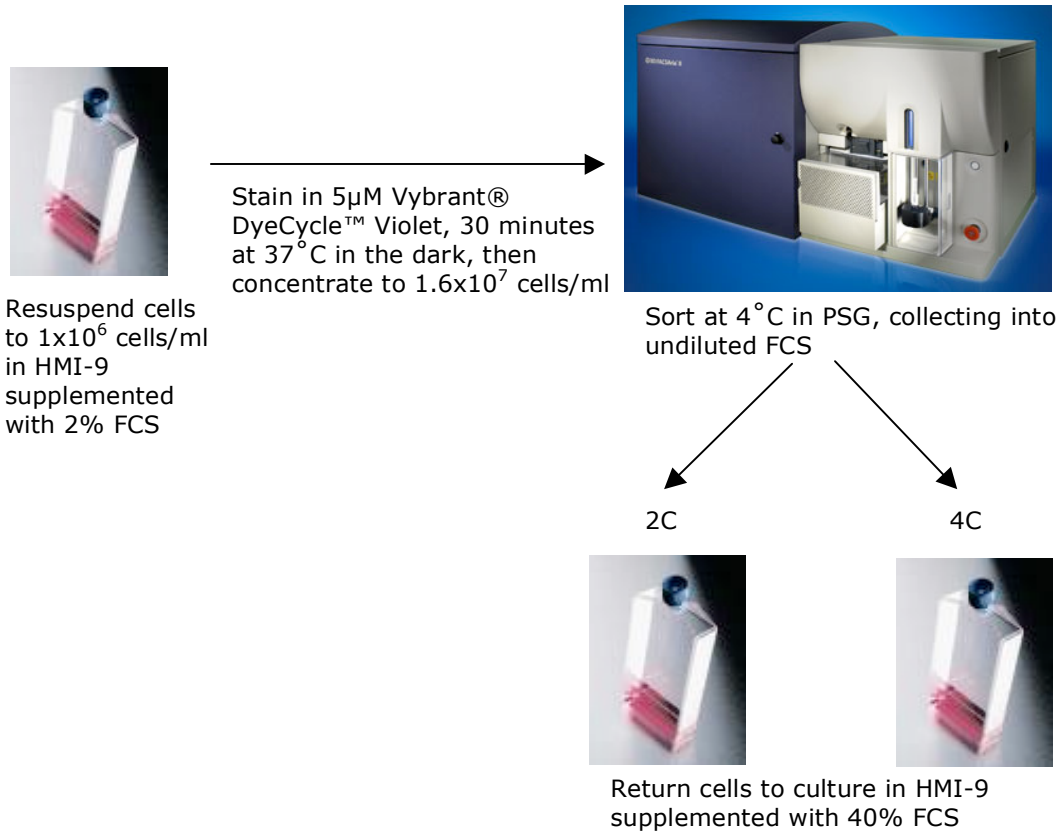


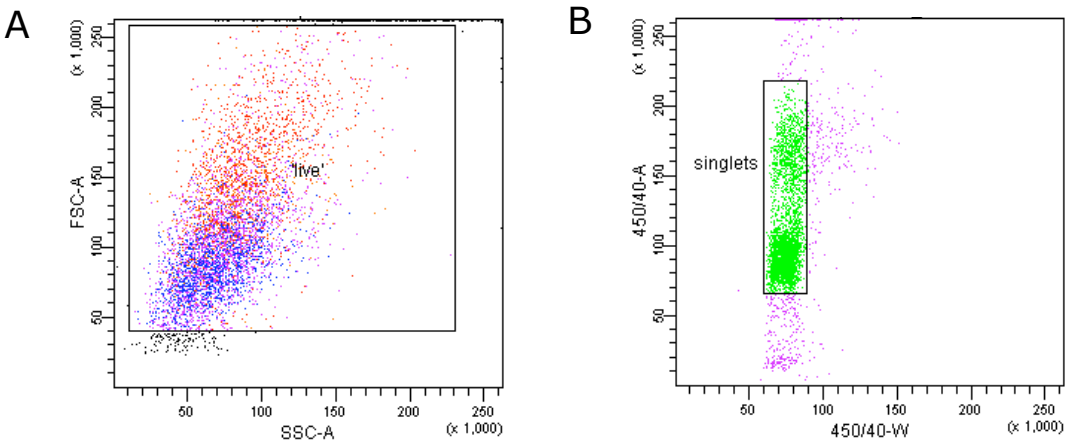

**Figure 5.1. Schematic diagram of a FACS sorting experiment**. Cells were placed into HMI-9 medium supplemented with 2% FCS and incubated with Vybrant® DyeCycle™ Violet for 30 minutes at 37˚C. The cells were concentrated into the staining medium so the density was >1.6x10$^7$ cells/ml and were sorted at 4˚C in a FACSAria machine using PSG as the sheath fluid. Two populations of cells with either a 2C or 4C DNA content were sorted into undiluted FCS, and once completed, the harvested populations were returned to culture in HMI-9 supplemented with 40% FCS.

The slightly disparate timings of replication of the kinetoplast DNA from the nuclear DNA (Woodward and Gull, 1990) should not influence the histograms generated, as kinetoplast DNA only amounts to approximately 5% of total cellular DNA, and so will not have a significant effect on the histograms (Siegel et al., 2008a). In order to limit the stress caused to cells through the procedure of staining and sorting, PSG was used as the FACS machine sheath fluid. In addition, sorted cells were collected into 1ml undiluted FCS in polypropylene tubes, the purpose of the FCS being 2-fold; firstly to help the cells recover from the dilution in PSG, and secondly, cells are prone to stick to the sides of the tube once sorted, so the FCS was used to coat the

wall of the tube and prevent adherence. The isolated populations were rapidly returned to culture following sorting, into pre-warmed HMI-9 medium supplemented with 40% FCS, to further reduce the trauma experienced by the cells. As the cells collected from the sorts were returned to culture, it was important to work under conditions as sterile as possible. Hence, the FACS machine was flushed with ethanol prior to use, the PSG sheath fluid was autoclaved and FCS in collection tubes was filter sterilised. In addition, once cells were returned to culture, the HMI-9 media supplemented with 40% FCS was prepared with fresh penicillin/streptomycin to reduce contamination. In spite of all these measures, cells returned to culture invariably showed bacterial contamination after approximately 24 hours. Presumably the addition of an antibiotic drug such as tetracycline to the media could prevent this problem, although this was not tested.

### 5.1.2 Gating of samples

It was essential that the samples prepared for a sort experiment were in a single-cell suspension to avoid mistaking clumped G1 stage cells for G2 cells. As trypanosomes are non-adherent and cells were not fixed, the cultures showed little clumping of cells, thus filtering of the cells prior to sorting was not required. Any cells that had stuck together were gated out of the analysis and were not sorted. Figure 5.2 shows the gating used for the sorts; primarily cells were plotted via FSC/SSC so that cells were distributed according to volume and density, intact cells were identified and a gate placed around them to exclude the damaged cells (Figure 5.2A). This population was then further gated by area/width to eliminate aggregated cells and cell debris such that only single cells (singlets) remained (Figure 5.2B). These gates removed around 20% of the total starting material, with the number of cells in G2/M and G0/G1 in roughly a 3:1 ratio (Figure 5.2C).




C

| Population | #Events | % Parent | % Total |
|---|---|---|---|
| All Events | 5,361 | | 100 |
| Intact | 4,691 | 87.5 | 87.5 |
| Singlets | 4,249 | 90.6 | 79.3 |
| G0/G1 | 1,881 | 44.3 | 35.1 |
| G2/M | 519 | 12.2 | 9.7 |

**Figure 5.2. Gating to acquire a singlet population**. (A) Scatter plot of all cells distributed via their volume to density (FSC/SSC) with a gate set up around the intact cells. (B) Distribution of the cells when plotted via their area/width with a gate set up to only sort non-aggregated cells. (C) Table with the proportion of cells remaining once dead and doublet cells had been removed, showing also that following sorting, G0/G1 and G2/M are recovered in an approximately 3:1 ratio.

Figure 5.2C shows that, of the 80% of cells remaining after gating, only 35% of cells were then collected in the G0/G1 fraction with just less than 10% in the G2/M fraction. The recovery rate in this experiment of 45% of the initial population was representative of all sorting experiments, which ranged from 39.5-52.8% from a total of seven experiments. Of the cells that were in the viable singlet population, over 40% did not fall between the gates for the G0/G1 or G2/m population, and so were removed.

### 5.1.3 Cell cycle specific toxicity of supravital dyes

In other cell types it has been shown that cells in S phase demonstrate higher sensitivity to DNA intercalating dyes such as Hoechst 33342 than cells in other cell cycle stages (Siemann and Keng, 1986). Therefore, cell cycle specific toxicity was examined in trypanosome cells through sorting the cells into G0/G1, S phase and G2/M populations and assessing viability as determined by cell growth for 24 hours

post sorting (Figure 5.3A). For this 3-way sorting, 5ml collection tubes were used, which unfortunately prevented each population being sorted into a single tube.

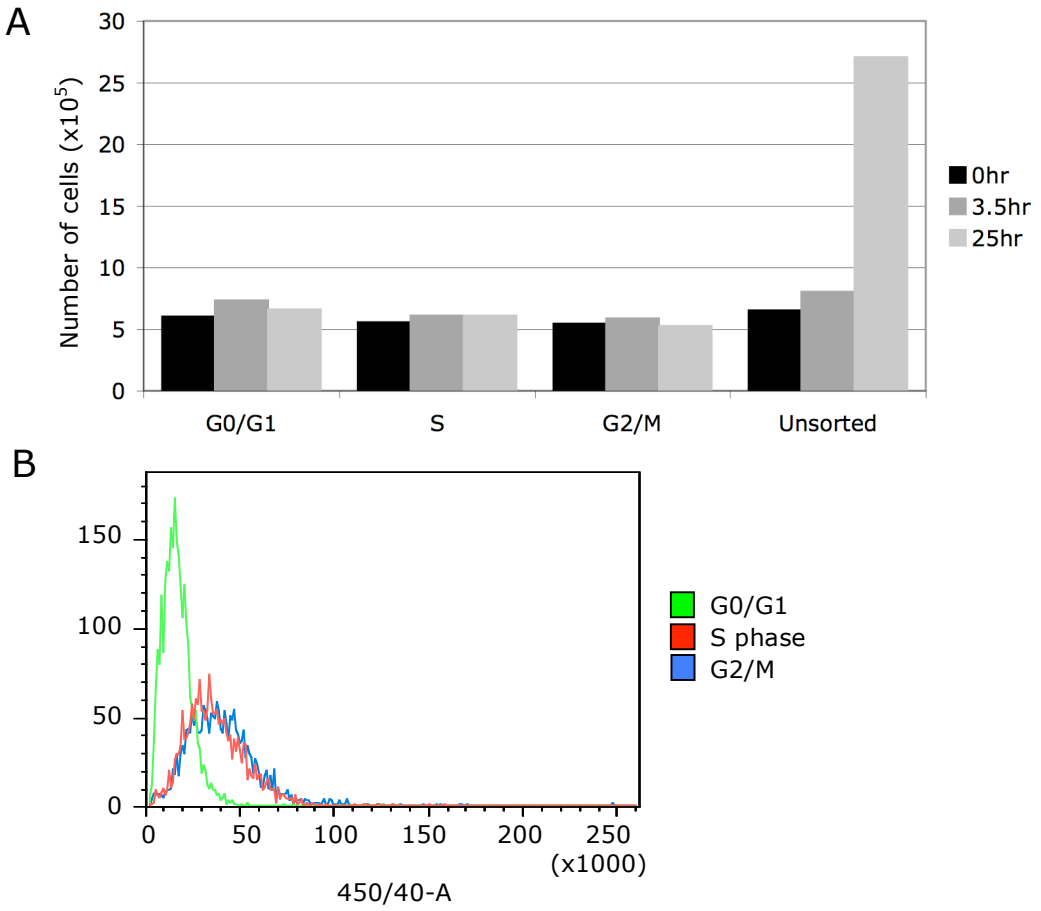



**Figure 5.3. Viability of sorted cells in different cell cycle stage**. (A) Each of the different populations show similar viability post-sorting, although all fail to show an increase in cell number compared to non-stained, unsorted cells. It is not clear whether the absence of increase in cell number is due to cell cycle arrest or an increase in cell death. However, if the reason is increase in cell death, then this is not associated with a particular cell cycle stage. (B) The post-sort profiles show that there was poor distinction between the S phase and G2/M populations.

Figure 5.3A shows that the cells from all the sorts grew poorly following the sort, with little cell growth seen 3.5 hours following the sort, and the G0/G1 and G2/M populations showing a small decrease in cell number at 25 hours post-sort. In parallel, the unsorted control cells were incubated in conditions mimicking the sort; ie they were incubated at 4˚C in HMI-9 medium supplemented with 2% FCS whilst the samples were sorted, before being returned to HMI-9 medium supplemented with

40% FCS at 37˚C. In contrast to the sorted cells, however, these cells show a clear increase in cell number after 25 hours, suggesting that the reason for the poor growth of the sorted cells was due to the stain, rather than the temperature changes and depletion of serum associated with the sorting process. The post-sorted populations taken from the S phase and G2/M fractions were almost indistinguishable on the DNA histograms (Figure 5.3B), thus the S phase population is unlikely to contain a very pure population of cells at this cell cycle stage. Nonetheless, as little difference in cell growth was seen between the G0/G1 populations and either the S phase or G2/M, it is likely that the dye does not display particular toxicity towards cells at a certain point within the cell cycle.

### 5.1.4 Isolating cell cycle specific populations

Due to the poor recovery of pure S phase populations, only the G0/G1 and G2/M populations were isolated for all future experiments. One benefit to using 2-way sorting was that it allowed the use of larger, 15ml collection tubes, reducing the number of tubes required to harvest each population. Figure 5.4A shows histograms generated from stained logarithmically growing cells and the gating used to isolate the two sorted populations.

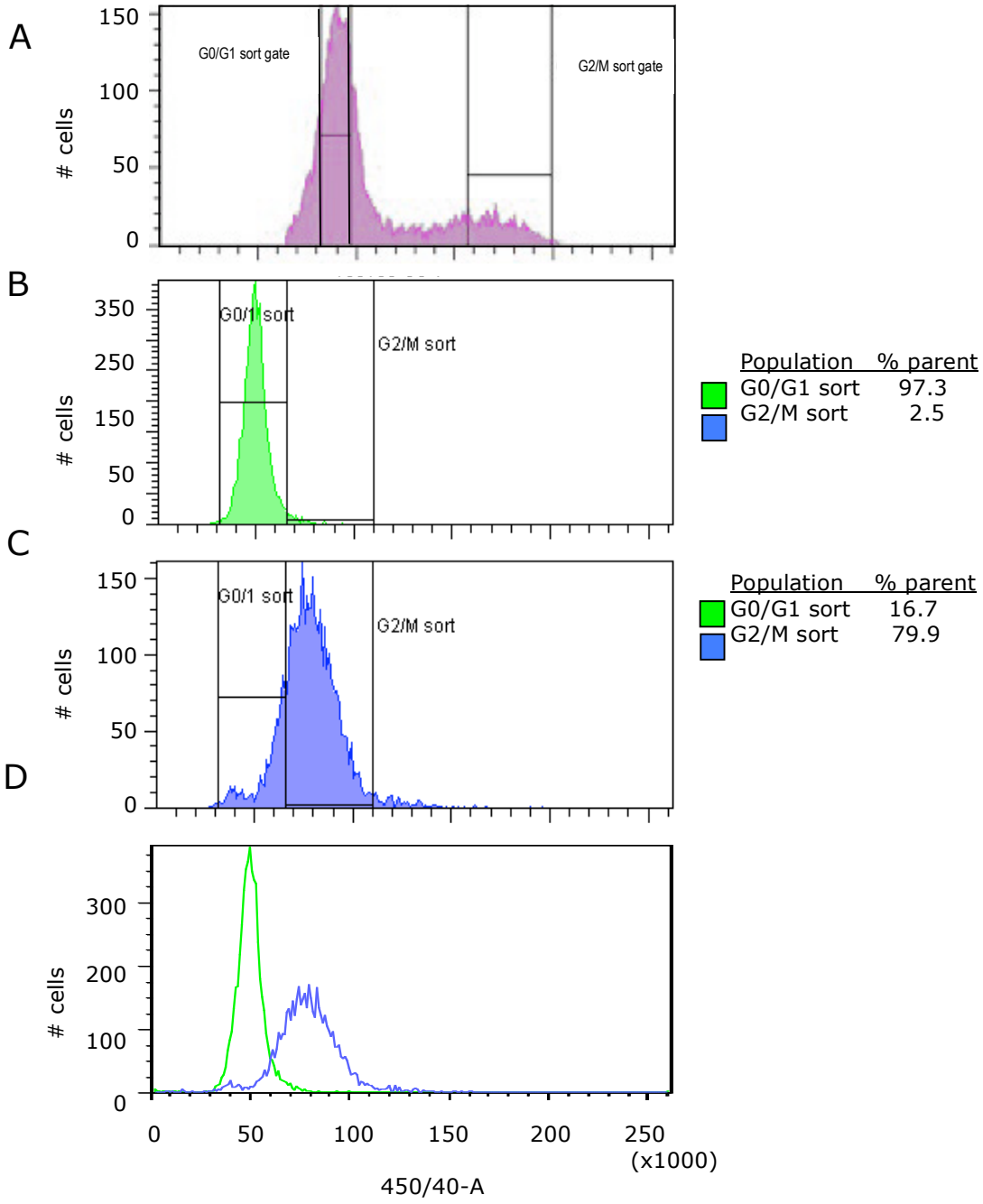



**Figure 5.4. Histograms of cells stained with Vybrant® DyeCycle™ Violet with gates set for stringent cell cycle stage specific recovery**. (A) Pre-sort profile of cells with gating showing from where the different cell-cycle populations were collected. (B) G0/G1 sorted cells, showing 97.3% cells with G1 content, (C) G2/M sorted cells with 79.9% of the population showing DNA content expected in G2/M cells, (D) overlay of B and C. The post-sort populations show a reduction in fluorescence intensity as the dye was gradually eliminated from the cells during the sort.

Nuclear G1 in *T. brucei* accounts for approximately 40% of the procyclic cell cycle, thus in a logarithmically growing culture, more cells will be in the G0/G1 cell cycle stage than in S phase and G2/M. As G0/G1 cells were in the majority in the starting population, it was possible to set extremely stringent gates around the very centre of the 2C peak (Figure 5.4A) without compromising high yields (Figure 5.4B). The 4C gate was set from the inner quartile peak to the outer peak, with a 79% pure G2/M population recovered, although overall fewer cells than were recovered for the G0/G1 population (Figure 5.4C). Overlaying the histograms of both sorted populations revealed clear distinction of the populations (Figure 5.4D), which was subsequently confirmed using microscopy. Of note is that the fluorescent intensity of the dye observed in the post-sorted population was much reduced from the initial population, with the average intensity for the G0/G1 population dropping from approximately 90,000 from the pre-sort population, down to 50,000 post-sort. This was likely due to the transient/dynamic nature of the dye binding, such that after 90 minutes, the dye had largely diffused out of the cells. The dye was also seen to be highly photosensitive, thus some of the reduction may have been due to dye bleaching from exposure to ambient light, although this was limited as far as possible.

### 5.1.5 Assessment of synchronisation

Since a method for reproducibly synchronising trypanosomes has yet to be established, especially for bloodstream form cells, the sorted populations were tracked for up to 17 hours post-sort to determine whether the cells grew synchronously. We were also interested to observe whether there was a lag period before the cells resumed replication, as has been seen for mammalian cells (Bradford et al., 2006). Hence, cells were sorted and slides prepared hourly to allow cell cycle scoring and so determine re-entry into the cell cycle. As a control to ensure that the culturing conditions and temperature changes were not having an effect on the cell cycle progression, cells that were unstained and unsorted, but were subjected to the same changes of media and incubations at 4˚C, were also analysed.

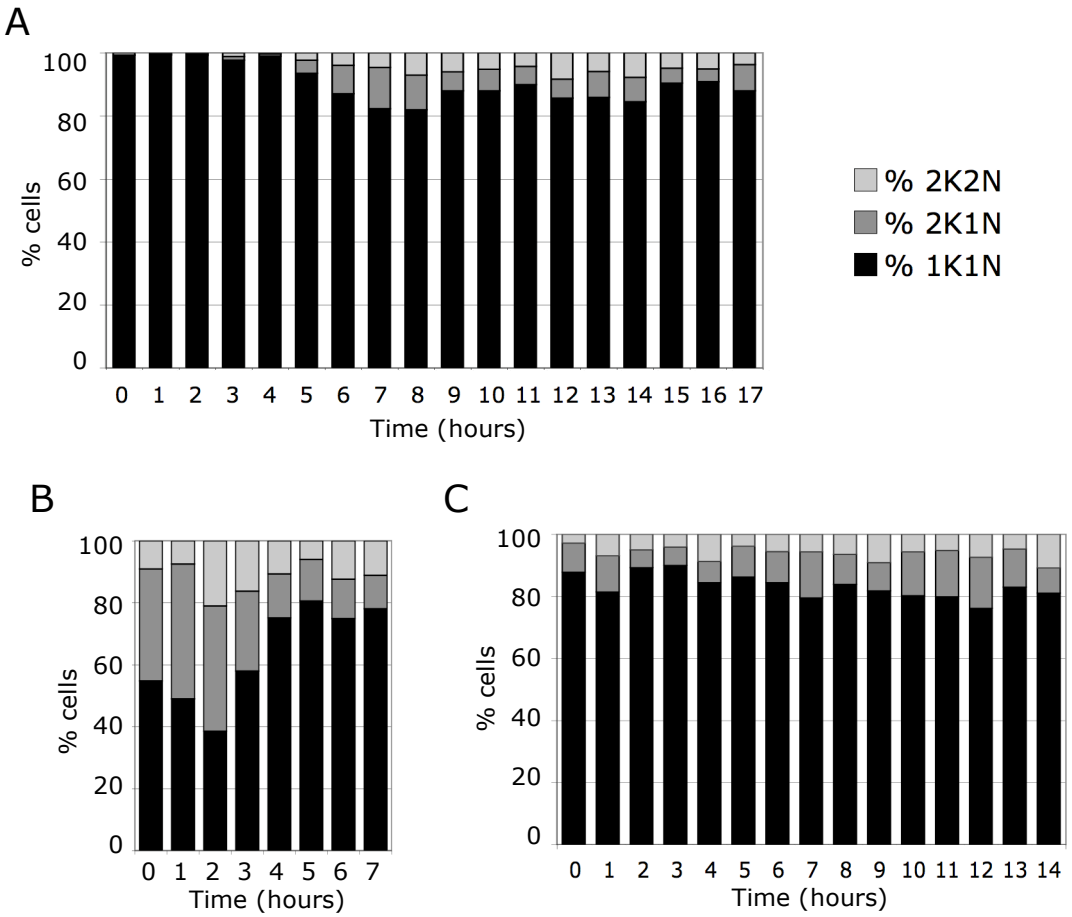



**Figure 5.5. Cell cycle progression of sorted cells**. Following a sort, every hour for up to 17 hours, samples were taken for microscopic analysis and the DNA visualised with DAPI so that the proportion of cells with 1K1N, 2K1N and 2K2N could be determined in (A) the G1/G0, and (B) the G2/M sorted populations. (C) Unsorted cells were also analysed as a control. The G2/M population was tracked for only 7 hours after recovery due to the lower number of cells harvested. For each time point, at least 250 cells were counted.

Figure 5.5A and B show that the sorted cells were not synchronised post-sort, with the cells from the G0/G1 population returning to the pre-sort ratios of cell cycle stage within seven hours, and the G2/M population returning to pre-sort ratios within five hours. Figure 5.5C, in contrast, is a good representation of the normal proportion of unsynchronised cells with the expected different nuclear/kinetoplast configurations in an actively dividing population. Hence, most cells within the culture were in the 1K1N configuration, which is seen in cells in G0, G1, S phase and early G2, thus accounting for around 70% of the entire cell cycle (Figure 5.6).

For the G0/G1 sorted population, 5 hour lag phase following sorting before the appearance of with a 2K1N configuration was observed. Cells at the end of G1 could show a second kinetoplast within ~3 hours, whilst those at the very start of G1 would take up to 6 hours before a second kinetoplast appeared (Figure 5.6). There are two possible explanations for this delay. The inability of the sorting to distinguish between cells at different positions within the G1 peak, means that cells immediately entering G1 and those about to commence DNA synthesis are indistinguishable on the FACS profile, so the first possibility for the delay in the appearance of cells with a 2K1N configuration could be because the sorted cells were predominantly in the early stages of G1. The second explanation could be that there was a delay in cells returning to a proliferative cell cycle. Such a delay would be possible, as similar studies on human cells also resulted in a lag time (Bradford et al., 2006) and could explain why cells with a 2K1N karyotype are not seen earlier. The former is the more likely reason, as the G2/M sorted cells change configuration immediately, therefore they do not suffer a lag time before resuming cell cycle progression.

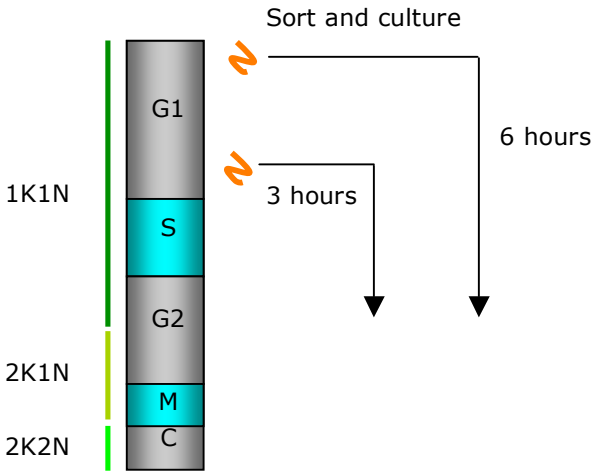



**Figure 5.6. Cells sorted from different points of G1 will progress through the cell cycle at different times**. Schematic diagram demonstrating that a cell at the start of G1 could take up to six hours before a second kinetoplast will be observed. Yet a cell towards the end of G1, which will also be sorted into the G0/G1 fraction, could produce a new kinetoplast after only three hours.

The G2/M sorted population comprised 36% cells with 2K1N and 9% with 2K2N, with the highest numbers of 2K2N cells seen after two hours, when the proportion of 2K2N cells rose to over 20%. This implies that the 1K1N cells present in the G2/M sorted population were in late S phase, thus two hours post-sort these cells would have progressed rapidly into a 2K1N configuration. Indeed, a reduction in cells with a 1K1N configuration was seen during the first three hours post-sort. Repeating the sort showed reproducible trends, with the maximum enrichment of 2K2N cells always being found two hours post-sort.

The proportion of cells in the different karyotypes in the sorted populations was highly reproducible between different experiments and different cell lines (Figure 5.7). In each instance the G0/G1 population showed a nearly 100% population of cells in the 1K1N configuration immediately following a sort (between 98-100%), while the G2/M population contained between 18-48% cells 2K1N and 6-26% 2K2N.

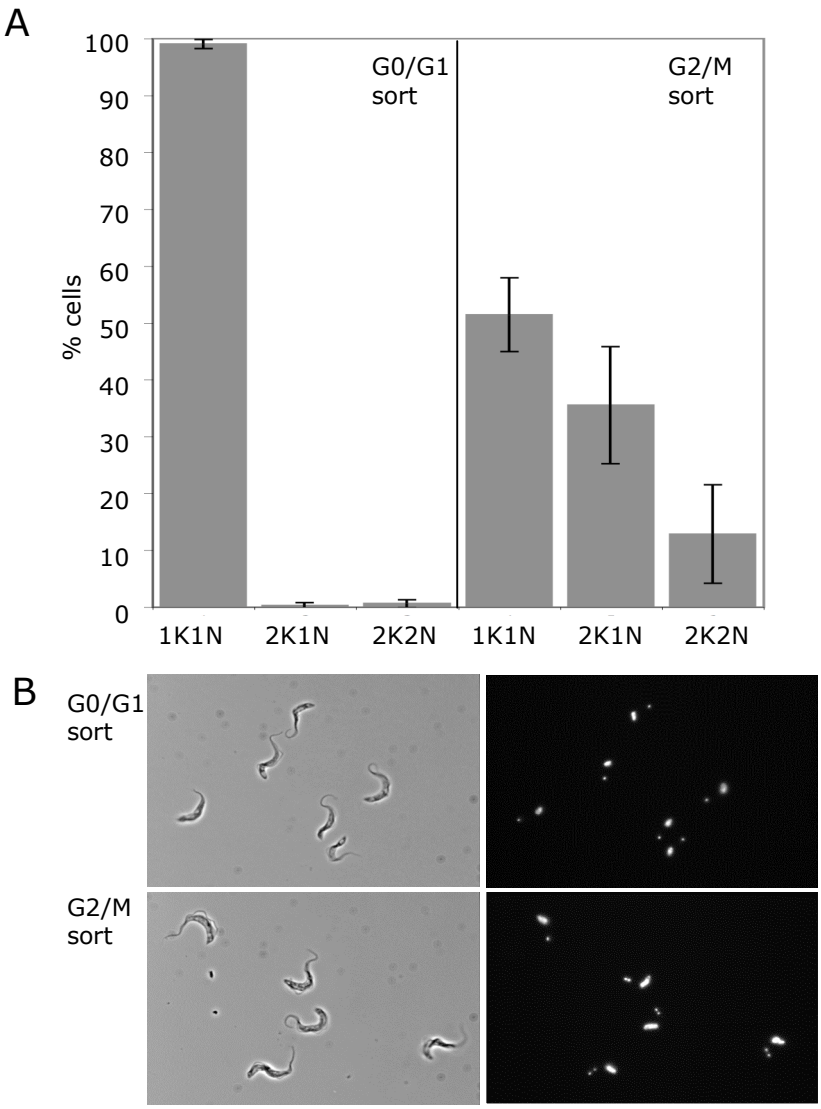



**Figure 5.7. Analysis of the recovered cells post sorting**. (A) Cell cycle configuration of cells immediately after sorting as determined from fluorescence microscopy of DAPI stained cells. At least 250 cells were counted for each slide with the data representing the mean of seven independent experiments with error bars showing the standard error. The G0/G1 sort generated a very pure population of cells with 1K1N reproducibly, whilst the G2/M sort yielded a mixed population showing enrichment in 2K1N and 2K2N cells. (B) Microscope images of cells immediately post-sort for G1/G0 (upper panel) and G2/M (lower panel) sorted populations. Phase contrast images show the morphology and integrity of the cells whilst the DAPI images showed that the G0/G1 population only had cells with 1 kinetoplast and 1 nucleus, whilst the G2/M population had many cells with 2 kinetoplasts. All cells appeared intact post sorting.

