A rational quest for drug targets in the protein kinome of *Trypanosoma brucei*

Fernando Fernandez-Cortes

MPharm, MSc

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

Wellcome Centre for Molecular Parasitology
Institute of Infection, Immunity and Inflammation
College of Medical, Veterinary and Life Sciences

University of Glasgow

December 2016

Abstract

Trypanosoma brucei is the protozoan parasite causing African trypanosomiasis, a neurological disease that affects humans and farm-stock in the impoverished sub-Saharan areas where tsetse fly transmission vector is endemic. Although it has great impact on public health and local economies, it has been neglected in drug discovery for almost a century. Current treatments are either toxic or of difficult administration, besides having serious risks of inducing resistance.

Protein kinases are the primary set of signaling proteins in eukaryotes, including *Trypanosoma brucei*. Their druggability has been widely exploited in cancer research, and has been established in the parasite too. A recent kinome-wide RNAi screen with 176 individual cell lines of mammalian infective bloodstream forms of *Trypanosoma brucei* identified protein kinases required for proliferation *in vitro*. In order to investigate which protein kinases are also essential virulence factors *in vivo,* lines were pooled, inoculated into mice and screened for loss of fitness after 48 hours RNAi compared to uninduced controls. The presence of trypanosomes in the bloodstream was assessed using RNAi target sequencing (RITseq) and compared to an *in vitro* control. This revealed 49 protein kinases with a significant loss of fitness *in vivo* in two independent experiments, and a strong correlation between *in vitro* and *in vivo* loss of fitness for the majority. However, depletion of nine protein kinases affected more pronouncedly the growth *in vivo* than *in vitro*. Amongst these protein kinases were several with putative functions related with stress responses mediated through the PI3K/TOR or MAPK signaling cascades including CK2A2, a promiscuous protein kinase whose activity can be stress-induced; two MAP3Ks, involved in cell integrity upon osmotic shock; VPS15, component of the PI3K complex with roles in autophagosome formation and vesicular trafficking; BUD32, transducer of the PI3K/TOR pathway involved in translational regulation; and FAZ20, a parasite-specific pseudo-kinase localizing to the flagellum attachment zone. The other three have been implicated in repair of alkylation-induced cellular damage: SRPK1, a stress response RNA splicing regulator; AUK2, which acts during mitosis; and CAMKL, an AMPK with calcium-binding domains putatively involved in metabolic regulation. Identification of these virulence-associated protein kinases

provides new insights in *T. brucei*-host interaction and reveals novel potential drug targets for protein kinase inhibitors.

This RNAi screen revealed that the evolutionary divergent NEK kinase Repressor of Differentiation Kinase 2 (RDK2) has severe loss of fitness both *in vitro* and *in vivo*. Depletion of RDK2 had been shown previously to promote differentiation from bloodstream to procyclic-like forms causing the parasite's death. Further investigation showed RDK2 to be an active protein kinase capable both of phosphorylating a substrate and to autophosphorylate. Protein kinase activity could be ablated by mutation of lysine 70 to methionine. Mutation of both serine residues (195 and 197), identified as sites of phosphorylation by phosphoproteomics, to alanine or glutamic acid, preventing and mimicking phosphorylation respectively, had no effect on protein kinase activity, suggesting they do not have a direct regulatory role on protein kinase activity. Introducing in the RNAi line a recoded *RDK2* whose transcript eluded interference, permitted to some extent the rescue of the induced phenotype, while introducing a recoded inactive mutant did not. This may suggest that the lack of kinase activity was responsible for the RNAi phenotype and not depletion of the protein alone. RDK2 RNAi-differentiated cells could be maintained in conditioned procyclic form media for more than a week. However, they were unable to proliferate. Overexpression of RDK2 blocked the differentiation mediated by sequential treatment with 8-pCPT-cAMP and citrate/cis-aconitate (CCA). RNAi experiments in combination with known differentiation cues, suggested that when differentiation is triggered by the CCA signalling pathway, RDK2 inactivation happens downstream of the phosphatase TbPTP1.

Differentiation caused by RDK2 inactivation could be traced in flow cytometry by the detection of EP procyclin expression. This was exploited in a cell-based mechanism-directed phenotypic screen for RDK2 inhibitors. A preliminary run with 518 drug-like molecules that had shown protein kinase inhibition, trypanocidal activity and/or activation of the EP procyclin promoter, unveiled 6 compounds triggering EP procyclin expression and parasite death in the low micromolar range. These compounds can be investigated further to assess whether RDK2 is their *in vivo* target.

Table of Contents

List of Tables

List of Figures

List of Accompanying Material

The attached DVD contains the following supplementary information:

Figure S1: Extract of a T-Coffee alignment with the complete protein kinome of *Trypansoma brucei*. In the first slide, a copy of Figure 1-10 is included. Described in Chapters 3 and 4.

Movie S1 and Movie S2: Microscopic images of RDK2 RNAi line. Live cells recovered 48 h post-induction and kept in SDM79 conditioned at 27^{o}C with tetracycline for 96 h. Described in Chapter 4.

Table S1: Content of each MSTL sub librarie included in the kinome-wide RITseq experiment. Described in chapter 3.

Table S2: Complete RITseq *in vivo* dataset.

Table S3: Complete RITseq *in vitro* dataset

Table_S4: T tetsts comparing the normalized reads induced and uninduced at each of the 5 time points of the *in vitro* RITseq (tabs 1-5), and of the two replicates of the *in vivo* RITseq. P values and induced/uninduced RNAi ratio after bootstrapping with 1000 replicates are provided. Described in chapter 3.

Table_S5: Fourteen barcoded oligos used for PCR enrichment at the *in vitro* RITseq, and at the second *in vivo* RITseq, Highlighting samples where they were used, and Illumina multiplexed samples where they were pooled. Described in chapter 3.

Table_S6: Results of the 10 μM primary screen looking for potential RDK2 inhibitors for the eight 96-well plates analysed (Tabs 1-8). Described in chapter 6.

Table_S7: Validation dose–response curve results for plate A (containing compounds DIF1-5 and DIF 10-12). It provides data after 24h and 48h of incubation (n=2). Including summary sheets. Described in chapter 6.

Table_S8: Validation dose –response curve results for plate B. (containing compounds DIF6-9 and DIF 13-16). It provides data after 24h and 48h of incubation (n=2). Including summary sheets. Described in chapter 6.

Acknowledgement

First of all I would like to thank Jeremy Mottram for his supervision over the last 4 years, for giving me the opportunity of working on this project and respecting and encouraging my ideas and interests (no matter how many Pharma companies or research labs I wanted to collaborate with). He knows how to provide his students with right resources. I am especially grateful for his help during the writing stage. It must be painful correcting spelling mistakes while moving a whole lab to a 'different country'.

I also need to thank, Richard McCulloch, who was available for discussions while the boss was busy. He has shown me that you can do science and still be a cool, enthusiastic person.

I cannot forget my assessors, Sylke Müller and Mike Barrett, for excellent ideas and productive discussions. Sylke coordinated Paramet, the EU-funded Marie Curie international training network that made this PhD project possible. It has been a real life-changing experience. I feel privileged for being a part of this network and attending the annual meetings, where we had the chance to discuss our projects with some of the best experts in the field in a very close environment. I really appreciate their open-minded training approach, where we learnt everything from clinical parasitology to *in silico* modelling and high throughput screening, as well as their efforts to deliver high quality communications and entrepreneurship workshops. I loved the way they encouraged our public engagement with such creative forms of outreach, as well as their support for the symposium I organised in Seville. I had many unforgettable moments with the Paramet crew, and I am pretty sure the fun is not over… as I am sure we will have the chance of collaborating in the future.

Mike Barrett deserves a special mention for hiring me knowing that I was still writing. I hope to give him some good results for what he gave me in patience and trust.

Special thanks to many people that helped making this project a reality, starting with Nath Jones who was a pioneer in developing the RNAi library, and showing me everything about culturing trypanosomes. Life in a lab would just be non-sense without people like Elaine Brown and Jim Scott working their magic, 1000 times thank you for existing. I will be up for as many boozy lunches as you fancy! Thank you to John Wilkes, who mapped the sequencing reads, he was another key contributor to the success of my research. Nick Dickens was another great idea-generator. My sporadic trips to the animal house were a pleasure because Ryan Ritchie was always there to make things easy. And Alan Scott was always ready to help purify a protein. The flow cytometry queen, Diane Vaughan, has been another fundamental factor in keeping the equipment always up and running. Finally, it was fun to collaborate with Eric Kalkman and Susan Baillie in the Scottish Bioscreening Facility, thank all you people.

I also need to thank my collaborators at GSK, especially Julio Martin and Imanol Peña, who were my drug dealers together with the Swiss Tropical crew, Tanja Wenzler, Pascal Maser, Reto Brun and Isabel Roditi. Julio was great at brain storming, and very kind to me, thanks indeed.

For a couple of months (or more…) putting all these parasites together would have been a real nightmare without my mate Tiago Serafim, who -with Ana Adrados, Elena Jimenez and Mari Serpeloni- was always ready for a late night pint… Late-night shift forever… I would not know how to work in a 9-18h time-frame. You guys were very important for this PhD and for my life as a whole. Wherever you go, be ready to receive squatters.

Another great contribution has been Helena de la Torre. She started as my student, helping in developing my overexpressor lines, and ended up becoming someone I hope to count among my friends for many years to come.

I send a general thank you to the level 6 at the Wellcome Centre for Molecular Parasitology. Not the place but the people. They have shared with me so much time and fun along these 4 years, we have burned Glasgow and discovered Scotland so often that I would need another book to put all my memories down. Special thanks for the great moments to the squashies, Sam and Dan, who are two damn good friends I hope to maintain forever. I am glad we could make that

pre-fatherhood weekend (or mid-week) at Czech. Amy, Cat, Andrea, Marko, Ben, Vivi, Antonio… all my mates that used to inhabit the end office. A special one to my flatmate, Pete, who showed me the Dutch way. I really enjoyed all those jamming nights, speaking cool and stupid stuff at home while respecting our "no-Science-chat-except-if-exciting-crap" rule. And finally to all the official and extra-official members of both "the wee family" and "the conspiracy club", you all helped making rainy Glasgow 'home'.

Now it is time to go back to the origins, for we must always remember who we are and why we are. I thank all the crazy guys from the school of Pharmacy in Seville that shared the dream of doing things with impact, hoping one day they may change the world: Mati, Chema, Pili, Tito, Marta… I know everyone is making wise movements in the right direction. Special thanks to Miguel Angel Caviedes, that beyond teaching me Microbiology, made me keen in exploring where research could take me. He is the closest thing I have ever had to an inspiring mentor in the academic world, and I will never forget. As I will never forget my first lab experience, exploring options beyond selling pills over the counter with my "brother" Juan, in Granada the most eye-opening experience.

Then I need to thank Michael Gaunt, Michael Miles and Martin Llewellyn, for providing me with my first international lab experience when I was a total amateur scientist. And to Manuel Fresno for hosting me at the great 226 lab in Madrid, where I learnt so much about parasitology under the supervision of Nuria Girones, and the great, genuine and unique Julien Santi-Rocca. All this path has been long and tortuous but awesome. I am glad, after all that mail exchange, Leonel Messi -as good at football as at tax evasion- did not pay me a PhD because I would have never made it to Glasgow.

But speaking about origins I need to thank my teachers at the primary school of my little Calzadilla de los Barros, for encouraging my curiosity. In especial to Angel Granero, who taught me the language I am using to write this thesis. And overall thanks to my parents, my grandma, and my brother who enjoyed this PhD from the distance as much as I enjoyed it from here. They are the etiologic agent of the way I think. They convinced me that every problem has some sort of solution if you think hard enough, and that life may have limits but I certainly should not be one to myself. I love you and this thesis is yours as it is mine.

Finally, if I ever graduate, the “P” and “h” of my PhD will be Ana’s, because although she does not give a damn about science she has started having a respect for it, and because she keeps my feet on the ground so I do not fall in the hamster wheel. For an insane number of reasons -and because I love you- this PhD thesis is also yours.

Fernando

This work was supported by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement n° 290080.

Author's Declaration

I here declare that this thesis and the results herein presented are a result of my own work, except where otherwise stated and acknowledged. None of the results herein presented have been used previously to obtain a degree at any university.

Fernando Fernandez-Cortes

Abbreviations

8-pCPT-cAMP: 8-(4-Chlorophenylthio)adenosine 3′,5′-cyclic monophosphate
AAT: Animal African Trypanosomiasis
Ala: Alanine
AMPK: AMP-activated protein kinase
aPKs: atypical protein kinases
Arg: Arginine
Asn: Asparagine
Asp: Aspartic acid
ATP: Adenosine triphosphate
AUK: aurora kinase
BARP: brucei alanine-rich protein
BBB: Blood Brain Barrier
BCR-ABL: breakpoint cluster region protein - Abelson murine leukemia viral oncogene homolog 1
BLASTp: Basic Local Alignment Search Tool for proteins
bp: base pairs
BSF: bloodstream forms
CAMK: Ca2+/calmodulin-dependent protein kinase
cAMP: cyclic adenosine monophosphate
CATT: card agglutination test for trypanosomiasis
CBPK: Calcium-binding protein K
CCA: cis-aconitate
CDK1: Cyclin-dependent kinase
CDS: Coding Sequence
CK1: Casein kinase 1
CLK: CDC2-like kinase
CMGC: family containing containing CDK, MAPK, GSK3 and CLK kinases
CNS: Central Nervous System
CRK: CDC2-Related Kinase
CRKs: CDC2-Related kinases
Cys: Cystein
DNA: Deoxyribonucleic acid
DNDi : Drugs for Neglected Diseases Initiative
DOT1B: disruptor of telomeric silencing-1B
DTM: differentiation trypanosome media
DYRKs: Dual Specificity YAK1-Related Kinases
EC50: Effective concentration 50%
EGFR: Epidermal growth factor receptor
ePKs: Eukaryotic protein kinases
ERK: Extracellular signal–regulated kinases
ES: expression sites
ESAG: expression site-associated genes
FAZ: flagellar attachment zone
FDA: Food and Drug Administration
FHA: forkhead-associated domain
FPLC: fast protein liquid chromatography
gDNA: genomic DNA

Glu: Glucose
Gly: Glycine
GPI: glycophosphatidyl inositol
GSK3: Glycogen synthase kinase 3
GTP: Guanosine triphosphate
HAT: Human African Trypanosomiasis
His: histidine
HMI11: Hirumi modify Iscove medium
HMM: hidden markov model
HpHgR: haptoglobin-haemoglobin receptor
IC50: Inhibition concentration 50%
ICAM:Intercellular Adhesion Molecule 1
IFA: Immunofluorescence assay
IgG: Immunoglobulin G
IgM: Immunoglobulin M
IMAC: immobilized metal affinity chromatography
ISG: invariant surface glycoproteins
ISWI: transcriptional-silencing Imitation SWItch
Kb: kilobase
Kda: Kilodalton
kDNA: kinetoplastid DNA
KKT: kinetoplastid kinetochore protein
Leu: Leucine
LIC: Ligation Independent Cloning
Lys: Lysine
MAPK: mitogen-activated protein kinase
Mb: megabase
MC: metal chromatography
Met: Methionine
MSB: metalloprotease
MKK1: Mitogen-activated protein kinase kinase
mRNA: messenger RNA
NCBI : National Center for Biotechnology Information
nDNA: nuclear DNA
NECT: Nifurtimox-Eflornithine Combined Therapy
NEK: NIMA-related kinase
P-site: phosphosite
PAD: proteins associated with differentiation
PATTEC: Pan African Tsetse and Trypanosomiasis Eradication Campaign
PBS: phosphate buffered saline
PBST:phosphate buffered saline tween
PCF: procyclic forms
PFAM: Protein family
PGDFR: Platelet-derived growth factor
PH: Pleckstrin Homology
Phe: Phenilalanine
PI3K : Phosphoinositide 3-kinase
PIP39: PTP1-interacting protein, 39 kDa
PK: Protein kinase

PKA: protein kinase A
PKC: Protein kinase C
PLC: Phospolipase C
PLK: Polo-like kinase
Pol: polymerase
Pro: Proline
PTB: phosphotyrosine-binding
PTU: polycistronic transcription units
PVDF: Polyvinylidene difluoride
RBP: RNA binding proteins
RDK: Repression of differentiation kinase
RISC: RNA-induced silencing complex
RNA: Ribonucleic acid
RITseq: RNAi target sequencing
RNAi: RNA interference
RNAit: RNAi target
RPC: receptor guanylate cyclases
rRNA: ribosomal RNA
RTKs: receptor tyrosine kinases
SAR: structure-activity relationship
SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser: Serine
sgRNA: small guide RNA
SH2: Src homology 2
SIF: Stumpy Induction Factor
SMART: Similar Modular Architecture Research Tool
SoMo: social motility
SRA: Serum Resistance-Associated Protein
SSR: strand switch region
TbLBPK: Trypanosoma brucei Lapatinib-Binding Protein Kinase
TbPCNA: proliferating cell nuclear antigen
TbPTP1: Trypanosoma brucei protein tyrosine phosphatase 1
TbVPS34: Trypanosoma brucei vacuolar protein sorting
TLF1: trypanosome lytic factor
TLK: Tousled-like kinase
TOR: target of rapamycin
TPP: target product profile
tRNA: transfer RNA
Trx: thioredoxin
Tyr: Tyrosine
UV: Ultra violet
Val: Valine
VCAM: Vascular cell adhesion protein
VEGFR : Vascular Endothelial Growth Factor Receptors
VSG: Variant Surface Glycoprotein
WHO: World Health Organization
ZFK: zinc finger kinase

“Consistency is the playground of dull minds”.

— Yuval Noah Harari.
Sapiens: A Brief History of Humankind

Chapter 1 General introduction

1.1 Trypanosoma brucei

The obligatory parasitic protozoan *Trypanosoma brucei* is an early divergent eukaryote widely used as a model organism to study many conserved and not conserved cellular functions. It is taxonomically clustered in the class "kinetoplastida" which groups flagellated protozoa characterized by a dense complex of mitochondrial circular DNA called the "kinetoplast". It is the causal agent of African Trypanosomiasis, a neurological disease of mammals that remains a major Public Health and economic problem in the areas of sub-Saharan Africa where the blood-feeding insect vector responsible for transmission, the tsetse fly, lives. African Trypanosomiasis has been neglected for years by the drug development industry as the potential market –large proportions of the population living in poor rural areas without access to decent health care systems- was not considered safe in order to recover the investment. Since the 2012 London Declaration on Neglected Tropical Diseases (inspired by the WHO 2020 roadmap), the raising of public-private partnerships, the connection of health care organizations with global experts trying to understand the disease, and the commitment of local and international authorities, society is in a position to achieve elimination in the not too distant future.

1.1.1 Kinetoplastids

The class Kinetoplastida are highly divergent eukaryotic protozoan belonging to the supergroup Excavata, group Discoba, subgroup Discicristata, and phylum Euglenozoa (Moreira, López-García and Vickerman, 2004), Figure 1-1. It is further divided in Prokinetoplastina and Metakinetoplastina, two sub classes, the last of which contains the order Trypanosomatida that includes three of the parasitic uniflagellates responsible for some of the diseases that cause greater economic and public health burden in the most impoverished areas of the world: *Trypansoma brucei, Trypanosoma cruzi* and *Leishmania species* (Lukeš *et al.*, 2014). *T. brucei* and *T. cruzi* genomes share an average of 57% identify, while *T. brucei* and *L. major* share about 44% (El-Sayed, Myler, Blandin, *et al.*, 2005). There is evidence that it is possible to develop therapies able to target the three

parasites with a single drug (Khare *et al.*, 2016). All kinetoplastids share divergent aspects of cell biology in respect to other eukaryotes as mitochondrial RNA editing, a particular genomic organization that will be discussed below, mRNA trans-splicing, or compartmentalized glycolysis (Simpson, Stevens and Lukeš, 2006; Lukeš *et al.*, 2014).

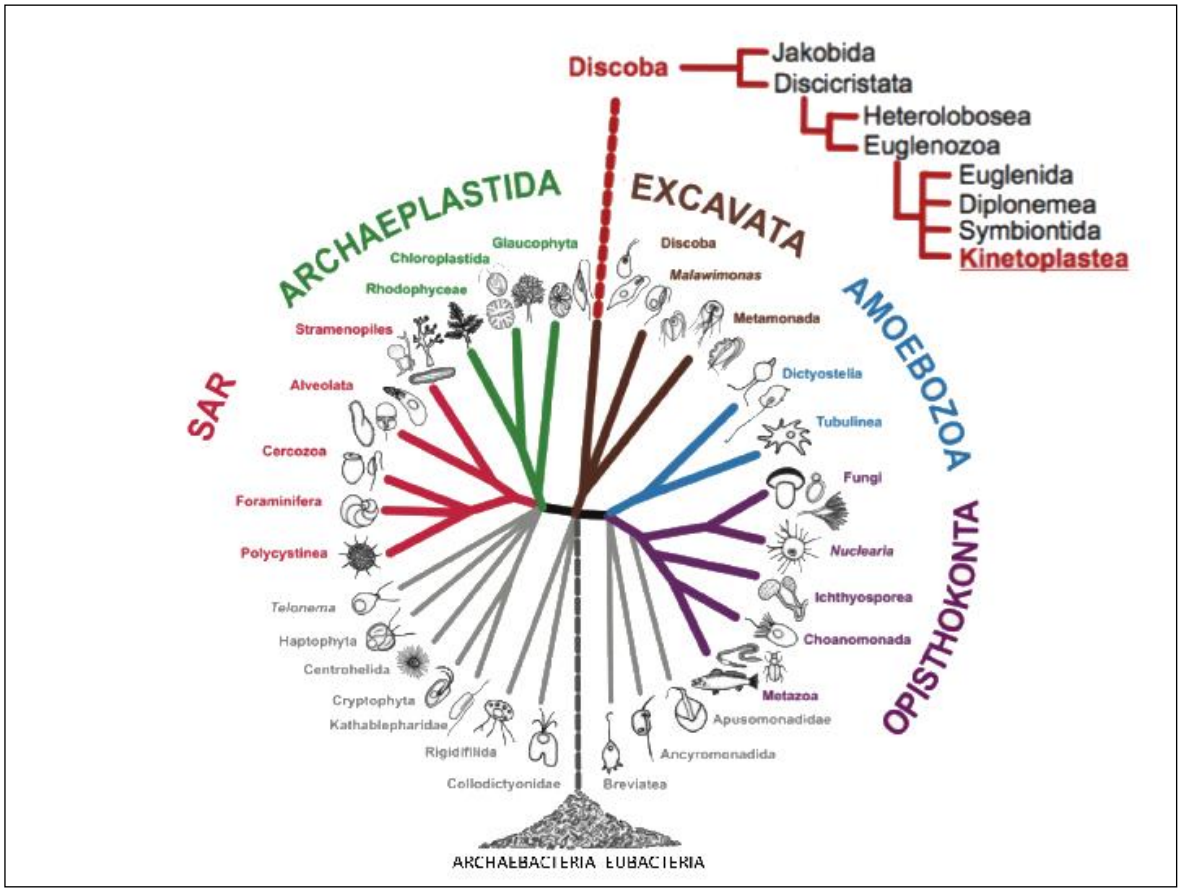


Figure 1-1. Eukaryotic phylogenetic tree highlighting the taxonomic classification of kinetoplastid parasites.
Modified from (Lukeš *et al.*, 2014)

1.2 African Trypanosomiasis

There are two subspecies described as a cause for Human African Trypanosomiasis (HAT) also called “sleeping sickness” due to the characteristic sleep-wake cycle disorders: *Trypanosoma brucei rhodesiense,* that has acute symptoms killing the patients in weeks or months, and is endemic in East and South Africa; and *Trypanosoma brucei gambiense*, that is a more progressive chronic debilitating disease, and is prevalent in Central and West Africa. *T. b. gambiense* causes 95% of the reported cases (Stuart *et al.*, 2008; Brun *et al.*, 2010).

There is only one subspecies responsible for the veterinary disease called Animal African Trypanosomiasis (AAT) or "Nagana": *Trypanosoma brucei brucei*. In addition, AAT can be caused by other three trypanosome species: *Trypanosoma vivax, Trypanosoma evansi,* and *Trypanosoma congolense*. These animal pathogens rarely produce disease in humans although few cases have been reported (Truc *et al.*, 1996; Joshi *et al.*, 2005; Deborggraeve *et al.*, 2008). They cause great financial losses in local economies, e.g. $1.3 billion were estimated in 1999 (Kristjanson *et al.*, 1999), as they compromise food security and life-stock farming due to increase in mortality, reduction in calving rates, draft power or meat and milk production (Holt *et al.*, 2016). *T. b. brucei* is commonly used as a laboratory model to study HAT.

1.2.1 Physiopathology

Even when the subspecies of the parasite cause different chronicity and virulence, they both develop two consecutive stages. Firstly, haemolymphatic, which happens when the parasite lives and replicates in the circulatory apparatus -blood and lymph- causing intermittent but chronic fever, itching, headache, lymphadenopathy, and sometimes hepato-splenomegaly. Secondly, meningo-encephalitic stage, that is when eventually the parasite penetrates the Central Nervous System (CNS) causing sleep disorders, coma and death if stays untreated (Brun *et al.*, 2010). It is unclear to what extent the neurological symptoms are caused by the parasite itself or by the immune response that it triggers (Kristensson *et al.*, 2010).

During the meningeal stage, *T. brucei* access to the cerebrospinal fluid at the circumventricular organs (together with inflammatory cells) throughout the fenestrated capillaries at the coroid plexus, where the blood brain barrier (BBB) is absent (Mhlanga, Bentivoglio and Kristensson, 1997; Kristensson *et al.*, 2010).

During the encephalitic stage, access to the brain parenchyma through the BBB has been documented in experiments with craniotomized mice. It was unclear whether it happened by crossing the vascular endothelium, or by invading the subarachnoid space, and spreading along the Virchov-Robin space (Frevert *et al.*, 2012). Experiments performed with non-invasive *in vivo*-imaging techniques,

however, were unable to reproduce this observation and parasites were retained in the cortical dura (Coles *et al.*, 2015).

Since the parasite expresses major surface metalloproteases (Grandgenett *et al.*, 2007), oligopeptidases, or cysteine peptidases such as brucipain (Lonsdale-Eccles and Grab, 2002) it has been proposed a proteolytic opening of the BBB, however there is no evidence of its disruption during transmigration of the parasites (Wolburg *et al.*, 2012).

It has been proposed also that brucipain may activate protease-sensitive G-protein coupled endothelial receptor, triggering Ca^{2+}-PKC signalling that would open tight junctions in the BBB (Grab *et al.*, 2009). Finally, passage from the perivascular space to the CNS may also happen with an increase of BBB permeability associated with lymphocyte-derived IFN-γ production. This has been proposed to happen in areas of the parenchymal basement membrane coated with laminin α4 (Masocha *et al.*, 2004), assisted by increased adhesion triggered by ICAM-1 and VCAM-1 expression, and secretion of endothelium activation factors (Grab and Kennedy, 2008)

Infection of the neurohypophisis, subformical organ, and the vascular organ of the lamina terminalis, causes also deregulation of vasopressin and luteinizing hormone causing oedema, impotency, and infertility (Kennedy, 2013).

Death happens by heart failure, usually after pancarditis in acute HAT mainly; or by opportunistic infection, malnourishment, and encephalitis, in chronic HAT (Schmidt, 1983).

Animal models for all subspecies of *T. brucei* show good correlation with observations made in post-mortem human samples of the disease: they all suffered neuroinflammation, progressing from meningitis to cerebral vasculitis and final encephalitis. They show infiltrates of lymphocytes, plasma cells and macrophages (accompanied by the parasites); activation of astrocytes and microglia is frequently reported; and little demyelinization occurs towards the terminal stages of the disease while nervous tissue is generally spared (Rodgers, 2010; Mogk *et al.*, 2016).

1.2.2 Epidemiology

Trypanosoma brucei is transmitted mainly by the bite of a blood-feeding vector, the tsetse fly: 31 species and subspecies of the genus *Glossina* (Jordan, 1977). The parasite spreads across all the regions where the insect lives (Brun *et al.*, 2010): mainly in rural areas of 30 sub-Saharan African countries Figure 1-2, although it has been reported in urban and periurban locations too (Robays et al. 2004). Additionally, for *T. b. gambiense* infection -which is an anthroponotic- a few cases have been identified after congenital transmission (Lindner and Priotto, 2010), or laboratory accidents (Herwaldt, 2001), and -at least one case reported- sexual interaction (Rocha *et al.*, 2004). Although it has never been documented, blood transfusions or organ donations would constitute a risk of infection too (World Health Organization, 2013). Efficiency of transmission in *T. b. rhodesiense* –which is a zoonosis and has its reservoir in domestic livestock- has been proven much higher: a single parasite can colonize the vector (Maudlin and Welburn, 1989) and a single tsetse bite can infect the mammal (Thuita *et al.*, 2008). Lower efficiency in *T. b. gambiense* is compensated by chronicity of the infection in the mammalian host making the parasite available in blood for longer (World Health Organization, 2013).

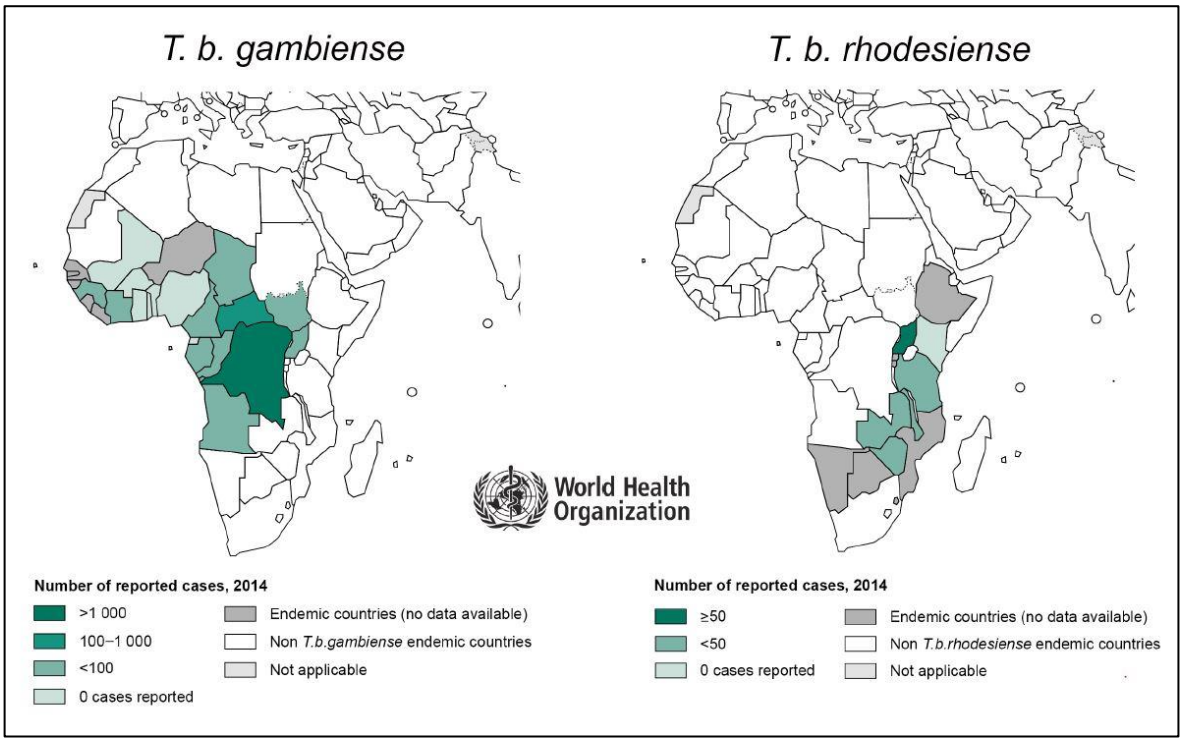


Figure 1-2. Distribution of human African trypanosomiasis caused by *T. b. gambiense* and *T. b. rhodesiense*.

Adapted from World Health Organization, Control of Neglected Tropical diseases site (http://www.who.int/trypanosomiasis_african/country/en/), accessed December 2016.

The World Health Organization (WHO) considered 2020 a realistic deadline for elimination as a public health problem of Human African Trypanosomiasis (HAT), which would mean reducing the incidence to less than 1 new case per every 10000 inhabitants (Savioli and Daumerie, 2012). In 2012, 13.1 million people lived still in areas where this was not true. Only by better usage of the drugs already available, and the coordinated work of local and international bodies, 2015 provided the lowest record of new reports of *T. b. gambiense* HAT in history: 3000 new cases in 24 countries of west and central Africa. Together with *T. b. rhodesiense* HAT data for east and southern Africa in 2014 (117 new cases in 13 countries), these are symptoms that WHO 2020 goals might be achievable (Drugs for Neglected Diseases initiative, 2015).

A model-based analysis of screening data predicted underestimation in 50% of the studies (Jo Robays *et al.*, 2004; Checchi *et al.*, 2012). The existence of animal reservoirs (El-Sayed, Myler, Bartholomeu, *et al.*, 2005), the climate change leading to vector redistribution (Moore *et al.*, 2012) and the sexual recombination described at least during the insect stages (Peacock, Bailey, *et al.*, 2014) have produced resurgence from lower levels in the past (Delespaux and de Koning, 2007; Brun *et al.*, 2011). Namely, after figures close to eradication in the 1960s, African independence and civil warfare led to interruption of control and surveillance measures implemented during the colonial period producing a peak of incidence (>300,000 estimated infections) in the 1990s (Paquet *et al.*, 1995; Ekwanzala *et al.*, 1996; Smith, Pepin and Stich, 1998; Barrett, 1999; Moore and Richer, 2001; Stanghellini and Josenando, 2001; Van Nieuwenhove *et al.*, 2001). *T. b. gambiense* tends to be epidemic (Ford, 1969; World Health Organization, 2013) while *T. b. rhodesiense* is more endemic, with a large variety of wild and domestic animal reservoirs

1.2.3 Diagnostics

A major problem affecting epidemiology of African Trypanosomiasis is the lack of systematic screens, which is limited to specific campaigns conducted in high-

and moderate-transmission areas. Many patients are identified by self-reporting to health care centres, something that is not possible in poor remote areas. For *T. b. gambiense,* there are serological diagnostic tools as the card agglutination test for trypanosomiasis (CATT). Otherwise diagnostics relays on evaluation of clinical symptoms and visual parasite identification. Thus, analysing samples of blood or chancre aspirates under the microscope requires skilled medical teams. In order to determine the disease stage, identification of the parasite in cerebrospinal fluid is conducted after lumbar puncture. There are no serological tests currently available for *T. b. rhodesiense,* most of these cases are identified attending to signs and symptoms (Sutherland *et al.*, 2015). Optimization of new rapid diagnostic tests is under way (Sternberg *et al.*, 2014), and a series of new antigens that may be useful also for *T. b. rhodesiense* diagnostics have been proposed (Biéler *et al.*, 2016).

1.2.4 Therapies and control strategies

The current therapies have limitations due to high toxicity and difficult administration patterns (Wilkinson and Kelly, 2009). For the early haemolymphatic stage, pentamidine (*T. b. gambiense*) and suramin (*T. b. rhodesiense*) are the treatments of choice. Once the parasite reaches the central nervous system (stage 2), for *T. b. gambiense,* the most satisfactory solution is NECT (Nifurtimox-Eflornithine Combined Therapy), less toxic than melarsoprol and of easier administration than eflornithine in independent therapy. However, for *T. b. rhodesiense,* melarsoprol is still the only available stage 2 option. Since the Drugs for Neglected Diseases Initiative (DNDi) has been in place, another two drugs have entered clinical trials (fexinidazole and SCYX-7158) however, fears of development of parasite resistance make essential appointment of novel treatment options (Jones *et al.*, 2015; Wyllie *et al.*, 2015). A lot need still to be done.

In *T. b. gambiense* infection control, mass-screening and treatment has been shown successful in reducing the prevalence as humans are the main reservoir for the parasite (Simarro *et al.*, 2006, 2015) but the above mentioned underestimation rates led to persistence in many cases (Jo Robays *et al.*, 2004; Checchi *et al.*, 2012).

In 2000, after the 36th summit of the Organisation of African Unity (OAU), several African Heads of State engaged with the Pan African Tsetse and Trypanosomiasis Eradication Campaign (PATTEC) whose objective was the total eradication of tsetse flies across the continent aiming to blocking transmission (Kabayo, 2002). Although commitment of local governments with the initiative have been cyclic along time, four approaches were used mainly for this purpose: artificial baits (insecticide-treated traps), insecticide-treated cattle, aerial spraying, and the sterile insect technique (Shaw *et al.*, 2013a), cost-effectiveness of the different approaches have been subject of debate in the field (Bouyer, Seck and Sall, 2013; Shaw *et al.*, 2013b).

1.3 Life cycle

T. brucei exhibits a complex life cycle that requires adaptation to many different environments, including different compartments within the blood-feeding insect vector responsible for transmission between different mammalian hosts: the tsetse fly (*Glossina sp.*). Figure 1-3.

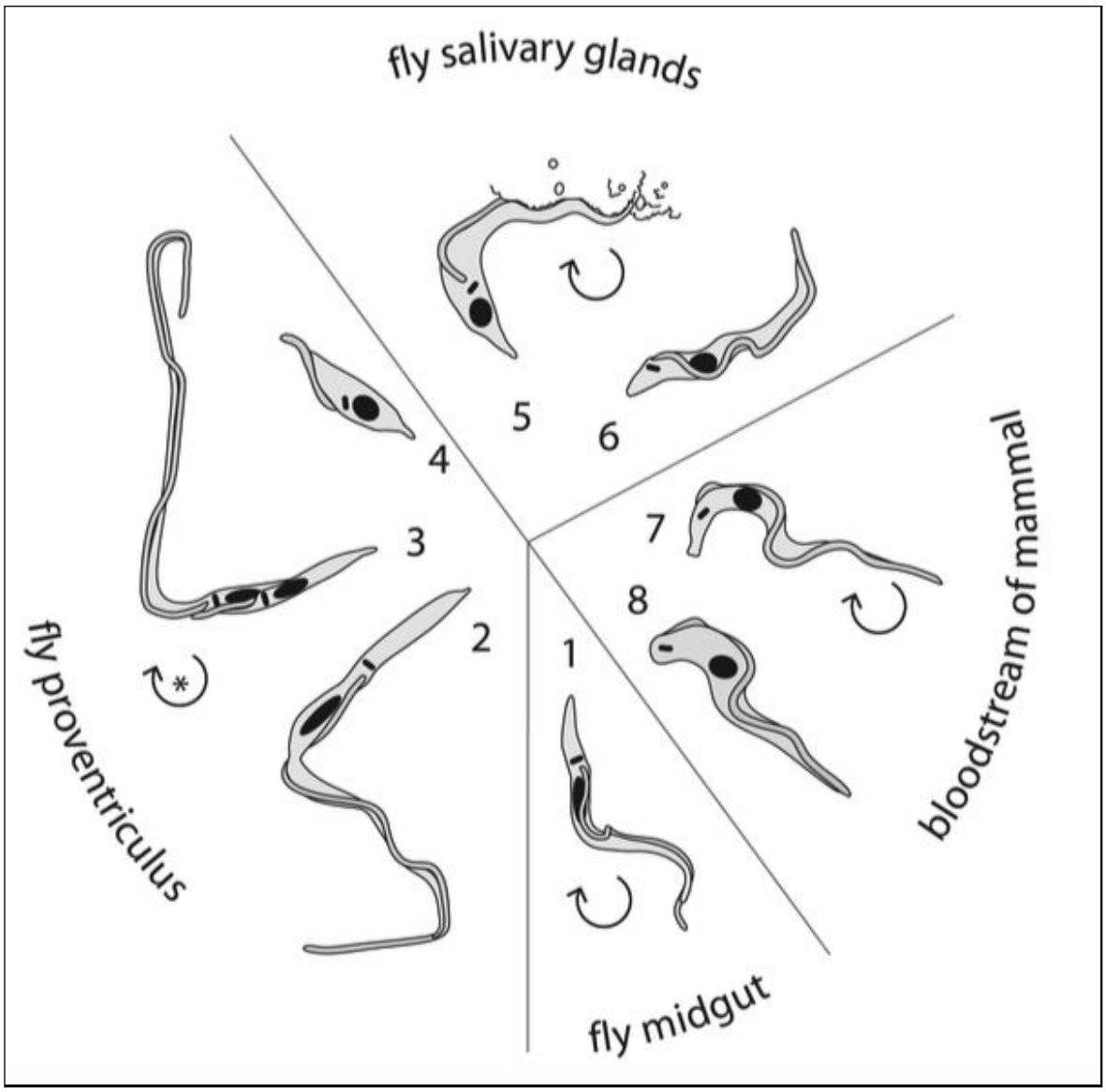


Figure 1-3. *Trypanosoma brucei* life-cycle.
(1) Procyclic trypomastigote, (2) elongated mesocyclic trypomastigote, (3) long epimastigote, (4) short epimastigote, (5) attached salivary gland epimastigote, (6) mammalian-infective metacyclic trypomastigote, (7) Slender bloodstream trypomastigote, and (8) stumpy bloodstream trypomastigote. Curved arrows = replicating stages. * = single asymmetric division. Adapted from Gluenz et al. (2015).

1.3.1 *T. brucei* in the mammalian host

Once in the mammal, *T. brucei* resides extracellularly in the bloodstream, tissue fluids, central nervous system, adipose tissue (Trindade *et al.*, 2016), and –as recently shown- in the skin (Capewell *et al.*, 2016). It presents two main life cycle stages: the slender, replicative, trypomastigote; and the stumpy, cell cycle-arrested form. These are two ends of an irreversible differentiation

process, which proceeds through many intermediate stages, and happens in response to density-dependent signalling (Rico *et al.*, 2013). All of them are coated with a densely packed single Variant Surface Glycoprotein (VSG) that assists the parasite in overcoming the immune response at a population level (Mugnier, Stebbins and Papavasiliou, 2016).

1.3.1.1 Slender to stumpy form differentiation

Differentiation to stumpy forms relies upon a quorum sensing mechanism in response to high densities reached by the parasite in the peak of parasitemia (Rico *et al.*, 2013). Many laboratory-adapted strains passaged repetitively from *in vitro* culture to mice without the natural mediation of the insect vector, lost their ability to differentiate into stumpy forms by monitoring their own concentration (Turner, 1990). These cells, selected before cryopreservation methods became available, can grow without limit until killing the host a few days after the infection. Their high virulence would be counter–selective in natural conditions as flies would not have time to transmit parasites before they were able to kill the host (Mony and Matthews, 2015).

Known signalling pathways involved in slender to stumpy form differentiation

The quorum sensing molecule has not been identified yet but has been proposed to be soluble, of low molecular weight, and heat stable, originated by *T. brucei* as an autocrine response (Vassella *et al.*, 1997). It has been termed Stumpy Induction Factor (SIF), and monomorphic parasites are not responsive to it (MacGregor *et al.*, 2012).

A 2-3 fold increase of cAMP levels has been detected in the peak of parasitemia which declines immediately after stumpy differentiation (Mancini and Patton, 1981), conditioned medium with SIF also increases cAMP in cultured pleomorphic parasites (Vassella *et al.*, 1997). Differentiation could only be triggered by hydrolysable analogues of cAMP, such as 8-(4-Chlorophenylthio)adenosine 3′,5′-cyclic monophosphate (8-pCPT-cAMP), suggesting that its metabolites (and not cAMP itself) are the responsible factors for triggering differentiation. The process requires phosphodiesterase activity (Oberholzer *et al.*, 2007). Unlike SIF,

cell permeable, hydrolysable cAMP analogues triggered stumpy differentiation also in monomorphic parasites: “stumpy*” (Laxman *et al.*, 2006). This suggests that monomorphic *T. brucei* conserves some of the downstream components of the SIF-mediated stumpy differentiation pathway.

A phenotypic screen detected RNA interference (RNAi) of several proteins causing insensitivity to 8-cPTP-cAMP, suggesting that they are involved in pathways regulating this signalling mechanism (Mony *et al.*, 2013).

Several protein kinases and phosphatases were found in this study with yeast orthologues having dormancy-triggering functions in response to glucose-starvation (Mony and Matthews, 2015). Among them there was a YAK (Tb927.10.15020) protein kinase, a dual specificity phosphatase (DsPhos, Tb927.7.7960), and the PKA regulatory subunit (PKA-R, Tb927.11.4610), all having yeast orthologues connected in a same pathway. In yeast, YAK translocates to the nucleus upon glucose starvation where it phosphorylates PKA-R, causing inactivation that leads to cell cycle arrest (Griffioen *et al.*, 2001); and DsPhos regulates PKA directing transcriptional changes also linked to glucose starvation (Zaman *et al.*, 2008). PKA (cAMP-dependent protein kinase) increases proliferation in glucose-rich media, activated by cAMP (Griffioen *et al.*, 2000). As it happens in yeast, glucose starvation following the peak of parasitemia, may be the cause for cAMP decrease and subsequent cell cycle arrest. PKA has been found in the flagellar matrix of *T. brucei* and RNAi impairs the parasite´s motility (Oberholzer *et al.*, 2011), a function regulated by PKA in other eukaryotes such as *Dictyostelium discoideum* (Simon *et al.*, 1992).

An AMPKα2 catalytic subunit (Tb927.3.4560) was also identified as required for stumpy development in the RNAi screen with yeast orthologues inhibiting TOR (Carling *et al.*, 2011). Interestingly, TbTOR4 downregulation triggers stumpy formation in *T. brucei* (Barquilla *et al.*, 2012). AMPK may act, like the yeast orthologue, as an energy sensor that reduces catabolism when AMP levels increase as a consequence of glucose starvation (Mony and Matthews, 2015). In fact, AMPKα2 has been shown to form complexes with AMPKα1 (Tb927.10.5310), whose activation via phosphorylation in a tyrosine residue at the activation loop, is mediated by cAMP analogues (that also inhibit TbTOR4) and oxidative stress;

and that once it is active, triggers differentiation to stumpy forms (Saldivia *et al.*, 2016).

MAPK (mitogen-activated protein kinase) cascades, are a three module system of protein kinases that convey extracellular signals activating transcription factors in other eukaryotes (Garrington and Johnson, 1999). In *T. brucei*, knockdown of a MAP3K (upstream component of these cascades), Tb927.2.2720, also renders cells insensitive to 8-pCTP-cAMP and SIF-mediated differentiation (Mony and Matthews, 2015). Finally, a NEK family of kinases, indistinguishable via RNAi (Tb927.10.5930/5940/5950), was also found retarding stumpy differentiation upon knockdown. NEKs (Never in mitosis A-related kinases) are a family of check-point regulators in the cell cycle (Moniz *et al.*, 2011; Fry *et al.*, 2012). This family of protein kinases is relatively expanded in *T. brucei* (Parsons *et al.*, 2005), likely having parasite specific functions as regulation of progression throughout the life cycle. In fact other NEK, RDK2, that will be analysed in detail in one of the chapters of this thesis, has been identified regulating this aspect of the parasite biology in subsequent stages of the life cycle (Jones *et al.*, 2014).

Several gene expression-regulation factors have been identified in the RNAi screen selecting lines after 8-pCTP-cAMP treatment as, for instance, the transcriptional-silencing Imitation SWItch (ISWI), which is a stumpy inducer (Hughes *et al.*, 2007; Stanne *et al.*, 2011). But more relevant according to *T. brucei* biology, they found RNA binding proteins (RBPs), which are the principally responsible for post-transcriptional regulation in the parasite (Clayton, 2014; Kolev, Ullu and Tschudi, 2014). RBP7A and RBP7B (Tb927.10.12090 and Tb927.10.12100, respectively) were found as required for slender-to-stumpy differentiation mediated by 8-pCTP-cAMP in the RNAi screen, with overexpression causing premature stumpy formation (Mony *et al.*, 2013). Other three hypothetical proteins found in this screen were identified as positive (Tb927.9.4080) or negative (Tb927.11.6600) transcriptional regulators in a whole genome overexpression analysis targeting RNA binding proteins (Erben *et al.*, 2014). The third, Tb927.11.2250, has not been validated yet as a stumpy inducer individually but it does bind RNA (Mony and Matthews, 2015).

Another 10 hypothetical proteins were identified in the RNAi screen, that have orthologues in other eukaryotes involved in mitochondrial import or

ubiquitination (Mony *et al.*, 2013). They may be involved in mitochondrion development, intrinsic to stumpy formation; or degradation of proteins that are no longer needed in that stage of the life cycle (Mony and Matthews, 2015).

Transition to stumpy forms is accompanied by silencing of the VSG expression site (ES) (Amiguet-Vercher *et al.*, 2004), while RNAi of VSG causes cell cycle arrest (Sheader *et al.*, 2005; Smith *et al.*, 2009). This prepares stumpy cells for quick coat exchange upon tsetse fly uptake, and may cause their death in the bloodstream upon VSG turnover (Sheader *et al.*, 2005), possibly due to lysis mediated by the alternative pathway of the complement (Ferrante and Allison, 1983). That would explain the 55 h life-span calculated in the bloodstream for this particular life-stage (MacGregor *et al.*, 2012). Depletion of three expression site-associated genes (ESAG1, 2, and 8) has also been described causing expression of stumpy-specific PAD transporters (Dean *et al.*, 2009), and producing cell cycle arrest (Batram *et al.*, 2014). It happens in a reversible manner and orchestrated by disruptor of telomeric silencing-1B (DOT1B)-mediated histone methylation.

Other factors that had been linked to stumpy formation before the RNAi screen was performed were: zinc finger kinase (ZFK), *in vitro* but not *in vivo* (Vassella *et al.*, 2001); and TbMAPK5, *in vivo* and *in vitro* although reducing the parasite survival in the mouse bloodstream (Domenicali Pfister *et al.*, 2006). While ZFK and MAPK5 knockdown only caused stumpy differentiation in pleomorphic parasites, TbTOR4 RNAi caused it also in monomorphic cells. It may suggest a role of TbTOR4 downstream the other two, in the part of the pathway retained by monomorphic strains which is sensitive to 8-pCTP-cAMP but not to SIF (Mony and Matthews, 2015).

1.3.1.2 Antigenic variation to evade the immune system

Slender forms replicate actively and travel in the blood colonizing the host, while expressing a single glycophosphatidyl inositol (GPI)-anchored Variant Surface Glycoprotein (VSG) per cell, which can be switched upon expansion of the population to create diversity (McCulloch and Field, 2015).

The parasite harbours >1000 different VSG genes placed at the subtelomeric regions within 15-20 expression sites (ES) located in the telomeres (see below). Only one ES is active at a given time, expressing a single VSG. 80% of VSG genes are pseudogenes with only about 400 complete open reading frames (Berriman *et al.*, 2005; Cross, Kim and Wickstead, 2014). Switches happen *in situ*, when one ES inactivates after a second becomes active; or by "gene conversion", when a different VSG jumps into the active ES (Pays *et al.*, 2001; Li, 2015). It is also possible homologous recombination leading to generation of novel variants, the so-called "mosaic" VSG (Mugnier, Stebbins and Papavasiliou, 2016). At a given time in an infection, >80 different VSG-expressing clones have been identified (Hall, Wang and David Barry, 2013; Mugnier, Cross and Papavasiliou, 2015), Figure 1-4.

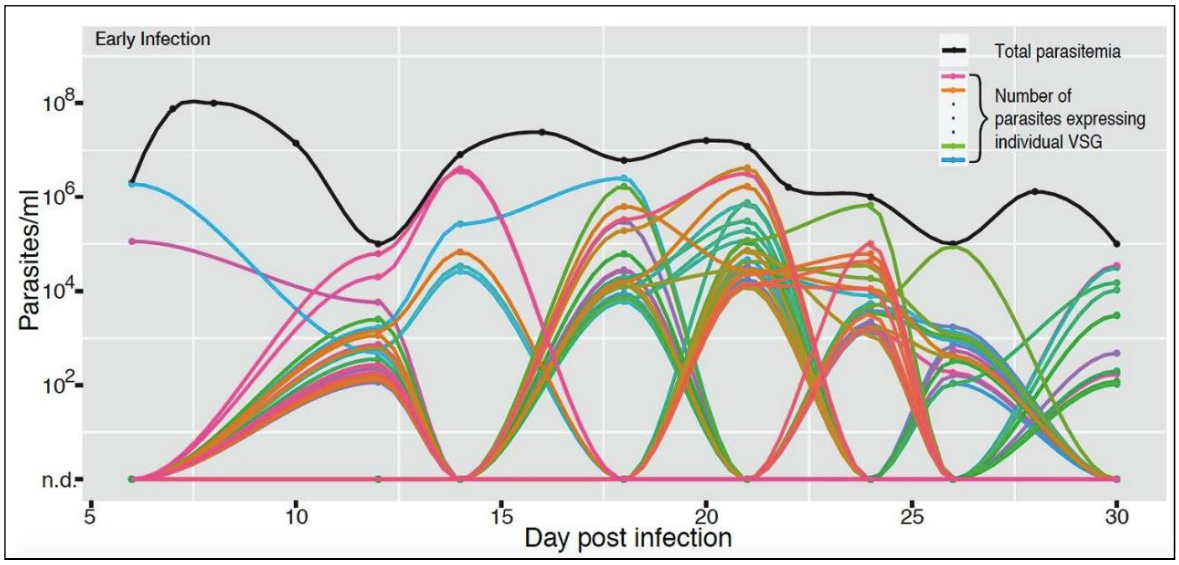


Figure 1-4. *In vivo* dynamics of VSG expression.
Based in RNA levels during the first 30 days of 4 mice infections, purifying parasites from blood every 3-4 days. Black line represents overall parasitemia levels. Coloured lines represent those VSG variants expressed that constituted more than 0.1% of the population. Adapted from Mugnier et al. (2015)

The host builds an adaptive immune response against at least the most abundant variants, leading to their clearance and enabling outgrowth of cells that have switched to an antigenically distinct VSG. Iteration of this process -coupled to stumpy differentiation- leads to the characteristic waves of parasitemia (Matthews, McCulloch and Morrison, 2015). Antigenic variation makes it extremely difficult to develop an effective vaccine (Deitsch, Lukehart and

Stringer, 2009; MacGregor *et al.*, 2012). Since current therapies have limitations due to high toxicity and difficult administration patterns, there is a need to explore new routes for disease treatment and parasite eradication.

Immunoglobulin M (IgM) is the critical antibody isotype for VSG recognition and clearance in acute infections developed by monomorphic parasites unable to undergo stumpy differentiation (Tyler *et al.*, 2001), but IgG anti-VSG is the main antibody responsible for opsonisation in pleomorphic strains and in the natural environment (Schwede *et al.*, 2015). Parasites expressing a given VSG are cleared within a week (Magez *et al.*, 2008), once the antibody titre overwhelms the endocytosis and degradation pathways (Engstler *et al.*, 2007). Protruding 12-15 nm from the surface, VSG shields most of invariable surface proteins (such as hexose transporters). However, IgG has been proven to access other surface antigens, some able to reach above the top of the VSG barrier, as the invariant surface glycoproteins (ISG) or ESAG4 transmembrane proteins (Hsia, Beals and Boothroyd, 1996; Schwede *et al.*, 2015). Antibodies against the invariable proteins are not sufficient to produce immunity (Ziegelbauer and Overath, 1993; Sullivan *et al.*, 2013).

Slender forms undergo VSG switching at a higher rate, stumpy are more efficient at antibody clearance by internalization and recycling of opsonized VSGs through the flagellar pocket located in their posterior end. They have been shown to be dragged there by means of hydrodynamic flow (Engstler *et al.*, 2007). Upon stumpy form differentiation, the haptoglobin-haemoglobin receptor (HpHgR) is downregulated, contributing to the immune defence of quiescent forms. HpHgR is the internalization route for the serum-borne trypanosome lytic factor (TLF1) (Vanhollebeke *et al.*, 2008), and protection against it is mediated by SRA, a VSG expression site (ES) component (Van Xong *et al.*, 1998). Depletion of HpHgR may compensate ES downregulation in stumpy forms.

1.3.2 *T. brucei* in the tsetse fly

T. brucei progression throughout the tsetse fly is the more challenging section of the life cycle, as less than 20% of the flies develop a mature infection under optimised laboratory conditions or in the wild (Haines, 2013). Susceptibility toward infection and capacity of the parasites to progress until metacyclogenesis

in the salivary glands is variable between different species of fly and parasite (Macleod *et al.*, 2007; Peacock *et al.*, 2012). Immune response clears 95% of trypanosomes in the fly (Aksoy, Gibson and Lehane, 2003). Once in the tsetse midgut the stumpy forms develop into procyclic trypomastigotes, which eventually cross the peritrophic membrane and proliferate in the ectoperitrophic space (Gibson and Bailey, 2003). Then, they migrate to the proventriculus (Van Den Abbeele *et al.*, 1999), where they differentiate into long mesocyclic trypomastigotes able to divide into short and long epimastigotes (Sharma *et al.*, 2008). Epimastigotes can re-enter the gut lumen migrating through the foregut with apparently only the short epimastigotes entering the salivary glands (Van Den Abbeele *et al.*, 1999). There, evidence of meiosis -causing genetic exchange- has been reported (Peacock *et al.*, 2011) but it is not needed for the final differentiation to mammalian-infective metacyclic trypomastigotes (Van Den Abbeele *et al.*, 1999), Figure 1-5. Overexpression *in vitro* of an RNA binding protein, RBP6 (Tb927.3.2930), has been proven to reproduce in axenic culture all the fly stages of the parasite, depending on the duration and level of expression (Kolev *et al.*, 2012).

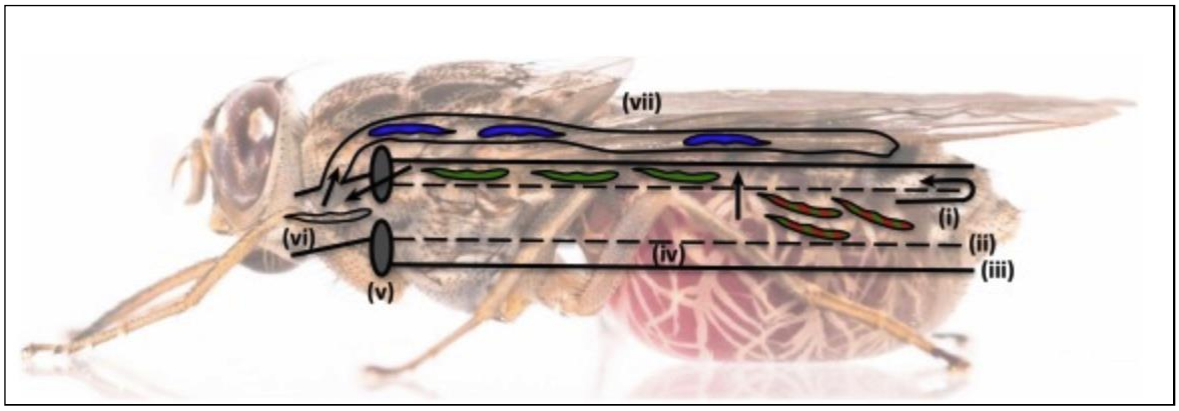


Figure 1-5. Schematic representation of the migration of *Trypanosoma brucei* throughout the different compartments of the tsetse fly.
(i) Midgut lumen, (ii) peritrophic matrix, (iii) midgut epitelium, (iv) ectoperitrophic space, (v) proventriculus, (vi) oesophagus, and (vii) salivary glands. Colours representing stage-specific surface antigens: EP procyclin (green), GPEET procyclin (red), *brucei* alanine-rich protein (BARP, blue). Adapted from (Imhof and Roditi, 2015).

1.3.2.1 Stumpy to procyclic form differentiation

Stumpy forms are cell cycle-arrested and pre-adapted for survival in the insect midgut upon parasite uptake. They express PAD1 and PAD2, carboxylate surface receptors (Dean *et al.*, 2009), whose transport to the membrane increases at the lower temperatures that the parasite reaches in the tsetse fly. Under exposure to the environment they find in the midgut -rich in intermediates of the tricarboxylic acid cycle- they differentiate to procyclic forms more readily than slender forms. This can be mimicked *in vitro* by citrate/cis-aconitate incubation, limited protease digestion, or mild acidic treatment, (Rico *et al.*, 2013). Differentiation involves a fast coat exchange of VSG for dipeptide (EP) and pentapeptide (GPEET) repeats-enriched procyclins, which bear a protease resistant C-terminal domain (Acosta-Serrano *et al.*, 2001) and whose expression is regulated by mitochondrial factors (Vassella *et al.*, 2004). In addition, stumpy forms present a more elaborated mitochondrion, which provides metabolic adaptation. In the tsetse midgut proline, instead of glucose, becomes the main carbon source (Vassella *et al.*, 2000). Transcripts for mitochondrial proteins and procyclins are stocked up in stumpy forms (Kabani *et al.*, 2009; Queiroz *et al.*, 2009), ready for translation upon exposure to the environmental cues driving procyclic formation in the fly.

Non-differentiated slender forms die in a few hours while procyclic forms continue the life cycle. In addition, stumpy cell cycle arrest provides at least two selective advantages: it avoids overwhelming the mammal causing premature death before fly transmission; and prevents DNA replication during the environmental shock at the fly midgut that could induce mutations (Rico *et al.*, 2013).

A phosphatase cascade is the best-described signalling pathway regulating this developmental transition. A tyrosine phosphatase, TbPTP1 (Szöor *et al.*, 2006), prevents stumpy forms from differentiating to procyclic until internalization of citrate or cis-aconitate (CCA) (Brun and Schonenberger, 1981), through the PAD transporters (Dean *et al.*, 2009). This can be mimicked *in vitro* by exposure to CCA, but also to mild acidic conditions (Rolin *et al.*, 1998), or conducting TbPTP1-specific inhibition with a tool compound (BZ3) (Szöor *et al.*, 2006). In the absence of these triggers, a glycosomal DxDxT phosphatase, TbPIP39 (Szöor

et al., 2010), is dephosphorylated by TbPTP1 in a feed back loop where the dephosphorylated enzyme empowers activity of its phosphatase. Upon TbPTP1 inactivation via the above-mentioned differentiation cues, TbPIP39 is phosphorylated and the parasite progresses to the procyclic form. This can be triggered also by exposure of the cell surface to limited protease digestion in a mechanism independent of the phosphatase cascade (Szöőr *et al.*, 2013). Exposure to both CCA and protease treatment at once, seems to have a synergistic effect (Sbicego *et al.*, 1999), so both triggers may have a combined effect *in vivo*.

A kinome-focused RNAi screen of the parasite *in vitro* (Jones *et al.*, 2014) that identified several protein kinases fundamental for cell cycle regulation, detected also two -RDK1 and RDK2- whose depletion triggered spontaneous differentiation from bloodstream to procyclic-like forms. As bloodstream form "repressors", they cannot be responsible for TbPIP39 phosphorylation because they have the opposite effect, but their role will be discussed in detail in the fourth chapter of this thesis.

6 hours upon initiating differentiation with CCA, VSG is released from the surface by GPI hydrolysis mediated by GPI-PLC and TbMSP-B (Gruszynski *et al.*, 2006). Then, immediately after a peak of intracellular cAMP, takes place the first cell division followed by ESAG4 loss, a transmembrane adenylate cyclase (Rolin *et al.*, 1993). Translation activates again 1-6 hours upon CCA exposure (Capewell *et al.*, 2013), DNA synthesis starts only after 8-10 hours, and procyclins express in the surface from 2 hours post induction (Matthews and Gull, 1994). There is another peak of adenylate cyclase activation immediately before cells start proliferating, between 8-19 hours (Rolin *et al.*, 1993).

Upon differentiation, EP and GPEET procyclins, that are resistant to GPI-PLC hydrolysis (Field, Menon and Cross, 1991), are co-expressed in what are called "early procyclic forms" but after 7-9 days, GPEET is downregulated coinciding with colonisation of the ectoperitrophic space ("late procyclic forms"). *In vitro* GPEET can be maintained by addition of glycerol, or removed quickly by hypoxia. GPEET is regulated post-transcriptionally by a sequence in the 3´untranslated region (Vassella *et al.*, 2000).

It has recently been observed that when early procyclics (GPEET+) are inoculated into agarose plates, to mimic the semi-solid nature of the surfaces they find in the tsetse fly, they are capable of developing social motility (SoMo). Isolated colonies grew until a threshold of about $1.6x10^6$ parasites and then created radial projections (Imhof and Roditi, 2015). These finger-like protrusions moved in masse from the inoculation site and, if streams protruding from different colonies came close, they shifted their direction to avoid contact, suggesting a repellent substance might be emitted from different streams. As it happens with stumpy form formation or procyclic forms grown in liquid culture, a density sensing mechanism must be in place holding cells together in a stream and also guiding them radially in finger-like shapes. It is the developmental state (early or late procyclic) and not the presence of the GPEET marker what causes SoMo (Imhof *et al.*, 2014). Several proteins are differentially expressed in early and late stages although none has been yet linked to the phenomenon. It has been proposed that this behaviour might have implications toward colonization of the ectoperitrophic space that happens around 3 days post infection (Gibson and Bailey, 2003). Less than 1% of the parasites can make it (Oberle *et al.*, 2010; Imhof and Roditi, 2015). Independent procyclic forms have been shown to fuse their flagellar membrane and exchange proteins in a form of cell-to-cell communication (Imhof *et al.*, 2016).

1.3.2.2 Procyclic trypomastigote-mesocyclic trypomastigote-epimastigote

The way *T. brucei* access the ectoperitrophic space is not well understood, but analysis of the peritrophic matrix structure suggests the need of proteases and lipases in order to cross the inner membrane (Evans and Ellis, 1983; Rose *et al.*, 2014). There is evidence of parasites crossing, -flagellum first, through holes in the membrane suggesting its damage (Ellis and Evans, 1977; Evans and Ellis, 1983). It is hypothesized that peritrophic membrane thickness and composition affects refractoriness or susceptibility of the fly toward infection (Dyer *et al.*, 2013). Flies are less susceptible in the first 48h upon emergence but, after this, the possibilities of getting infected are higher in “teneral” (non-blood-fed) flies than in those that have taken previous uninfected bloodmeals (Kubi *et al.*, 2006; Walshe, Lehane and Haines, 2011).

Ectoperitrophic procyclic forms elongate as they move forward toward the proventriculus until reaching a length-threshold of about 35 μm and then stop proliferating (Van Den Abbeele *et al.*, 1999; Sharma *et al.*, 2008). Cell cycle is arrested in G2 until parasites break into the proventriculus where they elongate further until 40-60 μm becoming mesocyclic trypomastigotes (Vickerman, 1985; Sharma *et al.*, 2008). Then, asymmetric division happens that generates the long and short epimastigotes (1/3 in size), of which only the short is believed to survive, re-enter the gut lumen, migrate through the foregut, and colonize the salivary glands (Van Den Abbeele *et al.*, 1999; Sharma *et al.*, 2008). None of the epimastigotes are able to proliferate in the proventriculus (Sharma *et al.*, 2009). The time that takes *T. brucei* to invade the salivary glands since they enter the fly after a blood meal can oscillate between 6 and 30 days depending on the species (Van Den Abbeele *et al.*, 1999; Aksoy, Gibson and Lehane, 2003).

1.3.2.3 Epimastigote-metacyclic trypomastigote

Upon colonization of the salivary glands, 4 morphologies of *T. brucei* have been defined: attached epimastigotes, pre-metacyclics, nascent metacyclics, and mature free metacyclics (Steiger, 1973; Vickerman, 1985). The lack of an efficient *in vitro* culture technique has limited our capacity to study these stages but evidence of procyclin transcripts without procyclin expression has been reported (Acosta-Serrano *et al.*, 2001; Urwyler *et al.*, 2005), and a particular surface antigen, isoform of the bloodstream stage alanine-rich protein (BARP), has been defined as stage-specific marker (Urwyler *et al.*, 2007). It is in that location where the parasite undergoes genetic exchange throughout meiosis, with kinetoplastid DNA (linked to the mitochondrion) inherited biparentally, suggesting that genetic exchange involves fusion of cell bodies (Peacock, Ferris, *et al.*, 2014).

Short epimastigotes are able to replicate in the salivary glands upon attachment to the host microvilli by intricate ramified flagellar protrusions, which branch between indentations and form hemidesmosome-like attachment plaques, forming an undulating membrane (Vickerman, 1969). Flagellar outgrowths retract while parasites develop into uncoated trypomastigotes (pre-metacyclic forms) that remain attached through the plaques and are capable of replication. The so called “nascent metacyclic” forms stop replicating, and express a surface

VSG coat accompanied by changes in the glycosome and mitochondrion while remaining attached. In the final step, the non-replicative mature metacyclic forms, expressing VSG and preadapted for conditions encountered in the mammalian bloodstream, detach from the salivary gland surface (Tetley and Vickerman, 1985) and are inoculated into a mammalian host during a tsetse fly blood meal.

1.4 Other fundamental aspects of *T. brucei* cell biology

1.4.1 Cell structure

T. brucei is a single-cell protozoan and constitutes an excellent biology model because it is equipped with mostly single-copy organelles, including the flagellar pocket, the flagellum, the Golgi apparatus, the mitochondrion, the kinetoplast, and the nucleus (Matthews, 2005), Figure 1-6.

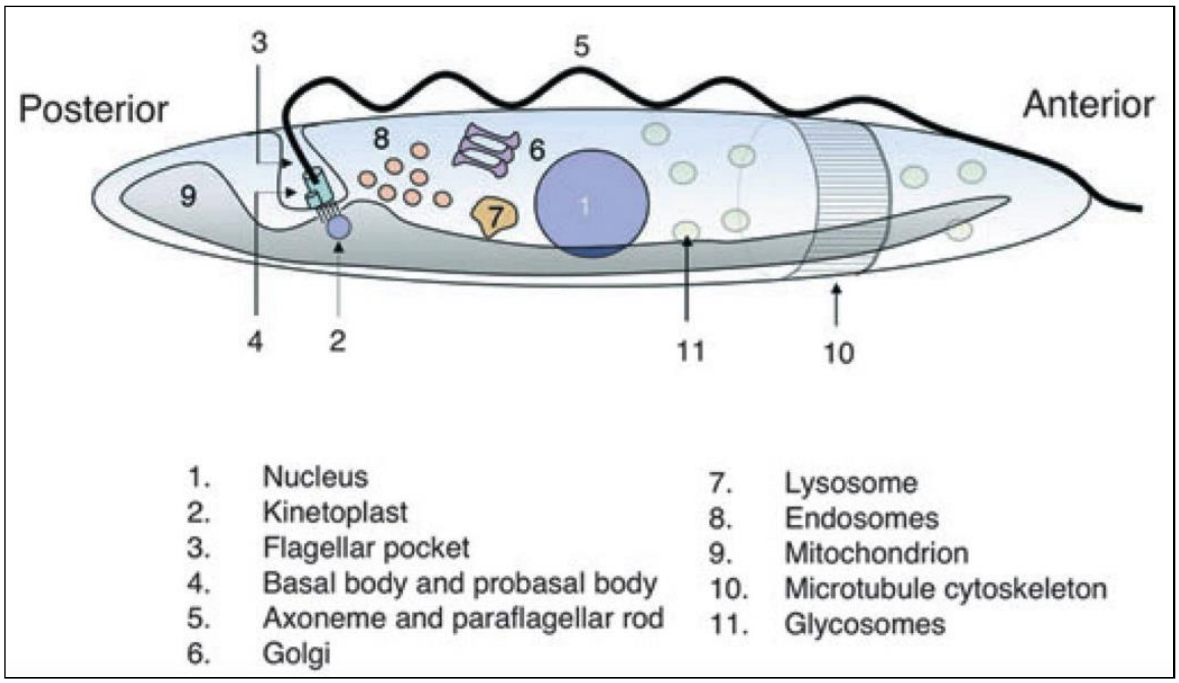


Figure 1-6. Cell structure of *Trypanosoma brucei*.
Adapted from (Matthews, 2005).

Along the *T. brucei* life cycle the parasite experiences extensive modifications in the metabolic, transcriptomic and proteomic profiles. From the morphological perspective, two main stages have been described that subdivide in slightly

different sub-stages: trypomastigotes (bloodstream, procyclic, mesocyclic and metacyclic) and epimastigote (long and short). The most obvious difference between these two forms is the location of the kinetoplast: posterior respect to the nucleous in trypomastigotes, and anterior in epimastigotes. All the sub-stages within each of these two types are slightly different but their general cellular structure is conserved (Wheeler, Gluenz and Gull, 2013).

They all share an elongated shape with a highly polarized microtubular structure that has its "plus" ends facing towards the posterior end, except for a particular microtubule quartet, which has the opposite orientation (Sunter and Gull, 2016).

The flagellum controls body length and is important for motility, signalling and cell division (Field and Carrington, 2009; Langousis and Hill, 2014). The canonical 9+2 axoneme structure with parafragellar rod of the flagellum starts in the basal body, which is connected to the kinetoplast via the tripartite attachment complex through the mitochondrial membrane. It exits the cell through the flagellar pocket, located in the posterior end of the cells (in trypomastigotes), or between the nucleus and the apical end (in epimastigotes) (Wheeler, Gluenz and Gull, 2013). The flagellar pocket, attached to the flagellum throughout a collar structure, is the only place where endo- and exocytosis takes place in the parasite (Overath and Engstler, 2004; Perdomo, Bonhivers and Robinson, 2016). The flagellum is connected to the whole cell body following a defined helical path by a complex cytoskeletal structure, the flagellar attachment zone (FAZ). Filaments in the FAZ are linked to the antiparallel microtubule quartet of the cell microtubular corset, leaving free the distal tip of the flagellum (Sunter and Gull, 2016).

1.4.2 Cell cycle

Cell-cycle of trypomastigotes in *T. brucei* is similar to that of other eukaryotes but with peculiarities due to the presence of single-copy organelles and the class-specific kinetoplast (Matthews, 2005; Vaughan and Gull, 2008). Organelles are duplicated by a template-based mechanism where a new cellular component is generated close to the old one (Vaughan and Gull, 2008). Regulatory molecular mechanisms are divergent too (Hammarton, Monnerat and Mottram, 2007; Li, 2012), for instance the checkpoint that prevents cytokinesis until

mitosis is completed is missing in bloodstream forms (Ploubidou *et al.*, 1999; Hammarton *et al.*, 2003), Figure 1-7.

In the G1 phase, the basal body duplicates and the new flagellum starts being synthetized (Woodward and Gull, 1990; Wheeler, Gluenz and Gull, 2013). Immediately after, the parasite duplicates the Golgi apparatus (Ho *et al.*, 2006).

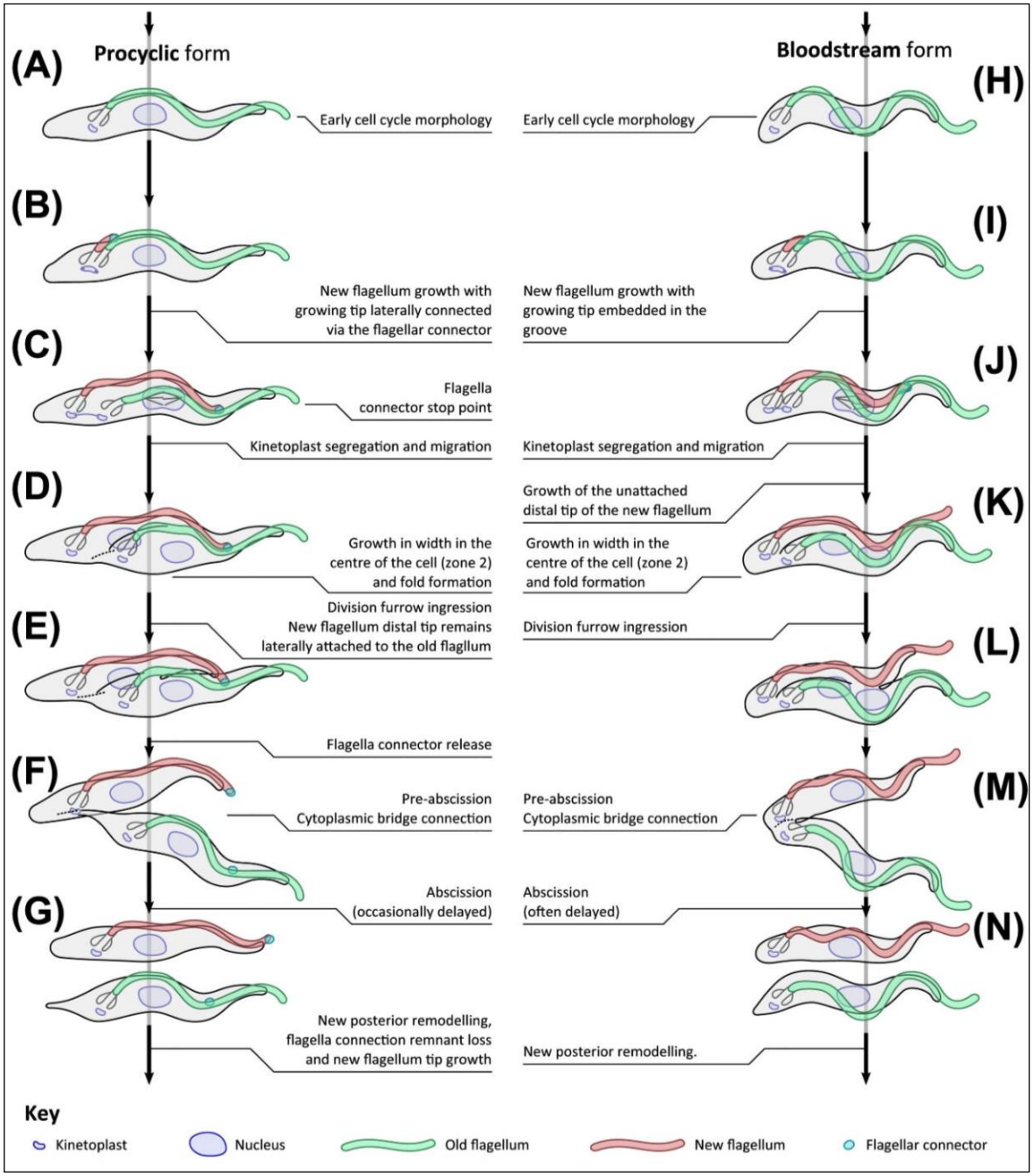


Figure 1-7. Comparative sequence of cell cycle events in procyclic and bloodstream trypomastigotes.
Adapted from (Gluenz *et al.*, 2015).

G1 cells have 1 nucleus and 1 kinetoplast (1N1K) but their relative position is different in bloodstream forms (BSF), where the kinetoplast sits at the terminal end of the cells; and in procyclic forms (PCF), where the kinetoplast is midway between the cell end and the nucleus. PCF are larger and wider than BSF, with the flagellum tip protruding at the apical end being shorter than BSF (Wheeler, Gluenz and Gull, 2013).

Kinetoplastid (kDNA) and nuclear DNA (nDNA) enter S-phase in an asynchronous but coordinated manner (Woodward and Gull, 1990). Initiation factors are different for kDNA, which replicates first and faster (Liu *et al.*, 2005; Jensen and Englund, 2012); and for nDNA (Tiengwe, Marques and McCulloch, 2014). When the nucleus enters S phase, the kinetoplast starts G2; and when the nucleus enters G2, basal bodies separate causing kinetoplast segregation (Liu *et al.*, 2005; Jensen and Englund, 2012). Golgi complex also segregates at this point (Hall *et al.*, 2006).

Before nuclear division, where chromosomes align and segregate by assembly of an evolutionary divergent kinetochore (Akiyoshi and Gull, 2014), cells have one nucleus and two kinetoplasts (1N2K) (Woodward and Gull, 1990). After chromosome segregation, upon mitosis, 2N2K cells are generated (Ogbadoyi *et al.*, 2000) but the relative position of nuclei and kinetoplasts vary from bloodstream and procyclic trypomastigotes: 1K1N1K1N in procyclic cells and 1K1K1N1N in bloodstream forms (Wheeler, Gluenz and Gull, 2013).

This accurate DNA replication pattern makes it easy to explore the time-wise implications of factors regulating the cell cycle with chemical or genetic tools (see below), as the K/N distribution phenotype precisely informs about the cell-cycle stage: 1N1K cells are in nuclear G1; cells with 1N2K are in nuclear G2; and cells with 2N2K have undergone mitosis but have not completed cytokinesis yet (Woodward and Gull, 1990; Siegel, Hekstra and Cross, 2008).

Whilst in G1, the new flagellum grows along the cell bringing with its base the new flagellar pocket and the new kinetoplast. In PCF cells, the tip of the new flagellum attaches to the old and slides towards the apical end up to a short distance from the top limit (Langousis and Hill, 2014). The two flagella only separate with cytoplasm/cell membrane invagination during cytokinesis (“furrow

ingression"). In BSF, the growing new flagellum is not connected to the old and protrudes free outside the cell, just before reaching the apical end (Hughes *et al.*, 2013), just before cytokinesis (Wheeler, Gluenz and Gull, 2013).

Cytokinesis starts immediately after mitosis. Unidirectional furrow ingression happens along the longitudinal axis backwards (Hammarton et al. 2007). The moment where the last cytoplasmic bridge breaks is known as abscission (Farr and Gull, 2012; Li, 2012; Wheeler, Gluenz and Gull, 2013).

1.4.3 Genome organization and gene expression

The haploid genome of *T. b. brucei* has 35 megabases (Mb) split between kDNA and nDNA although it can vary up to a 25% between different isolates (El-Sayed *et al.*, 2000; Berriman *et al.*, 2005).

1.4.3.1 Nuclear genome

The nuclear genome is organized in 11 megabase-size chromosomes (0.9-5.7 Mb), 2-4 intermediate size (300-900 kb), and about 150 minichromosomes (50-100 kb) (El-Sayed *et al.*, 2000; Berriman *et al.*, 2005), Figure 1-8. All of them have telomeres containing the same repeats (TTAGGG) observed in mammals (Dreesen, Li and Cross, 2006), which contain the bloodstream expression sites (ES) encoding VSG and expression site associated genes (ESAGs). Uniquely, these genes are transcribed by RNA polymerase I (Daniels, Gull and Wickstead, 2010; Ersfeld, 2011; Glover, Alsford and Horn, 2013; Morrison, McCulloch and Hall, 2014). *T. brucei* is the only known organism where RNA Pol I transcribes protein-coding genes (ESAG, VSG and procyclins) and not only rRNA genes (Günzl *et al.*, 2003; Haenni *et al.*, 2006, 2009).

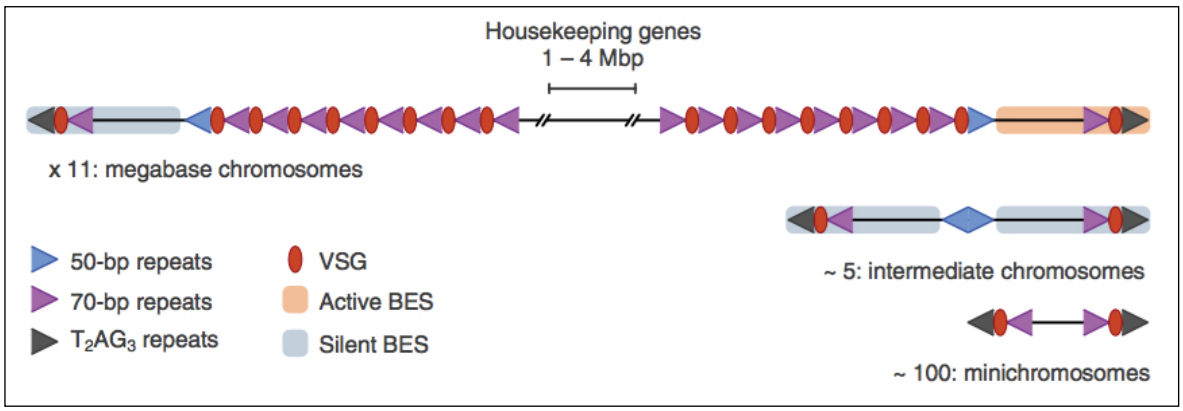


Figure 1-8. Schematic representation of nuclear genome organisation in the three types of chromosome.
Subtelomeres encode more than 2000 variants of VSG genes and pseudogenes, repetitive sequences found there are supposed to regulate monoallelic expression control and VSG swithching. Subtelomeres at the megabase chromosomes are hemizygous with the core region beingdiploid. There is only one active bloodstream expression site (BES) per genome. Adapted from (Glover et al. 2013).

Both, intermediate chromosomes and minichromosomes, encode *VSG* genes or pseudogenes that can recombine to generate novel "mosaic" VSGs, expanding the parasite´s repertoire (Glover, Alsford and Horn, 2013; Morrison, McCulloch and Hall, 2014). In total, more than 20% of the nuclear genome is devoted to generate antigenic variation (Berriman *et al.*, 2005).

The minichromosomes consist of a 20-80 Kbp-long palindromic core (made of 177 bp repeats), flanked by short subtelomeric regions containing non-repetitive DNA, and the telomeres (Wickstead, Ersfeld and Gull, 2004).

The megabase chromosomes were predicted originally to encode 9068 genes (Berriman *et al.*, 2005) but, after transcriptomic analysis, the figure expanded to 11425 (Siegel *et al.*, 2010). Their core regions contain ~7,500 housekeeping genes transcribed by RNA polymerase II (RNA Pol II), while the large hemizygous subtelomeres encode ~1,500 VSG genes and pseudogenes (Marcello and Barry, 2007; Cross, Kim and Wickstead, 2014). The centromeres have been mapped only for chromosomes 1 to 8 (Obado *et al.*, 2007). They colocalize with protein components of the divergent *T. brucei* kinetochore (Akiyoshi and Gull, 2014). A 500 Mb segmental duplication during evolution between chromosome 4 and 8 produced 75 genes uniquely duplicated in *T. b. brucei* (Jackson, 2007b).

House keeping genes are organized in polycistronic transcription units (PTU) that, unlike operons, directionally align head-to-tail genes that do not necessary belong to the same biological pathways (Daniels, Gull and Wickstead, 2010; Ersfeld, 2011; Jackson, 2014). Neighbouring PTUs are encoded in different DNA strands that can be convergent, divergent or head-to-tail in terms of their coding sense. The meeting point is known as strand switch region (SSR). Transcription, mediated by RNA pol II, is initiated from divergent SSRs (Daniels, Gull and Wickstead, 2010; Ersfeld, 2011). Transcriptional control is believed to be limited to codon adaptation (Horn, 2008) and gene duplication (Jackson, 2007b). Gene expression is mainly controlled at a post-transcriptional level. Several epigenetic modifications have been related with transcription initiation at head-to-tail or divergent SSR, i.e. histone H4 acetylated at lysine 10, histone H3 trimethylated at lysine 4 (Wright, Siegel and Cross, 2010) and histone H2 variants H2AZ and H2BZ (Siegel *et al.*, 2009). Termination is poorly understood though (Ersfeld, 2011), transcription is believed to stop at convergent SSR, which are rich in certain histone variants (H3V and H4V) and a chromatin modification of thymine, known as base J (Siegel *et al.*, 2009; Reynolds *et al.*, 2016). The tandem base J-H3V can be found both at the end of polycistronic units (transcribed by RNA pol II), and at subtelomeric regions silencing VSG variants (Reynolds *et al.*, 2016).

1.4.3.2 Mitocondrial genome (kinetoplast)

The mitochondrial genome (kinetoplast) consists of a complex, concatenated network of DNA including minicircles and maxicircles (Liu *et al.*, 2005). The maxicircles encode mitochondrial genes that, after transcription, require modification by RNA editing. The guide RNAs required for this process are encoded in the minicircles. RNA editing was first described in trypanosomes but is now known to occur in other organisms (Simpson *et al.*, 2004).

Evolutionary close dyskinetoplastic trypanosmes are *T. evansi* and *T. equiperdum,* for instance. *The* loss of the mitochondrial genome has been induced *in vitro* for *T. brucei* (Schnaufer, Domingo and Stuart, 2002; Dean *et al.*, 2013).

1.4.3.3 Post-transcriptional regulation

Polycistronic transcription units produced by RNA pol II are trans-spliced by addition of a 39 nucleotides-long spliced leader (SL) at the 5´end of each monocistronic mRNA (Michaeli, 2011); and polyadenylation, which is still not well understood (Clayton and Michaeli, 2011). The SL is the only sequence for which a RNA pol II promoter has been found in *T. brucei* (Das *et al.*, 2005; Schimanski, Nguyen and Gu, 2005). Together with translational regulation, processes determining the fate of the spliced transcripts are the main mechanisms controlling expression levels: mature RNA degradation in the nucleus, export to the cytoplasm, and degradation in the cytosol (Clayton, 2014). RNA decay or translation depends on a diverse network of RNA binding proteins that attach to the 3´ or 5´untranslated regions contributing to promote or prevent recruitment of the RNA degradation machinery, therefore disrupting translation (Clayton, 2014). Upon stress conditions, such as transmission to a new host, mature RNAs can be sequestered from the translation and degradation machinery in “stress granules” (Clayton, 2014).

1.5 Protein kinases

Protein kinases are the best-studied set of signalling proteins in the eukaryotic cell. They catalyse the transfer of the γ-phosphate of a purine nucleotide triphosphate (ATP or GTP) to the hydroxyl group of a substrate, establishing a monoester bond. Such modification alters the conformation of the substrate peptide affecting its activation and interaction networks.

Protein kinases play fundamental roles as central regulators in many intracellular biological functions including metabolism, transcription, cell cycle progression, cytoskeletal rearrangement, cell movement, apoptosis, and differentiation. But in complex organisms they regulate also intercellular communication during development, physiological responses, homeostasis, or in the functioning of the nervous and immune systems (Manning et al. 2002). Their dysregulation contributes to the aetiology of many medical conditions of which many forms of cancer are prominent representatives. They are the largest family of druggable targets in the proteome (Hopkins and Groom, 2002).

1.5.1 Classification

Protein kinases group into two main subdivisions: protein-serine/threonine kinases and protein-tyrosine kinases. In addition, the less abundant dual specificity kinases can phosphorylate both. A first in-depth analysis and classification was reviewed in the pre-genomic era (Hanks and Hunter, 1995) that was updated later with the publication of the kinome complement of the human genome (Manning et al. 2002).

The human kinome (Figure 1-9) is classified in 478 eukaryotic (ePKs) and 40 atypical protein kinases (aPKs). ePKs are divided in 9 broad groups comprising 184 families, 51 conserved also in flies, worms and yeast (Manning et al. 2002). There are some exclusive families and others asymmetrically expanded between different organisms, e.g. humans have 14 *Eph* receptor tyrosine kinases (RTKs) while *Drosophila melanogaster* only has one. 25 subfamilies have one single member in each of the 4 mentioned organisms, suggesting unduplicated functions. 10% are "pseudokinases" missing some of the key residues required for catalytic activity but conserving all the other features of the catalytic domain that will be described in the next section (Kroiher, Miller and Steele, 2001; Morrison, 2001; Zervas and Brown, 2002). Some of them might still function by non-canonical mechanisms but they mainly modulate other catalytic domains or have other implications in protein-protein interaction. In addition, 83 different alternative domains have been identified in human protein kinases (SH2, GTPase, lipid signalling, Ca^{2+} signalling…). Some of them regulate protein kinase activity, some link them to other signalling modules, and others assist protein localization. Finally, pseudogenes encoding disrupted protein kinase genes are also found in the genome, probably due to viral retrotransposition of processed transcripts.

Group	Families
AGC	14
CAMK	17
CK1	3
CMGC	8
Other	37
STE	3
Tyrosine kinase	30
Tyrosine kinase–like	7
RGC	1
Atypical-PDHK	1
Atypical-Alpha	1
Atypical-RIO	1
Atypical-A6	1
Atypical-Other	7
Atypical-ABC1	1
Atypical-BRD	1
Atypical-PIKK	1
Total	134

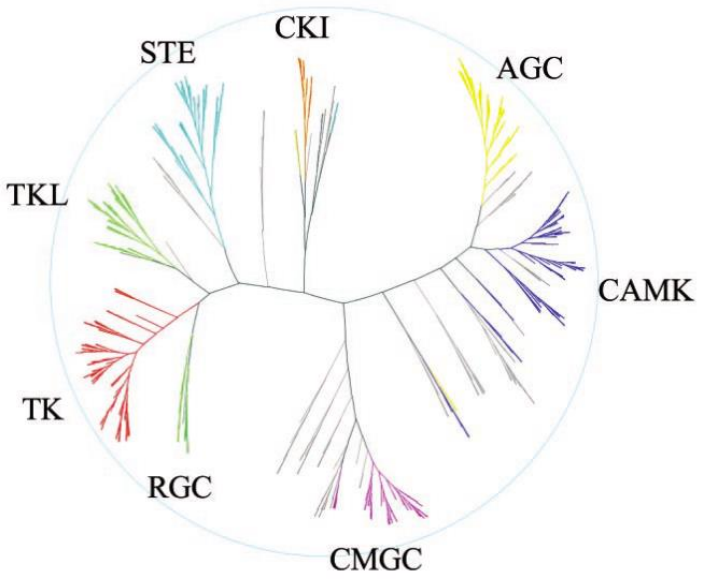


Figure 1-9. The human protein kinome.
Left hand panel: table with major groups of ePKs and aPKs, indicating the number of families that they contain. Right hand panel: phylogenetic dendrogram of 491 ePK human domains. Major groups are highlighted in colours. Modified from Manning et al. (2002).

1.5.2 The protein kinase catalytic domain: structure and functions

The protein kinase catalytic domain plays three fundamental roles: binding and orientation of the substrate peptide; binding and orientation of the ATP (or GTP) as a complex with a cation (Mg^{2+} or Mn^{2+}); and transfer of the γ-phosphate from the ATP (or GTP) to the acceptor hydroxyl group of a serine, threonine or tyrosine residue. There is no need for binding first substrate or ATP, but there is a thermodynamic preference for the nucleotide (Cook *et al.*, 1982; Grant and Adams, 1996). The phosphorylated residue in the substrate peptide, "phosphosite" (P-site), needs to be framed by a consensus recognition sequence with affinity for the catalytic site (Chan, Hurst and Graves, 1982; Pearson and Kemp, 1991).

The protein kinase catalytic domain structure (250-300 amino acids) is common to the entire eukaryotic tree including animals, plants, fungi and protozoa (Taylor and Radzio-Andzelm, 1994). It has 12 subdomains characterised by patterns of conserved residues nearly invariant throughout the protein kinase superfamily, these subdomains are never interrupted by large amino acid

insertions and fall into topologically similar core structures. The variable non-conserved insertions build up non-catalytically relevant features of the protein surface. The non-catalytic domain can modulate activity interacting with the protein kinase core, or binding other proteins that enhance or repress the catalytic function (Pawson and Gish, 1992; Newton, 1997; Sicheri and Kuriyan, 1997; Brushia and Walsh, 1999). Other factors affecting activation are fatty acid acylation (Blenis and Resh, 1993; Johnson and Cornell, 1999), isoprenylation (Inglese and Premont, 1996; Pitcher, Freedman and Lefkowitz, 1998), second messengers (Rasmussen, 1989; Liscovitch and Cantley, 1994; Lucas *et al.*, 2000), subcellular localization (Pawson and Scott, 1997; Garrington and Johnson, 1999; Scott and Pawson, 2000), and phosphorylation (Shenolikar, 1988; Hunter, 1995).

The first crystal structure of a catalytic protein kinase domain available was a mouse PKA with a substrate (Ben-David *et al.*, 1991; Rossomando *et al.*, 1992), soon obtained also with the complexes ATP-Mn^{2+} and ATP-Mg^{2+} (Pearson and Kemp, 1991). As new structures of other protein kinases were published it was confirmed that the topology of the catalytic domain was well conserved in eukaryotes. It is an overall two-lobe structure with a small NH_2-terminal lobe comprising subdomains I-IV, and a large COOH-terminal lobe including domains VIa –XI (Figure 1-10A and C). The hinge region is constituted by subdomain V, which protrudes towards both lobes creating the catalytic cleft. The NH_2-terminal lobe folds predominantly in antiparallel β-sheet secondary structure and its main function is anchoring and orientation of the ATP (or GTP). The COOH-terminal lobe is mainly α-helix configuration and is responsible for the phosphotransfer initiation. While the ATP nests deep into the pocket, the substrate remains in the periphery (Adams, 2001).

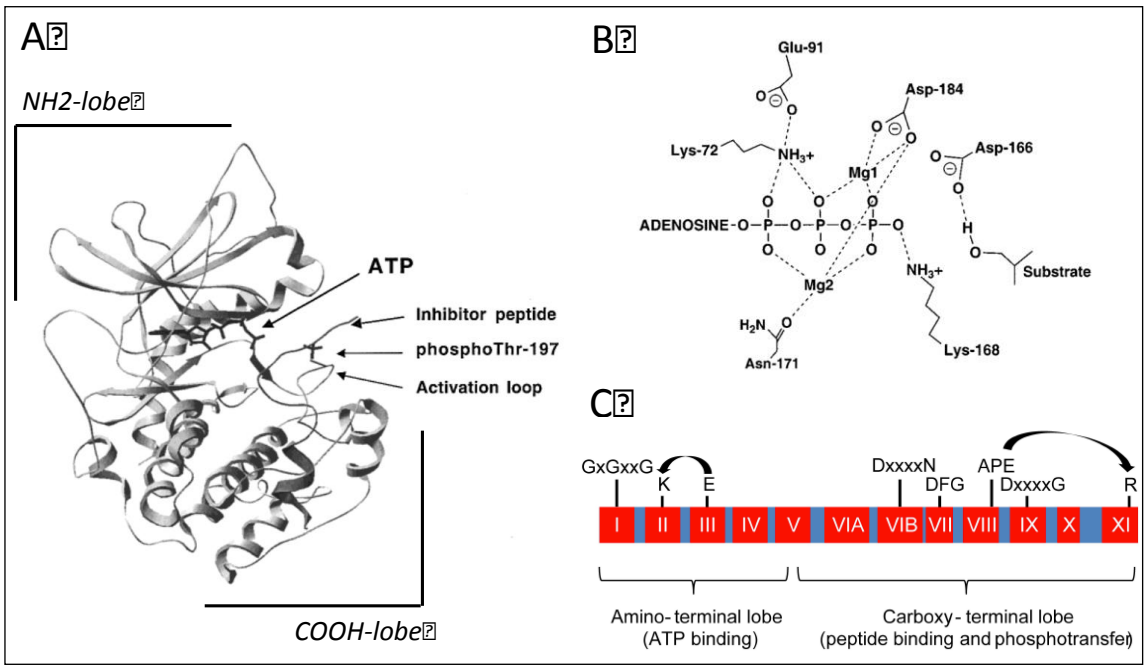


Figure 1-10. Protein kinase catalytic domain.
A. Ribbon diagram of PKA co-crystallized with ATP and a peptide inhibitor. NH_2-terminal lobe (mainly β-sheet) and COOH-terminal lobe (predominantly α-helix) are highlighted. Arrows point to the activation loop and its P-site, pThr-197, required for protein kinase activation. **B.** Key residue interactions in the active site of PKA. Mg1 and Mg2 represent the activating and inhibitory divalent metal ions, Mg2+, respectively (schematic representation based on A). Dotted lines indicate close contacts under 2.6 Å but are not drawn to scale. The hydroxyl of the substrate is shown based on a binary complex of PKA with a substrate peptide. **C.** ePK catalytic domain indicating the 12 conserved subdomains (Roman numerals). Positions of amino-acid residues and motifs highly conserved throughout the ePK superfamily are indicated above, using single-letter amino-acid code with x as any amino acid. A and B, adapted from Adams 2001. C, adapted from Hanks 2003.

The structure-activity relationship of the different conserved motives has been reviewed in the literature (Hanks and Hunter, 1995; Adams, 2001). Using as reference residues in PKA, subdomain I constitutes a clamp anchoring the non-transferable phosphates on the ATP through H-bonds with Ser53, Phe54 and Gly55. Leu49 and Val47 form a hydrophobic space to fit the adenine ring. There is a consensus sequence throughout most of protein kinases at this domain: Gly50-x-Gly-x-x-Gly-x-Val.

The main feature at subdomain II is Lys72, a residue essential for maximal activity whose substitution is commonly used to produce inactive mutants (Hixson and Krebs, 1979; Hathaway, Zoller and Traugh, 1981; Buhrow *et al.*, 1983; Weinmaster, Zoller and Pawson, 1986). It does not affect ATP binding but

is critical for correct orientation of the nucleotide. Lys72 forms a salt bridge with the invariable Glu91, which helps orientation and anchoring of ATP in the large α-helix of subdomain III, Figure 1-10B.

Subdomain IV and VIa are apparently not involved in catalysis and do not have any invariable residues. Subdomain V articulates the hinge region with half α-helix and half β-sheet. The consensus Glu121, Val123, and Glu127, in the joint, anchor ATP forming H-bonds with adenine and the ribose ring. Met120, Tyr122 and Val123, also help forming a hydrophobic pocket. Glu127 forms in addition an ion pair with an arginine in the peptide recognition sequence.

Subdomain VIb contains the "catalytic loop", the consensus His164-Arg-Asp-Leu-Lys-x-x-Asn. Asp166 may accept the proton from the substrate hydroxyl group, while Lys168 (sometimes an Arg) neutralize the negative charge of the phosphate during the transference. Asp166 also stabilize the loop establishing a H-bond with Asn171, which chelates the Mg^{2+} that bridges α- and β-phosphates, Figure 1-10B. Glu170 forms a H-bond with the ribose and an ion pair with an Arg in the peptide recognition consensus.

Subdomain VII has the consensus Asp184-Phe-Gly. Asp184 is essential as the kinase is not functional when replaced (Gibbs and Zoller, 1991). It forms an H-bond with Gly186 and chelates Mg^{2+} between α- and β-phosphates, assisting ATP orientation, Figure 1-10B.

Containing many peptide binding residues and a hydrophobic pocket made of Leu198, Cys199, Pro202 and Leu205, subdomain VIII has the characteristic consensus Ala206-Pro207-Glu208 (APE). Frequently this domain contains residues that require phosphorylation before the protein kinase becomes active (Chan, Hurst and Graves, 1982; Luo, Zhou and Lodish, 1995; Songyang *et al.*, 1995).

Subdomain IX, contains the conserved Asp220 that forms H-bonds with Arg165 and Tyr164 in the "catalytic loop" contributing to its stability and to substrate recognition. Subdomain X is the less conserved, with all CMGC kinases presenting large insertions. Finally, subdomain XI has the invariable His-x-Aromatic-Hydrophobic-10x-Arg280. This last residue forming an ion pair with Glu208 in the APE motive of subdomain VIII.

1.5.3 Protein kinases as drug targets

Protein kinases are a promising source for druggable targets with more than 100 inhibitors already in clinical trials (Naula, Parsons and Mottram, 2005; Barouch-bentov and Sauer, 2012) and successful drugs in the market, such as Imatinib® for chronic myelogenous leukemia (Deininger and Druker, 2003). It was estimated that in the 2000s over 30% of drug discovery programmes focused in kinase inhibitors (Pinna, 2003). Protein kinases represent the largest subset of druggable targets in the human proteome (22%) (Hopkins and Groom, 2002).

Given the high conservation of the protein kinase domain, one of the main challenges is achievement of kinase selectivity (Karaman *et al.*, 2008). Imatinib® was originally targeted against PGDFR but effectiveness has been later associated to BCR-ABL and cKIT inhibition (Lorusso and Eder, 2008). While inhibition of a single kinase may not be effective, targeting at once a number of them can have positive synergistic effects. If the objective is finding an efficient drug, finding selectivity with respect to the host is a more critical aspect than finding selectivity for a particular kinase. That is why target-based assays with the purified protein are losing priority compared to cell-based phenotypic screens. Intra-organism protein kinase-specificity can be crucial only if the objective is identification of a tool compound and it may require functionalization that increase molecular weight, risking its drug-like attributes (Gill *et al.*, 2007).

The prototypical kinase inhibitors are aromatic heterocycles with H-bond donor/acceptor motifs. The most effective ones fit easily into the protein kinase lipophilic pocket, so optimal physicochemical properties for inhibition challenge solubility and absorption properties. In order to circumvent this problem, chemical functionalization of the molecules with planarity-breaking and hydrophilic motives has been undertaken (Merritt *et al.*, 2014).

The increasing availability of structural data describing the kinase-substrate geometry, including orientation and key contacts for activity, is informing medicinal chemistry to develop more specific kinase inhibitors based on the structure-activity relationship (SAR) (Merritt *et al.*, 2014).

Two types of protein kinase inhibitors interact directly with the catalytic site in the active conformation. Type I (called DFG-in, due to position of the conserved triad at subdomain VII) compete directly with ATP for binding. Type II (DFG-out) bind in the surroundings of the ATP pocket but do not compete with the nucleotide for binding. Allosteric inhibitors (type III or type IV, depending on their distance to the catalytic cleft) change enzyme conformation preventing activation or blocking activity. As a general rule, non-competitive inhibitors show higher potencies because ATP concentrations are very high in the cells (1-10 mM), (Liu and Gray, 2006; Knight and Shokat, 2007)

1.5.4 Protein kinases in *Trypanosoma brucei*

Bioinformatics analyses based on multilevel Hidden Markov Model sequence alignment of the catalytic domain followed by manual curation showed 2% of the *Trypanosome brucei* proteome to have putative protein kinase activity (Parsons *et al.*, 2005; Nett *et al.*, 2009). Although other organisms, such as *Toxoplasma gondii*, possess a larger number of predicted protein kinases (356 protein kinases), the proportion they represent of the total proteome is comparatively smaller (1.5%). In absolute numbers, the *T. brucei* protein kinome is ~1.5 times larger than that of *Saccharomyces cerevisiae* or *Schizosaccharomyces pombe* (Miranda-Saavedra *et al.*, 2007), ~2 times the size of the one of the malaria parasite (Ward *et al.*, 2004), or ~1.7 times the protein kinase complement of the intracellular parasite *E. cuniculi* (Miranda-Saavedra *et al.*, 2007). This means that *T. brucei* exhibits a relative expansion on the protein kinome when compared with many other single celled eukaryotes.

T. brucei encodes 170 eukaryotic protein kinases (ePKs), 13 atypical protein kinases, and 12 pseudo-kinases predicted to be non-functional as they miss some of the canonical residues required for activity (Parsons *et al.*, 2005; Nett *et al.*, 2009; Urbaniak *et al.*, 2012; Jones *et al.*, 2014).

It is remarkable the lack in the parasite of receptor guanylate cyclases (RPC), which bear a protein kinase-like domain and are considered an evolutionary innovation of late metazoans; receptor and non-receptor tyrosine kinases; and tyrosine kinase-like kinases, when compared to other eukaryotes. Tyrosine phosphorylation has been reported though (Nett *et al.*, 2009). Phosphotyrosine

residues have been identified in several CMGC protein kinases and, at least, some of them are very likely due to conserved phosphorylation events conducted by dual specificity kinases such as Wee1, DYRKs, or GSK3. Wee1 mediates a tyrosine phosphorylation event in human CDK1 which is conserved in trypanosomes (Krupa, Preethi and Srinivasan, 2004; Nett *et al.*, 2009); the last two protein kinase types require autophosphorylation at a tyrosine to become active threonine/serine kinases and these phosphotyrosines have been identified by proteomics (Kentrup *et al.*, 1996; Cole, Frame and Cohen, 2004; Lochhead *et al.*, 2005).

T. brucei has relatively expanded the CMGC, STE and NEK families of kinases, which contain several highly divergent proteins that are very likely to play parasite-specific functions (Parsons *et al.*, 2005; Nett *et al.*, 2009).

Reversible phosphorylation is one of the main signalling mechanisms in the eukaryotic organisms and has been linked to regulation of most cellular processes (Cohen, 2000). Relevance of this signalling applies also to *T. brucei*, where 491 distinct phosphorylated proteins have been found in the cytosolic fraction of the parasite's bloodstream form. 75% of these were modified at serine residues, 21.5% at threonine, and 3.5% at tyrosine residues (Nett *et al.*, 2009). It is interesting that this distribution (75/21.5/3.5) is closer to prokaryotic organisms (70/20/10) than to human HeLa cells (86/12/2), for instance (Usacheva *et al.*, 2003; Macek *et al.*, 2007; Nett *et al.*, 2009).

Protein kinase involvement in intricate phosphorylation-dephosphorylation networks is also conserved in *T. brucei* as 44 protein kinases were found to be phosphorylated, 50% in more than one site (Nett *et al.*, 2009). Domains for phosphopeptide recognition are abundant in the *T. brucei* proteome, e.g. 14-3-3 proteins that bind to specific phosphoserine and phosphothreonine motifs (Inoue *et al.*, 2005). Among the protein kinases, there is a CAMK with a forkhead-associated domain (FHA), which binds phosphothreonine (Parsons *et al.*, 2005); and there is a polo-like kinase (TbPLK) with a phosphoserine/threonine-binding polo box domain (Kumar and Wang, 2006; Hammarton *et al.*, 2007).

Remarkably, phosphotyrosine-binding domains present in other eukaryotes such as Src homology 2 (SH2) or phosphotyrosine-binding (PTB), are missing in *T.*

brucei consistent with the lack of tyrosine kinases or tyrosine-like kinases (Nett *et al.*, 2009).

One of the best described phosphorylation cascades are MAPKs, conveying mainly extracellular signals and activating transcription factors (Widmann *et al.*, 1999). The canonical three-step module includes STEs (STE7s, STE11s and STE20s), one of the proportionally expanded groups of protein kinases in trypanosomes compared to other eukaryotes. Many of these protein kinases are phosphorylated in *T. brucei* (Nett *et al.*, 2009), but the unconventional transcription regulation in the parasite may suggest non-canonical functioning. In fact, MAPK cascades are missing in other protozoan parasites, such as those from the phylum Apicomplexa -as *Plasmodium* or *Toxoplasma*- or *E. cunicula* (Miranda-Saavedra *et al.*, 2007).

As protein kinases are central regulators of many cellular functions and have been proven good drug targets in humans, there are many studies assessing essentiality of these molecules in *T. brucei* with the prospect of targeting them in drug discovery programmes.

1.5.5 Approaches undertaken for functional analysis of protein kinases in kinetoplastids

Different approaches have been undertaken to determine the role of protein kinases in *T. brucei* and fellow kinetoplastid parasites, in an individualized manner or in high throughput campaigns. They are based in genetic manipulation, chemical inhibition or both.

1.5.5.1 Genetic approaches

At a DNA level, gene deletion of the two alleles encoding a particular protein (null mutant) has been extensively used to analyse the effects it produces in the cellular phenotype. For instance, MKK1, a MAPK, was determined not to be essential for bloodstream form fitness in culture by this method, however the null mutant was unable to colonise successfully the salivary glands in the tsetse fly indicating MKK1 relevance for progression of the life cycle (Morand *et al.*, 2012). A protein is considered potentially essential when the null mutant (double knockout) is impossible to generate, this happened analysing CK1 isoform 2

where the second allele was impossible to remove (Urbaniak, 2009). Yet, it raises concerns that the reason for such impossibility may be technical. This has been solved through development of conditional null mutants where deletion of the second allele is possible by inducible expression of the gene from an exogenous locus (Milne, Güther and Ferguson, 2001). Once the two endogenous alleles are removed, withdrawal of the induction factor generates the null mutant (Merritt and Stuart, 2013; Damerow *et al.*, 2016).

Another approach -that will be discussed extensively in the introduction to chapter 3 is RNA of interference (RNAi). Small, double stranded RNA molecules with homology to a gene of interest, can be expressed that bind the gene´s transcripts targeting them for degradation mediated by the RNAi machinery and preventing protein production (Ullu, Tschudi and Chakraborty, 2004). The system has been developed for inducible transcription of the RNAi probe (Alsford *et al.*, 2005) and, coupled to next generation sequencing (RNAi target sequencing), has been exploited for parallel phenotyping of complete genomes (Alsford *et al.*, 2011). Since its implementation in *T. brucei*, this system has been extensively applied for individualized gene studies and for high throughput screens (Glover *et al.*, 2014; Jones *et al.*, 2014).

Since *T. brucei* protein kinases are druggable (Urbaniak *et al.*, 2012), specific inhibition can be potentially achieved when a tool compound is available. Targeting a particular kinase with a compound can provide almost immediate inhibition (Knight and Shokat, 2007). To my knowledge not a single tool compound specific against a particular *T. brucei* protein kinase has been identified yet.

Combining chemical and genetic approaches, analogue-sensitive strategies have been developed where the conserved large gatekeeper residue at the ATP binding pocket of a selected kinase has been mutated to glycine or alanine (Knight and Shokat, 2007). This genetic modification permits ATP analogues bulky at the N6 position to fit in the pocket and cause specific inhibition of the mutant protein. It has been used in *T. brucei* against TbPLK (Lozano-Núñez *et al.*, 2013).

1.5.6 Protein kinases found with roles characterised throughout the above described approaches

Pointing in this direction, several studies involving RNAi or null mutant phenotyping have been conducted reporting many protein kinases to be essential for survival of the parasite or whose depletion caused a loss of fitness in the culture dish.

1.5.6.1 Essential protein kinases detected in screening campaigns

CRK12 and ERK8 were identified essential for *T. brucei* bloodstream forms with a luciferase-based RNAi screen targeting 30 protein kinases (Mackey *et al.*, 2011), CRK12 was also detected in a conditional null mutant screen with 20 ePKs belonging to the "STE" and "Other" families. GSK3-short and CLK1 were reported in an RNAi screening of the 21 *T. brucei* "CDXG"-motif-bearing kinases (binding platform for hypothemicin, a trypanocidal compound) (Nishino *et al.*, 2013), although GSK3-short had been reported a potential drug target through RNAi and compound screen targeting the human homolog before (Ojo *et al.*, 2008). More recently, all the 180 predicted protein kinases of *T. brucei* were systematically screened for cell viability in Alamar blue® using inducible RNAi identifying 41 protein kinases required for normal fitness *in vitro* (Jones *et al.*, 2014). A previous whole-genome screen based in RNAi target sequencing (RITseq) assessing cell growth detected depletion of 42 protein kinases detrimental for normal growth of the bloodstream forms in cultured conditions (16 overlapping with the kinome-wide Alamar blue screen) (Alsford *et al.*, 2011).

1.5.6.2 Protein kinases with roles in the parasite´s life cycle

Several protein kinases have been identified as critical for progression along the life cycle of the parasite. For instance, ZFK (Vassella *et al.*, 2001), MAPK5 (Domenicali Pfister *et al.*, 2006) and TbTOR4 (Barquilla *et al.*, 2012) have been identified as negative regulators of the slender-to-stumpy differentiation in *T. brucei* bloodstream forms while a genome-wide RNAi screen under stumpy-inducing conditions discovered positive regulators (whose depletion blocked this developmental transition): MEK kinase (Tb927.2.2720), NEK kinases Tb927.10.5930/40/50 (undistinguishable in this approach) and Dyrk/Yak kinase (Tb927.10.15020) (Mony *et al.*, 2013). A systematic RNAi screen of the protein

kinome of *T. brucei* unveiled involvement of 2 protein kinases (RDK1 and RDK2) in repression of bloodstream to insect-infective procyclic form differentiation (Jones *et al.*, 2014) while different studies proved NRK (Tb927.4.5390/8.6930) -a protein kinase enriched in stumpy forms- to be a requirement for procyclic differentiation (Alsford *et al.*, 2011; Domingo-Sananes *et al.*, 2015). The best described pathway implicated in this transition is a phosphatase cascade (Szöor *et al.*, 2006, 2010), inferring requirement of uncharacterized protein kinases antagonizing this regulation. A whole-genome RNAi parallel phenotyping experiment based in RNAi target sequencing, identified 16 protein kinases whose depletion did not cause a loss of fitness in bloodstream parasites but were required for normal fitness upon procyclic differentiation (Alsford *et al.*, 2011). Within the insect vector, two other protein kinases, MKK1 (Morand *et al.*, 2012) and AK3 (Ooi *et al.*, 2015), have been reported necessary for progression of the life cycle.

1.5.6.3 Cell cycle regulation

Another interesting feature of *T. brucei* biology also regulated by many protein kinases is a complex cell division cycle that requires accurate segregation of many single copy organelles and has different progression checkpoints compared to the mammalian host (Ploubidou *et al.*, 1999; Hammarton *et al.*, 2003; Hammarton, 2007). Divergence of *T. brucei* cell cycle regulation could provide a unique source of potential drug targets where many protein kinases play essential roles (Jones *et al.*, 2014). An evolutionary divergent kinetochore (highly conserved protein complex in most of Eukaryotes that is required for chromosome segregation) has been described for *T. brucei* identifying four protein kinases associated with the structure (KKT2, KKT3, KKT10 and KKT19) (Akiyoshi and Gull, 2014). RNAi of KKT3 and KKT10 caused loss of fitness *in vitro* (Alsford *et al.*, 2011; Jones *et al.*, 2014)

Based in homology with other eukaryotes, 12 *T. brucei* cyclin-dependent kinases orthologues, key cell cycle regulators in other eukaryotes, were identified. They were called in *T. brucei* CDC2-Related kinases (CRKs) (Mottram and Smith, 1995; Hammarton *et al.*, 2003).

Early experiments performing RNAi-mediated depletion showed CRK1 and CRK3 to be involved in G1/S and G2/M transition respectively (Tu and Wang, 2004); AUK1 and TLK1 to be required for mitosis (Li and Wang, 2006; Li, Gourguechon and Wang, 2007; Jetton *et al.*, 2009); and PLK1, PK50 and PK53 regulating cytokinesis (Hammarton *et al.*, 2007; Ma *et al.*, 2010). More recently, a systematic RNAi screen of the *T. brucei* protein kinome identified 20 other protein kinases required for normal progression along different stages of the cell division cycle (Jones *et al.*, 2014).

1.5.7 Kinase targets identified by small molecules

1.5.7.1 Covalent kinase inhibitors

21 *T. brucei* kinases contain a cysteine close to the DXG catalytic domain responsible of phosphate orientation in subdomain VII, "CDXG" kinases. Hypothemycin (Nishino *et al.*, 2013) is a trypanocidal antifungal compound, which establishes a covalent bond with that particular cysteine. Mass spectrometry identified in cell lysates 12 CDXG kinases that bind to hypothemycin, 2 of them were essential for the parasite survival upon RNAi. One of them, CLK1, was a new discovery.

1.5.7.2 Drug-mediated elution from an ATP-sepharose matrix

T. brucei cell lysates were passed through a column charged with an ATP-sepharose matrix so protein kinases were retained on them. After washing the columns, kinases were eluted by addition of 5 compounds developed against EGFR/VEGFR in humans. They competed with ATP for the active site of certain protein kinases that could be successfully eluted. 5 protein kinases (TbLBPK1-4 and TbCBPK1) were found in eluates obtained with the 3 inhibitors of human EGFR and/or VEGFR with trypanocidal activity: lapatinib, canertinib and AEE788 (Katiyar *et al.*, 2013).

1.5.7.3 Drug-mediated elution from a kinase inhibitor affinity matrix: "kinobeads"

A system developed by Cellzome (Bantscheff *et al.*, 2007), was applied to trypanosomes. A matrix composed by beads covered by immobilized pan-kinase inhibitors was exposed to *T. brucei* bloodstream form lysates in the presence or

absence of the screened kinase inhibitors: staurosporine, BMS-387032 and 3 hits each coming from target-based screens conducted against PK50, PK53 and GSK3-short (Urbaniak *et al.*, 2012). Detachment and quantification of protein kinases retained by the beads permitted comparisons of results obtained in the presence or absence of the inhibitors. Those protein kinases with specific affinity for the inhibitor tested would be competitively eluted from the kinobeads. Beyond identification, this technique permitted titration of the compound´s inhibitory activity and the experiment unveiled that protein kinases of *Trypanosoma brucei* are druggable.

1.5.8 Medicinal chemistry applied to *T. brucei* protein kinases

An ideal kinase inhibitor used in HAT treatment should have oral bioavailability and parasite penetrance. If expected to be efficient in stage 2 (meningo-encephalitic), it would require the ability to cross the blood-brain barrier (Wager, Chandrasekaran, *et al.*, 2010), which is linked to a limited molecular weight and certain restrictions considering polarity/lipophilic profiles. Extracellular nature of *T. brucei* results an advantage, as the potential compound does not need to penetrate the host cells in addition.

According to the Drugs for Neglected Diseases initiative (DNDi) the HAT target product profile (TPP) define 2 complementary strategies. TPP1 would be a safe, effective and practical stage 2 compound ideally active against both stages of the disease; and TPP2 a simple stage 1 treatment amenable to be used in mass screening and treatment campaigns in endemic areas. Ideally both compounds should be multitarget and able to treat both *gambiense* and *rhodesiense* infections. TPP1 should be administered in oral treatments under one week. Costs should not exceed 30€/course. TPP2 should also consist in a single daily dose, less than 3 days of treatment. Costs should be under 10€/course (Drugs for Neglected Diseases initiative, 2015).

1.5.8.1 Glycogen Synthase Kinase-3 short isoform (GSK3-short)

GSK3-short is the *T. brucei* protein kinase that has focused more efforts in screening campaigns looking for protein kinase inhibitors. Also investigated in antileishmaniasis programmes, is a perfect example of “piggy back” drug

discovery where a validated drug target in one organism is exploited in a different one (Ojo *et al.*, 2011). Two isoforms of GSK3 have been found in the *T. brucei* genome but RNAi of GSK3-short is the most deleterious (Ojo *et al.*, 2008). Analysis of the molecular model showed that substantial differences with the mammalian orthologue in the ATP binding pocket would permit designing compounds capable of selective inhibition in the parasite respect to the host. In the same work, a set of 48 commercially available protein kinase inhibitors and a collection of 255 known human GSK3β inhibitors were screened against the recombinant protein *ex vivo* and the parasite in culture. In the first set, 5 compounds were identified with acceptable efficacy both against the protein and *in vivo*. 3 showed lower EC50s in the parasite than in the protein *ex vivo*, suggesting that compounds may inhibit more than one target. In the GSKβ focused library they found better correlations between IC50s and EC50s.

Next, 16549 Pfizer compounds were screened against both human and *T. brucei* GSK3 proteins. 362 were found with efficacy under 1 μM (35 under 100 nM) and, among them 2 had 7-fold selectivity for the parasite protein kinase (Oduor *et al.*, 2011). Regardless of selectivity, 17 of the most active were screened against the parasite and the human cell in culture. 6 were identified with more than 6-fold selectivity for *T. brucei*. As a general rule there was a correlation between efficacy found *in vivo* and *ex vivo,* however a few exceptions were found due to problems of membrane permeability. In order to test possible multitarget inhibition, 13 were screened against a set of 40 human protein kinases with 1 showing high selectivity but no penetrance in the parasite cell; 1 inhibited only 2 protein kinases but showed no preference between human and parasite; and 1 inhibited 4 protein kinases for which human/parasite selectivity could not be tested. Each of the rest inhibited more than 10 human protein kinases.

A different screen with 4110 compounds, based in luminescence (Woodland *et al.*, 2013), detected many non-specific toxins and 11 compounds with activity against TbGSK3-short. Only one showed good EC50 values but it did not have selectivity for *T. brucei* compared to humans. In addition, GSK3-short was among the protein kinases non-specifically inhibited by two of the compounds tested in the kinobeads assay mentioned above (Urbaniak *et al.*, 2012). So far promising

leads have been appointed against this *T. brucei* target but they did not have a good selectivity-efficacy balance.

1.5.8.2 High throughput screening campaigns including kinase inhibitors

Several high throughput screening campaigns of compound libraries have been published linking trypanocidal activity with kinase inhibition (Diaz *et al.*, 2014; Peña *et al.*, 2015; Woodland *et al.*, 2015). They will be discussed in more detail in the introductory section of chapter 5.

1.5.8.3 Drug repurposing strategies

The wealth of knowledge available for human protein kinases, and their degree of conservation have been conceptually exploited in research of neglected tropical infectious diseases. Repurposing molecules that are already in the pipeline of drug discovery for other medical conditions can compensate for the lack of investment in diseases that affect populations in the poorest areas of the planet. This is especially relevant for compounds tailored to treat chronic diseases as cancer, which focuses most of the kinase inhibitor drug discovery campaigns. Compounds discarded due to toxic side effects caused during long lasting treatments, might not be such of a problem using the short administration patterns required for infectious diseases. In contrast, problems with acute toxicity must be taken into consideration (Merritt *et al.*, 2014).

Aurora kinase inhibitors

Aurora kinases are well known eukaryotic protein kinases that play critical roles in regulation of cell division. Their depletion disrupts formation of the mitotic spindle, chromosome segregation, and cytokinesis. Several inhibitors are in clinical trials for cancer. *T. brucei* has 3 aurora kinase genes, AUK1 is required for correct nuclear division, cytokinesis and normal growth both *in vitro* and *in vivo* (Li and Wang, 2006). AUK2 transcript depletion caused also loss of fitness in a whole-genome RNAi screen (Alsford *et al.*, 2011).

A couple of inhibitors of the human kinase were able to kill the parasite with good EC50 values but without selectivity *versus* human cells (Jetton *et al.*, 2009; Li, Umeyama and Wang, 2009). A rational modification of them based on their

structure-activity relationship (SAR) led to high efficacy with >300-fold selectivity over human cells (Patel *et al.*, 2014). Some of these analogues were also valuable against *Leishmania* parasites, although inactive against *T. cruzi*. SAR in the molecular model was also applied in two well-profiled molecules (one in clinical trials for cancer and its precursor) that had anti-trypanosome EC50s in the low micromolar range (Ochiana *et al.*, 2013). One compound showed good selectivity over human, and a good correlation was found between predicted affinity *in silico*, and efficacy over the parasite in culture. Compounds have not been tested in the recombinant protein yet.

Phosphoinositide-3 kinases and mTOR inhibitors

PI3K and mTOR play fundamental roles in growth and metabolism in human cells. They have been exploited as drug targets for anticancer therapy and in for treatment of inflammatory disease. There are inhibitors well advanced in the pipeline of drug discovery able to target these two protein kinases independently or in a cross-reactive manner.

PI3K orthologue in *T. brucei*, TbVPS34, localizes to the endosomal system and is required for receptor-mediated endocytosis, transport to the lysosomes, and exocytosis. Knockdown causes severe growth defect, which present with a block in cytokinesis and impaired segregation of the Golgi apparatus (Hall *et al.*, 2006). Unlike many other eukaryotes, mTOR has 4 orthologues in *T. brucei*, some of them are essential for fitness and progression along the life cycle in the parasite in culture (Barquilla, Crespo and Navarro, 2008; Barquilla and Navarro, 2009a, 2009b; De Jesus *et al.*, 2010; Saldivia *et al.*, 2013).