### 5.1.6 Viability of sorted cells

The viability of the cells post-sorting was essential in order to conduct experiments on the isolated populations. The redox indicator alamarBlue® was used as a means to track cell viability following a mock-sorting experiment. Thus, cell viability after treatment with Vybrant® DyeCycle™ Violet was tested through recreating the temperature and media changes that occur in the FACS experiment, although the cells were not sorted, only diluted into PSG and finally returned to HMI-9 supplemented with 40% FCS. The importance of temperature on viability was also tested. To prevent cells progressing through the cell cycle during the sort, the FACS machine was cooled to 4˚C prior to sorting the cells, thus it was estimated that the temperature of the samples would fall from 37˚C to below 20˚C. However, the fall in temperature might have had deleterious effects of the viability of the cells, therefore the viability of cells was assessed at both 37˚C and 4˚C.

Following the mock-sort, cells were incubated with alamarBlue® to test for metabolic activity with or without the Vybrant® DyeCycle™ Violet, at the experimental temperature of 4˚C and also as a control, at the preferred temperature for viability, 37˚C. Hence, 100µl volumes of the samples were loaded in triplicate into 96-well plates and the colour change occurring from the reduction of the alamarBlue® inspected after 24 hours. Control wells of unperturbed cells growing in HMI-9 supplemented with 20% FCS were also included to assess the extent of any loss of viability, and wells containing media without cells and without alamarBlue® were also included to allow quantitative analysis of the change in alamarBlue® absorbance.

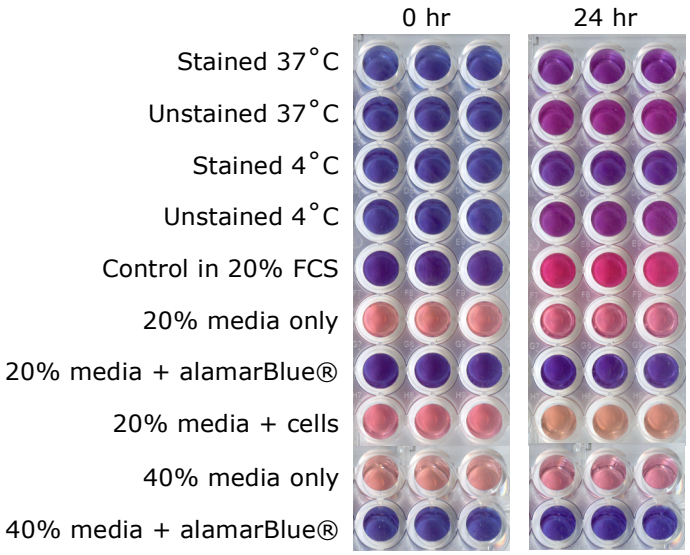



**Figure 5.8. Cell viability of Vybrant® DyeCycle™ Violet-stained cells measured using alamarBlue®.** alamarBlue® in its non-reduced state is blue as can be seen for the 0 hour wells, but once reduced through the action of cell metabolites the dye turns pink. On the 24 hour plate, it is possible to see that only the control wells have shown an obvious colour change. For the calculations to plot the reduction of the alamarBlue®, the control wells using media only, media + alamarBlue® only and media + cells only were included in order to subtract background changes in absorbance. The unsorted control cells were grown in HMI-9 supplemented with 20% FCS whilst the experimental cells were grown in 40% FCS supplemented medium, thus controls for each for included.

Figure 5.8 shows that the metabolic activity of cells treated with Vybrant® DyeCycle™ Violet, both and 4˚C and 37˚C is much less than in unstained cells. The rate of metabolism was then quantified by taking an absorbance reading of the reduction of alamarBlue® over a time course. Plates were loaded into a BioTek® Elx808 plate reader at a constant temperature of 37˚C, sealed to prevent evaporation of the sample, and readings taken every hour for 24 hours. In addition, cells were counted on a Beckman Coulter counter at 0 and 24 hours after resuspension in 40% FCS-supplemented media. Figure 5.9A shows that the differences between the metabolic activities of the cells at the different temperatures was very slight, but that treatment with Vybrant® DyeCycle™ Violet dye was inhibiting metabolism. Interestingly however, the cell counts revealed that cell growth did not correlate well with cell metabolism (Figure 5.8B). Nonetheless, from this data it was concluded

that treatment with Vybrant® DyeCycle™ Violet at a sorting temperature of 37˚C was more detrimental to cell growth than at 4˚C.

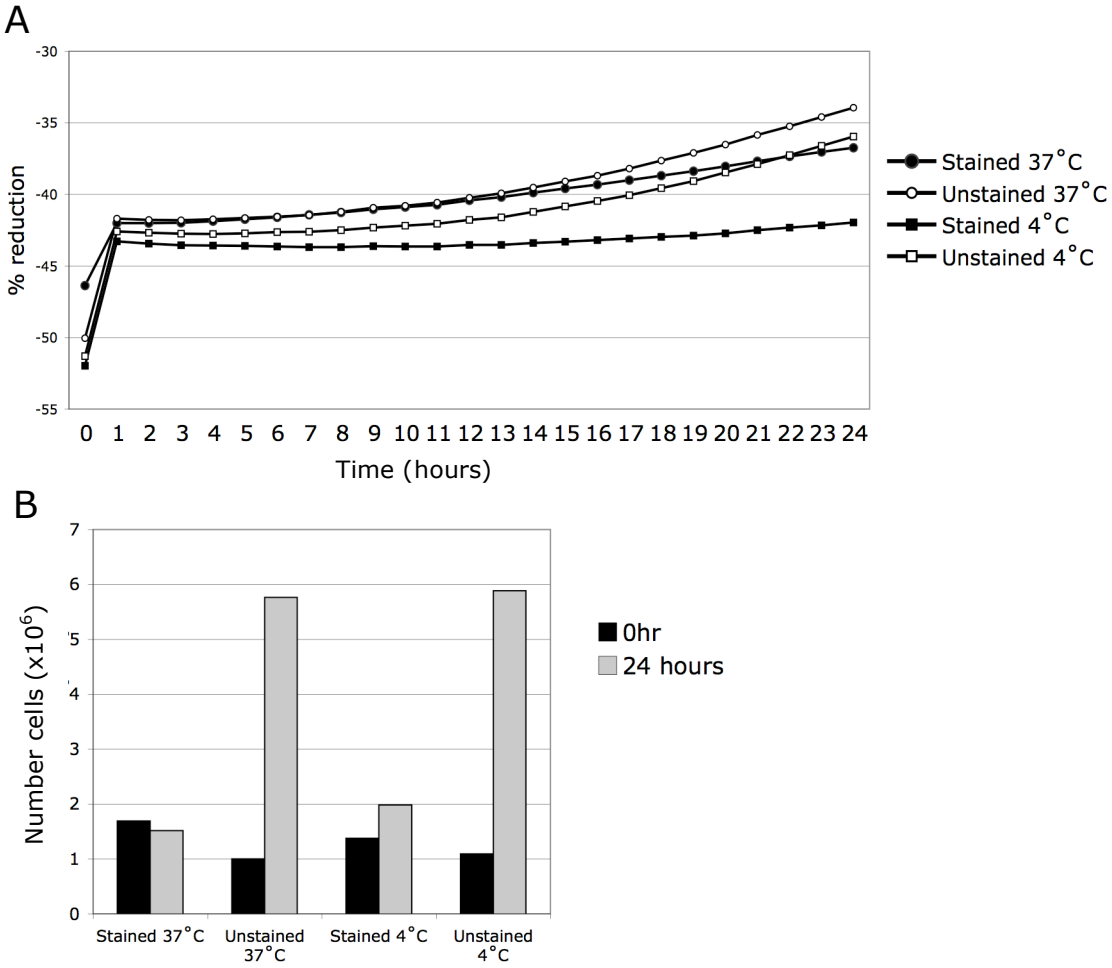



**Figure 5.9. Assessment of cell viability post sorting**. (A) Metabolic activity of cells as determined through reduction of alamarBlue®, comparing stained and unstained samples incubated at 4˚C (squares) and at 37˚C (circles). In both cases the stained sample showed lower metabolic activity, with the samples stained and incubated at 4˚C showing the lowest levels of metabolism. (B) Cell counts conducted on the same samples as in (A) taken at 0 and 24 hours.

Figure 5.9A shows the effects of Vybrant® DyeCycle™ Violet staining on the metabolic activity of cells incubated at either 4˚C, the preferred temperature to prevent cell cycle progression during the sort, and at 37˚C, being the preferred temperature of the cells for viability. Comparisons of both unstained samples incubated at 37˚C (empty circles) and at 4˚C (empty squares) reveals that whilst the cells incubated at 4˚C show a lower reduction of alamarBlue®, the increase seen

over the time course is exactly in parallel with the cells incubated at 37˚C. Thus, although the cells initially have a lower metabolic activity, they are proliferating at the same rate. However, comparison of the unstained samples with the stained cells at either temperature shows that the cells treated with the dye have impaired growth. The stained cells incubated at 37˚C diverge from the unstained cells after around 13 hours, whilst the stained cells incubated at 4˚C diverge much earlier at around 6-7 hours. These results would seem to suggest that the cells stained and sorted at 4˚C would be less healthy and grow more slowly than those incubated at 37˚C, however the cell counts carried out on the same samples contradict this hypothesis (Figure 5.9B). Using cell counts as the indicator of cell viability suggests that staining cells at 37˚C is detrimental, whereas staining cells at 4˚C, although still severely inhibiting growth, does not irreversibly prevent cells from replicating. Interestingly, the unstained samples showed similar growth at both 4˚C and 37˚C, thus demonstrating that the limited time that the cells spend at 4˚C during sorting should not have any impact on their subsequent proliferation.

Whilst relative absorbance of alamarBlue® has been shown to be a good indicator of cell number in primary cell cultures (O'Brien et al., 2000), we have found that this is not the case for trypanosomes. It is likely that the toxic properties of the Vybrant® DyeCycle™ Violet on the cells had indirect effects of the metabolism of cells, particularly the cells exposed to the dye and kept at 37˚C. Thus in these cells there is no clear correlation between cell number and alamarBlue® reduction. Other research groups have also questioned the value of using alamarBlue® on bloodstream form trypanosomes for studying the growth kinetics of drug-treated cells (Gould et al., 2008). Instead, alamarBlue® was considered better for end-point analyses, where the alamarBlue® was added only at the end of the treatment and used to compare the treated cells against controls (Gould et al., 2008). In addition, prolonged treatment of bloodstream form trypanosomes with alamarBlue® has been shown to reduce the cell doubling time, meaning results could be complicated by the effects of the alamarBlue® itself (Sykes and Avery, 2009). Thus for this experiment, using the cell counts as the more reliable indicator of cellular proliferation, it

appeared that prolonged incubation at 37˚C of cells treated with Vybrant® DyeCycle™ Violet was toxic to the cells, and resulted in disruption of their metabolism.

To further dissect the importance of temperature on the viability of sorted cells, instead of a mock sort, a more extreme version of the normal sorting experiment was conducted. For this experiment, in addition to the usual practice of regulating the temperature of the collection tubes, the temperature of the sheath fluid was also set prior to sorting, either cooled on ice to 4ºC or warmed to 37˚C. Using a sorting temperature of 4˚C, ten million cells were stained, resuspended in a 1ml volume of the staining media and sorted within 1 hour, with cell counts conducted for 24 hours afterwards. Attempting the same experiment but prewarming the sheath fluid and carrying out the sort at 37˚C with 10 million cells resulted in the recovery of only 87,000 cells, suggesting that the toxicity of the dye is high and even prolonging the time the cells spend in the dye from 30 to 90 minutes at 37˚C, results in significant cell death in the population.

These results confirmed that Vybrant® DyeCycle™ Violet was toxic to cells at 37˚C and thus the optimal method for use was to limit the time the cells were exposed to the dye and to keep the temperature of the cultures low during the sort procedure.

### 5.1.7 Differentiation capacity of sorted populations

The sorting experiments demonstrated that it was possible to recover enriched populations of cells in G2/M and very pure populations of cells in G0/G1. Therefore, we sought to discover the differentiation capacity of the two populations. Cells were sorted into G0/G1 and G2/M populations and re-cultured at a density of $5x10^5$ cells/ml and induced to differentiate through addition of 6mM cis aconitate with temperature reduction to 27˚C. Cell samples were taken from the cultures every four hours and fixed onto slides, and immunofluorescence analysis conducted to examine the expression of EP procyclin as a differentiation marker.

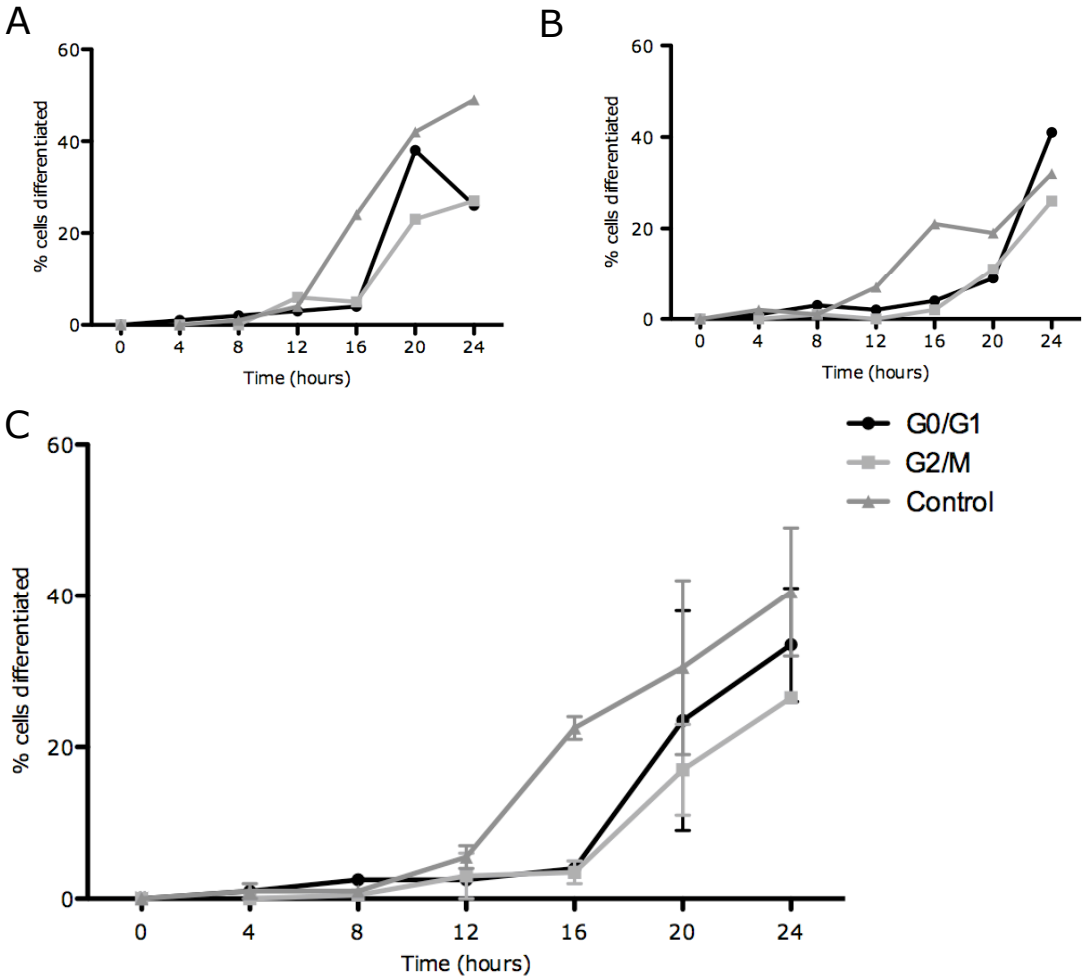



**Figure 5.10. Differentiation kinetics of cells sorted into different cell cycle stages**. (A) and (B) show the percentage of cells differentiated in two separate experiments of cells sorted into G0/G1 and G2/M populations and then differentiated through addition of cis aconitate and temperature reduction. Control cells are those which have neither been stained with Vybrant® DyeCycle™ Violet nor sorted. (C) The averages of the two experiments with error bars representing the standard error.

Figure 5.10 shows the percentage of cells that have differentiated during 24 hours after sorting of the cells into different cell cycle positions. Both the G0/G1 and G2/M populations differentiated with similar kinetics, although the control cells, which had not been stained with Vybrant® DyeCycle™ Violet or sorted, started to differentiate four hours earlier and did so in a more efficient manner. However, Figure 5.10A demonstrates the difficulties of interpreting data taken from slides, as the G0/G1 sorted population showed a decrease in the number of cells expressing EP

procyclin between 20-24 hours. This is almost certainly the result of a technical error such as poor staining of the entire slide or a section of the slide at the 24 hour time point, rather than a true biological event. This possibility is given more credence because the control cells in Figure 5.10B show a similar trend between 16-20 hours, yet those cells were unperturbed and should show normal differentiation kinetics. One of the limitations of detecting cell differentiation from slides is that it is possible to get uneven staining of slides, and the scoring of cells is subjective. A potentially superior method of detecting differentiation is using FACS. However, until (i) G0 populations can be separated from G1 populations, and (ii) increased cell viability post-sort can be obtained, it will still not possible to conclusively determine using this methodology, whether cell cycle position is a factor in differentiation.

### 5.1.8 Summary

The method described here for sorting cells according to their natural cell cycle position allows recovery of pure G0/G1 populations and of G2/M populations highly enriched for cells with 2K1N and 2K2N. This method is not capable of creating synchronised populations, but does however, allow significant enrichment of cells in a given cell cycle stage, with a yield appropriate for harvesting RNA or protein from the samples. This opens several opportunities for studying cell cycle regulated events. However, the problem associated with cell viability post sort render this approach in its current form of limited value for studying the cell-cycle dependence of differentiation. In order to determine with more precision the cell cycle position of the cells, other groups have used quantitative DAPI-imaging of the nucleus (Siegel et al., 2008a), or assessment of nuclear shape (Forsythe et al., 2008), both of which are time-consuming and thus not appropriate for experiments analysing large numbers of cells. Other methods include using antibodies against proteins expressed in a cell cycle specific manner, such as the antibody against the non-acetylated form of histone H4, which is only present during S phase (Siegel et al., 2008b). Unfortunately, there are no known markers for G0 in *T. brucei*, which would be the most useful for our experiments into the role of G0 in differentiation.

## 5.2 Target of Rapamycin

The target of rapamycin proteins in other organisms are known to be involved in G1 arrest. Therefore a *T. brucei* homologue was sought in order to determine whether this protein was required for the cell cycle arrest accompanying the formation of stumpy form cells. Also, it was hoped that perturbation of these molecules might assist analysis of the proposed link between cell cycle position and differentiation capacity.

### 5.2.1 Identification of Tor homologues

The full-length human mTor protein sequence was used as a template to query the *T. brucei* protein database using BLASTP at GeneDB (http://www.genedb.org/genedb/tryp/). The closest match was encoded by Tb10.6k15.2060, showing a sequence identity of 22% and the second most related protein was encoded by Tb927.4.420. Using only the rapamycin binding domain (FRB) domain of mTor as a template, the product of Tb927.4.420 showed the highest level of homology; thus this protein has the more conserved FRB domain, although its over-all homology to mTor was lower. Both these proteins were identified in the Barquilla et al (2008) study, as were two additional Tor-like proteins, so named because these proteins either lacked, or had additional, domains to those in the other Tor proteins characterised (Barquilla et al., 2008). However, the proposed TbTor-like 2 protein (Barquilla et al., 2008) lacks an FRB domain, and thus would be more accurately described as a PIKK, rather than a Tor, given that the only feature setting Tor proteins apart from PIKKs is the presence of a rapamycin binding domain. In accordance with Barquilla et al (2008), Tb10.6k15.2060 is hereafter referred to as TbTor1 and Tb927.4.420 as TbTor2.

Using TbTor1 as the template, the protein was used to query the entire protein database at NCBI. TbTor1 demonstrates most homology to the *L. major* orthologue, LmjF36.6320 (with 50% similarity), but after this the greatest homology was with *S. cerevisiae* Tor2 (with 27% similarity). The level of similarity was not constant across the length of the protein, however, with an increase in similarity being

detected towards the C terminus (Figure 5.11). The C terminus is the most evolutionarily conserved region due to the presence of the catalytic and regulatory elements. In comparison, the N terminus harbours the HEAT repeats, which might be considered more prone to evolutionary drift, particularly if the proteins they bind to are different or show low levels of conservation between species.

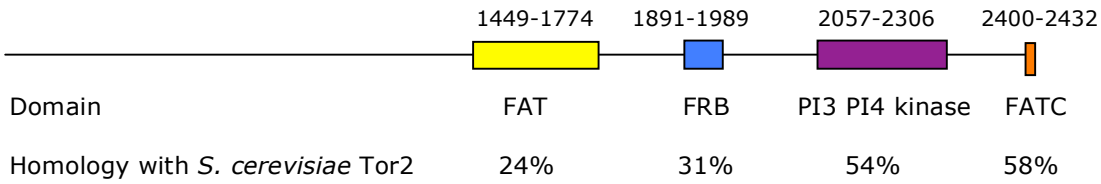



**Figure 5.11. Homology of TbTor1 with *S. cerevisiae* Tor2**. The overall identity of these two proteins is 27%, but the individual domains of the protein show increasing sequence identify towards to C terminus. The numbers above the domains relate to their position within the TbTor1 protein.

The gradual increase in conservation towards the C terminus of the protein is in keeping with results found in the levels of conservation between yeast Tor2 and human mTor (Brown et al., 1994, Sabatini et al., 1994), confirming the importance of the catalytic domains of the protein.

### 5.2.2 Lifecycle expression of TbTor1 and TbTor2

Due to the role of Tor in cell cycle arrest, the levels of TbTor1 and TbTor2 mRNA were investigated in different lifecycle stages to evaluate whether the genes were more highly expressed in the growth arrested stumpy form of the parasite. Hence, riboprobes were designed against unique sequences of the genes and used to probe Northern membranes containing RNA from *T. b. brucei* EATRO 2340 slender and stumpy form cells and cultured Lister 427 procyclic forms.

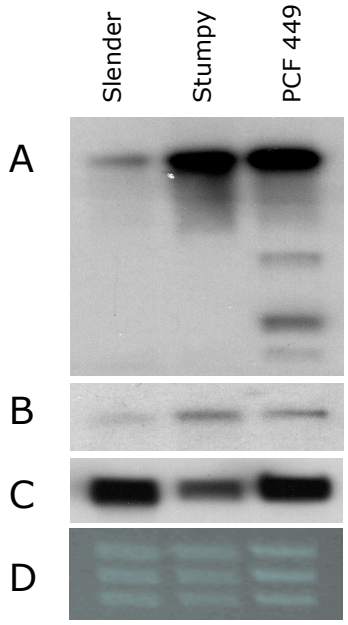



**Figure 5.12. Developmentally regulated expression of Tor homologues in *T. brucei*.** Northern blots of bloodstream form slender and stumpy EATRO 2340 cells derived from a rodent infection, and established procyclic form cells probed with (A) TbTor1, (B) TbTor2 and (C) alpha tubulin, used as a transfer control. (D) Ethidium bromide staining of the ribosomal RNA was used as a loading control. TbTor1 is up-regulated in stumpy form cells and maintained in procyclic form cells with additional smaller bands, possibly representing degradation products. TbTor2 follows the same trend, with very low levels of expression seen in bloodstream forms.

Figure 5.12 shows the life cycle specific expression of the *T. brucei* homologues of Tor1 and Tor2, with TbTor1 showing less expression in slender form cells, this being substantially increased in stumpy form cells and with similar levels seen in procyclic form cells (Figure 5.12A). The TbTor2 expression follows a very similar profile, although the hybridisation intensity was much lower than for TbTor1 (Figure 5.12B), perhaps suggesting lower expression. Tubulin expression was used here as a control and showed the expected profile, with down regulation in stumpy forms to coincide with growth arrest (Figure 5.12C). The TbTor1 Northern also shows numerous smaller bands in the procyclic form sample only, which might be the result of degradation of the mRNA. A BLAST search conducted against the region covered by the riboprobe did not show any homology with other genes, so this probe ought to show good levels of specificity. Therefore the smaller products should not have been

the result of cross hybridisation of the probe. Interestingly, the RNA expression profile seen here conflicted with the protein levels reported by Barquilla et al (2008), where the amount of TbTor1 and TbTor2 protein were found to be equal in both bloodstream form and procyclic forms (stumpy form cells were not included). Thus it is possible that the mRNA is differentially translated in the different cell cycle stages, although differences between the laboratory lines could also be responsible.

mTor mRNA has been shown to be expressed in all mammalian organs examined, however the level of expression varies dramatically (Kim et al., 2002, Murakami et al., 2004). In addition, the levels of both mTor mRNA and protein show no change throughout the cell cycle (Liu et al., 2007). Conversely, Tor transcript levels show transient up-regulation during development in zebra fish embryos (Makky et al., 2007). Therefore the developmental regulation of the TbTors is quite unusual.

### 5.2.3 RNA interference of TbTor1

The effect of TbTor1 depletion was assessed through knockdown of the gene expression using RNAi to determine whether absence of the protein would result in G1 arrest as is seen when TORC1 is disrupted in yeast cells (Barbet et al., 1996, Koltin et al., 1991, Kunz et al., 1993). Primers were designed using 'RNAit' (http://trypanofan.path.cam.ac.uk/software/RNAit.html) (Redmond et al., 2003) to amplify a unique section of the gene as described in Section 4.6, with the sequence ligated into the RNAi vector p2T7TaBlue (http://trypanofan.path.cam.ac.uk/trypanofan/vector/). Importantly, this region was carefully examined to ensure it would generate no cross-reaction with TbTor2. One clone was generated that showed inducible knock down of the transcript, and this showed severe growth and morphological defects (Figure 5.13).

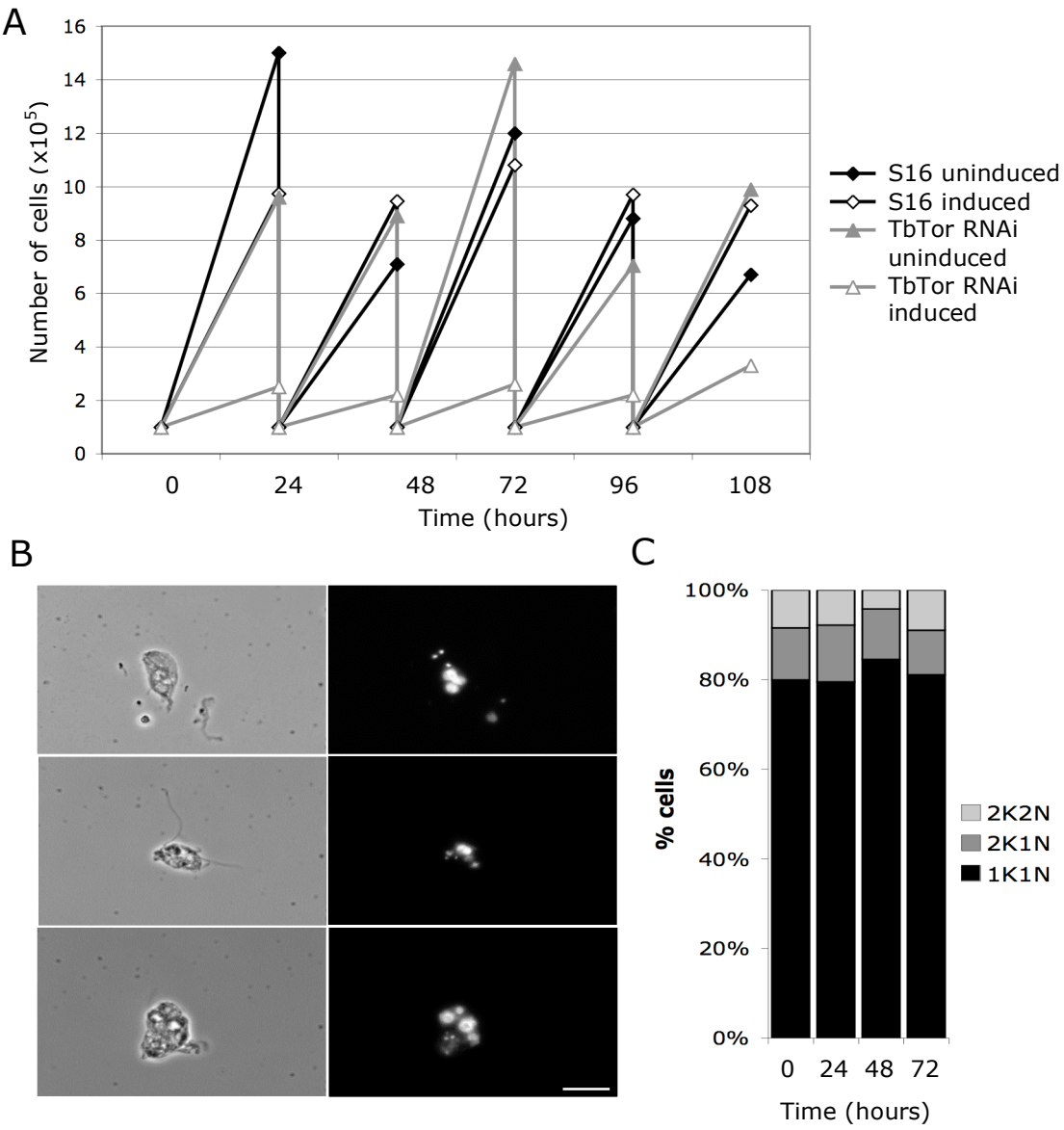



**Figure 5.13. Growth kinetics and morphological phenotype of bloodstream form cells depleted of TbTor1**. (A) The cells induced to perform RNAi against TbTor1 show an inhibition of growth even within 24 hours. The S16 parental cell lines were included as a control. (B) Images of cells 72 hours after induction of TbTor1 RNAi, showing phase contrast images and cells stained with DAPI to visualise the DNA. After this length of induction of TbTor1 RNAi, the majority of cells were dead, and those that remained intact demonstrated a morphological phenotype consisting of multiple kinetoplasts, nuclei, flagella and enlarged vacuoles. Scale bar represents 15μM. (C) Cell cycle configuration of TbTor1 depleted cells showed no G1 specific arrest. Cells with abnormal configurations were not scored, these consisting of approximately 10% of all cells.

Depletion of TbTor1 in bloodstream form cells resulted in a severe growth phenotype, with cells showing a huge reduction in cell number when compared with parental or uninduced cell lines (Figure 5.13A). In cells depleted for TbTor1, a number of cells with multiple flagella and aberrant numbers of nuclei and kinetoplasts were observed (Figure 5.13B). The cells also showed enlarged vacuoles with the cells 'rounding up'. However, since my observation of these morphological phenotypes, a publication from another group showed that depletion of TbTor2 resulted in cells that were enlarged, whilst knockdown of TbTor1 resulted in cell cycle arrest in G1, generating cells that were slightly smaller than wildtype (Barquilla et al., 2008). However Figure 5.13C shows that no accumulation of cells with a 1K1N configuration is observed upon depletion of TbTor1 up to 72 hours after induction of RNAi. For the purposes of kinetoplast and nucleus scoring, only cells with a normal configuration were counted to avoid skewing the results through the inclusion of aggregated dying cells. However, after 48 hours, approximately 10% of the population would generally be considered to be in an abnormal kinetoplast-nucleus configuration.

In order to investigate the differences between my data and those of Barquilla et al (2008), the same sequence used by Barquilla et al (2008) to induce RNAi was used to generate another RNAi cell line. Thus, the product amplified with the primers from that paper was inserted into p2T7TaBlue and transformed into single marker bloodstream form cells. One clone with inducible knock down of the transcript was generated using this sequence. Growth kinetic analysis of this TbTor1 RNAi cell lines showed a similar effect to that seen in the original TbTor1 depleted cell line (Figure 5.14). Furthermore, the cellular morphology generated upon RNAi induction was the same using either RNAi construct.

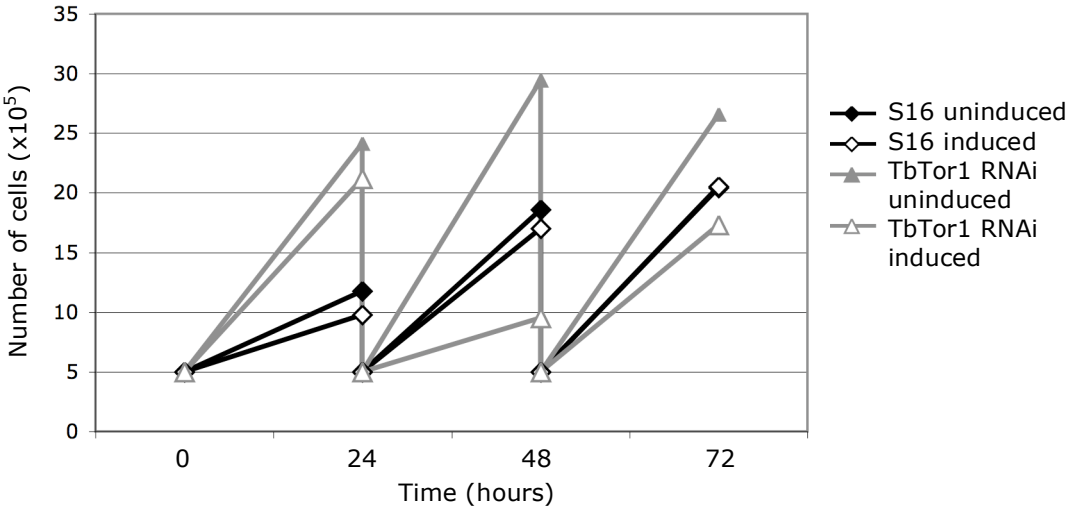



Figure 5.14. Growth kinetics of bloodstream form cells depleted of TbTor1 using the RNAi construct designed by Barquilla et al (2008). The growth kinetics of the clone generated to perform inducible RNAi using this construct do not have such an immediate growth defect, with RNAi-induced cells showing similar growth kinetics to the uninduced cell line after 24 hours. However, by 48 hours the depleted cells demonstrated a much reduced growth rate. In this experiment the parental S16 cells appear to grow much more slowly than the transfected cell lines, but this is because these are an independent stock from the S16 line used earlier (Figure 5.13).

Hence, depletion of TbTor1 in bloodstream form cells, using either of the RNAi constructs, caused both growth and morphological phenotypes. These phenotypic observations match closely those seen for TbTor2 knock down cell lines reported by Barquilla et al (2008), and were reproducible between the different constructs using different sections of the gene. This is in clear contrast to the TbTor1 knock-down cell lines described by Barquilla et al (2008) using the same construct, which showed only a minor morphological phenotype of a slightly reduced cell volume and an accumulation in G1 (Barquilla et al., 2008).

### 5.2.3.1 Confirmation of TbTor1 depletion in RNAi cell lines

In order to confirm that the phenotypes observed in both RNAi cell lines were due to the depletion of TbTor1, a TbTor1 antibody (kindly supplied by Miguel Navarro) was used to evaluate protein levels in the knock down cell lines.

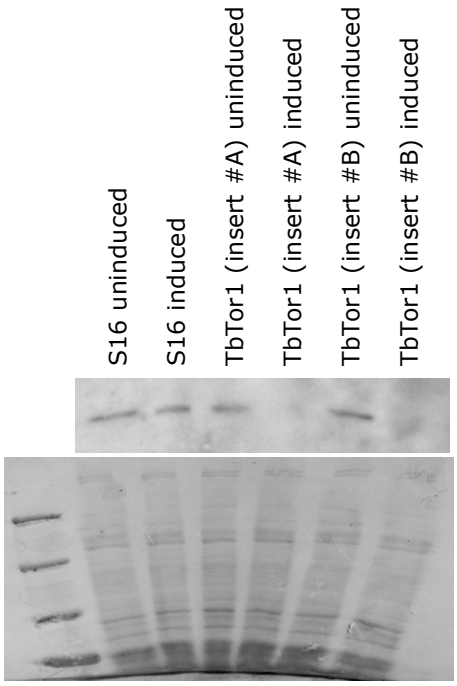



**Figure 5.15. Western blot of TbTor1 depleted bloodstream form cell lines 48 hours after induction of RNAi**. Both TbTor1 RNAi lines show depletion of TbTor1 protein after 48 hours of RNAi. Protein harvested from $3x10^6$ cells was loaded into each lane and the amido black-stained membrane was used to confirm equal loading of samples.

The Western blot in Figure 5.15 shows that both the induced TbTor1 cell lines were efficiently performing RNAi, as the protein had disappeared after 48 hours. This confirmed that the phenotypes seen in both RNAi cell lines were the result of depletion of TbTor1. It is concerning that Barquilla et al (2008) have conflicting phenotypes, particularly as the Western blots shown in their paper also confirm the depletion of TbTor1 and the same sequence was used to direct RNAi. The extent of depletion between the cell lines generated by Barquilla has not been compared against those generated here, and thus the depletion efficiency could possibly account for the difference in phenotype.

### 5.2.4 Rapamycin

#### 5.2.4.1 Treatment of *T. brucei* with rapamycin

As human Tor1 is present in both the TORC1 and TORC2, it was possible that the dramatic phenotypes seen from the TbTor1 depletion were the result of disruption of

two pathways; one of which would be involved in cell cycle control and the other in cytoskeletal regulation. Rapamycin is a selective inhibitor of the TORC1; presumably the conformational arrangement of Tor1 in TORC1 allows access to the rapamycin binding domain, whilst this is inaccessible when it is present in TORC2 (Loewith et al., 2002). Therefore, it was anticipated that rapamycin treatment would result in specific inhibition of just one pathway, most likely the TORC1 pathway involved in cell growth. However, mapping alignments of the FRB domains of TbTor1 with those from other organisms showed the absence of one of the highly conserved serine residues (Figure 5.16), thus it was not certain whether trypanosomes would be sensitive to rapamycin. This conserved serine residue is even present in the FRB domain of the Tor homologue in *C. elegans* which is insensitive to rapamycin, therefore the absence of this residue in *T. brucei* was predicted to indicate rapamycin insensitivity. Interestingly, the Tor2 protein in *T. brucei* and its homologue in *L. major* have an alanine residue in place of the serine, which has been shown in mutational studies to still allow rapamycin binding (Chen et al., 1995, Zheng et al., 1995).

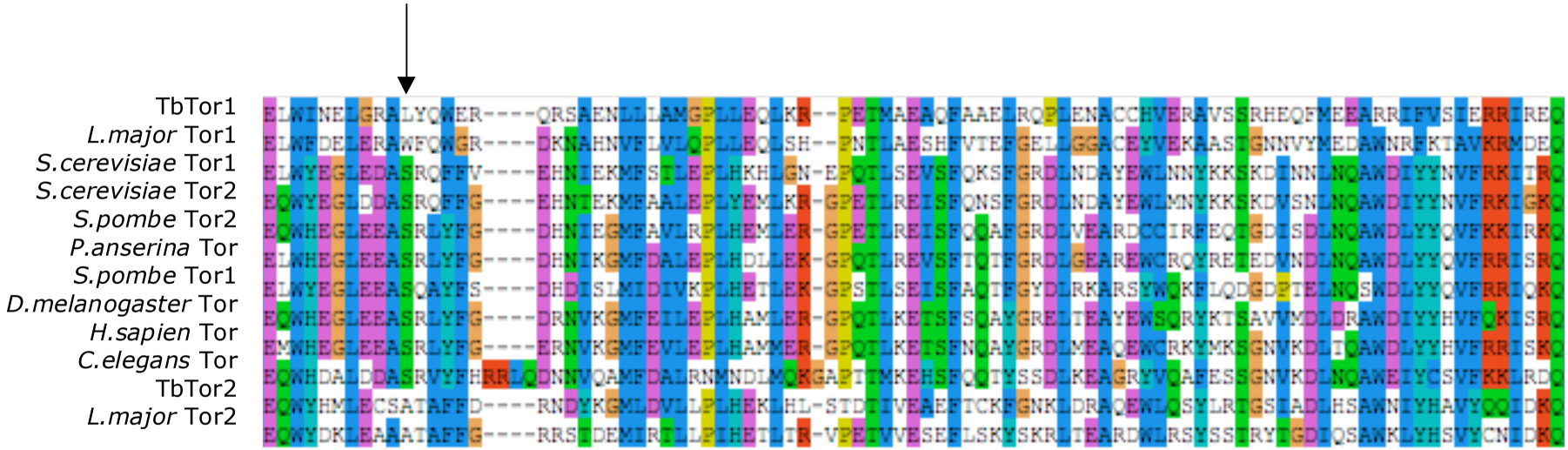



**Figure 5.16. Clustalx alignments of rapamycin binding domains**. The FRB domains of Tor1 and Tor2 in yeast, *Leishmania major* and *Trypanosoma brucei*, and the single Tor proteins of human, *Drosophila* and *Podspora anserina* show high levels of homology, ranging from 19-66% similarity. The conserved serine residue that is required for rapamycin binding (arrow) is present for all organisms except for *L. major* and *T. brucei*, although the Tor2 proteins in both these organisms have an alanine at this position, shown to be the only other residue that allows rapamycin binding.

Preliminary analysis of the effect of rapamycin on trypanosomes had shown no effect using a maximum dose of 400nM, although the rapamycin in this case had been dissolved in DMSO. However, the Barquilla et al (2008) study showed a dramatic growth and morphological phenotype using low concentrations of rapamycin (IC50 = 152nM), thus the experiment was revisited. For this experiment, a rapamycin titration was carried out on bloodstream form *T. brucei* using the same conditions as the Barquilla et al (2008) study: ie the drug was dissolved in ethanol and a range of concentrations from 10nM to 10µM was used. Finally, the assay was allowed to continue for 72 hours

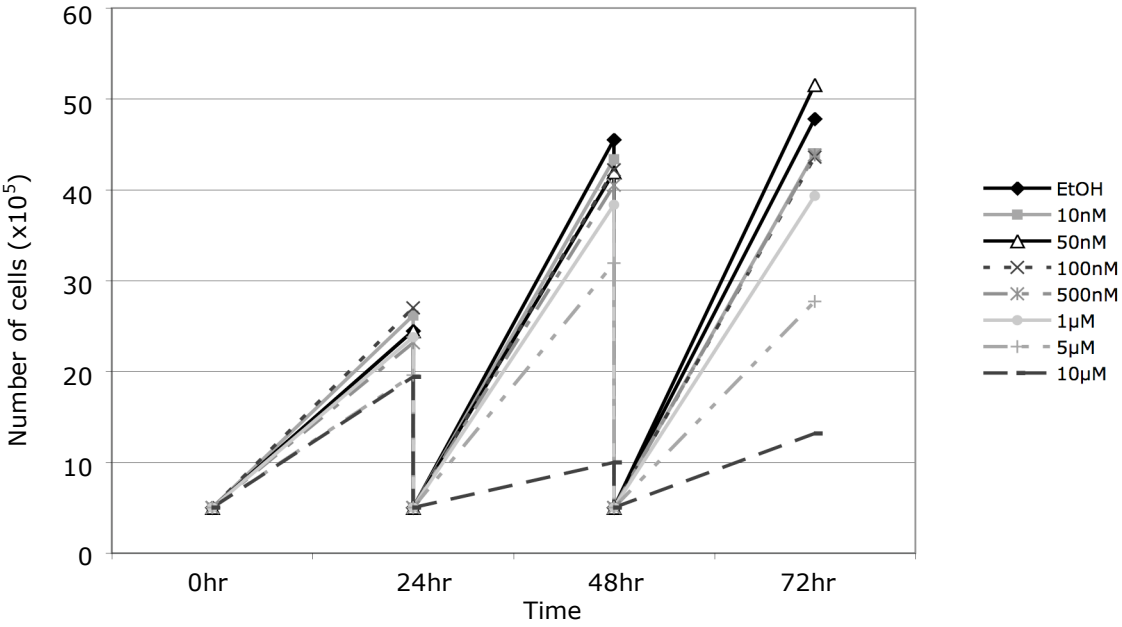



**Figure 5.17. Growth kinetics of bloodstream form cells treated with rapamycin**. A slight growth inhibition of the cells was seen after 24 hours for the highest concentration of rapamycin. By 48 hours the growth inhibition from 10µM rapamycin was more pronounced and the cells treated with 5µM rapamycin were also showing inhibited growth. By 72 hours the growth inhibition at the 5µM dose was very much more apparent and there were slight differences from the lower doses in comparison with the drug solvent alone. At the 72 hour time point the cells treated with 50nM rapamycin appeared to have a higher cell count than the control.

Figure 5.17 shows that only the highest doses of rapamycin (≥5µM) showed any significant growth inhibitory effect of bloodstream form cells. From these

experiments very little difference in cell number was seen below a rapamycin concentration of 5μM, in contrast to the data from Barquilla et al (2008), where significant inhibition was seen at concentrations of >100nM. It cannot even be certain that such a high concentration of drug is physiologically relevant, as at such high doses non-specific effects are possible. Interestingly, an increase in cell growth was seen in the cultures treated with 50nM rapamycin, which was a reproducible effect seen in every assay, this not becoming evident until 72 hours of treatment. The use of low doses of rapamycin has been seen to cause lifespan extension in yeasts and mammals (Powers et al., 2006, Harrison et al., 2009). In yeast, rapamycin concentrations between 300pg/ml to 1ng/ml have been shown to increase lifespan via a transcription factor responsible for the specific regulation of stress-induced genes, Msn2 (Powers et al., 2006). Msn2 is under the control of Tor and, in normal logarithmic growth, this protein is located in the cytoplasm, but upon depletion of Tor, Msn2 shuttles into the nucleus and induces transcription of a cohort of genes involved in stress response (Powers et al., 2006). Using concentrations of rapamycin below 300pg/ml had no effect on the lifespan of yeast cells (Powers et al., 2006), which is similar to these results seen in *T. brucei*, where the 10nM dose did not appear to cause an increase the cell numbers. Hence, this positive effect of low doses of rapamycin in *T. brucei* appears to be effective over a narrow concentration, similar to the effects on lifespan in yeast.

To calculate the dose at which the rapamycin caused a growth reduction of 50% in bloodstream form cells in comparison to the drug vehicle alone, the cumulative cell numbers from the growth kinetic experiments (Figure 5.17) were calculated after 72 hours of drug treatment and a non-linear regression curve fitted.

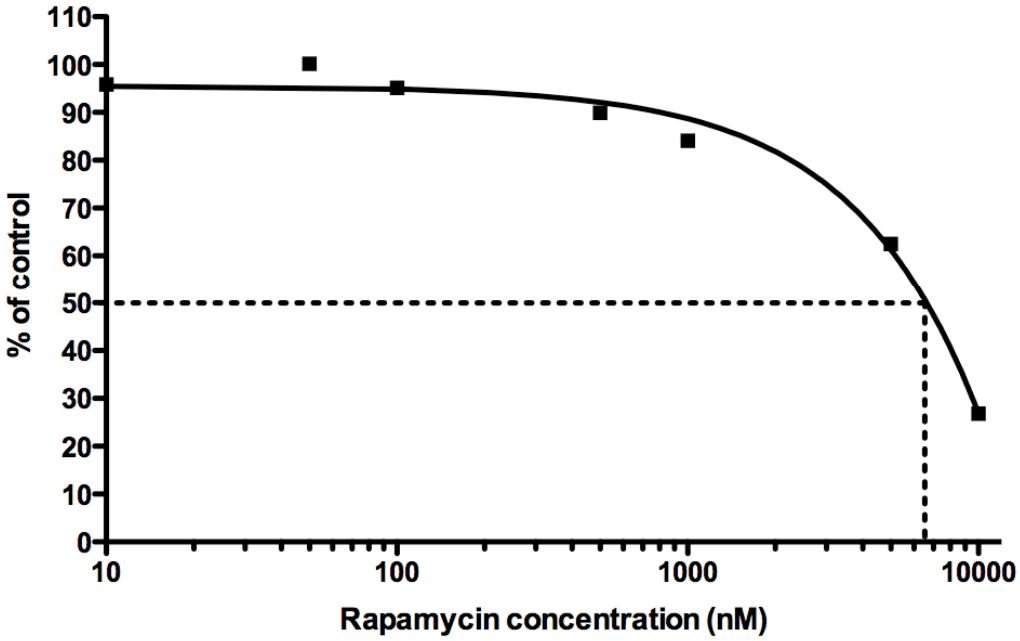



**Figure 5.18. Dose-response curve of rapamycin on bloodstream form cells**. The cumulative number of cells after 72 hours of growth was plotted as a percentage of the ethanol control. The dose of rapamycin causing a 50% reduction in growth was determined as ~6.5μM.

The dose required to cause a 50% inhibition of growth in comparison to control cells was estimated from the dose-response curve (Figure 5.18) and was approximately 6.5μM, over 40 fold higher than that calculated by Barquilla et al (2008).

### 5.2.4.2 Assessment of rapamycin activity

The data presented here suggested that the sensitivity of *T. brucei* to rapamycin was much lower than that discovered by Barquilla et al (2008). However, in order to ensure that the lack of rapamycin sensitivity shown by the cells in my assay was not due to low rapamycin activity, the drug sensitivity titration assay was repeated using different batches of the drug. This always showed reproducible sensitivity at levels higher than reported by Barquilla et al (2008). Furthermore, to validate the rapamycin sensitivity, its activity was tested against yeast cells. Although convincing studies titrating the sensitivity of *S. cerevisiae* cells in response to rapamycin are not available, one study has shown that cells treated with 200nM rapamycin undergo condensation of the nuclear DNA within six hours,

demonstrating that this concentration of the drug was sufficient to show a physiological effect on the cells (Tsang et al., 2007).

Therefore, overnight cultures of *S. cerevisiae* BY4741 were diluted to an optical density of 0.005 (measured using an Eppendorf BioPhotometer at wavelength 600nm (OD600)), into pre-warmed YPD medium and incubated at 30˚C until the OD600 reached 0.01, such that the cells were considered to be in logarithmic growth. Then 3ml of the culture was placed into glass vials and rapamycin added to the cells to the following concentrations: untreated, ethanol only (drug vehicle), 10nM, 50nM, 100nM, 500nM, 1µM and 10µM. Importantly, these drug dilutions were aliquots of the same rapamycin batch used against trypanosomes in the assays described earlier. These cultures were incubated at 30˚C with shaking for eight hours, and OD600 recorded every two hours. In addition, to allow a longer time course experiment and to assess the kinetics of the drug action, 100µl samples of the drugged cultures were also loaded in triplicate into 96 well plates and their OD595 was recorded on a BioTek® Elx808 plate reader every hour for 12 hours.

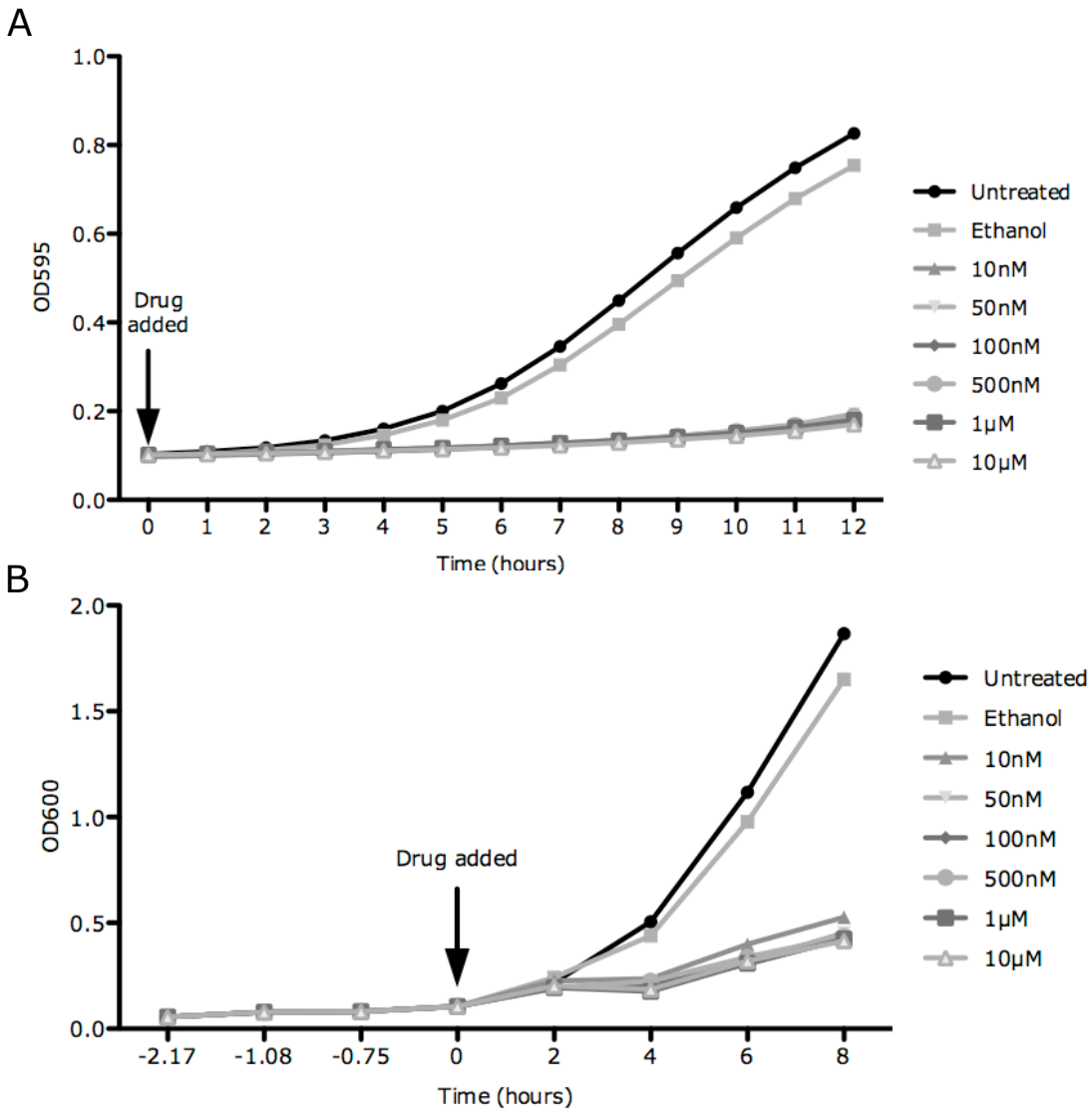



**Figure 5.19. Growth kinetics of rapamycin treated *S. cerevisiae*.** The growth kinetics of *S. cerevisiae* liquid cultures treated with rapamycin were assessed through (A) analysis of 3ml cells grown in a shaking incubator with OD600 readings taken every two hours for eight hours. Overnight cultures of cells were diluted to an OD600 of 0.005 and then monitored regularly until the OD600 reached 0.01 at which point the cultures were split into different vials and the drug added (arrow). (B) Automated readings from a plate reader at OD595 taken hourly for 12 hours of the same cells grown in 96 well plates. These cultures were taken at the time 0 point of Figure 5.19A, at the point at which the rapamycin was added. Both methods show a minor growth perturbation when ethanol alone is added to the cultures, but a severe growth effect is seen with the addition of rapamycin, even at the lowest dose of 10nM.

Figure 5.19A shows that rapamycin was active against liquid cultures of *S. cerevisiae* even at the lowest dose of 10nM. Figure 5.19B shows a longer time course experiment, although in this case the density of the cultures cannot be considered fully quantitative as the cells were grown in sealed wells with 100µl culture per well, such that the cells would not have had enough oxygen to grow optimally. Nonetheless, this figure is included as a qualitative indicator to allow comparisons between the treated and untreated cultures. Figure 5.19A gives a more accurate measure of growth of the yeast cells in rapamycin containing media and shows that even 10nM of the drug causes cells to arrest in growth after four hours of treatment.

The growth of yeast was also analysed on solid medium containing rapamycin. Thus, logarithmically growing liquid cultures of yeast were spotted onto YPD agar plates containing rapamycin. Overnight cultures diluted to OD600 0.005 and allowed to reach OD600 0.01 were then serially diluted into sterile water and 10µl of each dilution was spotted onto plates, which were incubated at 30˚C for 48 hours.

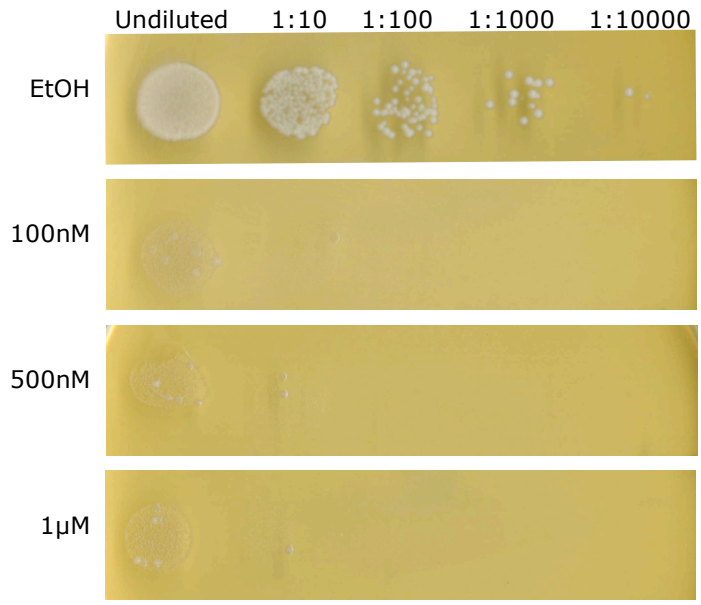



**Figure 5.20. *S. cerevisiae* shows high sensitivity to rapamycin**. Logarithmically growing cultures were subjected to 10-fold serial dilutions and spotted onto the YPD plates with either the drug vehicle (ethanol) or the indicated concentrations of rapamycin. The cells were viable and form robust colonies on the ethanol plates but only extremely small cells can be seen in the undiluted spots for all other doses of rapamycin suggesting that the sensitivity of *S. cerevisiae* cells is well below 100nM. This result confirms that the rapamycin is active.

The growth assay on the solid agar medium confirms the results seen in Figure 5.20, with 100nM rapamycin sufficient to severely inhibit cell growth. Thus, both figures demonstrate that the rapamycin is active against yeast at very low concentrations, hence discounting the possibility that the lack of an effect of rapamycin on *T. brucei* was due to inactivity of the drug.

### 5.2.5 Summary

At least two homologues of Tor are present in the *T. brucei* genome. Depletion of TbTor1 does not result in G1 arrest as is seen in some other organisms and in the Barquilla et al (2008) study, neither do the depleted cells have a smaller volume that wildtype cells. Instead, the phenotype seen in the TbTor1 depleted cell line was very similar to that seen for the TbTor2 RNAi cell line created by Barquilla et al (2008), regardless of which region of the gene the RNAi effect was directed against. Also in contrast to the Barquilla et al (2008) study, bloodstream form cells did not appear to

show sensitivity to rapamycin within physiologically relevant concentrations. The IC50 calculated by Barquilla et al (2008) was 152nM, whilst the results presented here show that rapamycin only starts to inhibit growth over 1μM, with the dose required for a 50% inhibition of cell growth approximately 6.5μM; over a 40 fold increase. The dose calculated here is much more similar to the EC50 of rapamycin in *Leishmania major* of 4.9μM (Madeira da Silva et al., 2009), which seems logical as the *Leishmania* ssp Tor homologues show high sequence identity to the TbTors. Due to the rapid lethality of TbTor1 depletion in bloodstream form cells and their failure to accumulate in G1, it was not possible to determine if this protein is required during differentiation.

The dramatic phenotype seen in bloodstream form cells depleted of TbTor1 was surprising considering that the mRNA expression levels were relatively low in this lifecycle stage. Depletion of TbTor1 in procyclic form cells might be expected to show a more severe phenotype; however, another report shows a reduction in cell growth only initiating 72 hours after induction of RNA interference, associated with the appearance of cells with aberrant numbers of nuclei and kinetoplasts (Monnerat et al., 2009). Thus, the phenotype seen in procyclic form cells depleted of TbTor shows no similarity to that seen by Barquilla et al (2008) when RNAi was conducted in bloodstream form cells, but was much more similar to the phenotype observed in the results shown here. However, it was surprising that the growth phenotype was not seen until 72 hours after induction of RNAi considering that this gene was so highly expressed in this life cycle stage as detected by Northern blot (Figure 5.12). One explanation may be that the protein is more stable in procyclic form cells and thus the delay in growth phenotype extends from the increased time for turnover of the residual protein. Alternatively, the efficiency of transcript depletion may have been less in the procyclic form study.

## *5.3 Cdh1*

### 5.3.1 Identification of TbCdh1

In yeast, the APC/C interacting protein Cdh1 is responsible for G0 regulation, in addition to functions in the maintenance of G1 and the correct progression of S phase (Sigl et al., 2009, Brandeis and Hunt, 1996, Kitamura et al., 1998). Thus, the T. brucei genome was searched for homologues, using the human Cdh1 protein as a template for a BLASTP query. Tb927.8.6500 was identified as showing the highest levels of similarity, however this protein has previously been annotated as the T. brucei Cdc20 homologue (Kumar and Wang, 2005). In order to determine whether Tb927.8.6500 was more similar to Cdh1 or Cdc20, the *Homo sapiens* Cdh1 or Cdc20 were both aligned against Tb927.8.6500, which was found to show 33% identity to Cdc20 and 34% identity to Cdh1. As the alignments did not identify definitively which protein Tb927.8.6500 was more similar to, a phylogenetic tree was generated using the Cdc20 and Cdh1 orthologues from a number of eukaryotes alongside Tb927.8.6500 to see in which branch it was situated. All the Cdh1 and Cdc20 protein sequences from *Homo sapiens*, *Sus scrofa, Mus musculus, Gallus gallus, Xenopus, Drosophila melanogaster, Xenopus*, *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* were downloaded from GeneDB or NCBI and aligned on Clustalx version 2.0.12 (Figure 5.22). Pairwise alignments were created using parameters allowing a gap opening penalty of 10 and a gap extension penalty of 0.1 to prevent the presence of too many gaps and thus reduce the possibility of false matches (Hall, 2001). The multiple alignment parameters were set using the gap opening penalty of 15 and gap extension penalty of 0.3 and the delay divergent sequence was set to 25%. These are the parameters recommended by in the ClustalX tutorial guide (Hall, 2001). The alignments were uploaded into PAUP* and a phylogram was generated using the Neighbour Joining distance method.