A promising compound in the pipeline for drug discovery against cancer, proved to be effective against *T. brucei*, *Leishmania*, and *T. cruzi*. Phenotype resulting from compound treatment correlated with biological implications of RNAi against both targets in *T. brucei*. It also showed efficacy in mice infections (Diaz-Gonzalez *et al.*, 2011). SAR studies led to development of analogues aiming to increase selectivity *versus* human, solubility and predicted permeability to the CNS (Wager, Hou, *et al.*, 2010). These conditions were achieved for two of these compounds although they showed lower potency than the originals (Seixas *et al.*,

2014). These compounds were shown in a phosphoproteomics study inhibiting PI3K pathways.

Tyrosine kinase inhibitors

Tyrosine kinases are absent in the *T. brucei* kinome but tyrosine phosphorylation has been reported, possibly due to dual specificity kinases (Nett *et al.*, 2009). Tyrphostin, a non-specific tyrosine kinase inhibitor, was shown to be a disruptor of receptor-mediated endocytosis of transferrin in *T. brucei* (Subramanya and Mensa-Wilmot, 2010). This suggested that these chemotypes could have utility in *T. brucei* drug development.

With this rationale, three tyrosine kinase inhibitors beyond phase I of clinical trials for breast cancer, lapatinib, canertinib and AEE788, were studied in *T. brucei*. Although they targeted EGFR, which has no homologues in the parasite, the three drugs were trypanocidal. An ATP-sepharose affinity chromatography coupled to mass spectrometry after elution with the compounds unveiled 4 candidate tyrosine kinase inhibitor-binding protein kinase targets (Katiyar *et al.*, 2013).

8 analogues of these drugs designed through SAR were found with drug-like EC50s, 2 showing selectivity over human cells, with 1 bioavailable in oral dose. The compound induced a block in cytokinesis but physicochemical properties suggested low CNS penetrance (Patel *et al.*, 2013). In a more recent study, lapatinib inhibit parasite growth compared to controls *in vitro* and *in vivo*. 3 mice (out of 4) survived longer than the untreated controls with one cured from the infection (Behera, Thomas and Mensa-Wilmot, 2014).

Other kinase inhibitor libraries screened against TbERK8

It was mentioned above that TbERK8 (MAPK6) was demonstrated to be essential in a luciferase-based RNAi screen (Mackey *et al.*, 2011). Interesting, the human orthologue is not indispensable. TbERK8 is an active kinase that phosphorylates the putative replication factor TbPCNA and that can be targeted with inhibitors of human ERK8 (Valenciano, Knudsen and Mackey, 2016). A library of 274 kinase inhibitors FDA approved or advanced in the pipeline of drug discovery was

screened identifying 10 novel hits that are potent against the parasite in culture. 4 inhibited recombinant TbERK8, and one did it specifically (Valenciano *et al.*, 2016).

1.6 Aims of this thesis

My intention in this PhD was to attempt a methodical approximation to the preclinical stages of the drug discovery process in a tropical neglected disease. I designed the experimental plan so I covered the subjects of drug target identification, drug target characterization, and mechanism-directed active compound screening making the most out of the resources available at my institution: the Welcome Centre for Molecular Parasitology.

As such, the objective of this thesis was to identify protein kinases of biological relevance in *T. brucei* with potential value as drug targets for chemotherapy. In chapter 3, I will show how I exploited RNAi target sequencing technology in order to understand which are required for establishment of the mammalian infection. Among these protein kinases, depletion of several caused a loss of fitness in the parasite at the mammalian bloodstream that did not occur in the *in vitro* culture. Most of these protein kinases had orthologues in other organisms playing roles in resistance to environmental shocks and different stress factors. I will show proof of requirement of some of these kinases to withstand serum-mediated lysis and osmotic shock treatment.

In chapter 4, I will describe in more detail RDK2: a protein kinase found essential for parasite fitness both *in vitro* and *in vivo*. At the culturing conditions of bloodstream forms, RDK2-targeting RNAi mutants acquire characteristics of the insect-infective procyclic forms, 24 h prior to cell death. Interestingly, the procyclic-like phenotype could be maintained alive for more than one week in conditioned procyclic form culturing media. However parasite density did not increase suggesting a cell cycle arrest. Recombinant RDK2 was produced and proven to be an active protein kinase capable of phosphorylating a substrate and of autophosphorylation. Ectopic expression of RNAi-refractory mutants of RDK2 in cell lines depleted of the wildtype did not permit to establish with clarity whether the phenotype was due to kinase inhibition or to other consequences derived of RDK2 absence.

In chapter 5, I present the results of a collaborative project with GSK and the Swiss TPH where I screen against the complete parasite 518 compounds for activity matching the phenotypic outcome of RDK2 depletion. I identify 6 compounds producing signs of differentiation followed by parasite death. In chapter 6, I discuss the gaps of knowledge I have covered and the new questions and research perspectives that arise from this work.

Chapter 2 Materials and methods

2.1 Ethics statement

Animal procedures followed the guidelines and were approved by The Home Office of the UK government. The procedures here presented were covered by the project license PPL60/4442 entitled ‘‘Molecular Genetics of Trypanosomes and Leishmania’’. The University of Glasgow ethics committee also approved all protocols.

2.2 General Bioinformatics

2.2.1 Sequence retrieval

DNA and protein sequences from the different genes analysed in this thesis were extracted from TriTrypDB.org, version 29 (12 Oct 2016) while this thesis is being written. This includes not only *T. brucei brucei* genes but also *T. cruzi* and *L. major* orthologues analysed in the RITseq chapter. Protein sequences from other organisms used in BLASTp for function predictions were taken from NCBI database (http://www.ncbi.nlm.nih.gov/).

2.2.2 General sequence manipulation

Sequence manipulation was conducted in CLC genomics workbench 7 (Qiagen), including design of primers, cloning strategies, *in silico* translation, hairpin folding and protein size predictions. Reverse translation was applied to the RDK2 protein sequence in order to generate an RNAi refractory version of the gene, using *T. brucei* 927 codon bias to prioritise the less frequent codons. The gene product was assessed manually to make sure that no regions longer than 20 nucleotides were left with identity to wild-type genes. Then they were used in an RNAit search (Redmond, Vadivelu and Field, 2003) to ensure that there was no homology to wild-type *RDK2* and thus would not be downregulated by the RDK2 RNAi target sequence. Identification of different domains predicted for RDK2 was made by alignment to three hidden markov model (HMM) databases:

SMART (Letunic, Doerks and Bork, 2015), PFAM (Finn *et al.*, 2014) and Superfamily (Gough *et al.*, 2001).

2.3 Molecular biology

2.3.1 DNA preparation

2.3.1.1 Genomic DNA

Genomic DNA was purified with QIAamp DNA minikit (Qiagen) or DNAeasy Blood and Tissue (Qiagen). Concentration and purity was measured in Nanodrop 1000 (Thermo Scientific) in a general basis. The source of *T. brucei brucei* strain for general cloning proposes was Lister 427 wild type. DNA was isolated from all other cell lines used in this thesis at some point. For RNAi target sequencing (RITseq), Nanodrop accuracy was not considered sufficient and fluorescence-based methods to measure double stranded DNA concentration were used: Qubit™ fluorometric quantitation (Thermo Scientific) for the first replicate of the *in vivo* RITseq, and Quant-iT™ PicoGreen® dsDNA Assay Kit (Thermo Scientific) for the second replicate *in vivo* and the three *in vitro* experiments. DNA preparation was conducted both by hand or using robotics (Qiacube, Qiagen).

2.3.1.2 Plasmid DNA

Plasmids were purified from overnight liquid bacterial cultures in LB medium with selection drugs using QIAprep Miniprep kit (Quiagen). Concentration and purity was measured with Nanodrop 1000 (Thermo Scientific). If required, successful cloning was assessed via test digestion, PCR or sequencing.

Table 2-1. List of plasmids used in this thesis outside the pTL library.

Mottram Lab Code	Insert	Function	Destination	Selection
pGL2084	Gateway-modified pRPa^{iSL} (Jones *et al.*, 2014)	Tet-inducible RNAi	RRNA ΔHyg	Ampicillin
pGL2220	pRPa^{x6mycx} (Alsford and Horn, 2008)	Tet-inducible Overexpression	RRNA ΔHyg	Ampicillin
pGL2323	*Rdk2-6myc* (pGL2220-modified)	Tet-inducible Overexpression	RRNA ΔHyg	Ampicillin
pGL2324	*6myc-Rdk2* (pGL2220-modified)	Tet-inducible Overexpression	RRNA ΔHyg	Ampicillin
pRM481	(Proudfoot and McCulloch, 2005)	Constitutive expression	Tubulin	Ampicillin
pGL2271	*Rdk2^{REC}-6xHA* (pRM481-modified)	Constitutive expression	Tubulin	Ampicillin
pGL2297	*Rdk2$^{REC/K70M}$-6HA* (pGL2271-modified)	Constitutive expression	Tubulin	Ampicillin

pGL2471	*Rdk2^{REC7A50C}-6HA* (pGL2271-modified)	Constitutive expression	Tubulin	Ampicillin
pGL2587	*6HA-Rdk2^{REC}* (pGL2271-modified)	Constitutive expression	Tubulin	Ampicillin
pGL2099	*Trx-Rdk2* (pET32 Xa/LIC-modified)	Recombinant expression	*E. coli*	Ampicillin
pGL2215	*Trx-Rdk2^{K70M}* (pGL2099-modified)	Recombinant expression	*E. coli*	Ampicillin
pGL2654	*Trx-Rdk2^{S195A}* (pGL2099-modified)	Recombinant expression	*E. coli*	Ampicillin
pGL2655	*Trx-Rdk2^{S195E}* (pGL2099-modified)	Recombinant expression	*E. coli*	Ampicillin
pGL2656	*Trx-Rdk2$^{S195A/S197A}$* (pGL2654-modif.)	Recombinant expression	*E. coli*	Ampicillin
pGL2657	*Trx-Rdk2$^{S195A/S197E}$* (pGL2654-modif.)	Recombinant expression	*E. coli*	Ampicillin
pGL2658	*Trx-Rdk2$^{S195E/S197E}$* (pGL2655-modif.)	Recombinant expression	*E. coli*	Ampicillin
pGL2659	*Trx-Rdk2$^{S195E/S197A}$* (pGL2655-modif.)	Recombinant expression	*E. coli*	Ampicillin

2.3.2 DNA synthesis

DNA synthesis of complete genes, such as recoded *RDK2* (Figure 4-8), was conducted at Dundee Cell products. Routine synthesis of oligonucleotides was produced by Eurofins genomics.

2.3.3 Polymerase chain reaction (PCR)

Most of PCR reactions conducted for this thesis required the high fidelity polymerases Phusion HF polymerase and Q5 HF polymerase (New England Biolabs (NEB). Reactions were conducted in 25 or 50 μl following manufacturer's instructions. Annealing temperatures were calculated based on the melting temperature of the diverse primers based on results from NEB Tm calculator. When necessary, gradient PCRs were set up to select the optimal annealing conditions.

2.3.3.1 RITseq PCR

In order to enrich the DNA prepared from samples coming from any of the RITseq experiments in the RNAi target, a PCR was optimized using OL4161 and derivative barcoded primers. The PCR program was: 3 mins at 98°C; followed by 28 cycles comprising: 10 seconds at 98°C, 10 seconds at 61°C (64°C if barcoded primers), and 30 seconds at 72°C; with a final extension step at 72°C for 10 mins.

Table 2-2. List of primers used for optimization of the RITseq PCR

Mottram Lab code	Binding site	Sequence
OL4161	AttL1 Fw1	TAATGCCAACTTTGTACAAA
OL4162	AttL1 Fw2	TTGTACAAAAAAGCAGGC
OL4163	Stuffer Rv1	CTCAACCCTATCTCATCG

OL4164	AttL2 Rv3a	ACCTGTTCGTTGCAACAA
OL4165	AttL2 Rv3b	CGTTGCAACAAATTGATA
OL4212	AttL2 Rv4	TAATGCCAACTTTGTACAAG

Binding sites depicted in Figure 3-4.

Table 2-3. List of primers used for the barcoded RITseq experiment

Mottram Lab code	RITseq barcode	Sequence
OL4440b	RIT-1-F	ATCACGTAATGCCAACTTTGTACAAA
OL4442	RIT-2-F	CGATGTTAATGCCAACTTTGTACAAA
OL4444	RIT-3-F	TTAGGCTAATGCCAACTTTGTACAAA
OL4446	RIT-4-F	TGACCATAATGCCAACTTTGTACAAA
OL4448	RIT-5-F	ACAGTGTAATGCCAACTTTGTACAAA
OL4450	RIT-6-F	GCCAATTAATGCCAACTTTGTACAAA
OL4452	RIT-7-F	CAGATCTAATGCCAACTTTGTACAAA
OL4454	RIT-8-F	ACTTGATAATGCCAACTTTGTACAAA
OL4456	RIT-9-F	GATCAGTAATGCCAACTTTGTACAAA
OL4458	RIT-10-F	TAGCTTTAATGCCAACTTTGTACAAA
OL4460	RIT-11-F	GGCTACTAATGCCAACTTTGTACAAA
OL4462	RIT-12-F	CTTGTATAATGCCAACTTTGTACAAA
OL4464	RIT-13-F	AGTCAATAATGCCAACTTTGTACAAA
OL4466	RIT-14-F	AGTTCCTAATGCCAACTTTGTACAAA
OL4468	RIT-15-F	ATGTCATAATGCCAACTTTGTACAAA
OL4470	RIT-16-F	CCGTCCTAATGCCAACTTTGTACAAA
OL4472	RIT-17-F	GTCCGCTAATGCCAACTTTGTACAAA
OL4474	RIT-18-F	GTGAAATAATGCCAACTTTGTACAAA
OL4476	RIT-19-F	GTGGCCTAATGCCAACTTTGTACAAA
OL4478	RIT-20-F	GTTTCGTAATGCCAACTTTGTACAAA
OL4480	RIT-21-F	CGTACGTAATGCCAACTTTGTACAAA
OL4482	RIT-22-F	GAGTGGTAATGCCAACTTTGTACAAA
OL4484	RIT-23-F	ACTGATTAATGCCAACTTTGTACAAA
OL4486	RIT-24-F	ATTCCTTAATGCCAACTTTGTACAAA

Primers highlighted in blue showed homogeneous amplification and were used in the screen, Figure 3-5.

2.3.4 PCR clean up

When required, PCR products were purified using QIAquick PCR Purification Kit (Qiagen). This technique was applied for purification of digested/linearized PCR constructs or plasmids when it was required to create sticky ends in a PCR product amplified with restriction sites or it was not necessary to isolate a particular DNA fragment.

2.3.5 Agarose gel electrophoresis

DNA was analysed in UltraPure agarose (Invitrogen) gels made in 0.5 x TBE buffer (20mM Tris, 20 mM boric acid, 0.5 mM EDTA, pH 7.2). They were used 1% (w/v) for PCR, and 0.7% (w/v) for plasmid analysis. Gels were supplemented with SYBR™ Safe DNA Gel Stain (Invitrogen) to allow DNA visualization in a Gel Doc™ XR + Gel Documentation System (Bio-Rad). Gels were run for 1 h at 100 V, using 1 Kb plus ladder (Invitrogen) as size standards.

2.3.6 Gel purification

When necessary, DNA was purified from agarose by selecting bands of interests in a DarkReader blue light transilluminator using a scalpel blade to remove the gel fragment. Excised bands were processed with MinElute Gel Extraction kit (Qiagen) by hand or using an automated protocol in a Qiacube (Qiagen).

2.3.7 DNA cloning

Different approaches have been followed in this thesis for DNA cloning purposes.

2.3.7.1 Restriction/Ligation traditional cloning

All restriction digestions were conducted using restriction enzymes from NEB, following manufacturers recommendations. Linearized backbone vectors were purified from agarose gels with MinElute Gel Extraction kit (Qiagen), unless the drop out was small enough to flow through a PCR purification column. When inserts were amplified from a DNA source (gDNA or plasmid) using Phusion or Q5 polymerases (NEB) including restriction sites in the primers, they were purified with QIAquick PCR Purification Kit (Qiagen) prior to being digested. After digestion a sample was analysed in agarose gel and the rest was processed in a second round of PCR cleanup to discard the small cleaved fragments removed to create the sticky ends. If the inserts were a product of digestion of another plasmid, they were run in an agarose gel and the right-sized band purified with MinElute Gel Extraction kit (Qiagen). Backbone and inserts carrying complementary cohesive ends were ligated using T4 ligase (Promega), incubating reaction for 4 h at room temperature or at 16 ^{o}C overnight.

Table 2-4. List of primers used for digestion/ligation cloning in this thesis.

Mottram lab code	Gene	Objective	Destination	System	Sequence
OL4172	*Rdk2*	6myc (N) *Avr*II Fw	pGL2220	Tet-ind OE	GACTCCTAGGACGGAACATAAGCGTCC
OL4173	*Rdk2*	6myc (N) *Bam*HI Rv	pGL2220	Tet-ind OE	GACTGGATCCTCAACGAGGCACGGCAT
OL4174	*Rdk2*	6myc (C) *Hind*III Fw	pGL2220	Tet-ind OE	GACTAAGCTTATGACGGAACATAAGCG
OL4175	*Rdk2*	6myc (C) *Xba*I Rv	pGL2220	Tet-ind OE	GACTTCTAGAACGAGGCACGGCATTCT

Red=landing for the restriction enzyme/Blue=restriction sites

2.3.7.2 Gateway cloning

The majority of plasmids included in the pTL library were produced using Gateway cloning (Invitrogen), a few needed to be remade in the system in order to include the universal AttL flanks where the RITseq PCR primers sit (Figure 3-2). Primers to amplify the RNAi target sequence were redesigned in CLC Genomics Workbench including AttB gateway flanks and synthetized by Eurofins genomics (Table 2-5). PCR products were generated with Phusion polymerase (NEB) using gDNA from *T. brucei brucei* Lister 427 wild type strain as a template. Reactions were set in 50 μl following manufacturers instructions. AttP-containing pGL2084 was the backbone vector (Jones *et al.*, 2014). Within both AttP-flanked RNAit acceptor sites this plasmid contains the *ccdB* gene, which is toxic unless cloned in One Shot® *ccdB* Survival™ 2 T1R Competent Cells (Invitrogen) and maintained at 25°C. Every batch of plasmid used required assessment by test digestion as deletion of *ccdB* happened frequently as a mechanism of resistance. AttB-amplified PCR products were cloned into AttP-containing pG2084 using Gateway® BP Clonase® II Enzyme mix (Invitrogen). Resulting lines were analysed in Alamar blue (see below).

Table 2-5. List of primers to make with Gateway cloning lines that were made with ligation.

Mottram Lab code	Gene ID	Name	Sequence
TLO41	Tb927.10.4990	CRK3 Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTaaaggctctcgagaagaggg
TLO42		CRK3 Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTagcgctggtaacacctgagt
TLO417	Tb927.11.12410	KKT10 Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTtaactcatccaccattgcca
TLO418		KKT10 Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTagatgcatatgacgtgacgc
TLO255	Tb927.10.14300	STE11 Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTtgcattttcaagcaaacagc
TLO256		STE11 Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTgcagaccgaaccagtacatt
TLO419	Tb927.7.3580	NEK 11 Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTcctgaagctgggagactttg
TLO420		NEK 11 Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTgttttgttattacctcgggcg
TLO421	Tb927.10.16160	Orphan Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTaaatggactttaccctccgc
TLO422		Orphan Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTctgcctcggctaattctgac

TLO423	Tb927.3.3290	Orphan Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTtgtattgcatgtgggaggaa
TLO424		Orphan Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTggcacttgtataccacgcct
TLO425	Tb927.3.3080	Nek 6 Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTgaagaagttgcaggtcaggc
TLO426		Nek 6 Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTtcatcaaggaaactcgaccc
TLO427	Tb927.3.3190	Nek 7 Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTgccctcggtgtcattcttta
TLO428		Nek 7 Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTtatgctctcccgattcatcc
TLO429	Tb927.8.1670	Nek 13 Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTcacgtgaagagaagcagcag
TLO430		Nek 13 Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTaatttcgactgaaacgtccg
TLO431	Tb927.9.4910	TbPDK1 Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTggtccgcaggaccttatgta
TLO432		TbPDK1 Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTaccacaggcccagtaatcag
TLO433	Tb927.6.2250	TbRAC Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTctccagagtttctgcttggg
TLO434		TbRAC Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTccactacttcatcgtccggt
TLO435	Tb927.11.6690	Orphan Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTatctccagacccactgatgg
TLO436		Orphan Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTaggcaagggcacatacaatc
TLO437	Tb927.3.2440	TbPKDH Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTggcgaactcatttgttggtt
TLO438		TbPKDH Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTttgtttccactttgtgcagc

Fw= Forward primer; Rv= Revers primer; red= AttLB1 site; blue= AttLB2 site

2.3.7.3 Ligation independent cloning

Expression vector pGL2099 (expressing Trx:RDK2) was created by ligation independent cloning using pET32 Xa/LIC (Novagen following manufacturer instructions. RDK2 was amplified with primers containing overhangs homologous to the vector: 15 nucleotides in the 5' hybridization site and 17 nucleotides in the 3' end (Table 2-6).

Table 2-6. List of primers for ligation independent cloning of RDK2 in pET32 Xa/LIC (Novagen)

Mottram Lab Code	Gene ID	Function	Sequence
OL3746	Tb927.4.5310	RDK2 LIC/Xa pET32 Fw	GGTATTGAGGGTCGCatgacggaacataagcgtcc
OL3747		RDK2 LIC/Xa pET32 Rv	AGAGGAGAGTTAGAGCtcaacgaggcacggcattct

Upper case=vector specific/Lower case=kinase specific

The resulting PCR product was purified and treated with T4 DNA polymerase in the presence of dGTP as the only nucleotide available. This permitted the enzyme to act with its 3'→5' exonuclease activity until finding a position where dGTP could be inserted creating sticky ends that would anneal with linearized vector. Vector and T4 polymerase-treated insert were annealed for 5 mins at 22 ^{o}C and then, 25 mM EDTA was added to the mix for a further 5 mins at the same temperature. pGL2099 was transformed in NovaBlue GigaSingles™ competent cells (Novagen) as described below.

2.3.8 Site directed mutagenesis

Recoded *Rdk2* (plasmid pGL2271) was mutated to create RDK2^{A50C} and RDK2^{K70M} variants. The recombinant protein (pGL2099) was also mutated to produce RDK2^{K70M}, RDK2$^{S195A/S197A}$, RDK2$^{S195A/S197E}$, RDK2$^{S195E/S197A}$, and RDK2$^{S195E/S197E}$. To create the double phosphosite mutations, intermediate plasmids were generated encoding RDK2^{S195A}, and RDK2^{S195E} respectively. All these mutations were induced with Q5® Site-Directed Mutagenesis Kit (NEB). Primers were designed using the NEBaseChanger™ online tool and used in a first-stage PCR of the vector that introduces the selected deletions/insertions included in the primers. In a second stage, methylated-specific DpnI treatment removed the original template as it was the only one methylated by the bacterial host, and the kinase-ligase tandem circularized the mutated plasmid by 5' phosphorylation and ligation. Primers designed for this purpose can be found in Table 2-7.

Table 2-7. List of primers designed to introduce mutations using Q5® Site-Directed Mutagenesis Kit (NEB).

Mottram Lab Code	Target gene	Mutation	Sequence
OL4061	*Rdk2^{REC}*	K70M Fw	GGACTTATATTTGCAGCGATGGTAATGGATACGAACAAC
OL4062	*Rdk2^{REC}*	K70M Rv	GTTGTTCGTATCCATTACCATCGCTGCAAATATAAGTCC
OL4921	*Rdk2^{REC}*	A55C Fw	CTTTGGTTCTtgcTGGCGTGTAGAGG
OL4922	*Rdk2^{REC}*	A55C Rv	CTACCTTGCCCTAATAAC
OL5172	*Trx-Rdk2*	S195E Fw	ATACGAGGACgaGCTGTCAAATCC
OL5173	*Trx-Rdk2*	S195E or S195A Rv	TGACGACTGAAGCCAAAG
OL5174	*Trx-Rdk2^{S195E}*	S197E Fw	GGACGAGCTGgaAAATCCTGTTG
OL5175	*Trx-Rdk2^{S195E}* or *Trx-Rdk2^{S195A}*	S197E Rv	TCGTATTGACGACTGAAG
OL5176	*Trx-Rdk2^{S195E}*	S197A Fw	GGACGAGCTGgCAAATCCTGT
OL5177	*Trx-Rdk2^{S195E}*	S197A Rv	TCGTATTGACGACTGAAGC
OL5178	*Trx-Rdk2*	S195A Fw	ATACGAGGACgCGCTGTCAAATC
OL5179	*Trx-Rdk2^{S195A}*	S197E Fw	GGACGCGCTGgaAAATCCTGTTG
OL5180	*Trx-Rdk2^{S195A}*	S197A Fw	GGACGCGCTGgCAAATCCTGT
OL5181	*Trx-Rdk2^{S195A}*	S197A Rv	TCGTATTGACGACTGAAGCCAAAG

2.3.9 Transformation in bacteria

Plasmids were amplified by transformation in *E. coli*, with the objective of expanding the plasmid and/or generating frozen stocks. As a general rule, 50 μl of competent cells were incubated with 2 μl of plasmid for 20 mins on ice. Then,

tubes were heat shocked for 30 seconds at 42 ^{o}C, and left for 2 mins recovering on ice. After this, 150 μl of SOC media were added on top, cells were incubated for 1 h at 37 ^{o}C, and spread on the surface of pre-warmed LB agar plates supplemented with selective antibiotics. After overnight incubation at 37 ^{o}C, 2-3 colonies were isolated and expanded in overnight 5-10 ml liquid LB cultures supplemented with selective antibiotics. 800 μl of cells were used for cryopreservation and the rest for plasmid purification using QIAprep Miniprep kit (Quiagen).

Different plasmids required different bacterial hosts for expansion. Routine cloning with Tet-inducible RNAi or overexpression plasmids in Table 2-1, was conducted in *E. coli* DH5α at 37^{o}C (Invitrogen). pGL2084, that contained the toxic *ccd*B gene, were grown in *E. coli* One Shot® *ccd*B Survival™ 2 T1^{R} competent cells (Invitrogen). Finally, recombinant expression vectors in Table 2-1 were grown in codon-enriched *E. coli* Rosetta™ (DE3) p-lysS strain (Novagen). pGL2099 was grown immediately after cloning in NovaBlue GigaSingles™ competent cells (Novagen) including in the Xa/LIC kit, see "Ligation independent cloning" section.

For cryopreservation 800 μl of cells from an overnight liquid LB culture of transformed bacteria were diluted with 800 μl of LB 40% glycerol/2% peptone and kept at -80 ^{o}C for long-term storage. For over one week, colonies in the plates could also be used.

2.3.10 DNA sequencing

2.3.10.1 Routine sequencing

Routine sequencing for cloning validation or sequence identification was conducted at Eurofins genomics after PCR amplification with a high fidelity polymerase (Phusion or Q5, New England Biolabs) sending premixed samples including the desired primer and PCR products directly at the concentration indicated by the service provider. For longer reads, PCR products were purified and subcloned in pGEM-T easy vector (Promega) and sent with sequencing primers to eurofins genomics. Cloning products –such as newly made vectors- were sent directly with sequencing primers. Both polymerases produced blunt-

end products so A-overhangs were added with *Taq* (NEB) in a single cycle of 30 mins at 72°C. Reactions were later purified and ligated in pGEM-T easy vector following manufacturer instructions. Alternatively, PCR products were subcloned with StrataClone Blunt PCR Cloning Kit (Agilent Technologies), which permits blunt-end ligation. Cloned products were transformed in MAX efficiency *Escherichia coli* DH5α (Invitrogen) and plated overnight. Colonies were picked and grown in liquid LB overnight cultures, plasmids purified the next day with QIAprep Miniprep kit (Quiagen), and send for sequencing with appropriate primers.

2.3.10.2 RNAi target sequencing

The first *in vivo* RNAi target sequencing replicate was conducted at Glasgow Polyomics. 200 ng of DNA per sample were fragmented, size selected (average read length after sequencing ~106 nucleotides) and processed according to Ion Torrent library prep protocols. Multiplexed sample was subjected to emulsion PCR and Ion Torrent sequencing.

The second *in vivo* replicate and the three *in vitro* experiments were sequenced at Eurofins genomics using Illumina HiSeq, avoiding the fragmentation and size selection steps. Samples for this second set of experiments were enriched in the PCR cassette using 14 distinct bar-coded primers (Table 2-3). In this manner they could be pooled to provide equal mass of every sample and 400 ng of each pool of 14 were sent for sequencing.

2.4 Protein manipulation

2.4.1 Protein sample preparation

2.4.1.1 Bacterial cultures

0.5 ml of bacterial cultures were pelleted at 12000xg for 2 mins per condition, and supernatant removed. As a general rule, if analysing protein expression, one sample was taken for the uninduced control and two samples for the IPTG induced (see below). The complete uninduced sample and one of the induced were resuspended in 60 μl of 1x protein loading buffer (250 μl 4x NuPAGE® LDS sample buffer [Invitrogen], 750 μl 1x PBS and 25 μl β-mercaptoethanol), 10

seconds of sonication were applied, and samples were denatured at 95 ^{o}C for 10 mins. Samples could be stored at -20 ^{o}C until being used. For the other induced sample, it was resuspended in 45 μl of PBS, subjected to 3 cycles of sonication (10 seconds each) in order to obtain the soluble fraction and centrifuged again 12000xg for 10 mins. Supernatant (soluble fraction) was transferred to a different tube and supplemented with 15 μl of 4x NuPAGE® LDS sample buffer (Invitrogen). Pellet (insoluble fraction) was supplemented with 60 μl of 1x protein loading buffer.

2.4.1.2 Trypanosome cultures

As a general rule, between 10^6-2.5x10^6 parasites were used per sample. They were centrifuged at 1500 x g for 10 mins and lysed with 20 μl 1x protein loading buffer per sample before being denatured at 95 ^{o}C. Samples could be kept at -20 ^{o}C until needed.

2.4.2 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Sodium dodecyl sulphate polyacrylamide gel electrophoresis was used to separate proteins according to size. NuPAGE™ Novex™ 10% Bis-Tris Protein Gels, 1.0 mm, 12-well (Invitrogen) were generally used in a XCell SureLock™ Mini-Cell Electrophoresis System (Invitrogen). NuPAGE® MES SDS Running Buffer or NuPAGE® MOPS SDS Running Buffer (Invitrogen) were used as running buffer. Samples were defrosted and denatured again at 95 ^{o}C before loading 20 μl per lane. Electrophoresis was run for 45 mins at 150 V.

To visualize proteins after electrophoresis, polyacrylamide gels were rinsed in milliQ water and SimplyBlue™ SafeStain ready-to-use reagent (Life technologies) was applied for about one h followed by three 20-minute rinses with distillate water. To fasten the process, the reaction was warmed up for 10 seconds in the microwave at maximum power.

2.4.3 Western blotting

For western blot, proteins separated by SDS-PAGE were transferred to a Polyvinylidene difluoride (PVDF) Biotrace membrane (Pall) in an X-Cell II Blot

Module (Invitrogen). PVDF membrane was activated in methanol for 30 seconds, rinsed in milliQ water then equilibrated in transfer buffer [20 % (v/v) methanol, 20 mM Tris, 15 mM glycine]. The transfer chamber was assembled and filled with transfer buffer with all pads and filter papers soaked previously in it. Transfer was performed at 30 V for 2 h with the outer chamber filled with water to provide a coolant. As a general rule, membrane was rinsed in 1x PBST (PBS, 0.01 % Tween-20 [Sigma]) for 10 mins on a rocker then incubated for 1 h in blocking solution (1x PBST, 5 % Milk powder [Marvel]) or if required, overnight at 4 ^{o}C. After, the membrane was rinsed once for 10 mins in 1x PBST then it was moved to blocking buffer containing the required primary antisera and incubated for 1 h. Then, membrane was washed three times for 10 mins with 1xPBST, and incubated for another hour with a secondary antibody against the primary serotype conjugated to horseradish peroxidase (HRP). Again, the membrane was washed three times for 10 mins with 1 x PBST then the SuperSignal West Pico Chemiluminescent Substrate (Thermo-Fisher) or ECL Prime Western Blotting Detection Reagent (Amersham) were pipetted evenly onto the membrane surface, which was subsequently incubated in the dark for 5 mins. The membrane was then exposed in a gel doc, or using medical X-ray blue film MXB (Carestream) in a Kiran cassette for ~one second to overnight and films were developed using a Kodak M-25-M X-omat processor.

2.4.3.1 Antibody detection

Table 2-8. List of antibodies used for western blot in this thesis.

Antibody	Serotype	Clone	Concentration	Supplier
αVSG 221 variant	Rabbit IgG	N/A	1:20000	R. McCulloch
αVSG crossreactive*	Rabbit IgG	N/A	1:20000	D. Horn
αEF1α	Mouse IgG	CBP-KK1	1:25000	Millipore
αEP procyclin	Mouse IgG	TBRP1/247	1:500	Cedarlane
αMyc	Mouse IgG	4A6	1:7000	Millipore
αHA	Mouse IgG	HA-7	1:1000	Sigma
αMouse IgG	Goat IgG	N/A	1:5000	Novex
αRabbit IgG	Goat IgG	N/A	1:5000	Novex

*Required hypotonic lysis to maintain folding of the crossreactive determinant (10 mM Tris base, pH 7.5 with HCl on ice) and centrifugation 12000xg 2 mins to separate supernatant separating soluble and insoluble fractions.
^The αHA antibody was incubated overnight at 4 ^{o}C
Abbreviations: VSG (Variant Surface Glycoprotein), EF1α (Elongation Factor 1α, Ig (Immunoglobulin).

2.4.4 Protein expression and purification

All proteins expressed for this thesis were mutants derived from the same Trx-RDK2 protein and were expressed under the same conditions. Codon-enriched *Escherichia coli* Rosetta™ (DE3) p-lysS strain, with T7 lysozyme to supress basal T7 expression, was used as the bacterial host. Protein expression was induced with 1 mM IPTG in 1L log-phase cells (OD600=0.5) at 17°C overnight.

Proteins were derived from pET32 Xa/LIC expression system (Novagen). They all express the protein of interested fused to a N-terminal thioredoxin tag for solubility and a 6xHis tag for purification using ion metal affinity chromatography (IMAC).

After centrifugation of an overnight induced culture, 14000g for 15 mins pellets were lysed with B-PER (Thermo Scientific). Then they were sonicated (5 cycles of 10 seconds on ice), centrifuged again to collect supernatant, and filtered through a 0.22 μm minisart (Sigma), and loaded for IMAC in a 5 ml HisTrap™ HP column (GE Healthcare), precharged with Ni^2 Sepharose™ High Performance medium on an AKTA purifier (GE Healthcare). All buffers used in the AKTA system were previously filtered through a 0.22 μm Stericup® filtration unit (Millipore). The column was equilibrated with 20 mM NaH_2PO_4, 500 mM NaCl, 5 mM imidazole, pH 7.4. After binding the 6xHis-tag, the column was washed with the same buffer but containing 35 mM imidazole. Finally, elution was conducted with the same buffer but including 500 mM imidazole monitoring UV absorbance in 1 ml fractions to detect the protein content. Spaced fractions across the elution peak were sampled and run in a 12% SDS-PAGE and stained with SimplyBlue™ SafeStain ready-to-use reagent (Life technologies) to assess purity of Trx-RDK2 (and derivatives) in order to pool the best fractions. Imidazole was removed from the pooled fractions in a PD-10 desalting column. To purify the protein further, the sample was loaded into a pre-equilibrated anion exchange HiTrap Q XL column in the AKTA purifier, washed with 10 column volumes of 25 mM Tris-HCl pH 7.9 and eluted in the same buffer with a linear gradient 0-1M NaCl over 10 column volumes. Fast Protein Liquid Chromatography (FLPC) was monitored with UV absorbance, in order to sample fractions matching the elution peak for SDS-PAGE and SimplyBlue™ SafeStain ready-to-use reagent (Life technologies). Proteins were collected in 6-8 ml of buffer and were concentrated with 30 KDa molecular weight cut-off Vivaspin Turbo 15 columns (Sartorius).

Protein concentration was assessed both in Nanodrop 1000 (Thermo Scientific) and via Bradford assay (Bradford, 1976). Nanodrop was chosen for systematic measurements as it was more straightforward and provided the same range of concentrations.

All proteins were stored by immersion in liquid nitrogen and storage at -80^{o}C with 10% glycerol.

2.4.5 Protein kinase assay

Purified, recombinant Trx-RDK2 and derivative proteins (active, dead mutants and phosphosite mutations) were adjusted to 100 μg/ml in 20 μl of kinase assay buffer (KAB) (50 mM MOPS, 20 mM $MgCl_2$, 10 mM DTT, 10 mM EGTA), i.e. 2 μg of protein per reaction tube. 2 reactions were prepared per protein to be assayed with and without a substrate. 20 μl of kinase assay mastermix (KAM) were added to 100 μM ATP, 0.37 MBq [γ-32P]-ATP, and 10 mg/ml of α-casein (Sigma-Aldrich) as a substrate if applicable in KAB. Final assay concentration of the protein was 50 μg/ml. Reactions were incubated at 30^{o}C for 30 mins with shaking and stopped by the addition of 40 μl 2x SDS-PAGE buffer and boiling for 3 mins. 20 μl of this was separated by SDS-PAGE, the gel was dried and placed into a cassette with a phosphor-storage screen (Molecular Dynamics) for 8 to 72 h before visualising the screen in a Typhoon 9400 imager (GE Healthcare). Alternatively, it was exposed for 72 h at -80 ^{o}C with a medical X-ray blue film MXB (Carestream) in a cassette with a calcium tungstate-intensifying screen (Kiran). Films were developed in a Kodak X-Omat processor.

2.4.6 Xa factor cleavage

In order to choose the right conditions, before scaling up, Xa factor (Novagen) cleavage properties were assayed in a miniaturized screen testing 0, 0.1, 0.2, and 0.5 U/μl against 10 μg of Trx-RDK2. Incubation times were optimized also testing 2, 4, 8, and 16 h in 50 μl reactions. 10 μl aliquots were mixed with 10 μl 2x loading buffer, boiled for 5 mins, separated SDS-PAGE and stained with SimplyBlue™ SafeStain ready-to-use reagent (Life technologies). After analysing results, 6 h at 0.5 U/μl was used at a large scale to cleave Trx-RDK2 and Trx-$RDK2^{K70M}$.

2.5 Parasite maintenance

Cell lines contained in the RNAi library are listed in Table S1. Other *Trypanosoma brucei brucei* clones used in this screen are listed in Table 2-9.

Table 2-9. List of cell lines used in this thesis besides the pTL library.

Cell line	Parental + Transfected plasmid	Aim
STL768	sTL629 + pGL2271	*Rdk2*^{REC}-6HA
STL854	sTL 629 + pGL2297	*Rdk2*$^{REC/K70M}$-6HA
STL967	sTL629 + pGL2471	*Rdk2*$^{REC/A50C}$-6HA
STL968	sTL629 + pGL2271	*Rdk2*^{REC}-6HA
STL965	2T1 + pTL247 (newly made)	Tet-inducible *RDK2* RNAi
STL966	2T1 + pTL247 (newly made)	Tet-inducible *RDK2* RNAi
STL969	sTL629 + pGL2587	6HA-*Rdk2*^{REC}
STL970	sTL965 + pGL2587	6HA-*Rdk2*^{REC}
sTL797	2T1 + pGL2323 clone 1	Tet-inducible *RDK2* overexpression (C-term 6xMyc)
sTL798	2T1 + pGL2323 clone 2	Tet-inducible *RDK2* overexpression (C-term 6xMyc)
sTL801	2T1 + pGL2324	Tet-inducible *RDK2* overexpression (N-term 6xMyc)

sTL629 and pTL247 are the RNAi line and vector targeting *RDK2* transcripts in the pTL library. The rest of plasmids (labelled pGLXXXX) are described in Table 2-1. pTL147 is the RNAi vector targeting RDK2

2.5.1 *Trypanosoma brucei brucei* 2T1 bloodstream forms

Monomorphic *Trypanosoma brucei brucei* 2T1 bloodstream forms (Alsford and Horn, 2008) were cultured in HMI-11 [HMI-9 (GIBCO) supplemented with 10% (v/v) foetal bovine serum (GIBCO), penicillin 20 U ml^{-1} and streptomycin 20 mg ml^{-1} (Pen/Strep solution, Sigma-Aldrich)], at pH 7.4, 37 ºC and 5% CO_2 in vented flasks. Selective antibiotics were used as follows: 0.2 μg ml^{-1} puromycin, 0.5 μg ml^{-1} phleomycin (InvivoGen). Parasites were maintained at a concentration of less than 2 x 10^6 parasites ml^{-1}.

2.5.2 Tetracycline-inducible RNAi or overexpression lines

Cell lines generated upon transfection of 2T1s with a plasmid contained in the "pTL" RNAi library (Jones *et al.*, 2014), or a pGL2220-derived (pRPa^{x6mycx}) overexpression construct (Alsford and Horn, 2008), were cultured in HMI11 complemented with 0.5 μg ml^{-1} phleomycin (InvivoGen), and 2.5 μg ml^{-1} hygromycin B (Calbiochem). RNAi/overexpression was induced *in vitro* with tetracycline (Sigma-Aldrich) at 1 μg ml^{-1}, and RNAi in mice with doxycycline

hyclate (Sigma-Aldrich) added in drinking water at 0.2 g L^{-1} laced with 50 g L^{-1} of sucrose. Parasites were maintained at a concentration of less than 2 x 10^6 parasites ml^{-1} at 37ºC and 5% CO_2 in vented flasks.

2.5.3 RDK2 RNAi lines complemented with a recoded mutant

RDK2 RNAi lines transfected with a recoded mutant using a pRM481-derived vector (Proudfoot and McCulloch, 2005), were maintained in HMI11 complemented with 0.5 μg ml^{-1} phleomycin (InvivoGen), 2.5 μg ml^{-1} hygromycin B (Calbiochem), and 10 μg ml^{-1} blasticidin (InvivoGen). Parasites were maintained at a concentration of less than 2x10^6 parasites ml^{-1} at 37 ºC and 5% CO_2 in vented flasks.

2.5.4 *T. brucei brucei* Lister strain 427 wild type procyclic forms

T. brucei brucei 427 wild type procyclic forms were kept in SDM79 (GIBCO) supplemented with 10% (v/v) foetal bovine serum (GIBCO), penicillin 20 U ml^{-1}, streptomycin 20 mg ml^{-1} (Pen/Strep solution, Sigma-Aldrich), at pH 7.4. Cultures were maintained in a range of concentrations between 5 x 10^5-10^7 parasites ml^{-1} at 27 ^oC in non-vented flasks.

2.6 Transfection of bloodstream trypanosomes

Plasmids designed for *T. brucei* genetic modification were linearized before transfection to enhance homologous recombination (Lee and Van der Ploeg, 1990). All the vectors were subjected to an *Asc*I-mediated double cut performed during an overnight digestion at 37 ^oC. 5 μg of linearized plasmid were precipitated by addition to the 50 μl digestion of 105 μl "ethanol/salt mix" (150 mM of sodium acetate in 70% ethanol), incubation for 15 mins in dry ice, and centrifugation at 13000xg at room temperature. Supernatant was discarded, and DNA pellet washed with 100 μl of 70% ethanol. After 5 mins centrifugation at 13000xg and room temperature, supernatant was discarded, and tubes left to dry in a laminar flow hood. DNA was resuspended with 11 μl of MilliQ water, 1 μl used to verify linearization in an agarose gel, and 10 μl used for transfection.

10^7 log-phase parasites per transfection were purified and washed with PBS using two rounds of centrifugation at 1500 x g for 10 mins before being resuspended in 100 μl of Amaxa Nucleofector T-cell buffer (Lonza). The 10 μl of linearized vector were added to the cells and transferred to an electroporation cuvette before shocking them with programme X-001 in a Lonza Nucleofector II machine. Cells were then transferred to 10 ml of prewarmed HMI11 media. From this, 1.5 ml were diluted (1:20) to 30 ml with HMI11 and plated in a 96-well plate. From the 1:20, 4 ml were diluted (1:5) to 20 ml making a 1:100 dilution from the original concentration, and in parallel, 2 ml were diluted (1:10) to 20 ml representing a 1:200 dilution from the original concentration. Both dilutions, 1:100 and 1:200, were also plated in 96-well plates. All wells were stamped with 100 μl of parasite dilutions and kept for 6 h at 37 ^oC. After this period, 50 μl with selection drugs at 3x the required concentration were added on top. Clones were recovered after one week of drug selection.

2.7 *T. brucei brucei* cryopreservation

Frozen stocks were kept when necessary by resuspension of 2.5 x 10^6 parasites in 1 ml of HMI11 (for bloodstream forms), or SDM79 (for procyclic forms), supplemented with 10% (v/v) glycerol. Tubes were placed for one week at -80^oC wrapped in cotton, and later moved to liquid nitrogen.

2.8 *In vitro* growth analysis

2.8.1 RNAi, overexpression lines, and derivatives

In vitro growth curves started with parasites at 1 x 10^4 cells ml^{-1} were counted daily with a haemocytometer. If log phase needed to be sustained, cultures were diluted 1:100 after 48 h, when fit parasites reached ~10^6 cells ml^{-1}. In such cases cumulative curves were represented by applying the inverted dilution factor (x100). Growth was assessed in triplicate comparing tet-induced cells (RNAi or overexpression) to uninduced controls using T-test. Results were considered significant when $P<0.05$. Curves were plotted with GraphPad Prism 6.0.

2.8.2 RDK2 RNAi phenotype growth assessment

Cells collected 48 h after tetracycline induction in HMI11 at 37 ^{o}C were placed at 27 ^{o}C with 10^5 parasites ml^{-1}. They were tested renewing or not tetracycline and maintained in different conditions: the same medium, in fresh HMI11, in fresh SDM79 or in conditioned SDM79. All media were 10% (v/v) foetal bovine serum. Conditioned SDM79 was prepared by supplementation of fresh medium with the product of spinning 1500 x g for 10 mins a *T. brucei brucei* Lister 427 procyclic culture in stationary phase and passaging supernatant through a Minisart filter of 0.2 μm pore size. Curves were plotted with GraphPad Prism 6.0.

2.9 Life cycle differentiation assays

2.9.1 Slender to stumpy* form differentiation

Slender to stumpy* differentiation was triggered by incubation for 24 h with 100 μM 8-(4-Chlorophenylthio) adenosine 3′,5′-cyclic monophosphate sodium salt (pCPT-cAMP), purchased from Sigma-Aldrich (Laxman *et al.*, 2006). Compound was washed off after 24 h to avoid cell toxicity. Differentiation was measured by an increase in response to procyclic differentiation treatment (see next section), by detecting levels of EP procyclin expression in Flow cytometry.

2.9.2 Bloodstream to procyclic form differentiation

Bloodstream-to-procyclic form differentiation was triggered in different conditions. The standard (Overath, Czichos and Haas, 1986) was used to test RDK2 overexpression effect after incubation with pCPT-cAMP. It consisted in treating with 3 mM of each sodium citrate and cis-aconitic acid (CCA treatment), both purchased from Sigma-Aldrich, in differentiation trypanosome media (DTM) and adjusting the pH to 7.4, at 27 ^{o}C.

Other conditions tested with RDK2 RNAi to understand the role that this protein kinase plays in the signalling networks controlling differentiation were: 3 mM CCA treatment alone or preincubated with 100 μM 8-pCPT-cAMP in HMI11, tested both at 27 and 37 ^{o}C; and treatment with 150 μM 3-(3,5-Dibromo-4-hydroxy-benzoyl)-2-ethyl-benzofuran-6-sulfonicacid-(4-(thiazol-2-ylsulfamyl)-phenyl)-amide (short-named “BZ3”, purchased from Calbiochem®, Merk Millipore).

2.10 Making the RNAi library pool

The pooling strategy had as a first step the mixing of RNAi lines in 9 '*Mixed Stabilate Trypanosome Libraries*' (MSTL 1-8), each containing 19-25 cell lines (detailed grouping in Table S1 and Table 3-1). Then MSTLs were mixed to make the final pool. This intermediate step reduced risks of overgrowth and cross contamination. After the first replicate of the *in vivo* RITseq a few lines were found that had lost their stem loop construct or contained the wrong one. They were repaired and included in MSTL9 for a new MSTL used in the second replicate of the *in vivo* experiment and the *in vitro* experiment (Table 3-1). This part of the work was conducted in collaboration with Dr. Tiago D. Serafim.

To make the MSTL pools, independent cell lines were defrosted diluting glycerol to 1% with HMI-11 and no antibiotics. After 24 h recovery at 37 ^{o}C and 5% CO_2 atmosphere, cells were diluted again adding selective drugs (hygromycin and phleomycin). After another day, cultures were set at 2 x 10^5 cells ml^{-1} and ~24 h after- they were counted to pool 6 x 10^6 cells for each line in a common flask that was diluted to 50 ml. Cells were allowed to acclimatize for 1 h at 37 ^{o}C and the mixed culture was centrifuged 1500xg for 10 mins. Parasites were resuspended in 24 ml of HMI-11, 10% glycerol to generate 24 cryostabilates per MSTL.

For the final pool, each sub-library was defrosted, diluted 1:10 in HMI-11 without drug selection and cells were accurately counted in order to identify the 'limiting sublibrary': the MSTL culture containing a smaller amount of cells per kinase line after post-thaw recovery, see Table 3-1. The 'limiting culture' was used complete. Volumes to mix in the final pool from the other MSTLs were calculated to enable addition of as many parasites per line as were added using the whole 'limiting MSTL'. The culture was acclimatized at 37 ^{o}C 5% CO_2 for 45 mins, cells pelleted at 1200 x g for 10 mins and resuspended at 5 x 10^5 cells ml^{-1}. After 6 h incubation 37 ^{o}C 5% CO_2, cells were counted, washed and resuspended with PBSG to produce 200 µl inoculums of 5 x 10^4 cells ml^{-1}. Injections were performed intraperitoneally in mice. The rest of the pool was frozen to perform the *in vitro* experiments.

2.11 Animals and *in vivo* growth analysis

Female CD1 outbred mice (6–8 weeks old) were obtained from Charles River (Edinburgh, Scotland). Animals were given water and food *ad libitum*. Infections were evaluated daily by assessing parasitemia levels with a haemocytometer by tail vein sampling diluting blood in 0.83% (w/v) ammonium chloride to lyse red blood cells. Curves were plotted with GraphPad Prism 6.0.

During the *in vivo* growth curves mice were euthanized with carbon dioxide and cervical dislocation if parasitemia exceeded 10^8 cells ml $^{-1}$. Growth was assessed in triplicate comparing tet-induced mice to uninduced controls using T-test. Results were considered significant when $P<0.05$.

2.11.1 Parasite purification from blood with DEAE cellulose

For the *in vivo* RITseq, whole blood from euthanized mice after 72 h of infection was collected in anticoagulant CBSS/Heparin buffer (25 mM HEPES, 120 mM NaCl, 5.4 mM KCl, 0.55 mM CaCl, 0.5 mM $MgSO_4$, 5.6 mM Na_2HPO_4, 11.1 mM D-glucose). The blood was centrifuged first at 200xg for 10 mins to remove the majority of red blood cells, parasites remained be in the supernatant that was applied immediately to diethylaminoethyl-cellulose (DEAE) columns (Lanham and Godfrey, 1970).

Whatman® anion exchange DEAE-cellulose (Schleicher & Schuell) was suspended in phosphate buffered saline glucose (PSG: 44 mM NaCl, 57 mM Na_2HPO_4, 3 mM KH_2PO_4, and 55 mM glucose at pH8) until settled. Columns were prepared in 20 ml syringes with a thin layer of glass wool at the bottom. Then, they were loaded with 10 ml DEAE slurry settled overnight. Before operation, they were washed with 3 column volumes of PSG buffer. Then blood was applied and parasites eluted with 30 ml PSG buffer pH 8.0. $\pm 2 \times 10^7$ parasites/sample were harvested in the eluent. Cells were collected by centrifugation and DNA isolated from the parasites using QIAamp DNA kit (Qiagen).

2.12 *In vivo* RNAi Target sequencing sample preparation

Experiment 1 (Ion Torrent) and Experiment 2 (Illumina MiSeq) were processed in a slightly different manner.

The library pool used in Experiment 1 was the product of mixing MSTLs 1-8 (see Table 3-1). A single universal cassette-specific primer (OL4161: 5'-TAATGCCAACTTTGTACAA**A**-3') was used to PCR enrich RNAi inserts amplified from 10 ng of genomic DNA obtained per sample in a 50 µl reaction using Q5® High-Fidelity DNA polymerase (NEB, Ipswich, USA). The PCR program was: 3 mins at 98°C; followed by 28 cycles comprisinging: 10 seconds at 98°C, 10 seconds at 61°C and 30 seconds at 72°C; with a final extension step at 72°C for 10 mins. PCR products were cleaned up with Minielute PCR purification kit (Qiagen), and sent for sequencing as described above.

The library pool used at Experiment 2 was a mix of MSTLs 1-9 (see Table 3-1). PCR conditions were the same but annealing temperature was elevated to 64°C as additional 6 nucleotide-long labels were added to the single universal primers (list of oligonucleotides and multiplex distribution in Table 2-3). This enabled pooling of the samples and using a single commercial barcode before sending for sequencing as described above.

2.13 *In vitro* RNAi Target sequencing

The library pool used for the *in vivo* RITseq was defrosted, grown for 24 h and diluted to contain 1 x 10^5 cells ml^{-1} in 100 ml. The culture was then split into two 50 ml flasks (tetracycline induced and uninduced control) and grown for 120 h bringing down the concentration to 1 x 10^5 cells ml^{-1} every day. Before dilution, 1 x 10^7 cells were sampled daily for DNA purification, PCR enrichment and sequencing was conducted as for *in vivo* replicate 2. Time 0 sample was also processed and the whole experiment was conducted in triplicate.

2.14 RITseq data analysis

2.14.1 Processing sequencing data

This part of the work was carried out mainly by Dr. Jonathan Wilkes.

2.14.1.1 Reads mapping

Artificial chromosomes were generated *in silico* by concatenating the 183 amplicons representing all the RNAi targets included in this study, joining the 3'

end of one with the 5' end of the next using a random sequence of 15 bases. The coordinates of each sequence were recorded. The artificial chromosome sequences were indexed to use in Bowtie2 (short-read alignment software) (Langmead and Salzberg, 2012), using default parameters.

For *in vivo* Exp.1, single end reads generated from each sample were selected by the presence of a 9 base long tag [GCCAACTTT] present within the universal primers, permitting 1 mismatch (insertion, deletion or substitution). Approximately 20% of the reads contained the RNAi cassette.

For *in vivo* Exp.2 and *in vitro* RITseq, single end reads were identified and clustered according to our home-made 6 base-long barcodes (Table 2-3) before being processed as in Experiment1 finding 65-70% of reads containing the RNAi target. The absence of a fragmentation step conferred a substantial advantage.

Selected reads were mapped to the artificial chromosomes with Bowtie2 (aligning in local mode). The ".sam" format files generated by Bowtie2 were parsed and the coordinates of the appropriate chromosome, to which the read mapped, were recorded. The mapped reads were assigned to the appropriate protein kinase sequence by reference to the index generated above. The read was assigned if the read sequence overlapped the 3' or 5' end of the protein kinase sequence - or if it laid entirely within the protein kinase sequence.

Bowtie2 mapping was used also to compare whole-genome/kinome-wide RITseq coverage (Figure 3-17). Paired ends were sequenced in both experiments, however only the ends containing the 9mer tag in the universal flank were considered for the analysis.

2.14.1.2 Statistical analysis

In each individual sample, accumulated reads mapping to a particular protein kinase sequence were expressed in Reads Per Million of mapped reads (RPM), i.e. reads normalized by multiplying raw counts 10^6 times, dividing by the sum of total valid reads accepted for the analysis in the whole sample and rounding to the next integer. A bootstrap analysis with 1000 repeats pairing randomly induced and control normalized counts per gene ID enabled calculation of

median induced/control ratios and 95% intervals of confidence. Regression of both *in vivo* replicates was calculated and plotted using GraphPad prism 6.0. T test comparing induced and uninduced normalized reads per gene identifier in each of the *in vivo* replicates, and in each of the time points of the *in vitro* RITseq were calculated. Volcano plots were generated from them. Means and ranges were calculated and plotted with Graphpad too.

2.14.2 Phylogenetic tree with *T. brucei*, *T. cruzi* and *L. major* orthologues

Coding sequences of 49 protein kinases with a loss-of-fitness phenotype and their *Trypanosoma cruzi* and *Leishmania major* orthologues were extracted from the TritrypDB platform (Aslett *et al.*, 2010). They were all aligned against the protein kinase hidden Markov model (HMM), PF00069 (obtained from PFAM (Finn *et al.*, 2014), using Hmmer [HMMER 3.1 (February 2013); http://hmmer.org/] and trimmed to include only sequences conformant with the profile. Resulting clustalW file was converted to ungapped FASTA format. Conformant sequences were aligned using three additional and complementary programs: T-coffee (Notredame, Higgins and Heringa, 2000), ClustalW (Larkin *et al.*, 2007) and MUSCLE (Edgar, 2004). Results obtained from all four alignment tools were combined with T-coffee and the output was processed to remove low quality (<5) and low occupancy (<60%) regions. The resulting high quality alignment was used in Splits Tree (Huson and Bryant, 2006) to generate the final Neighbour-Joining phylogenetic tree (Figure 3-13). An extract from this alignment was also used to depict the putative lack of kinase activity of Tb927.11.9290 (Pseudo-Orphan/FAZ20), Figure 3-18.

Amino acid sequences from the seven protein kinases triggering a greater loss of fitness upon depletion *in vivo* than *in vitro* were aligned using protein-protein Basic Local Alignment Search Tool (BLASTp) (Altschul SF *et al.*, 1990; Altschul *et al.*, 1997) (NCBI) against *Homo sapiens, Saccharomyces cerevisiae, Caenorhabditis elegans, Trypanosoma brucei, Trypanosoma cruzi* and *Leishmania major* non-redundant protein databases Table 3-6.

2.15 Serum tests

200 μl of log-phase parasites in HMI11 at 10^6 cells ml^{-1} were exposed to fresh rat serum for 3 h (10% and 50% concentration). Relative survival was measured of induced and uninduced clones compared to no serum controls (Figure 3-15). The whole experiment was performed in triplicate. Significance of results was assessed with T-test.

2.16 Mild osmotic shock test

2 x 10^6 log-phase cells were collected and resuspended in 500 μl of a solution 55 mM KCl/1 mM glucose to induce cell swelling during 5 mins incubation in ice. Then, a second 5 mins ice incubation was allowed after addition on top of 500 μl of a 263 mM KCl/1.75 mM $MgCl_2$ shrinking solution. Remaining cells were washed once and resuspended in 1 ml of HMI11 media before counting. Survival of the induced cells compared to uninduced controls was plotted (Figure 3-16), and significant results after T-tests indicated with a star. Modified from 'swell dyalisis' protocol (Voorheis and Martin, 1980).

2.17 Alamar blue

2.17.1 Assessing RNAi phenotypes of new-made lines in the pTL library

As done for the rest of the pTL library (Jones *et al.*, 2014), cell cultures seeded at 2 x 10^4 parasites ml^{-1} per cell line were divided into two pools: induced with 1 μg/ml of tetracycline (Sigma), and uninduced. Then 200 μl of each were plated in triplicate in a 96 well plate (Corning) and incubated for 72 h. For the last 24 h, 20 μl of Alamar blue (Resazurin 0.49 mM in PBS) was added and mixed. The plate was then read at $\lambda_{excitation}$ 530 nm and $\lambda_{emission}$ 590 nm in an Envision Plate Reader (Perkin Elmer). The ratio of the values for induced wells over uninduced wells was calculated and used to produce a mean and a standard deviation. Values under 0.8 were considered a loss of fitness.

2.18 Flow cytometry

2.18.1 Differentiation analysis

All this processing was done using ice-cold buffers and centrifuges at 4°C. 3 x 10^6 parasites were pelleted per analysed condition by centrifugation at 1500xg for 10 mins, and washed with phosphate buffered saline (PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 2 mM KH_2PO_4). Then, cells were incubated 1 h on ice with a primary mouse IgG targeting EP procyclin (Clone TBRP1/247, Cedarlane), used at 1:500. After three washes with PBS, a secondary rabbit IgG anti-mouse conjugated to allophycocyanin (APC) (Clone M1-14D12, Affimetrix) was used at 1:1000 , and incubated on ice 1 h in the dark. Three more washes were conducted with PBS and cells were transferred to flow cytometry tubes before fixing with 2% paraformaldehyde (PFA) for at least 5 mins. If convenient, cells were kept in the fridge for several days at this stage. After 3 rounds of PBS washing parasites were resuspended in 200 μl of PBS and passed through nylon filtration cloth of 200 μm opening (Nitex). Samples were analysed in a MACSQuant® Analyzer 10 (Miltenyi Biotec) recording 100000 events. Trypanosomes were gated in the forward versus side scatter plot, single cells selected by comparison of width and height for the forward scatter signal, and EP procyclin positive population selected taking as a negative reference the uninduced controls or the parental line. Results were analysed in FlowJo v10 (FlowJo, LLC).

2.18.2 Cell cycle analysis

For cell cycle analysis, DNA of cells was stained with propidium iodide (Sigma) or Hoechst 33342 (Thermo Scientific). Fluorescence intensity of these molecules for single cell events is proportional to their DNA content. The distribution of the whole analysed sample was estimated across the distinct cell cycle stages based in the histogram profile of DNA-associated fluorescence in the entire population.

2.19 Immunofluorescence assay (IFA)

10^6 parasites per sample were pelleted at 1000xg for 10 mins, washed in PBS, and, after resuspending in 25 μl PBS, placed in one well of a 12-well slide coated

with poly-L lysine solution 0.1% w/v (Sigma-Aldrich) incubated for 4 mins. Each well used in the slide was marked in the margins with a PAP pen (Abcam) making a hydrophobic barrier that avoided buffers diffusing between wells. Cells were then fixed with 4% paraformaldehyde (PFA) and incubated for 15 mins. Sometimes they were left overnight at this stage in a wet chamber. After three washes with PBS, cells were treated with 100 mM glycine to neutralize excess of PFA. After three more PBS washes, if needed, permeabilisation was conducted with 25 µl 1x PBS/Triton X-100 (Thermo Scientific) for 10 mins. Then 25 µl of primary antibody were incubated for 1 h in a wet chamber, followed three more washes. After, 25 µl of the secondary antibody were incubated for 1 h. Finally after three washes, DAPI Fluoromount-G (SouthernBiotech) was added and cover slip sealed with nail varnish after 4 mins incubation.

Table 2-10. List of antibodies used in immunofluorescence assay (IFA).

Antibody	Serotype	Clone	Concentration	Supplier
αVSG 221 variant	Rabbit IgG	N/A	1:8000	R. McCulloch
αMyc	Mouse IgG	4A6	1:7000	Millipore
αEP procyclin	Mouse IgG	TBRP1/247	1:500	Cedarlane
αMouse Alexa Fluor® 488	Goat IgG	N/A	1:1000	Thermo
αRabbit Alexa Fluor® 594	Goat IgG	N/A	1:1000	Thermo

2.19.1 Cell cycle analysis

IFA slides stained with 4',6-diamidino-2-phenylindole (DAPI) were imaged at a magnification that maximized the number of well-isolated parasites that it was possible to count in a single picture. Images were analysed in Fiji ImageJ updated with a cell counting plugin. 200 parasites were counted per condition classified in G0-G1 (1 nucleus, 1 kinetoplast = 1N1K), G2 (1N2K), M (2N2K) and other aberrant combinations.