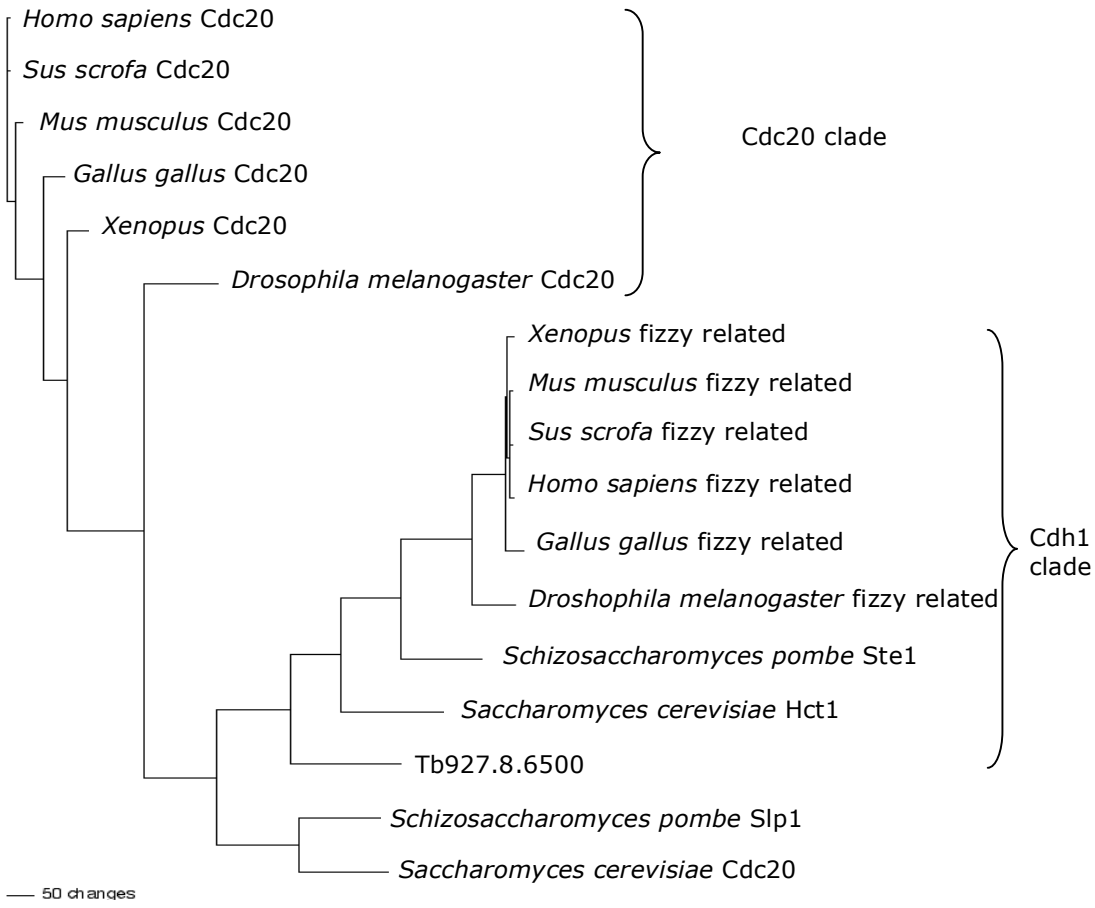
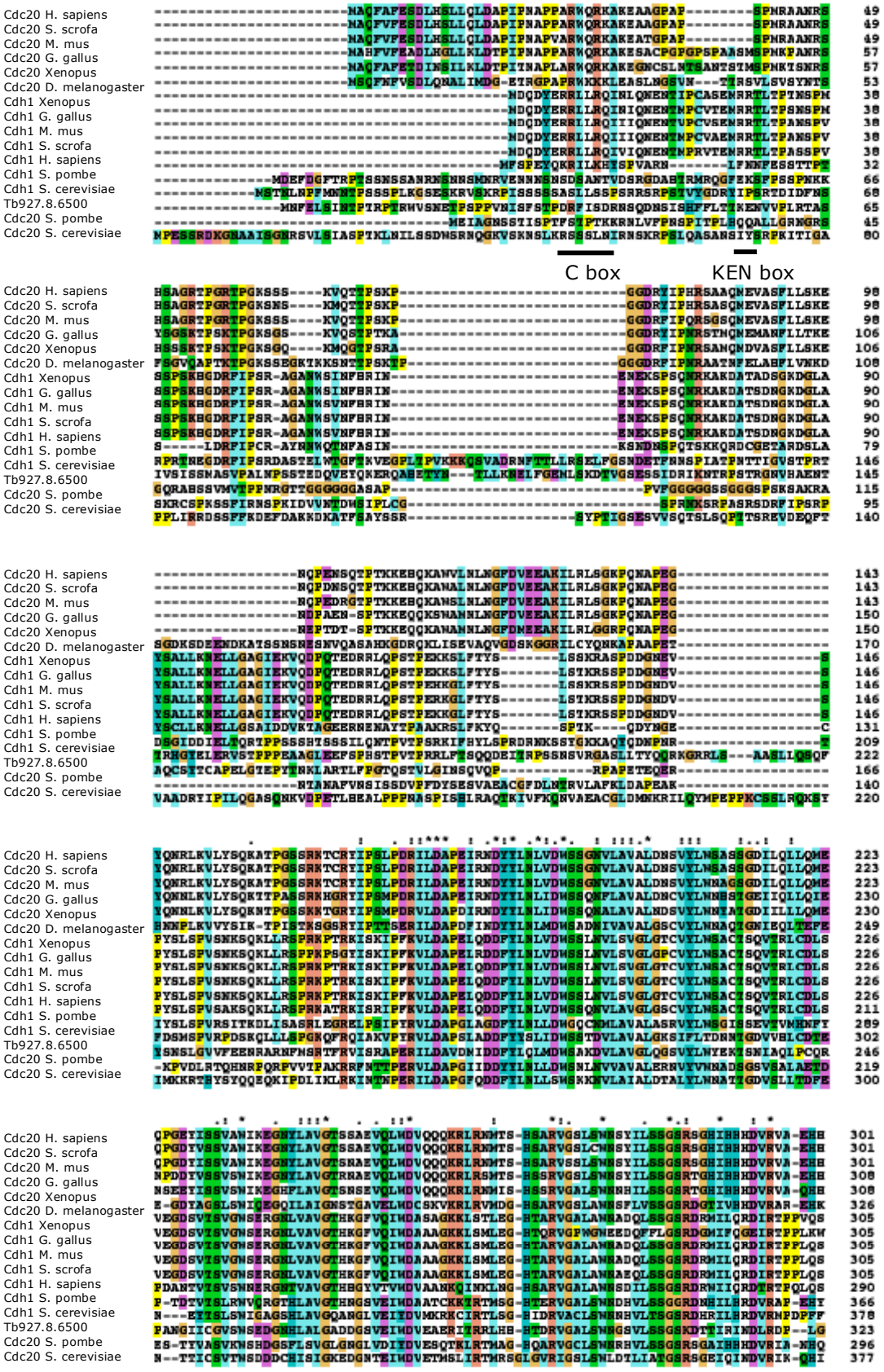



**Figure 5.21. Tree of Cdh1 and Cdc20 homologues**. Unrooted phylogram prepared in PAUP* of the Cdh1 and Cdc20 homologues from chicken, human, frog, pig, *Drosophila*, budding yeast and fission yeast and the putative Cdh1 homologue from *T. brucei*. The Cdc20 orthologues from yeast do not fit well with any of the other sequences, although all sequences from all other organisms lie distinctly in either the Cdc20 clade or Cdh1 clade, with Tb927.8.6500 falling into the Cdh1 clade. However, this sequence is also the most similar of all organisms to the Cdc20 orthologues in yeast, so it is difficult to make a firm declaration of which clade the *T. brucei* protein should be a member. The scale bar represents 50 substitutions in the amino acid sequence.

The phylogram in Figure 5.21 shows two distinct groups comprising the Cdh1 and Cdc20 proteins, with the Cdc20 homologues in *S. cerevisiae* and *S. pombe* forming a separate group within the Cdh1 clade, nearest to the Cdh1 homologues of yeast. Tb927.8.6500 falls into the Cdh1 clade, however it is still highly related to the two yeast Cdc20 homologues. This is similar to a tree compiled by Sigrist and Lehner

(1997) containing the sequences of Cdh1 and Cdc20 from budding yeast, frog, pig, human and *Drosophila*, where both the Cdh1 and Cdc20 homologues from *S. cerevisiae* were present on the same branch. Similarly, a tree showing the relationships between the Cdc20 and Cdh1 proteins in these organisms, in addition to an extensive number of plant species, also showed that *S. pombe* Cdc20 was highly divergent when compared to all other sequences (Tarayre et al., 2004).


Cdc20 H. sapiens
Cdc20 S. scrofa
Cdc20 M. mus
Cdc20 G. gallus
Cdc20 Xenopus
Cdc20 D. melanogaster
Cdh1 Xenopus
Cdh1 G. gallus
Cdh1 M. mus
Cdh1 S. scrofa
Cdh1 H. sapiens
Cdh1 S. pombe
Cdh1 S. cerevisiae
Tb927.8.6500
Cdc20 S. pombe
Cdc20 S. cerevisiae
C box
KEN box

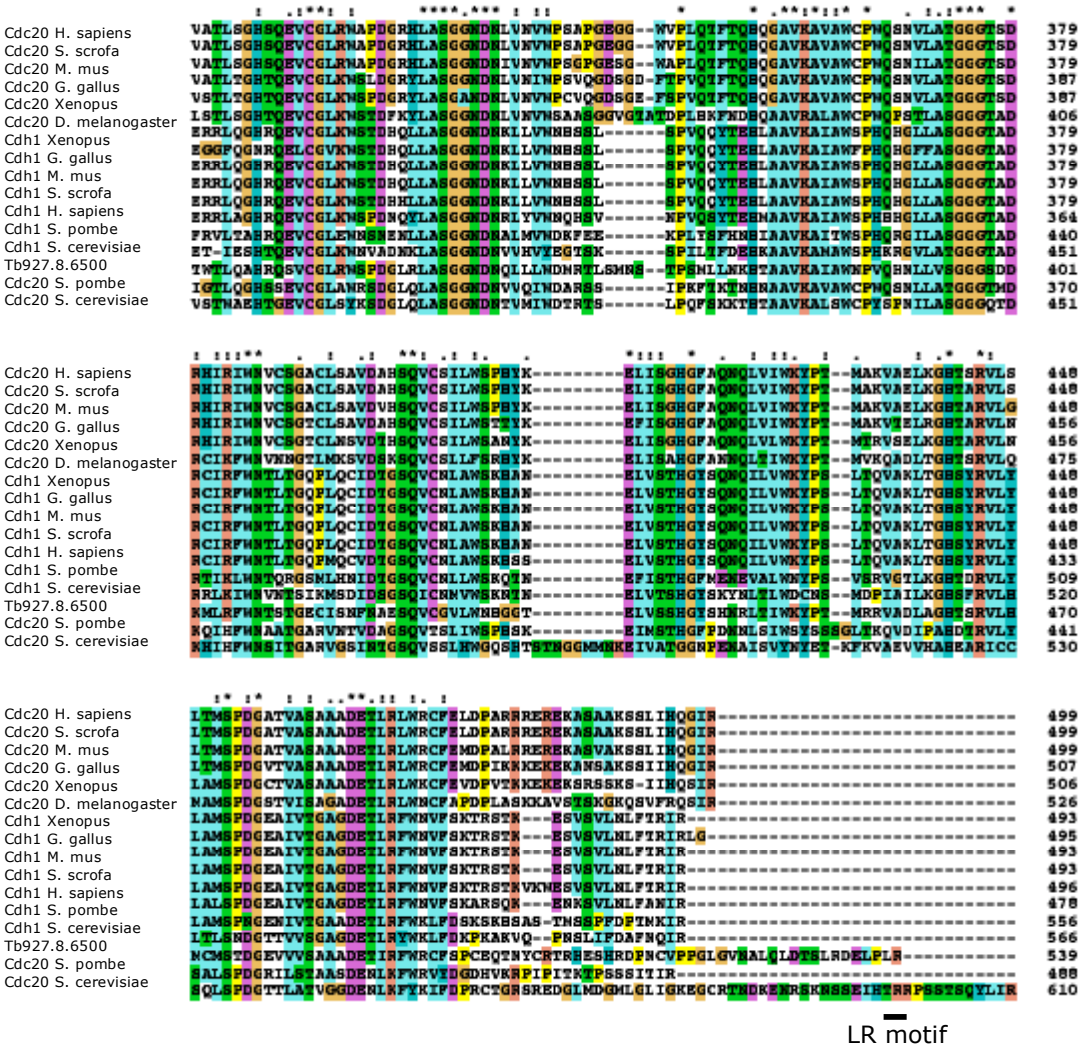



**Figure 5.22. Clustalx alignment of Tb927.8.6500 with Cdh1 and Cdc20 orthologues from other organisms**. The full length sequences of Cdh1 and Cdc20 proteins from yeast, *Drosophila*, *Xenopus*, chicken, mouse, human and pig aligned against Tb927.8.6500. Motifs present in the Tb927.8.6500 protein sequence are labelled.

The sequence of Tb927.8.6500 was analysed in detail to determine the conserved domains (Figure 5.22). The Cdc20 and Cdh1 proteins have a number of shared and specific domains. The shared domains include the seven WD40 domains, which are present in Tb927.8.6500 and a C box at the N terminal domain which is required by both proteins for binding the APC/C (Schwab et al., 2001). Tb927.8.6500 has a fairly well conserved C box with the sequence DRFISDR, which correlates well with the consensus sequence of DR(F/Y)IPxR generated from the C boxes found in higher eukaryotes. The position of the C box in Tb927.8.6500 is closer to the N terminus

than is seen in other proteins and is much more apparent than a C box in *S. pombe*. Cdh1 also has two destruction (D) boxes, with consensus sequence RxxLxxxxN upstream of the C box (Listovsky et al., 2004). Nothing similar to the consensus sequence was found at this position in Tb927.8.6500 due to the close proximity of the C box to the N terminus, also no D boxes were found elsewhere in the sequence. In contrast, whilst the Cdc20 protein in *S. cerevisiae* also has two D boxes (Prinz et al., 1998), these are absent in mammalian Cdc20, which instead has a KEN box that is also recognised by Cdh1 for degradation (Pfleger and Kirschner, 2000). A KEN motif was found in Tb927.8.6500 at position 55-57. Thus even though the KEN motif in *T. brucei* is much closer to the N terminus than in mammalian Cdc20, in both organisms the KEN boxes are both positioned a similar distance away from the C box: 14 and 13 residues respectively, although in *T. brucei* the KEN box is downstream of the C box, whereas in mammals it is upstream. Mammalian Cdc20 also contains a CRY box, which also allows degradation through recognition by Cdh1 (Reis et al., 2006), however this motif is not present in Cdc20 homologues in yeast, and a CRY box was not identified in Tb927.8.6500. Tb927.8.6500 differs from both human Cdh1 and Cdc20 in that the final two residues do not conform to the IR motif that is so conserved in other organisms, and is necessary in this organism for binding to the APC/C. Rather, an LR motif, which is also seen in the *S. pombe* Cdc20, was present. Another motif was discovered in plants, ~50 residues downstream of the C box that was present only in the Cdh1 homologues and not Cdc20 proteins. The consensus sequence for this motif was generated from Cdh1 proteins in eukaryotes as YxxLL(K/R)x(E/A)L(F/L)G, which was designated the CSM (Cdh1 specific motif) (Tarayre et al., 2004). The CSM motif also appears to be absent in Tb927.8.6500.

| Domain | Mammalian Cdh1 | *S. cerevisiae* Hct1 | Mammalian Cdc20 | Tb927.8.6500 |
|---|---|---|---|---|
| **WD40** | 7 | 7 | 7 | 7 |
| **D box** | 2 | 0 | 0 | 0 |
| **KEN box** | 0 | 0 | 1 | 1 |
| **C box** | Yes | Yes | Yes | Yes |
| **IR motif** | Yes | Yes | Yes | No – LR motif |
| **CSM** | Yes | Yes | No | No |

**Table 5.1. Conserved domains in Cdh1 or Cdc20**. Many of the domains associated with Cdh1 in other organisms were not conserved in Tb927.8.6500, however one motif specific to the Cdc20 proteins was found in Tb927.8.6500.

Despite the presence of the KEN box and lack of a CSM motif, we shall cautiously consider Tb927.8.6500 to be TbCdh1. This is based on the positioning of this protein with Cdh1 in the phylogenetic tree, the fact that the protein shows slightly more similarity to human Cdh1 than Cdc20 and the absence of a CRY motif.

### 5.3.2 Expression profile of TbCdh1

The mRNA levels of Cdh1 and Cdc20 in higher eukaryotes fluctuate throughout the cell cycle with both Cdh1 and Cdc20 transcripts highest in G2 cells (Fang et al., 1998). Likewise, the protein levels of Cdc20 parallel the mRNA profiles, however the profile of Cdh1 protein expression is controversial. Some studies suggest that Cdh1 protein is constant throughout the cells cycle (Fang et al., 1998), and others report that the protein level also fluctuates with the mRNA level (Kramer et al., 2000). The reasons behind the variation between different groups is suggested to be due to the quality of antibody and the method of cell synchronisation used, which may be having direct or indirect effects on Cdh1 protein expression (Kramer et al., 2000). It would not be implausible that the Cdh1 protein would be much more stable throughout the cell cycle than Cdc20, as Cdh1 is largely regulated at the post-translational level through phosphorylation, whereas the Cdc20 protein is degraded by the APC/$C^{Cdh1}$ (Kramer et al., 1998). In order to determine the mRNA levels in *T. brucei*, the expression profile of TbCdh1 as determined by the microarray was investigated (see Chapter 4), and validated using Northern blots using the same mRNA samples as used for the microarray (Figure 5.22B). In addition, the expression profile was determined in another pleomorphic *T. brucei brucei* strain, *T. brucei brucei* EATRO 2340, contrasting the levels in slender, stumpy and procyclic form cells (Figure 5.22C).

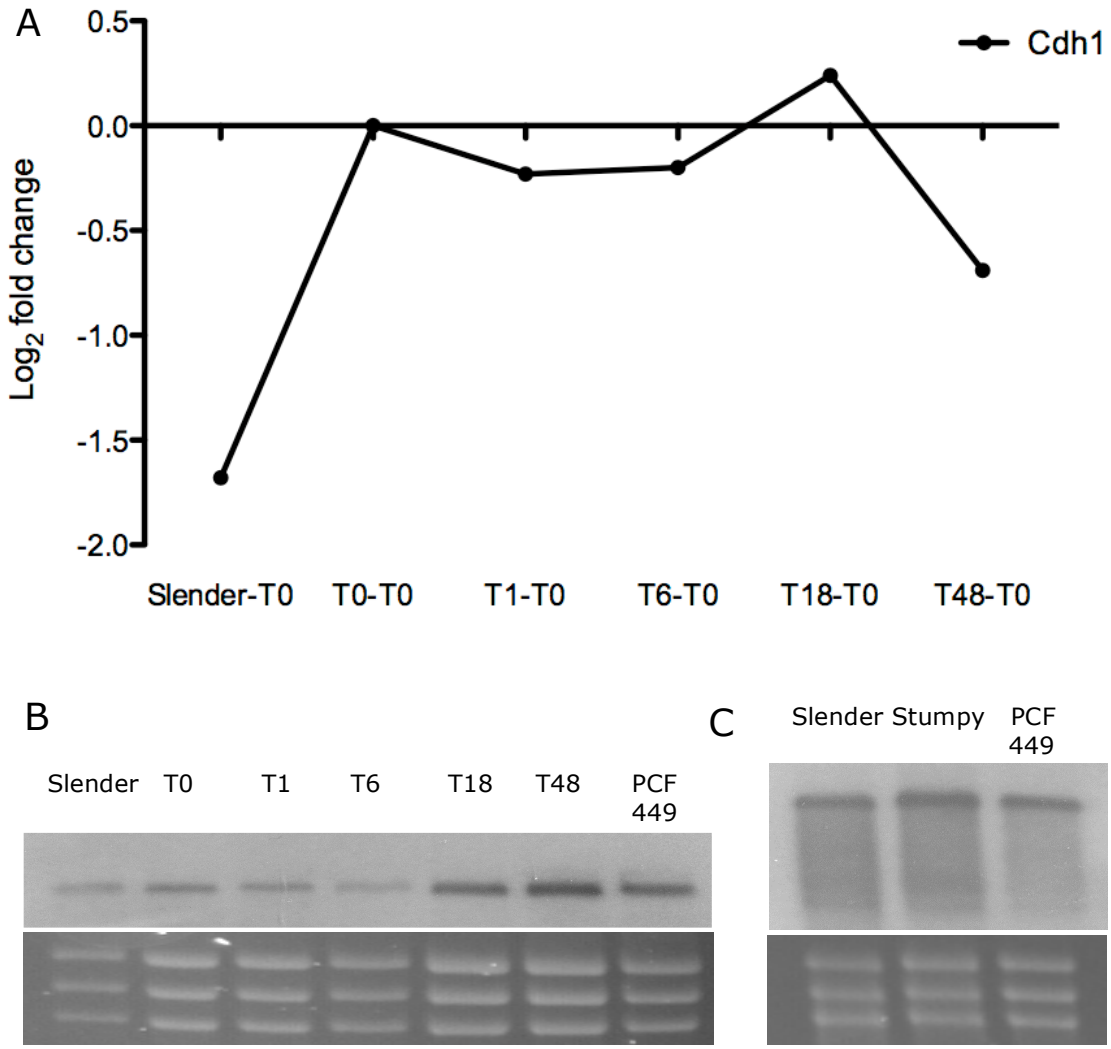



**Figure 5.23. Expression profile of TbCdh1**. (A) Microarray profile of TbCdh1, showing transient elevation during differentiation with a marked increase from slender to stumpy, peaking at 18 hours before dropping at 48 hours. (B) Northern analysis using RNA generated for the microarray (from a *T. b. brucei* AnTat1.1 cell line), shows elevation during differentiation, with the peak at 48 hours before decreasing in established procyclic forms. (C) Northern analysis of slender and stumpy form *T. b.brucei* EATRO 2340 cells with established procyclic form cells shows the greatest expression is in stumpy forms with lower levels in slender form cells.

The TbCdh1 transcript showed enrichment during differentiation on the microarray (Figure 5.23A), although this result was not quite so apparent looking at results from Northern membranes probed with a riboprobe against TbCdh1 (Figure 5.23B).

However, in that case, as can be seen from the rRNA levels in the ethidium bromide stained gel, the gel was unevenly loaded. The expression profile of a second pleomorphic cell line, *T. b .brucei* EATRO 2340, showed a slight increase in transcript level in stumpy form cells when compared against slender and procyclic form cells (Figure 5.23C). It is possible the increase in TbCdh1 in differentiating cells is because they are in G0 stage, as Cdh1 from other species is known to maintain the G0 state (Inbal et al., 1999). Cdh1 also plays an important role in G1, and it would be expected that this transcript remained high in cells entering the first cell division following release from G0 arrest, but would be lower in asynchronously growing populations, which would fit well with the profile revealed by the microarray (Figure 5.23A). However both Cdh1 and Cdc20 show peak expression levels at the end of G2 phase (Fang et al., 1998), which would fit with the maximal expression in the differentiation time course at 18 hours, as previous studies examining the cell cycle following differentiation shows that the G2/M phase would occur between 13-14 hours into the time course (Matthews and Gull, 1994). These kinetics are in keeping with the cell cycle scoring analysis carried out, showing that 2K1N cells appeared after 6 hours through differentiation (Figure 3.7). Therefore, the 18 hour time point would be expected to have a higher proportion of cells towards to end of G2, than any other time point sampled.

### 5.3.3 Depletion of TbCdh1

TbCdh1 was depleted in bloodstream form monomorphic cells using the p2T7TAblue vector to express double stranded RNA of a unique sequence of the gene and thus induce the RNAi system. Two clones were generated, both showing a growth phenotype with the induction of RNAi, although the non-induced cells also showed poorer growth than parental cells lines suggesting that dsRNA was being expressed at a low level (Figure 5.24A). After 72 hours following induction of RNAi, the cells returned to normal growth kinetics – likely reflecting a selection for a minority of cells deficient in RNAi and hence able to outgrow the Cdh1 depleted population. TbCdh1 depleted cells also showed gross morphological phenotypes with cells with multiple flagella, consistent with defects in cytokinesis (Figure 5.24B).

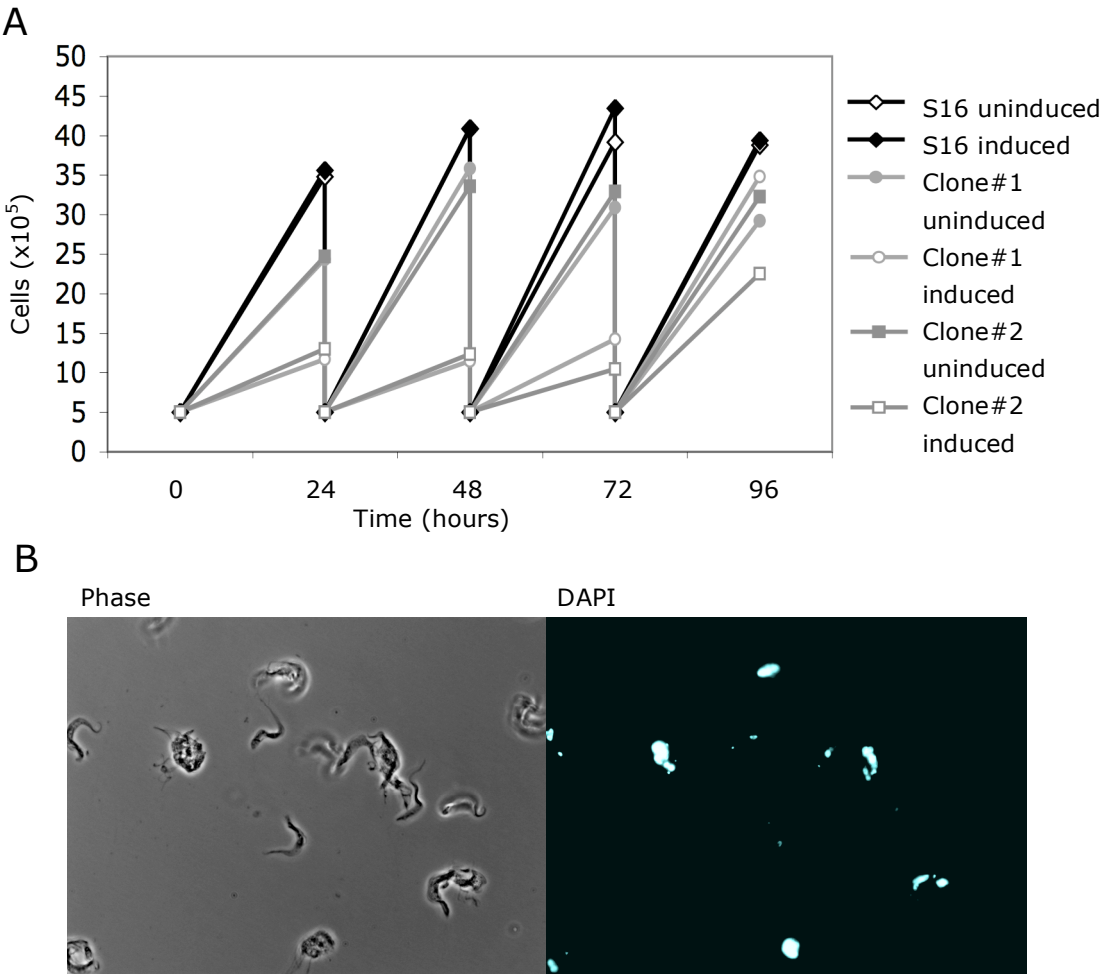



**Figure 5.24. Growth kinetics and morphological phenotype of TbCdh1 depleted bloodstream form cell lines**. (A) Two clones with inducible knock down of TbCdh1 were generated and both showed impaired growth within the first 24 hours after induction of RNAi. After 72 hours these knock down cell lines appeared to recover, as RNAi deficient cells out competed the TbCdh1 depleted cells. The growth rate of the uninduced cell lines was slower than the parental single marker cells probably due to leaky expression from the RNAi vector. (B) Bloodstream form cell lines after 48 hours of TbCdh1 depletion. Many cells showed aberrant numbers of kinetoplasts and nuclei.

Depletion of TbCdh1 resulted in severe growth effects with cells failing to grow within the first 24 hours after the induction of TbCdh1 RNAi (Figure 5.24A). The reason for the poor growth was high levels of cell death, with cells showing aberrant cell division to form cells with multiple nuclei, kinetoplasts and flagella (Figure 5.23B) leading to death. In other organisms Cdh1 is non-essential, and depletion of

the protein leads to only minor cell cycle abnormalities, mostly altering the length of cell cycle phases ((Brandeis and Hunt, 1996, Floyd et al., 2008, Visintin et al., 1997, Yamaguchi et al., 2000). Therefore, this result showing a clear essential role of TbCdh1 in *T. brucei* may mean this protein is more similar to Cdc20, which is an essential protein in other organisms, or it may be because *T. brucei* has only one APC/C co-activator and thus the role of this protein is much more wide-ranging that Cdh1 in other organisms. Certainly, however, depletion of the protein did not cause a specific cell cycle arrest, thereby limiting its usefulness in analysis of G0/G1 control in trypanosomes.

### 5.3.4 Antibody

In order to see whether the protein levels were also altered during differentiation and to find out the localisation of the protein in the cells, an antibody was raised against TbCdh1. The N terminus was targeted, as the C terminus forms complex tertiary structures and thus might be harder to access. A peptide sequence corresponding to amino acids 9-24 was synthesised and injected into two rabbits on four occasions, with a final bleed taken after 28 days by Eurogentec. Although time constraints limited the analysis that could be carried out with this reagent, this will allow in depth analysis of TbCdh1 in the future.

### 5.3.5 Summary

It is difficult to definitively describe Tb927.8.6500 as either the homologue of Cdh1 or Cdc20, although considering that it appears to be the sole activator homologue of APC/C, it is perhaps inadvisable to name it either, as the role it plays will almost certainly differ from that of either protein in other species. Certainly this protein is essential in the bloodstream form parasite as depletion results in rapid lethality, rendering it difficult to assist in studies examining a potential role in G0 and differentiation.