Chapter 3 Protein kinase-focused RNAi screen defines virulence factors of *T. brucei* required for proliferation in mammals

3.1 Introduction

3.1.1 RNA of interference (RNAi)

The field of ribonucleic acid (RNA) biology has become an exciting research area within the last twenty years as an increasing number of non-coding functions have been unveiled for these molecules. The functional repertoire of the 'small RNA' universe known in the early nineties -with roles in translation (tRNA and 5S RNA), pre-mRNA processing (by small nuclear RNAs), and protein translocation across the endoplasmatic reticulum (7SL RNA)- experienced a dramatic expansion after the discovery of 20-30 nucleotide-long small silencing RNA molecules (ssRNA) able to switch off gene expression: the so-called pathway of RNA interference (RNAi). It was characterised for the first time in plants (Napoli, 1990; van der Krol *et al.*, 1990), termed RNAi and assigned to double stranded RNA in the nematode *Caenorhabditis elegans* (Fire *et al.*, 1998), and was soon described for many other eukaryotes, including *Trypanosoma brucei* (Ngô *et al.*, 1998). Two main types of ssRNAs were identified in protozoa: small interfering RNA (siRNA), binding specific mRNAs and targeting them for degradation; and micro RNA (miRNA), annealing mainly to the 3' UTR of certain transcripts and inhibiting their translation (Atayde, Tschudi and Ullu, 2011; Kolev, Tschudi and Ullu, 2011). In *T. brucei* miRNA have been predicted but not detected though (Mallick, Ghosh and Chakrabarti, 2008; Atayde, Tschudi and Ullu, 2011). In addition, metazoan also have a third type: PIWI-interacting small RNAs (piRNAs) (Siomi *et al.*, 2011).

The RNAi pathway in *Trypanosoma brucei* has been reviewed in the literature (Atayde, Tschudi and Ullu, 2011; Kolev, Tschudi and Ullu, 2011). It involves two Dicer-like proteins: TbDCL1 (cytoplasmic) and TbDCL2 (nuclear) (Shi, Tschudi

and Ullu, 2006; Patrick *et al.*, 2009). Both have RNase III activity and nuclease-cleave double stranded RNA (dsRNA) into ssRNA. TbDCL1 activity is assisted by cofactor TbRIF5 (Barnes *et al.*, 2012). The resulting 20-30 nucleotide dsRNA fragments are protected from nucleases through methylation in their 3' end by HEN1 activity (Horwich *et al.*, 2007; Patrick *et al.*, 2009). After separation of the two strands mediated by exonuclease TbRIF4 (Barnes *et al.*, 2012), the targeted transcript is specifically bound by the complementary half of the ssRNA and loaded into another protein, AGO1 'slicer'. This assembly is called RNA-induced silencing complex (RISC) and is responsible for the ultimate mRNA cleavage that limits gene expression (Durand-Dubief and Bastin, 2003; Shi *et al.*, 2004).

The siRNA pathway has several endogenous functions: maintenance of genome integrity through transposon and retrotransposon silencing; post-transcriptional regulation; defence against viral invasion; chromosome segregation; and heterochromatin formation. The RNAi pathway has also been exploited as a reverse genetics tool, becoming a real breakthrough in the post genomic era. Genetic manipulation enabled the exploitation of RNAi machinery to selectively target genes for transcript depletion. If their role was unknown, observation of the resulting phenotypes permitted examination of their putative functions. If a gene's role had been already identified, RNAi could help understanding of how its absence or lack of activity affected other factors in the cell (Atayde, Tschudi and Ullu, 2011; Kolev, Tschudi and Ullu, 2011).

Unlike other Eukaryotes, *T. brucei* lacks RNA-dependent RNA polymerase (RdRp), meaning it is unable to expand or preserve exogenous dsRNA after replication. Thus, the effect after transfection of dsRNA is only transient (Siomi and Siomi, 2009; Kolev, Tschudi and Ullu, 2011). Full exploitation of RNAi was therefore not possible until development of a system for inducible expression of dsRNA. This involved integration of a DNA template in a silent genomic locus, from which the desired dsRNA could be inducibly expressed under control of a strong promoter controlled by a tetracycline operator (Orth *et al.*, 2000; Shi *et al.*, 2000). Since then, two different dsRNA expression strategies have been used: the 'stem-loop', and the use of 'opposed promoters'. In the first system, the DNA sequence integrates in 'sense' and is followed by the complementary sequence in 'antisense' separated by a linker region. This allows formation of a double

stranded hairpin upon transcription by self-annealing of the RNA transcript. In the second approach, the desired RNAi target integrates in the genome flanked by opposed promoters. After transcription, complementary strands will result that anneal, producing the expected dsRNA. Both strategies were compared, showing stem-loop to have a better efficiency for mRNA depletion (Durand-Dubief, Kohl and Bastin, 2003). For the Tet-operator to be effective, cell lines need to be modified to express the tetracycline repressor from *Escherichia coli*, which blocks transcription of the gene located downstream of the promoter until displaced by tetracycline (Wirtz and Clayton, 1995).

Many variants of these systems, combining different genomic loci and promoters, have been applied to *T. brucei* research. The most commonly used has been the EP/GPEET procyclin promoter; however, this is down-regulated in mammal-infective forms. In bloodstream forms one of the most widely used promoters is T7, which requires parasites to be modified to express the T7 RNA polymerase (Wirtz *et al.*, 1999). High levels of expression of this polymerase are toxic for trypanosomes. Therefore, weak promoters must be used in order to achieve balanced levels between efficacy and toxicity. RRNA promoters have been proposed instead as a robust endogenous alternative that can be used in *T. brucei* regardless the life stage and permitting performance of differentiation studies (Alsford *et al.*, 2005).

Among the silent loci most successfully used to target inducible expression cassettes are the rRNA spacers. However, the *T. brucei* genome contains nine rRNA loci, each providing different transcription levels. After targeting a GFP reporter with the hygromicin resistance cassette (*HYG*), the best expressing locus was selected. Then, GFP was replaced with a puromycin resistance gene (*PAC*), while leaving *HYG* with only its 3´end as a non-functional tag. Such a strategy enabled accurate targeting of any RNAi construct to the tagged *RRNA* locus by engineering it to contain the missing 5´end of *HYG*. After successful homologous recombination, RNAi clones recover hygromycin resistance, while normally losing *PAC* (Alsford *et al.*, 2005).

3.1.2 RNAi screens in *Trypanosoma brucei*

Since RNAi technology became available for *T. brucei* research, several high throughput functional and phenotypic screens have been undertaken using a variety of RNAi libraries. A software was released providing an algorithm to rapidly select and design RNAi targets, called 'RNAit' (Redmond, Vadivelu and Field, 2003). RNAit identifies the best possible primer pair (e.g. temperature of melting, self-annealing, GC content) to amplify a gene region of the desired RNAi target size, and aligns the predicted amplicon against the parasite´s genome in BLASTn (Altschul *et al.*, 1997) in order to certify specificity and avoid off-target effects of RNAi.

A key tactical development was the first attempt of simultaneous (or parallel) phenotyping with a library produced by introduction of sonicated, random genomic fragments as RNAi targets. It was performed in insect-infective procyclic forms to obtain a 5-fold genomic coverage. Unfortunately, that could not be achieved at the time in mammal-infective bloodstream forms due to a limited efficiency of transfection. The library was used to identify genes reducing expression of the main surface glycoprotein, EP-procyclin, by detecting a loss of affinity for concanavalin A (Morris *et al.*, 2002). Final identification of RNAi that successfully reduced EP-procyclin expression required clonal isolation of the selected parasites and RNAi target amplification and sequencing. This approach has been applied successfully in procyclic forms for identification of drug transporters (Schumann Burkard, Jutzi and Roditi, 2011) and mRNA regulators (Schumann Burkard et al. 2013).

The first systematic screen in bloodstream forms targeted the 197 genes encoded in Chromosome 1, analyzing in independently developed cell lines a range of phenotypic characteristics, including loss of fitness (Subramaniam *et al.*, 2006). The approach developed by Morris et al. (2002) was also adapted to bloodstream forms in order to identify components of cAMP signalling pathways (Gould et al. 2013). But two major breakthroughs then emerged: increased affordability of Next Generation Sequencing (NGS), and development of the Sce* *T. brucei* cell line that allows hyper-efficient transfection in mammal infective bloodstream parasites (Glover and Horn, 2009). The Sce* *T. brucei* line expresses a site-specific meganuclease, I-SceI, able to induce double stranded DNA breaks

in a locus tagged with a targeting sequence. Attraction of the homologous recombination machinery to repair the double-stranded break increased efficiency of recombination of exogenous constructs directed to the same location (Glover and Horn, 2009). The library of sonicated genomic fragments used previously in procyclic forms (Morris *et al.*, 2002) could now be transfected into bloodstream forms with a predicted 11x coverage; NGS increased the throughput of RNAi target identification in parasites resisting any given selective pressure; and the technique could be used for quantitative analysis by comparison of normalized reads in the RNAi induced samples with those in the non-induced controls. The overall approach was termed 'RNAi target sequencing' (RITseq) and became the most popular strategy for high throughput dissection of trypanosome gene function (Alsford *et al.*, 2011, 2012).

The main limitation of the RITseq approach was the need for huge numbers of parasites to ensure whole genome coverage. For mice infections, we need to reduce the inoculum to the minimum in order to avoid overwhelming the host, causing early mouse death (Turner, Aslam and Dye, 1995). Moreover, if the library is made in a monomorphic cell line, which is unable to produce cell cycle-arrested stumpy forms in response to the still uncharacterized stumpy induction factor (Vassella *et al.*, 1997).

3.1.3 A kinome wide *in vitro* screen: the pTL library

An RNAi screen of the complete protein kinome of *T. brucei* was recently published in bloodstream form cells, looking for potential parasite-specific drug targets, regulators of the cell cycle, and protein kinases involved in life stage maintenance (Jones *et al.*, 2014). 183 individual cell lines were generated, in duplicate, ('pTL' library) containing tet-inducible stem loop RNAi cassettes. The vector used, pGL2084, was a modified version of pRPa^{iSL} (Alsford and Horn, 2008), containing ligation-free Gateway® sites (AttP1 and AttP2), Figure 3-1. This allowed directional cloning of both fragments as a stem loop (sense and antisense) in a single reaction. The approach increased throughput, avoiding two rounds of restriction digestion and ligation per construct. In addition, the system avoided problems of incompatibility when restriction sites needed for cloning in pRPa^{iSL} were present in the RNAi target.

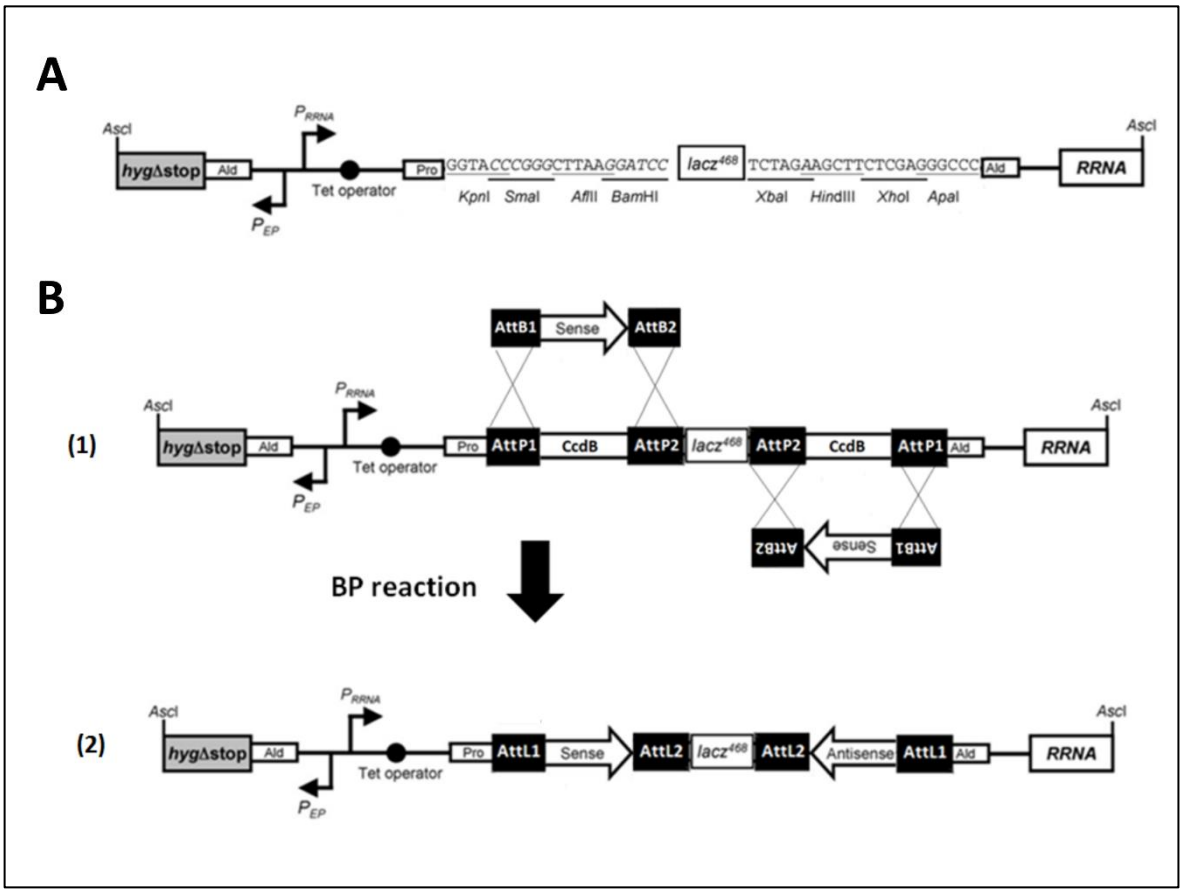


Figure 3-1. Gateway® based cloning strategy used to make the pTL library. **A.** AscI-linearized section of the backbone vector pRPa^{iSL} as it is targeted for homologous recombination to the best expressing *RRNA* locus, tagged in 2T1 cells with the terminal section of the hygromycin resistance gene (Alsford *et al.*, 2005). Includes aldolase (Ald) and procyclin (Pro) intergenic regions; RRNA (P_{RRNA}) and EP procyclin (P_{EP}) promoters; Tet operator for induction with tetracycline in 2T1 recipient cells (expressing Tet repressor); a lacz^{468} stuffer region to produce the hairpin; and multicloning site to assist directional ligation of the RNAi inserts. **B.** Gateway®-modified version of (A), including *ccdB* toxic gene and AttP flanks. Below, schematic representation of the BP reaction for directional cloning. Sequences containing the RNAi target amplified with primers including AttB flanks. Modified from Alsford et al. 2005.

Gateway® cloning is based on the site-specific recombination system of phage lambda (Katzen, 2007). After synthesis or amplification of the RNAi target with primers flanked by AttB1 and AttB2 sites, BP clonase mediates recombination. pGL2084 contains a *ccdB* gene, which is toxic for standard *E. coli* strains, between each pair of AttP1 and AttP2 sites. Successful recombination replaces it with the RNAi target and enables positive selection.

In the published kinome RNAi study cell survival was analysed in culture with Alamar Blue®, comparing uninduced controls with induced knockdowns after 72h

of RNAi induction. 45 protein kinases were identified showing a loss of fitness *in vitro*, which was classified as “slow growth”, “cell cycle arrest” or “death” phenotypes. As protein kinases of *T. brucei* have been shown to be druggable (Urbaniak *et al.*, 2012), they can be considered potential drug targets. 21 of these protein kinases had never been reported to cause a loss of fitness before. 24 cell lines in the library displayed some kind of cell cycle defect and two acquired procyclic-like phenotypes following RNAi induction (e.g. kinetoplast repositioning and EP procyclin expression).

3.2 Research aims

Many biological roles for protein kinases have been revealed in *T. brucei*, but they have been observed mainly in the artificial conditions of *in vitro* culture. In this chapter I address a key question of relevance both for biology and drug discovery: which protein kinases are required for survival of the parasite in the environmental conditions encountered in the mammalian bloodstream? Potential drug targets identified by RNAi studies *in vitro* require studies in mice to validate the targets. Multiplexing *in vitro* studies to genome-wide level has been implemented but not yet applied *in vivo*. Improving this would reduce the number of animals required to perform these studies. The adaptation of RITseq to simultaneously analyse the 177 cell lines contained in the pTL library, potentially using a single mouse, would provide new insights in *T. brucei*-host interaction and reveal novel potential drug targets for protein kinase inhibitors.

3.3 Results

3.3.1 Kinome-wide *in vitro* and *in vivo* RNAi screening; the approach

In order to increase our capacity for screening the kinome RNAi library, we made a pool of the 177 cell lines contained in the ‘pTL’ library (see Table S1), which target 183 protein kinases because 6 are double knockdowns (Jones *et al.*, 2014). This pool allowed parallel phenotyping of the whole population in a single culture (*in vitro*) and in a single mouse (*in vivo*) through RITseq (Figure 3-2).

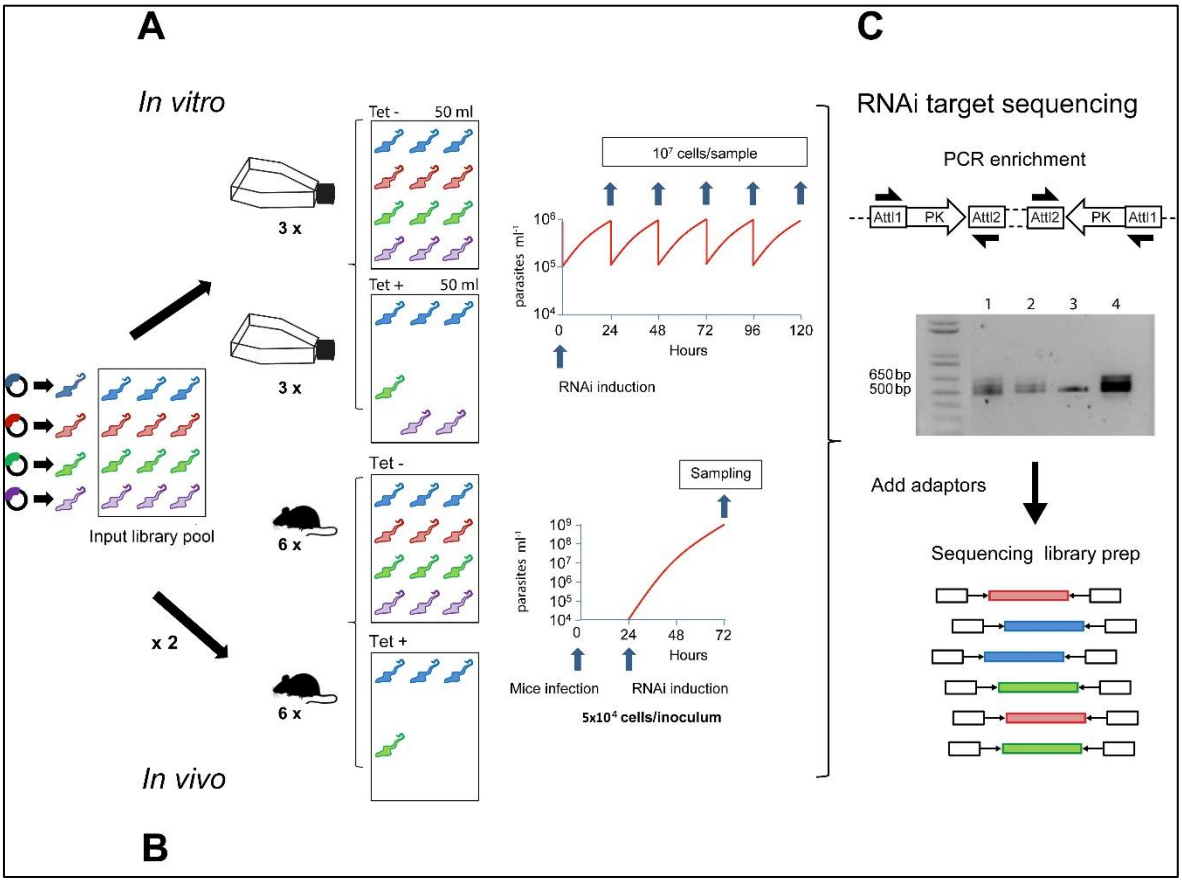


Figure 3-2. *In vitro* and *in vivo* phenotyping of a *T. brucei* kinome RNAi library. Schematic representation of the experimental workflow.
A. A pre-inoculation pooled kinome RNAi library was diluted to contain 1 x 10^5 cell ml^{-1} in 100 ml and grown in culture for 24 h in triplicate. The culture was then split into two flasks, one was induced with tetracycline (Tet+) and the other remained uninduced (Tet-). 1 x 10^7 cells were sampled every 24 h over 5 days for DNA purification and cultures diluted daily to contain 1x10^5 cells ml^{-1}. **B.** 5 x 10^4 *T. brucei* bloodstream form parasites of the pooled kinome RNAi library were injected intraperitoneal into 12 CD1 mice and, 24 h post inoculation, RNAi was induced with doxycycline in 6 animals (Tet+) and 6 were left uninduced (Tet-). 48 h and 72 h post RNAi induction, parasites were purified from blood and genomic DNA prepared. **C.** PCR enrichment in the RNAi target was carried out with a single primer binding at the universal Gateway® AttL sites flanking the 500-600 bp protein kinase-specific inserts. The gel example shows RNAi target distribution in 4 different samples: Tet- (1), Tet + (2), a single line control (3), and the input library pool (4). The experiment was performed twice and PCR-enriched samples from each biological replicate were sequenced with different platforms.

At the required time points, parasites were isolated from *in vitro* cultures or mice blood, using DEAE cellulose (Lanham and Godfrey, 1970). Then, genomic DNA was purified from all samples (including input library), and polymerase chain reaction (PCR) was used to enrich the RNAi target. This PCR was conducted with a single primer that hybridized to the universal regions of the RNAi

cassette, located at both flanks of the protein kinase-specific sequence (Figure 3-2).

Cytocidal or cytostatic phenotypes were identified by a quantitative reduction in reads mapping to a particular RNAi target in the induced samples relative to their representation in the uninduced controls. For both the *in vitro* and the *in vivo* experiments, reads were filtered to include only those that contained a 9-nucleotide tag encoded in the universal PCR primers and were then mapped against the concatenated sequences of the 177 protein kinase RNAi targets.

For the *in vitro* experiment, three biological replicates were conducted, collecting samples daily during 120 h (5 days). Cells were diluted after sampling to maintain the exponential phase of growth, and tetracycline was renewed as it has stability for 4 days at 37 ºC according to manufacturer instructions.

For the *in vivo* experiment, two biological replicates were performed. Each replicate included 6 control mice, and 6 where RNAi was induced with doxycycline, which is more bioavailable and stable than tetracycline *in vivo* (Agwuh and MacGowan, 2006). Working with monomorphic parasites has the inconvenience of cells being unable to monitor and control their own density and thus rapidly causing death of the host. Ethical commitments oblige us to kill mice when parasitemia reaches 10^8 cells ml^{-1}. With this in mind, infections were initiated with 5 x 10^4 parasites. This provided >250 parasites per kinase line, assuming we were able to keep homogeneous representation in the final pool; and 3 days growth before reaching the parasitemia limit. One mouse has roughly 1.4 ml blood, with parasites growing faster and reaching higher concentrations than in culture. The choice of 6 replicates helped us reach statistically significant numbers and prevented eventual fails during establishment of the infection, as happened for one of the uninduced mice within the first replicate (see Table S2). Infections were allowed to establish for 24 h before adding doxycycline in the drinking water. RNAi inductions were then carried out for 48 h before killing mice and recovering parasites from the blood.

3.3.2 Pooling strategy

It was critical to, as much as possible, ensure equal representation of every cell line in the final parasite protein kinase library pool. In order to do so, a pooling strategy was adopted by dividing the library into 9 sub-groups, named "Mixed Stabilate Trypanosome Library" (MSTL1-9). Each MSTL contained between 18-27 RNAi lines in order to ease handling and avoid overgrowing of cells. Cultures were always maintained below $2x10^6$ cells x ml^{-1}. For each cell line, the clone producing a stronger loss of fitness in the Alamar blue screen (Jones *et al.*, 2014) was selected. Cells were defrosted, and passaged in fresh media until all lines in the batch presented normal fitness, assessed by shape and mobility under the microscope. When new cryostocks were needed, they were generated. Once they were in optimal conditions, cells were pooled and frozen. 24 cryovials were generated per sublibrary, and 2.5 $x10^5$ cells/kinase were included in each one. This task was undertaken in collaboration with Dr Tiago D. Serafim.

One tube was defrosted per sublibrary in order to estimate the survival rate. The operation was conducted three times, with a space of several months between replicates. The number of parasites per kinase recovered in each case was very reproducible for each MSTL, Figure 3-3. However, the percentage of recovery per kinase line between the different libraries was variable, from 30% in MSTL4 to over 90% in MSTL7. Since we put effort on ensuring homogeneous fitness, we made the assumption that cell loss upon a freezing/thawing cycle was equitable for all the lines included in a stabilate. It is positive that the "limiting sublibrary", *i.e.* the one that contained the minimum amount of parasites per kinase line, was the same in every replicate, with the exception of the third that included MSTL9 and a new version of MSTL2 (MSTL2*), Figure 3-3.

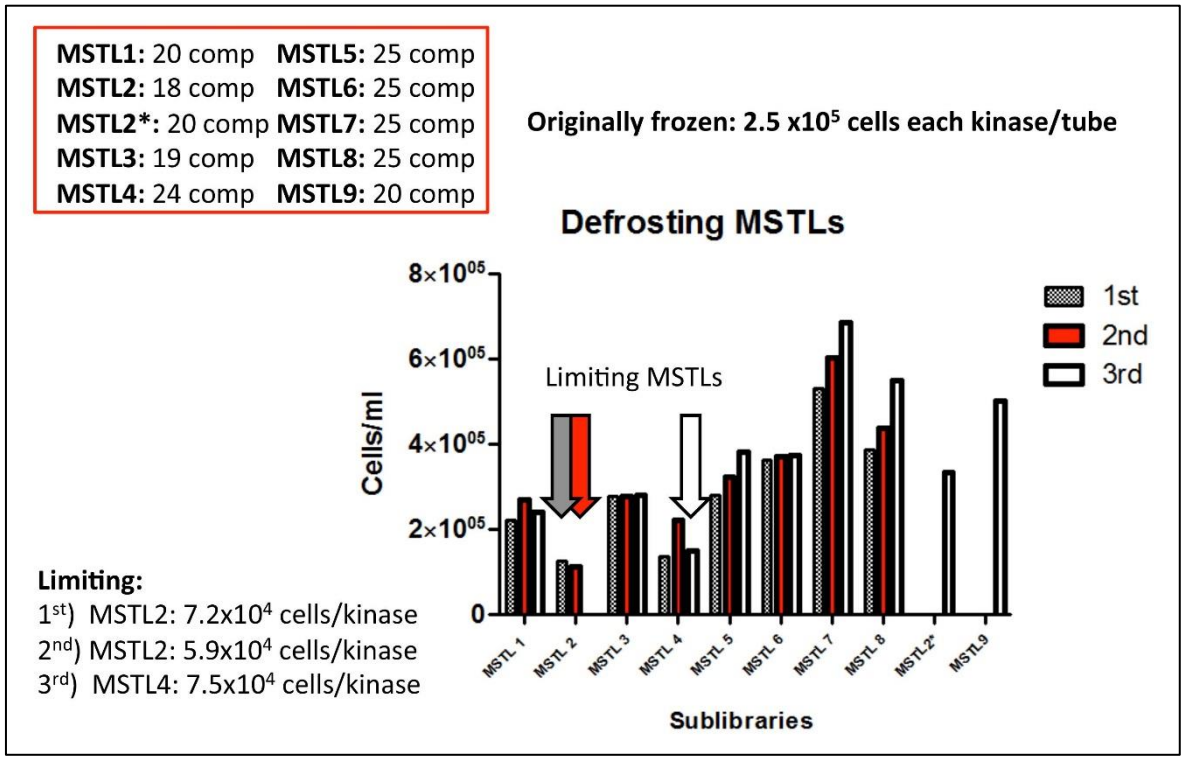


Figure 3-3. Defrosting survival assay for sublibraries (MSTL1-9) used in the *T. brucei* protein kinome RITseq.
3 tubes out of the 24 generated for each MSTL were defrosted with a space of two months between replicates. Highlighted in a red box, are number of kinase lines contained per MSTL. Codes for the cell lines and plasmids, together with alamar blue phenotypes and RNAi target sequences, can be found in Table S1. The third replicate was slightly different as MSTL2 was remade to contain two extra cell lines, and MSTL9 was added to include lines remade after they were not detected in the first replicate for the *in vivo* screen; details are provided below.

The final pool was made 7 h before using the library to inoculate mice or to seed the *in vitro* experiments. This kept the cells in culture for less than 1 population doubling (8 h in 2T1s). After defrosting all MSTLs, glycerol was diluted and survivors accurately counted to identify the "limiting sublibrary". Number of cells was maximized by adding to the mix the complete culture for that particular MSTL, and, for the rest, volumes contributing the same number of cells per kinase line than the "limiting sublibrary". Details for the two pools used for the *in vivo* and *in vitro* RITseq experiments are provided in Table 3-1. MSTL2 was the limiting library in the first pool and MSTL4 in the second (as MSTL2 was remade). 2 cell lines were added individually to the pool in replicate 1 in order to see if any difference was observable in our detection capacity compared to the premixed lines: Tb927.9.6560, which produced a loss of fitness in the alamar blue *in vitro* screen (Jones *et al.*, 2014); and Tb927.6.2980, which

did not. The number of reads collected for them in the final mix after RITseq was in the same range as the other pooled lines, and the expected phenotypes after RNAi were reproduced (fit and not fit)(Tables S2 and S3). These cell lines were included in MSTL2* for the second replicate. In replicate 2, another 5 lines were added individually, as they could not be found in the first pool after sequencing. Details will be discussed below.

Table 3-1. Input library composition for *in vitro* and *in vivo* RITseq experiments determining limiting sublibrary.

Pool 1. *In vivo* Replicate 1

Sublibrary	Components	Post-thawing concentration (cells/ml)	Culture volume (ml)	Total cells	Cells/ kinase line		Volume to mix (ml)
MSTL1	20	2,20E+05	11,1	2,44E+06	1,22E+05		6,6
MSTL2	18	1,26E+05	10,3	1,30E+06	7,21E+04	**Limiting**	**ALL: 10,3**
MSTL3	24	2,77E+05	10,8	2,99E+06	1,25E+05		6,3
MSTL4	19	1,37E+05	10,5	1,44E+06	7,57E+04		10,1
MSTL5	25	2,80E+05	11	3,08E+06	1,23E+05		6,5
MSTL6	25	3,63E+05	10,4	3,78E+06	1,51E+05		5
MSTL7	25	5,30E+05	10,7	5,67E+06	2,27E+05		3,4
MSTL8	25	3,87E+05	10,7	4,14E+06	1,66E+05		4,7
sTL206	Tb927.9.6560	6,00E+04					1,21
sTL209	Tb927.6.2980	9,70E+04					0,749

Pool2. *In vivo* Relicate 2, and *in vitro* Replicates 1-3

Sublibrary	Components	Post-thawing concentration (cells/ml)	Culture volume (ml)	Total cells	Cells/ kinase line		Volume to mix (ml)
MSTL1	20	2,40E+05	9,5	2,28E+06	1,14E+05		6,3
MSTL2*	20	3,33E+05	9,6	3,20E+06	1,60E+05		4,5
MSTL3	24	2,80E+05	9,6	2,69E+06	1,12E+05		6,4
MSTL4	19	1,50E+05	9,5	1,43E+06	7,50E+04	**Limiting**	**ALL: 9.5**
MSTL5	25	3,80E+05	9,5	3,61E+06	1,44E+05		4,95
MSTL6	25	3,73E+05	9,6	3,58E+06	1,43E+05		5,03
MSTL7	25	6,86E+05	9,5	6,52E+06	2,61E+05		2,74
MSTL8	25	5,50E+05	9,6	5,28E+06	2,11E+05		3,41
MSTL9	20	5,00E+05	9,7	4,85E+06	2,43E+05		3
sTL102	Tb927.11.12310	1,08E+06					0,069
sTL284	Tb927.3.5180	1,05E+06					0,071
sTL598	Tb927.11.7160	7,80E+05					0,096
sTL635	Tb927.11.5860	8,00E+05					0,093
sTL641	Tb927.9.2240	8,83E+05					0,084

The pooled culture was diluted to 50 ml and acclimatized at 37 ^{o}C 5% CO_2 for 45 min, cells were counted, pelleted at 1200 g for 10 minutes and resuspended at

$5x10^5$ cells ml^{-1}. After 6 h incubation 37 ^oC 5% CO_2, cells were counted again and used for the *in vitro* or *in vivo* experiments.

3.3.3 Optimization of PCR to enrich the RNAi cassette

Gene-specific RNAi inserts were designed using a web-based tool, RNAit (Redmond, Vadivelu and Field, 2003), and ranged between 450 and 650 base pairs. Such restricted size, in addition to the Gateway® flanks, AttL1 and AttL2, were exploited to design a universal PCR, which enabled homogenous enrichment of the RNAi targets from genomic DNA and increased efficiency of the sequencing reaction (Figure 3-2C). It was critical to use a high fidelity polymerase to ensure amplified sequences were indistinguishable from their templates, so we could unequivocally assign sequencing reads to the original templates.

After alignment in CLC genomics workbench, AttL1 and AttL2 displayed in nearly identical sequences, differing in only 4 bases. Based on that, a variety of primers were designed, trying to achieve optimal PCR amplification (Figure 3-4A): two forward primers (Fw1 and Fw2) binding to AttL1; two reverse (Rv1 and Rv2) hybridizing in the stuffer region; and two reverse (Rv3a and Rv3b) binding to AttL2, Figure 3-4B.

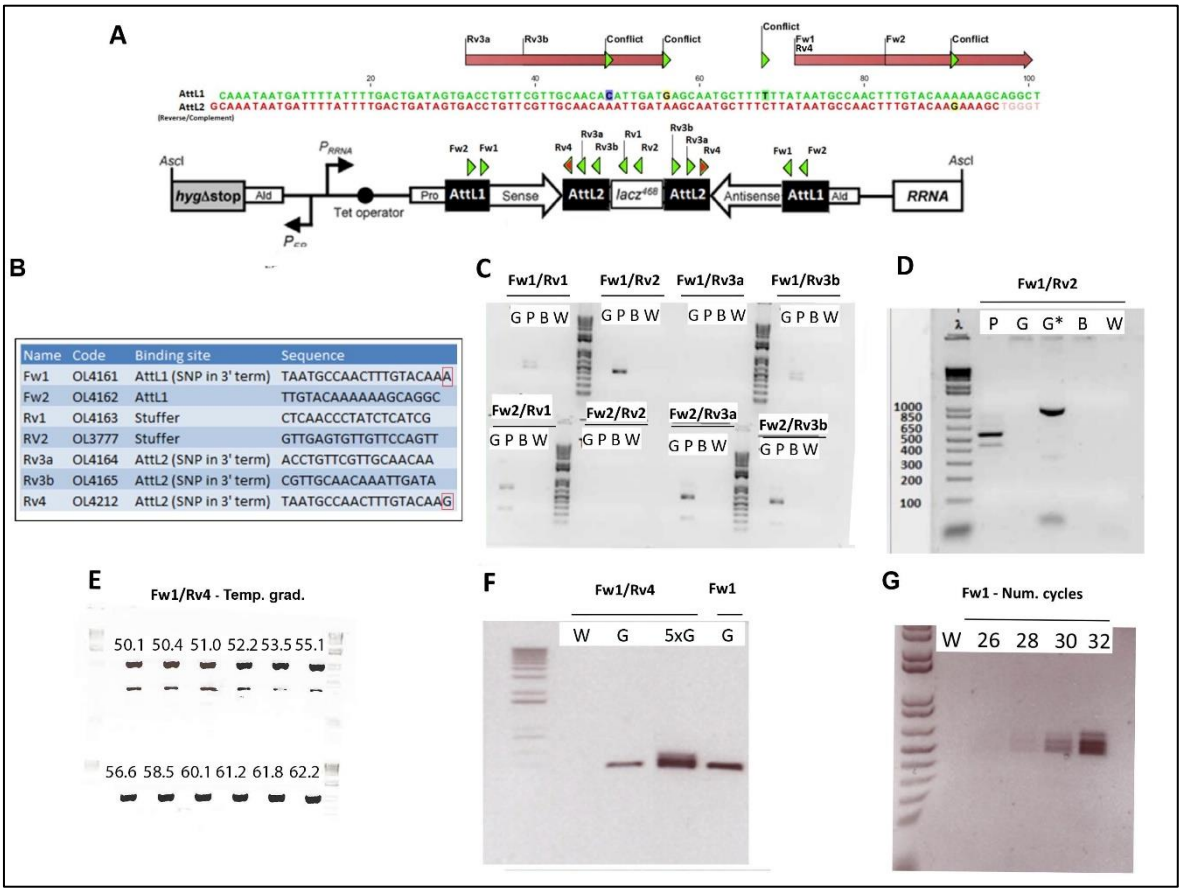


Figure 3-4. PCR optimization for enrichment in the RNAi insert.
A. (Above) Sequence alignment of the AttL sites within the universal Gateway® flanks. Highlighted on top, forward (Fw) and reverse (Rv) primers designed exploiting the single nucleotide conflicts between sequences. **(Below)** Schematic representation of the RNAi cassette (see **Figure 3-1**) indicating binding sites for all the oligonucleotides tested. **B.** List of all primers tested during the optimization process. **C.** Phusion® high fidelity polymerase reaction (60 ^{o}C annealing/30 cycles) with different combinations of primers. Depending on the template, lanes are labelled: G=gDNA from an RNAi line (sTL0082); P= DNA from its untransfected plasmid (pTL50); B=genomic DNA from parental 2T1s; W=water control. **D.** Fw1/Rv2 PCR with Phusion® high fidelity polymerase (60 ^{o}C annealing/30 cycles). Lanes labelled as in (C) with G*= positive control for PCR with gDNA from sTL0082 using primers to amplify the kinase domain of Tb927.9.10920. **E**. Fw1/Rv4 gradient PCR using pTL50 DNA and Phusion polymerase for 30 cycles, annealing temperature highlighted above each lane. **F.** Q5 high fidelity polymerase reaction (61^{o}C annealing/30 cycles) using Fw1/Rv4 and single primer Fw1. Lanes labelled as in (B) with 5xG= gDNA from 5 RNAi lines selected from the library. **G.** Fw1 (OL4161) Q5 high fidelity polymerase reaction (61 ^{o}C annealing/input kinase pool as a template) increasing number of PCR cycles.

All combinations were tested with Phusion High Fidelity polymerase at 60 ^{o}C. As a DNA template we used a plasmid from the library (pTL50) and genomic DNA (gDNA) from the cell line generated upon pTL50 transfection into 2T1 cells (sTL0082), Figure 3-4C. The construct and cell line targeted ATR by RNAi, a

protein kinase involved in DNA repair, which has been validated previously somewhere else: the plasmid through restriction digestion and sequencing; and the cell line via antibiotic resistance, RNA/protein depletion, and decrease in resistance to genotoxic effects caused by methyl methanesulfonate (Stortz, PhD thesis, 2016). As a control for PCR, primers were also tested against gDNA from parental 2T1 cells, and water. We only could obtain a single defined PCR product that matched the predicted 628-bp with Fw1/Rv2. This size includes a small AttL1 portion (starting from Fw1), plus the RNAi target with the complete Attl2 flank and the section of the stuffer before Rv2 binding site. PCR amplification was only possible with the plasmid template, not with gDNA.

In order to verify that gDNA from sTL0082 was not degraded we used PCR conditions and primers validated to amplify the kinase domain of Tb927.9.10920 (kindly provided by Nathaniel Jones) in parallel to the Fw1/Rv2 PCR, Figure 3-4D. This time, more DNA was loaded in the gel in order to increase resolution. We observed, on the one hand, that genomic DNA was fine because a band could be identified at the expected size for the genomic control (882 bp), and on the other hand, that a secondary amplicon was produced with Fw1/Rv2. After sequencing the secondary band, we found it corresponded to the 488 bp-long product of Fw1 (OL4161) acting both as forward and reverse primer, binding to AttL2 with a single mismatch at the 3´end (G instead of A), Figure 3-4B.

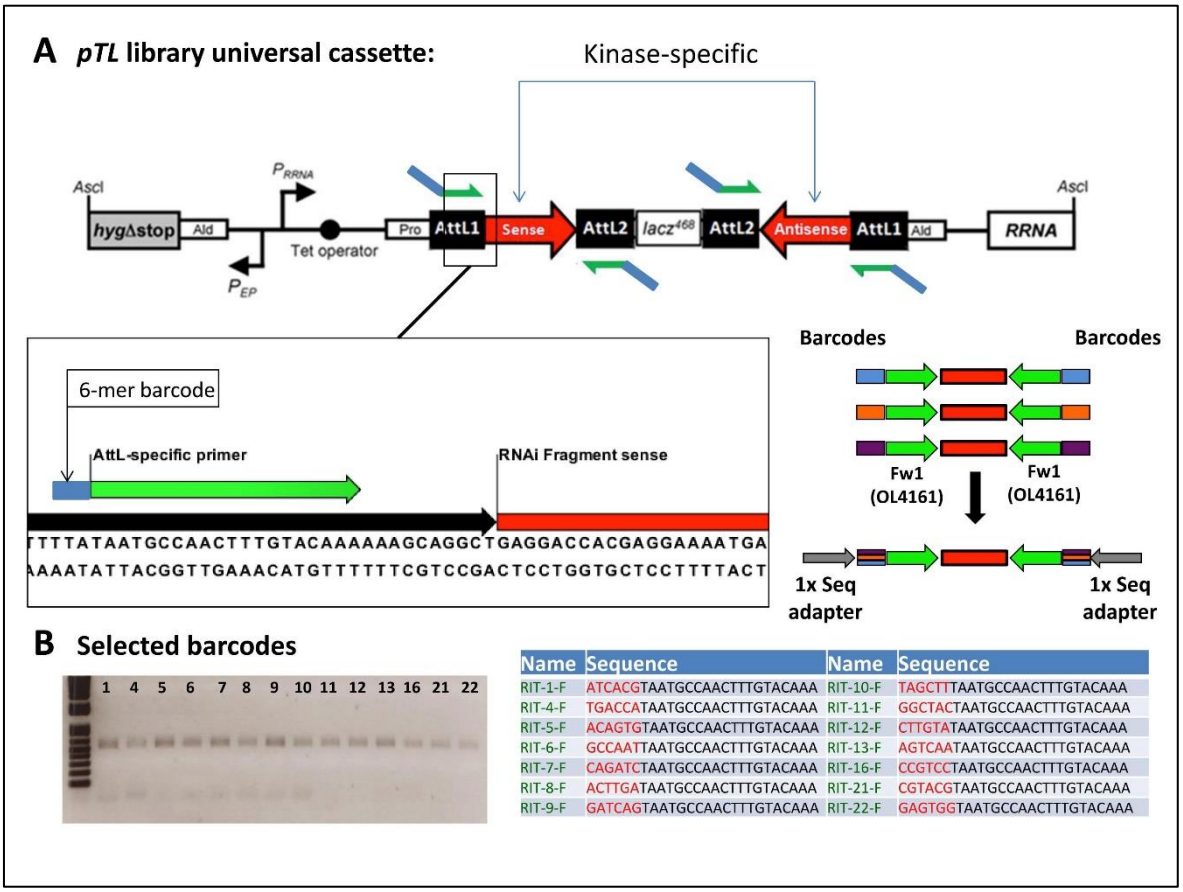
To solve this problem, a reverse primer targeting the site where Fw1 hybridized unspecifically to AttL2 was designed (Rv4). This enabled production of a single band, and 61 ^{o}C was selected as the annealing temperature after a gradient PCR, using pTL50 as a template, Figure 3-4E.

Applying the same conditions with Fw1/Rv4 and the more efficient Q5 High fidelity polymerase, we were finally able to PCR amplify the RNAi target from sTL0082 gDNA alone and pooled with gDNA of other 5 lines included in MSTL1, which targeted: Tb927.2.2260 (584 bp PCR product), Tb927.8.7450 (490 bp), Tb927.11.850 (428 bp), and Tb927.1.1000 (437 bp). In parallel, a PCR was performed in the same conditions with Fw1 alone, concluding that its efficiency was better than using Fw1/Rv4 combined, Figure 3-4F.

In Figure 3-2C, which illustrates the overall approach, we displayed results of PCR amplification with Fw1 of four different samples: *in vivo* uninduced mouse (1), *in vivo* induced mouse (2), gDNA from sTL0082 (3), and input library pool (4).

It was necessary to adjust the optimal number of cycles for the single primer PCR in order to avoid saturation levels. This can be caused by accumulation of double stranded DNA produced during the exponential phase, which saturates the polymerase. Against common thinking, saturation does not happen due to an exhaustion of reagents (e.g. primers, dNTPs) (Kainz, 2000). Although primers were designed to bind in the universal flanks, slight differences in size and GC content may cause some amplicons to be favoured over others. Saturation would overenrich the most physio-chemically privileged RNAi targets in detriment of the less favoured, potentially compromising the basis of the RITseq screen, even when this compares induced/uninduced reads per gene ID. 28 cycles was chosen as this produced enough enrichment while far from saturation (Figure 3-4G)

Addition of 6-nucleotide tags to the single primer used to perform the PCR (Fw1= OL4161) permitted expanding our multiplexing capacity and reducing overall costs of the experiment. Pooling equal masses of PCR products generated with distinct barcoded primers, from gDNAs purified from different experimental conditions or replicates, permitted creating a single sequencing sample. Reads could be assigned later to the original experimental condition *in silico* (Figure 3-5).



Name	Sequence	Name	Sequence
RIT-1-F	ATCACGTAATGCCAACTTTGTACAAA	RIT-10-F	TAGCTTTAATGCCAACTTTGTACAAA
RIT-4-F	TGACCATAATGCCAACTTTGTACAAA	RIT-11-F	GGCTACTAATGCCAACTTTGTACAAA
RIT-5-F	ACAGTGTAATGCCAACTTTGTACAAA	RIT-12-F	CTTGTATAATGCCAACTTTGTACAAA
RIT-6-F	GCCAATTAATGCCAACTTTGTACAAA	RIT-13-F	AGTCAATAATGCCAACTTTGTACAAA
RIT-7-F	CAGATCTAATGCCAACTTTGTACAAA	RIT-16-F	CCGTCCTAATGCCAACTTTGTACAAA
RIT-8-F	ACTTGATAATGCCAACTTTGTACAAA	RIT-21-F	CGTACGTAATGCCAACTTTGTACAAA
RIT-9-F	GATCAGTAATGCCAACTTTGTACAAA	RIT-22-F	GAGTGGTAATGCCAACTTTGTACAAA

Figure 3-5. PCR bar-coded strategy for sample enrichment of the RNAi target. **A.** Schematic representation of the strategy. 6-mer labels added to OL4161 permits pooling equal masses of PCR products obtained from different experimental conditions (cond.1, 2, 3) and use of a single set of Illumina sequencing adapters. **B.** 1% agarose gel showing homogeneous amplification of gDNA purified from the input library with the 14 selected barcoded versions of OL4161. Sequences for each of the primers is depicted in the table.

The PCR enrichment strategy required 13 plasmids and their respective cell lines to be remade so that they could include the universal Gateway® sites flanking the RNAi cassette. The original plasmids had been made using traditional restriction digestion-ligation cloning and would not have been suitable for PCR amplification of the RNAi target sequence by OL4161. As with the rest of the library (Jones *et al.*, 2014), alamar blue phenotypes after 72 h of RNAi induction were also analysed (Table 3-2). Among these lines a few showed a loss of fitness: the 2 clones made for CRK3, KKT10 and AGC1, one clone for KKT19, PDK1 and the orphan kinase Tb927.11.6690.

Table 3-2. Alamar blue (AB) ratios (Tet+/Tet-) after 72h for cell lines remade in the Gateway® system to enable PCR enrichment in the pooled library.

Plasmid	Gene ID	PK name	Cell line	Mean AB (Tet+/Tet-)	SD
pTL21	Tb927.10.4990	CRK3	sTL0024	0,511	0,011
			sTL0097	0,516	0,023
pTL23	Tb927.6.2250	AGC (RAC)	sTL0031	0,997	0,022
			sTL0034	0,974	0,203
pTL236	Tb927.10.16160	Orphan	sTL590	1,049	0,040
			sTL591	0,996	0,036
pTL237	Tb927.3.3290	Orphan	sTL593	0,956	0,095
			sTL594*	1,031	0,031
PTL234	TB927.11.12420	CMGC/CLK2 (KKT19)	sTL538	1,023	0,058
			sTL596	0,820	0,122
pTL128	Tb927.10.14300	STE (STE11)	sTL616	0,984	0,045
			sTL617	0,976	0,022
PTL233	TB927.11.12410	CMGC/CLK (KKT10)	sTL535	0,618	0,009
			sTL594	0,721	0,172
PTL244	Tb927.3.2440	AGC1	sTL618	0,606	0,012
			sTL619	0,801	0,019
pTL235	Tb927.7.3580	NEK-11	sTL708	0,945	0,002
			sTL709	1,020	0,042
pTL238	Tb927.3.3080	NEK-6	sTL710	1,042	0,028
			sTL711	0,985	0,014
pTL240	Tb927.8.1670	NEK-13	sTL712	0,979	0,016
			sTL713	0,994	0,036
pTL241	Tb927.9.4910	AGC (TbPDK1)	sTL672	0,896	0,191
			sTL714	1,004	0,044
pTL243	Tb927.11.6690	Orphan	sTL715	0,502	0,014
			sTL716	1,047	0,043

SD= standard deviation. Mean AB highlighted in red= loss of fitness phenotypes. PK = Protein kinase.

3.3.4 RNAi target sequencing analysis

In order to score loss of fitness in the RITseq, ratios were determined from reads mapped to each protein kinase gene in the induced sample relative to the uninduced sample (and in the uninduced sample relative to the pre-inoculation pool, or time 0) with a bootstrap analysis including 1000 repeats (see material and methods for details). This approach enabled calculation of the median and estimation of 95% confidence intervals for each gene (see summary of results, Table 3-3). Median ratios under 0.6 were considered loss of fitness phenotypes based on the scores for the protein kinase with an upper limit for the 95% confidence interval under 0.75 in both *in vivo* replicates. When the 0.6 (induced/uninduced) threshold was applied to all data sets it was determined to

be discriminatory both *in vitro* and *in vivo*, showing a progressive increase of the number of targets identified with time (Figure 3-6, A and B).

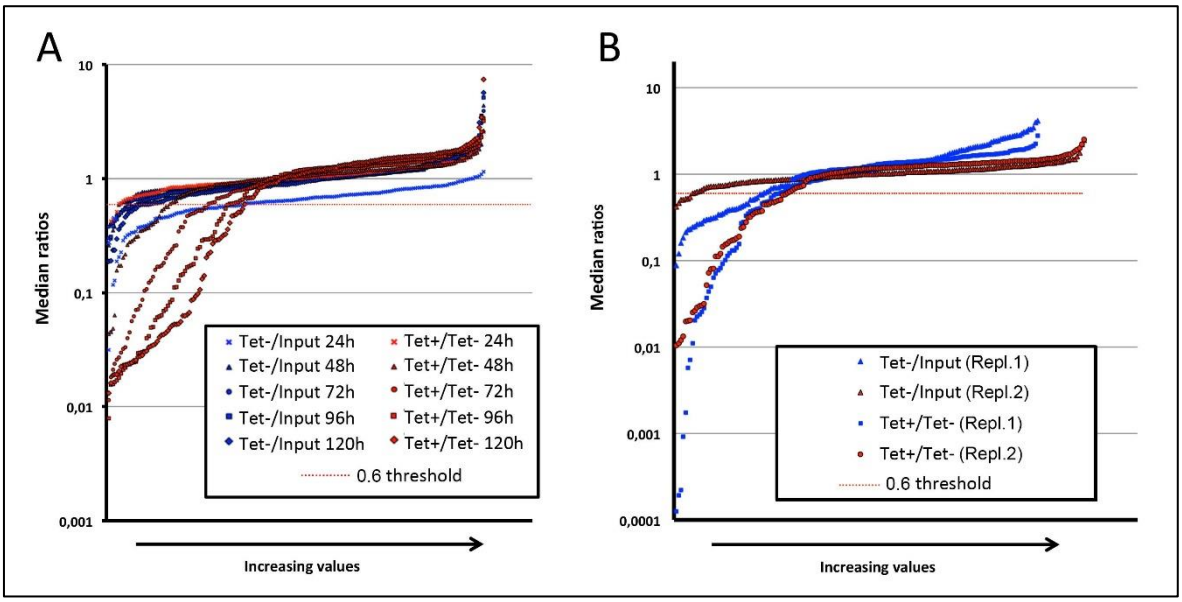


Figure 3-6. Ranking of *in vitro* and *in vivo* RITseq results.
A. Bootstrap median ratios (Tet-/input and Tet+/Tet-) of triplicates for each time point of the *in vitro* RNAi screen, arranged in increasing order. Reads were normalized as 10^6 times raw count mapped to a gene ID divided per sum of total reads included for the analysis in a particular sample, rounded to the next integer. **B.** Bootstrap median ratios (control/input and induced/control) of each *in vivo* experiment arranged in increasing order.

To add an extra layer of statistical analysis, P values were calculated with T-tests performed per gene and time point, Table S4. Volcano plots, representing induced/uninduced ratios *versus* P values, also highlighted an increase in the number of significant loss of fitness phenotypes with time *in vitro* (Figure 3-7) and a better separation between induced and uninduced results *in vivo*.

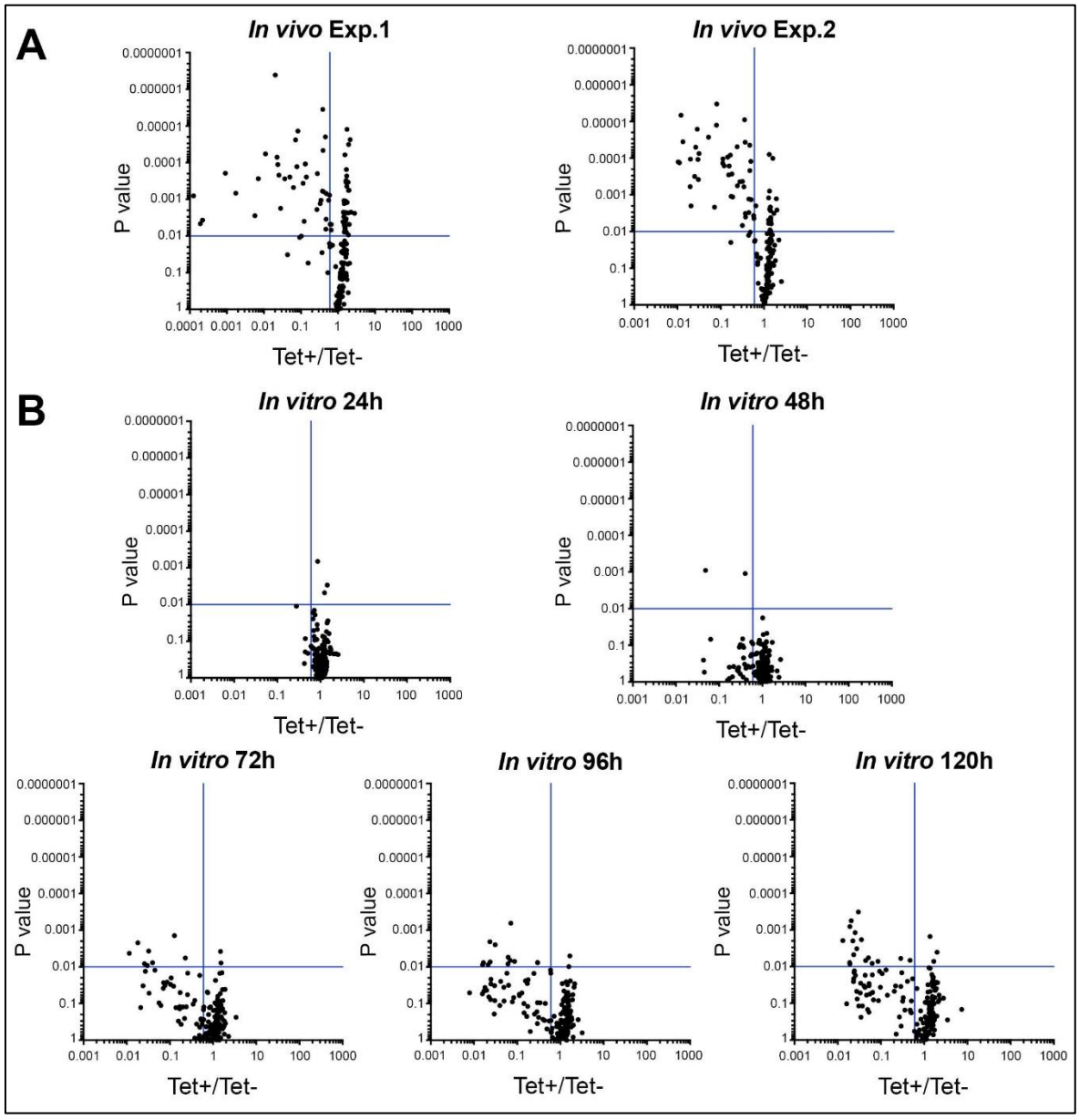


Figure 3-7. Volcano plots for the *in vivo* (A) and *in vitro* (B) kinome-wide RITseq experiments.
Statistical significance (P values, calculated via T-test) is depicted *versus* induced/uninduced ratios. Blue lines determining thresholds of loss of fitness (ratios < 0.6, vertical) and of significance (P < 0.01, horizontal).

RITseq data for all protein kinase genes are summarized in Table 3-3. Every protein kinase line is labelled with gene identifier, protein kinase family, and the MSTL library in which it was included. The table contains: *in vivo* results (median bootstrap ratios) for replicates 1 and 2 and the average between them; *in vitro* results (median bootstrap ratios) for every time point of the *in vitro* RITseq (24 h-120 h); and details of loss of fitness detected in the RNAi screens covering the protein kinome available in the literature (Alsford *et al.*, 2011; Jones *et al.*, 2014). Growth defect or cell cycle phenotypes, if defined by Jones

et al., are noted. Gene identifiers of lines that did not produce a loss of fitness in the alamar blue screen but were affected *in vivo* are highlighted in red. Results obtained in the whole genome RITseq developed by Alsford et al. are also inscribed in the table after 72 h and 144 h, with significant loss of fitness noted as “TRUE”.