It might be that TbCdh1 has different degrons to those in other homologues, especially considering the sequences involved in targeting the mammalian APC/C coactivators for destruction differ significantly from those in yeast. The seven APC/C subunit homologues found in *T. brucei* show sequence identity ranging from

14-25% with those in *S. cerevisiae*, with TbCdc27, the subunit that binds to the co-activators, showing only 20% identity to the yeast homologue (Kumar and Wang, 2005). Thus, it is highly likely that the mechanism of binding might differ. The expression profile of TbCdh1 can also not aid in distinguishing between Cdc20 and Cdh1, as high Cdc20 expression is observed in yeast arrested in G1with alpha mating factor, even though normal G1 the levels are quite low (Prinz et al., 1998), which may explain the high levels of mRNA seen in stumpy form cells.

### *5.4 Chapter summary*

This chapter describes the approaches taken to examine G0 in *T. brucei*. Unfortunately, neither of the proteins identified as homologues to G0 regulator proteins in other species demonstrated a specific role in G0 control in trypanosomes. However, the divergent functions of these proteins from those seen in other organisms will prove interesting for further study from an evolutionary perspective. The method described for isolation of cell cycle specific populations was unable to create synchronised populations, and was not suitable for examination of G0 as it was not possible to distinguish between G0 and G1 populations. Nevertheless, this technique will allow future investigation of cell cycle regulated proteins, and is particularly useful for examination of G1 enriched populations as no other method for isolation of this cell cycle stage currently exists.

# 6 – Discussion

The aims of this PhD were to study the underlying genetic events occurring throughout differentiation from the stumpy form into the procyclic form and to analyse the role of G0 in the initiation of this process. This was addressed in several ways using a whole-genome microarray, individual characterisation of two pathways involved in G0 in higher eukaryotes and through development of a method for isolating cells in specific cell cycle stages.

## *6.1 Microarray validation*

The aim of the microarray was to analyse the expression profiles of genes through the differentiation process. The main objective was to generate data that was amenable to statistical interrogation, and to that end five biological replicates were examined for each time point. A consensus does not exist regarding which statistical methods are most appropriate for microarray validation, however it has been shown that applying different statistical analyses to microarray data can be performed and still give the same results (Minning et al., 2009). Thus our choice of statistics for this analysis were not as important as the level of significance applied, and the statistical analysis of the data was deliberately stringent to reduce the possibility of false positive results. Furthermore, extensive validation was also carried out to confirm the microarray data using a combination of methods including qPCR, Northern Blots and mass spectrometry, in addition to verification against the published literature (Chapter 4). The combination of the stringency of the statistic threshold applied and the extent of the validation of the microarray data should overcome many of the shortcomings of the microarray slides. Comparison of the data retrieved from the microarray against that published was generally very consistent; significantly the published literature historically has focussed on the most highly expressed transcripts, which might show better sensitivity on the microarray.

Puzzlingly, a poor rate of correlation was found between the qPCR and the microarray, even though the same RNA samples were used for both assays (Figure

4.12). In microarray studies, results gained from qPCR and microarrays often show disparate results, leading some researchers to reject qPCR as a method for microarray validation (Allison et al., 2006); however, superior methods of validation have not been suggested in its stead as no other method can recreate the extent of normalisation used in a microarray. Other limitations in our ability to find correlation between the data generated from the microarray and qPCR and those generated from Northern Blots could be in part due to using oligo-dT primers to make the cDNA for the qPCRs and microarray, such that only mRNA, and not rRNA, was transcribed into cDNA. The presence of rRNA in the Northern blots makes the results generated from this method of validation unrepresentational of the profiles based on poly A+ mRNA used for both the qPCR and microarrays. One option to make Northern Blots more representative of microarrays would be to use the same probe as was used for the microarray, so that any cross-reactivity on the microarray would be duplicated on the Northern Blot (Kane et al., 2000). It would also be possible to normalise the Northern data using a second probe against a transcript showing unchanging expression across the time points. However the current, qualitative, system of processing Northern blots in our laboratory using film does not allow accurate quantification of signal, therefore a quantitative method for Northern blotting is required before such a control strategy is revealing. Alternatively, a spiked transcript could be used as a control, thus overcoming the problem of finding a suitable internal control showing a non-fluctuating profile, which had proved to be an obstacle for the qPCR experiments (Kane et al., 2000).

Due to the different criteria for designing probes for microarrays, Northern Blots and qPCR, in our analyses, the probe used for the validation very rarely matched the probe that was used in the microarray. This may have influenced the validation in multiple ways as, should these mRNAs be degraded over the course of differentiation, using a probe that bound closer to either the 3' or 5' end of the sequence would alter the probability of detecting it. Such a factor might be an important issue for genes showing transient differentially regulated expression profiles, which was the case for a number of the transcripts used for validation of the

microarray, thus these transiently changing transcripts might be more prone to errors and were therefore not the best candidates for verification. One factor that accounts for low correlation between qPCR and microarray results is a low p value of the differential regulation of a transcript, thus the best candidates for validation are those transcripts showing differentiation expression with a p value of $p<0.0001$ (Morey et al., 2006). A better method for qPCR validation of these data, therefore, would have been to look at genes differentially expressed at a p value of $p<0.0001$ and only examined the comparisons with this high probability of significant differential expression. Our method was instead to find genes where at least one comparison had a p value of $p<0.05$ and examine all time points for this transcript, even those where the difference in expression level did not reach this level of significance.

### 6.1.1 Evaluation of performance of arrays

The microarray slides used for the lifecycle analysis were the publicly available JCVI Trypanosoma brucei v3 arrays. Most of the 70mer probes on the array are unique and will only bind to a single target, however some probes will cross-hybridise with genes that are extremely homologous to the extent that a 100% unique target would be impossible to design.

Tile arrays are sometimes viewed as more accurate since several different probes are designed for each gene and can be distributed across the slide, circumventing problems associated with varying locational hybridisation. Shortly after the publication of the microarray data reported here, another paper was published describing the gene expression profiles of slender and stumpy form cells using a custom-made NimbleGen chip (Jensen et al., 2009). This microarray comprised eight different 60mer targets per gene, present in triplicate, however the genes annotated as 'hypothetical, unlikely' on the database were not included on this array. There was a good amount of correlation between the two sets of data, and the transcripts showing divergent results between the arrays could have arisen due to strain specific differences or may have been the result of the different procedures for isolation of the parasites. As the NimbleGen arrays are designed to such stringent criteria, no validation of the results was carried out for these experiments. The two

data sets will be useful for researchers to get information about a gene of interest in two different strains and at slightly different time points; our microarray examines the continuous process of differentiation from stumpy form cells whilst the Jensen study examines separate points in differentiation. The different studies also included slightly different cohorts of transcripts; the 'hypothetical unlikely' genes were excluded from the NimbleGen array, yet are represented in our microarray, therefore our microarray provides information on an additional subset of the genome which might prove to be of great interest and biological relevance. Similarly, the Jensen et al (2009) arrays included a number of genes encoded by the mitochondrial genome, thus the separate datasets are able to complement each other with the different data they provide.

One point highlighted from the two studies is the extent to which differences in the experimental design between our microarray and that of Jensen et al (2009) could potentially skew the expression profiles seen, through the manner of parasite isolation. The parasites used on the NimbleGen array were used for RNA extraction immediately after purification from blood, in contrast to the cells used in our arrays, which were allowed one hour to recover from the purification process, during which time a number of cellular processes could have become activated.

A third microarray study published at the same time as both these studies also looked at the process of differentiation using the JCVI Trypanosoma brucei v3 arrays, but took the slightly different approach of studying the differentiation of cells in vitro (Queiroz et al., 2009). Even if of pleomorphic origin, slender form cells in culture do not take on a clear stumpy morphology and thus the authors cultured the slender cells to a high density before inducing differentiation with cis aconitate (Queiroz et al., 2009). This study will be of use to other researchers who do not have access to true stumpy form cells, and then the results can be cross referenced against the two in vivo differentiated cells to compare the results in a more biologically relevant context. Moreover this in vitro study seemed to suggest that slender cells progress

through a stumpy transition before differentiation to procyclic forms (Queiroz et al., 2009). This supports the notion that stumpy form cells are essential for transmission.

### 6.1.2 Future genome-wide analyses

Some of the most interesting and important features of differentiation may have been lost from our analysis because we concentrated on the differentiation associated with the stumpy to procyclic transition. Indeed, there are likely to be many genes involved in the differentiation from slender cells, through intermediate stage and then into stumpy forms. For example, genes that we have detected to 'spike' in stumpy and immediately decline in abundance, may have been already up-regulated in intermediate or early stumpy form. Further, we may have entirely missed genes that were only transiently required for these early events. It would benefit the examination of differentiation to be able to study the events leading up to cells becoming stumpy, but as yet the process of slender cells transforming into stumpy forms is ill defined. Additionally, for microarray studies a homogenous population is beneficial (Ness, 2007), whereas slender cells are unlikely to proceed into stumpy form cells in a synchronous manner. The problems associated with these heterogeneous populations were highlighted in the slender samples used for our microarrays, which showed far more variability than the uniform stumpy populations. It would also be interesting to investigate whether the genes we have identified as showing up-regulation during differentiation, are also required for the differentiation events that occur during the metacyclic to slender transition, as these cells are also released from a G0 arrest (Shapiro et al., 1984), or whether these genes are unique for preparing the parasite for tsetse fly survival.

Improvements in technology are occurring rapidly and even tile arrays are being surpassed by the newest deep sequencing methods. One such methodology emerging as a successor of microarray technology is RNA-Seq, in which cDNA is fragmented and sequenced from one or both ends, resulting in a high coverage quantitative measurement of transcript abundance (Nagalakshmi et al., 2008). This technology does not suffer the problem of cross hybridisation as microarrays do, thus very little background is present. It is also a much more sensitive technology, so that it can be

used to identify transcripts present in low abundance such as those which are rapidly degraded (Nagalakshmi et al., 2008). Experiments on yeast looking at expression profiles throughout meiosis have shown that RNA-Seq has advantages over microarrays in terms of identifying novel sequences including mitochondrial transcripts (Wilhelm et al., 2008), something that would be useful for *T. brucei* research as the microarrays publicly available do not contain mitochondrial gene probes. Microarray results can be biased based on where the probes target; if they lie towards the 3' end, then transcripts that are degraded might be appear to be present at higher levels than in reality. This problem is resolved using RNA-Seq studies, where it is possible to distinguish between full-length transcripts showing high expression and those in the process of being actively degraded, thereby defining the 'degradome' (Lister et al., 2009).

The use of long 70mer oligos in the microarrays is beneficial, as unique probes can be generated for transcripts that are similar; however it has been reported that only 15 consecutive bases are required for cross hybridisation of transcripts showing over 50% similarity when using 50mer oligo probes (Kane et al., 2000). RNA-Seq is capable of distinguishing between much more similar transcripts than microarrays, which would be helpful for *T. brucei* as many genes in tandem repeats have coding regions too similar to allow discrimination on the microarray, such as the PAD array, where the probes on our array against PAD4 also had cross-reactivity with PAD6 and PAD8. In general however, the extent of discrimination between highly similar transcripts on these 70mer arrays was good. The three PGK genes show up to 94% sequence identity, and the probe for PGKB had 29 consecutive residues of identity with PGKC, however these three transcripts all showed distinct profiles on the microarray, therefore the stringency of the hybridisation conditions used were suitable for differential binding. In the case of the PAD genes however, whilst the PAD4 probe showed 84% identity with PAD1 and 27 residues were consecutive, this transcript did not hybridise to the probe, yet the probe also showed 97% identity with PAD6 and 94% with PAD8 and cross reacted with both these transcripts.

In an opinion paper, Duncan (2004) suggests that a better use of microarrays in the study of gene expression in kinetoplastids would be in the comparisons of the mRNA associated with the polyribosome, which will distinguish the transcripts being actively translated and will be a more accurate representation of the genes required under certain conditions (Duncan, 2004). Such an approach has been attempted previously in *T. brucei*, but was unsuccessful due to poor recovery of pure polyribosome-associated mRNA (Brems et al., 2005). The concept has been demonstrated in *T. cruzi,* where increased association of one stage regulated mRNA with the polyribosome has explained why the protein and mRNA levels of this gene were incongruent in two life cycle stages, demonstrating that association of mRNAs with ribosomes is an important mechanism of regulation in kinetoplastids (Nardelli et al., 2007). Future experiments optimised to recover polyribosome associated mRNA in *T. brucei* will help to narrow down the genes important in different life cycle stages.

Studies of the importance of the probe design in signal output have shown that the annealing temperature of probes is far more important for signal intensity than the length of the probe, particularly as it appears that only the final 40 nucleotides of the probe are able to perform hybridisation; the sequence closest to the slide surface are shielded from the cDNA (Wei et al., 2008). This would suggest that a better method for probe design would be to design probes that have similar annealing temperatures, rather than similar lengths as was the criteria in the TIGR microarray. However, other more complex attributes also alter binding efficiency, such as C/G content, therefore all probes would need similar percentage of C/G to avoid bias being generated by the strength of hybridisation (Wei et al., 2008). Applying all these conditions when designing probes would be extremely time consuming and, in many cases, impossible. Hence, the approach used in the JVCI microarrays, and the NimbleGen tiling arrays, are likely to represent the most appropriate for the platform used, although other technologies do not suffer from these constraints.

Other studies that would strengthen the knowledge of the developmental effects in *T. brucei* would be the investigation of the proteome of differentiating parasites. Rather than a stand-alone data set, proteomic data would complement the transcriptome data, and would add another dimension of information about gene regulation. A proteomic study of *Leishmania infantum* cells at different stages of their life-cycle identified a number of proteins with differential expression, and analysis of the transcript levels of these proteins using a microarray showed that the transcript regulation approximately mirrored that of the proteins. However, the protein information not only showed the levels of expression but also provided data on the modifications of the proteins in the different life cycle stages and inherent bias from the over representation of highly expressed proteins was avoided through the use of prefractionated samples (McNicoll et al., 2006). Nonetheless, proteomic studies are usually of lower sensitivity than transcriptome analysis, rendering most useful for only relatively abundant proteins.

### 6.1.3 Oligo motif scoring

The identification of a conserved hexamer with a positional bias in the 3'UTR of a significant cohort of stumpy-enriched genes will provide a very interesting avenue for further research. The sequence of the motifs, TCTTAC and TTCTTA are not unusual as 3'UTR regulatory motifs, as RNA binding motifs are often AU rich (Haile and Papadopoulou, 2007), including those found in the COX subunits, PGKB and EP procyclin. RNA binding proteins in *T. brucei* have been shown to be involved in regulation of gene expression, with the recently characterised DRBD3 shown to bind to a conserved motif in transcripts that are up-regulated in procyclic form cells (Estevez, 2008). A genome-wide search for potentially conserved cis acting RNA elements between *T. brucei* and *L. braziliensis* was recently conducted, which identified 21 potential regulatory elements in the 3'UTR of functionally related genes (Mao et al., 2009). The hexamer from the stumpy-enriched genes was not identified in this study, but because the stumpy enriched genes do not share functional attributes this group would have been overlooked in the analysis. The Mao (2009) study highlights the variety of conserved sequences in non-coding regions of the genome, and indeed in PGKC, six of the conserved sequences were

discovered in the 3'UTR suggesting that more than one sequence might be regulating the same gene, as we have also found for genes with copies of both SM1 and SM2.

In order to understand the function of the identified conserved motifs, the 3'UTR, or regions of it, would conventionally be fused onto a reporter construct such as CAT or luciferase so that the role played by the 3'UTR in stability or translation could be examined. Another essential experiment would be to identify the proteins that recognise these motifs. Methods are available to isolate RNA binding proteins, such as the StretpoTag system, in which the motif of interest is fused onto a streptomycin-binding sequence, which is used to attach the construct onto a streptomycin-coated column (Bachler et al., 1999). RNA binding proteins are then isolated through separating cell extracts over the column and eluting those that bind to the motif through release of the streptomycin-binding sequence by washing with streptomycin (Bachler et al., 1999). The method has been optimised to allow binding of proteins with low affinities for their targets, and the motif of interest can range up to 600nt, allowing an entire 3'UTR to be studied if necessary (Dangerfield et al., 2006). Estevez successfully used a similar method to identify the TbDRBD3 RNA binding protein, whereby all proteins interacting with the various regions of the 3'UTR of PGKB, were identified (Estevez, 2008). However, in that case the selected protein was found to not regulate the target transcript, highlighting that target specificity remains a significant challenge.

One caveat of using a such a ligand selection approach to identifying RNA binding proteins is that it is only suitable for proteins recognising primary motifs, additionally this methodology is conducted in vitro so might not show true biological associations. Should the above method fail to identify the RNA binding protein recognising SM1 and SM2, an in vivo method could be considered. A method for identifying the RNA binding proteins binding to UTRs has been developed for use in live cells, through generating a peptide nucleic acid (PNA) oligonucleotide designed to bind to regions of the 3'UTR of interest and associate with the RNA binding proteins also present nearby (Zielinski et al., 2006, Zeng et al., 2006). The PNA

oligo contains a cell membrane-penetrating peptide to allow entry into the cells, and a photoactivatable amino acid adduct, which, when excited with UV, will induce cross linking of the surrounding proteins, thus linking the RNA and RNA binding protein with itself (Zeng et al., 2006). The PNA oligo and its binding partners are then isolated using a strepavidin magnetic bead fused with the sense sequence of the oligo target and purified, before the proteins are separated on an SDS-PAGE gel and identified using mass spectrometry (Zielinski et al., 2006). The problems of such an analysis include the expense of the reagents, particularly as oligos would preferably be designed for the analysis of both the SM motifs.

### 6.1.4 Applications of the microarray to the field

The microarray data generated here will provide researchers with a greater depth of knowledge of the expression profile of their gene of interest. It will also highlight genes for further study that show regulated expression at stages distinct from the conventional cultured bloodstream and procyclic forms. These genes which have been shown to be transiently up-regulated in stumpy forms hold potential as new stumpy markers in addition to PAD1 (Dean et al., 2009), and may provide information of transmission biology. Targeting such genes for drug therapy is also an attractive option, as many of the genes up-regulated in stumpy forms are unique to kinetoplasts and thus would not incur any cross-reactivity with the host. If such genes could be targeted to force slender form cells to differentiate into stumpy form cells, the parasite infection would arrest with a single VSG coat and could thus be cleared entirely by the immune response. Should an inhibitory RNA binding protein be found that was responsible for the stage specific degradation of transcripts required only in the stumpy form, this would be an excellent possible target. The next experiments to conduct based on the microarray data will be further analysis of the stumpy enriched genes and their biological functions, and investigation of the regulation potential of the SM1 and SM2 motifs.

One of the transcripts found to be strongly up-regulated in stumpy form cells, a predicted chloride channel, was knocked down using RNAi in monomorphic cells, but no growth phenotype or differentiation effect was observed (Figure 4.14). This

may be due to redundancy of the proteins, as this chloride channel has several related genes in the genome. However it may be that the transcript depletion methods used thus far are not suitable for dissecting the true function of this gene in pleomorphic cells. For example, depletion of ZFK or TbMAP5 in pleomorphic cells resulted in rapid formation of stumpy form cells, but no effect was detected in monomorphic cells (Domenicali Pfister et al., 2006, Vassella et al., 2001). Thus, follow up experiments should include conducting the RNAi experiment in pleomorphic cells, which has proved successful for a number of genes (Dean et al., 2009) (Szoor et al, submitted). However, RNAi experiments cannot be conducted directly on stumpy form cells, as they are quiescent, therefore the RNAi would need to be induced in slender forms, which would then be left to transform into stumpy forms. Another option will be the degron system that has been used successfully in *Leishmania* (Madeira da Silva et al., 2009), in which an unstable domain from FKBP12 is fused onto the protein of interest, and maintained in the presence of a rapamycin-based drug. Release from the drug then allows destruction of the unstable domain, thus effectively depleting the protein (Edwards and Wandless, 2007). Further characterisation of the genes showing significantly increased expression in stumpy form cells must also be conducted through over expression of the wildtype proteins, or mutant protein, allowing investigations of their function by a dominant negative approach.

At present there are no characterised markers of G0 in *T. brucei*, and only one recently discovered marker exists now for stumpy form cells (Dean et al., 2009). Previously the only markers to distinguish a stumpy form cell were morphology, mitochondrial activation and growth arrest. The scarcity of knowledge relating to gene expression in quiescent cells and stumpy form cells means that whether the transcripts identified as being significantly up-regulated in stumpy form cells are later found to be transcripts specific for G0 maintenance rather than lifecycle stage specific, they will still be invaluable markers. Markers for stumpy form cells will be potentially useful to dissect out differences between true stumpy form cells and stumpy-like cells that can be generated through treatment with drugs such as

troglitazone, the cysteine protease inhibitor Z-Phe-Ala-$CHN_2$ and hydrolysis products of cAMP analogues (Denninger et al., 2007, Scory et al., 2007, Laxman et al., 2006). With better markers, or markers that distinguish slender, intermediate, and stumpy cells, it would be possible to generate refined mathematical models to determine the kinetics of stumpy cell formation during a *T. brucei* infection (Fenn and Matthews, 2007, MacGregor and Matthews, 2008).

## *6.2 Synchronisation of* T. brucei

The cell cycle enrichment method described in Chapter 5 will be a good alternative to current synchronisation techniques used in *T. brucei* which arrest the cell cycle progression through treatment of cultures with hydroxyurea, followed by release into fresh media, thereby allowing cells to continue replicating DNA. Such 'batch-release' approaches for whole culture synchronisation should be used with caution, as such severe disruption of the cell cycle is likely to have complicated follow-on effects on the released cells. Studies on human cell lines synchronised through release from cell cycle block caused by drug treatment, have shown that the DNA content of the cells did not correlate with other features of cell cycle progression such as cell size and levels of cell cycle regulation proteins. This suggested that the cell cycle arrest merely represented a superficial arrest of DNA synthesis (Urbani et al., 1995). In consequence, an interpretation of the events occurring in such cells following release into fresh medium would be impossible, as the cells were in an abnormal cell cycle arrest.

### 6.2.1 'Selective' methods of synchronisation

Instead of using chemicals to trigger cell-cycle checkpoints or arrest proliferation, 'selective' methods to fractionate cells from the entire cell population on the basis of their cell cycle positions. Several alternative methods of selective synchronisation exist. In other organisms for example, cell sorting has been used to isolate populations in different cell cycle stages through fusing a fluorescent tag onto a known cell cycle regulated protein (Coquelle et al., 2006). Theoretically several proteins could be tagged with markers fluorescing at different wavelengths allowing proteins specific to different cell cycle stages to be selected simultaneously in order to isolate pure populations. At present very few cell cycle regulated proteins are

known in *T. brucei*, and those that are known generally show maximal expression in S phase cells (Archer et al., 2009). Such experiments would require that extensive genetic manipulation would be needed to generate the tagged cell lines, potentially complicating the results and adding additional transfections into an experiment. Moreover, one concern when using cell cycle regulated genes as markers is that these genes may themselves be aberrantly regulated when cells are perturbed by batch-release protocols (Cooper, 2004). In addition, the same cell cycle markers may not be useful in bloodstream form and procyclic form cells, as these different life cycle stages can show different regulation of cell cycle proteins (Hammarton et al., 2003, Hammarton et al., 2004, Kumar and Wang, 2005, Janzen et al., 2006, Li and Wang, 2006, Gourguechon and Wang, 2009). Instead of a fluorescent marker fused to the protein, if antibodies exist against a cell cycle regulated protein, then this could be used to sort cells. A candidate would be the histone variant H4K4, which is only present in its non-modified form in S phase immediately after it is produced, and antibodies have been developed that are able to distinguish this form from the acetylated form found throughout the remainder of the cell cycle (Siegel et al., 2008b). However, this approach can only be successful if the cells can be stained without requiring fixation so that viable populations can be recovered. This is unlikely to be feasible for cell cycle regulated proteins, which are usually intracellular.

The various selective methods of synchronisation described thus far have a range of benefits for analysis of cell cycle proteins. However a superior method still would be the 'baby machine' endorsed by Cooper used for the production of cells synchronised on the basis of both cell cycle and age (Helmstetter et al., 2003). This process allows the collection of the daughter cells from each cell division and is a more gentle process than FACS, which can stress cells due to shearing force and dilution in buffer solutions. It is yet to be determined whether this would be possible for *T. brucei*, which require motion for division (Ralston et al., 2006, Monnerat et al., 2009) and thus might be impeded if adhered to a filter.

Another selective method for isolating populations in a certain cell cycle stage is elutriation, where cells are forced through filters via centrifugation, with the size of cell recovered dependent on the speed of the rotor, such that increasing the speed recovers larger cells. Earlier methods for elutriation had the disadvantage that the process pushed cells into G0 arrest (Zickert et al., 1993), presumably through serum starvation, and thus the cells generated were not representative of cells going through a normal cell cycle progression. Additionally, cells recovered following elutriation have shown abnormal levels of cyclin mRNAs (Mikulits et al., 1997). Modification of the protocol has been shown to minimise the stresses on the cells, and these include returning cells into conditioned media following elutriation or for adherent cells, choosing the correct density of cells to allow rapid re-adherence to the culture vessel. Nonetheless, even using these optimised conditions all recovered populations showed a 30 minute lag period before resuming proliferation (Mikulits et al., 1997). In addition, conditions must be optimised for each cell type, and the length of the centrifugation is never less than two hours (Mikulits et al., 1997). Moreover, with these selective methods for isolating cell cycle specific populations, the sterility of all equipment is of paramount importance and contamination of the cultures is more of a possibility than with batch-release methods. Probably, the best method of selection varies for each cell type, although it should be noted that elutriation is optimised for mammalian and yeast cells, all of which have a circular morphology, whereas trypanosomes, with a more elongated structure, might prove more difficult to separate according to cell size. In addition, the process of elutriation subjects the cells to extreme stresses caused by the sustained centrifugation, whilst those experienced by the FACS sorted cells will be much lower, and the exposure to the laser has been shown to have no negative effect on cells (Siemann and Keng, 1986).

One of the concerns about the Vybrant® DyeCycle™ Violet was that it might show cell cycle specific toxicity as has been seen with Hoescht staining, which shows exclusive toxicity to cells in S phase (Siemann and Keng, 1986). This was not found to be the case with Vybrant® DyeCycle™ Violet staining of *T. brucei* bloodstream form cells (Chapter 5); however the dye did show a low level of toxicity to cells

irrespective of cell cycle position, which could be reduced through limiting the time the cells were exposed to the dye and keeping the cultures cooled to below 20˚C. Using these conditions it was possible to recover around half of the cells initially stained and the recovered populations showed good motility post sorting and would therefore be amenable to further experimentation. This synchronisation method was particularly suitable for isolating pure G1 populations, with a much poorer recovery of S phase cells; this sorting experiment was therefore limited to harvesting G0/G1 and G2/M populations only, in order to recover as pure a population as possible. When more accurate knowledge of the levels of cell cycle associated proteins in *T. brucei* are known, it would be a worthwhile exercise to examine their levels in Vybrant® DyeCycle™ Violet synchronised cells to assess the extent, if any, of cell-cycle perturbation during sorting.

### 6.2.2 Applications of synchronising *T. brucei* cells

Although little is known of cell cycle regulation in *T. brucei*, especially in bloodstream form cells, many of the mechanisms discovered differ significantly from those in higher eukaryotes. A method for selecting cells in specific cell cycle stages will allow the study of the levels of expression of genes in different cell stages and allow comparisons between those seen in bloodstream form and procyclic form. Synchronised populations could be used in microarray, or other whole-genome studies in the future, to determine the cell cycle regulated transcripts. It remains to be determined whether the Vybrant® DyeCycle™ Violet method of cell cycle specific enrichment will be suitable for use in sorting pleomorphic bloodstream form cells, which may prove more sensitive to the physical strains of the sorting process than the culture-adapted monomorphic cells used for these analyses.

## *6.3 Target of rapamycin*

The repertoire of protein kinases has been detailed for *T. brucei, T. cruzi* and *L. major* (the TriTryps) and has been shown to differ from that of higher eukaryotes (Parsons et al., 2005). Thus the unusually high number of Tor homologues in the *T. brucei* genome is not so unexpected; Barquilla et al identified four Tor homologues in the *T. brucei* genome (Barquilla et al., 2008). However, from the absence of a rapamycin binding domain from Tor-like 2, this final member of the Tor family is

more likely to represent a PIKK member rather than a Tor. There are no homologues in the genomes of the TriTryps or *Plasmodium* species of the normal eukaryotic tyrosine kinases, although certain atypical tyrosine kinases are present, and it is likely that kinases with activity against more than one residue are responsible for the tyrosine phosphorylation occurring in these kinetoplastids (Parsons et al., 2005) (Fenn and Matthews, 2007). Tor proteins show phosphorylational activity against both serine and threonine residues and the functions of the Tor-like proteins will be interesting to study as these appear to be unique to trypanosomes.

### 6.3.1 TbTor1

Tor has been shown to act in regulation of G1 arrest when in the TORC1 complex in a wide range of organisms. Depletion of TbTor1 using RNA interference caused a severe growth phenotype in bloodstream form cells, as cells depleted of TbTor1 showed significantly reduced cell growth compared with wildtype cells (Figure 5.13). These cells also showed dramatic morphological phenotypes with the appearance of cells with multiple nuclei, kinetoplasts and flagella suggesting that the cell cycle has been impaired in these cells. These results conflict with the results of Barquilla et al (2008), who detected these phenotypes in TbTor2 depleted cells, but found that TbTor1 depletion resulted in a slight decrease in cell volume and cell cycle arrest. In my analysis, the same growth and morphological phenotype also was observed in TbTor1 depleted cells in cell lines transfected using the same sequence to induce RNAi as was used in the Barquilla et al (2008) paper.

### 6.3.2 Rapamycin sensitivity in *T. brucei*

Bio-informatic analysis of the rapamycin binding domains of the TbTor proteins highlighted the lack of conserved residues in TbTor1, regarded as essential for rapamycin interaction in other organisms (Chen et al., 1995). The FRB domain of TbTor2 has an alanine at the site of the conserved serine residue, which has been shown to still allow rapamycin binding (Chen et al., 1995, Zheng et al., 1995). Nonetheless, initial treatments of bloodstream form cultures with rapamycin up to 400nM showed no effect on growth rate in contrast to what might be expected, since the drug triggers G1 arrest in all sensitive organisms. Following the publication of the rapamycin sensitivity of the same *T. brucei* cell line (Barquilla et al., 2008), the

experimental conditions used in the paper were used to reassess sensitivity in our laboratory. Whilst a growth inhibitory effect was observed, this was not until micromolar levels of the drug were used (Figure 5.18), in contrast to Barquilla et al (2008) where the dose required to cause a 50% inhibition of cell growth was calculated as 152nM (Barquilla et al., 2008). The possibility that the lack of sensitivity to rapamycin in our laboratory was due to low activity of the drug was ruled out, as a titration of the same aliquots of rapamycin were used against *S. cerevisiae* both in liquid culture and on agar plates containing the drug (Figures 5.19 and 5.20). Both experiments showed that yeast cells were highly sensitive to doses of rapamycin as low as 10nM, confirming that inactivity of the rapamycin in our experiments was not the reason for the discrepancies between the levels of sensitivity seen. The EC50 calculated from the experiments presented here was similar to that seen in *Leishmania major* (Madeira da Silva et al., 2009), which is consistent with the FRB domain of a Tor homologue in *L. major* showing strong similarity to TbTor2. It is not certain, however, whether the growth phenotypes observed with the high doses of rapamycin were due to a specific effect from inactivating a TORC complex or whether they were a non-specific effect resulting from using such high concentrations of the drug.

The low sensitivity of *T. brucei* to rapamycin does not make rapamycin a good candidate for drug therapy to cure human African trypanosomiasis as was suggested by Barquilla et al 2008. This drug is prescribed for the treatment of certain tumours as it prevents proliferation of the cancerous cells and also for transplant patients, since it prevents the proliferation of cells of the immune system and thus prevents immune rejection (Ballou and Lin, 2008). For transplant patients, the rapamycin concentration in the blood must be maintained between 12-20µg/ml (equivalent to a maximum of 21.9nM) (Committee, 2009), and our experiments show that the growth of half the parasites is inhibited at 6.5µM. Hence, the concentration of rapamycin necessary to clear a trypanosome infection would be sufficient to also inhibit the immune system. This would be an unfavourable side effect in trypanosomiasis patients, who largely live in poor areas in sub-Saharan Africa and

are exposed to numerous pathogens. However, TbTor1 could still provide a good drug target as depletion of the protein causes cell death, and the difference between TbTor1 and mammalian Tor would suggest that drugs targeting TbTor1 would not cross-react with the host's Tor pathway.

A reproducible observation during the rapamycin titration assays was the increase in cell number of the cultures treated with 50nM rapamycin after 72 hours (Figure 5.17). This discovery suggested that rapamycin was directly targeting TbTor1, therefore the growth defects seen at high concentrations of the drug probably do result from specific inhibition of a Tor protein. The growth benefit only becomes visible in the saw-tooth graph rather than the dose-response curve, because the increase only occurs after 72 hours and thus is not observed when the data are presented according the cumulative cell number. In the Barquilla et al (2008) study, the growth of rapamycin treated cells was displayed only as cumulative data, so it is unknown whether a similar phenomenon was detected in their experiments. However, the authors did report an increase in cell numbers in the early stages of TbTor1 depletion, prior to their growth arrest, suggesting that depletion of TbTor1 renders cells more resistant to stress (Barquilla and Navarro, 2009a). Similar results have been seen in other organisms, including vertebrates, where treatment of mice with low doses of rapamycin leads to increased lifespan (Harrison et al., 2009). Additionally in *C. elegans,* which is resistant to rapamycin, and in *Drosophila*, depletion of Tor leads to an increase in lifespan (Vellai et al., 2003, Kapahi et al., 2004). Experiments examining the effect of Tor depletion in budding yeast, through a combination of methods, showed that treatment with doses of rapamycin between 300pg/ml and 1ng/ml increased the lifespan of non-proliferating cells and that the addition of the drug caused the transcription factor Msn2 to relocate into the nucleus (Powers et al., 2006). This transcription factor is specific for genes involved in the stress response, and the rapamycin-treated cells also showed increased resistance to stress caused by high temperature and the presence of free radicals (Powers et al., 2006).

However, autophagy is also implicated in the increased survival of rapamycin-treated cells (Diaz-Troya et al., 2008). TORC1 is known to regulate the phosphorylation of the autophagy related gene ATG13 in yeast, and when this protein is in its phosphorylated state, it is unable to associate with the other autophagy related protein ATG1, thus preventing the formation of an active complex that leads to autophagy (Chang and Neufeld, 2009). Treatment with rapamycin, or depriving cells of essential amino acids, inactivates TORC1 allowing ATG13 to associate with ATG1 and thus activate autophagy (Chang and Neufeld, 2009). Tor inactivation (through either nitrogen depletion or treatment with rapamycin) also results in the translocation of Gln3 to the nucleus where it specifically transcribes genes required for autophagy (Chan et al., 2001). TbTor1 has already been implicated to play a role in autophagy as double-membraned structures corresponding to those seen in autophagy in other organisms are seen upon depletion of TbTor1. The authors claimed it was impossible to infer the role of autophagy in these phenotypes due to a lack autophagy markers for *T. brucei*, although double-membraned detected in TbTor1 depleted cells certainly implied that autophagy was occurring (Barquilla et al., 2008, Barquilla and Navarro, 2009a). Whilst bloodstream form *T. brucei* cells depleted of TbTor1 appear to show enhanced resistance to the stresses caused by growing the cells in media leached of nutrients, depletion of TbTor1 rapidly leads to growth arrest and thus it is not possible to ascribe this stress resistance to TbTor1 (Barquilla et al., 2008).

Another morphological marker of autophagy in *T. brucei* is the co-localisation of glycosomes into the lysosomes, a phenomenon that occurs during differentiation, demonstrating that pexophagy (autophagy of glycosomes) is occurring, thus this phenotype could also be examined in the TbTor1 depleted cells. Certainly, no molecular assays are available to quantify autophagy in *T. brucei*; ATG8 is used as a marker in a number of other species, where immunofluorescence using an antibody directed against this protein reveals punctate staining of autophagic vesicles at the initiation of autophagy (Chang and Neufeld, 2009). However, even though an ATG8 homologue exists in *T. brucei*, such a punctate pattern is not seen under conditions

inducing autophagy (Barquilla and Navarro, 2009a). Nonetheless, in a bio-informatic search, out of 40 yeast proteins involved in autophagy queried, 20 homologues were found in the *T. brucei* genome (Herman et al., 2006), therefore it should be possible to discover a genetic marker for autophagy.

In *Drosophila*, Tor directly regulates the ATG1/ATG13 complex through phosphorylation (Chang and Neufeld, 2009), yet there is no clear homologue of ATG1 or ATG13 in the *T. brucei* genome (Herman et al., 2006). Similarly, an alternative mechanism of Tor-dependent autophagy regulation in higher eukaryotes, through the induction of a transcription factor to specifically transcribe autophagy related genes, is unlikely to show conservation in *T. brucei* due to the polycistronic nature of transcription. Therefore, if autophagy is under TbTor1 regulation, it is conducted through a novel process.

### 6.3.3 Future directions

It was not possible to determine the reason for the differences seen between the effects of rapamycin presented here and those observed by Barquilla et al (2008), moreover as the same supplier of the rapamycin was used and the same solvent used to dissolve the drug in both studies. Also inexplicable are the differences in phenotype seen between the two studies in relationship to TbTor1 depletion, again peculiar as one of the clones generated used the same sequence for RNAi as was used in the Barquilla et al (2008) study. The results showing the poor response of *T. brucei* to rapamycin was consistently observed between batches of the drug, using both ethanol and DMSO as the solvent, and the high sensitivity of the *S. cerevisiae* positive control confirms that the rapamycin used for the *T. brucei* titration was active.

From the perspective of this PhD, based on the results I have generated, it seems that depletion of TbTor1 does not cause a specific G1 arrest, limiting the possibility that this protein would direct the cell cycle arrest required for stumpy cell formation. Further study of the TbTor pathway would be of great interest to fields beyond

parasitology due to the unique feature of at least three Tor homologues in *T. brucei* and possibly also in *Leishmania*. The early branching of trypanosomes from the other organisms that have characterised Tor pathways will allow interrogations into the evolution of these pathways, however the exact roles of the TbTors must first be confirmed.

Very little is known about autophagy in *T. brucei*, or indeed to what extent this process occurs, as the parasite appears to contain only a subset of the autophagy genes present in other species (Herman et al., 2006). Tor plays a major regulatory role in autophagy in a diverse array of organisms and so knowledge of the Tor pathway in *T. brucei* may allow greater investigation of the autophagy pathway in this organism.

### *6.4 Cdh1*

It has proved difficult to designate Tb927.8.6500 as either the homologue of Cdc20 or Cdh1, however the most interesting result from the analysis of APC/C coactivators was that there were no other suitable candidates for either protein in the rest of the genome, so either there is only the one activator of APC/C in *T. brucei*, or the other activators are evolutionarily very different. This is in contrast to the situation seen in a number of other organisms, which, in addition to having two APC/C coactivators for regulation of the mitotic cell cycle, also have more than one isoform of Cdh1. Examples include mammalian Cdh1, which is active in two alternatively spliced isoforms, showing differential expression between tissues and different localisation within cells (Zhou et al., 2003b). Furthermore, four Cdh1 homologues are present in the chicken genome with non-overlapping functions (Wan and Kirschner, 2001) and two Cdh1 isoforms are present in plants, although only one is capable of binding to the APC/C (Tarayre et al., 2004). The sequence of TbCdh1 is intriguing in having some Cdc20 specific domains, but not others. It is possible that the KEN box in TbCdh1 is sufficient for degradation and thus no other motifs, such as a CRY box, are required. Alternatively, TbCdh1 may have other sequences that recruit proteins to degrade it; deletion studies will be able to address these questions.

Depletion of TbCdh1 is lethal in bloodstream form cells as cells lose coordination of the cell cycle, resulting in cells with aberrant numbers of nuclei, kinetoplasts and flagella (Section 5.7). This phenotype suggests that TbCdh1 plays an important role in cell cycle regulation and shows more similarity to the results of Cdc20 depletion, rather than Cdh1 depletion, in higher eukaryotes. In yeast and mammals Cdh1 is not required for survival, however cells depleted of Cdh1 are unable to perform G1 arrest in response to the normal triggers of G1 arrest such as nutrient starvation or treatment with mating pheromone for yeast cells (Kitamura et al., 1998). It is possible that the severity of depletion of TbCdh1 occurs because it is the only activator of APC/C, and thus it has more diverse functions than the homologue in other organisms and could be required for other points in the cell cycle.

Knowledge of the role of TbCdh1 in regulation of the cell cycle is essential to understand more about the APC/C in *T. brucei*, and now that an antibody has been designed against this protein, the proteins that interact with it can be identified. Using an antibody against TbCdh1 for pull-down assays will rapidly allow the discovery of the proteins that are required for entry and exit of certain cell cycle stages. Further study of TbAPC/C might give more insight into cell cycle regulation if the proteins that interact with it during the different cell cycle stages can be identified. The role of TbCdh1 in differentiation also needs to be investigated and over expression of the protein will lend more knowledge about the function of the protein.

# Bibliography


ACOSTA-SERRANO, A., VASSELLA, E., LINIGER, M., KUNZ RENGGLI, C., BRUN, R., RODITI, I. & ENGLUND, P. T. (2001) The surface coat of procyclic Trypanosoma brucei: programmed expression and proteolytic cleavage of procyclin in the tsetse fly. *Proc Natl Acad Sci U S A,* 98**,** 1513-8.

ALEXANDER, K. & PARSONS, M. (1991) A phosphoglycerate kinase-like molecule localized to glycosomal microbodies: evidence that the topogenic signal is not at the C-terminus. *Mol Biochem Parasitol,* 46**,** 1-10.

ALIBU, V. P., STORM, L., HAILE, S., CLAYTON, C. & HORN, D. (2005) A doubly inducible system for RNA interference and rapid RNAi plasmid construction in Trypanosoma brucei. *Mol Biochem Parasitol,* 139**,** 75-82.

ALLISON, D. B., CUI, X., PAGE, G. P. & SABRIPOUR, M. (2006) Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet,* 7**,** 55-65.

ALVAREZ, B. & MORENO, S. (2006) Fission yeast Tor2 promotes cell growth and represses cell differentiation. *J Cell Sci,* 119**,** 4475-85.

ARCHER, S. K., LUU, V. D., DE QUEIROZ, R. A., BREMS, S. & CLAYTON, C. (2009) Trypanosoma brucei PUF9 regulates mRNAs for proteins involved in replicative processes over the cell cycle. *PLoS Pathog,* 5**,** e1000565.

ARNDT, C., CRUZ, M. C., CARDENAS, M. E. & HEITMAN, J. (1999) Secretion of FK506/FK520 and rapamycin by Streptomyces inhibits the growth of competing Saccharomyces cerevisiae and Cryptococcus neoformans. *Microbiology,* 145 ( Pt 8)**,** 1989-2000.

BACHLER, M., SCHROEDER, R. & VON AHSEN, U. (1999) StreptoTag: a novel method for the isolation of RNA-binding proteins. *RNA,* 5**,** 1509-16.

BALLOU, L. M. & LIN, R. Z. (2008) Rapamycin and mTOR kinase inhibitors. *J Chem Biol,* 1**,** 27-36.

BARBET, N. C., SCHNEIDER, U., HELLIWELL, S. B., STANSFIELD, I., TUITE, M. F. & HALL, M. N. (1996) TOR controls translation initiation and early G1 progression in yeast. *Mol Biol Cell,* 7**,** 25-42.

BARQUILLA, A., CRESPO, J. L. & NAVARRO, M. (2008) Rapamycin inhibits trypanosome cell growth by preventing TOR complex 2 formation. *Proc Natl Acad Sci U S A,* 105**,** 14579-84.

BARQUILLA, A. & NAVARRO, M. (2009a) Trypanosome TOR as a major regulator of cell growth and autophagy. *Autophagy,* 5**,** 256-8.

BARQUILLA, A. & NAVARRO, M. (2009b) Trypanosome TOR complex 2 functions in cytokinesis. *Cell Cycle,* 8**,** 697-9.

BARRY, J. D., GRAHAM, S. V., FOTHERINGHAM, M., GRAHAM, V. S., KOBRYN, K. & WYMER, B. (1998) VSG gene control and infectivity

strategy of metacyclic stage Trypanosoma brucei. *Mol Biochem Parasitol,* 91**,** 93-105.

BASS, K. E. & WANG, C. C. (1991) The in vitro differentiation of pleomorphic Trypanosoma brucei from bloodstream into procyclic form requires neither intermediary nor short-stumpy stage. *Mol Biochem Parasitol,* 44**,** 261-70.

BENJAMINI, Y. & HOCHBERG, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Satistical Society,* 57**,** 289-300.

BERBEROF, M., PAYS, A., LIPS, S., TEBABI, P. & PAYS, E. (1996) Characterization of a transcription terminator of the procyclin PARP A unit of Trypanosoma brucei. *Mol Cell Biol,* 16**,** 914-24.

BERBEROF, M., VANHAMME, L., TEBABI, P., PAYS, A., JEFFERIES, D., WELBURN, S. & PAYS, E. (1995) The 3'-terminal region of the mRNAs for VSG and procyclin can confer stage specificity to gene expression in Trypanosoma brucei. *Embo J,* 14**,** 2925-34.

BIENEN, E. J., SARIC, M., POLLAKIS, G., GRADY, R. W. & CLARKSON, A. B., JR. (1991) Mitochondrial development in Trypanosoma brucei brucei transitional bloodstream forms. *Mol Biochem Parasitol,* 45**,** 185-92.

BIRNBOIM, H. C. & DOLY, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res,* 7**,** 1513-23.

BLANCO, M. A., PELLOQUIN, L. & MORENO, S. (2001) Fission yeast mfr1 activates APC and coordinates meiotic nuclear division with sporulation. *J Cell Sci,* 114**,** 2135-43.

BLANCO, M. A., SANCHEZ-DIAZ, A., DE PRADA, J. M. & MORENO, S. (2000) APC(ste9/srw1) promotes degradation of mitotic cyclins in G(1) and is inhibited by cdc2 phosphorylation. *Embo J,* 19**,** 3945-55.

BORST, P. & SABATINI, R. (2008) Base J: discovery, biosynthesis, and possible functions. *Annu Rev Microbiol,* 62**,** 235-51.

BOUCHER, N., WU, Y., DUMAS, C., DUBE, M., SERENO, D., BRETON, M. & PAPADOPOULOU, B. (2002) A common mechanism of stage-regulated gene expression in Leishmania mediated by a conserved 3'-untranslated region element. *J Biol Chem,* 277**,** 19511-20.

BRADFORD, J., WHITNEY, P., HUANG, T., CHEUNG, C., YUE, S. & GODFREY, W. (2006) Cell cycle analysis in live cells using novel Vybrant Dyecycle stains with violet, blue, and green excitation.

BRANDEIS, M. & HUNT, T. (1996) The proteolysis of mitotic cyclins in mammalian cells persists from the end of mitosis until the onset of S phase. *Embo J,* 15**,** 5280-9.

BREIDBACH, T., NGAZOA, E. & STEVERDING, D. (2002) Trypanosoma brucei: in vitro slender-to-stumpy differentiation of culture-adapted, monomorphic bloodstream forms. *Exp Parasitol,* 101**,** 223-30.

BREMS, S., GUILBRIDE, D. L., GUNDLESDODJIR-PLANCK, D., BUSOLD, C., LUU, V. D., SCHANNE, M., HOHEISEL, J. & CLAYTON, C. (2005) The transcriptomes of Trypanosoma brucei Lister 427 and TREU927 bloodstream and procyclic trypomastigotes. *Mol Biochem Parasitol,* 139**,** 163-72.

BRIGGS, L. J., MCKEAN, P. G., BAINES, A., MOREIRA-LEITE, F., DAVIDGE, J., VAUGHAN, S. & GULL, K. (2004) The flagella connector of Trypanosoma brucei: an unusual mobile transmembrane junction. *J Cell Sci,* 117**,** 1641-51.

BRINGAUD, F. & BALTZ, T. (1993) Differential regulation of two distinct families of glucose transporter genes in Trypanosoma brucei. *Mol Cell Biol,* 13**,** 1146-54.

BROWN, E. J., ALBERS, M. W., SHIN, T. B., ICHIKAWA, K., KEITH, C. T., LANE, W. S. & SCHREIBER, S. L. (1994) A mammalian protein targeted by G1-arresting rapamycin-receptor complex. *Nature,* 369**,** 756-8.

BRUN, R. & SCHONENBERGER (1979) Cultivation and in vitro cloning or procyclic culture forms of Trypanosoma brucei in a semi-defined medium. Short communication. *Acta Trop,* 36**,** 289-92.

BRUN, R. & SCHONENBERGER, M. (1981) Stimulating effect of citrate and cis-Aconitate on the transformation of Trypanosoma brucei bloodstream forms to procyclic forms in vitro. *Z Parasitenkd,* 66**,** 17-24.

BURTON, J. L., TSAKRAKLIDES, V. & SOLOMON, M. J. (2005) Assembly of an APC-Cdh1-substrate complex is stimulated by engagement of a destruction box. *Mol Cell,* 18**,** 533-42.

CAMPBELL, D. A., THOMAS, S. & STURM, N. R. (2003) Transcription in kinetoplastid protozoa: why be normal? *Microbes Infect,* 5**,** 1231-40.

CECCHI, G., PAONE, M., FRANCO, J. R., FEVRE, E. M., DIARRA, A., RUIZ, J. A., MATTIOLI, R. C. & SIMARRO, P. P. (2009) Towards the Atlas of human African trypanosomiasis. *Int J Health Geogr,* 8**,** 15.

CHAN, T. F., BERTRAM, P. G., AI, W. & ZHENG, X. F. (2001) Regulation of APG14 expression by the GATA-type transcription factor Gln3p. *J Biol Chem,* 276**,** 6463-7.

CHANG, Y. Y. & NEUFELD, T. P. (2009) An Atg1/Atg13 complex with multiple roles in TOR-mediated autophagy regulation. *Mol Biol Cell,* 20**,** 2004-14.

CHEN, J., ZHENG, X. F., BROWN, E. J. & SCHREIBER, S. L. (1995) Identification of an 11-kDa FKBP12-rapamycin-binding domain within the 289-kDa FKBP12-rapamycin-associated protein and characterization of a critical serine residue. *Proc Natl Acad Sci U S A,* 92**,** 4947-51.

CHOI, J., CHEN, J., SCHREIBER, S. L. & CLARDY, J. (1996) Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science,* 273**,** 239-42.

CHOWDHURY, A. R., ZHAO, Z. & ENGLUND, P. T. (2008) Effect of hydroxyurea on procyclic Trypanosoma brucei: an unconventional mechanism for achieving synchronous growth. *Eukaryot Cell,* 7**,** 425-8.

CLAYTON, C. E. (1985) Structure and regulated expression of genes encoding fructose biphosphate aldolase in Trypanosoma brucei. *Embo J,* 4**,** 2997-3003.

CLAYTON, C. E. (1987) Import of fructose bisphosphate aldolase into the glycosomes of Trypanosoma brucei. *J Cell Biol,* 105**,** 2649-54.

CLAYTON, C. E. (2002) Life without transcriptional control? From fly to man and back again. *Embo J,* 21**,** 1881-8.

CLIFFE, L., KIEFT, R., MARSHALL, M., SIEGEL, T. N., CROSS, G. A. & SABATINI, R. (2009) The enrichment of base J at polycistronic transcriptional termination sites in the genome of Trypanosoma brucei is regulated by JBP1. *Kinetoplastid Molecular Cell Biology.* Woods Hole.

COLASANTE, C., ROBLES, A., LI, C.-H., SCHWEDE, A., BENZ, C., VONCKEN, F., GUILBRIDE, D. L. & CLAYTON, C. (2007) Regulated expression of glycosomal phosphoglycerate kinase in Trypanosoma brucei. *Molecular and Biochemical Parasitology,* 151**,** 193.

COMMITTEE, J. F. (2009) *British National Formulary,* London, British Medical Association and Royal Pharmaceutical Society of Great Britain.

COOPER, S. (2003) Rethinking synchronization of mammalian cells for cell cycle analysis. *Cell Mol Life Sci,* 60**,** 1099-106.

COOPER, S. (2004) Rejoinder: whole-culture synchronization cannot, and does not, synchronize cells. *Trends Biotechnol,* 22**,** 274-6.

COOPER, S. & SHEDDEN, K. (2003) Microarray analysis of gene expression during the cell cycle. *Cell Chromosome,* 2**,** 1.

COQUELLE, A., MOUHAMAD, S., PEQUIGNOT, M. O., BRAUN, T., CARVALHO, G., VIVET, S., METIVIER, D., CASTEDO, M. & KROEMER, G. (2006) Enrichment of non-synchronized cells in the G1, S and G2 phases of the cell cycle for the study of apoptosis. *Biochem Pharmacol,* 72**,** 1396-404.

CRESPO, J. L., DIAZ-TROYA, S. & FLORENCIO, F. J. (2005) Inhibition of Target of Rapamycin Signaling by Rapamycin in the Unicellular Green Alga Chlamydomonas reinhardtii. *Plant Physiol,* 139**,** 1736-1749.

CROSS, G. A. & MANNING, J. C. (1973) Cultivation of Trypanosoma brucei sspp. in semi-defined and defined media. *Parasitology,* 67**,** 315-31.

CUNNINGHAM, M. P. & VICKERMAN, K. (1962) Antigenic analysis in the Trypanosoma brucei group, using the agglutination reaction. *Trans R Soc Trop Med Hyg,* 56**,** 48-59.

CUTLER, N. S., PAN, X., HEITMAN, J. & CARDENAS, M. E. (2001) The TOR signal transduction cascade controls cellular differentiation in response to nutrients. *Mol Biol Cell,* 12**,** 4103-13.

DANGERFIELD, J. A., WINDBICHLER, N., SALMONS, B., GUNZBURG, W. H. & SCHRODER, R. (2006) Enhancement of the StreptoTag method for isolation of endogenously expressed proteins with complex RNA binding targets. *Electrophoresis,* 27**,** 1874-7.

DAS, A., GALE, M., JR., CARTER, V. & PARSONS, M. (1994) The protein phosphatase inhibitor okadaic acid induces defects in cytokinesis and organellar genome segregation in Trypanosoma brucei. *J Cell Sci,* 107 ( Pt 12)**,** 3477-83.

DE LA CRUZ, J., KRESSLER, D., ROJO, M., TOLLERVEY, D. & LINDER, P. (1998) Spb4p, an essential putative RNA helicase, is required for a late step in the assembly of 60S ribosomal subunits in Saccharomyces cerevisiae. *RNA,* 4**,** 1268-81.

DEAN, S., MARCHETTI, R., KIRK, K. & MATTHEWS, K. R. (2009) A surface transporter family conveys the trypanosome differentiation signal. *Nature,* 459**,** 213-7.

DENNINGER, V., FIGARELLA, K., SCHONFELD, C., BREMS, S., BUSOLD, C., LANG, F., HOHEISEL, J. & DUSZENKO, M. (2007) Troglitazone induces differentiation in Trypanosoma brucei. *Exp Cell Res,* 313**,** 1805-19.

DIAZ-TROYA, S., PEREZ-PEREZ, M. E., FLORENCIO, F. J. & CRESPO, J. L. (2008) The role of TOR in autophagy regulation from yeast to plants and mammals. *Autophagy,* 4**,** 851-65.

DIEHL, S., DIEHL, F., EL-SAYED, N. M., CLAYTON, C. & HOHEISEL, J. D. (2002) Analysis of stage-specific gene expression in the bloodstream and the procyclic form of Trypanosoma brucei using a genomic DNA-microarray. *Mol Biochem Parasitol,* 123**,** 115-23.

DOMENICALI PFISTER, D., BURKARD, G., MORAND, S., RENGGLI, C. K., RODITI, I. & VASSELLA, E. (2006) A Mitogen-activated protein kinase controls differentiation of bloodstream forms of Trypanosoma brucei. *Eukaryot Cell,* 5**,** 1126-35.

DRAGHICI, S. (2003) *Data analysis tools for DNA microarrays,* London, Chapman & Hall/CRC.

DUBE, P., HERZOG, F., GIEFFERS, C., SANDER, B., RIEDEL, D., MULLER, S. A., ENGEL, A., PETERS, J. M. & STARK, H. (2005) Localization of the coactivator Cdh1 and the cullin subunit Apc2 in a cryo-electron microscopy model of vertebrate APC/C. *Mol Cell,* 20**,** 867-79.

DUNCAN, R. (2004) DNA microarray analysis of protozoan parasite gene expression: outcomes correlate with mechanisms of regulation. *Trends Parasitol,* 20**,** 211-5.

EDWARDS, S. R. & WANDLESS, T. J. (2007) The rapamycin-binding domain of the protein kinase mammalian target of rapamycin is a destabilizing domain. *J Biol Chem,* 282**,** 13395-401.

EL-SAYED, N. M., HEGDE, P., QUACKENBUSH, J., MELVILLE, S. E. & DONELSON, J. E. (2000) The African trypanosome genome. *International Journal for Parasitology,* 30**,** 329.

ENGSTLER, M. & BOSHART, M. (2004) Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in Trypanosoma brucei. *Genes Dev,* 18**,** 2798-811.

ENGSTLER, M., PFOHL, T., HERMINGHAUS, S., BOSHART, M., WIEGERTJES, G., HEDDERGOTT, N. & OVERATH, P. (2007) Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes. *Cell,* 131**,** 505-15.

ERSFELD, K., DOCHERTY, R., ALSFORD, S. & GULL, K. (1996) A fluorescence in situ hybridisation study of the regulation of histone mRNA levels during the cell cycle of Trypanosoma brucei. *Mol Biochem Parasitol,* 81**,** 201-9.

ESTEVEZ, A. M. (2008) The RNA-binding protein TbDRBD3 regulates the stability of a specific subset of mRNAs in trypanosomes. *Nucleic Acids Res,* 36**,** 4573-86.

FANG, G., YU, H. & KIRSCHNER, M. W. (1998) Direct binding of CDC20 protein family members activates the anaphase-promoting complex in mitosis and G1. *Mol Cell,* 2**,** 163-71.

FENN, K. & MATTHEWS, K. R. (2007) The cell biology of Trypanosoma brucei differentiation. *Curr Opin Microbiol,* 10**,** 539-46.

FEVRE, E. M., WISSMANN, B. V., WELBURN, S. C. & LUTUMBA, P. (2008) The Burden of Human African Trypanosomiasis. *PLoS Negl Trop Dis,* 2**,** e333.

FLOYD, S., PINES, J. & LINDON, C. (2008) APC/CCdh1 Targets Aurora Kinase to Control Reorganization of the Mitotic Spindle at Anaphase. *Current Biology,* 18**,** 1649.

FORSYTHE, G. R., MCCULLOCH, R. & HAMMARTON, T. C. (2008) Hydroxyurea-induced synchronisation of bloodstream stage Trypanosoma brucei. *Mol Biochem Parasitol*.

GALE, M., JR., CARTER, V. & PARSONS, M. (1994a) Cell cycle-specific induction of an 89 kDa serine/threonine protein kinase activity in Trypanosoma brucei. *J Cell Sci,* 107 ( Pt 7)**,** 1825-32.

GALE, M., JR., CARTER, V. & PARSONS, M. (1994b) Translational control mediates the developmental regulation of the Trypanosoma brucei Nrk protein kinase. *J Biol Chem,* 269**,** 31659-65.

GARCIA-SALCEDO, J. A., GIJON, P. & PAYS, E. (1999) Regulated transcription of the histone H2B genes of Trypanosoma brucei. *Eur J Biochem,* 264**,** 717-23.

GIBSON, W. C., SWINKELS, B. W. & BORST, P. (1988) Post-transcriptional control of the differential expression of phosphoglycerate kinase genes in Trypanosoma brucei. *Journal of Molecular Biology,* 201**,** 315.

GILINGER, G. & BELLOFATTO, V. (2001) Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. *Nucleic Acids Res,* 29**,** 1556-64.

GINGER, M. L., BLUNDELL, P. A., LEWIS, A. M., BROWITT, A., GUNZL, A. & BARRY, J. D. (2002) Ex vivo and in vitro identification of a consensus promoter for VSG genes expressed by metacyclic-stage trypanosomes in the tsetse fly. *Eukaryot Cell,* 1**,** 1000-9.

GINGRAS, A. C., RAUGHT, B. & SONENBERG, N. (2001) Regulation of translation initiation by FRAP/mTOR. *Genes Dev,* 15**,** 807-26.

GOPAL, S., CROSS, G. A. & GAASTERLAND, T. (2003) An organism-specific method to rank predicted coding regions in Trypanosoma brucei. *Nucleic Acids Res,* 31**,** 5877-85.

GOTHEL, S. F. & MARAHIEL, M. A. (1999) Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. *Cell Mol Life Sci,* 55**,** 423-36.

GOULD, M. K., VU, X. L., SEEBECK, T. & DE KONING, H. P. (2008) Propidium iodide-based methods for monitoring drug action in the kinetoplastidae: comparison with the Alamar Blue assay. *Anal Biochem,* 382**,** 87-93.

GOURGUECHON, S. & WANG, C. C. (2009) CRK9 contributes to regulation of mitosis and cytokinesis in the procyclic form of Trypanosoma brucei. *BMC Cell Biol,* 10**,** 68.