Details for individual lines can be found in Table S1 classified by MSTL, including plasmid code, RNAi target sequence, and alamar blue ratios *in vitro* (Jones *et al.*, 2014)

Table 3-3. Summary of RNAi target sequencing results *in vivo* and *in vitro*

Repl. 1 and Repl. 2 = median Tet+/Tet- for each *in vivo* replicate
24h, 48h, 72h, 96h, 120h = median Tet+/Tet- for each time point at *in vitro*
C= cytokinesis; M= mitosis; K= kinetoplast duplication/segregation; U= unclassified (Jones et al. 2014)
*=Phenotypes in Jones et al. (2014) not found in Alsford et al. (2011)
Yes = significant loss of fitness (Alsford et al. 2011) / No = no significant loss of fitness (Alsford et al. 2011)

In vivo or in vitro RITseq Tet+ / Tet- <0.1 <0.3 <0.6 | In vivo not in vitro | Loss of fitness (Jones et al.2014; Alsford et al. 2011)

1 / 6	Family/Name	Sub library	*In vivo* Tet+/Tet- (48 h)			*In vitro* Tet+/Tet-					Jones 72h		Alsford 72h		Alsford 144h	
Gene ID			Repl 1	Repl 2	Mean	24h	48h	72h	96h	120h	Growth	Cell cycle	Tet+/Tet-	Signif.	Tet+/Tet-	Signif.
Tb927.8.5950	STE/STE7/MKK4	MSTL3	2.796	1.994	2.395	0.668	0.401	0.759	2.571	2.807	none	unknown	1.190	No	1.888	No
Tb927.10.1910	STE/STE11	MSTL3	1.594	2.215	1.904	0.940	1.044	0.798	1.288	1.963	none	unknown	1.411	No	1.288	No
Tb927.11.11110	Other/NEK/NEK20	MSTL8	1.683	2.060	1.871	1.107	1.052	1.323	1.241	1.131	none	unknown	0.808	No	0.243	Yes
Tb927.10.12040	CMGC/MAPK/MAPK11	MSTL7	2.124	1.534	1.829	1.417	1.269	1.533	1.732	1.377	none	unknown	0.696	No	0.516	Yes
Tb927.4.3770	CAMK	MSTL5	1.908	1.685	1.796	1.072	1.246	2.294	1.560	1.225	none	unknown	1.144	No	1.883	No
Tb927.7.3650	STE/STE11	MSTL3	1.895	1.669	1.782	0.442	1.324	0.868	0.848	1.080	none	unknown	1.760	No	3.461	No
Tb927.8.3770	CMGC/MAPK	MSTL4	1.776	1.773	1.775	1.072	1.027	1.224	1.422	1.445	none	unknown	1.058	No	1.375	No
Tb927.11.9270	AGC/ZFK	MSTL5	1.987	1.436	1.712	1.733	1.036	0.877	1.552	1.565	none	unknown	1.295	No	1.523	No
Tb927.3.4860	STE/STE7/MKK1	MSTL3	2.243	1.062	1.652	1.394	2.636	1.121	0.547	3.456	none	unknown	0.723	No	0.875	No
Tb927.8.3550	CMGC/MAPK/MAPK3	MSTL4	1.929	1.373	1.651	1.057	0.920	1.236	1.292	0.902	none	unknown	1.647	No	1.910	No
Tb927.3.1570	Orphan	MSTL6	1.653	1.604	1.628	0.862	1.302	1.486	1.654	1.322	none	unknown	1.436	No	1.456	No
Tb927.10.14780	STE/STE11	MSTL3	2.052	1.196	1.624	1.301	2.203	1.690	1.219	0.670	none	unknown	0.084	Yes	0.028	Yes
Tb927.5.3320	Pseudo-Orphan	MSTL6	1.679	1.561	1.620	0.995	0.890	1.287	1.582	2.215	none	unknown	0.711	No	1.761	No
Tb927.8.5500	Orphan	MSTL6	1.932	1.282	1.607	0.837	1.289	1.438	1.032	1.015	none	unknown	2.280	No	2.745	No
Tb927.9.10040	Pseudo-Orphan	MSTL6	1.699	1.507	1.603	1.042	0.987	1.302	0.996	0.779	none	unknown	0.098	Yes	0.098	Yes
Tb927.10.8730	aPK-ABC1	MSTL1	1.875	1.288	1.582	0.709	0.750	1.019	1.071	1.184	none	unknown	0.518	No	1.186	No
Tb927.10.11940	aPK-ABC1	MSTL1	1.313	1.837	1.575	0.842	1.306	1.434	1.306	2.269	none	unknown	0.291	Yes	1.725	No
Tb927.10.1380	Other/NEK/NEK19	MSTL8	1.745	1.401	1.573	1.239	1.334	1.134	1.490	1.554	none	unknown	1.228	No	0.795	No
Tb927.5.1650	CMGC/DYRK	MSTL7	Absent	1.552	1.552	1.359	1.270	0.694	1.548	1.984	none	unknown	1.308	No	2.879	No
Tb927.11.4780	aPK-PDK	MSTL1	1.664	1.411	1.537	1.059	1.194	1.117	1.838	1.609	none	unknown	2.426	No	2.117	No
Tb927.7.6680	STE/STE11	MSTL3	1.710	1.361	1.536	0.422	0.729	1.267	0.860	1.621	none	unknown	0.994	No	1.465	No
Tb927.6.2030	STE/STE11	MSTL3	2.018	1.049	1.534	1.387	1.643	1.045	0.948	1.606	none	unknown	0.396	Yes	1.224	No
Tb927.11.8150	pseudo Other/ULK	MSTL8	1.124	1.932	1.528	2.304	0.976	3.410	1.490	1.771	none	unknown	0.775	No	0.749	Yes
Tb927.10.16030	CMGC/MAPK/MAPK2	MSTL4	1.441	1.612	1.527	1.240	1.696	1.215	0.603	0.915	none	unknown	0.425	No	1.606	No
Tb927.3.2690	AGC/RSK	MSTL5	1.659	1.390	1.524	1.070	1.136	1.337	1.212	1.239	none	unknown	1.254	No	2.106	No

2 / 6	Family/Name	Sub library	*In vivo* Tet+/Tet- (48 h)			*In vitro* Tet+/Tet-					Jones 72h		Alsford 72h		Alsford 144h	
Gene ID			Repl 1	Repl 2	Mean	24h	48h	72h	96h	120h	Growth	Cell cycle	Tet+/Tet-	Signif.	Tet+/Tet-	Signif.
Tb927.5.3420	aPK-ABC1	MSTL1	1.711	1.325	1.518	0.844	1.068	1.245	1.295	1.504	none	unknown	0.417	No	0.523	Yes
Tb927.10.15880	Orphan	MSTL6	1.728	1.305	1.517	1.022	1.269	1.415	1.575	1.667	none	unknown	1.176	No	2.440	No
Tb927.10.5270	STE/STE7/MKK5	MSTL3	1.830	1.201	1.515	0.957	0.919	0.375	0.600	0.225	none	unknown	0.969	No	1.393	No
Tb927.11.8940	CAMK/LDK	MSTL5	1.506	1.514	1.510	1.034	1.421	1.658	1.299	1.476	none	unknown	0.772	No	2.635	No
Tb927.3.1850	Orphan	MSTL6	1.558	1.449	1.503	1.316	0.957	1.289	1.491	1.623	none	unknown	3.515	No	5.157	No
Tb927.7.1900	CMGC/CDK/CRK7	MSTL4	1.645	1.342	1.493	0.876	0.987	1.319	1.495	1.503	none	unknown	0.047	Yes	0.000	Yes
Tb927.10.2390	CK1/CK1	MSTL5	1.496	1.417	1.457	1.189	1.452	1.609	1.542	1.912	none	unknown	0.410	Yes	1.024	No
Tb927.5.4430	CAMK	MSTL5	1.632	1.273	1.453	0.985	1.358	1.526	1.098	1.049	none	unknown	1.089	No	0.278	Yes
Tb927.4.800	aPK-PIKK/FRAP/TOR3	MSTL1	1.476	1.428	1.452	0.959	1.235	1.221	1.247	1.503	none	unknown	0.910	No	1.446	No
Tb927.9.1500	Other/NEK/NEK15	MSTL8	1.471	1.426	1.448	1.010	1.315	1.400	1.316	1.223	none	unknown	1.050	No	1.428	No
Tb927.11.7010	Other/CAMKK	MSTL2	1.449	1.425	1.437	0.729	1.286	1.141	0.998	1.667	none	unknown	0.373	Yes	1.479	No
Tb927.4.2460	Pseudo-Orphan	MSTL6	1.603	1.268	1.435	0.900	1.193	1.699	1.298	1.381	none	unknown	0.920	No	1.116	No
Tb927.10.460	Other/NEK/NRKC	MSTL8	1.175	1.694	1.434	1.251	1.181	1.222	0.975	1.448	none	unknown	0.792	No	1.173	No
Tb927.8.870	CAMK/CAMKL	MSTL5	1.371	1.482	1.426	1.147	0.918	0.922	1.325	0.851	none	unknown	0.518	No	1.814	No
Tb927.8.4260	Orphan	MSTL6	1.600	1.218	1.409	1.150	1.049	1.866	1.488	1.142	none	unknown	1.179	No	1.667	No
Tb927.6.2450	AGC/putative PKB	MSTL5	1.435	1.378	1.407	0.879	1.301	1.106	1.245	1.527	none	unknown	0.086	Yes	0.321	Yes
Tb927.2.2260	aPK-PIKK/ATM	MSTL1	1.568	1.240	1.404	1.075	1.045	1.069	1.099	0.888	none	unknown	1.014	No	1.692	No
Tb927.9.12400	STE/STE11/KinER3	MSTL3	1.332	1.473	1.403	1.020	1.020	0.771	3.201	7.400	none	unknown	0.487	Yes	1.135	No
Tb927.11.14270	STE/STE7	MSTL3	1.512	1.292	1.402	1.216	0.959	1.018	1.900	1.582	none	unknown	0.968	No	2.735	No
Tb927.9.2350	Orphan	MSTL6	1.347	1.455	1.401	1.433	1.184	1.177	1.549	1.809	none	unknown	1.548	No	1.502	No
Tb927.8.1780	STE/STE7	MSTL3	1.631	1.168	1.399	0.934	0.885	1.306	1.770	0.827	none	unknown	0.535	No	0.918	No
Tb927.8.7110	Other/NEK*/NEK12.1	MSTL8	1.352	1.432	1.392	1.244	1.156	1.334	1.528	1.542	none	unknown	1.092	No	0.607	Yes
Tb927.10.350	CMGC/DYRK/PK4	MSTL7	1.462	1.303	1.383	0.799	1.162	1.796	1.191	1.495	none	unknown	0.959	No	1.683	No
Tb927.10.9900	aPK-ABC1	MSTL1	1.483	1.279	1.381	0.958	1.263	0.931	1.420	1.376	none	unknown	1.384	No	0.530	Yes
Tb927.10.14800	CMGC/RCK/MAPK9/RCK2	MSTL7	1.530	1.205	1.368	1.139	0.970	1.018	1.482	1.439	none	unknown	1.084	No	1.958	No
Tb927.7.3190	Pseudo-Orphan	MSTL6	1.361	1.365	1.363	1.141	1.211	1.396	1.719	1.377	none	unknown	0.294	Yes	1.378	No
Tb927.10.9600	CMGC/DYRK	MSTL7	1.354	1.351	1.352	1.181	1.124	1.142	1.568	1.751	none	unknown	0.382	Yes	1.036	No
Tb927.8.6810	STE/STE11	MSTL3	Absent	1.350	1.350	1.199	1.110	1.442	1.469	1.626	none	unknown	0.977	No	1.648	No
Tb927.9.2140	Orphan	MSTL6	1.672	1.029	1.350	1.159	0.760	1.822	1.181	1.194	none	unknown	0.848	No	2.081	No
Tb927.2.4200	CMGC/CLK	MSTL7	1.241	1.456	1.349	0.883	1.190	1.385	1.657	1.854	none	unknown	0.991	No	1.569	No
Tb927.4.2500	Other/PEK/EIF2K2	MSTL2	1.435	1.253	1.344	1.650	1.050	1.156	1.468	1.574	none	unknown	1.035	No	1.583	No
Tb927.1.1530	STE	MSTL3	1.416	1.272	1.344	1.294	1.133	1.734	0.865	0.486	none	unknown	1.022	No	1.368	No
Tb927.2.2120	Other/NEK/NEK22	MSTL8	1.304	1.375	1.339	1.215	1.266	1.398	1.210	1.109	none	unknown	1.817	No	0.873	No
Tb927.3.1610	CMGC/CLK	MSTL7	1.537	1.140	1.338	1.040	1.182	1.659	1.463	1.173	none	unknown	0.630	No	1.631	No
Tb927.3.3190	Other/NEK/NEK7	MSTL8	1.397	1.279	1.338	1.009	1.195	1.574	1.648	2.046	none	unknown	0.793	No	1.592	No
Tb927.11.6690	Orphan	MSTL8	1.491	1.181	1.336	0.848	1.329	1.168	1.802	1.421	none	unknown	0.683	No	2.114	No

3 / 6	Family/Name	Sub library	*In vivo* Tet+/Tet- (48 h)			*In vitro* Tet+/Tet-					Jones 72h		Alsford 72h		Alsford 144h	
Gene ID			Repl 1	Repl 2	Mean	24h	48h	72h	96h	120h	Growth	Cell cycle	Tet+/Tet-	Signif.	Tet+/Tet-	Signif.
Tb927.6.1780	CMGC/MAPK/MAPK4	MSTL4	1.246	1.418	1.332	1.122	1.444	1.465	1.195	1.288	none	unknown	0.848	No	0.659	No
Tb927.3.2060	STE	MSTL3	Absent	1.320	1.320	1.000	1.532	1.314	1.369	1.086	none	unknown	1.141	No	2.391	No
Tb927.10.5310	CAMK/CAMKL/TbAMPKa1	MSTL5	1.266	1.370	1.318	0.976	1.128	1.273	1.395	1.395	none	unknown	0.815	No	1.104	No
Tb927.5.3160	Orphan	MSTL6	1.162	1.467	1.314	1.011	1.438	1.272	1.994	1.296	none	unknown	0.789	No	1.547	No
Tb927.11.14500	Pseudo-Orphan	MSTL6	1.690	0.929	1.309	1.193	1.072	1.240	0.768	0.735	none	unknown	0.222	Yes	0.712	No
Tb927.6.3430	Other/NEK/NEK10	MSTL8	1.400	1.218	1.309	1.294	1.088	1.381	1.027	1.589	none	unknown	1.048	No	1.238	No
Tb927.10.13480/10.13490	CAMK/CAMKL	MSTL5	1.388	1.223	1.306	0.833	1.110	1.302	1.097	1.541	none	unknown	0.200	No	0.067	Yes
Tb927.9.1570	CAMK/	MSTL5	Absent	1.302	1.302	0.914	1.005	1.301	1.268	1.115	none	unknown	0.828	No	0.374	Yes
Tb927.3.1630	CK1/CK1/CK1	MSTL5	Absent	1.302	1.302	1.163	1.323	1.109	1.285	1.575	none	unknown	1.327	No	1.903	No
Tb927.2.5230	pseudo CAMK	MSTL8	Absent	1.287	1.287	0.836	1.273	1.132	1.514	1.764	none	unknown	0.091	Yes	0.162	Yes
Tb927.4.5390/8.6930	Other/NEK/NRKB	MSTL8	1.394	1.175	1.284	0.903	1.198	0.994	1.178	1.041	none	unknown	1.590	No	3.504	No
Tb927.5.3150	Orphan/KinER5	MSTL6	1.360	1.201	1.280	1.366	0.911	0.893	1.392	1.235	none	unknown	1.348	No	1.449	No
Tb927.9.9320	CMGC/RCK-MAPK/MOK	MSTL7	1.405	1.148	1.276	1.087	1.124	1.087	1.402	1.334	none	unknown	0.370	Yes	2.420	No
Tb927.6.2980	Other/PEK/TbEIFK3	MSTL2	1.392	1.161	1.276	1.001	0.921	1.167	1.311	1.392	none	unknown	0.938	No	1.060	No
Tb927.10.3340	Orphan	MSTL6	Absent	1.274	1.274	1.108	1.094	1.161	1.423	1.059	none	unknown	1.123	No	1.704	No
Tb927.11.6990	Other/CAMKK	MSTL2	1.159	1.385	1.272	1.283	0.893	1.190	1.434	1.240	none	unknown	0.161	Yes	0.086	Yes
Tb927.7.5220	STE/STE11/Forkhead Kinase	MSTL3	Absent	1.269	1.269	1.185	1.212	1.320	1.443	1.056	none	unknown	0.494	Yes	0.837	Yes
Tb927.10.15020	CMGC/DYRK/YAK1	MSTL7	1.319	1.215	1.267	1.308	0.938	1.107	1.269	1.691	none	unknown	1.368	No	1.871	No
Tb927.7.2420	CMGC/GSK/GSK3-long	MSTL4	1.366	1.161	1.263	1.033	0.983	1.347	1.252	1.536	none	unknown	1.134	No	1.382	No
Tb927.3.3080	Other/NEK/NEK6	MSTL8	1.262	1.263	1.262	0.930	1.269	1.831	1.357	1.328	none	unknown	0.715	No	1.611	No
Tb927.11.790	Orphan/	MSTL6	0.987	1.510	1.249	1.646	1.052	1.702	1.966	1.714	none	unknown	2.732	No	2.846	No
Tb927.4.840	aPK-PDK/	MSTL1	1.064	1.433	1.248	0.804	1.234	1.906	1.104	1.707	none	unknown	1.053	No	0.693	No
Tb927.3.4670	CMGC/CDK/CRK10	MSTL4	1.214	1.278	1.246	1.013	1.538	1.098	1.221	1.519	none	unknown	0.244	Yes	0.000	Yes
Tb927.11.16790	CMGC/CDKL/ECK1	MSTL7	1.131	1.340	1.236	1.100	0.931	0.622	0.338	0.219	none	unknown	0.828	No	0.345	Yes
Tb927.4.4970	aPK/	MSTL1	Absent	1.231	1.231	1.233	1.283	1.294	1.238	1.156	none	unknown	1.027	No	2.147	No
Tb927.8.4000	aPK-ABC1	MSTL1	1.130	1.325	1.228	0.740	1.001	0.914	1.441	1.191	none	unknown	2.159	No	1.319	No
Tb927.10.10870	CMGC	MSTL7	1.472	0.965	1.218	0.969	1.283	0.992	1.164	0.505	none	unknown	0.771	No	0.924	No
Tb927.1.1000	aPK	MSTL1	1.078	1.320	1.199	1.389	1.168	1.505	1.398	1.545	none	unknown	1.085	No	0.075	Yes
Tb927.11.15010	Other/NEK/NEK21	MSTL8	Absent	1.193	1.193	0.843	1.116	0.988	1.374	1.316	none	unknown	1.589	No	2.145	No
Tb927.4.4590	AGC	MSTL5	1.227	1.149	1.188	1.051	1.005	0.952	1.409	1.714	none	unknown	0.855	No	2.328	No
Tb927.11.3140	CMGC/DYRK/DYRK2	MSTL7	1.317	1.053	1.185	1.384	0.823	0.698	0.427	0.681	none	unknown	0.201	Yes	1.049	No
Tb927.8.5390	CMGC/CDK/CRK4	MSTL4	1.126	1.243	1.185	1.089	1.053	1.282	1.694	1.302	none	unknown	0.343	Yes	0.514	Yes
Tb927.10.1940	Pseudo-CAMK/CAMKL	MSTL8	1.363	0.991	1.177	1.058	1.008	1.133	1.150	0.686	none	unknown	1.490	No	1.265	No
Tb927.10.3230	CMGC/MAPK	MSTL4	1.154	1.178	1.166	1.002	1.246	1.555	2.134	1.372	none	unknown	1.491	No	3.566	No
Tb927.11.7210	Other/PEK/TbEIFK1	MSTL2	1.134	1.185	1.160	0.673	0.978	1.170	1.931	1.361	none	unknown	1.816	No	2.639	No
Tb927.8.6490	Other/CAMKK	MSTL2	1.078	1.234	1.156	1.162	1.154	1.214	1.195	1.482	none	unknown	0.922	No	1.167	No

4 / 6 Gene ID	Family/Name	Sub library	*In vivo* Tet+/Tet- (48 h)			*In vitro* Tet+/Tet-					Jones 72h		Alsford 72h		Alsford 144h	
			Repl 1	Repl 2	Mean	24h	48h	72h	96h	120h	Growth	Cell cycle	Tet+/Tet-	Signif.	Tet+/Tet-	Signif.
Tb927.1.3130	Orphan	MSTL6	1.029	1.256	1.142	1.014	1.285	1.305	1.399	1.422	none	unknown	0.951	No	3.457	No
Tb927.10.16160	Orphan	MSTL8	1.140	1.142	1.141	1.090	1.134	1.233	1.431	1.545	none	unknown	0.750	No	1.766	No
Tb927.2.2720	Orphan/FlaK	MSTL6	0.961	1.315	1.138	0.858	0.949	1.147	1.031	0.595	none	unknown	1.296	No	0.854	Yes
Tb927.5.2820	Other/NEK/NEK9	MSTL8	1.178	1.081	1.130	1.196	0.816	1.115	1.287	1.225	none	unknown	1.093	No	1.147	No
Tb927.8.1670	Other/NEK*/NEK14	MSTL8	1.124	1.130	1.127	0.885	1.110	1.116	1.295	1.661	none	unknown	1.134	No	3.114	No
Tb927.10.5950/10.5940	Other/NEK	MSTL8	1.096	1.141	1.118	0.878	1.246	1.178	1.258	1.182	none	unknown	0.532	No	0.117	No
Tb927.10.14420	Other/NEK/NEK16	MSTL8	1.038	1.193	1.116	0.894	1.029	0.813	0.897	0.842	none	unknown	0.392	Yes	0.576	Yes
Tb927.8.1100	STE/STE11	MSTL3	1.072	1.127	1.099	1.996	0.927	0.523	0.396	1.011	none	unknown	1.197	No	1.766	No
Tb927.6.2250	AGC/RAC	MSTL8	1.174	1.023	1.099	0.958	1.019	1.375	1.259	1.590	none	unknown	1.614	No	3.154	No
Tb927.4.1700	CK1/TTBK	MSTL5	1.013	1.153	1.083	1.288	1.129	0.792	1.703	0.893	none	unknown	1.782	No	1.642	No
Tb927.11.8170	CMGC/CDKL	MSTL7	0.980	1.184	1.082	1.259	1.661	1.102	1.273	1.724	none	unknown	0.550	No	0.624	Yes
Tb927.3.5650	STE/kinER4	MSTL3	1.046	1.100	1.073	0.914	1.032	1.505	1.449	1.582	none	unknown	0.898	No	0.923	No
Tb927.3.4560	CAMK/CAMKL/TbAMPKa2	MSTL5	0.928	1.165	1.046	1.095	1.221	1.323	0.656	0.267	none	unknown	1.235	No	1.533	No
Tb927.7.3580	Other/NEK/NEK11	MSTL8	Absent	1.039	1.039	1.239	0.803	0.601	0.642	0.318	none	unknown	0.699	No	0.765	Yes
Tb927.9.3120	STE/STE11/KinER2	MSTL3	1.157	0.898	1.028	0.795	2.431	0.788	1.206	1.169	none	unknown	1.671	No	2.614	No
Tb927.8.1690	Other/NEK/NEK14	MSTL8	Absent	1.011	1.011	0.903	0.813	0.781	0.620	0.680	none	unknown	1.134	No	3.114	No
Tb927.10.3900	CAMK/CAMKL	MSTL5	Absent	1.005	1.005	0.725	0.789	0.806	0.292	0.106	none	unknown	1.659	No	0.760	No
Tb927.9.4910	AGC/TbPDK1	MSTL7	0.880	1.076	0.978	1.287	1.553	1.935	2.066	2.111	Slow*	No	1.583	No	2.645	No
Tb927.3.3290	Orphan/	MSTL7	Absent	0.962	0.962	1.057	1.119	1.124	0.712	0.475	none	unknown	2.451	No	2.435	No
Tb927.7.3210	Pseudo-Orphan	MSTL6	0.870	0.930	0.900	0.911	0.827	0.481	0.313	0.304	Slow	U	0.801	No	0.398	Yes
Tb927.6.4220	CMGC/MAPK/MAPK5	MSTL4	Absent	0.870	0.870	0.890	0.774	1.228	1.238	1.055	none	unknown	0.758	No	0.333	Yes
Tb927.6.3110	CMGC/CDK/CRK11	MSTL4	0.913	0.680	0.797	0.746	0.853	0.608	0.445	0.409	none	unknown	1.153	No	2.590	No
Tb927.7.4090	CMGC/DYRK	MSTL7	0.710	0.833	0.772	1.677	1.102	1.443	0.831	1.001	none	unknown	1.030	No	0.652	Yes
Tb927.11.14070	STE/STE11/DFK/KinER1/RDK1	MSTL3	1.078	0.443	0.761	0.512	1.187	1.364	0.200	0.293	Slow	No	1.555	No	1.788	No
Tb927.4.3420	Other/WEE/WEE1	MSTL2	0.651	0.699	0.675	1.003	0.846	0.667	0.486	0.064	Slow*	No	1.882	No	1.629	No
Tb927.11.12420	CMGC/CLK/CLK2/KKT19	MSTL7	0.589	0.738	0.664	0.889	1.147	1.213	0.613	0.326	No test	unknown	0.000	Yes	0.200	No
Tb927.10.13010	AGC/PKAC3	MSTL5	0.606	0.717	0.662	1.052	1.013	0.931	0.896	0.908	none	unknown	1.680	No	0.822	No
Tb927.9.12880	STE/STE20/SLK2	MSTL3	0.371	0.911	0.641	0.856	0.262	0.086	0.050	0.023	Arrest	No	0.970	No	0.811	Yes
Tb927.11.1180	CMGC/CDK/CRK6	MSTL4	0.657	0.608	0.633	0.911	1.062	0.558	0.244	0.289	Slow*	No	0.805	No	1.143	No
Tb927.9.11100/9.11030	AGC/PKAC1	MSTL5	0.562	0.649	0.606	0.899	0.427	0.169	0.042	0.036	Slow	C	0.303	No	0.045	Yes
Tb927.2.1820	CAMK/CAMKL/CAMKL	MSTL5	0.591	0.586	0.588	1.030	0.538	0.525	0.166	0.131	none	unknown	0.779	No	0.145	Yes
Tb927.6.5100	Other/PLK/PLK-like	MSTL2	0.458	0.713	0.585	0.862	1.344	0.711	0.606	0.539	none	unknown	0.739	No	1.108	No
Tb927.3.3920	Other/AUR/AUK2	MSTL2	0.485	0.622	0.554	0.934	1.191	0.825	0.823	0.358	none	unknown	0.105	Yes	0.000	Yes
Tb927.10.15300	Other/CAMKK/CAMKK	MSTL2	0.616	0.446	0.531	0.647	0.369	0.183	0.077	0.141	Slow	No	0.668	No	0.698	Yes
Tb927.11.9290	Pseudo-Orphan/ORPHAN/FAZ20	MSTL6	0.478	0.577	0.527	1.012	1.097	0.687	1.081	0.504	none	unknown	0.738	No	0.649	Yes

5/ 6	Family/Name	Sub library	*In vivo* Tet+/Tet- (48 h)			*In vitro* Tet+/Tet-					Jones 72h		Alsford 72h		Alsford 144h	
Gene ID			Repl 1	Repl 2	Mean	24h	48h	72h	96h	120h	Growth	Cell cycle	Tet+/Tet-	Signif.	Tet+/Tet-	Signif.
Tb927.3.5400	aPK/RIO1	MSTL1	0.405	0.507	0.456	1.371	0.522	0.229	0.087	0.044	Arrest*	No	1.299	No	2.017	No
Tb927.7.2750	CAMK/CAMK3	MSTL5	0.532	0.375	0.453	1.484	0.621	0.201	0.179	0.088	Slow	No	0.489	No	0.256	Yes
Tb927.1.1930	aPK-PIKK/FRAP/TOR4	MSTL1	Absent	0.450	0.450	0.956	0.612	0.491	0.294	0.204	Slow	G1/S	0.274	Yes	0.301	Yes
Tb927.11.850	aPK/Bud32	MSTL1	0.395	0.498	0.447	1.138	0.690	0.259	0.069	0.025	none	unknown	2.099	No	1.386	No
Tb927.10.8420	aPK-PIKK/FRAP/TOR1	MSTL1	0.380	0.472	0.426	0.954	0.359	0.104	0.062	0.041	Slow	U	0.677	No	0.480	No
Tb927.10.1070	CMGC/CDK/CRK1	MSTL4	0.348	0.469	0.409	0.921	0.469	0.240	0.061	0.093	Slow	G1/S	0.113	Yes	0.007	Yes
Tb927.9.6560	pseudo Other/NAK/OTHER/NAK	MSTL8	0.331	0.483	0.407	1.186	0.202	0.056	0.040	0.056	Arrest	No	0.024	Yes	0.000	Yes
Tb927.11.9190	Other/VPS15/VPS15	MSTL2	0.472	0.319	0.396	1.488	1.199	0.561	0.143	0.097	none	unknown	1.583	No	0.831	No
Tb927.7.960	CMGC/SRPK/SRPK1	MSTL7	0.387	0.357	0.372	0.837	0.997	0.671	0.588	0.472	none	unknown	1.021	No	1.539	No
Tb927.6.2840	aPK/RIO2	MSTL1	0.278	0.371	0.325	0.767	0.294	0.128	0.031	0.021	Arrest	No	0.340	Yes	0.146	Yes
Tb927.10.10350	STE/STE11/STE11/Bck1p	MSTL3	0.270	0.367	0.318	0.863	1.075	1.167	0.871	0.954	none	unknown	1.915	No	1.907	No
Tb927.11.10520	Orphan/KKT2	MSTL6	Absent	0.243	0.243	0.991	0.310	0.107	0.155	0.043	none	unknown	0.122	Yes	0.231	Yes
Tb927.7.5770	AGC/NDR/PK53	MSTL5	0.131	0.344	0.237	0.884	0.893	0.417	0.294	0.085	Slow	K/C	0.536	No	2.000	No
Tb927.10.2040	STE/STE11/CDC15-like	MSTL3	0.092	0.316	0.204	0.445	0.366	0.218	0.171	0.037	Death*	No	0.805	No	1.210	No
Tb927.7.3880	CMGC/DYRK/DYRK1	MSTL7	0.140	0.239	0.190	0.759	0.776	0.345	0.128	0.062	none	unknown	0.353	No	0.052	Yes
Tb927.10.14770	CAMK/CAMK1 (AKB1)	MSTL5	Absent	0.174	0.174	0.821	1.203	0.612	0.175	0.067	none	unknown	0.524	No	1.320	No
Tb927.2.2430	Other/CK2/CK2A2	MSTL2	0.063	0.277	0.170	0.602	1.074	0.525	0.253	0.046	none	unknown	1.280	No	1.873	No
Tb927.11.14680	aPK-PIKK/ATR	MSTL1	0.156	0.181	0.169	1.101	0.708	0.173	0.058	0.063	Arrest*	M	0.742	No	0.983	No
Tb927.7.6220	CAMK/CAMK2	MSTL5	0.114	0.188	0.151	1.003	1.445	0.847	0.112	0.031	Arrest*	C	1.314	No	1.243	No
Tb927.10.7780	CMGC/MAPK/KFR1	MSTL4	0.121	0.168	0.145	1.009	0.881	0.493	0.083	0.053	Slow*	No	1.000	No	0.821	No
Tb927.2.4510	CMGC/CDK/CRK9	MSTL4	0.131	0.156	0.144	0.709	0.216	0.040	0.024	0.028	Death	K	0.205	Yes	0.285	Yes
Tb927.11.2040	STE/STE11/STE11a	MSTL3	0.102	0.171	0.136	0.902	0.043	0.021	0.020	0.053	Arrest*	No	0.826	No	1.404	No
Tb927.11.4470	Other/ULK/ULK	MSTL2	0.078	0.153	0.115	1.079	1.046	0.504	0.286	0.043	Arrest	K/C	0.179	Yes	0.415	Yes
Tb927.10.14300	STE/STE11/STE11/MRK1	MSTL7	Absent	0.112	0.112	0.986	1.148	0.734	0.299	0.288	none	unknown	0.639	No	0.346	Yes
Tb927.6.4970	CMGC/SRPK/SRPK2	MSTL7	0.050	0.145	0.098	1.129	0.350	0.169	0.111	0.053	Arrest*	No	0.702	No	0.680	Yes
Tb927.7.7360	CMGC/CDK/CRK2	MSTL4	0.082	0.111	0.097	0.696	0.349	0.122	0.071	0.013	Slow*	No	1.673	No	2.052	No
Tb927.4.5310	Other/NEK/NEK12.2 /RDK2	MSTL8	0.071	0.120	0.096	0.956	0.309	0.071	0.030	0.019	Death	K/C	0.028	Yes	0.094	Yes
Tb927.10.13780	CMGC/GSK/GSK3-short	MSTL4	0.037	0.080	0.058	0.611	0.418	0.077	0.025	0.023	Arrest	M/C	0.054	Yes	0.000	Yes
Tb927.8.5730	STE/STE20/SLK1	MSTL3	0.044	0.072	0.058	0.689	0.172	0.033	0.018	0.049	Arrest*	No	0.946	No	0.559	No
Tb927.3.690	CMGC/RCK/RCK1/TbMAK	MSTL7	0.028	0.081	0.055	0.874	0.636	0.138	0.023	0.023	Death	C	0.307	Yes	0.380	Yes
Tb927.4.420	aPK-PIKK/FRAP/TOR2	MSTL1	Absent	0.052	0.052	1.228	1.017	0.376	0.063	0.030	Death	No	1.164	No	0.730	Yes
Tb927.10.4940	AGC/NDR/PK50	MSTL5	0.023	0.032	0.027	0.746	0.064	0.024	0.016	0.024	Arrest	C	1.714	No	0.752	No
Tb927.9.14430	Other/CK2/CK2A1	MSTL2	0.026	0.026	0.026	1.571	0.602	0.256	0.029	0.034	Arrest	K	0.650	No	0.280	Yes
Tb927.9.1670	Other/AUR/AUK3	MSTL2	0.020	0.030	0.025	1.481	0.723	0.158	0.049	0.033	Slow	M/C	0.814	No	0.283	Yes

6 / 6	Family/Name	Sub library	*In vivo* Tet+/Tet- (48 h)			*In vitro* Tet+/Tet-					Jones 72h		Alsford 72h		Alsford 144h	
Gene ID			Repl 1	Repl 2	Mean	24h	48h	72h	96h	120h	Growth	Cell cycle	Tet+/Tet-	Signif.	Tet+/Tet-	Signif.
Tb927.9.10920	Orphan/PK6/KKT3	MSTL6	0.024	0.020	0.022	1.268	0.567	0.045	0.021	0.024	Arrest	No	0.301	Yes	0.209	Yes
Tb927.10.4990	CMGC/CDK/CRK3	MSTL7	0.011	0.030	0.021	0.276	0.048	0.027	0.016	0.026	Death	M	0.000	Yes	0.000	Yes
Tb927.7.6310	Other/PLK/PLK1	MSTL2	Absent	0.020	0.020	0.636	0.045	0.018	0.030	0.019	none	unknown	0.000	Yes	0.030	Yes
Tb927.3.2440	AGC/AGC1	MSTL8	0.007	0.029	0.018	0.966	0.851	0.472	0.145	0.051	Death	C	0.450	No	0.203	Yes
Tb927.11.5340	CMGC/CMGC1	MSTL7	0.002	0.025	0.013	0.719	0.389	0.071	0.017	0.019	Slow*	M	1.128	No	2.507	No
Tb927.4.5180/8.7220	Other/TLK*/TLK1&2	MSTL2	0.001	0.020	0.010	1.103	0.283	0.032	0.016	0.028	Death	M/C	0.214	Yes	0.615	No
Tb927.10.5140	CMGC/MAPK/MAPK6 (ERK8)	MSTL4	0.006	0.013	0.010	1.165	0.157	0.011	0.025	0.027	Death	C	0.036	Yes	0.023	Yes
Tb927.5.790/5.800	CK1/CK1*/CK1.1/CK1.2	MSTL5	0.000	0.012	0.006	0.662	0.278	0.026	0.008	0.022	Death	K/C	0.102	Yes	0.022	Yes
Tb927.11.12410	CMGC/CLK/CLK1/KKT10	MSTL7	0.000	0.011	0.006	1.239	0.567	0.062	0.024	0.034	Death	M/C	0.000	Yes	0.000	Yes
Tb927.11.8220	Other/AUR/AUK1	MSTL2	0.000	0.010	0.005	0.732	0.172	0.030	0.022	0.016	Death	M	0.222	Yes	0.029	Yes

Arranged in increasing *in vivo* loss of fitness (decreasing mean Tet+/Tet- between replicates 1 and 2).

3.3.5 *In vitro* results and correlation with the literature

The *in vitro* RITseq screen comprised three replicates at 5 different time points, corresponding to 5 consecutive days upon RNAi induction (Figure 3-2). Examining library behaviour within such a timeframe permits comparisons with other screens available in the literature (Jones et al. 2014, at 72h; Alsford et al. 2011, at 73 and 144h), thereby assessing reliability of our approach and providing a more detailed scrutiny of the loss of fitness progression with time. As this RITseq was conducted in parallel to the *in vivo* screen (that will be described below), following the same sample processing and data analysis treatment, it also provides a more meaningful dataset to analyse divergence between *in vitro* and *in vivo* behaviour than the previously published studies (which either did not rely on a single, defined RNAi target for each protein kinase gene, or analysed growth by alamar blue and not relative abundance of the protein kinase RNAi cells in the population). The results of read mapping (summary and raw data in Table 3-3 and Table S3 respectively) show an increase with time of both the number of protein kinases that displayed loss of fitness (as defined by induced/uninduced ratios <0.6, Figure 3-6A), and the extent of loss of individual protein kinases from the population (Table 3-3).

We showed in Figure 3-7 that, compared to the *in vivo* data set, *in vitro* ratios with a P value supporting a significant result are less abundant, see Table S4. However, when normalised reads were plotted (uninduced *versus* induced) for the RNAi lines showing a loss of fitness at each of the time points analysed *in vitro*, intrinsic reproducibility was still acceptable, Figure 3-8A-E. Little read-count variability could be observed per gene and condition (Tet+ or Tet-) across different mice. The percentage of the kinome whose depletion caused a growth defect increased with time: at 24 h just 5 kinases showed a phenotype, 3.4% of the kinome; after 48 h this number increased to 16.4%, then 26.5% at 72h, 32.8% at 96h and 36.7% after 120h of RNAi induction. Amongst these protein kinases that show loss of fitness, those that may be interesting are highlighted later, in comparison with the *in vivo* results, but results are summarized in Table 3-3.

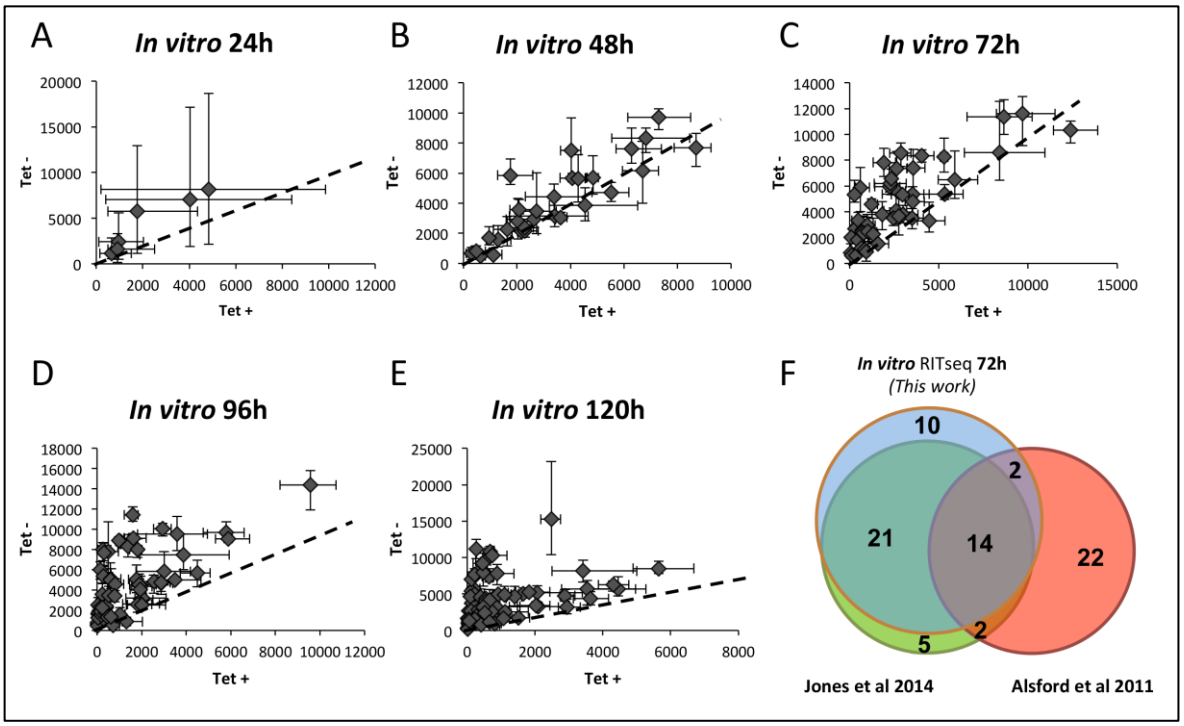


Figure 3-8. Reproducibility of the *in vitro* RITseq screen.
A-E. Breakdown of results for loss of fitness RNAi lines defined by Tet+/Tet- <0.6 within each of the individual *in vitro* biological replicates. Depicted average and range of normalized counts for the 3 cultures analyzed per condition (induced or uninduced) at every time point and per RNAi target. **F.** Venn diagram showing the overlap between the number of loss-of-fitness protein kinase depletions identified with the kinome-wide *in vitro* RITseq screen after 72 h of RNAi induction and the two *in vitro* studies available in the literature (Alsford *et al.*, 2011; Jones *et al.*, 2014) at the same time point.

Comparing the *in vitro* screen with available data in the literature, 38 out of the 47 lines found at the 72 h time point that showed loss of fitness had also been identified in the 72 h *in vitro* Alamar blue screen performed by Jones *et al.* (2014) with the same library (Figure 3-8F). Overlap with the genome-wide RITseq by Alsford et al. (2011) was less pronounced, with only 18 of the protein kinases identified in this study being detected by Alsford, who additionally found another 23. Overlap between the kinome-wide and whole-genome screens is more evident for protein kinase depletions causing the most deleterious growth phenotypes. For instance, protein kinases located at the bottom of Table 3-3, such as AUK1, KKT10, CK1.2 or ERK8, were also identified by Alsford et al. (2011) after three and six days of induction. Each of these has been validated as essential kinases in the literature (Merritt *et al.*, 2014) and, in the kinome-wide RITseq produced the strongest *in vivo* loss of fitness (Tet+/Tet- < 0.001). For

each, *in vitro* loss was observable from 48 h onwards and they were defined by Jones et al. (2014) as “death” phenotypes with the same library.

In order to test further whether clones growing as a pool in culture acted the same as individually in culture, 14 of the lines used in the pool were grown independently to make triplicate growth curves. Ratios of induced and uninduced parasite counts were calculated and compared to RITseq induced/uninduced ratios per protein kinase gene in every independent time point (24 h, 48 h, 72 h and 96 h). Paired (2-tailed) t tests were performed comparing ratios of the growth curve and the RITseq triplicates. Results are plotted in Figure 3-9, showing good correspondence between the two datasets, with only a few exceptions from 72 h (KKT2, DYRK and Bud32) where differenceswere found to be significant.

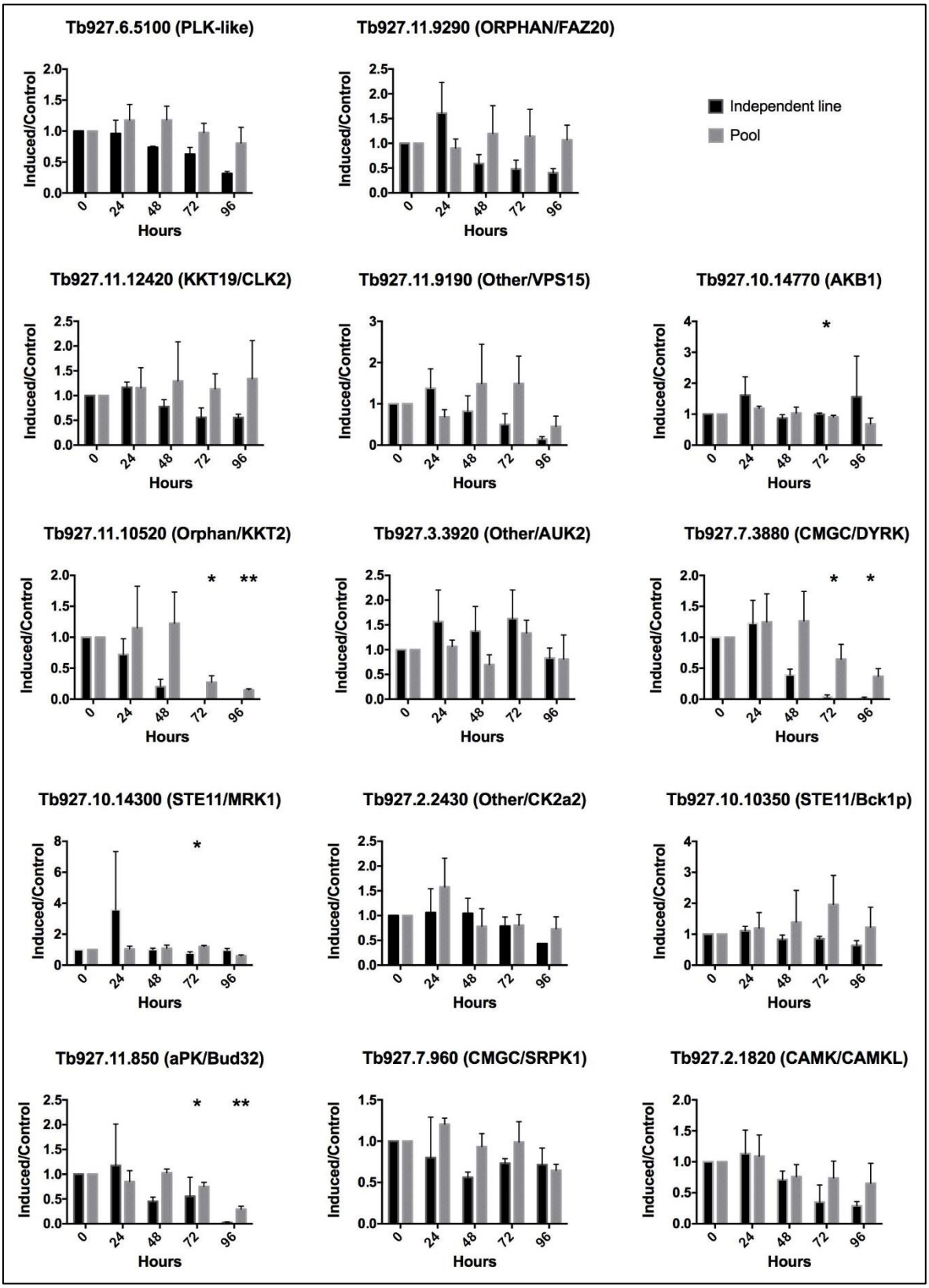


Figure 3-9. Individual cell line *versus* library pool, comparative of loss of fitness.

Depiction of Tet+/Tet- ratios obtained per time point for 14 different protein kinase lines, using growth curves counting cells in a neubauer hemocytometer (independent lines) and after RITseq analysis (pool). * $p < 0.05$ and ** $p < 0.01$ using T test comparing individual and pooled results.

3.3.6 *In vivo* results and correlation with the literature

The *in vivo* experiment (comprising 12 mice, 6 uninduced and 6 induced) was repeated to provide two biological replicates. 72 h post infection (48 h post RNAi induction), parasites were recovered with DEAE cellulose (Lanham and Godfrey, 1970) and gDNA was purified from them.

A few differences are significant between the experimental procedures in the two replicates, which were sequenced in different platforms. PCR enrichment carried out in Replicate 1 was conducted with non-barcoded OL4161 primer. PCR products from each sample were processed separately, being sheared, size-selected (~106 nucleotides) and labelled with Ion Torrent adapters according to manufacturer's instructions. After emulsion PCR and DNA sequencing, *in silico* processing permitted filtration of reads selecting those containing a 9-nucleotide tag (GCCAACTTT). This was indicative of PCR amplification with OL4161, which certified its origin in the RNAi cassette. 20% of the reads were finally usable. Induced/control ratios acquired values 1 fold lower than these of Replicate 2 (below). In addition, infection of one of the mice was unsuccessful, and so we only used 5 uninduced controls.

In the second replicate we optimized the approach. 14 different 6-nucleotide-long barcodes were attached to the OL4161 primer, Figure 3-5. A distinct barcoded primer was used to enrich the RNAi target in *T. brucei* gDNA coming from different mouse infections, Table S5. As discussed before, this approach enabled pooling 14 PCR products, and tagging them with a single pair of Illumina adapters. Reads could be assigned to the original samples later, during the data analysis process. We reasoned, given the RNAi amplicon size, it was not necessary to fragment the sample. This enabled, after *in silico* filtration, an enhanced recovery of 65-70% useful reads per experimental condition. All the 12 mice were successfully infected this time. An extra library, MSTL9, was added to the pool used in Replicate 2, introducing cell lines that were missed or contained the wrong construct after analysis of Replicate 1.

A total of 157 RNAi cell lines were detected in the first *in vivo* experiment, including a number of cell lines that were generated using the traditional restriction/ligation system and were remade to include the Gateway® sites

required for the PCR (Table 3-2). However, 24 further cell lines were found to be missing, Figure 3-10A. A schematic representation of the troubleshooting strategy undertaken with all these lines is depicted in Figure 3-10B. When an RNAi target sequence was missing, the individual line was defrosted, gDNA purified and OL4161 PCR conducted to see whether they contained an insert. If the insert was present, we sequenced it. If the insert was absent, we defrosted the second clone available (transfected with the same construct) and repeated the strategy. If the insert in ether clone was not as expected, we sequenced the original plasmid. If the plasmid was the right one, we transfected again. If the plasmid was mislabelled, we generated the plasmid from scratch. Three of the missing lines were in the batch remade from the older ligation system (Table 3-2). Once the missing protein kinase RNAi cells were developed (Table 3-4) and added to the second experiment as an additional sub-library, MSTL9, we were able to detect 177 protein kinases. Five protein kinases missing in Replicate 1 contained the right insert, and so they were added again to the pool for the second experiment: Tb927.11.12310 (CRK12), Tb927.3.5180 (aPK/A6), Tb927.11.7160 (CRK8), Tb927.11.5860 (Orphan) and Tb927.9.2240 (Pseudo-Orphan). After RITseq, they were still in low abundance, making assignment of reads inaccurate and were therefore excluded from the analysis; see Table 3-1 and raw data (Tables S2-S3). Analysing the *in vivo* data based on the 0.6 induced/uninduced threshold, 53 protein kinase genes were identified that showed a relative depletion of reads in the induced samples, indicating a loss of fitness 48 h after RNAi induction in one or both of the experiments. Forty genes showed such loss of fitness in both of the biological replicates (Table 3-3), with 6 seen only in Replicate 1 and eight only in Replicate 2 (6 of which were absent in the first experiment).

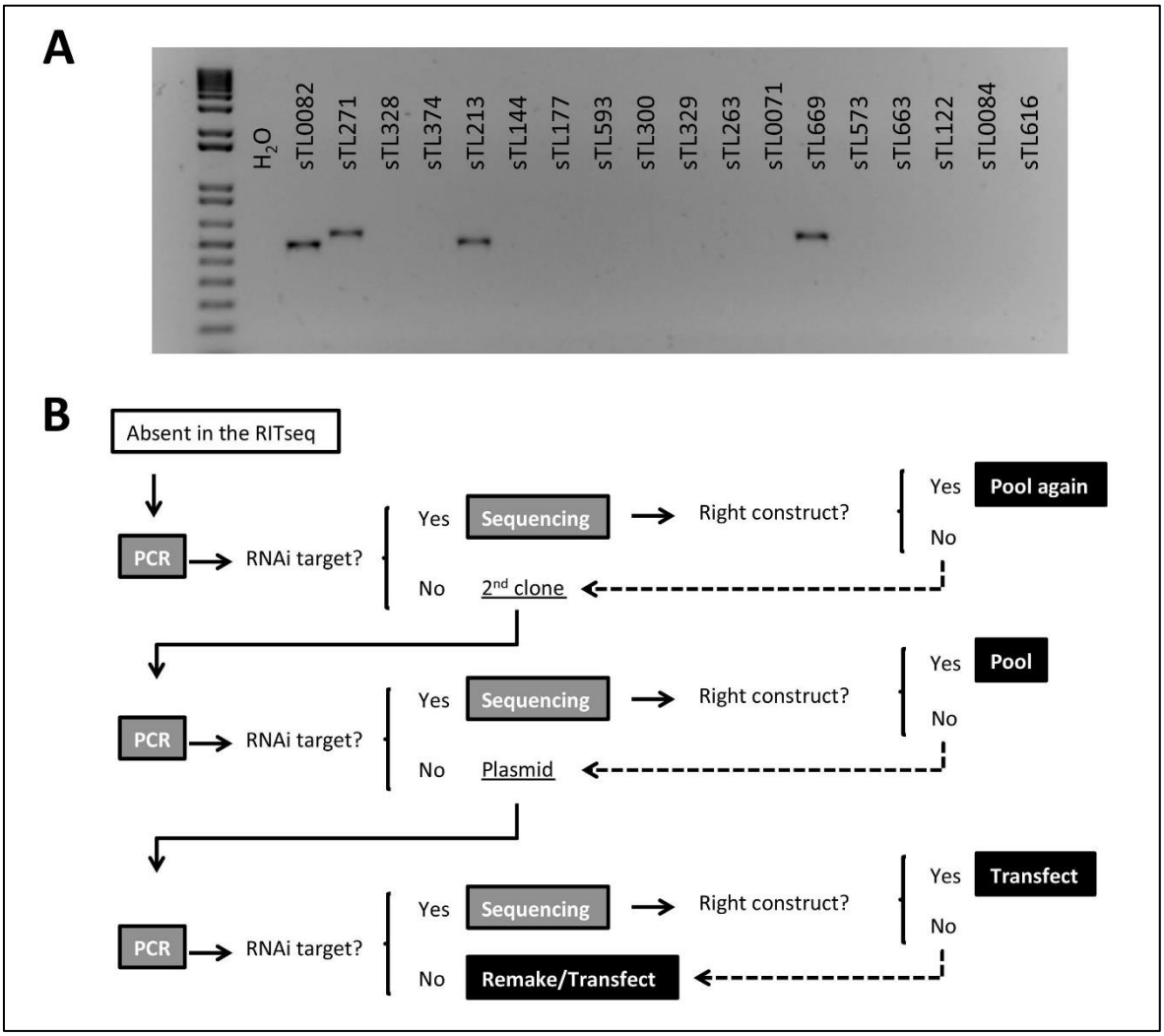


Figure 3-10. Defective RNAi targets in the pTL library.
A. 1% Agrose gel showing RNAi enriching PCR (OL4161), using the water and positive controls (sTL0082, ATR) for a number of protein kinases undetectable in the *in vivo* RITseq Replicate 1. **B**. Schematic representation of the trouble shooting strategy used to solve this problem before the second library pool.

Table 3-4. List of RNAi targets that were not detected in the first *in vivo* experiment, and troubleshooting strategy.

RNAi target	Cell line	PTL	Family/Name	Sublibrary	Insert	Action	MSTL9
Tb927.8.1690	sTL0006	4	Other/NEK14	MSTL8	No insert	Remade/retransfected	STL769
Tb927.11.15010	sTL699	10	Other/NEK21	MSTL8	Wrong insert	Remade/retransfected	STL762
Tb927.10.3340	sTL0084	13	Orphan	MSTL6	No insert	2nd clone used	sTL0089
Tb927.10.1460	sTL0071	38	Orphan	MSTL6	No insert	Missing	Missing
Tb927.11.10520	sTL0069	39	Orphan	MSTL6	Wrong insert	Remade/retransfected	STL783
Tb927.5.1650	sTL122	88	CMGC/DYRK	MSTL7	No insert	Remade/retransfected	STL772
Tb927.6.4220	sTL144	99	CMGC/MAPK5	MSTL4	No insert	2nd clone used	sTL145
Tb927.7.5220	sTL177	118	STE11/Forkhead	MSTL3	No insert	Remade/retransfected	STL774
Tb927.8.6810	sTL195	124	STE/STE11	MSTL3	No insert	Remade/retransfected	STL776
Tb927.10.14300	sTL616	128	STE/STE11	MSTL7	No insert	Remade/retransfected	STL785
Tb927.4.4970	sTL374	140	aPK	MSTL1	No insert	2nd clone used	sTL375
Tb927.8.7450	sTL263	141	aPK	MSTL1	No insert	Missing	Missing
Tb927.4.420	sTL300	148	aPK-PIKK/TOR2	MSTL1	No insert	2nd clone used	sTL301
Tb927.7.6310	sTL213	165	Other/PLK	MSTL2	Wrong insert	Remade/retransfected	STL777
Tb927.3.1630	sTL271	173	CK1/CK1	MSTL5	Wrong insert	Remade/retransfected	STL767
Tb927.2.5230	sTL225	176	pseudo CAMK	MSTL8	Wrong insert	Remade/retransfected	STL780
Tb927.10.14770	sTL329	179	CAMK/AKB1	MSTL5	No insert	Plasmid retransfected	STL793
Tb927.9.1570	sTL227	181	CAMK	MSTL5	No insert	2nd clone used	sTL228
Tb927.10.3900	sTL328	187	CAMK/CAMKL	MSTL5	No insert	Plasmid retransfected	STL791
Tb927.3.2060	sTL351	203	STE	MSTL3	Wrong insert	Remade/retransfected	STL787
Tb927.7.3580	sTL708	235	Other/NEK11	MSTL8	No insert	2nd clone used	sTL709
Tb927.3.3290	sTL593	237	Orphan	MSTL7	No insert	2nd clone used	sTL594*

pTL= plasmid code

3.3.7 Reproducibility of the *in vivo* data

Linear regression across the two *in vivo* experiments showed good reproducibility for the complete dataset (R^2=0.77; p<0.0001, Figure 3-11A). Indeed, reproducibility was substantially better if only the 49 RNAi targets with an average induced/uninduced ratio <0.6 between the 2 biological replicates were compared (R^2=0.82; p<0.0001, Figure 3-11A and B). Regression values for cell lines with fitness ratios over the 0.6 threshold decreased to R^2=0.22. The intrinsic reproducibility across mice for both control and induced groups per gene identifier was also excellent within the two biological replicates (see means and ranges displayed in Figure 3-11C).

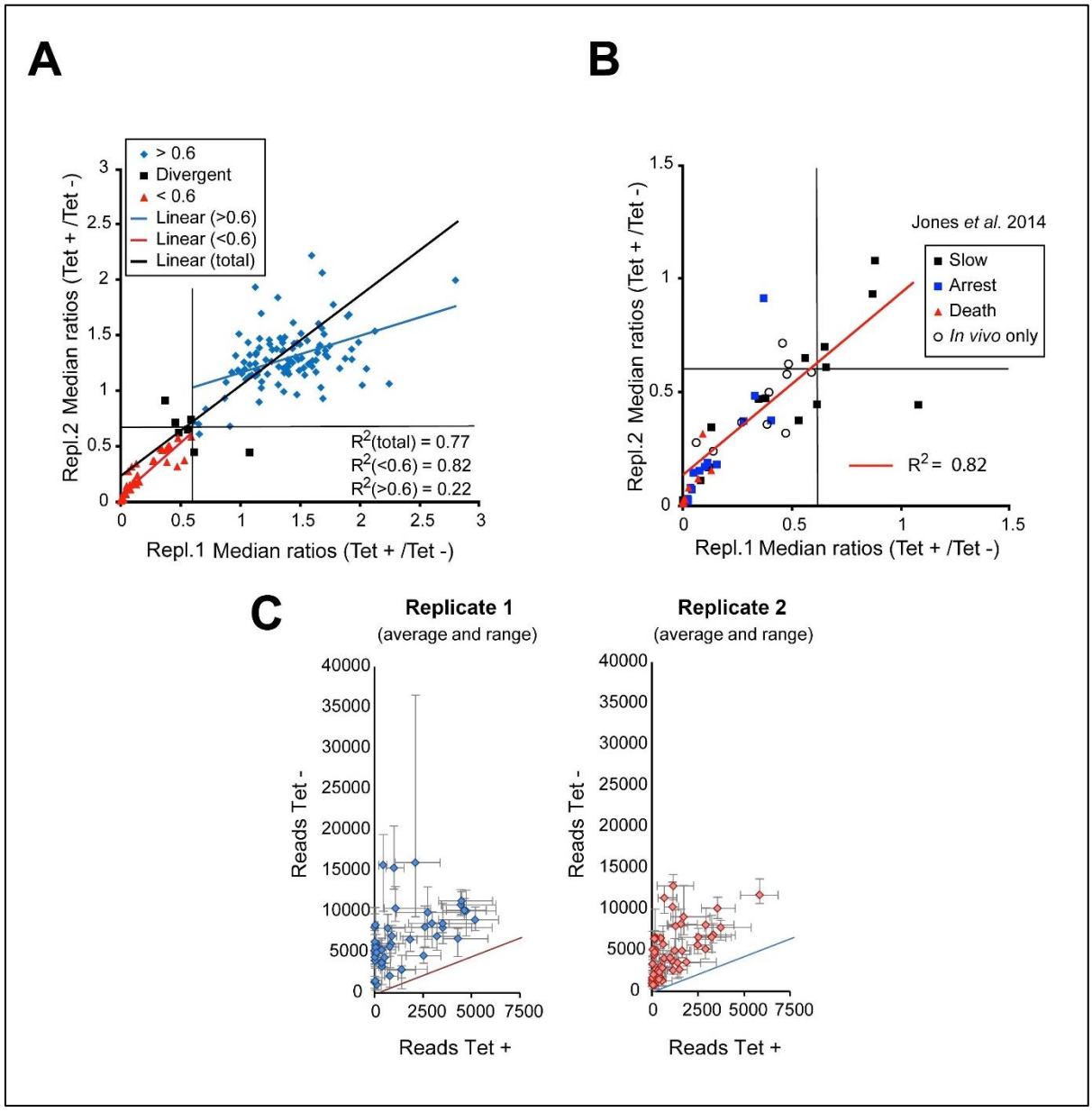


Figure 3-11. Intrinsic reproducibility of the *in vivo* RITseq and correlation between *in vivo* and *in vitro* experiments.
A. Bootstrap medians of ratios calculated between induced and uninduced *in vivo* samples plotted per RNAi target assessing regression of both biological replicates. Highlighted in red are those with a loss of fitness in the two experiments, in black those with a loss of fitness in one only, and in blue those without a loss of fitness. **B.** As panel A but focusing on the RNAi cell lines giving a loss of fitness (ratio <0.6) in at least one of the 2 biological replicates. Includes depicted correlation with the *in vitro* screen previously published by Jones et al. (2014) with the same library (slow growth, cell cycle arrest or death) highlighting protein kinases whose depletion produced loss of fitness only *in vivo*. **C.** Breakdown of results for loss of fitness RNAi lines within each of the individual *in vivo* biological replicates. Depicted average and range of normalized counts for the 6 mice analyzed per condition (Tet+ or Tet-) and per RNAi target.

Thirty-five out of 42 RNAi lines displaying loss fitness in the *in vitro* alamar blue screen (Jones *et al.*, 2014) also had a loss of fitness *in vivo*, with the most pronounced 'death' or 'cell cycle arrest' phenotypes *in vitro* also having the most significant loss of fitness *in vivo* (Figure 3-11B and Figure 3-12). Among the seven RNAi lines displaying loss of fitness *in vitro* but not *in vivo* (this study), six were classified as "slow growth" and they only had significant growth retardation after 72 h of RNAi induction, which would not have been apparent in the 48 h timescale of the *in vivo* screen. In the whole genome *in vitro* RNAi screen described previously (Alsford *et al.*, 2011), loss of bloodstream *T. brucei* fitness was assessed in culture after 3 and 6 days of RNAi induction. Only 19 (46%) of the loss of fitness-associated protein kinase genes observed in the Alsford study were also seen in the 48 h *in vivo* RITseq screen (Figure 3-12B).

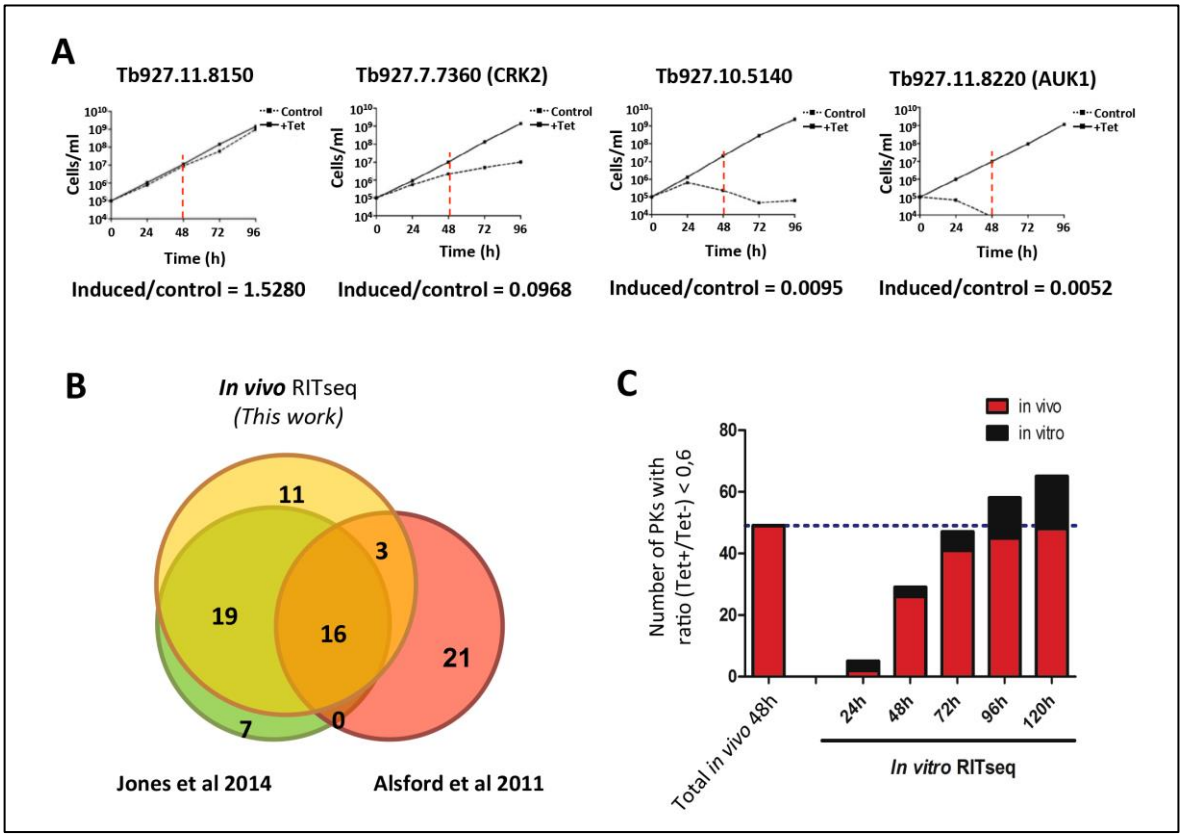


Figure 3-12. Comparison of *in vitro* and *in vivo* kinome-wide screens. **A.** *In vitro* growth curves for four independent RNAi cell lines with (from left to right) no growth defect, slow growth phenotype, growth arrest and cell death. Median ratios of induced/uninduced (Tet+/Tet-) for the *in vivo* RITseq are shown. **B.** Venn diagram showing the number of protein kinases with a loss of fitness in the present *in vivo* RITseq and two published *in vitro* RNAi studies (Alsford *et al.*, 2011; Jones *et al.*, 2014). **C.** Overlap between the number of cell lines with a loss-of-fitness phenotype detected *in vivo* after 48 h of induction and those detected with the *in vitro* RITseq over 120 h of RNAi.

In order to validate the observed correlation between the published *in vitro* loss-of-fitness phenotypes and those observed *in vivo*, growth analysis of 4 cell lines showing no growth defect, slow growth, arrest and death, as defined in Jones et al. (2014), was performed (Figure 3-12A). A good correlation was observed between the *in vitro* growth of individual lines and the median ratios of induced/control for the *in vivo* RITseq, in concordance with the analysis performed with the *in vitro* RITseq data to assess correlation between pooled and independent cell lines (Figure 3-9).

3.3.8 Comparison of the kinome wide RITseq *in vivo* and *in vitro*

Comparing the protein kinases that showed loss of fitness *in vitro* and *in vivo* in the kinome-wide RITseq revealed pronounced overlap, though it was clear that there was a delay in the manifestation of loss of fitness *in vitro* when compared with the *in vivo* infection (Figure 3-12C). Data for all genes can be observed in detail in Table 3-3. A total of 49 protein kinases caused a fitness defect *in vivo* after 48h of RNAi. However, *in vitro,* the numbers required longer to rise. After 24 h RNAi, 5 protein kinases had a reduced count in the induced compared with the uninduced *in vitro*. Three of them, all STE11 kinases, recovered later and the reduced counts may therefore be a product of stochastic fluctuation. Interestingly, one of these three is RDK1, reportedly triggering differentiation to procyclic forms after 72 h of RNAi induction (Jones *et al.*, 2014); this protein kinase showed a phenotype later, with an increasing loss of fitness from 96 h. The other two protein kinases, Tb927.10.2040 (STE11/CDC15-like) and CRK3, maintained an *in vitro* loss of fitness for the complete 120 h time frame. After 48 h RNAi, 29 cell lines had an *in vitro* loss of fitness phenotype, 26 of which were also found *in vivo,* providing a 59% overlap between experiments. At the 72 h time point, 47 cell lines had an *in vitro* loss of fitness phenotype, 41 of which had been already observed *in vivo* after 48 h of induction (83% overlap). After 96 h RNAi, we found 58 cell lines with a loss of fitness, 45 of which had already a phenotype *in vivo* at 48 h (91% overlap). By 120 h there was 97% overlap between RNAi lines showing loss of fitness *in vitro* and *in vivo* (Figure 3-12C). Only STE11/Bck1p showed a phenotype *in vivo* after 48 h that neither could be observed *in vitro* (in the whole 120 h experiment), nor was reported in the literature before.

The 49 genes with a mean < 0.6 of the median induced/uninduced ratios *in vivo* obtained in each of the two independent experiments were followed up in the next analyses (Table 3-5 and Figure 3-13).

Table 3-5. Meta-analysis of the 49 protein kinases with a loss of fitness *in vivo*.

Gene ID	Group/Family	Name	Loss of ftiness reported in the literature Including Jones (2014) & Alsford (2011)	Value for the HAT box Peña (2015)	Gene ID of *L. major* orthologues	Gene ID of *T. cruzi* orthologues
AGC	Tb927.3.2440	AGC1	Yes	-	LmjF.25.2340	0
	Tb927.7.5770	PK53	Yes	-	LmjF.06.1180	0
	Tb927.10.4940	PK50	Yes	-	0	TcCLB.510291.40
aPK	Tb927.3.5400	RIO1	Yes	-	LmjF.29.0240	TcCLB.510141.4
	Tb927.11.850	Bud32	Novel	-	LmjF.25.0870	TcCLB.506925.260
	Tb927.6.2840	RIO2	Yes	-	LmjF.30.1360	TcCLB.511753.70
	Tb927.11.14680	ATR	Yes	-	LmjF.32.1460	TcCLB.506223.120
	Tb927.1.1930	TOR4	Yes	-	LmjF.20.1120	TcCLB.508997.20, TcCLB.511435.49, TcCLB.506475.14, TcCLB.509989.10
	Tb927.10.8420	TOR1	Yes	Human orthologue targeted	LmjF.34.4530, LmjF.36.6320	TcCLB.508549.30, TcCLB.508551.10
	Tb927.4.420	TOR2	Yes	-	LmjF.34.4530, LmjF.36.6320	TcCLB.508549.30, TcCLB.508551.10
CAMK	Tb927.7.2750	CAMK3	Yes	-	LmjF.22.0810	TcCLB.510525.10
	Tb927.10.14770	AKB1	Yes	-	LmjF.19.0140	TcCLB.506211.220
	Tb927.7.6220	CAMK2	Yes	Human orthologue targeted	LmjF.17.0060	TcCLB.506513.50
	Tb927.2.1820	CAMKL	Novel	Divergent from human orthologue	LmjF.33.1710	TcCLB.510257.130
CK1	Tb927.5.790/5.800	CK1.1/CK1.2	Yes	Human orthologue targeted	LmjF.35.1010	TcCLB.504929.10, TcCLB.504929.15, TcCLB.504929.15, TcCLB.508541.230, TcCLB.508541.240
CMGC	Tb927.11.5340	CMGC1	Yes	-	LmjF.24.0670	TcCLB.504125.90
	Tb927.10.1070	CRK1	Yes	Human orthologue targeted	LmjF.21.1080, LmjF.36.0550, LmjF.18.1080,	TcCLB.504181.40, TcCLB.506677.20, TcCLB.509805.10

Gene ID	Group/Family	Name	Loss of ftiness reported in the literature Including Jones (2014) & Alsford (2011)	Value for the HAT box Peña (2015)	Gene ID of *L. major* orthologues	Gene ID of *T. cruzi* orthologues
					LmjF.35.3960, LmjF.35.4010	
	Tb927.2.4510	CRK9	Yes	Divergent from human orthologue	LmjF.27.1940	TcCLB.509099.150
	Tb927.7.7360	CRK2	Yes	-	LmjF.05.0550	TcCLB.510609.70
	Tb927.10.4990	CRK3	Yes	Human orthologue targeted	LmjF.21.1080, LmjF.36.0550, LmjF.18.1080, LmjF.35.3960, LmjF.35.4010	TcCLB.504181.40, cCLB.506677.20, TcCLB.509805.10
	Tb927.11.12410	CLK1/KKT10	Yes	-	LmjF.09.0400. LmjF.09.0410	TcCLB.511127.320, TcCLB.511127.330
GMGC	Tb927.7.3880	DYRK1	Novel	Divergent from human orthologue	LmjF.14.1070, LmjF.21.1650	TcCLB.506869.60
	Tb927.10.13780	GSK3-short	Yes	Human orthologue targeted	LmjF.18.0270	0
	Tb927.10.7780	KFR1	Yes	-	LmjF.36.6470	TcCLB.504167.30
	Tb927.10.5140	MAPK6 (ERK8)	Yes	Human orthologue targeted	LmjF.36.0720	TcCLB.510295.50
	Tb927.3.690	RCK1	Yes	-	LmjF.19.0180, LmjF.27.0100	TcCLB.506211.180, TcCLB.509231.20
	Tb927.7.960	SRPK1	Novel	-	LmjF.26.0980	TcCLB.477079.10, TcCLB.508215.3
	Tb927.6.4970	SRPK2	Yes	-	LmjF.30.3580	0
ORPHAN	Tb927.11.10520	KKT2	Yes	Parasite-specific	LmjF.36.5350	TcCLB.510285.70
	Tb927.9.10920	PK6/KKT3	Yes	Parasite-specific	LmjF.35.4050	0
OTHER	Tb927.3.3920	AUK2	Yes	-	LmjF.29.1330	TcCLB.503685.10
	Tb927.9.1670	AUK3	Yes	Divergent	LmjF.26.2440	TcCLB.506715.10, TcCLB.510349.80
	Tb927.11.8220	AUK1	Yes	Human orthologue targeted	LmjF.28.0520	0
	Tb927.10.15300	CAMKK	Yes	-	LmjF.19.0590	TcCLB.506505.40
	Tb927.2.2430	CK2A2	Novel	-	LmjF.02.0360	TcCLB.503513.10
	Tb927.9.14430	CK2A1	Yes	Human orthologue targeted	LmjF.35.1730	TcCLB.510761.60

Gene ID	Group/Family	Name	Loss of ftiness reported in the literature Including Jones (2014) & Alsford (2011)	Value for the HAT box Peña (2015)	Gene ID of *L. major* orthologues	Gene ID of *T. cruzi* orthologues
OTHER	Tb927.4.5310	NEK12.2 /RDK2	[16,17]	Divergent from human orthologue	LmjF.31.2960	0
	Tb927.6.5100	PLK-like	Novel	-	0	TcCLB.506945.350
	Tb927.7.6310	PLK1	[Yes	Human orthologue targeted	LmjF.17.0790	TcCLB.506513.160
	Tb927.4.5180/8.7220	TLK1&2	Yes	-	LmjF.31.2860	TcCLB.510597.9
	Tb927.11.4470	ULK	Yes	Human orthologue targeted	LmjF.13.0440	TcCLB.504089.54
	Tb927.11.9190	VPS15	Novel	-	LmjF.28.1760	0
PSEUDO	Tb927.9.6560	OTHER/NAK	Yes	-	LmjF.15.0770	TcCLB.504883.10
	Tb927.11.9290	ORPHAN/FAZ20	Novel	Parsite-specific	LmjF.28.1650	0
STE	Tb927.10.2040	STE11/CDC15-like	Yes	-	LmjF.21.0270	TcCLB.508949.50
	Tb927.10.10350	STE11/Bck1p	Novel	Human orthologue targeted	LmjF.05.0390	0
	Tb927.11.2040	STE11a	Yes	-	LmjF.27.1370	0
	Tb927.10.14300	STE11/MRK1	Novel	-	LmjF.32.0120	TcCLB.506525.140
	Tb927.8.5730	SLK1	Yes	Human orthologue targeted	LmjF.16.0300	TcCLB.506743.160

In red, not identified *in vitro* with the same library by Jones et al (2014).

3.3.9 Classification of protein kinases required for optimal growth *in vivo* after 48 h of induction

The 49 protein kinases whose depletion caused an *in vivo* loss of fitness are distributed across all the major protein kinase families. There are 40 ePKs, 7 aPKs and 2 putative pseudo-kinases. Two ePKs and one of the pseudo-kinases are considered "orphan" (without clear identity with known orthologues in *Saccharomyces cerevisiae*, *Caenorhabditis elegans, Drosophila melanogaster* or *Homo sapiens)* and are likely to be kinetoplastid-specific. All of the 49 protein kinases have orthologues in either *Leishmania major* or *Trypanosoma cruzi* or both, and 13 are closely related to human orthologues targeted by a compound in the HAT box with potential to be modified for specificity (Peña *et al.*, 2015), Figure 3-13, Table 3-5. Among these 13 we found, for example, AUK1, which showed the most intense *in vivo* phenotype in our *in vivo* screen (Table 3-3); this protein kinase is involved in cell cycle progression and is encoded by a single copy gene in kinetoplastids (Brown *et al.*, 2004) with active inhibitors found both in *T. brucei* and *L. major* (Patel *et al.*, 2014). We also see CK1.2, CRK1, CRK3, GSK3-short, ERK8 or PLK, all of which have been already proven interesting drug targets in *T. brucei* (Merritt *et al.*, 2014). Finally there is Bck1p, which is one of the novel reports in this thesis, with a loss of fitness *in vivo*, which materializes more quickly and more intensely than *in vitro,* Table 3-3.

After extracting their sequences from TritrypDB, and alignment in three different platforms (see material and methods), protein kinase domains were used to build a neighbour-joining phylogenetic tree, which demonstrates variable levels of sequence identity between orthologues present in different kinetoplastid parasites. Thirty-five of the protein kinases that were identified in the Alamar blue kinome-wide RNAi screen have been discussed previously (Jones *et al.*, 2014), and so details will not be provided here. Three of these genes were identified in the whole-genome RITseq (Alsford *et al.*, 2011): PLK1 and KKT2, which have been shown to be essential in other studies (Hammarton *et al.*, 2007; Akiyoshi and Gull, 2014), and AUK2, which only showed a loss of fitness *in vitro* after 120 h of induction (Table 3-3).

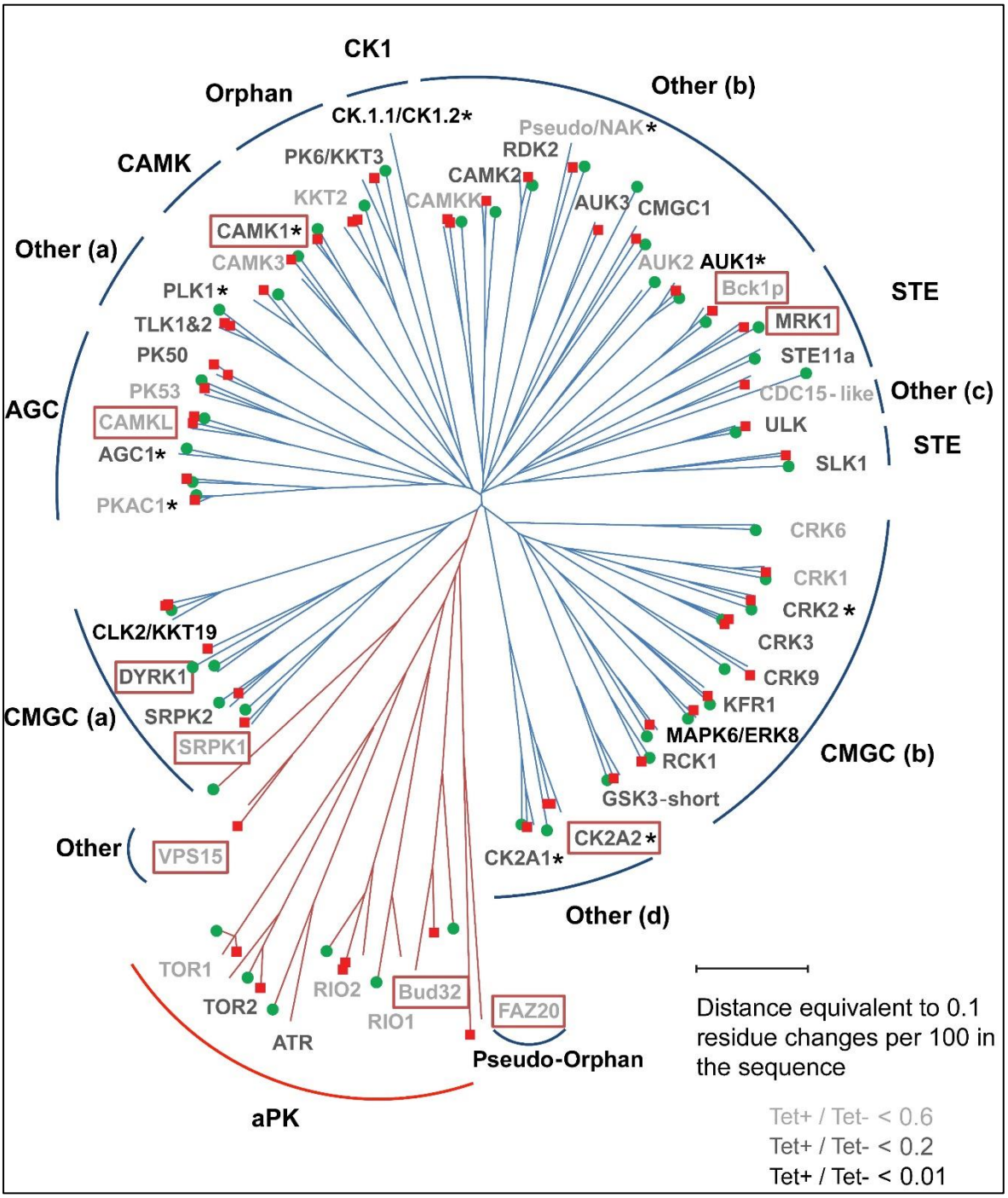


Figure 3-13. Neighbour-Joining phylogenetic tree of *T. brucei* protein kinases showing RNAi dependent loss-of-function *in vivo* with their *T. cruzi* and *L. major* orthologues.

Red bars represent aPKs and blue bars ePKs. Protein kinase lines causing a phenotype *in vivo* not reported before *in vitro* (Alsford *et al.*, 2011; Jones *et al.*, 2014) after RNAi induction are inscribed within a red box; those with a human orthologue targeted by a compound in the HAT box (Peña *et al.*, 2015) are labelled an asterisk; *L. major* and *T. cruzi* orthologues are labelled with a green circle and a red square, respectively. Grey scale describes intensity of the loss of fitness according to the mean Tet+/Tet- ratio.

3.3.10 Protein kinases essential *in vivo* but not *in vitro*

In addition to AUK2, we identified 11 other cell lines where depletion of the protein kinase led to a loss of fitness *in vivo*, but not in either the kinome- (Jones *et al.*, 2014) or the genome- wide (Alsford *et al.*, 2011) *in vitro* RNAi studies (Table 3-5 and Figure 3-13). In order to validate the observation *in vitro*, independent growth curves were performed in culture for each of the individual RNAi cell lines used in the inoculum pool. Where a growth defect was not detected after 48 h of RNAi induction *in vitro*, parasitemia was assessed in mice after infection with the selected clone for 96 h (Figure 3-14). Tb927.10.14770 (AKB1) was not examined as has been shown to be an active kinase involved in regulation of cytokinesis and cell division in bloodstream and procyclic forms upon knock down from 48 h onwards (Inoue *et al.*, 2015). The reason why it was not detected by Jones et al. is that the original cell line used in Replicate 1 did not contain an insert (Table 3-4). Another protein kinase gene, Tb927.7.3880 (DYRK), had a clear and unexpected growth defect after 48 h of RNAi induction *in vitro* (Figure 3-14A), and so this gene was not considered further. For the other protein kinase genes, where no growth defect was detected after 48 h of RNAi induction *in vitro*, parasitemia was assessed for 72-96 h in mice, with or without induction of RNAi, after inoculation of each individual selected cell line (Figure 3-14B).

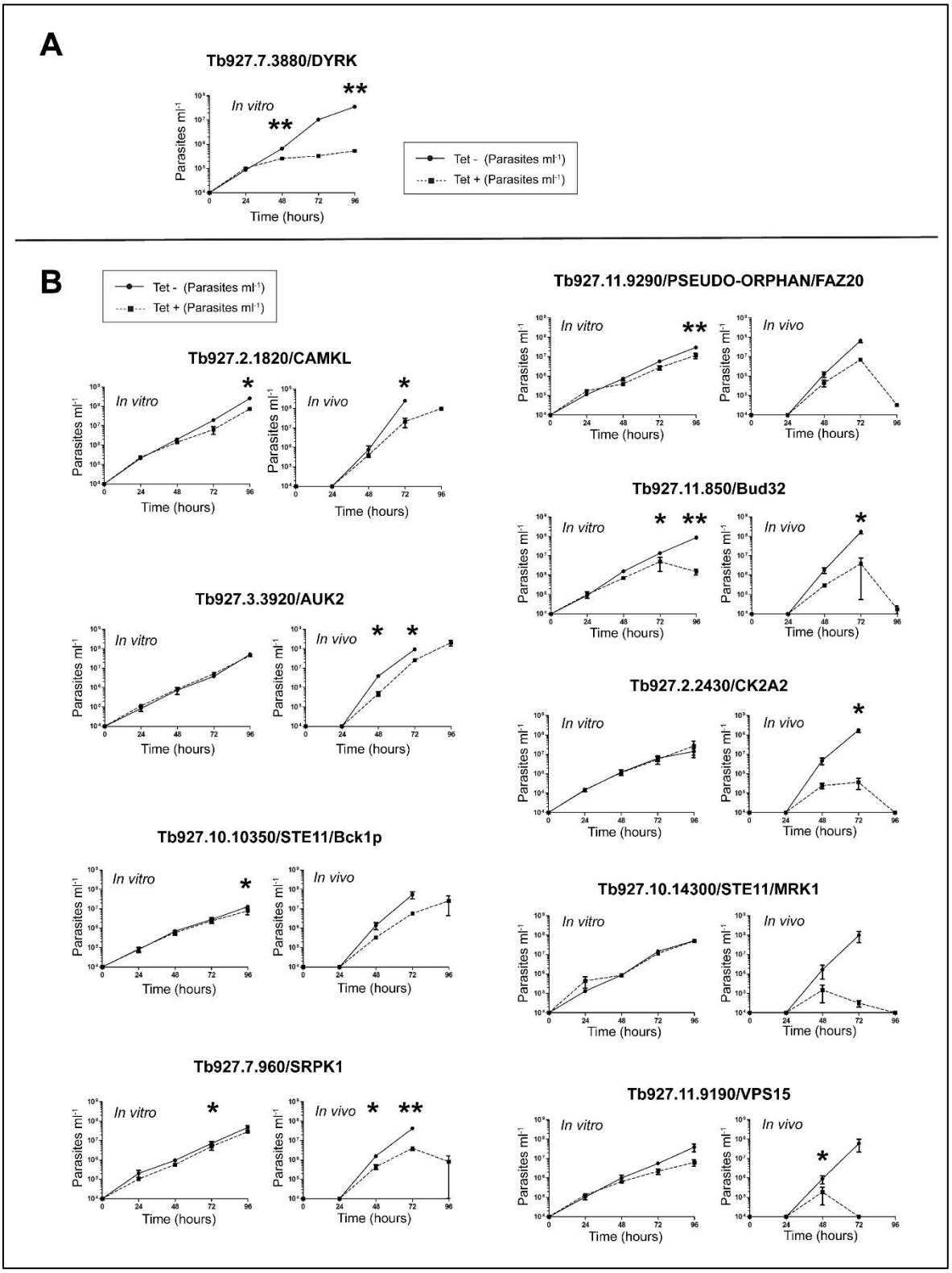


Figure 3-14. Validation of RITseq *in vivo*-mainly loss of fitness phenotypes. For each RNAi cell line, cumulative *in vitro* growth curves are shown on the left hand side; and *in vivo* parasitemia on the right hand side. Growth curves were built in triplicate. *In vitro* cultures started at 10^4 cells x ml^{-1}. Mice were culled if parasitemia reached 10^8 cells x ml^{-1}. **A.** Tb927.7.3880 (DYRK). **B.** 9 other cell lines with a loss of fitness phenotype *in vivo* that is more pronounced than the one observed *in vitro*. * $p < 0.05$ and ** $p < 0.01$ using T-test.

BLAST alignment with protein sequences in other eukaryotes allowed a tentative proposal of the putative function of each protein kinase based on gene ontology (Table 3-6). Tb927.11.9290 (Pseudo-unique), Tb927.2.1820 (CAMKL), Tb927.11.850 (aPK/Bud32) and Tb927.11.9190 (Other/VPS15) displayed some evidence for growth retardation after RNAi *in vitro* (Figure 3-14B) but, in each case, the extent of growth impairment or death was more severe *in vivo* and arose more quickly. For each of the remaining six cell lines there was no evidence of RNAi-induced growth retardation *in vitro*. For two of the genes, Tb927.3.3920 (AUK2) and Tb927.10.10350 (STE11/Bck1p), RNAi caused growth retardation *in vivo* from 48 h, indicating a loss of fitness. For three other genes, Tb927.2.2430 (Other/CK2A2), Tb927.7.960 (CMGC/SRPK1) and Tb927.10.14300 (STE11/MRK1), the loss of fitness after *in vivo* RNAi was even more severe, since a decrease in parasitemia or clearance was seen before 96 h (Figure 3-14B). Taken as a whole, these growth curves validate the kinome-wide RITseq both *in vivo* and *in vitro* for nine of eleven predicted protein kinases (Table 3-3).

Table 3-6. Predicted function of *in vivo* only protein kinases based on gene ontology with other eukaryotes.

pBLAST	Organism	Identity	E value	Protein	Putative function
Tb927.2.2430 (CK2A2)	*T. brucei*	100%	0	casein kinase II, α chain (CK2A2)	· Proven kinase activity. · Mainly nucleolar but ubiquitous (Jensen *et al.*, 2007), controlled by distribution and abundance (Pinna, 2003): stress-induced mobilization (Gerber *et al.*, 2000; Davis *et al.*, 2002). · Gene expression/protein synthesis: survival, differentiation and proliferation (>300 substrates) (Meggio and Pinna, 2003). · Downregulation causes autophagy through PI3K/AKT/TOR pathway (Olsen, Svenstrup and Guerra, 2012; Hales, Taub and Matherly, 2014; Sanchez-Casalongue *et al.*, 2015). · Intersects with mitogen and stress-activated protein kinases (Shi *et al.*, 2009; Jacks and Koch, 2010).
	L. major	60%	$4x10^{-148}$	putative casein kinase II, α chain	
	T. cruzi	73%	0	casein kinase II, α chain	
	S. cerevisiae	48%	$4x10^{-95}$	Cka2p	
	C. elegans	48%	$1x10^{-87}$	Casein kinase II subunit α	
	Human	51%	10^{-109}	Chain A, Crystal Structure Of Ck2	
Tb927.10.14300 (STE11/MRK1)	*T. brucei*	100%	0	MEKK-related kinase 1, putative (MRK1)	· Member of the HOG pathway (Zhi *et al.*, 2013) · Response to stress signalling, mainly hyperosmotic shock (through actin recovery pathway (Zhi *et al.*, 2013) · Essential for promastigote survival, in Leishmania (Agron, Reed and Engel, 2005)
	L. major	37%	$6x10^{-174}$	MEKK-related kinase 1, putative (MRK1)	
	cruzi	50%	0	MEKK-related kinase 1, putative (MRK1)	
	S. cerevisiae	36%	$1x10^{-43}$	Ssk2p	
	C. elegans	34%	$7x10^{-37}$	Germinal Center Kinase family	
	Human	36%	$1x10^{-46}$	MAPKKK19 isoform 3	
Tb927.10.10350 (STE11/Bck1p)	*T. brucei*	100%	0	protein kinase, putative	· Contains 8 amino terminal "MORN motif" sequences: subcellular localization, membrane fusion, fission and mobility in other organisms (Lee, Han and Hur, 2010). · Cell integrity via MAPK pathway in response to growth factors and stress (hyperosmotic, oxidative and sheer fluid
	L. major	60%	0	protein kinase, putative	
	T. cruzi	73%	0	protein kinase, putative	
	S. cerevisiae	41%	$8x10^{-55}$	Bck1p	
	C.elegans	39%	$2x10^{-44}$	hypothetical protein F59A6.1	

	Human	45%	$4x10^{-65}$	MEK kinase 3	shocks) (Widmann *et al.*, 1999) · Direct activation of the stress-activated protein kinase (SAPK) and extracellular signal-regulated protein kinase (ERK) pathways (Ellinger-Ziegelbauer *et al.*, 1997)
Tb927.11.850 (aPK/Bud32)	*T. brucei*	100%	0	protein kinase, putative	· KEOPS/EKC complex: translational regulation throughout tRNA modification (Srinivasan *et al.*, 2011; Rojas-Benítez, Ibar and Glavic, 2013). · Transducer for TOR activation that, in turn, has also been related with autophagy and endocytosis under stress conditions (Ibar *et al.*, 2013; Rojas-Benítez, Ibar and Glavic, 2013).
	L. major	48%	$1x10^{-70}$	protein kinase, putative	
	T. cruzi	63%	$2x10^{-103}$	protein kinase, putative	
	S. cerevisiae	32%	$4x10^{-24}$	Bud32p	
	C.elegans	35%	$2x10^{-40}$	Uncharacterized protein CELE_F52C12.6	
	Human	40%	$2x10^{-44}$	TP53 regulating kinase	
Tb927.11.9190 (aPK/VPS15)	*T. brucei*	100%	0	protein kinase, putative	· PI3K complex mamber: autophagosome formation/regulation of protein and vesicular trafficking and sorting (Abraham, 2004; Araki *et al.*, 2013; Liu *et al.*, 2014; Anding and Baehrecke, 2015).
	L. major	33%	$2x10^{-64}$	protein kinase, putative	
	T. cruzi	50%	0	protein kinase, putative	
	S. cerevisiae	23%	$6x10^{-32}$	Vps15p	
	C.elegans	33%	$7x10^{-23}$	Vacuolar Protein Sorting factor	
	Human	24%	$6x10^{-38}$	PI3K regulatory subunit 4	
Tb927.7.960 (CMGC/SRPK)	*T. brucei*	100%	0	protein kinase, putative	· Nucleus/cytoplasm localization · Phosphorylation of splicing factors containing serine/arginine-rich domains · Regulation of constitutive/alternative splicing (Zhou *et al.*, 2012). Stress-response mechanism in eukaryotes (Zhong *et al.*, 2009).
	L. major	53%	$8x10^{-173}$	protein kinase, putative	
	T. cruzi	51%	0	protein kinase, putative	
	S. cerevisiae	40%	$5x10^{-26}$	Sky1p	
	C.elegans	39%	$6x10^{-24}$	SR Protein Kinase	
	Human	41%	$2x10^{-27}$	SFRS protein kinase 1, isoform CRA_c	
Tb927.11.9290 (Pseudo-Orphan/FAZ20)	*T. brucei*	100%	0	protein tyrosine kinase, putative	· Parasite-specific · Localizes to the tip of the flagellum attachment zone in both the old and the new cell Generated during cytokinesis (Zhou, Hu and Li, 2016) · None predicted kinase activity (Figure 3-18) and (Taylor *et al.*, 2013)
	L. major	32%	$2x10^{-35}$	protein kinase, putative	
	T. cruzi	34%	$2x10^{-110}$	protein kinase, putative	
	S. cerevisiae	26%	$8x10^{-7}$	Gin4p	
	C.elegans	25%	$3x10^{-4}$	unc-82	
	Human	25%	$4x10^{-4}$	NUAK family SNF1-like kinase 2	
Tb927.3.3920 (AUK2)	*T. brucei*	100%	0	protein kinase, putative	· Associates to the spindle poles · Regulates entry in mitosis and it is often overexpressed in human cancers (D'Assoro, Haddad and Galanis, 2015) · DNA repair candidate (Stortz, 2016)
	L. major	49%	$9x10^{-89}$	protein kinase, putative	
	T. cruzi	60%	0	protein kinase, putative	
	S. cerevisiae	31%	$1x10^{-33}$	Kin82p	
	C.elegans	34%	$4x10^{-50}$	Aurora/IPL1-related protein kinase 2	
	Human	37%	$7x10^{-54}$	Chain A, Structure Of Aurora-2 (AUKA)	
Tb927.2.1820 (CAMKL)	*T. brucei*	100%	0	protein kinase, putative	· Energy balance maintenance (Salminen A, Kauppinen A and Kaarniranta K, 2016) · Phosphorylation of metabolic enzymes and transcription factors · Epigenetic regulation · Calcium binding (contains EF hands)
	L. major	57%	0	protein kinase, putative	
	T. cruzi	63%	0	protein kinase, putative	
	S. cerevisiae	40%	$8x10^{-57}$	Ampk Homolog Snf1	
	C.elegans	61%	$7x10^{-59}$	Ampk subunit alpha-2	
	Human	38%	$1x10^{-55}$	Ampk subunit alpha-2	

Results of BLASTp analysis aligning against a non-redundant protein sequence database. E = Expect value.

3.3.11 Increased susceptibility of the *in vivo*-impaired RNAi mutants to fresh serum exposure

A companion study (Stortz, PhD thesis, 2016) showed three of the *in vivo* specific kinases to have an important role in maintaining cellular integrity, with RNAi induced parasites shown to be more sensitive to alkylation-induced cellular damage with methyl methanesulfonate: Tb927.2.1820 (CAMKL), Tb927.3.3920 (AUK2), and Tb927.7.960 (CMGC/SRPK1). For AUK2 at least, this phenotype is due to a role in genome maintenance or transmission. One of the most obvious and distinctive elements encountered by parasites during infection in mice when compared with the *in vitro* culture conditions is exposure to mammalian serum factors, with the VSG coat providing protection against complement-mediated lysis (Matthews, McCulloch and Morrison, 2015). As genome fragility affecting subtelomeric regions can disrupt VSG expression (Glover, Alsford and Horn, 2013), we exposed the three candidates to fresh rat serum for 3 h (10% and 50% concentration) and analysed relative survival of induced and uninduced cells when compared to controls that were not exposed to rat serum (Table 3-5). In addition, we tested Tb927.7.3880 (DYRK) that was found to produce a growth defect both *in vivo* and *in vitro*.

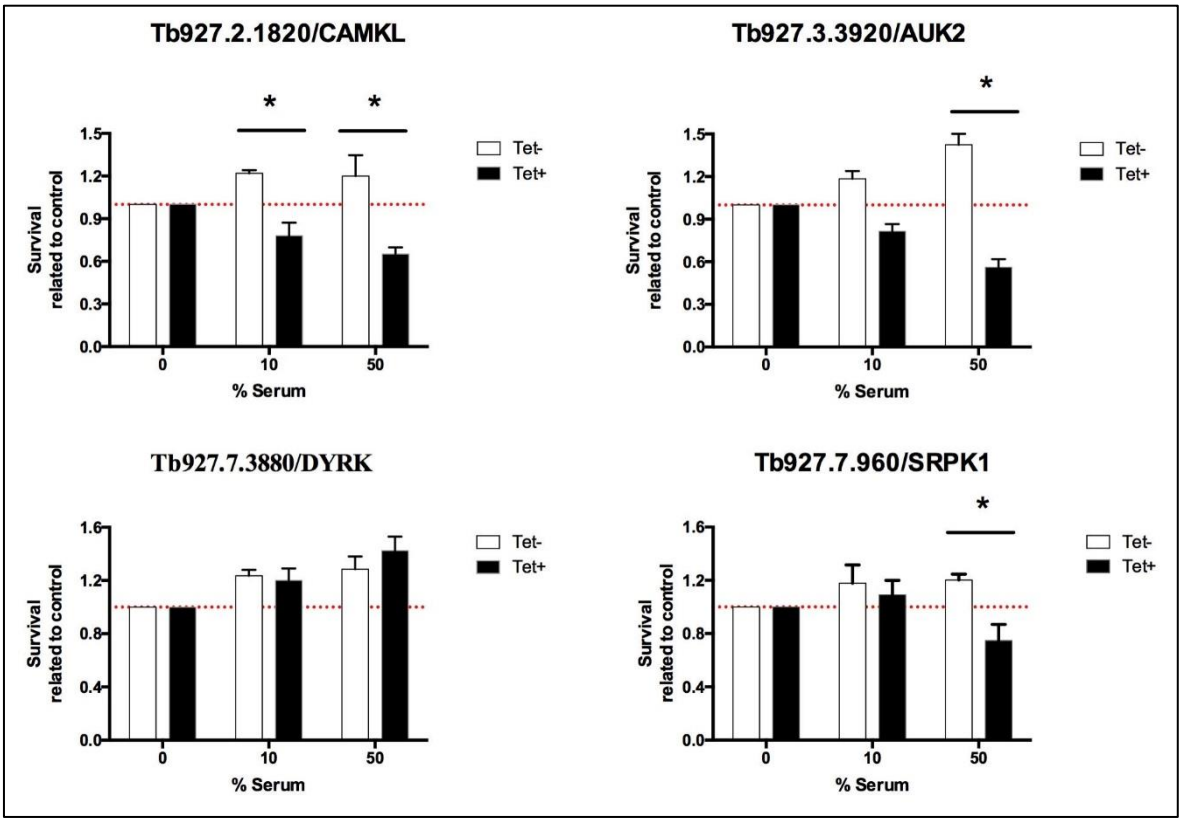


Figure 3-15. Survival assay to rat serum exposure of four *in vivo*-specific protein kinases related to DNA repair.
Parasites at 10^6 cells ml^{-1} were exposed to fresh rat serum for 3h (10% and 50% concentration). Relative survival was measured by normalizing counts for induced (48 h RNAi) and uninduced parasites to their respective no-serum controls. Induced and uninduced results were compared with T test. * = $p<0.05$ (n= 3).

RNAi induced cells, targeting the three MMS damage response protein kinases, showed a decrease in survival as serum concentration increased relative to uniduced controls. In contrast, RNAi against Tb927.7.3880 (DYRK), which caused impaired cell growth *in vitro,* caused no such increased serum sensitivity.

3.3.12 STE11 (MAP3K) kinases involved in osmotic shock resistance

Two of the MAP3K protein kinases (Tb927.10.14300 and Tb927.10.10350) that are essential only *in vivo* (Figure 3-13) have sequence identity to yeast protein kinases that regulate osmotic stress (Table 3-6). To test whether they might be involved in resistance to osmotic changes encountered during circulation in the mammalian host, a modified 'Swell dialysis' assay was performed (Voorheis and Martin, 1980). Controlled buffering conditions were used to produce mildly

turgid cells and RNAi-induced cells were assessed for survival under gentle osmotic shock compared with uninduced controls (Figure 3-16). RNAi against each predicted STE11 protein kinase resulted in a statistically significant reduction in cell survival under these conditions. RNAi of Tb927.10.14300 resulted in a 20% reduction in cell viability after osmotic shock at both 48 h and 72 h after RNAi induction. RNAi of Tb927.10.10350 resulted in a more severe phenotype with a 50% reduction in cell viability after osmotic shock at 48 h and 72 h after RNAi induction.

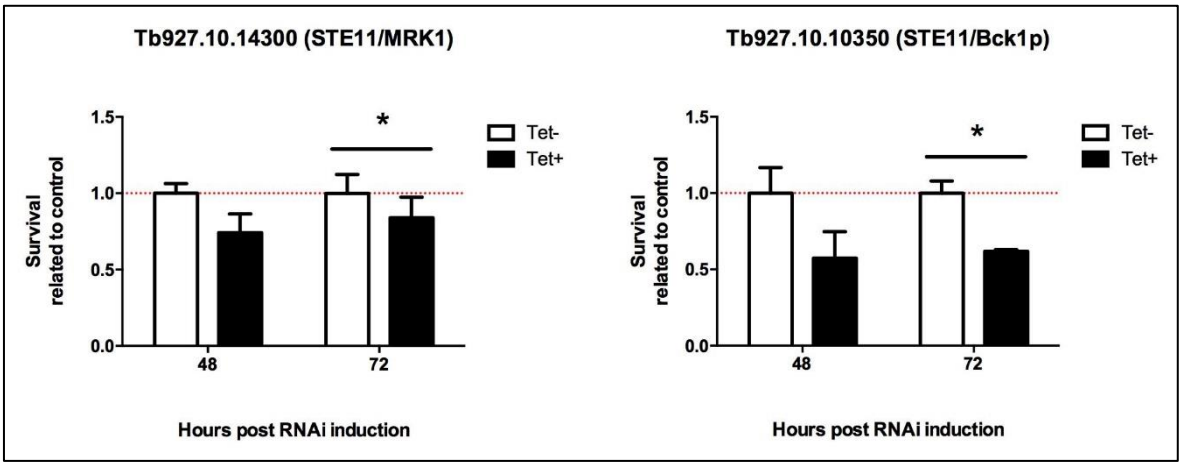


Figure 3-16. Survival assay to osmotic shock of two MAP3Ks (STE11s) found in the *in vivo* specific set.
2 x 10^6 cells were subjected to a modified swell dialysis protocol (Voorheis and Martin, 1980): cell swelling (5 min at 4 ºC in 0.5 ml 55 mM KCl/1 mM glucose) followed by cell shrinking adding on top 0.5 ml 263 mM KCl/1.75 mM Mg_2Cl (5 min at 4ºC). Survival of the induced cells compared to uninduced controls is plotted and compared with T test. *= p <0.05.

3.3.13 Kinome focused approach compared to genome-wide RITseq screenings

The rationale behind the kinome wide screen was that the RITseq approach applied to a reduced cohort of cell lines that had been individually developed and validated, instead of with a large library of thousands of clones, would permit sufficient coverage with the reduced inoculum size required for a mouse infection.

As overlap in the growth phenotypes seen in the two RITseq approaches was less than a 50% of identified genes, one interesting aspect was to understand how different coverage and resolution was. The whole genome RITseq data (Alsford

et al., 2011) was downloaded from the European Nucleotide Archive and processed through our pipeline. Three genes were selected, one showing a loss of fitness in both screens (Tb927.3.690), one identified only in the whole genome (Tb927.8.5390), and one identified only in the kinome-wide screen (Tb927.11.5340). As expected, whole genome reads scattered across the whole genomic sequence of each protein kinase, while the kinome focused mapped accurately in the flanks of the homology region targeted by the RNAi. Read depth in the focused library was at least 100x read depth found in the whole genome library (Figure 3-17).

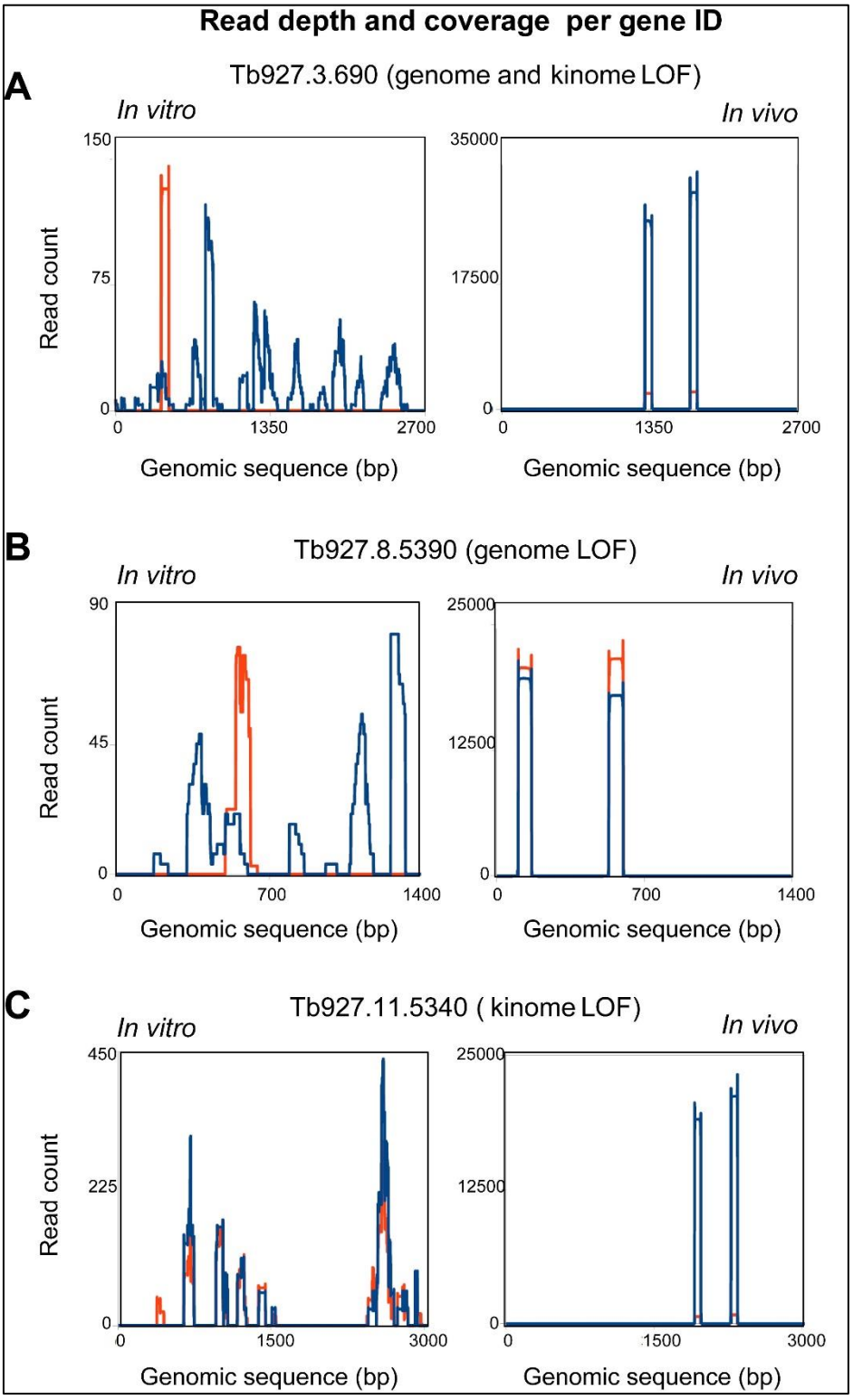


Figure 3-17. Comparative kinome focused/whole genome RITseqs (Alsford *et al.*, 2011)**.**
A. Compared read count resolution for a gene that causes loss of fitness (LOF) both in the kinome-wide and in the genome-wide screening (Tb927.3.690); **B.** for a gene that is a LOF only in the genome-wide (Tb927.8.5390); and **C.** for a gene that is a LOF in the kinome-wide (Tb927.11.5340). Red=reads coming from an induced sample. Blue= reads coming from an uninduced sample.

3.4 Discussion

In this study we aimed to define the cohort of protein kinases that are essential for *T. brucei* to survive in the mammalian bloodstream. We modified whole-genome parallel phenotyping based on RITseq technology (Alsford *et al.*, 2011) and applied it to a defined collection of individually designed and validated RNAi cell lines. This method was used to run *in vitro* and *in vivo* experiments that permitted the identification of 49 essential protein kinases, including 9 protein kinases that are uniquely required for survival in the environment of the mammalian circulatory system, Figure 3-14. Protein sequence alignment for each gene suggested these protein kinases are involved in parasite stress-response pathways, based on orthology with other eukaryotes. Validation of individual cell lines identified protein kinases that play roles in protecting the parasite from mammalian serum factors and osmotic stresses. These findings expand the repertoire of essential and potentially druggable protein kinases in this pathogen and begin to explore the mechanisms by which the parasite survives the multiple stresses of life inside its host.

3.4.1 Adaptation of RITseq technology to pre-existing RNAi libraries

Viability of the pooled approach is validated by the 91% overlap found between the *in vitro* RITseq at 72 h and results for individual cell lines obtained in the alamar blue screen for the same library at the same time point (Jones *et al.*, 2014), supported by directly comparing growth of fourteen independent lines and the RITseq results obtained with the pooled approach at every time point. The same can be inferred from the *in vivo* experiment results, where loss of fitness matched observations made with the individual lines by Jones et al (2014), supported by analysis of 4 independent cell lines selected to represent the range of loss of fitness phenotypes observed. The adaptation of genome-wide RITseq technology to smaller, defined libraries was advantageous for *in vivo* experiments for several reasons. The coefficient of determination between the RNAi induced/uninduced read mapping ratios obtained between the two experiments (R^2=0.77) was good, especially when focused on those with a loss of fitness RNAi phenotype (R^2=0.82). Intrinsic reproducibility across replicates *in vitro* and, remarkably, across different mice within the same experiment was

also good, as evidenced by the reduced standard deviation of the average read counts. Inclusion of a single clone per RNAi target may allow an advantage in terms of phenotypic homogeneity compared with strategies where a pool of RNAi vectors are transfected into parasites. The second approach generates an unknown number of clones per target with different levels of depletion after induction, which may mask certain phenotypes.

We show that focusing RITseq to smaller libraries enhances resolution of the loss of fitness readout. While whole-genome sequencing maps reads across ~7500 genes, our approach distributes the same reads along 183 genes (177 rationally targeted fragments, including six double knockdowns), enabling a ~1000-fold increase in sequence read depth. This increase in resolution meant that we were able to detect 56 protein kinases that result in a loss of fitness upon depletion with the kinome-library *in vivo* or *in vitro*, while only 19 of these protein kinases, with the most severe growth defects, were identified in a whole-genome library RITseq screen (Alsford *et al.*, 2011) that assessed loss of fitness in culture after 3 and 6 days of RNAi. A similar observation has been made in a screen of 99 mitochondrial proteins with procyclic forms (Mbang-Benet *et al.*, 2015). There, 44 RNAi phenotypes (some of them supported by individualized studies published before) differed from whole-genome RITseq results, but both studies were coincident in detection of the most severe phenotypes.

Although whole genome RITseq screens may have lower resolution, they assess in parallel all the protein families, thus being a valuable tool to identify pathways and networks whose activity is required under any selective pressure or culturing condition. This has been successfully applied, for instance, to identify proteins required for stumpy and procyclic differentiation (Alsford *et al.*, 2011; Mony *et al.*, 2013), proteins whose depletion triggers drug resistance (Baker, Alsford and Horn, 2011; Alsford *et al.*, 2012), or resistance to human serum factors (Alsford *et al.*, 2014).

Both approaches are complementary: strong phenotypes caused by any gene can be identified using genome wide libraries, while gene families can be quantitatively assessed in higher resolution with the focused-library approach. Importantly for this study, reduced library size and associated RITseq resolution allowed a 72 h *in vivo* model of infection in mice, which would be difficult with

the larger parasite numbers required to cover the whole genome without using lots or bigger animals, as they would cause an early host death by overwhelming the bloodstream (Turner, Aslam and Dye, 1995). The comparison between the *in vivo* and *in vitro* RITseq screens presented in this chapter also highlights that phenotypes observed *in vitro* could be observed generally before and with a higher statistical significance (maybe due to a larger number of replicates) *in vivo* than *in vitro* .

3.4.2 Nine protein kinases causing a loss of fitness *in vivo* unveiled as novel virulence factors

Trypanosomes are highly adaptable organisms, capable of withstanding immune assault by the host and responding to the many changing conditions they experience during their life cycles: temperature, pH, osmotic imbalance, availability of carbon and nitrogen sources. This study reveals that depletion of 9 protein kinases limits *T. brucei* growth in the mouse bloodstream, while the same cells remain fully competent for growth in culture (Figure 3-14), suggesting they have roles as virulence factors required to overcome challenges encountered only *in vivo*. The precise functions provided by the protein kinases remain to be identified, but a common theme from sequence-based functional predictions (Table 3-6) is the management of stress. It has been shown in bacteria that, in addition to 'direct' or *sensu stricto* virulence factors (such as toxins or invasins), elements involved in stress management are also 'indirect' or 'contributory' elements required for establishment of the infection (Sleator and Hill, 2002). A connection between stress resistance and virulence has been found for many other infectious eukaryotic organisms (Hengge-Aronis *et al.*, 1993; Mattick *et al.*, 2000; Loh, Torres and Cover, 2007; Cameron *et al.*, 2012; Burbank and Stenger, 2016), including *Leishmania* (Zhang and Matlashewski, 1997; McCall and Matlashewski, 2010).

Nine independent cell lines were tested *in vivo* and *in vitro* individually, showing consistency with the pooled approach. Six of them showed significant loss of fitness at the RITseq *in vitro* only after 120 h, with Bud32 and MRK1 showing also a phenotype after 96 h RNAi. Among these novel virulence factors of *T. brucei* there is an orphan pseudo-kinase with putative parasite-specific functions and another eight protein kinases showing sequence similarity with protein kinases in

other eukaryotes, suggesting diverse functions including gene expression, alternative splicing, protein synthesis/translational regulation, and stress adaptation linked with autophagosome formation.

A gene first included in the *in vivo*-only set, Tb927.10.14770 (CAMK/AKB1), was not reported at the *in vitro* alamar blue screen (Jones *et al.*, 2014) because the clone used in the original library did not contain an RNAi insert. After being remade for the second pool (re-transfecting the plasmid, pTL179), it was used for the *in vivo* and for the *in vitro* RITseqs, causing a rapid and severe loss of fitness *in vivo* and *in vitro* only after 96 h. AKB1 has been proven to be an active kinase, downregulating cytokinesis and cell division in bloodstream and procyclic forms upon knock down *in vitro* from 48h (Inoue *et al.*, 2015). The phenotype there also involved morphological abnormalities. Inoue et al. showed overexpression to cause a similar phenotype in procyclic forms, with the AKB1 kinase activity responsible for this (as a kinase-dead overexpressor did not produce a similar effect).

3.4.2.1 An orphan *pseudo-kinase* with parasite specific functions

One of the nine protein kinases differed from the rest, in that it appears to be a highly divergent putative pseudo-kinase (Tb927.11.9290/FAZ20), since it lacks several of the key residues and subdomains required to be an active protein kinase (Hanks and Hunter, 1995), Figure 3-18. In the small ATP-binding NH2-terminal lobe, FAZ20 misses two of the three conserved glycines needed in subdomain I to anchor the non-transferable phosphates of the nucleotide. In addition, the large COOH terminal lobe, which binds the substrate and initiates the phosphotransference in most protein kinases, is truncated. FAZ20 also lacks the DxxxxN catalytic loop in subdomain VIB (which accepts the proton from the substrate´s attacking COOH group) and the DFG motif at subdomain VII (responsible for γ-phosphate orientation). Finally, it lacks the APE motif in subdomain XIII, which forms an ion pair with an arginine in subdomain XI, while subdomain XI is totally missing.

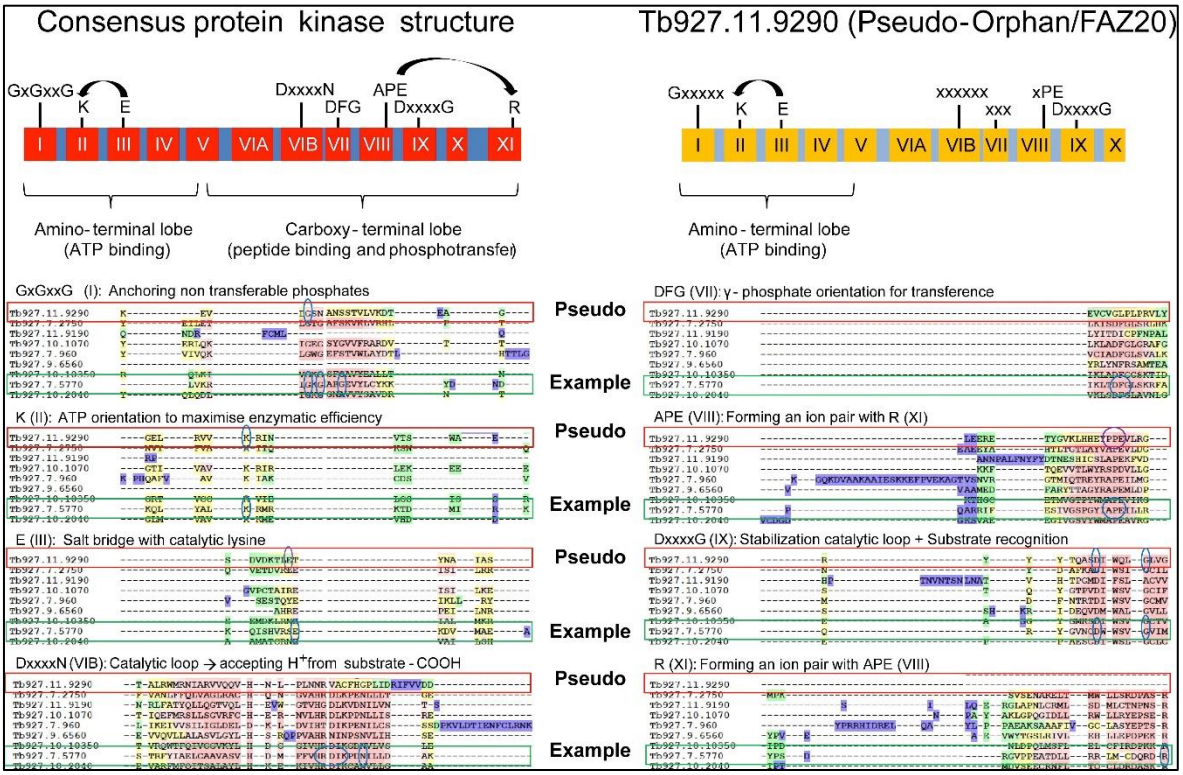


Figure 3-18. FAZ20 is a putative pseudo-kinase.
Extract of the alignment performed for the phylogenetic tree displaying amino acid sequence alignment of Tb927.11.9290 (Pseudo-Orphan/FAZ20) with other protein kinases and its lack of some of the critical subdomains and motives conferring to ePKs their canonical kinase activity. Protein kinase model inspired in the one depicted at Hanks, 2003 (Hanks, 2003).

Tb927.11.9290 (FAZ20) was already identified as an 'orphan' kinase in the seminal article describing the *T. brucei* protein kinome, suggesting an evolutionary divergence with protein kinases in other eukaryotes (Parsons *et al.*, 2005). This is supported by observations here (Table 3-6) of poor BLASTp E-values after alignment with similar protein kinases in other reference organisms (apart from the closely related *Leishmania* and *T. cruzi* parasites). Such a structural divergence may indicate parasite-specific functions, as may be inferred from FAZ20 localization to the tip of the flagellum attachment zone in both the old and the new cells generated during cytokinesis (Zhou, Hu and Li, 2016). Through BioID (proximity-dependent biotin identification) and immunolabelling, its presence has been detected close to CIF1 and CIF2. These two factors play a direct role in the canonical anterior end-initiated cytokinesis, in cooperation with PLK1 and AUK1 (both essential also in our *in vivo* screen). RNAi-mediated depletion of CIF1 or CIF2, as well as CIF2 overexpression, causes

atypical initiation of cytokinesis from the posterior end, indicating the need of a tight regulation (Zhou, Hu and Li, 2016).

Altogether, these data may suggest that FAZ20 involvement in the process of cell division may become particularly crucial in the context of the *in vivo* infection, where linkage with stress resistance mechanisms, as observed for other protein kinases detected in the screen, may ensure correct cytokinesis under challenges encountered in the bloodstream that are not reproduced *in vitro*. FAZ20 classification as a putative pseudo-kinase (Hanks, 2003; Taylor *et al.*, 2013) in the context of the BioID and localization results (Zhou, Hu and Li, 2016) may also suggest scaffolding roles for this protein.

3.4.2.2 Two MAP3Ks (STE11) involved in osmotic shock resistance

In other eukaryotes, Mithogen Activated Protein Kinases (MAPKs) are transcription activators acting at the end of a three module phosphorylation cascade, providing a response to extracellular stimuli in order to ensure a quick adaptation to environmental conditions: growth factors, cytokines, irradiation, osmolarity or shear stress caused by the hydrodynamic flux over the cells as they proliferate in the bloodstream. MAP3Ks (STE11s) are the first step of these cascades; they interact with MAP2Ks (STE7s), which in turn act upon MAPKs that, in the end, are responsible for transcription factor activation (Widmann *et al.*, 1999).

In *T. brucei*, this pathway may have non-canonical roles, as transcription mechanisms are divergent from other eukaryotes. RNA polymerase II initially produces large polycistronic transcripts that are trans-spliced to produce mature mRNAs, mediated by the 39–nt-long “Splice leader” (SL). The SL transcript is unique, being the only gene for which an RNA Pol II promoter has been found in *T. brucei* so far (Das *et al.*, 2005; Schimanski, Nguyen and Gu, 2005). RNA polymerase II is believed to initiate transcription from strand switch regions, regulated by epigenetic mechanisms (Siegel *et al.*, 2009). Codon content (Horn, 2008) and gene duplication (Jackson, 2007b) are shown to have an influence in the final expression levels. Expression of pooled mRNA is finally controlled by mRNA-binding factors (Wurst *et al.*, 2012) .

Tb927.10.14300 (STE11/MRK1) has sequence similarity with members of the yeast High Osmolarity Glycerol (HOG) pathway, which responds to osmotic challenge through actin recovery (Zhi *et al.*, 2013), while the MORN motif-containing protein Tb927.10.10350 (STE11/Bck1p) resembles human MEKK3, a positive regulator of the stress-activated protein kinase (SAPK) pathway (Ellinger-Ziegelbauer *et al.*, 1997) putatively involved in cell integrity maintenance under stress conditions. The increased susceptibility to osmotic shock seen after RNAi induction of these two STE11 protein kinases when subjected to a modified ´Swell dialysis´(Voorheis and Martin, 1980) protocol indicates that they may provide similar roles in *T. brucei*.

Two osmosensor systems are described in yeast connected to the reactive MAPK cascade (HOG pathway): a 2-component transduction system, which deactivates the cascade in low-osmolarity conditions, and an integral membrane protein that activates the pathway (reviewed by Widmann et al. 1999). Interestingly, the 8 Membrane Occupation and Recognition Nexus (MORN) motifs present in Tb927.10.10350 (STE11/Bck1p) may suggest membrane association (Lee, Han and Hur, 2010), but this needs to be studied in more detail. MORN motifs are a highly conserved 14-amino acid sequences found in many eukaryotes. The number of repeats and their position varies in the different proteins. They have been linked to lipid interaction and subcellular localization (Ma et al. 2006), but functions of these motifs are still unclear in the literature.

Immediate downstream interaction partners of STE11s (STE7s) were not identified in this screen, which may suggest redundancy amongst the *T. brucei* STE7 cohort, or perhaps a non-canonical mode of action for the *T. brucei* STE11 protein kinases. This is likely, as transcriptional regulation in trypanosomes is different from other eukaryotes. Canonical MAPK signalling modules are also missing in other parasites, such as *Encephalitozoon cuniculi,* or the complete Apicomplexa phylum (Miranda-Saavedra *et al.*, 2007).

Two other MAPKs (ERK-like), often involved in transcriptional and non-transcriptional regulation in response to external stimuli, were found to be essential both *in vivo* and *in vitro*: KFR1, previously reported to be regulated by interferon γ in order to promote proliferation in *T. brucei* (Hua and Wang, 1997); and MAPK6 (ERK8), which regulates cytokinesis (Mackey *et al.*, 2011).

Interestingly, within the first 24 h of the *in vitro* RITseq (Table 3-3), 5 protein kinase lines were identified with a relative loss of fitness and 4 of them were STE11s. If this is the product of stochastic fluctuation, or indicates common functions of *T. brucei* STE11s, remains to be investigated.

Three protein kinases recovered later and two, a CDC15-like protein (together with CRK3), maintained the loss of fitness for the whole 120 h of the screen. CDC15 is an essential component of the actomyosin ring of yeast, and required for spindle pole separation during the anaphase of cell division (Ren *et al.*, 2014). It has an F-BAR domain, which links membranes to the actin cytoskeleton and has a strong cell cycle-dependent phospho-regulation mediated by Pom1, a DYRK kinase (Ullal *et al.*, 2015). Interestingly, knockdown of CRK3, the other protein kinase found steadily misrepresented in the induced population with respect to the control along the 120 h of the *in vitro* experiment, has been related to G2/M arrest (Tu and Wang, 2004), so they could be involved in the same pathway.

One of the three STE11 that displayed recovery, however, was RDK1 and shows again a loss of fitness after 96 h that increases at 120 h. RDK1 has been reported to cause differentiation to procyclic forms after 72 h of RNAi (Jones *et al.*, 2014). Inducing anachronistic differentiation to procyclic forms in the bloodstream entails VSG loss and has been proposed a novel therapeutic strategy (Wenzler *et al.*, 2016). This rationale underpins also the compound screening presented in chapter 5 of this thesis.

3.4.2.3 Three stress-related protein kinases linked to the PI3K/TOR pathway

Orthologues of three other *in vivo*-only essential protein kinases cluster together within the PI3K/TOR pathway: Tb927.11.9190 (VPS15), Tb927.11.850 (Bud32) and Tb927.7.960 (SRPK1). Tb927.11.9190 (VPS15), is required for stress-induced and developmentally-triggered autophagosome formation in other eukaryotes (Liu *et al.*, 2014; Anding and Baehrecke, 2015) and assembles within the PI3K complexes (Abraham, 2004; Araki *et al.*, 2013). VPS15 activates VPS34, a PI3K in yeast. Depletion of the *T. brucei* orthologue, TbVPS34 (which could potentially be mimicked by lack of activation if VPS15 is missing), impairs correct segregation of the Golgi apparatus during cytokinesis, receptor-mediated

endocytosis, transport to the lysosome, and VSG export to the membrane (Hall *et al.*, 2006). Examining these pathways after VPS15 depletion would be good follow up work.