GRAVES, A. R., CURRAN, P. K., SMITH, C. L. & MINDELL, J. A. (2008) The Cl-/H+ antiporter ClC-7 is the primary chloride permeation pathway in lysosomes. *Nature,* 453**,** 788-92.

GUTHER, M. L., LEE, S., TETLEY, L., ACOSTA-SERRANO, A. & FERGUSON, M. A. (2006) GPI-anchored proteins and free GPI glycolipids of procyclic form Trypanosoma brucei are nonessential for growth, are required for colonization of the tsetse fly, and are not the only components of the surface coat. *Mol Biol Cell,* 17**,** 5265-74.

HAILE, S. & PAPADOPOULOU, B. (2007) Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Current Opinion in Microbiology,* 10**,** 569.

HALL, B. G. (2001) *Phylogenetic trees made easy: A how-to manual for molecular biologists,* Sunderland, Sinauer Associates.

HAMMARTON, T. C. (2007) Cell cycle regulation in Trypanosoma brucei. *Molecular and Biochemical Parasitology,* 153**,** 1.

HAMMARTON, T. C., CLARK, J., DOUGLAS, F., BOSHART, M. & MOTTRAM, J. C. (2003) Stage-specific Differences in Cell Cycle Control in Trypanosoma brucei Revealed by RNA Interference of a Mitotic Cyclin. *J. Biol. Chem,* 278**,** 22877-22886.

HAMMARTON, T. C., ENGSTLER, M. & MOTTRAM, J. C. (2004) The Trypanosoma brucei cyclin, CYC2, is required for cell cycle progression

through G1 phase and for maintenance of procyclic form cell morphology. *J Biol Chem,* 279**,** 24757-64.

HARPER, J. W., BURTON, J. L. & SOLOMON, M. J. (2002) The anaphase-promoting complex: it's not just for mitosis any more. *Genes Dev,* 16**,** 2179-206.

HARRISON, D. E., STRONG, R., SHARP, Z. D., NELSON, J. F., ASTLE, C. M., FLURKEY, K., NADON, N. L., WILKINSON, J. E., FRENKEL, K., CARTER, C. S., PAHOR, M., JAVORS, M. A., FERNANDEZ, E. & MILLER, R. A. (2009) Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature,* 460**,** 392-5.

HART, D. T., MISSET, O., EDWARDS, S. W. & OPPERDOES, F. R. (1984) A comparison of the glycosomes (microbodies) isolated from Trypanosoma brucei bloodstream form and cultured procyclic trypomastigotes. *Mol Biochem Parasitol,* 12**,** 25-35.

HARTMANN, C., BENZ, C., BREMS, S., ELLIS, L., LUU, V. D., STEWART, M., D'ORSO, I., BUSOLD, C., FELLENBERG, K., FRASCH, A. C., CARRINGTON, M., HOHEISEL, J. & CLAYTON, C. E. (2007) Small trypanosome RNA-binding proteins TbUBP1 and TbUBP2 influence expression of F-box protein mRNAs in bloodstream trypanosomes. *Eukaryot Cell,* 6**,** 1964-78.

HARTMANN, C. & CLAYTON, C. (2008) Regulation of a transmembrane protein gene family by the small RNA-binding proteins TbUBP1 and TbUBP2. *Mol Biochem Parasitol,* 157**,** 112-5.

HAY, N. & SONENBERG, N. (2004) Upstream and downstream of mTOR. *Genes Dev,* 18**,** 1926-45.

HEITMAN, J., MOVVA, N. R. & HALL, M. N. (1991) Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. *Science,* 253**,** 905-909.

HELMSTETTER, C. E., THORNTON, M., ROMERO, A. & EWARD, K. L. (2003) Synchrony in human, mouse and bacterial cell cultures--a comparison. *Cell Cycle,* 2**,** 42-5.

HENDRIKS, E. F. & MATTHEWS, K. R. (2005) Disruption of the developmental programme of Trypanosoma brucei by genetic ablation of TbZFP1, a differentiation-enriched CCCH protein. *Mol Microbiol,* 57**,** 706-16.

HERMAN, M., GILLIES, S., MICHELS, P. A. & RIGDEN, D. J. (2006) Autophagy and related processes in trypanosomatids: insights from genomic and bioinformatic analyses. *Autophagy,* 2**,** 107-18.

HERMAN, M., PEREZ-MORGA, D., SCHTICKZELLE, N. & MICHELS, P. A. (2008) Turnover of glycosomes during life-cycle differentiation of Trypanosoma brucei. *Autophagy,* 4**,** 294-308.

HIDE, G. & TAIT, A. (2006) Molecular epidemiology of African sleeping sickness. *Parasitology,* 136(12), 1491-500.

HIRUMI, H. & HIRUMI, K. (1989) Continuous cultivation of Trypanosoma brucei blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *J Parasitol,* 75**,** 985-9.

HOBBS, M. R. & BOOTHROYD, J. C. (1990) An expression-site-associated gene family of trypanosomes is expressed in vivo and shows homology to a variant surface glycoprotein gene. *Mol Biochem Parasitol,* 43**,** 1-16.

HOLZER, T. R., MCMASTER, W. R. & FORNEY, J. D. (2006) Expression profiling by whole-genome interspecies microarray hybridization reveals differential gene expression in procyclic promastigotes, lesion-derived amastigotes, and axenic amastigotes in Leishmania mexicana. *Mol Biochem Parasitol,* 146**,** 198-218.

HOTZ, H. R., LORENZ, P., FISCHER, R., KRIEGER, S. & CLAYTON, C. (1995) Role of 3'-untranslated regions in the regulation of hexose transporter mRNAs in Trypanosoma brucei. *Mol Biochem Parasitol,* 75**,** 1-14.

HU, C., RIO, R. V., MEDLOCK, J., HAINES, L. R., NAYDUCH, D., SAVAGE, A. F., GUZ, N., ATTARDO, G. M., PEARSON, T. W., GALVANI, A. P. & AKSOY, S. (2008) Infections with immunogenic trypanosomes reduce tsetse reproductive fitness: potential impact of different parasite strains on vector population structure. *PLoS Negl Trop Dis,* 2**,** e192.

HUYNH, M. A., STEGMULLER, J., LITTERMAN, N. & BONNI, A. (2009) Regulation of Cdh1-APC Function in Axon Growth by Cdh1 Phosphorylation. *J. Neurosci,* 29**,** 4322-4327.

INBAL, N., LISTOVSKY, T. & BRANDEIS, M. (1999) The mammalian Fizzy and Fizzy-related genes are regulated at the transcriptional and post-transcriptional levels. *FEBS Lett,* 463**,** 350-4.

INOKI, K., OUYANG, H., LI, Y. & GUAN, K. L. (2005) Signaling by target of rapamycin proteins in cell growth control. *Microbiol Mol Biol Rev,* 69**,** 79-100.

JACINTO, E., LOEWITH, R., SCHMIDT, A., LIN, S., RUEGG, M. A., HALL, A. & HALL, M. N. (2004) Mammalian TOR complex 2 controls the actin cytoskeleton and is rapamycin insensitive. *Nat Cell Biol,* 6**,** 1122-8.

JACKSON, D. G., SMITH, D. K., LUO, C. & ELLIOTT, J. F. (1993) Cloning of a novel surface antigen from the insect stages of Trypanosoma brucei by expression in COS cells. *J Biol Chem,* 268**,** 1894-900.

JANZEN, C. J., HAKE, S. B., LOWELL, J. E. & CROSS, G. A. (2006) Selective di- or trimethylation of histone H3 lysine 76 by two DOT1 homologs is important for cell cycle regulation in Trypanosoma brucei. *Mol Cell,* 23**,** 497-507.

JAQUENOUD, M., VAN DROGEN, F. & PETER, M. (2002) Cell cycle-dependent nuclear export of Cdh1p may contribute to the inactivation of APC/C(Cdh1). *Embo J,* 21**,** 6515-26.

JENSEN, B. C., SIVAM, D., KIFER, C. T., MYLER, P. J. & PARSONS, M. (2009) Widespread variation in transcript abundance within and across developmental stages of Trypanosoma brucei. *BMC Genomics,* 10**,** 482.

KANE, M. D., JATKOE, T. A., STUMPF, C. R., LU, J., THOMAS, J. D. & MADORE, S. J. (2000) Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Res,* 28**,** 4552-7.

KAPAHI, P., ZID, B. M., HARPER, T., KOSLOVER, D., SAPIN, V. & BENZER, S. (2004) Regulation of lifespan in Drosophila by modulation of genes in the TOR signaling pathway. *Curr Biol,* 14**,** 885-90.

KENNEDY, P. G. E. (2008) The continuing problem of human African trypanosomiasis (sleeping sickness). *Annals of Neurology,* 64**,** 116-126.

KIM, D. H., SARBASSOV, D. D., ALI, S. M., KING, J. E., LATEK, R. R., ERDJUMENT-BROMAGE, H., TEMPST, P. & SABATINI, D. M. (2002) mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell,* 110**,** 163-75.

KITAMURA, K., MAEKAWA, H. & SHIMODA, C. (1998) Fission yeast Ste9, a homolog of Hct1/Cdh1 and Fizzy-related, is a novel negative regulator of cell cycle progression during G1-phase. *Mol Biol Cell,* 9**,** 1065-80.

KOLTIN, Y., FAUCETTE, L., BERGSMA, D. J., LEVY, M. A., CAFFERKEY, R., KOSER, P. L., JOHNSON, R. K. & LIVI, G. P. (1991) Rapamycin sensitivity in Saccharomyces cerevisiae is mediated by a peptidyl-prolyl cis-trans isomerase related to human FK506-binding protein. *Mol. Cell. Biol.,* 11**,** 1718-1723.

KOUMANDOU, V., NATESAN, S., SERGEENKO, T. & FIELD, M. (2008) The trypanosome transcriptome is remodelled during differentiation but displays limited responsiveness within life stages. *BMC Genomics*, 23(9), e298.

KRAFSUR, E. S. (2009) Tsetse flies: Genetics, evolution, and role as vectors. *Infection, Genetics and Evolution,* 9**,** 124.

KRAFT, C., VODERMAIER, H. C., MAURER-STROH, S., EISENHABER, F. & PETERS, J. M. (2005) The WD40 propeller domain of Cdh1 functions as a destruction box receptor for APC/C substrates. *Mol Cell,* 18**,** 543-53.

KRAMER, E. R., GIEFFERS, C., HOLZL, G., HENGSTSCHLAGER, M. & PETERS, J. M. (1998) Activation of the human anaphase-promoting complex by proteins of the CDC20/Fizzy family. *Curr Biol,* 8**,** 1207-10.

KRAMER, E. R., SCHEURINGER, N., PODTELEJNIKOV, A. V., MANN, M. & PETERS, J. M. (2000) Mitotic regulation of the APC activator proteins CDC20 and CDH1. *Mol Biol Cell,* 11**,** 1555-69.

KUMAR, P. & WANG, C. C. (2005) Depletion of anaphase-promoting complex or cyclosome (APC/C) subunit homolog APC1 or CDC27 of Trypanosoma brucei arrests the procyclic form in metaphase but the bloodstream form in anaphase. *J Biol Chem,* 280**,** 31783-91.

KUNZ, J., HENRIQUEZ, R., SCHNEIDER, U., DEUTER-REINHARD, M., MOVVA, N. R. & HALL, M. N. (1993) Target of rapamycin in yeast, TOR2, is an essential phosphatidylinositol kinase homolog required for G1 progression. *Cell,* 73**,** 585-96.

LANHAM, S. M. & GODFREY, D. G. (1970) Isolation of salivarian trypanosomes from man and other mammals using DEAE-cellulose. *Exp Parasitol,* 28**,** 521-34.

LAXMAN, S., RIECHERS, A., SADILEK, M., SCHWEDE, F. & BEAVO, J. A. (2006) Hydrolysis products of cAMP analogs cause transformation of Trypanosoma brucei from slender to stumpy-like forms. *Proc Natl Acad Sci U S A,* 103**,** 19194-9.

LEE, S., COMER, F. I., SASAKI, A., MCLEOD, I. X., DUONG, Y., OKUMURA, K., YATES, J. R., 3RD, PARENT, C. A. & FIRTEL, R. A. (2005) TOR complex 2 integrates cell movement during chemotaxis and signal relay in Dictyostelium. *Mol Biol Cell,* 16**,** 4572-83.

LI, H., TSANG, C. K., WATKINS, M., BERTRAM, P. G. & ZHENG, X. F. (2006) Nutrient regulates Tor1 nuclear localization and association with rDNA promoter. *Nature,* 442**,** 1058-61.

LI, M. & ZHANG, P. (2009) The function of APC/CCdh1 in cell cycle and beyond. *Cell Division,* 4**,** 2.

LI, Y., LI, Z. & WANG, C. C. (2003) Differentiation of Trypanosoma brucei may be stage non-specific and does not require progression of cell cycle. *Mol Microbiol,* 49**,** 251-65.

LI, Z. & WANG, C. C. (2006) Changing roles of aurora-B kinase in two life cycle stages of Trypanosoma brucei. *Eukaryot Cell,* 5**,** 1026-35.

LISTER, R., GREGORY, B. D. & ECKER, J. R. (2009) Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. *Curr Opin Plant Biol,* 12**,** 107-18.

LISTOVSKY, T., OREN, Y. S., YUDKOVSKY, Y., MAHBUBANI, H. M., WEISS, A. M., LEBENDIKER, M. & BRANDEIS, M. (2004) Mammalian Cdh1/Fzr mediates its own degradation. *Embo J,* 23**,** 1619-26.

LIU, B., LIU, Y., MOTYKA, S. A., AGBO, E. E. & ENGLUND, P. T. (2005) Fellowship of the rings: the replication of kinetoplast DNA. *Trends Parasitol,* 21**,** 363-9.

LIU, Y., HIDAYAT, S., SU, W. H., DENG, X., YU, D. H. & YU, B. Z. (2007) Expression and activity of mTOR and its substrates in different cell cycle phases and in oral squamous cell carcinomas of different malignant grade. *Cell Biochem Funct,* 25**,** 45-53.

LOEWITH, R., JACINTO, E., WULLSCHLEGER, S., LORBERG, A., CRESPO, J. L., BONENFANT, D., OPPLIGER, W., JENOE, P. & HALL, M. N. (2002) Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. *Mol Cell,* 10**,** 457-68.

LONG, X., SPYCHER, C., HAN, Z. S., ROSE, A. M., MULLER, F. & AVRUCH, J. (2002) TOR deficiency in C. elegans causes developmental arrest and intestinal atrophy by inhibition of mRNA translation. *Curr Biol,* 12**,** 1448-61.

LORENZ, M. C. & HEITMAN, J. (1995) TOR mutations confer rapamycin resistance by preventing interaction with FKBP12-rapamycin. *J Biol Chem,* 270**,** 27531-7.

LYTHGOE, K. A., MORRISON, L. J., READ, A. F. & BARRY, J. D. (2007) Parasite-intrinsic factors can explain ordered progression of trypanosome antigenic variation. *Proc Natl Acad Sci U S A,* 104**,** 8095-100.

MACGREGOR, P. & MATTHEWS, K. R. (2008) Modelling trypanosome chronicity: VSG dynasties and parasite density. *Trends Parasitol,* 24**,** 1-4.

MADEIRA DA SILVA, L., OWENS, K. L., MURTA, S. M. F. & BEVERLEY, S. M. (2009) Regulated expression of the Leishmania major surface virulence factor lipophosphoglycan using conditionally destabilized fusion proteins. *Proc Natl Acad Sci U S A,* 106**,** 7583-7588.

MAIR, G. R., BRAKS, J. A., GARVER, L. S., WIEGANT, J. C., HALL, N., DIRKS, R. W., KHAN, S. M., DIMOPOULOS, G., JANSE, C. J. & WATERS, A. P. (2006) Regulation of sexual development of Plasmodium by translational repression. *Science,* 313**,** 667-9.

MAKKY, K., TEKIELA, J. & MAYER, A. N. (2007) Target of rapamycin (TOR) signaling controls epithelial morphogenesis in the vertebrate intestine. *Developmental Biology,* 303**,** 501.

MAO, Y., SHATERI NAJAFABADI, H. & SALAVATI, R. (2009) Genome-wide computational identification of functional RNA elements in Trypanosoma brucei. *BMC Genomics,* 10**,** 355.

MARTIN, K. L. & SMITH, T. K. (2006) The glycosylphosphatidylinositol (GPI) biosynthetic pathway of bloodstream-form Trypanosoma brucei is dependent on the de novo synthesis of inositol. *Mol Microbiol,* 61**,** 89-105.

MATHIEU-DAUDE, F., WELSH, J., DAVIS, C. & MCCLELLAND, M. (1998) Differentially expressed genes in the Trypanosoma brucei life cycle identified by RNA fingerprinting. *Mol Biochem Parasitol,* 92**,** 15-28.

MATSUO, T., OTSUBO, Y., URANO, J., TAMANOI, F. & YAMAMOTO, M. (2007) Loss of the TOR kinase Tor2 mimics nitrogen starvation and activates the sexual development pathway in fission yeast. *Mol Cell Biol,* 27**,** 3154-64.

MATTHEWS, K. R. & GULL, K. (1994) Evidence for an interplay between cell cycle progression and the initiation of differentiation between life cycle forms of African trypanosomes. *J Cell Biol,* 125**,** 1147-56.

MATTHEWS, K. R. & GULL, K. (1997) Commitment to differentiation and cell cycle re-entry are coincident but separable events in the transformation of African trypanosomes from their bloodstream to their insect form. *J Cell Sci,* 110 ( Pt 20)**,** 2609-18.

MATTHEWS, K. R. & GULL, K. (1998) Identification of stage-regulated and differentiation-enriched transcripts during transformation of the African trypanosome from its bloodstream to procyclic form. *Mol Biochem Parasitol,* 95**,** 81-95.

MATTHEWS, K. R., SHERWIN, T. & GULL, K. (1995) Mitochondrial genome repositioning during the differentiation of the African trypanosome between life cycle forms is microtubule mediated. *J Cell Sci,* 108 ( Pt 6)**,** 2231-9.

MATYSKIELA, M. E. & MORGAN, D. O. (2009) Analysis of activator-binding sites on the APC/C supports a cooperative substrate-binding mechanism. *Mol Cell,* 34**,** 68-80.

MAYHO, M., FENN, K., CRADDY, P., CROSTHWAITE, S. & MATTHEWS, K. (2006) Post-transcriptional control of nuclear-encoded cytochrome oxidase subunits in Trypanosoma brucei: evidence for genome-wide conservation of life-cycle stage-specific regulatory elements. *Nucl. Acids Res,* 598(34)**,** 5312-5324.

MCCANN, A. K., SCHWARTZ, K. J. & BANGS, J. D. (2008) A determination of the steady state lysosomal pH of bloodstream stage African trypanosomes. *Molecular and Biochemical Parasitology,* 159**,** 146.

MCKEAN, P. G. (2003) Coordination of cell cycle and cytokinesis in Trypanosoma brucei. *Curr Opin Microbiol,* 6**,** 600-7.

MCLINTOCK, L. M., TURNER, C. M. & VICKERMAN, K. (1993) Comparison of the effects of immune killing mechanisms on Trypanosoma brucei parasites of slender and stumpy morphology. *Parasite Immunol,* 15**,** 475-80.

MCNICOLL, F., DRUMMELSMITH, J., MULLER, M., MADORE, E., BOILARD, N., OUELLETTE, M. & PAPADOPOULOU, B. (2006) A combined proteomic and transcriptomic approach to the study of stage differentiation in Leishmania infantum. *Proteomics,* 6**,** 3567-81.

MICHELOTTI, E. F., HARRIS, M. E., ADLER, B., TORRI, A. F. & HAJDUK, S. L. (1992) Trypanosoma brucei mitochondrial ribosomal RNA synthesis, processing and developmentally regulated expression. *Mol Biochem Parasitol,* 54**,** 31-41.

MIKULITS, W., DOLZNIG, H., EDELMANN, H., SAUER, T., DEINER, E. M., BALLOU, L., BEUG, H. & MULLNER, E. W. (1997) Dynamics of cell cycle regulators: artifact-free analysis by recultivation of cells synchronized by centrifugal elutriation. *DNA Cell Biol,* 16**,** 849-59.

MILNE, K. G., PRESCOTT, A. R. & FERGUSON, M. A. (1998) Transformation of monomorphic Trypanosoma brucei bloodstream form trypomastigotes into procyclic forms at 37 degrees C by removing glucose from the culture medium. *Mol Biochem Parasitol,* 94**,** 99-112.

MINNING, T. A., WEATHERLY, D. B., ATWOOD, J., 3RD, ORLANDO, R. & TARLETON, R. L. (2009) The steady-state transcriptome of the four major life-cycle stages of Trypanosoma cruzi. *BMC Genomics,* 10**,** 370.

MISSEL, A., LAMBERT, L., NORSKAU, G. & GORINGER, H. U. (1999) DEAD-box protein HEL64 from Trypanosoma brucei: subcellular localization and gene knockout analysis. *Parasitol Res,* 85**,** 324-30.

MISSEL, A., SOUZA, A. E., NORSKAU, G. & GORINGER, H. U. (1997) Disruption of a gene encoding a novel mitochondrial DEAD-box protein in Trypanosoma brucei affects edited mRNAs. *Mol Cell Biol,* 17**,** 4895-903.

MONNERAT, S., CLUCAS, C., BROWN, E., MOTTRAM, J. C. & HAMMARTON, T. C. (2009) Searching for novel cell cycle regulators in Trypanosoma brucei with an RNA interference screen. *BMC Res Notes,* 2**,** 46.

MOREY, J. S., RYAN, J. C. & VAN DOLAH, F. M. (2006) Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biol Proced Online,* 8**,** 175-93.

MORROW, K. J. (2007) Cytometry drives new business models. *Genetic Engineering & Biotechnology News,* 27(17).

MURAKAMI, M., ICHISAKA, T., MAEDA, M., OSHIRO, N., HARA, K., EDENHOFER, F., KIYAMA, H., YONEZAWA, K. & YAMANAKA, S. (2004) mTOR is essential for growth and proliferation in early mouse embryos and embryonic stem cells. *Mol Cell Biol,* 24**,** 6710-8.

MUTOMBA, M. C. & WANG, C. C. (1996) Effects of aphidicolin and hydroxyurea on the cell cycle and differentiation of Trypanosoma brucei bloodstream forms. *Mol Biochem Parasitol,* 80**,** 89-102.

MUTOMBA, M. C. & WANG, C. C. (1998) The role of proteolysis during differentiation of Trypanosoma brucei from the bloodstream to the procyclic form. *Mol Biochem Parasitol,* 93**,** 11-22.

NAGALAKSHMI, U., WANG, Z., WAERN, K., SHOU, C., RAHA, D., GERSTEIN, M. & SNYDER, M. (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science,* 320**,** 1344-9.

NARDELLI, S. C., AVILA, A. R., FREUND, A., MOTTA, M. C., MANHAES, L., DE JESUS, T. C., SCHENKMAN, S., FRAGOSO, S. P., KRIEGER, M. A., GOLDENBERG, S. & DALLAGIOVANNA, B. (2007) Small-subunit rRNA processome proteins are translationally regulated during differentiation of Trypanosoma cruzi. *Eukaryot Cell,* 6**,** 337-45.

NATESAN, S. K., PEACOCK, L., MATTHEWS, K., GIBSON, W. & FIELD, M. C. (2007) Activation of endocytosis as an adaptation to the mammalian host by trypanosomes. *Eukaryot Cell,* 6**,** 2029-37.

NESS, S. A. (2007) Microarray analysis: basic strategies for successful experiments. *Mol Biotechnol,* 36**,** 205-19.

NICKLIN, P., BERGMAN, P., ZHANG, B., TRIANTAFELLOW, E., WANG, H., NYFELER, B., YANG, H., HILD, M., KUNG, C., WILSON, C., MYER, V. E., MACKEIGAN, J. P., PORTER, J. A., WANG, Y. K., CANTLEY, L. C., FINAN, P. M. & MURPHY, L. O. (2009) Bidirectional transport of amino acids regulates mTOR and autophagy. *Cell,* 136**,** 521-34.

NOLAN, D. P., ROLIN, S., RODRIGUEZ, J. R., VAN DEN ABBEELE, J. & PAYS, E. (2000) Slender and stumpy bloodstream forms of Trypanosoma brucei display a differential response to extracellular acidic and proteolytic stress. *Eur J Biochem,* 267**,** 18-27.

O'BRIEN, J., WILSON, I., ORTON, T. & POGNAN, F. (2000) Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity. *Eur J Biochem,* 267**,** 5421-6.

PALENCHAR, J. B. & BELLOFATTO, V. (2006) Gene transcription in trypanosomes. *Mol Biochem Parasitol,* 146**,** 135-41.

PANIGRAHI, A. K., OGATA, Y., ZIKOVA, A., ANUPAMA, A., DALLEY, R. A., ACESTOR, N., MYLER, P. J. & STUART, K. D. (2009) A comprehensive analysis of Trypanosoma brucei mitochondrial proteome. *Proteomics,* 9**,** 434-50.

PARSONS, M., WORTHEY, E. A., WARD, P. N. & MOTTRAM, J. C. (2005) Comparative analysis of the kinomes of three pathogenic trypanosomatids: Leishmania major, Trypanosoma brucei and Trypanosoma cruzi. *BMC Genomics,* 6**,** 127.

PATTERSON, T. A., LOBENHOFER, E. K., FULMER-SMENTEK, S. B., COLLINS, P. J., CHU, T. M., BAO, W., FANG, H., KAWASAKI, E. S., HAGER, J., TIKHONOVA, I. R., WALKER, S. J., ZHANG, L., HURBAN, P., DE LONGUEVILLE, F., FUSCOE, J. C., TONG, W., SHI, L. & WOLFINGER, R. D. (2006) Performance comparison of one-color and two-color platforms within the MicroArray Quality Control (MAQC) project. *Nat Biotechnol,* 24**,** 1140-50.

PAYS, E., HANOCQ-QUERTIER, J., HANOCQ, F., VAN ASSEL, S., NOLAN, D. & ROLIN, S. (1993) Abrupt RNA changes precede the first cell division during the differentiation of Trypanosoma brucei bloodstream forms into procyclic forms in vitro. *Mol Biochem Parasitol,* 61**,** 107-14.

PEARCE, L. R., HUANG, X., BOUDEAU, J., PAWLOWSKI, R., WULLSCHLEGER, S., DEAK, M., IBRAHIM, A. F., GOURLAY, R., MAGNUSON, M. A. & ALESSI, D. R. (2007) Identification of Protor as a novel Rictor-binding component of mTOR complex-2. *Biochem J,* 405**,** 513-22.

PESIN, J. A. & ORR-WEAVER, T. L. (2008) Regulation of APC/C activators in mitosis and meiosis. *Annu Rev Cell Dev Biol,* 24**,** 475-99.

PETERS, J. M. (2006) The anaphase promoting complex/cyclosome: a machine designed to destroy. *Nat Rev Mol Cell Biol,* 7**,** 644-56.

PFAFFL, M. W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res,* 29**,** e45.

PFGR TRYPANOSOMA BRUCEI VERSION 3. *PATHOGEN FUNCTIONAL GENOMICS RESOURCE CENTER*.

PFLEGER, C. M. & KIRSCHNER, M. W. (2000) The KEN box: an APC recognition signal distinct from the D box targeted by Cdh1. *Genes Dev,* 14**,** 655-65.

PHILLIPS, J. & EBERWINE, J. H. (1996) Antisense RNA Amplification: A Linear Amplification Method for Analyzing the mRNA Population from Single Living Cells. *Methods,* 10**,** 283-8.

PLOUBIDOU, A., ROBINSON, D. R., DOCHERTY, R. C., OGBADOYI, E. O. & GULL, K. (1999) Evidence for novel cell cycle checkpoints in trypanosomes: kinetoplast segregation and cytokinesis in the absence of mitosis. *J Cell Sci,* 112**,** 4641-4650.

PONTING, C. P., OLIVER, P. L. & REIK, W. (2009) Evolution and functions of long noncoding RNAs. *Cell,* 136**,** 629-41.

POWERS, R. W., 3RD, KAEBERLEIN, M., CALDWELL, S. D., KENNEDY, B. K. & FIELDS, S. (2006) Extension of chronological life span in yeast by decreased TOR pathway signaling. *Genes Dev,* 20**,** 174-84.

PRINZ, S., HWANG, E. S., VISINTIN, R. & AMON, A. (1998) The regulation of Cdc20 proteolysis reveals a role for APC components Cdc23 and Cdc27 during S phase and early mitosis. *Curr Biol,* 8**,** 750-60.

PRIOTTO, G., KASPARIAN, S., MUTOMBO, W., NGOUAMA, D., GHORASHIAN, S., ARNOLD, U., GHABRI, S., BAUDIN, E., BUARD, V., KAZADI-KYANZA, S., ILUNGA, M., MUTANGALA, W., POHLIG, G., SCHMID, C., KARUNAKARA, U., TORREELE, E. & KANDE, V. (2009) Nifurtimox-eflornithine combination therapy for second-stage African Trypanosoma brucei gambiense trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. *Lancet,* 374**,** 56-64.

QUEIROZ, R., BENZ, C., FELLENBERG, K., HOHEISEL, J. D. & CLAYTON, C. (2009) Transcriptome analysis of differentiating trypanosomes reveals the existance of multiple post-transcriptional regulons. *BMC Genomics,* 10.

RALSTON, K. S., LERNER, A. G., DIENER, D. R. & HILL, K. L. (2006) Flagellar motility contributes to cytokinesis in Trypanosoma brucei and is modulated by an evolutionarily conserved dynein regulatory system. *Eukaryot Cell,* 5**,** 696-711.

RAMOS, C. S., FRANCO, F. A., SMITH, D. F. & ULIANA, S. R. (2004) Characterisation of a new Leishmania META gene and genomic analysis of the META cluster. *FEMS Microbiol Lett,* 238**,** 213-9.

RAPE, M., REDDY, S. K. & KIRSCHNER, M. W. (2006) The processivity of multiubiquitination by the APC determines the order of substrate degradation. *Cell,* 124**,** 89-103.

RAUGHT, B., GINGRAS, A. C. & SONENBERG, N. (2001) The target of rapamycin (TOR) proteins. *Proc Natl Acad Sci U S A,* 98**,** 7037-44.

REDMOND, S., VADIVELU, J. & FIELD, M. C. (2003) RNAit: an automated web-based tool for the selection of RNAi targets in Trypanosoma brucei. *Mol Biochem Parasitol,* 128**,** 115-8.

REIS, A., LEVASSEUR, M., CHANG, H. Y., ELLIOTT, D. J. & JONES, K. T. (2006) The CRY box: a second APCcdh1-dependent degron in mammalian cdc20. *EMBO Rep,* 7**,** 1040-5.

REUNER, B., VASSELLA, E., YUTZY, B. & BOSHART, M. (1997) Cell density triggers slender to stumpy differentiation of Trypanosoma brucei bloodstream forms in culture. *Mol Biochem Parasitol,* 90**,** 269-80.