The PI3K signalling pathway in yeast involves also Bud32 (orthologues called PRPKs in humans and *Drosophila*), which is part of the KEOPS/EKC complexes. There, it acts as a transducer for TOR activation that, in turn, has also been linked to autophagy and endocytosis under stress conditions (Ibar *et al.*, 2013; Rojas-Benítez, Ibar and Glavic, 2013). Through this PI3K/TOR pathway, Bud32 has been linked to translational regulation throughout tRNA modification (Srinivasan *et al.*, 2011; Rojas-Benítez, Ibar and Glavic, 2013). Finally, Tb927.7.960 (SRPK1), regulated via PI3K/AKT pathway (in parallel to TOR), is responsible for activation/inhibition of splicing factors through phosphorylation (Zhou *et al.*, 2012). Localization and availability is also a stress response.

Identification of these TOR-related protein kinases is of interest, because TOR pathways are central for parasite-host interaction in *T. brucei* (Saldivia *et al.*, 2013). While most eukaryotes encode 2 TOR paralogues, trypanosomatids are the only known organisms containing four, each integrated in a different protein complex, of which three are required for normal fitness both *in vivo* and *in vitro* (here and Jones *et al.*, 2014). TbTOR1 regulates transcription and translation. Its depletion leads to the induction of autophagy, thereby enhancing survival upon nutritional stress, which is linked to developmental progression (Barquilla and Navarro, 2009a). This material recycling process enables production of new organelles and has also been observed in *T. cruzi* and *Leishmania* (Figueiredo, Rosa and Soares, 2000; Besteiro *et al.*, 2006; Alvarez *et al.*, 2008; Duszenko *et al.*, 2011). TbTOR2 is required for actin polarization, which in turn is essential for secretion and endocytosis (Barquilla and Navarro, 2009a, 2009b). TbTOR3 responds to osmotic shock by control of polyphosphates and acidocalcisomes (De Jesus *et al.*, 2010), and in *Leishmania* has implications for *in vivo* infection (Madeira da Silva and Beverley, 2010). The acidocalcisome is crucial for autophagy in *T. brucei* (Li and He, 2014). The last paralogue, TbTOR4, is required for *T. brucei* proliferation and life stage maintenance in bloodstream forms as its depletion triggers differentiation from replicative slender to stumpy cell-cycle-arrested forms (Barquilla *et al.*, 2012). How the three *in vivo*-only

protein kinases identified here might act within the expanded TOR signaling network awaits further analysis.

3.4.2.4 A pleiotropic, stress-responsive protein kinase: CK2A2

In the set of *in vivo*-only protein kinases, CK2A2 deserves a special mention as it is the only one that has been previously studied in *T. brucei* (Park *et al.*, 2002; Jensen *et al.*, 2007). This protein has a proven protein kinase activity and localizes mainly to the nucleolus, although remains ubiquitous throughout the cell. With more than 300 known substrates in humans (Meggio and Pinna, 2003), regulation of eukaryotic CK2 is mainly controlled by distribution and abundance (Pinna, 2003), with stress-induced mobilization extensively reported (Gerber *et al.*, 2000; Davis *et al.*, 2002; Yamane and Kinsella, 2005). Downregulation of CK2 has also been implicated in the control of autophagy through intersection with the PI3K/AKT/TOR pathway in eukaryotes (Olsen, Svenstrup and Guerra, 2012; Hales, Taub and Matherly, 2014; Sanchez-Casalongue *et al.*, 2015). Interestingly, CK2 signalling also intersects in humans with mitogen and stress-activated protein kinases (Shi *et al.*, 2009; Jacks and Koch, 2010), and has been implicated as a key kinase controlling transcription elongation from yeast to mammals (Basnet *et al.*, 2014).

Regulation of VSG mono-allelic expression is not well understood yet, but there is evidence for epigenetic control affecting transcription conducted by RNA polymerase I (Günzl *et al.*, 2015). Mammalian CK2A2 has been related with regulation of RNA polymerase I activity in the literature before (Panova *et al.*, 2006; Bierhoff *et al.*, 2008).

There is also evidence in the literature of inducible expression of a different VSG causing DOT1B-mediated histone methylation that attenuates the active expression site (Batram *et al.*, 2014). When this happens, which can also be forced by depletion of ESAG genes, the cell acquires a cell cycle-arrested and PAD-expressing stumpy-like phenotype. This state reverts to normal slender VSG221-expressing cell morphology when the secondary VSG stops being induced. Among its many functions, CK2A2 has been linked to epigenetic regulation in many eukaryotes. Such membrane coat disruption, together with

the associated temporary cell cycle arrest, could be an explanation for the exacerbated loss of fitness *in vivo* compared to *in vitro*.

Casein kinase 2 was one of the earliest described eukaryotic protein kinases. It does not cluster within any of the main families of protein kinases but it is distantly related with the CMGC family (Burnett and Kennedy, 1954). It is present mainly as a tetrapeptide with 2 α (catalytic) subunits and 2 β (regulatory) subunits. The yeast counterpart is constitutively active and the β subunit only modifies its interaction with other partners. *T. brucei* has 2 isotypes for the catalytic subunit (there is 43% identity between them, a lot lower than between their yeast or human orthologues) (Park *et al.*, 2002). In contrast with CK2A1 (aka catalytic subunit CK2α), attempts carried out to characterise CK2A2 (aka catalytic subunit CK2α') gave mixed results. Studies based on *in vitro* translation could not prove CK2α' interaction with CK2β, probably due to misfolding, as it lacked kinase activity. However, kinase activity (with casein as a substrate) was proven in CK2α', which was obtained by tandem affinity purification. Unfortunately, low concentration precluded identification of interaction partners over background contaminants (Jensen *et al.*, 2007). RNAi lines targeting each of the isoforms showed a slight growth defect, visible after 120 h. Difficulties in obtaining the purified CK2A2 stymied further analysis of this protein, but the degree of loss of fitness unveiled here, leading to parasite clearance in the mouse bloodstream, suggests it can be a promising drug target. There is at least one compound in the GSK HAT box targeting the closely related human CK2A1 (Peña *et al.*, 2015), which may be active against both TbCK2A2 (whose effect would be evident mainly *in vivo*) and TbCK2A1 (*in vivo* and *in vitro*), therefore having an effect *in vivo* greater than *in vitro*.

3.4.2.5 Three protein kinases engaged with mechanisms of resistance to DNA damage

It is intriguing that three protein kinases (AUK2, CAMKL and SRPK1) only essential for *in vivo* growth have also been identified in a companion RNAi screen for factors required for *T. brucei* to withstand or repair damage caused by DNA alkylation *in vitro* (Stortz, 2016). In what way, if at all, the two phenotypes might overlap is unclear, but potentially these *T. brucei* protein kinases, associated with serum resistance (Figure 3-15), may be involved in

aspects of genome maintenance. Confirming such hypothesis would require further investigation. From the four tested in Figure 3-15, only the one causing a clear loss of fitness *in vitro* (DYRK) was unaffected for serum exposure upon RNAi. Overlap between death caused by serum exposure and death caused by RNAi would be a reason for not finding a difference between serum treated and non-treated controls in DYRK. However, relative growth observed in the Tet+ samples as serum concentration increased discarded this as a cause. Verification of serum resistance as a common theme in other *in vivo*-specific candidates from our screen still remains to be done.

AUK2 resembles AUKA, which in other eukaryotes complexes with the spindle poles and regulates entry in mitosis. It is often overexpressed in human cancers (D'Assoro, Haddad and Galanis, 2015). Unlike yeast, which only has one AUK, or plants which have two, *T. brucei* (like mammals) contains three members of this important family of protein kinases involved in cell cycle regulation (Carmena, Ruchaud and Earnshaw, 2009).

CAMKL has great identity with AMPKs in other eukaryotes but it also bears two calcium binding domains (EF hands). AMPKs have functions in maintaining the balance of energy through phosphorylation of metabolic enzymes and transcription factors (Salminen A, Kauppinen A and Kaarniranta K, 2016).

SRPK1 has been discussed above, with Bud32 and VPS15, in the set of protein kinases putatively linked to the PI3K pathway. It had orthologues regulating gene expression by phosphoryating SR-motive-containing splicing factors in response to stress.

Interestingly, the fourth gene we analysed for serum resistance, Tb927.7.3880 (DYRK), has recently been co-purified with several SR-rich proteins in a pull-down of CRK9, a protein kinase critical for RNA splicing (Badjatia *et al.*, 2016). One of these SR proteins, CYC12, is an L-type cyclin that forms a complex with CRK9 and has a human orthologue phosphorylated by DYRK1A (Graaf *et al.*, 2004). This suggests that a similar pathway can be in place for *T. brucei*, and SRPK1 might be involved in the network.Tb927.7.3880 (DYRK) was in the original set of *in vivo* loss of fitness not reported *in vitro* by Jones et al. (2014) but upon validation showed a growth defect (Figure 3-14A). It is a dual specificity kinase

predicted to require tyrosine autophosphorylation in order to activate serine-threonine kinase activity. This mechanism suggests that they may have also dual sensitivity to protein kinase inhibitors (Lochhead *et al.*, 2005).In yeast, DYRKs have functions in activation of cytokinesis and growth by mediating interaction of their cellular machineries with actin and microtubules (Bähler and Nurse, 2001). A role for DYRKs regulating cytokinesis has been also detected in *C. elegans* during embryogenesis (Pang *et al.*, 2004) and in neurogenesis in more complex metazoans, including *Drosophila* (Tejedor *et al.*, 1995) and mammals (Guimera *et al.*, 1999; Fotaki *et al.*, 2002).

3.4.2.6 A PLK-like protein kinase without polo-box domains

One protein kinase with an induced/uninduced median ratio under 0.6 detected only *in vivo* was missed in the original analysis because ratios in the *in vivo* Replicate 1 were not calculated with the bootstrapping approach at first. Tb927.6.5100 has homology with polo-like kinases in other eukaryotes, but lacks the characteristic polo box domains at the C-terminal end of the protein that allosterically inhibit the kinase domain until binding the substrate (Elia *et al.*, 2003). We do not have validation data for this loss of fitness phenotype, so it will not be discussed further.

3.4.3 Potential of 49 protein kinases with *in vivo* value as drug targets

T. brucei protein kinases can be targeted with chemical inhibitors in the low micromolar and nanomolar range of concentration (Urbaniak *et al.*, 2012). Despite conservation of the protein kinase catalytic domain, protein kinase inhibitors can be designed for selectivity and have been proven good drug candidates (Cohen, 2002; Weinmann and Metternich, 2005; Cohen and Tcherpakov, 2010). This work identifies 49 protein kinases that are required for *in vivo* fitness. For those causing strong phenotypes, leading to parasite clearance (Table 3-3), it may provide evidence of their potential as drug targets, though a more thorough genetic validation is required, as lethal phenotypes observed upon protein kinase depletion via RNAi are not necessarily due only to protein kinase inhibition. For instance, the absence of the protein in a given

interaction network can cause harm by itself, something that could not be reproduced with a chemical inhibitor.

One method to analyse this would be integration into the RNAi line of a recoded version of the targeted gene. The resulting transcript would be unharmed by the RNAi machinery, while producing the original protein. If this recovers the wild type phenotype, we would confirm that RNAi is on target. If, on the other hand, cells are unable to recover when the recoded gene has key mutations for activity, e.g. substitution of the catalytic lysine in the second subdomain (Hanks and Hunter, 1995), it would confirm that the same phenotype could be obtained with a targeted kinase inhibitor. One protein kinase in the set, RDK2, has been targeted with this approach and results are described in the next chapter.

Another important consideration is the timing and severity of the outcome. Upon chemical inhibition, the derived phenotype can happen immediately after addition of the compound. RNAi requires transcription machinery to be active and, as transcript depletion avoids production of fresh protein, time is required to deplete the cell of available stocks. To examine transcript and protein levels with qRT-PCR and western blot in order to understand the timing is critical.

In addition to the obvious possibility of designing protein kinase-specific screens targeting them individually, some protein kinases identified here open an opportunity to develop drug-repurposing strategies. A phylogenetic comparative between the *T. brucei* and *Homo sapiens* kinomes showed how 13 out of the 49 protein kinases found to be important *in vivo* had close human orthologues targeted by at least one of the compounds in the 'HAT box' (Peña *et al.*, 2015). Interestingly, among them is Tb927.10.10350 (STE11/Bck1p), from the *in vivo* only set. These molecules could be chemically modified to achieve *T. brucei* specificity. This has been also suggested for aurora kinases (Patel *et al.*, 2014) and PI3K/TOR kinase inhibitors (Diaz-Gonzalez *et al.*, 2011). Both pathways have putative components detected also in the *in vivo*-only set, Table 3-6.In contrast, the other 8 *in vivo* validated protein kinases are highly divergent from any human kinase. Three 'orphan' kinases in this set are very likely to play parasite-specific functions. Selectively targeting these may be easier than other more conserved protein kinases.

3.4.3.1 *Leishmania/Trypanosoma cruzi* extrapolation

Most of these 49 protein kinases have close orthologues in *Trypanosoma cruzi* and *Leishmania major* (see Figure 3-13/Table 3-5). Extrapolation of results based on gene ontology using *T. brucei* as a model may be a valuable perspective to identify potential drug targets, since to date these organisms lack efficient reverse genetics tools for kinome- or genome-wide studies. Examples of protein kinases with a loss of fitness in our study, which also play essential roles for survival of *Leishmania* include MPK1 (called KFR1 in *T. brucei* (Wiese, 1998), CK1 (Knockaert *et al.*, 2000; Rachidi *et al.*, 2014), CRK3 (Hassan *et al.*, 2001; Walker *et al.*, 2011) and GSK3-short (Xingi *et al.*, 2009; Ojo *et al.*, 2011). The last three are genetically validated targets and have had HTS screens where inhibitors have been identified. In addition, one of the MAP3Ks observed to have an only *in vivo* fitness defect, MRK1 (with a role in osmotic resistance), has a *L. major* orthologue proven essential for promastigote survival (Agron, Reed and Engel, 2005). All the main hits causing a loss of fitness in Table 3-3 are potentially worthy of further consideration in this manner. For example, the *L. donovani* orthologue of AUK1 (LdAIRK) has been tested showing implications in cell cycle but requires further analysis (Chhajer *et al.*, 2016). KKT10, the kinetoplastid-specific kinetochore member, may be targeted with specificity, like CK2A1 and CK2A2,with human orthologues targeted by compounds with trypanocidal activity in the HATbox (Peña *et al.*, 2015), suggesting potential for drug repurposing. A CK2-targetting compound may have enhanced *in vivo* efficacy, according to results shown in this chapter. Supporting its potential, a CK2 orthologue in *L. braziliensis* has been shown to be secreted, helping macrophage association prior to invasion (Zylbersztejn *et al.*, 2015). Others, like DYRK1, CRK9, AUK3, or RDK2, which are highly divergent from any human orthologues according to a phylogenetic comparative made by Peña et al (2015), may be interesting kinetoplastid-specific targets too, Table 3-5.

3.4.4 Potential applications of RITseq used in other pre-existing libraries

With this study we prove that RITseq can be applied to any pre-existing RNAi library, expanding our capacity to search for genes involved in specific aspects of *T. brucei* biology for which a screen can be devised. Reduced library size may

be an advantage by increasing resolution and requiring less starting material in order to obtain significant results, as we have shown by *in vivo* infections. The barcoded primer approach reduced costs and increased our multiplexing capacity. Possibilities are endless, but we have already used this library to identify protein kinases required for kinetoplast maintenance, cell cycle regulation, withstanding or repairing alkylation-mediated DNA damage, and differentiation of bloodstream to procyclic forms.

Technically it would be possible to use this library to find the target of a kinase inhibitor whose time to kill at a given concentration and EC50 is well profiled. We should be able to find a time frame where a synergistic effect due to overlap of kinase inhibition and kinase depletion can produce a reduction the IC50. The population, after a certain incubation time, would be enriched in the parasites that do not target with RNAi the liable protein kinase.

The library could also be used for the identification of the protein kinase responsible for a particular phosphorylation event, if a specific antibody is available against the phosphorylated peptide. The library, induced and uninduced, would be fixed before staining with such antibody, which could be directly or indirectly coupled to a fluorophore. Then, cells would be processed through fluorescence-activated cell sorting (FACS) so the non-phosphorylated (and non-stained) population could be isolated. Sequencing the RNAi targets of the unstained population would show which kinase depletions were responsible for the lack of phosphorylation. The main challenge would be choosing the right timing in case the responsible kinase caused a strong loss of fitness. Induction should be sustained long enough for the protein kinase to be depleted so the phosphorylation defect becomes evident. However, it cannot be so long that the cell line of interest is cleared if the protein kinase is essential.

This rationale can be subsequently applied to libraries of genetically modified parasites of many different makes (e.g. overexpression, knock out, fluorescent or luminescent tagging) as soon as we can use universal primers to enrich in a quantifiable molecular label, e.g. a barcode, an RNAi cassette or an overexpression cassette. For instance, an RNAi library made in a bioluminescent pleomorphic parasite could be used to test which proteins are needed to invade the central nervous system by infecting mice. We could track the progress of the

parasites throughout the host organism to certify invasion of the central nervous system with an *in vivo* imaging system (IVIS), before purifying DNA from the brain tissue or haematoencephalic fluid. Genomic DNA would be enriched in the cassette and sequencing reads assigned to it filtered *in silico*. No protein kinase has been reported to date as specifically required to cross the blood brain barrier (BBB) in the parasite, although this signalling is fundamental in the endothelial cells to permit this to happen. Examples of *T. brucei*-secreted factors proven to have a direct role in BBB crossing are two cathepsin-B-like cysteine-proteases (brucipains). In addition, others have been proposed that may have a role are the closely related oligopeptidase B or acid phosphatases expressed in the surface (Lonsdale-Eccles and Grab, 2002). Focused RNAi libraries targeting protein kinases or peptidases could be used in this manner.

The screen described in this thesis is not foolproof. False negatives in RITseq studies are rare, but can be caused by a failure to produce dsRNA or leaky phenotypes giving a reduced count (due to cell line depletion) both in the induced and uninduced pools (Glover *et al.*, 2014). Starting with a pool of validated cell lines is time-consuming but gives a strategic advantage. False positives can happen due to rare off-target effects. With time, loss in complexity in the induced sample could cause unequal expansion of different clones in the library, due to uneven access to energy resources. Deletion of the RNAi target from the genome after induction has been also detected.

3.4.5 Final conclusion

All together, we showed in this chapter that application of next generation sequencing to pre-existing libraries of RNAi lines enhances greatly our capacity to evaluate the impact of protein depletion on the cell phenotype when subjected to selective pressures of any kind that mimic relevant cellular processes. As a proof of concept we used the library to understand what protein kinases in *T. brucei* are indispensable for mammalian infection, something of interest both for drug discovery and biology. We validated *in vivo* the effect in loss of fitness that had been observed *in vitro* for more than 40 protein kinases. Interestingly, we identified at least 9 which showed a degree of essentiality in the context of a mouse infection that could not be reproduced in culture, therefore behaving as potential virulence factors. Most of these *in vivo*-

indispensable protein kinases showed phylogenetic relationship with orthologues in other organisms that had a role in regulation of stress response pathways, including osmotic shock control, gene expression or autophagy mediated by the PI3K/TOR axis. We demonstrated the contribution of some of these *in vivo*-only kinases towards serum resistance, and tolerance to osmotic shock.

Severity of the phenotype in more than 20 protein kinases in this screen highlights their potential value as drug targets. We discussed how kinase inhibitor chemotypes are generally enriched in the available collections of trypanocydal compounds. All of them had orthologue proteins in *L. major* and *T. brucei*, and they may be also essential in those organisms. Until proper high-content genetic dissection tools, equivalent to RNAi, are developed in these parasites, extrapolation from observations made in *T. brucei* can be a successful approach.

Given the reduced size of the library, this technique provided the advantage -compared to whole-genome approaches- of increasing greatly our resolution in terms of read depth, permitting more compelling results using a smaller number of parasites, as required for mice inoculation. Possibilities of RITseq and similar genetic approaches are endless.

Chapter 4 Repressor of differentiation kinase 2

4.1 Introduction

When the kinome-wide RNAi library of *T. brucei* bloodstream forms was screened for regulators of differentiation, two protein kinases were found that upon depletion triggered spontaneous differentiation from bloodstream trypomastigote to procyclic-like forms. This differentiation event was characterised by cell elongation, kinetoplast repositioning and EP procyclin expression (Jones *et al.*, 2014). These two protein kinases were named RDK1 (Tb927.11.14070) and RDK2 (Tb927.4.5310), denoting Repressor of Differentiation Kinase as they seem to play a role in bloodstream stage maintenance. Interestingly, depletion of RDK2 by RNAi knockdown led to an important loss of fitness both *in vitro* (Jones *et al.*, 2014) and *in vivo*, as described in the RITseq experiment reported in Chapter 3.

4.1.1 Repressor of differentiation kinases 1 and 2

RDK1 (STE11-like/Tb927.11.14070) is a putative MAP3K: the first of a three-member cascade of protein kinases (MAP3K>MAP2K>MAPK) responsible for transcription activation in response to extracellular signals in many eukaryotes. RDK1 has 2-3 predicted N-terminal transmembrane domains, which appear functional as the protein localises to the membrane fraction of parasite lysates and is detected by IFA in the whole cell body, including the flagellum (Oberholzer *et al.*, 2011; Jones *et al.*, 2014). Although induction of RNAi does not lead to cell death, procyclic form differentiation (entailing VSG loss) would make the parasites susceptible to complement-mediated lysis (Ferrante and Allison, 1983; Wenzler *et al.*, 2016), suggesting RDK1 inhibition may be a potential therapeutic strategy. After 72 h of RNAi in standard bloodstream form culture conditions, 20% of cells expressed EP procyclin on the surface. RNAseq data of parasites at this time point post-RNAi revealed that 236 genes (out of 275 differentially transcribed if compared to non-induced controls) matched mRNA profiles of procyclic forms, even though only one fifth of the population expressed the phenotype.

The progenitor *T. brucei* line used in RDK1 RNAi analysis was 2T1, which is derived from Lister 427 and is considered monomorphic, as long term *in vitro*/*in vivo* passage -carried out before the development of effective cryopreservation methods- rendered cells unable to produce stumpy forms in response to high parasite densities (Rico *et al.*, 2013). Suspicion existed that the 20% of procyclic-like forms obtained upon RNAi induction would match the reduced differentiation-competent subset of the monomorphic culture, also known as stumpy* forms (Alsford and Horn, 2008; Jones *et al.*, 2014). This hypothesis is supported by observation of a 2-fold increase in the number of EP-procyclin-positive cells after RDK1 depletion in cells co-treated with 8-pCPT-cAMP. As with other cell-permeable hydrolysable analogues of cAMP, 8-pCPT-cAMP induces differentiation to stumpy* forms in monomorphic cultures (mimicking the high density-mediated effect observed in pleomorphic forms) (Reuner *et al.*, 1997; Laxman *et al.*, 2006). 8-pCPT-cAMP also increased transcription of *PAD1,* a stumpy stage specific marker (Dean *et al.*, 2009), 2-fold in 2T1 *T. brucei*. In addition, differentiation triggered by RNAi of RDK1 was enhanced, generating 60% EP positive cells, by treating the induced culture with a TbPTP1 inhibitor (BZ3); this additive effect has been proposed to be a response to two parallel pathways affecting two different subsets of cells (Jones *et al.*, 2014).

RDK2 (NEK12.2/Tb927.4.5310) belongs to the NEK family of protein kinases, which is expanded in *T. brucei* with respect to humans (Parsons *et al.*, 2005). The expanded NEK family may have parasite-specific functions, with RDK2 providing a clear example since it seems to regulate life-stage maintenance. In addition to triggering differentiation, knock down of RDK2 causes cell cycle arrest, which is associated with a block in cytokinesis (after 48h of RNAi induction), followed by cell death (after 72h) (Jones *et al.*, 2014).

RDK2 (NEK12.2) has 87% sequence identity with NEK12.1 (Tb927.8.7110), Figure 4-1. The two genes may be the evolutionary product of a duplication event that has been reported for the 3' regions of chromosomes 4 and 8 (Jackson, 2007a). Unlike the majority of protein kinases, NEK12.1 has a smaller gate-keeper residue in the ATP binding pocket (alanine instead of methionine), making the pocket putatively larger. This variation may allow substrate identification, by labelling the γ-phosphate of large ATP analogues and determining the substrates

by tag detection, or protein kinase inhibition, with bulky analogues that would compete with ATP for the binding pocket of NEK12.1 only (Bishop *et al.*, 2000; Elphick *et al.*, 2007). The original cell line showing differentiation to procyclic-like forms prior to cell death upon RNAi induction targeted an identical section of both genes (RDK2 and NEK12.1), causing a double knockdown, so specific RNAi constructs were designed by joining small divergent fragments longer than 20 nucleotides (Jones *et al.*, 2014). Only the RDK2-specific RNAi line showed cell elongation, kinetoplast repositioning and EP procyclin expression phenotypes characteristic of the dual RNAi line. Transcription analysis (quantitative RT-PCR) revealed that NEK12.1 was independently downregulated with its specific RNAi while the original construct affected both genes. It has been suggested (Jones *et al.*, 2014) that reduced expression of NEK12.1 mRNA may be a natural event that occurs during BSF-PCF differentiation (Jensen *et al.*, 2009; Kolev *et al.*, 2010).

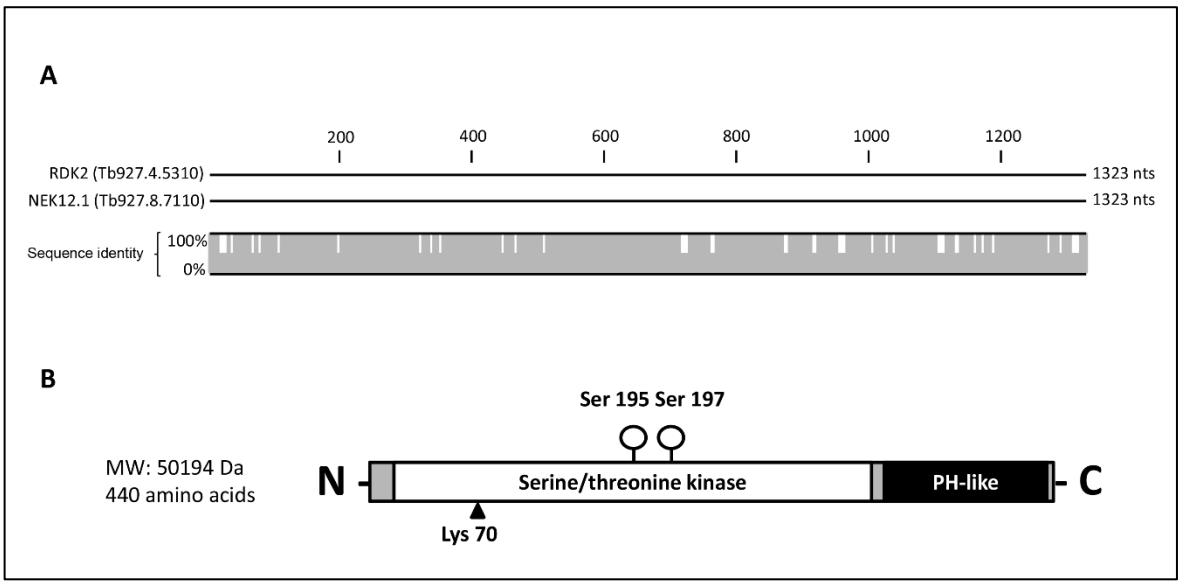


Figure 4-1. Nucleotide sequence identity of *RDK2* and *NEK12.1*, and predicted protein structure.
A. Alignment between *RDK2* and *NEK12.1* performed in CLC genomics workbench. Sequence identity between the two is displayed, highlighting divergent regions. **B.** Schematic representation (to scale) of RDK2 functional features, according to an alignment with three different hidden markov model databases: SMART (Letunic, Doerks and Bork, 2015), PFAM (Finn *et al.*, 2014) and Superfamily (Gough *et al.*, 2001).

RNAi lines targeting RDK2 (Jones *et al.*, 2014), showed a reduction in transcript abundance after 24 h of RNAi induction to 20% of the uninduced line levels. In that study, an alamar blue fluorescence induced/uninduced ratio of 0.676 was recorded after 72 h of induction, indicating a loss of fitness (Jones *et al.*, 2014).

RDK2 essentiality was detected in Alsford et al's whole genome screen at days 3 and 6 and there, its presence was not required for procyclic form survival or for differentiation (Alsford *et al.*, 2011). In the kinome-wide RITseq presented in this thesis (chapter 3), RDK2 depletion caused a strong loss of fitness both *in vivo* and *in vitro*, from 48 h onwards as indicated by median ratios between induced and uninduced normalized reads under the 0.6 established threshold. The intensity and significance of the observed loss of fitness during the *in vitro* RITseq increased with time (Table 3-3. Summary of RNAi target sequencing results *in vivo* and *in vitro*).

RDK2 RNAi induction in the presence of 8-pCPT-cAMP, cold shock or BZ3 did not cause differentiation of more cells than the induced RNAi line alone. However, the proportion of cells differentiated by RDK2 knockdown was equivalent to the proportion obtained by RDK1 depletion empowered by these treatments. This suggests that RDK2 blocks differentiation of bloodstream to procyclic forms at least upstream of RDK1.

The *RDK2* gene putatively encodes a polypeptide of 440 amino acids (Figure 4-1B), with the protein having a predicted molecular weight of 50.2 KDa. Alignment against three different hidden Markov model databases –SMART (Letunic, Doerks and Bork, 2015), PFAM (Finn *et al.*, 2014) and Superfamily (Gough *et al.*, 2001)- reveals an N-terminal serine-threonine protein kinase domain (41-299) and a C-terminal PH-like domain (348-437). Two phosphoserines (S195/S197) have been reported in a phosphoproteomics screen of the bloodstream form cytosolic fraction (Nett *et al.*, 2009). The key lysine responsible for γ-phosphate orientation, whose mutation generally ablates kinase activity, is residue Lys 70 (Figure 4-1**).**

4.2 Research aims

With the objective of characterizing further a potential drug target among the essential protein kinases identified in the RITseq screen presented in Chapter 3, RDK2 was selected, since the RNAi phenotype encompassed both cell death and differentiation, suggesting an interesting biological role in the parasite. This chapter provides a mechanistic characterization of RDK2 and investigates in more detail the role it plays in differentiation.

4.3 Results

4.3.1 RDK2 RNAi phenotype description

For all analyses we used the RDK2-targeting RNAi line used in the RITseq screen described in chapter 3, sTL629 (Table S1). It was created by transfection of 2T1 bloodstream forms (Alsford *et al.*, 2005) with *Asc*I-linearized pTL247, a plasmid derived from pGL2084 (Jones *et al.*, 2014), as shown in Figure 3-1. The RNAi insert was synthetized by combining divergent fragments of *RDK2*, longer than 20 nucleotides, which were not found in *NEK12.1*. An *in silico* simulation with CLC Genomics Workbench, suggested that the hairpin produced upon RDK2 stem loop transcription is highly favoured energetically (ΔG = -1391.8 kcal/mol) (Figure 4-2A). RNAi knockdown of *RDK2* RNA was validated by qRT-PCR from cells taken 24 h after RNAi induction. A growth curve showed growth arrest after 48 h, and cell death from 72 h onwards (Figure 4-2B). Analysis by flow cytometry and IFA permitted identification of the procyclic-like phenotype described by Jones et al (2014): from 36 h cells showed differentiation, as defined by elongation of the cell body, kinetoplast repositioning and EP procyclin expression, confirmed in the surface by indirect IFA without membrane permeabilization (Figure 4-2C-D).

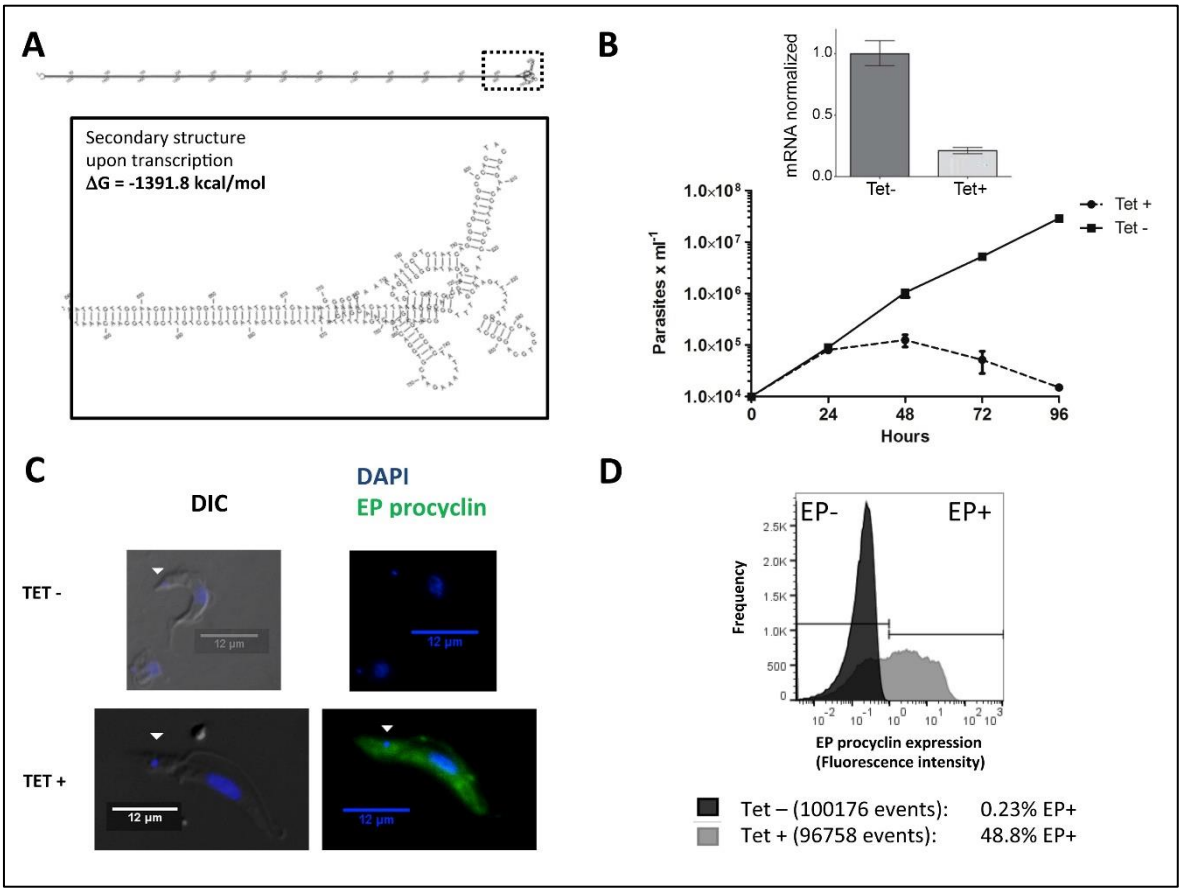


Figure 4-2. RNAi phenotype of RDK2 RNAi.
A. Highly favoured double stranded RNA hairpin predicted for the *RDK2*-targeting stem loop RNAi construct (pTL246) upon transcription. Secondary structure calculated with CLC Genomics Workbench. **B.** Triplicate growth curve starting at 10^4 parasites x ml^{-1} of sTL629 (*RDK2*-targeting RNAi line) comparing RNAi induced (Tet+) cells to uninduced controls (Tet-). The controls were diluted (1:100) after 48h to maintain exponential growth, and the dilution factor (100x) was applied to the graph, generating a cumulative analysis. Embedded qRT-PCR profile of sTL629 24 h post induction, taken from Jones et al. (2014). **C.** Immunofluorescent images of uninduced (Tet-) sTL629 bloodstream forms compared to RNAi induced cells after 48 h of growth (Tet+). On the left hand panel, differential interference contrast microscopy (DIC); on the right hand panel, fluorescence microscopy after staining DNA with DAPI (blue), and visualising EP procyclin using indirect immunofluorescence (green) without cell permeabilization. Arrows indicate kDNA **D.** Flow cytometry profiles of EP procyclin expression for uninduced (Tet-) sTL629 bloodstream forms compared to RNAi induced cells after 36 h of growth (Tet+).

The percentage of cells with EP procyclin expression detected at every time point by IFA were 10% larger than the percentages detected in flow cytometry: for instance, at 36 h, 50% were positive in flow cytometry, compared with 60% in IFA; at 72 h, 70% were positive in flow cytometry and almost 80% EP procyclin positive in IFA, indicating that IFA was more sensitive for detection.

Cell cycle analysis of the RNAi phenotype was in agreement with data published by Jones et al (2014), Figure 4-3A-B. Enrichment in mitotic cells containing two nuclei and two kinetoplasts (2N2K) was indicative of a block or delay in cytokinesis between 36 h and 48 h after RNAi. At these time points there was a reduction of 1N1K and 1N2K cells compared to uninduced cells (or time 0), and an increase in "other" (abnormal) cell types. At 72 h there was an accumulation of 1N1K parasites, consistent with a slightly reduced proportion of abnormal cells and a drastic reduction in 2N2K cells relative to the 48 h time point.

The block/delay in cytokinesis between 36-48 h was coincident with the first manifestation of growth arrest and differentiation, when 60-70% of cells became EP procyclin positive, Figure 4-3A-B. After 72 h of RNAi induction, ~88% of cells were EP procyclin positive and population numbers reduced as enrichment in 1N1K parasites was seen, Figure 4-3A-B.

A more detailed analysis of the phenotype, separating EP positive and EP negative populations, found no substantial difference in cell cycle distribution at 48h, but a slight increase in the 2N2K subset of the undifferentiated population after 72 h (although calculated from a very reduced sample size, n=38), Figure 4-3C-D.

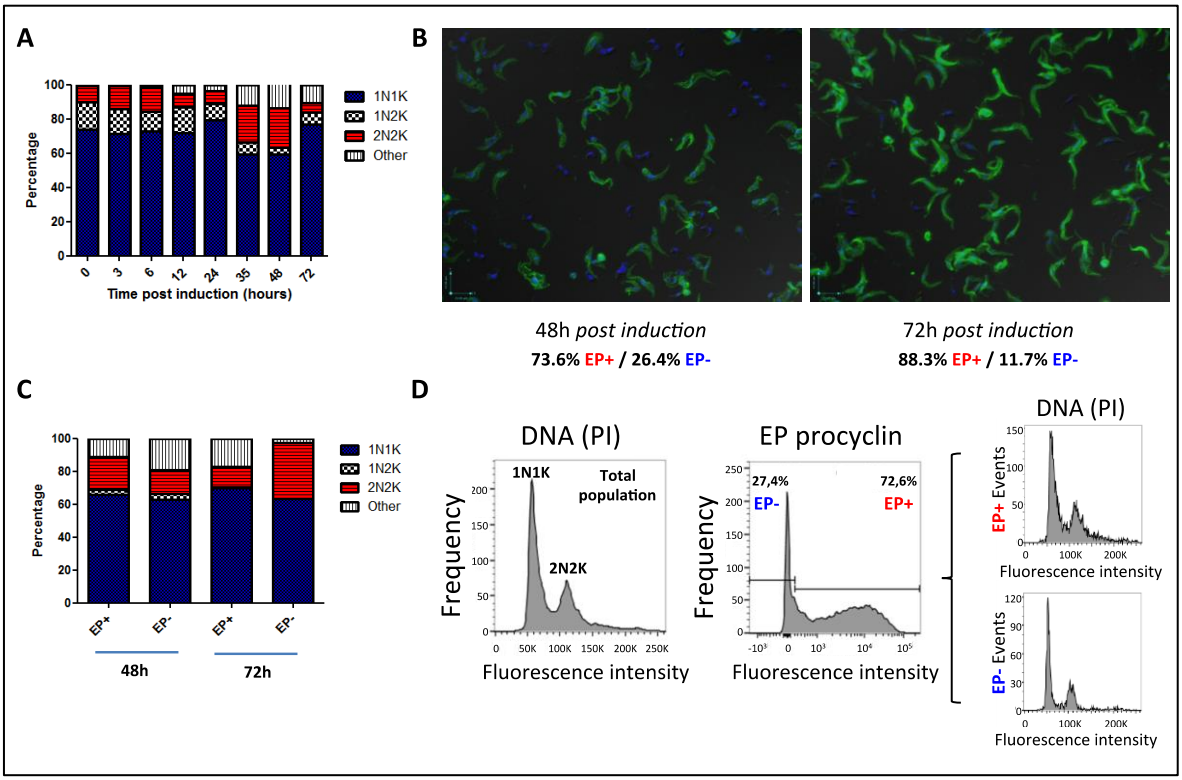


Figure 4-3. Cell cycle analysis after RDK2 RNAi.
A. Bar diagram depicting percentage of cells with increasing time after RNAi induction containing: one nucleus and one kinetoplast (1N1K); one nucleus and two kinetoplasts (1N2K); two nuclei and two kinetoplasts (2N2K); and other aberrant combinations (other). 200 cells were counted per time-point using microscopy where DNA was stained with DAPI. **B.** Microscopic images of cells at 48 h (left) and 72 h (right) post RNAi induction. DNA was labelled with DAPI (blue) and EP procyclin with a specific antibody in indirect immunofluorescence (green). Flourescent images are merged with DIC to distinguish the parasite shape. **C.** Discrete cell cycle analysis for EP positive and EP negative cells using slides imaged in B. **D.** After 48h of induction, flow cytometry histograms assigning events to fluorescence intensity of: DNA (left), using propidium iodide (PI); EP procyclin expression (middle), using a specific antibody in indirect immunofluorescence; and DNA (PI) for individualized cell cycle analysis of EP positive (right above) and EP negative populations (right below).

Optimal growth of procyclic forms in culture occurs between 10^5 and 5 x 10^7 cells x ml^{-1}. In order to explore further the growth phenotype of cells induced to differentiate by RDK2 RNAi, different culturing conditions were tested by collection of parasites 48 h post-induction and concentration adjustment to start a new culture at 1-5 x 10^5 cells ml^{-1}. To ascertain whether the differentiation effect upon RNAi was reversible, different conditions were tested (Figure 4-4A-D).

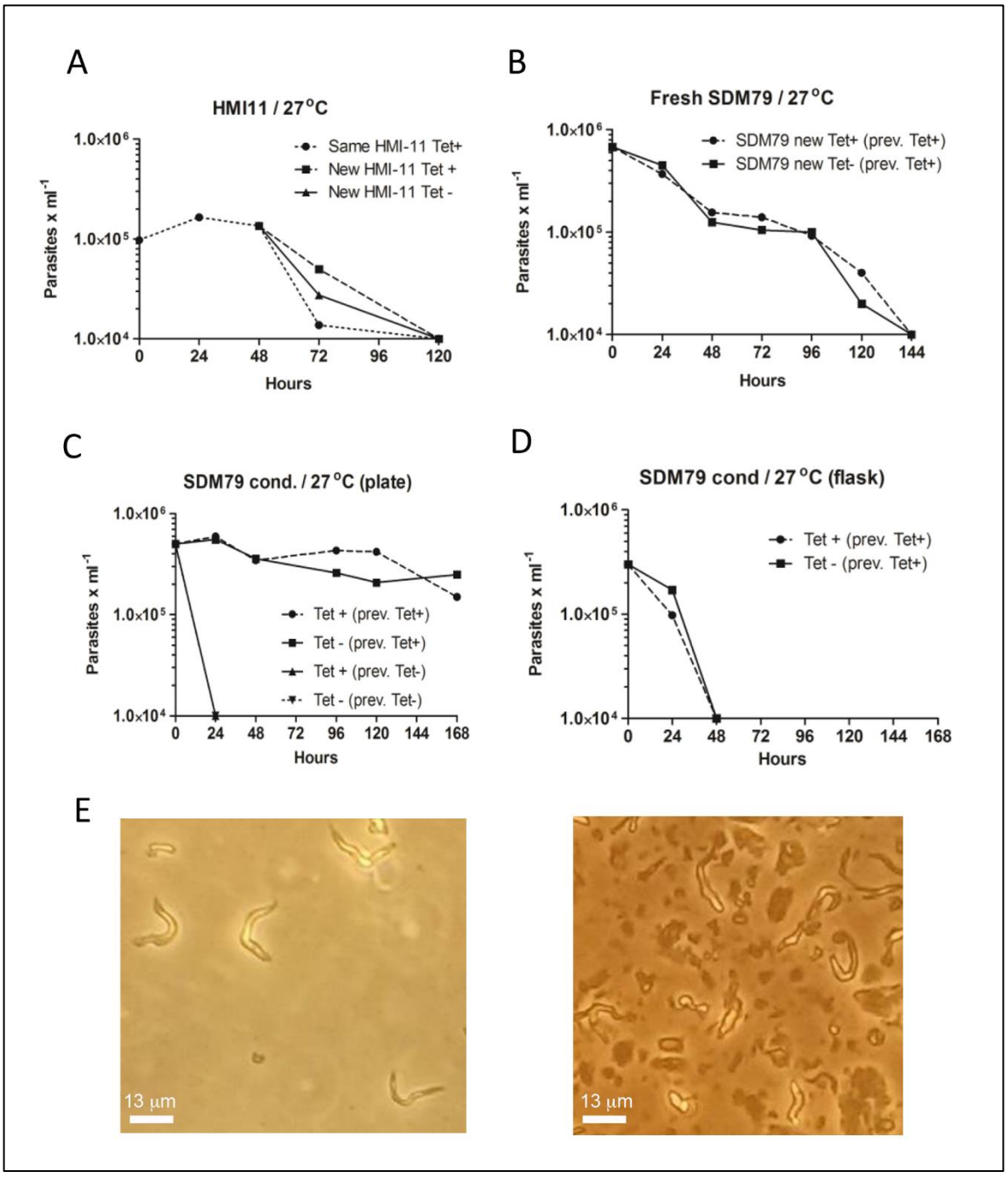


Figure 4-4. Growth properties after RDK2 RNAi.
A. Growth curve of cells recovered 48 h after induction of RNAi with tetracycline (Tet+) and placed at 27 °C in the same medium or in fresh HMI11, with and without (Tet-) renewing tetracycline. **B.** Growth curve of cells recovered 48 h after induction of RNAi with tetracycline and placed at 27°C in fresh SDM79, with and without addition of tetracycline. **C.** Same as (B) but placing the cells in conditioned SDM79 medium and testing parasites that were never RNAi induced (with and without addition of tetracycline). **D.** Cells after 48 h of induction placed at 27 °C in conditioned SDM79 using flasks instead of plates, with and without addition of tetracycline. **E.** Microscopic images of live cells recovered 48 h post-induction and kept in SDM79 conditioned at 27°C with tetracycline for 96 h; videos are provided in attached DVD (Movie_S1 and Movie_S2).

If placed at 27 ^{o}C in HMI-11, cell numbers were maintained for about 48 h (Figure 4-4A). Later, cell numbers decreased until cell death was evident by 72 h. Death was quicker if the cells were kept in the same medium than if moved to fresh. Within fresh medium, those cells under sustained tetracycline induction resisted cell death comparatively better than those where the drug was removed. If placed at 27 ^{o}C but changed to fresh SDM79, the standard medium used to culture procyclic form parasites, cells died slowly over a period of 6 days. In these conditions, no substantial difference was found between maintaining cells in tetracycline or removing it from the medium (Figure 4-4B).

It is well known that procyclic forms are dependent on density-related environmental signals for optimal growth, which in some strains has been circumvented by removal of free radicals (Archer, 2009). When cells were collected after 48 h of RNAi induction and placed in conditioned SDM79 medium (by addition of 15% SDM79 filter sterilized from a dense procyclic culture, see M&Ms), they managed to survive for at least one week, with numbers tending to decrease slowly (Figure 4-4C). Where tetracycline induction was sustained, cells did comparatively better. If non-induced bloodstream forms were exposed to these conditions, cell cultures were dead in 24 h.

Interestingly, all the above experiments were conducted in 6-well plates, where cells appeared to be attached to the bottom (Figure 4-4E). When cells were grown in SDM79 conditioned medium using vertical flasks, cells were not detectable after 48 h (Figure 4-4D).

4.3.2 Expression of recombinant RDK2 and assessment of protein kinase activity

In order to examine whether RDK2 has protein kinase activity, recombinant expression was conducted in *E. coli*. Inspired by successful experiments with the structurally similar NEK12.1 (Jones, 2012), the pET-32 Xa/LIC expression system (Novagen) was selected. This system permits a strong IPTG-inducible expression under control of the T7 promoter, with the protein of interest fused to an N-terminal thioredoxin tag (Trx), which increases solubility (LaVallie *et al.*, 1993) (Figure 4-5A). In addition, the fusion protein contains a six-histidine tag for purification via immobilized metal affinity chromatography (IMAC). A construct

(named pGL2099) was generated by insertion of the *RDK2* coding sequence through ligation-independent cloning. The resulting fusion protein, Trx-RDK2, increased RDK2 size by 17.2 KDa to give a total predicted size of 67.4 KDa (Figure 4-5A). In addition, site directed mutagenesis was conducted in pGL2099 by modification of codon 70 to produce a methionine instead of the catalytic lysine, which has been shown to be critical for the phosphate transference in protein kinases (Kamps and Sefton, 1986). The resulting Trx-RDK2^{K70M} expressing construct was denoted pGL2215 (Figure 4-5A).

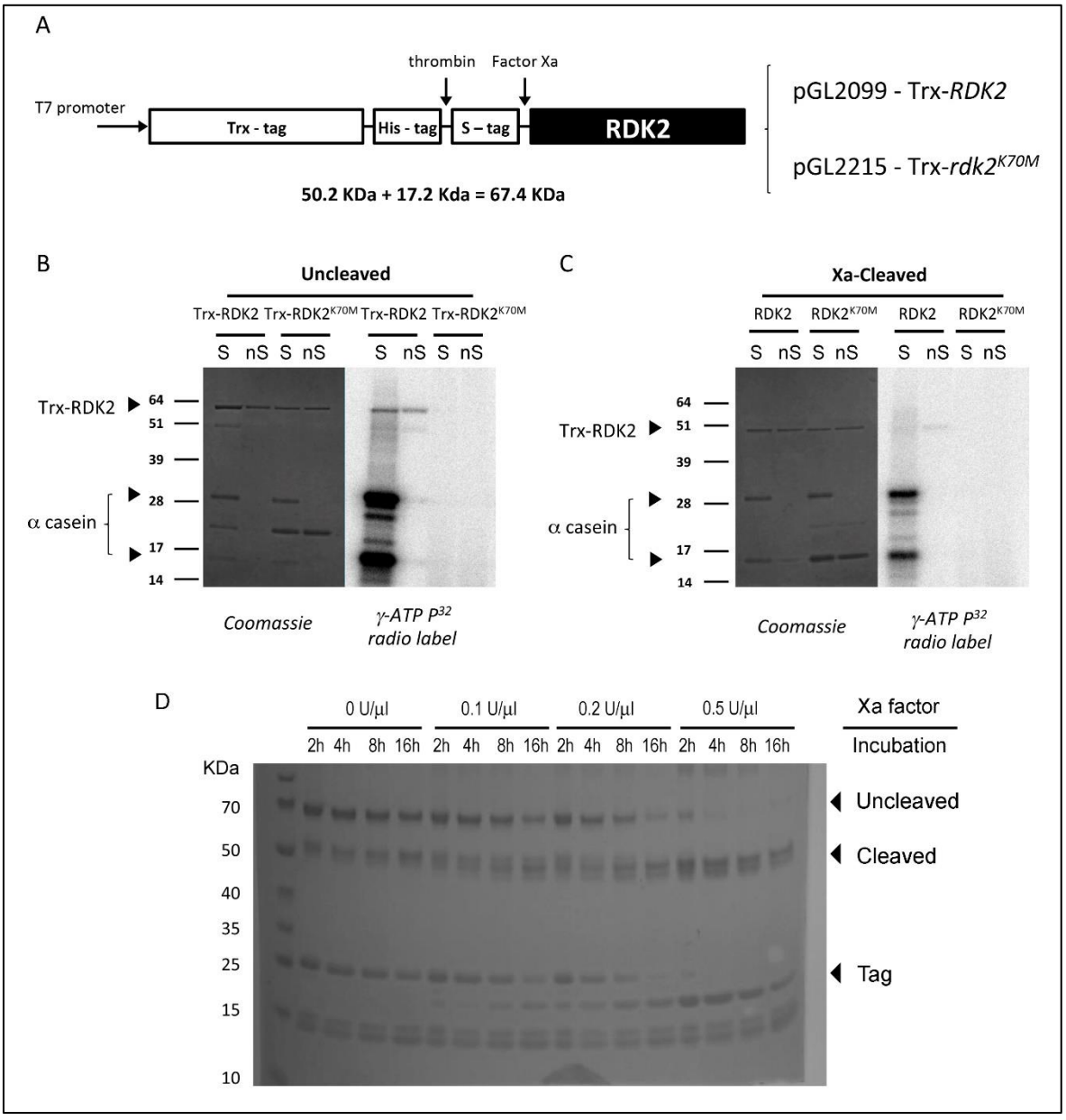


Figure 4-5. Recombinant RDK2 expression and kinase assay.
A. Schematic representation of recombinant Trx-RDK2 (pGL2099) and Trx-RDK2^{K70M} (pGL2215) as expressed in pET-32 Xa/LIC (Novagen) under control of the T7 promoter. N-terminal thioredoxin (Trx), His and S tags are added for

solubility and purification proposes. Cleavage sites of thrombin and factor Xa are highlighted. **B.** Protein kinase assay performed with [$\gamma-^{32}$P]-ATP radiolabel in the active protein (Trx-RDK2) and the protein kinase 'dead' mutant (Trx-RDK2^{K70M}), both carried out in the presence (S) and absence (nS) of the generic substrate α-casein. Left panel represents the Coomassie blue stained gel and right panel represents a storage phosphor screen imaged in a Typhoon 9400, 72 h of exposure to the gel. **C.** Same as B, but protein kinase activity assay carried out after cleaving the tag with factor Xa. **D.** Coomassie blue stained gel of Trx-RDK2 exposed to different units (U) x μl^{-1} of factor Xa at four different incubation times to optimize cleavage of the tag.

Kinase activity was assayed using radioactive [γ-^{32}P]-ATP and the generic substrate, α-casein, which had previously been used for NEK12.1 (Jones, 2012). The Trx-RDK2 protein was capable of phosphorylating α-casein, whilst the Trx-RDK2^{K70M} was not (Figure 4-5B). Interestingly, Trx-RDK2 also carried out autophosphorylation and this effect was greater in the absence of an exogenous substrate (Figure 4-5B, arrowed). In order to see if the protein maintained its activity after the Trx tag was cleaved, trying to understand if the tag was the target of autophosphorylation, different incubation times and concentrations of Xa factor were tested, Figure 4-5D. The best conditions, 6 h incubation at 0.5 U/μl, were scaled up and applied to both Trx-RDK2 and Trx-RDK2^{K70M} that, after cleavage, were subjected to the kinase assay. The autophosphorylation signal was still observable, although not as intense if the substrate was present, Figure 4-5C.

To attempt to purify RDK2, codon-enriched *Escherichia coli* Rosetta^{TM} (DE3) p-lysS strain, with T7 lysozyme to supress basal T7 expression, was used as the bacterial host. Protein expression was induced with 1 mM IPTG in log-phase cells (OD600=0.5) during an overnight incubation at 17 ^{o}C (see M&M). Protein was purified from bacterial lysates with two rounds of fast protein liquid chromatography (FPLC). This permitted simultaneous fractionation of the elution and monitoring of protein content with UV absorbance, Figure 4-6.

In the first FPLC purification, a Ni^{2} Sepharose column was used to retain the 6xHis-tagged protein. After profuse washing with 35 mM imidazole buffer, elution was conducted by increasing imidazole to 500 mM. As this salt has more affinity for Nickel than the His tag, the Trx-RDK2 could be successfully released from the matrix, Figure 4-6A. After removing residual imidazole from the eluate in a desalting column, a second FPLC round was conducted with an anion

exchange column, Figure 4-6. Every step of purification was assessed by SDS-PAGE and Coomassie blue staining, Figure 4-6. The purest fractions after ion exchange were pooled and concentrated to conduct a kinase assay. Trx-RDK2, and all the mutant versions used in this chapter, could be frozen at -80 ^{o}C with 10% glycerol maintaining their capacity for phosphorylation.

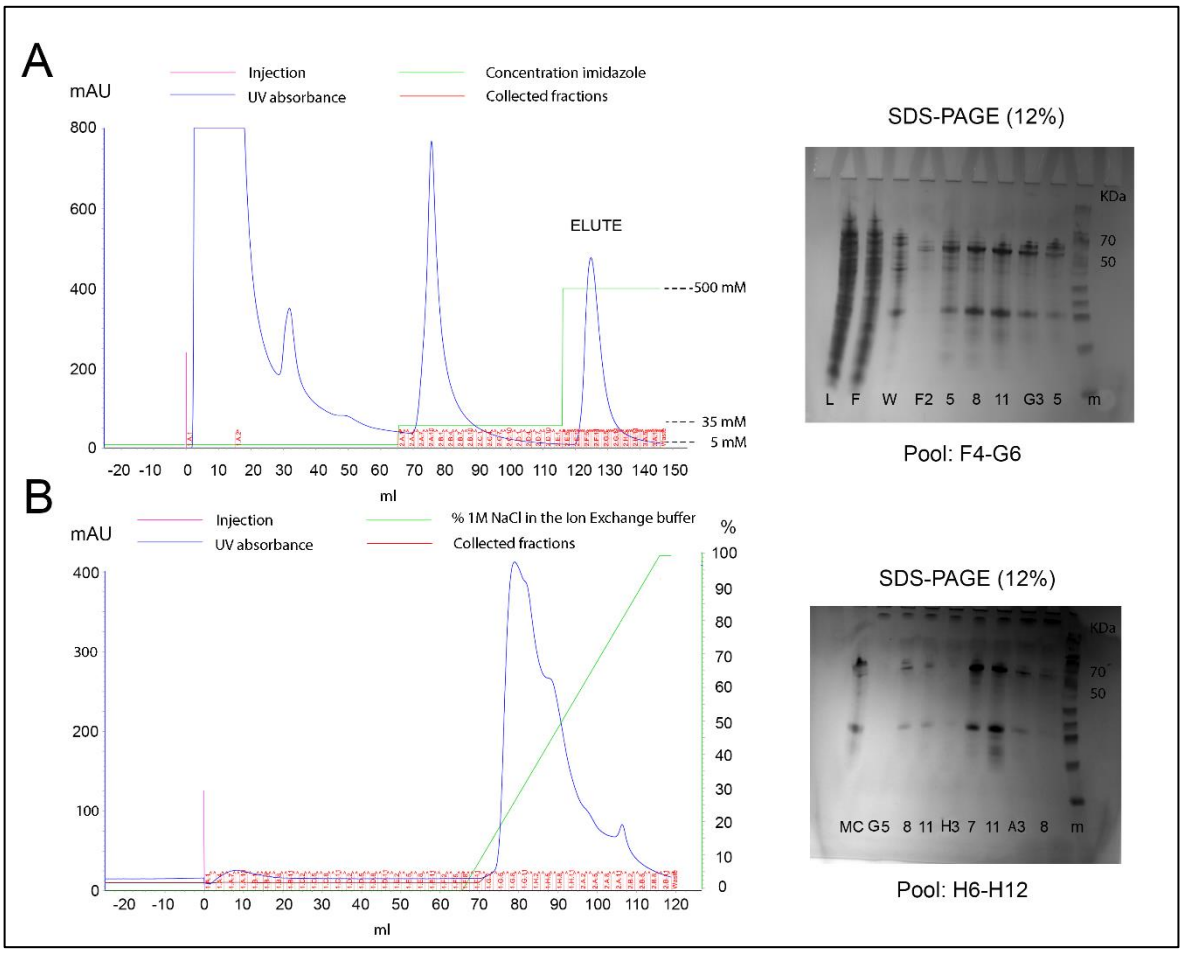


Figure 4-6. Trx-RDK2 purification.
A. Immobilised metal-ion affinity chromatography using a Ni^{2} Sepharose column in an ÄKTA purifier. Left panel, UV absorbance during binding of the His-tagged protein with 5 mM imidazole (0-65 ml), washing with 35 mM imidazole (65-115 ml) and elution with 500 mM imidazole (115-150 ml). 1 ml fractions were collected across the wash and elution peaks. Right panel, Coomassie staining of a 12% SDS-PAGE, loading the whole bacterial lysate before (L) and after filtration (F); a fraction sample from the washing peak (W); a series of fractions across the elution peak (F2,5,6 and 11, and G3 and 5). Fractions F4-G6 (7ml) were pooled and processed through a desalting column. **B**. Ion exchange chromatography using a Q sepharose column in an ÄKTA purifier. Left panel, desalted samples were passed through the column and washed with 10 volumes ion exchange buffer without NaCL (70 ml) and a gradient 0-1 M of NaCl was applied across other 10 column volumes for elution of the protein. Fractions were collected from all the process. Right panel, Coomassie staining of a 12% SDS-PAGE loading the desalted sample from the metal chromatography (MC); and fractions collected across the elution peak (1-G5, 8, and 11; f1-H3, 7 and 11; and 2-A3,

and 2-A8). Fractions 1-H6 to 1-H12 were pooled, kept in 10% glycerol, and used for subsequent kinase assays. m= protein standards.