ROCHETTE, A., RAYMOND, F., CORBEIL, J., OUELLETTE, M. & PAPADOPOULOU, B. (2009) Whole-genome comparative RNA expression profiling of axenic and intracellular amastigote forms of Leishmania infantum. *Mol Biochem Parasitol,* 165**,** 32-47.

RODGERS, J. (2009) Human African trypanosomiasis, chemotherapy and CNS disease. *Journal of Neuroimmunology,* 211**,** 16.

RODITI, I., FURGER, A., RUEPP, S., SCH¸RCH, N. & B¸TIKOFER, P. (1998) Unravelling the procyclin coat of Trypanosoma brucei. *Molecular and Biochemical Parasitology,* 91**,** 117.

RODITI, I., SCHWARZ, H., PEARSON, T. W., BEECROFT, R. P., LIU, M. K., RICHARDSON, J. P., BUHRING, H. J., PLEISS, J., BULOW, R., WILLIAMS, R. O. & ET AL. (1989) Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of Trypanosoma brucei. *J Cell Biol,* 108**,** 737-46.

ROLIN, S., PAINDAVOINE, P., HANOCQ-QUERTIER, J., HANOCQ, F., CLAES, Y., LE RAY, D., OVERATH, P. & PAYS, E. (1993) Transient adenylate cyclase activation accompanies differentiation of Trypanosoma brucei from bloodstream to procyclic forms. *Mol Biochem Parasitol,* 61**,** 115-25.

ROSSI, A., PICA-MATTOCCIA, L., CIOLI, D. & KLINKERT, M. Q. (2002) Rapamycin insensitivity in Schistosoma mansoni is not due to FKBP12 functionality. *Mol Biochem Parasitol,* 125**,** 1-9.

ROZEN, S. & SKALETSKY, H. J. (2000) Primer3 on the WWW for general users and for biologist programmers. . *Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press***,** 365-386.

SABATINI, D. M., ERDJUMENT-BROMAGE, H., LUI, M., TEMPST, P. & SNYDER, S. H. (1994) RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion and is homologous to yeast TORs. *Cell,* 78**,** 35-43.

SARBASSOV, D. D., ALI, S. M., KIM, D. H., GUERTIN, D. A., LATEK, R. R., ERDJUMENT-BROMAGE, H., TEMPST, P. & SABATINI, D. M. (2004) Rictor, a novel binding partner of mTOR, defines a rapamycin-insensitive and raptor-independent pathway that regulates the cytoskeleton. *Curr Biol,* 14**,** 1296-302.

SAXENA, A., LAHAV, T., HOLLAND, N., AGGARWAL, G., ANUPAMA, A., HUANG, Y., VOLPIN, H., MYLER, P. J. & ZILBERSTEIN, D. (2007) Analysis of the Leishmania donovani transcriptome reveals an ordered

progression of transient and permanent changes in gene expression during differentiation. *Mol Biochem Parasitol,* 152**,** 53-65.

SBICEGO, S., VASSELLA, E., KURATH, U., BLUM, B. & RODITI, I. (1999) The use of transgenic Trypanosoma brucei to identify compounds inducing the differentiation of bloodstream forms to procyclic forms. *Molecular and Biochemical Parasitology,* 104**,** 311.

SCHWAB, M., LUTUM, A. S. & SEUFERT, W. (1997) Yeast Hct1 is a regulator of Clb2 cyclin proteolysis. *Cell,* 90**,** 683-93.

SCHWAB, M., NEUTZNER, M., MOCKER, D. & SEUFERT, W. (2001) Yeast Hct1 recognizes the mitotic cyclin Clb2 and other substrates of the ubiquitin ligase APC. *Embo J,* 20**,** 5165-75.

SCORY, S., STIERHOF, Y. D., CAFFREY, C. R. & STEVERDING, D. (2007) The cysteine proteinase inhibitor Z-Phe-Ala-CHN2 alters cell morphology and cell division activity of Trypanosoma brucei bloodstream forms in vivo. *Kinetoplastid Biol Dis,* 6**,** 2.

SEED, J. R. & WENCK, M. A. (2003) Role of the long slender to short stumpy transition in the life cycle of the african trypanosomes. *Kinetoplastid Biol Dis,* 2**,** 3.

SHAPIRO, S. Z., NAESSENS, J., LIESEGANG, B., MOLOO, S. K. & MAGONDU, J. (1984) Analysis by flow cytometry of DNA synthesis during the life cycle of African trypanosomes. *Acta Trop,* 41**,** 313-23.

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.

SIEGEL, T. N., HEKSTRA, D. R. & CROSS, G. A. (2008a) Analysis of the Trypanosoma brucei cell cycle by quantitative DAPI imaging. *Mol Biochem Parasitol,* 160**,** 171-4.

SIEGEL, T. N., HEKSTRA, D. R., KEMP, L. E., FIGUEIREDO, L. M., LOWELL, J. E., FENYO, D., WANG, X., DEWELL, S. & CROSS, G. A. (2009) Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei. *Genes Dev,* 23**,** 1063-76.

SIEGEL, T. N., KAWAHARA, T., DEGRASSE, J. A., JANZEN, C. J., HORN, D. & CROSS, G. A. (2008b) Acetylation of histone H4K4 is cell cycle regulated and mediated by HAT3 in Trypanosoma brucei. *Mol Microbiol,* 67**,** 762-71.

SIEMANN, D. W. & KENG, P. C. (1986) Cell cycle specific toxicity of the Hoechst 33342 stain in untreated or irradiated murine tumor cells. *Cancer Res,* 46**,** 3556-9.

SIGL, R., WANDKE, C., RAUCH, V., KIRK, J., HUNT, T. & GELEY, S. (2009) Loss of the mammalian APC/C activator FZR1 shortens G1 and lengthens S phase but has little effect on exit from mitosis. *J Cell Sci*.

SIMARRO, P. P., JANNIN, J. & CATTAND, P. (2008) Eliminating human African trypanosomiasis: where do we stand and what comes next? *PLoS Med,* 5**,** e55.

SKAAR, J. R. & PAGANO, M. (2008) Cdh1: a master G0/G1 regulator. *Nat Cell Biol,* 10**,** 755-7.

SMARGIASSO, N., GABELICA, V., DAMBLON, C., ROSU, F., DE PAUW, E., TEULADE-FICHOU, M. P., ROWE, J. A. & CLAESSENS, A. (2009) Putative DNA G-quadruplex formation within the promoters of Plasmodium falciparum var genes. *BMC Genomics,* 10**,** 362.

SMYTH, G. K. (2005) Bioinformatics and Computational Biology Solutions Using R and Bioconductor. *Statistics for Biology and Health.* Springer New York.

SMYTH, G. K. & SPEED, T. (2003) Normalization of cDNA microarray data. *Methods,* 31**,** 265-73.

STAN, R., MCLAUGHLIN, M. M., CAFFERKEY, R., JOHNSON, R. K., ROSENBERG, M. & LIVI, G. P. (1994) Interaction between FKBP12-rapamycin and TOR involves a conserved serine residue. *J Biol Chem,* 269**,** 32027-30.

STEKEL (2003) *Microarray Bioinformatics*, Cambridge University Press.

SUBRAMANIAM, C., VEAZEY, P., REDMOND, S., HAYES-SINCLAIR, J., CHAMBERS, E., CARRINGTON, M., GULL, K., MATTHEWS, K., HORN, D. & FIELD, M. C. (2006) Chromosome-Wide Analysis of Gene Function by RNA Interference in the African Trypanosome. *Eukaryotic Cell,* 5**,** 1539-1549.

SUTTERWALA, S. S., HSU, F. F., SEVOVA, E. S., SCHWARTZ, K. J., ZHANG, K., KEY, P., TURK, J., BEVERLEY, S. M. & BANGS, J. D. (2008) Developmentally regulated sphingolipid synthesis in African trypanosomes. *Mol Microbiol,* 70**,** 281-96.

SYKES, M. L. & AVERY, V. M. (2009) Development of an Alamar Blue viability assay in 384-well format for high throughput whole cell screening of Trypanosoma brucei brucei bloodstream form strain 427. *Am J Trop Med Hyg,* 81**,** 665-74.

SZOOR, B., WILSON, J., MCELHINNEY, H., TABERNERO, L. & MATTHEWS, K. R. (2006) Protein tyrosine phosphatase TbPTP1: A molecular switch controlling life cycle differentiation in trypanosomes. *J Cell Biol,* 175**,** 293-303.

TAKAHASHI, T., HARA, K., INOUE, H., KAWA, Y., TOKUNAGA, C., HIDAYAT, S., YOSHINO, K., KURODA, Y. & YONEZAWA, K. (2000) Carboxyl-terminal region conserved among phosphoinositide-kinase-related kinases is indispensable for mTOR function in vivo and in vitro. *Genes Cells,* 5**,** 765-75.

TARAYRE, S., VINARDELL, J. M., CEBOLLA, A., KONDOROSI, A. & KONDOROSI, E. (2004) Two classes of the CDh1-type activators of the anaphase-promoting complex in plants: novel functional domains and distinct regulation. *Plant Cell,* 16**,** 422-34.

TETLEY, L., TURNER, C. M., BARRY, J. D., CROWE, J. S. & VICKERMAN, K. (1987) Onset of expression of the variant surface glycoproteins of

Trypanosoma brucei in the tsetse fly studied using immunoelectron microscopy. *J Cell Sci,* 87 ( Pt 2)**,** 363-72.

THOMAS-CHOLLIER, M., SAND, O., TURATSINZE, J. V., JANKY, R., DEFRANCE, M., VERVISCH, E., BROHEE, S. & VAN HELDEN, J. (2008) RSAT: regulatory sequence analysis tools. *Nucleic Acids Res,* 36**,** W119-27.

TIMMS, M. W., VAN DEURSEN, F. J., HENDRIKS, E. F. & MATTHEWS, K. R. (2002) Mitochondrial development during life cycle differentiation of African trypanosomes: evidence for a kinetoplast-dependent differentiation control point. *Mol Biol Cell,* 13**,** 3747-59.

TSANG, C. K., LI, H. & ZHENG, X. S. (2007) Nutrient starvation promotes condensin loading to maintain rDNA stability. *Embo J,* 26**,** 448-58.

TYLER, K. M. (2003) Maintainence of parasitaemia - is it to die for? *Kinetoplastid Biol Dis,* 2**,** 2.

URBANI, L., SHERWOOD, S. W. & SCHIMKE, R. T. (1995) Dissociation of nuclear and cytoplasmic cell cycle progression by drugs employed in cell synchronization. *Exp Cell Res,* 219**,** 159-68.

URWYLER, S., STUDER, E., RENGGLI, C. K. & RODITI, I. (2007) A family of stage-specific alanine-rich proteins on the surface of epimastigote forms of Trypanosoma brucei. *Mol Microbiol,* 63**,** 218-28.

VAN DEN ABBEELE, J., CLAES, Y., VAN BOCKSTAELE, D., LE RAY, D. & COOSEMANS, M. (1999) Trypanosoma brucei spp. development in the tsetse fly: characterization of the post-mesocyclic stages in the foregut and proboscis. *Parasitology,* 118 ( Pt 5)**,** 469-78.

VAN MEIRVENNE, N., JANSSENS, P. G. & MAGNUS, E. (1975) Antigenic variation in syringe passaged populations of Trypanosoma (Trypanosoon) brucei. I. Raionalization of the experimetnal approach. *Ann Soc Belg Med Trop,* 55**,** 1-23.

VANHAMME, L. & PAYS, E. (1995) Control of gene expression in trypanosomes. *Microbiol Rev,* 59**,** 223-40.

VASSELLA, E., DEN ABBEELE, J. V., BUTIKOFER, P., RENGGLI, C. K., FURGER, A., BRUN, R. & RODITI, I. (2000) A major surface glycoprotein of trypanosoma brucei is expressed transiently during development and can be regulated post-transcriptionally by glycerol or hypoxia. *Genes Dev,* 14**,** 615-26.

VASSELLA, E., KRAMER, R., TURNER, C. M., WANKELL, M., MODES, C., VAN DEN BOGAARD, M. & BOSHART, M. (2001) Deletion of a novel protein kinase with PX and FYVE-related domains increases the rate of differentiation of Trypanosoma brucei. *Mol Microbiol,* 41**,** 33-46.

VASSELLA, E., OBERLE, M., URWYLER, S., RENGGLI, C. K., STUDER, E., HEMPHILL, A., FRAGOSO, C., BUTIKOFER, P., BRUN, R. & RODITI, I. (2009) Major surface glycoproteins of insect forms of Trypanosoma brucei are not essential for cyclical transmission by tsetse. *PLoS One,* 4**,** e4493.

VASSELLA, E., REUNER, B., YUTZY, B. & BOSHART, M. (1997) Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway. *J Cell Sci,* 110 ( Pt 21)**,** 2661-71.

VAUGHAN, S. & GULL, K. (2008) The structural mechanics of cell division in Trypanosoma brucei. *Biochem Soc Trans,* 36**,** 421-4.

VELLAI, T., TAKACS-VELLAI, K., ZHANG, Y., KOVACS, A. L., OROSZ, L. & MULLER, F. (2003) Genetics: influence of TOR kinase on lifespan in C. elegans. *Nature,* 426**,** 620.

VICKERMAN, K. (1965) Polymorphism and mitochondrial activity in sleeping sickness trypanosomes. *Nature,* 208**,** 762-6.

VICKERMAN, K. (1985) Developmental cycles and biology of pathogenic trypanosomes. *Br Med Bull,* 41**,** 105-14.

VICKERMAN, K., TETLEY, L., HENDRY, K. A. & TURNER, C. M. (1988) Biology of African trypanosomes in the tsetse fly. *Biol Cell,* 64**,** 109-19.

VILELLA-BACH, M., NUZZI, P., FANG, Y. & CHEN, J. (1999) The FKBP12-rapamycin-binding domain is required for FKBP12-rapamycin-associated protein kinase activity and G1 progression. *J Biol Chem,* 274**,** 4266-72.

VISINTIN, R., PRINZ, S. & AMON, A. (1997) CDC20 and CDH1: a family of substrate-specific activators of APC-dependent proteolysis. *Science,* 278**,** 460-3.

WALRAD, P., PATEROU, A., ACOSTA-SERRANO, A. & MATTHEWS, K. R. (2009) Differential trypanosome surface coat regulation by a CCCH protein that co-associates with procyclin mRNA cis-elements. *PLoS Pathog,* 5**,** e1000317.

WAN, Y. & KIRSCHNER, M. W. (2001) Identification of multiple CDH1 homologues in vertebrates conferring different substrate specificities. *Proc Natl Acad Sci U S A,* 98**,** 13066-71.

WANG, J., WU, Y., YANG, G. & AKSOY, S. (2009) Interactions between mutualist Wigglesworthia and tsetse peptidoglycan recognition protein (PGRP-LB) influence trypanosome transmission. *Proc Natl Acad Sci U S A,* 106**,** 12133-8.

WEI, H., KUAN, P. F., TIAN, S., YANG, C., NIE, J., SENGUPTA, S., RUOTTI, V., JONSDOTTIR, G. A., KELES, S., THOMSON, J. A. & STEWART, R. (2008) A study of the relationships between oligonucleotide properties and hybridization signal intensities from NimbleGen microarray datasets. *Nucleic Acids Res,* 36**,** 2926-38.

WEI, Y., TSANG, C. K. & ZHENG, X. F. (2009) Mechanisms of regulation of RNA polymerase III-dependent transcription by TORC1. *Embo J,* 28**,** 2220-2230.

WELBURN, S. C. & MAUDLIN, I. (1999) Tsetse-trypanosome interactions: rites of passage. *Parasitol Today,* 15**,** 399-403.

WILHELM, B. T., MARGUERAT, S., WATT, S., SCHUBERT, F., WOOD, V., GOODHEAD, I., PENKETT, C. J., ROGERS, J. & BAHLER, J. (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature,* 453**,** 1239-43.

WILUSZ, J. E., SUNWOO, H. & SPECTOR, D. L. (2009) Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev,* 23**,** 1494-504.

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.

WOODWARD, R. & GULL, K. (1990) Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of Trypanosoma brucei. *J Cell Sci,* 95 ( Pt 1)**,** 49-57.

WURMBACH, E., YUEN, T. & SEALFON, S. C. (2003) Focused microarray analysis. *Methods,* 31**,** 306-16.

YAMAGUCHI, S., MURAKAMI, H. & OKAYAMA, H. (1997) A WD repeat protein controls the cell cycle and differentiation by negatively regulating Cdc2/B-type cyclin complexes. *Mol Biol Cell,* 8**,** 2475-86.

YAMAGUCHI, S., OKAYAMA, H. & NURSE, P. (2000) Fission yeast Fizzy-related protein srw1p is a G(1)-specific promoter of mitotic cyclin B degradation. *EMBO J,* 19**,** 3968-77.

YANG, H., HENNING, D. & VALDEZ, B. C. (2005) Functional interaction between RNA helicase II/Gu(alpha) and ribosomal protein L4. *Febs J,* 272**,** 3788-802.

ZENG, F., PERITZ, T., KANNANAYAKAL, T. J., KILK, K., EIRIKSDOTTIR, E., LANGEL, U. & EBERWINE, J. (2006) A protocol for PAIR: PNA-assisted identification of RNA binding proteins in living cells. *Nat Protoc,* 1**,** 920-7.

ZHENG, X. F., FLORENTINO, D., CHEN, J., CRABTREE, G. R. & SCHREIBER, S. L. (1995) TOR kinase domains are required for two distinct functions, only one of which is inhibited by rapamycin. *Cell,* 82**,** 121-30.

ZHOU, Y., CHING, Y. P., CHUN, A. C. & JIN, D. Y. (2003a) Nuclear localization of the cell cycle regulator CDH1 and its regulation by phosphorylation. *J Biol Chem,* 278**,** 12530-6.

ZHOU, Y., CHING, Y. P., NG, R. W. & JIN, D. Y. (2003b) Differential expression, localization and activity of two alternatively spliced isoforms of human APC regulator CDH1. *Biochem J,* 374**,** 349-58.

ZICKERT, P., WEJDE, J., SKOG, S., ZETTERBERG, A. & LARSSON, O. (1993) Growth-regulatory properties of G1 cells synchronized by centrifugal elutriation. *Exp Cell Res,* 207**,** 115-21.

ZIEGELBAUER, K. & OVERATH, P. (1992) Identification of invariant surface glycoproteins in the bloodstream stage of Trypanosoma brucei. *J Biol Chem,* 267**,** 10791-6.

ZIEGELBAUER, K., QUINTEN, M., SCHWARZ, H., PEARSON, T. W. & OVERATH, P. (1990) Synchronous differentiation of Trypanosoma brucei from bloodstream to procyclic forms in vitro. *Eur J Biochem,* 192**,** 373-8.

ZIELINSKI, J., KILK, K., PERITZ, T., KANNANAYAKAL, T., MIYASHIRO, K. Y., EIRIKSDOTTIR, E., JOCHEMS, J., LANGEL, U. & EBERWINE, J. (2006) In vivo identification of ribonucleoprotein-RNA interactions. *Proc Natl Acad Sci U S A,* 103**,** 1557-62.

# Appendix A – Primer sequences

## *Primers used in qPCR*

| Gene ID | Forward primer | Reverse primer | PCR efficiency |
|---|---|---|---|
| Tb10.26.0220 | TCCAACCGATAACACGACAG | CGATATGACCGACACGTCAC | 90.8% |
| Tb927.7.7160 | AGGCATCCATCGAGTACAGC | TATCCTTCCGCAACACCTTC | 98.2% |
| Tb10.6k15.3640 | ACGGCCTCGTTGATACACTC | CAACATTCCACCGACCATC | 100.4% |
| Tb927.3.2710 | GCAAGTCCATCACACAGGAG | GAAGAGGCTACGGACACACC | 95.3% |
| Tb927.1.4450 | AGCAGCAGGTTATGGTGGAG | ATACAACGATTCCGGTGAGC | 91.3% |
| Tb10.389.0540 | CCAGCCTTCTCAATCTCCAG | GGCCACAGTTGGATAGCTTG | 98.5% |
| Tb10.70.6180 | TCGAGGCAGGTTGAGTAAGG | CAGAGGATGAAGCGAAGTCC | 96.3% |
| Tb927.5.3400 | TCTGTGACATGCGAAGGAAG | CTTCGCCAATCTTCTCAACC | 97.1% |
| Tb10.389.0650 | ACGCTAGCACAACCAGAAGC | GACCGACCAGGTCTTCTACG | 100.8% |
| Tb927.8.1270 | AGAGGACGTGGAAGATGACG | GCAACTTACCAGCCTTCTGC | 99.1% |
| Tb927.6.4180 | GGCCGAAGGTTACACAAGAG | CGATCTCCACCTTCTTCCAG | 103.7% |
| Tb927.7.3170 | AAGATGACAGGAACGCGAAG | AGGTCACGAAGGTCATCCAC | 103.7% |
| Tb10.61.3040 | CATCGTCGCCAACATAATTG | GCACACAATTCCACCAACAG | 106.1% |
| Tb10.6k15.0300 | ACGAACTGTGCCGTAGGAAG | ACTCGCACGTAGTGGTGAAG | 91.2% |
| Tb09.160.2110 | ATGCTAACAACGGCGAAGG | CTCTCGGCTAATCACATCAGG | 92.5% |

### *Cloning primers*

| Experiment | Primer name | Sequence |
|---|---|---|
| RNAi TbTor1 | Tor_F | ccggatccgCTACTCAACCTGTCGGAGC |
| | Tor_R | ccctcgagACTGCCTCCACATTTCCAAC |
| RNAi TbTor1 (Barquilla 2008) | Tor_Barquilla F | ggatccCGGATTGTTGCGGTGAAAAC |
| | Tor_Barquilla R | ctcgagATGCTTTATGCTTGCCTTCC |
| RNAi TbTor2 | Tor_B_F | ccggatccgGTCGGACTTCGTTCAAGAGC |
| | Tor_B_R | ccctcgagAAACCAACCGTGTCCTTCAG |
| RNAi TbTor2 (Barquilla 2008) | Tor2_Barquilla F | ggatccTACGATGCGCGGAAAGTGGAGG |
| | Tor2_Barquilla R | ctcgagAGTCTACCGTTTTCTTCCTGTT |
| RNAi Tb10.26.0220 | Tb10.26.0220 F | ACAGTGATCCCCGTTACTGC |
| | Tb10.26.0220 R | TTGTGCCACCCACAAAACTA |
| RNAi Cdh1 | Chd1 F | ACTCGCTGGGTGTGGTATTC |
| | Cdh1 R | CAAACCACACACACTTTGCC |

# Appendix B - Solutions

**YPD**

| | |
|---|---|
| Bacto-yeast extract | 1% w/v |
| Bacto-peptone | 2% w/v |
| Dextrose | 2% w/v |

Autoclave.

For plates, 2% agar was added to the medium prior to autoclaving, and plates were stored at 4˚C.

## *Northern blot solutions*

[10x] MOPS

| | |
|---|---|
| MOPS | 46.26 g/l |
| Sodium-Acetate (pH 7.0) | 50 mM |
| EDTA | 10 mM |

Autoclave. Solution stored in the dark at 4 °C.

**RNA Gel Loading Buffer**

| | |
|---|---|
| Formamide | 30 % |
| Formaldehyde | 6.1 % |
| Bromophenol blue | 0.01 % |
| MOPS | 1 × |
| glycerol | 10 % |

**Maleic Acid Buffer**

| | |
|---|---|
| Maleic acid | 100mM |
| NaCl | 150mM |

pH7.5 with NaOH, autoclaved

**10% Block Solution**

| | |
|---|---|
| Blocking reagent | 10 % (w/v) |
| Maleic Acid Buffer | 1 × |

**Wash Buffer**

| | |
|---|---|
| Maleic Acid Buffer | 1 × |
| Tween 20 | 0.3 % |
| **Detection Buffer** | |
| Tris-HCl | 100 mM |
| NaCl | 100 mM |
| pH 9.5 | |

## *Western blot solutions*

| | |
|---|---|
| **[10x] Running Buffer** | |
| Tris-HCl, pH8.3 | 0.25 M |
| Glycine | 1.92 M |
| SDS | 1% |
| **Laemmli Sample Buffer** | |
| Tris-HCl, pH6.8 | 62.5 mM |
| SDS | 2 % |
| Glycerol | 10 % |
| Bromophenol blue | 20 μg |
| **[10x] Blot Transfer Buffer Stock** | |
| Tris (do not pH) | 0.025 M |
| Glycine | 0.15 M |
| **1× Completed Transfer Buffer** | |
| SDS | 0.2 % |
| [10x] Transfer Buffer Stock | 10 % |
| Methanol | 20 % |
| **Ponceau Stain** | |
| Ponceau S | 0.4% |
| TCA | 3% |

**DAPI working stock**

| | |
|---|---|
| 4',6-diamidino-2-phenylindole (DAPI) | 10 μg ml-1 |

**MOWIOL Mounting Medium**

| | |
|---|---|
| Glycerol | 25 % w/v |
| MOWIOL | 10 % w/v |
| Tris pH8.5 | 0.1 M |

MOWIOL dissolved through incubation at 50˚C with regular stirring. Clarified by centrifuging at 5000xg for 15 minutes. Aliquots were stored at -20°C.

**Phosphate Buffer Solution (PBS)**

| | |
|---|---|
| NaCl | 137 mM |
| KCl | 3 mM |
| Na2HPO4 | 16 mM |
| KH2PO4 | 3 mM |

pH7.6

## *Small scale plasmid preparation solutions*

**Solution I**

| | |
|---|---|
| glucose | 50 mM |
| EDTA | 10 mM |
| Tris-HC1 (pH 8.0) | 25 mM |

**Solution II**

| | |
|---|---|
| NaOH | 0.2 M |
| Sodium dodecyl sulfate (SDS). | 1% |

**Solution III**

| | |
|---|---|
| Potassium acetate | 3 M |
| Glacial acetic acid | 8.7% |

pH 5.2

**TAE 1×**

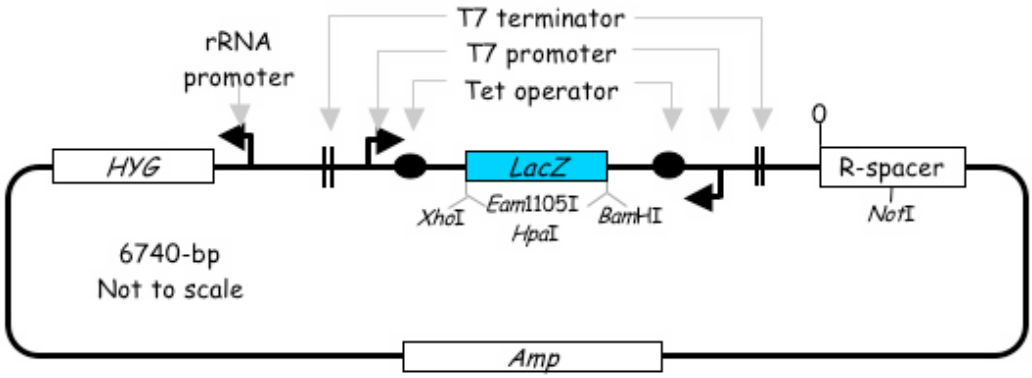
| | |
|---|---|
| Tris-Acetate | 40 mM |
| EDTA | 1 mM |
| pH 8.0 | |

**ZPFMG**

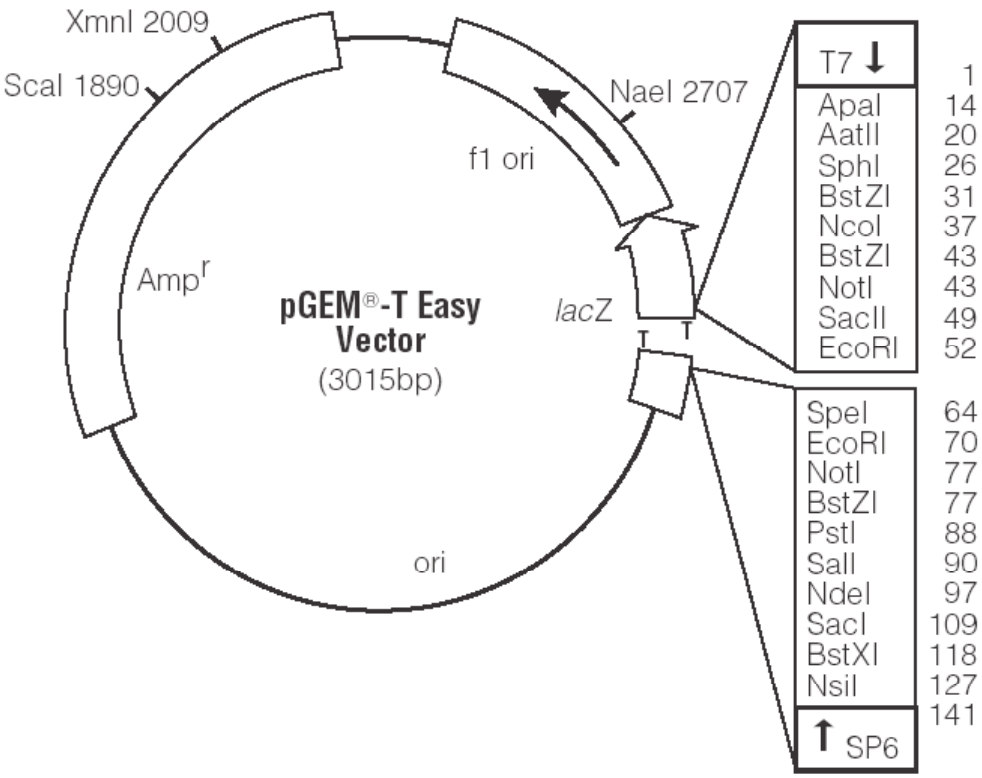
| | |
|---|---|
| NaCl | 132mM |
| KCl | 8mM |
| $Na_2HPO_4$ | 8mM |
| $KH2PO_4$ | 1.5mM |
| $MgOAC.4H_20$ | 0.5mM |
| $CaOAc/Cl_2$ | 90µM |
| Glucose | 0.5% |

# Appendix C – Vector maps

### *p2T7TABlue*




### *pGem-T-easy*



## Appendix D - Publications

Two publications have so far been generated from the work presented in this thesis. These are inserted here:

KABANI, S., FENN, K., ROSS, A., IVENS, A., SMITH, T. K., GHAZAL, P. & MATTHEWS, K. (2009) Genome-wide expression profiling of in vivo-derived bloodstream parasite stages and dynamic analysis of mRNA alterations during synchronous differentiation in Trypanosoma brucei. BMC Genomics 10(1): 427.

KABANI, S., WATERFALL, M. & MATTHEWS, K. R. (2009) Cell-cycle synchronisation of bloodstream forms of Trypanosoma brucei using VybrantDyeCycle Violet-based sorting. Mol Biochem Parasitol 169(1): 59-62.