A study has reported the phosphoproteome of cytosolic fractions from *T. brucei* bloodstream form cells (Nett *et al.*, 2009). There, peptides were identified indicating phosphorylation of two serine residues (S195 and S197) in the middle of the RDK2 protein kinase domain (Figure 4-7**A)**. In order to ascertain whether these residues are required for kinase activity, site-directed mutageneses were performed to allow RDK2 recombinant protein to be expressed with mutations that would mimic the phosphorylated and dephosphorylated states in all the four possible combinations (Trx-RDK2$^{S195E/S197E}$, Trx-RDK2$^{S195E/S197A}$, Trx-RDK2$^{S195A/S197E}$, and Trx-RDK2$^{S195A/S197A}$). Alanine (A) prevents phosphorylation, while glutamic acid (E) simulates the phosphorylated state by introduction of a bulky group including a negative charge (Dissmeyer and Schnittger, 2011). For purification of these mutants we used exactly the same procedure used for Trx-RDK2 (purification data not shown). In the absence of the phosphoserines, only Trx-RDK2$^{S195A/S197E}$ was unable to phosphorylate α-casein and, whilst autophoshorylation was still detectable in all mutants (indicative of other residues not reported by Nett et al. being phosphorylated), it appeared reduced in Trx-RDK2$^{S195A/S197E}$ (Figure 4-7B). Although results for this mutant were encouraging at first, smell of soap and excess of foam during the cell-lysis step were indicative of detergent contamination in the flask used to centrifuge the bacterial culture. If that were true, soap-triggered denaturalization of the protein could justify a loss of the kinase activity instead of a change of conformation caused by the S195A/S197E mutation. Although a faint autophosphorylation was detectable, it would be necessary to repeat the experiment before claiming a mutation-based loss of the kinase activity.I

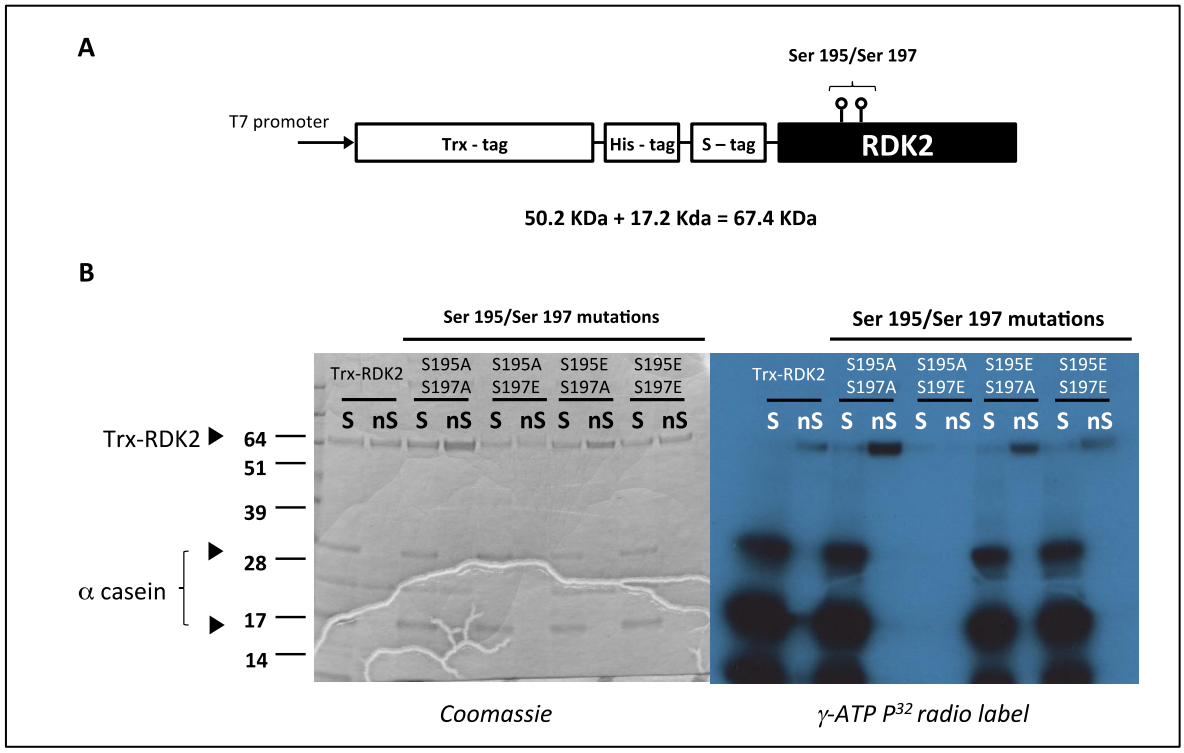


Figure 4-7. RDK2 S195/S197 phosphosite mutations.
A. Schematic representation of recombinant Trx-RDK2 highlighting phosphoserines found in a phosphoproteomic analysis of the cytosolic fraction in bloodstream trypanosomes (Nett *et al.*, 2009). **B.** Mutants to mimic permanent phosphosphorylation (S to E) or permanent dephosphorylation (S to A) in the four possible combinations were expressed and subjected to a kinase assay. Left panel represents the Coomassie blue stained gel and right panel represents the x-ray film after 72 h of exposure to the radiolabel ([γ-^{32}P]-ATP). Kinase assay conducted with the active protein (Trx-RDK2) and the four phosphosite mutants both in the presence (S) and absence (nS) of the generic substrate α-casein.

4.3.3 Molecular validation of the RNAi phenotype

Aiming for molecular validation of the RDK2 RNAi knockdown phenotype, the protein was next re-expressed in the induced RNAi line as an RNAi-resistant variant. To achieve this, the *RDK2* gene was recoded *in silico* by reverse translation to DNA of RDK2's complete amino acid sequence in CLC genomics Workbench, giving priority to *T. brucei* infrequent codons. The resulting gene was manually curated in order to certify that no section longer than 25 nucleotides remained identical to the original CDS of *RDK2* while still producing the right protein (Figure 4-8A). This recoded version of *RDK2* (*rdk2*^{REC}) was cloned into plasmid pRM481 (Proudfoot and McCulloch, 2005), which targets inserts to the tubulin array of genes located in chromosome 1 via homologous recombination. After *Asc*I linearization and transfection into the RNAi line

(sTL629), the new construct would produce stable constitutive expression, conferring blasticidin resistance to permit selection of successful transfectants. Transcripts from *rdk2*^{REC} should not be recognised by the RNAi machinery, avoiding cleavage via RISC and therefore producing RDK2 protein, whilst endogenous *RDK2* transcripts would be degraded. In order to detect expression by western blot, pRM481 was modified to enable fusion of RDK2 with a C- or N-terminal 6xHA tag. Modified pRM481 expressing *rdk2*^{REC}-6x*HA* (C-Terminal tagged) was renamed pGL2271, whilst the 6x*HA*-*rdk2*^{REC} version (N-Terminal) was called pGL2587. Vector maps are depicted in Figure 4-8B.

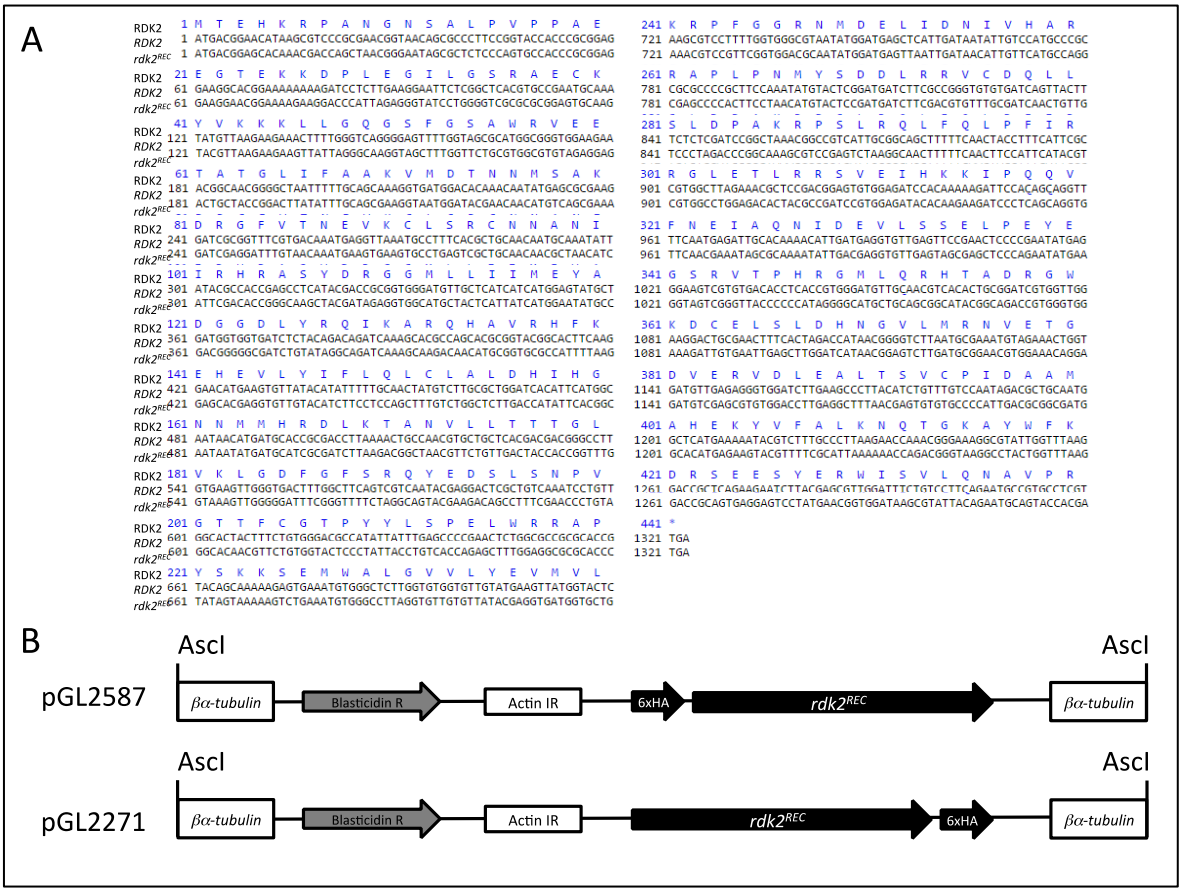


Figure 4-8. RNAi refractory recoded *RDK2* complementation strategy.
A. RDK2 protein sequence (RDK2) aligned to the DNA sequence annotated for *Trypanosoma brucei brucei* TREU927 in TrytripDB (*RDK2*) and the recoded DNA sequence (*rdk2*^{REC}) created by reverse translation with infrequent *T. brucei* codons in CLC Genomics Workbench. **B.** Schematic representation of modified versions of the *Asc*I linearized tubulin-targeting pRM481 by introduction of *rdk2*^{REC} with N-terminal (pGL2587) or C-terminal (pGL2271) 6xHA tag. The construct includes a blasticidin resistance cassette as a selectable marker and an actin intergenic region for trans-splicing and polyadenylation between the blasticidin and *RDK2* ORFs.

A blasticidin resistant clone (renamed sTL768) was selected after transfection of sTL629 with pGL2271 where protein of the expected size (57.2 KDa, with the 6xHA tag) was detectable in western blot with anti-HA antiserum, indicative of successful recombination. Growth curves and EP procyclin expression (measured by flow cytometry) showed the recoded RDK2 could complement for loss of the endogenous gene, since cell growth was equivalent with or without RNAi and differentiation was undetectable after RNAi (Figure 4-9A).

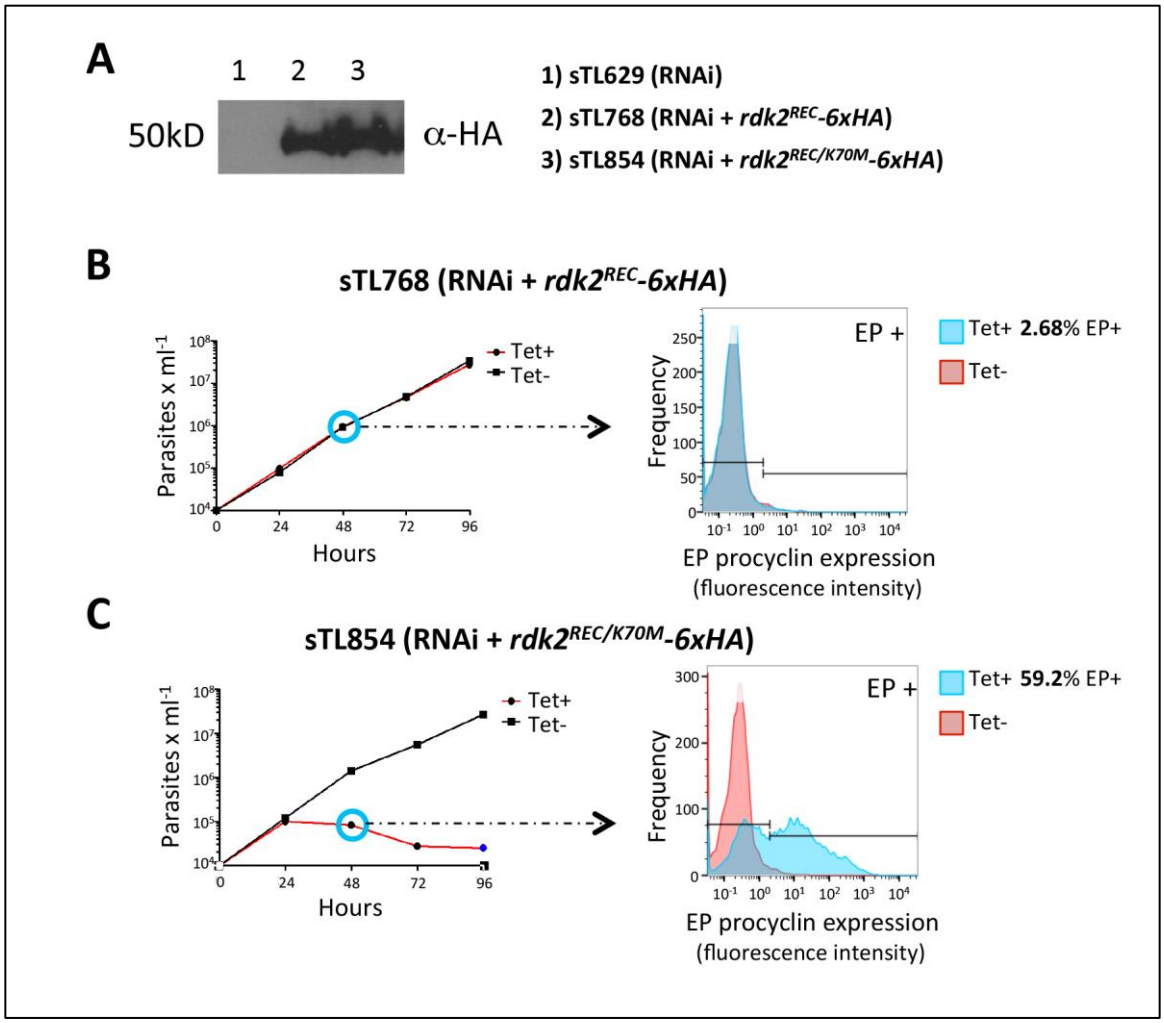


Figure 4-9. Complementation assays with RNAi refractory *rdk2*^{REC} and *rdk2*$^{REC/K70M}$ after RNAi induction.
A. Anti-HA western blot comparing the untagged RNAi line (sTL629) with *rdk2*^{REC}-expressing (sTL768) and *rdk2*$^{REC/K70M}$ -expressing (sTL854) RNAi cells. **B.** Cumulative growth curve (left) and flow cytometry histograms (right) for EP procyclin expression in the *rdk2*^{REC}-expressing clone (sTL768). **C.** Same as (B) but for the *rdk2*$^{REC/K70M}$ -expressing clone (sTL854).

In line with experiments performed with the recombinant protein (Figure 4-5) a kinase dead version, *rdk2*$^{REC/K70M}$-6x*HA*, was produced by site directed

mutagenesis in the recoded plasmid (construct named pGL2297). After transfection and confirmation of protein expression, it was observed that *rdk2*$^{REC/K70M}$-6x*HA* expression was unable to complement for loss of the wild type RDK upon RNAi, since differentiation to procyclic–like forms and a growth defect were detected, Figure 4-9B). These data indicate that RDK2 kinase activity, and not additional functions, are lost when the protein is missing after RNAi.

RDK2 is a NEK kinase, a family whose structural properties have been exploited in *Plasmodium falciparum* to achieve specific inhibition using a tool compound (NCL-00016066) upon substitution of the alanine residue corresponding to codon 55 in RDK2 for a cysteine (RDK$^{REC/A55C}$) (Mitcheson *et al.*, 2016). Aiming to obtain chemical validation of the phenotype, such protein modification was conducted in pGL2271 by site directed mutagenesis, generating the construct pGL2472. However, 6xHA-expressing clones obtained after pGL2474 transfection were unable to rescue the RNAi phenotype as before (Figure 4-10A). Unfortunately, when pGL2271 or pGL2587 were re-transfected in sTL629 (or newly generated RDK2-targeting RNAi clones) as a control, the perfect recovery observed previously was found not to be reproducible (Figure 4-10B-C). Clones obtained with both N-terminal and C-terminal fusion versions of recoded RDK2 were only partially able to rescue the wild type phenotype, both in terms of EP procyclin expression and growth.

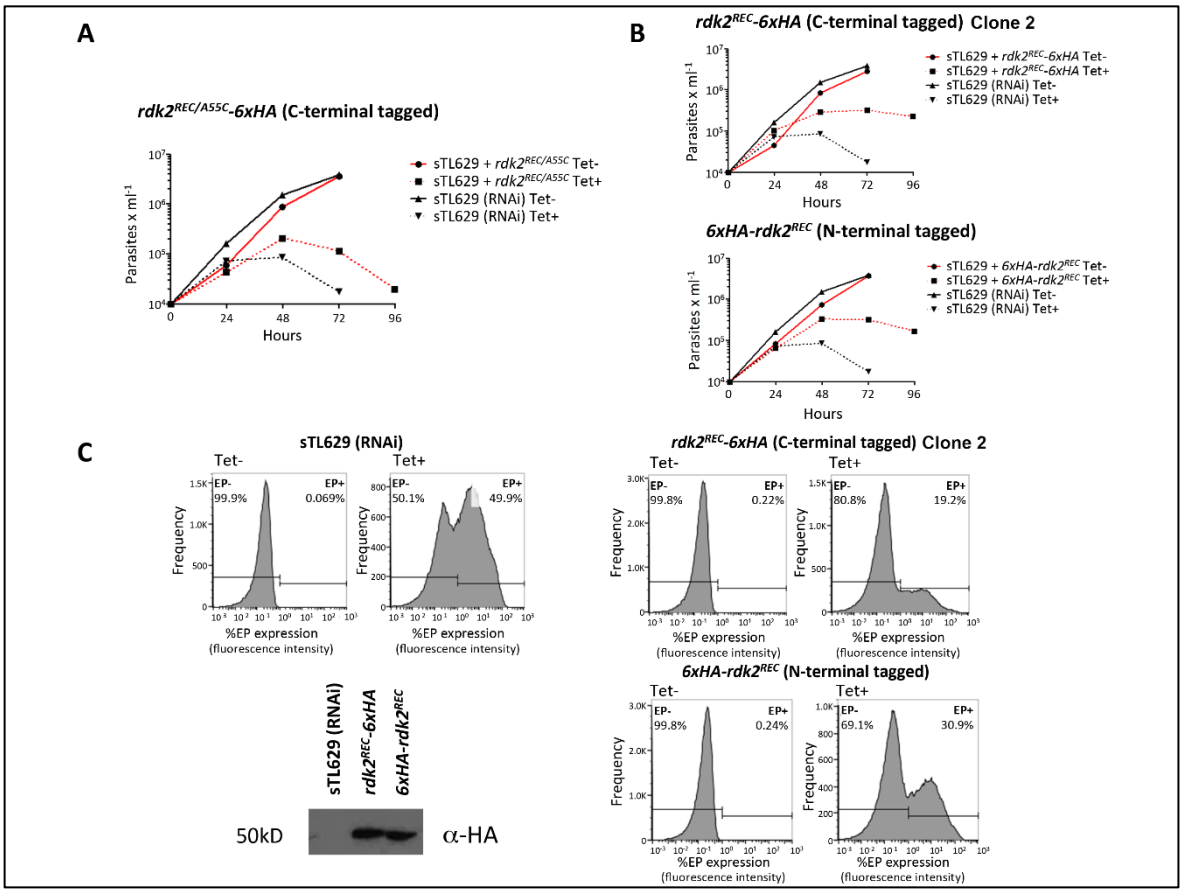


Figure 4-10. Complementation assays with clones recovered after sTL629 transfection with RNAi refractory *rdk2*$^{REC/A55C}$ and *rdk2*^{REC}.
A. Growth curve of a *rdk2*$^{REC/A55C}$-expressing clone, induced (Tet+) and uninduced (Tet-) compared to backbone RNAi phenotype (sTL629 induced and uninduced). **B.** Growth curve of a *rdk2*^{REC}-expressing clone 2, which was C-terminal tagged (top), and *rdk2*^{REC}-expressing clone which was N-terminal tagged (bottom). Induced (Tet+) and uninduced (Tet-), compared to backbone RNAi phenotype (sTL629 induced and uninduced). **C.** Anti-HA western blot showing expression of the *rdk2*^{REC}-expressing clones showed in (B) compared to backbone sTL629 (left bottom) and flow cytometry histogram profiles of EP procyclin expression for the same lines showing uninduced controls (Tet-, left panels) versus induced (Tet+, right panels).

Although promising, the original complementation results (Figure 4-9) were not compelling as perfect recovery observed with clone 1 was not reproducible and later attempts to achieve this always rendered cell lines capable of partial recovery of both phenotypes (growth and differentiation), Figure 4-10. The pronounced difference in the extent of complementation of RDK2 RNAi by recoded, wild type RKD2 in Figure 4-9 and Figure 4-10 is perplexing and remains unexplained.

4.3.4 RDK2 overexpression blocks *in vitro*-triggered differentiation

In order to test further the hypothesis that RDK2 acts as a repressor of bloodstream to procyclic developmental progression, the protein was overexpressed in cells and then subjected to standard differentiation conditions.

To inducibly overexpress the protein, the complete *RDK2* gene sequence was cloned into pRPa^{x6mycx} (Alsford and Horn, 2008), vector known as pGL2220 in the Mottram laboratory. Using the same backbone structure as the pRPa^{iSL} RNAi plasmid described in the RITseq chapter (Figure 3-1), pGL2220 targets a gene of interest to a specific *RRNA* locus selected for efficient transcription. Successful recombination renders cells resistant to hygromycin and generally sensitive to puromycin. 2T1 cells express TetR, which binds the Tet operator blocking transcription until addition of tetracycline (Alsford *et al.*, 2005).

The *RDK2* coding sequence was PCR-amplified from genomic DNA (obtained from *T. brucei brucei* Lister strain 427 wild type cells) and cloned into pGL2220. By addition of suitable restriction sites to the primers, RDK2 was fused to a 6xmyc tag, enabling expression to be tracked. Constructs were made using C-terminal (*RDK2*-6x*myc*, pGL2323) and N-terminal (6x*myc*-*RDK2*, pGL2324) tagging, Figure 4-11A.

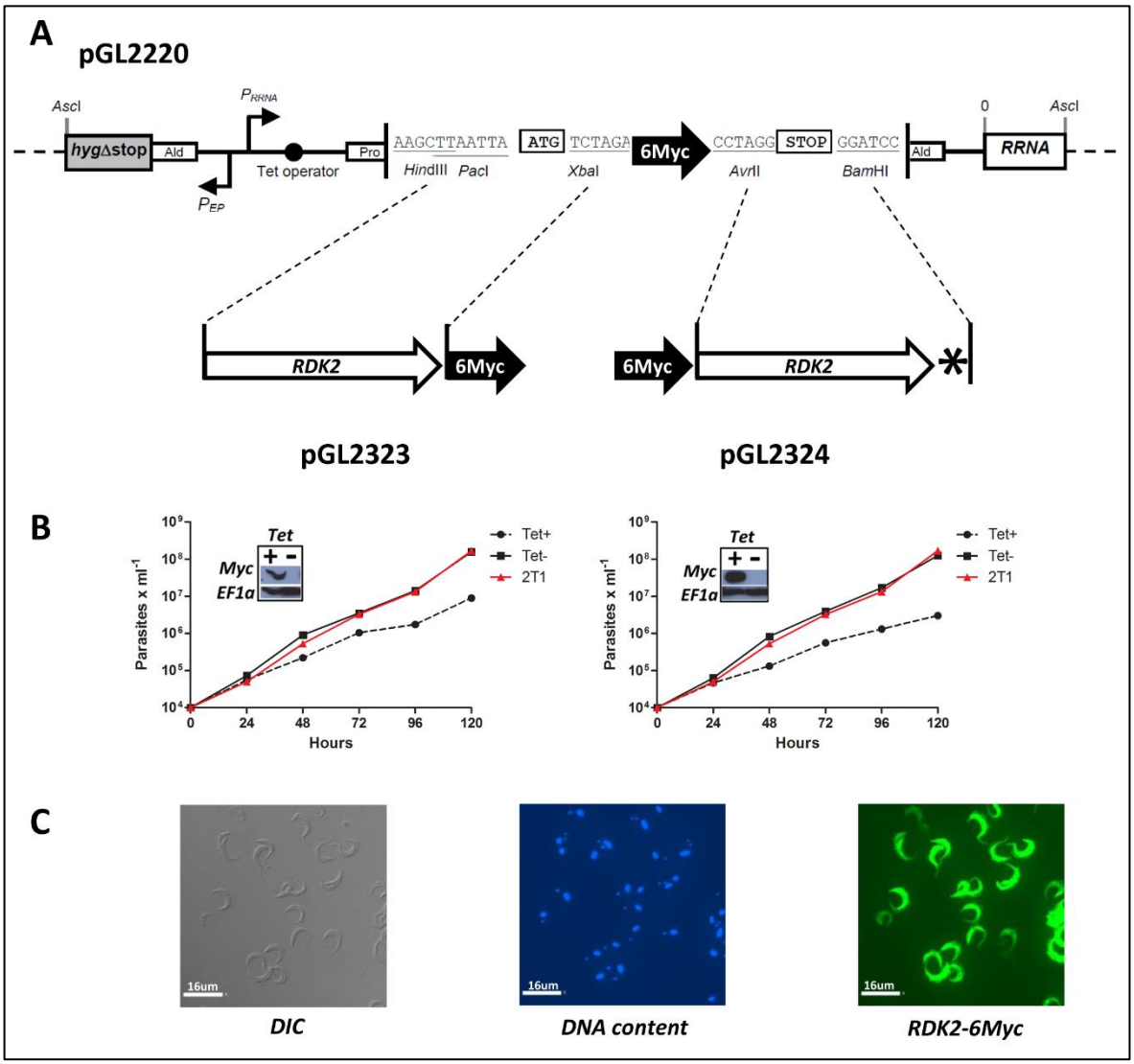


Figure 4-11. RDK2 inducible overexpression.
A. Schematic representation of pRPa^{x6mycx} (pGL2220) and cloning strategy to generate C-terminal (pGL2323, without RDK2 stop codon) and N-terminal (pGL2324, with RDK2 stop codon) tagged constructs. * = stop codon. **B.** Western blot determining expression of RDK2-6Myc and 6Myc-RDK2 in clones obtained upon transfection in 2T1 bloodstream cells, and growth curves comparing progenitor 2T1s with induced (Tet+) and uninduced (Tet-) overexpression. **C.** Microscopic images of the *RDK2-6Myc* over-expresser phenotype after 24 h of induction, showing from left to right: shape (DIC); DNA stained with DAPI (blue); and RDK2 with anti-myc indirect immunofluorescence (green). This work has been undertaken in collaboration with Helena de la Torre.

Plasmids were linearized with *Asc*I and transfected into 2T1 cells. Hygromycin resistant cells recovered were tested for expression and growth under induced and uninduced conditions compared with progenitor 2T1 cells. Western blotting with an antibody that recognises the 6-myc tag showed expression of the fusion protein after 24 h of induction, but not in the uninduced lines, Figure 4-11B. Using the same antibody, IFA showed distribution of the protein across the whole

cell body after 24 h of induction. Parasites looked morphologically normal under the microscope, Figure 4-11C. Growth curves showed a slower growth in the induced cells compared to the uninduced, which grew at the same rate as the parental line, Figure 4-11B. This work was produced in collaboration with Helena de la Torre.

In order to understand how RDK2 overexpression affected life stage maintenance, differentiation assays were carried out *in vitro*. It has been already mentioned that stumpy* formation can be induced in monomorphic trypanosomes, such as 2T1, by treatment with hydrolysable and cell permeable cAMP analogues (Laxman *et al.*, 2006). Differentiation from bloodstream to procyclic forms, which is more efficient starting from stumpy* parasites, can be triggered *in vitro* by treatment with 3 mM citrate/cis-aconitate (CCA) in differentiation trypanosome media (DTM) at 27 ^{o}C (Overath, Czichos and Haas, 1986).

EP procyclin expression after differentiation treatment was assessed in flow cytometry for Tet-induced RDK2 overexpressers compared to uninduced, monitoring basal levels of EP procyclin expression in comparison with non-treated controls for both conditions (Tet+ and Tet-). The assay consisted in addition of 100 μM 8-pCPT-cAMP for 24 h at 37^{o}C in HMI11, washing the compound out, suspending the cells in DTM with 3 mM CCA, and incubating for 24 h at 27^{o}C, before analysing results, Figure 4-12A.

Cells untreated for differentiation (Tet+ and Tet-) showed a background of 0.33% and 0.53% EP procyclin positive populations. After treatment for differentiation, the positive population was ~5% for 6Myc-RDK2 over-expressers (Tet+), and >60% for the uninduced line (Figure 4-12).

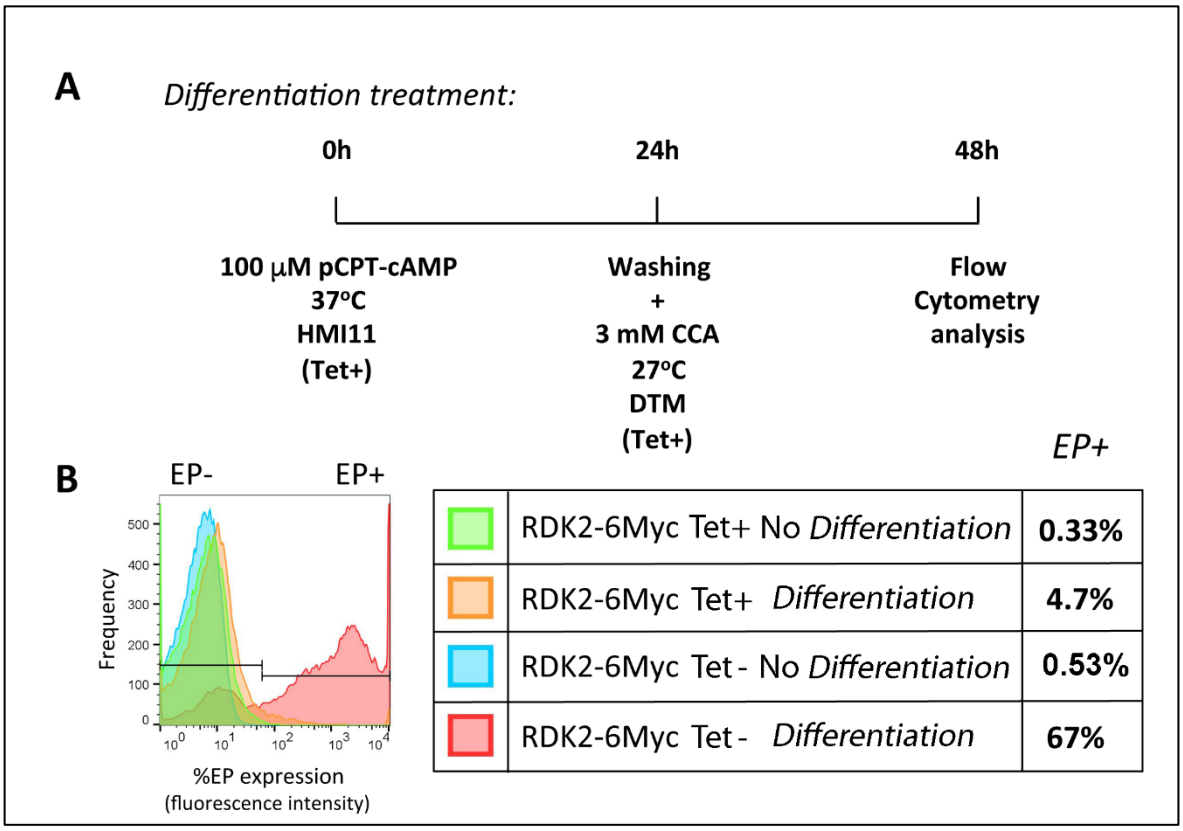


Figure 4-12. Bloodstream to procyclic form differentiation assay with RDK2 over-expresser.
A. Treatment regimen to induce differentiation. **B.** Flow cytometry histograms for EP procyclin expression comparing RDK2 over-expressers, Tet+ and Tet- treated or not for differentiation.

All together these data suggest that RDK2 overexpression efficiently blocks CCA-mediated bloodstream to procyclic form differentiation.

4.3.5 Exploring further RDK2 involvement in known differentiation pathways

In order to learn more about the role RDK2 plays in repressing differentiation, and how it fits within the known molecular networks described in the general introduction, five different conditions were tested after 48 h of RNAi induction compared to uninduced cells, Figure 4-13**A**. The percentage of EP procyclin expression, measured in flow cytometry, was used as a marker for differentiation. All experiments were conducted in HMI11 media, with the distinct differentiation treatments entirely conducted at 37 °C or transferring the cells for the last 24 h to 27 °C.

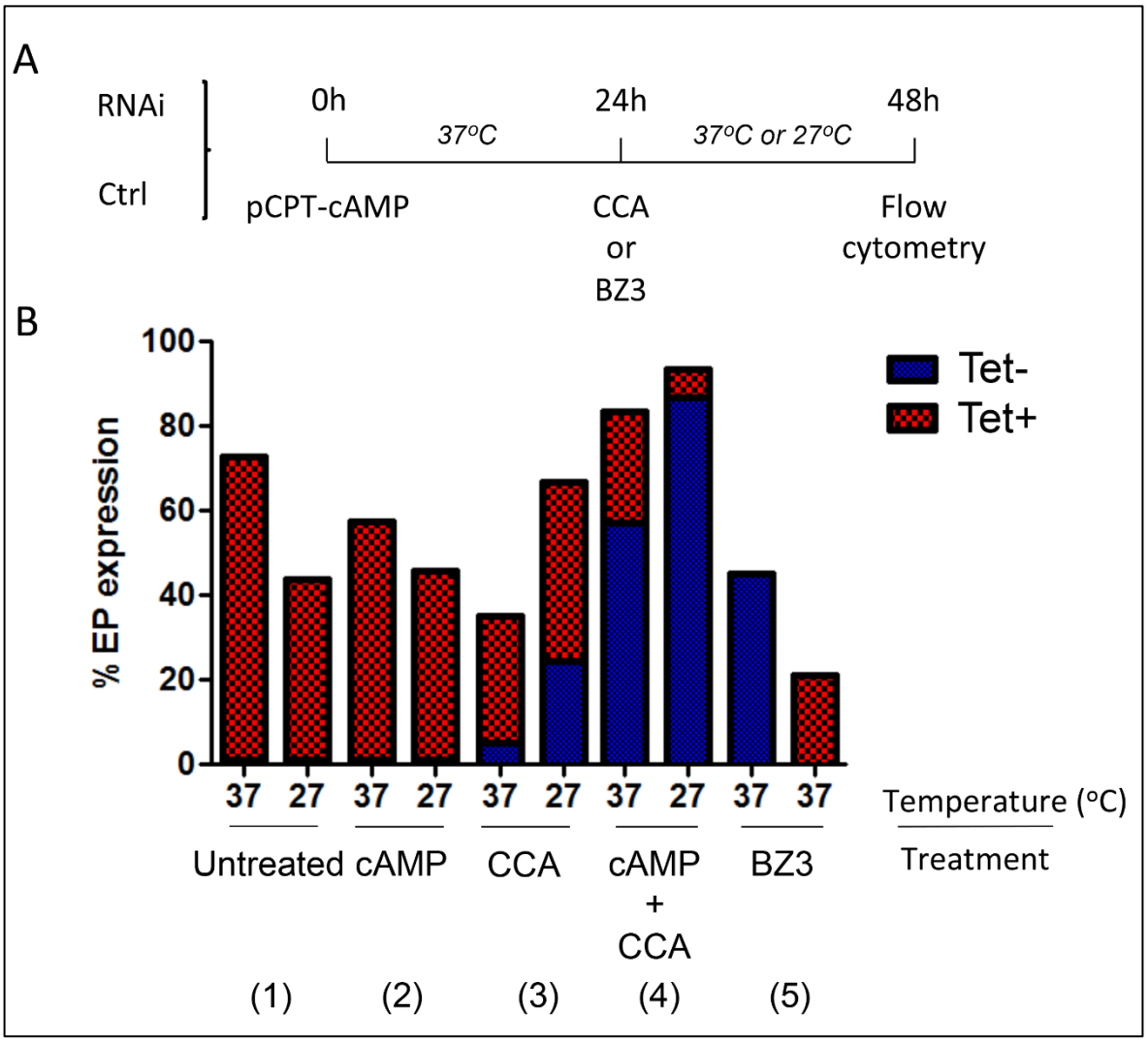


Figure 4-13. RDK2 interaction with the citrate/cis-aconitate pathway of differentation *in vitro*.
A. Schematic representation of the treatment regimen. RNAi was induced for 48h, or uninduced and grown for an equivalent time (ctrl), and cells analysed in flow cytometry to detect EP procyclin expression under 5 different conditions: (1) untreated; (2) 3 mM CCA during the last 24 h of RNAi induction; (3) 100 μM 8-pCPT-cAMP for the first 24 h of RNAi induction, washed and incubated for a further 24 h with tetracycline (RNAi induced only); (4) 100 μM 8-pCPT-cAMP for the first 24 h of RNAi induction, washed and incubated for a further 24 h with 3 mM CCA and tetracycline (RNAi induced only); and (5) 150 μM of BZ3 for the last 24 h of RNAi induction. Last 24 h for all conditions were incubated at 27 °C or 37 °C except BZ3 (5), which was only tested at 37 °C. **B.** Bar diagram depicting the percentage of cells displaying EP procyclin expression in uninduced (blue) and RNAi induced (red) cells for each temperature condition. BZ3 (5) data is split in two columns as only one temperature was analysed.

First, the percentage EP expression was assessed in untreated cells (1). At 37 °C, 70% of RNAi induced were EP positive and 0% of uninduced, validating data shown previously in this chapter. Placing cells at 27 °C for the last 24 h, only ~45% of RNAi induced cells were positive.

Then cells were treated with 100 μM 8-pCPT-cAMP (2), washed, suspended in fresh HMI11, and incubated for further 24 h. When analysed, RNAi produced ~60% positive procyclin population at 37 °C and ~45% at 27 °C. Uninduced cells were still 0% positive.

Next, cells were incubated for the last 24 h with 3 mM CCA (3). At 37 °C, <40% of RNAi cells were EP positive and uninduced cells differentiated to ~5%. Interestingly, both values increased at 27°C: >60% differentiated in the RNAi induced, and ~25% in the uninduced.

The two previous treatments were combined for condition (4). 100 μM 8-pCPT-cAMP was incubated for the first 24 h, washed away, and 3 mM CCA added with fresh medium for the last day. At 37 °C more than 80% of cells were procyclin positive in the RNAi-induced, and almost 60% in the uninduced. Interestingly, at 27 °C, both induced and uninduced lines showed ~90% of cells expressing EP procyclin.

The last condition, treatment with 150 μM of TbPTP1 inhibitor BZ3 (5), was conducted only at 37 °C. RNAi induced cells differentiated to produce ~20% of the population being procyclin positive, while uninduced produced ~45%.

4.4 Discussion

4.4.1 RDK2 is an active kinase

This work demonstrates that RDK2 is an active protein kinase, since recombinant protein is able to phosphorylate a generic substrate and is capable of autophosphorylation. Furthermore, the protein kinase activity was ablated by mutation to methionine of the catalytic lysine in subdomain II (Kamps and Sefton, 1986). Two RDK2 phosphosites (S195 and S197) reported in a phosphoproteome screen (Nett *et al.*, 2009) were also mutated. Trx-RDK2$^{S195A/S197E}$ mutant showed only a very faint autophosphorylation signal and no capability of α-casein phosphorylation. While the kinase assay was performed at the same time and with the same materials as the other serine-modified versions, for this mutant, the elution peak in the chromatographic purification was substantially different. This may have been due to soap contamination

during the protein extraction, as the flasks used for centrifugation of this particular bacterial culture were insufficiently rinsed and produced abundant foam that smelled of soap. This could be enough to cause, for example, inadequate folding. It would be necessary to repeat the protein purification in order to test this possibility. However if the lack of activity is real, it may be due to the specific orientation of these residues respect to the active site. It would be good to model the protein in order to visualize their orientation. It is intriguing why being so close to each other, S195A/S197E is so relevant whilst the other combinations are not. According to the alignment we performed for the phylogenetic tree (extract in Figure S1), these two phosphoserines are in the N-terminal flank of subdomain VIII that, as we indicated in the general introduction, often contains residues that require phosphorylation before the PK becomes active (Chan, Hurst and Graves, 1982; Luo, Zhou and Lodish, 1995; Songyang *et al.*, 1995). RDK2 is a NEK kinase and other members of this family, as human NEK9, also require autophosphorylation in the activation loop in order to be activated (Meirelles *et al.*, 2014)

Autophosphorylation may be targeted to residues that were not detected in the phosphoproteomics screen published by Nett et al. (2009) because when the therein-reported serines 195 and 197 were mutated to alanine or glutamic acid, the signal was still apparent. However, the tag was not cleaved in those mutants and we cannot discard the tag itself being at least a secondary target of phosphorylation. On the one hand, the autophosphorylation was fainter in cleaved Trx-RDK2, but on the other, the band corresponding to the tag did not look phosphorylated. As lack of phosphorylation in the cleaved tag may be product of inadequate conformation when detached from the protein, cleaving the tag in the mutants to see whether we still see autophosphorylation would be a good approach.

In addition to their impact on kinase activity directly, RDK2 phosphosites may have other functional implications, such as interaction with other proteins (Cohen, 2000). Indeed, 50% of *T. brucei* PKs were found to be phosphorylated in more than one site (Nett *et al.*, 2009). Examples of protein kinases in *T. brucei* affected by phosphorylation in other proteins are TbPLK1, which contains a phosphoserine/threonine-binding POLO box (Hammarton *et al.*, 2007), or CAMK,

which contains a phosphothreonine-binding forkhead associated domain (Parsons *et al.*, 2005). Nonetheless, autophosphorylation is required for activation in some cases, such as DYRK kinasess (Lochhead *et al.*, 2005). In most eukaryotes, MAPK cascades are a chain of phosphorylation events that mainly activate transcription factors (Widmann *et al.*, 1999). In differentiation pathways of the *T. brucei*, for instance, TbPIP39 phosphatase is activated by tyrosine phosphorylation and deactivated by dephosphorylation (Szöor *et al.*, 2010).

4.4.2 Pleckstrin homology domain

In terms of localization, function and structure, this study is missing a thorough exploration of the exact role that the C-terminal Pleckstrin Homology (PH)-like domain of RDK2 plays, Figure 4-1B. It would be interesting to see whether its presence is required for the kinase activity, by expression of a truncated recombinant protein containing just the PK domain; and if complementation with such a mutant in the RNAi induced line affects *in vivo* function, e.g. by mislocalization or deficient interaction with other partners. Truncation of PH domains has been successfully used in other eukaryotes to understand their activation mechanism (Han *et al.*, 1998)

PH domains are 100-120 amino acids in length and are generally folded in a β-sandwich core structure (Lemmon, Ferguson and Abrams, 2002). Crystallography studies suggest that PH domains could block substrate access to the active site of protein kinases in certain structural contexts (Soisson *et al.*, 1998; Worthylake, Rossman and Sondek, 2000). In other organisms, such as mammalian neutrophils (Servant *et al.*, 2016) and *Dictyostelium* (Jin *et al.*, 2000; Funamoto *et al.*, 2001), PH domains have also been linked with protein polarization and chemotaxis mediated by PI3K activity through specific binding to phosphoinositides, acting as substrates or products. There is also evidence of PH carrier proteins, such as pleckstrin, where activity is regulated by phospholipid-dependent protein kinase C (PKC)-mediated phosphorylation in three adjacent sites to the PH domain (Ma and Abrams, 1999). Whether such phosphorylation sites may be present in RDK2, in addition the two reported serines (195 and 197) in the PK domain, is unknown. However, the phosphoproteomic screen (Nett *et al.*, 2009) only examimed the cytosolic fraction of *T. brucei* bloodstream forms and, in order to provide unambiguous

validation, only considered peptides showing at least three fragment ions. Some low abundant peptides or phosphorylation events elsewhere in the cell may therefore have escaped analysis.

Around 10% of described PH domains have high affinity and specificity for phosphoinositides, with small K_D values that make them behave as membrane targeting sequences (Paterson *et al.*, 1995; Lemmon, Ferguson and Abrams, 2002). Others, of lower affinity, act more in protein-protein interaction, such as binding the $\beta\gamma$ subunits of G proteins (Touhara *et al.*, 1994; Pitcher *et al.*, 1995), actin (Yao *et al.*, 1999), RACK1 (Rodriguez *et al.*, 1999) or PKC (Yao *et al.*, 1997).

Given the role of RDK2 in the pathways regulating *T. brucei* differentiation, it is possible the kinase influences membrane and cytoskeletal remodelling through the PH domain. For example, expression of procyclins on the surface (EP and GPEET) (Vassella *et al.*, 2000), and VSG removal (Ferguson *et al.*, 1988; Grandgenett *et al.*, 2007) may affect RDK2 activity and localization. The glycosylphosphatidylinositol (GPI) anchors that attach the different surface glycoproteins to the membrane (Schwede *et al.*, 2015) are an example of a structure that may interact with RDK2 via the PH domain. This deserves to be investigated further.

4.4.3 Molecular validation of the RDK2 RNAi phenotype

Molecular validation of the phenotype observed upon RDK2 RNAi was conducted with an RNAi-refractory recoded version of the gene able to complement the endogenous protein after RNAi-mediated loss. A first clone (sTL768) was capable of nearly 100% recovery of the wild type phenotype, with clear expression of the HA-tagged recoded protein, indicative of the RNAi not producing off-target effects. When the same construct was introduced with a mutation in the catalytic lysine, the phenotype could not be rescued. This finding would validate RDK2 as a potential drug target for kinase inhibitors, since the lack of activity, and not depletion of the protein itself, appears responsible for the phenotype. Differentiation to procyclic forms, preceding cell death, EP procyclin expression could be exploited as a reporter of RDK2 inhibition in a phenotypic screen of

trypanocydal compounds. This approach was attempted in the following chapter of this thesis.

Unfortunately, the above evidence for complementation is compromised by the observation that other clones, which also showed protein expression in western blot, were only partially able to rescue the wildtype phenotype.

One possible explanation is the approach used to achieve RDK2 depletion independently from NEK12.1. RDK2-specific fragments longer than 20 nucleotides (divergent from NEK12.1) were artificially assembled. Cleavage via Dicer (Patrick *et al.*, 2009) of the resulting double-stranded RNA may produce unwanted small silencing RNA fragments encompassing the artificial junctions and binding undesired transcripts for degradation. However, RNAi having off target effects is difficult to envisage since the phenotype observed with the potentially conflictive RDK2-specific construct was identical to that observed with the construct that targeted also NEK12.1 (Jones *et al.*, 2014). Another possible explanation for this complementation failure would be a need of tightly controlled levels of expression for the protein kinase to act correctly. Indeed, inducible overexpression of RDK2 caused a noticeable growth defect. Finally, the infrequent codons used to recode the gene (Figure 4-8**A**) may bind less abundant tRNAs, leading to suboptimal expression levels. This has been proposed to be part of the unconventional gene expression regulation in *T. brucei* (Horn, 2008). Supporting this idea, a large number of parasites -10^7 parasites/lane- were required to obtain the western blot signals in Figure 4-9A and Figure 4-10C. Profiling RNA levels for *RDK2*^{REC} by qRT-PCR could shed light on this.

An alternative validation system for RDK2 function would be deletion of the gene. If loss of fitness driving to cell death is still a consequence of RDK2 depletion, a knockout line would be impossible to make, but a conditional null mutant is a possibility. A successful approach for this is replacing 1 gene allele with a selectable marker and introducing a Tet-inducible copy of the gene in the *RRNA* locus (Damerow *et al.*, 2016). The second endogenous allele is then removed while complementing the gene expression from the *RRNA* copy. Finally, a null mutant is induced by removing tetracycline. For RDK2, we would need to explore the untranslated regions to see if it is possible to remove specifically *RDK2*, since homology with *NEK12.1* may be a challenge.

Alternatively, it might be worth implementing the CRISPR/CAS9 genome editing technology, which has been probably the last revolution in the molecular biology field. Since this system has been developed and applied successfully in *Trypanosoma cruzi* (Lander *et al.*, 2015), *Leishmania major* (Sollelis *et al.*, 2015) and *Leishmania donovani* (Zhang and Matlashewski, 2015), an inducible version could also by attempted in *Trypanosoma brucei*. The approach would have great potential as it would be used in the same manner than RNAi is used today, with potential for whole genome screening. It has been recently done in *Toxoplasma gondii* (Sidik *et al.*, 2016). Using a 2T1 background, the sgRNA sequence could be introduced in the *RRNA* locus under control of the Tet-operator and the Cas9 expressed from the tubuline array. We would need to deal with the problem of modulating Cas9 to avoid off target cleavage and toxicity. In other organisms they solved that problem developing inducible systems (Dow *et al.*, 2015).

Finding a specific inhibitor of RDK2 to attempt chemical validation of the phenotype would provide an alternative approach. RNAi refractory *rdk2*$^{REC/A55C}$ was designed with the objective of specific RDK2 inhibition with compound NCL-00016066 (Mitcheson *et al.*, 2016) in the RNAi line depleted of endogenous RDK2. However, the inability to reproduce the initial complementation with the recoded gene forced a stop in this analysis. Nonetheless, Chapter 5 describes a screening campaign devoted to find an RDK2 inhibitor; further details are provided there.

4.4.4 RDK2 expression levels

The main limitation of RNAi is that it degrades transcripts and it is not clear if all aspects of the proteome are equivalent to the transcriptome. Kinetoplastids encode genes in polycistronic units that, after transcription, are spliced and polyadenylated into mRNAs that can be stored in pools awaiting translation until, for example, signalling triggered by environmental cues initiates a developmental adaptation (Cassola, De Gaudenzi and Frasch, 2007). This system, where RNA stability plays a fundamental role, enables a fast life cycle progression when survival of the parasite requires it. Endogenous tagging of RDK2 would be a useful strategy to understand better the RNAi effect on the cellular pools of the protein: how long before the transcript absence becomes severely depleted can the stocks of protein that were available prior to RNAi

last? Expression of EP procyclin can be detected by 36 h, but Jones et al reported, for the cell line targeting RDK2 used here, successful transcript depletion from 24 h onwards (Jones *et al.*, 2014).

Detailed profiling of *T. brucei* protein expression upon bloodstream to procyclic differentiation triggered by CCA has been recently described (Dejung *et al.*, 2016). In that work, slender and stumpy bloodstream forms are compared to established procyclic forms and six different time points along the differentiation programme (2 h, 4 h, 6 h, 12 h, 24 h and 48 h). RDK2 was highly expressed in bloodstream forms, with a significant increase upon stumpy differentiation that was sustained during the first 24 h post CCA treatment. In the 48 h time point, expression drastically decreased, being completely absent in established procyclic forms. A peak in transcript abundance found in slender forms preceded the peak of protein expression observed in stumpy forms. This 'transcriptomic anticipation' pattern was repeated later, as a decrease in RNA levels was detectable between 18 h and 24 h within the programme of differentiation (Kabani *et al.*, 2009; Queiroz *et al.*, 2009), whilst not being evident at a protein level until 48 h. It is interesting how this matches the 36-48h time frame required for the RNAi phenotype to arise (Figure 4-2B-D) and supports the idea of RDK2 transcription arrest, inferred from the transcriptomic datasets (Kabani *et al.*, 2009; Queiroz *et al.*, 2009), being a consequence of CCA treatment.

A possible hypothesis to take these data into account is that major remodelling following life cycle progression affects RDK2's interaction network and also the signalling regulating its expression levels. Here the PH domain could play an interesting role; for instance, it may dynamically target RDK2 to the membrane surroundings, which is object of some of the major remodelling events, or it may mediate interactions with other, unknown proteins. In this regard, it is interesting that another fundamental protein involved in membrane remodelling following CCA treatment, GPI-PLC, has a similar expression profile to RDK2 in the quantitative proteomics screen (Dejung *et al.*, 2016).

Overexpression experiments showed RDK2 to be localised throughout the BSF cell, although proper deconvolution of the fluorescent signal would be necessary to determine if it is located in the membrane environment. Also, the

overexpression may cause abnormal localization and it would be better to conduct such analysis tagging the original locus or raising an antibody against RDK2 in order to observe natural levels of expression.

4.4.5 RDK2 depletion and known pathways of bloodstream-to-procyclic differentiation.

If we assume that the RNAi effect is exclusively a consequence of RDK2 depletion, as the first complementation experiment seemed to suggest, it is still not easy to disentangle the exact role this protein kinase plays in the differentiation cascades. From the three pathways that have been described to trigger bloodstream-to-procyclic form differentiation *in vitro* (CCA treatment, mild acidic conditions and mild protease digestion), the first two have been proven to act through a phosphatase cascade involving TbPTP1 and TbPIP39 (Szöőr *et al.*, 2013). The same effect can be reproduced by inhibition of TbPTP1 with a tool compound, BZ3 (Szöor *et al.*, 2006). In the data shown in Figure 4-13B, after CCA treatment at 37 ^{o}C, there were only ~5% of differentiation-competent stumpy* cells but this effect increased to ~60% if pre-treated with 8-pCPT-cAMP. These are the same changes observed for uninduced cells in the RDK2 over-expresser line, and for the parental 2T1 strain. These observations suggest that the uninduced mutants are good controls. Tetracycline-inducible constructs based on the pRPa vector (Alsford and Horn, 2008), used here for RNAi and overexpression, can only act on the ~95% of cells that are slender replicative forms, because in differentiation-competent stumpy* forms the transcription machinery required for production of the double stranded RNA or mRNA is quiescent (Brecht and Parsons, 1998; Capewell *et al.*, 2013).

Interestingly, in cells overexpressing RDK2, if stumpy* formation was induced by incubation with cAMP analogue for 24 h, washed, and incubated with CCA at 27 ^{o}C for further 24 h, only ~5% were procyclin positive (see Figure 4-11). Assuming that only stumpy* cells are able to progress in the life cycle upon CCA treatment, we can hypothesize that RDK2 overexpression blocks stumpy* formation. This interpretation contradicts the elevated expression of RDK2 observed in stumpy forms (Dejung *et al.*, 2016). However, this dichotomy could be explained if the increase in expression of RDK2 were a consequence, and not

a cause, of stumpy formation. Further experiments to evaluate stumpy markers, such as PAD transporter staining, are needed to test this hypothesis.

Figure 4-13B shows that RDK2 depletion alone at 37 ^{o}C triggers procyclic form differentiation to similar levels (~60%) obtained by CCA treatment in stumpy* cells enriched with 8-pCPT-cAMP. However, the reason why CCA treatment decreased the extent of differentiation obtained with RDK2 RNAi remains intriguing. The extra EP procyclin expression found if RNAi induced and CCA treated cells are incubated at 27 ^{o}C (~25%) matches the extra expression observed at 27 ^{o}C when the "residual" stumpy* population in the uninduced culture was treated with CCA. One explanation may be that formation of procyclic form cells in these conditions represents the additive effect of differentiation of slender or intermediate forms by RNAi and differentiation produced by CCA only in the residual stumpy* population. Slender and stumpy states are two ends of a developmental progression, which has many intermediate stations (Mony and Matthews, 2015). Under the proposed model (Figure 4-14), it may be that some of the transcriptionally active (RNAi susceptible) intermediate cells are already receptive to CCA by expression of PAD transporters (Dean *et al.*, 2009), and that the concomitant effect with RDK2 depletion is counter-productive. PAD expression should be analysed to test this.

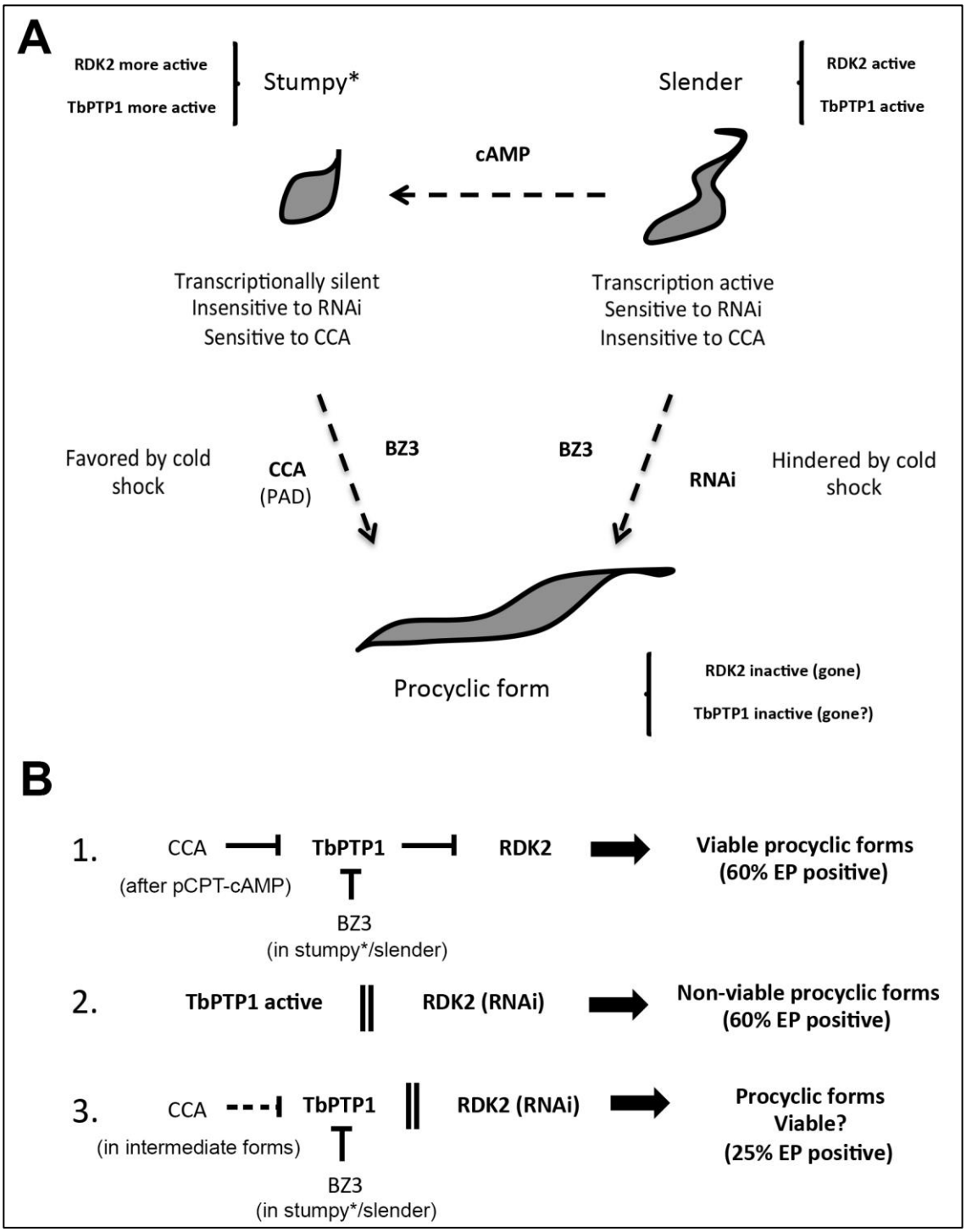


Figure 4-14. Model of RDK2 RNAi-mediated differentiation interconnection with CCA-TbPTP1 pathway.
A. Stumpy* cells are transcriptionally silent and insensitive to RNAi action, while expressing PAD transporters able to convey CCA signalling. Slender replicative forms are sensitive to RNAi but do not have PAD transporters to convey the CCA signal. BZ3 inhibits TbPTP1, which is a downstream effect of CCA, regardless the presence of PAD transporters. **B.** (1) Proposed CCA differentiation pathway involving RDK2 inactivation as a downstream effect, which produces viable procyclic forms. (2) Effect of RDK2 RNAi depletion when TbPTP1 is active, producing non-viable procyclic forms. (3) Effect of concomitant RDK2 RNAi-depletion and BZ3-mediated inactivation, which produced a reduced procyclic positive population of untested viability. Such effect could be triggered by CCA

in a fraction of the population expressing PAD transporters while maintaining active transcription (intermediate forms).

As mentioned above, under natural CCA-mediated action, RDK2 abundance remains high for the first 24 h on the differentiation programme; after 48 h it is reduced to levels found in slender forms and it is absent in established procyclic forms (Dejung *et al.*, 2016). RDK2 presence may be a requirement for the CCA-mediated mechanism during the first 48 h of treatment, while RDK2 depletion in standard bloodstream conditions may trigger the mechanism through a different pathway that renders imperfect procyclic cells unable to survive for long, Figure 4-14B. Perhaps, CCA action changes the normal interactome of RDK2 in bloodstream forms, which may be required in the RNAi-triggered mechanism in order to produce a differentiation effect. Analysing RDK2 interaction partners through BioID (Roux *et al.*, 2012) or immunoprecipitation linked to mass spectrometry in slender and stumpy bloodstream forms would be very informative.

TbPTP1 inhibition with BZ3 at 37 ^{o}C, as shown Figure 4-13B, reinforces the above interpretation, since this treatment, which mimics a downstream effect of CCA treatment (Szöor *et al.*, 2006), reduces the extent of RNAi-mediated differentiation (~20%) to similar levels of CCA alone. In contrast, the percentage of uninduced cells differentiated with BZ3 -which does not require PAD transporters to penetrate the cell- was similar to what it is achievable after consecutive treatment with 8-pCPT-cAMP and CCA. Altogether, these data suggest that the conflict between RDK2 RNAi-mediated differentiation and the CCA pathway happens downstream of TbPTP1.

To sum up, the additive effect with RNAi observed after 8-pCPT-cAMP+CCA treatment can be explained by CCA-differentiation of the extended stumpy* population produced via 8-pCPT-cAMP. Due to transcription silencing, RNAi would have no penetrance in that population. The rest of differentiated cells would be a product of a conflict between the RNAi depletion and PTP1 inhibition mechanisms. On the other hand, under BZ3 treatment, we only see differentiation for the reduced population, as a result of the conflict between RDK2 RNAi and TbPTP1 inhibition. If RDK2 is not present, TbPTP1 inhibition is not that effective in triggering differentiation and, if TbPTP1 is not active, RDK2

RNAi is not capable of differentiating cells. The interactions within this intricate network would change between slender and stumpy stages.

The reason why 8-pCPT-cAMP/CCA treatment permits an additive effect with RNAi, but BZ3 does not, would be that CCA requires PAD to penetrate stumpy* cells (where RNAi is uneffective), while BZ3 acts without needing stumpy* formation and therefore can interfere with changes following RDK2 RNAi. CCA would still be able to interfere with RNAi in intermediate cells, which are expressing PAD receptors but are transcriptionally active.

The transport of PAD receptors to the parasite surface increases at 27 ^{o}C (Dean *et al.*, 2009), which may explain why our 8-pCPT-cAMP-treated uninduced cells are more responsive to CCA at this temperature, Figure 4-13. We have no data of combined 8-pCPT-cAMP and BZ3 treatment to see whether it expands the procyclin positive population in uninduced cells, as occurs with CCA. If the only cellular rearrangement important for differentiation in stumpy* cells is targeting PAD transporters to the membrane, it would not increase the extent of differentiation triggered by BZ3.

Although TbPTP1 inhibition (artificially produced via BZ3 treatment) mediates differentiation caused by mild acidic treatment, we have not tested directly the effect of RDK2 depletion under lower pH conditions. On the other hand, divergent differentiation pathways activated by mild protease digestion (Szöőr *et al.*, 2013) might be of particular interest as, although RDK2 inactivation and depletion seems to be a consequence of TbPTP1 inactivation, early depletion must cause differentiation through a slightly different route. Protease treatment, to which procyclic forms are resistant (Acosta-Serrano *et al.*, 2001), has been described to cause differentiation faster and in a more synchronous manner than CCA. Interestingly, this effect happens mainly in stumpy forms, as this treatment rapidly kills slender forms (Sbicego *et al.*, 1999). A hypothesis that needs to be tested is that proteases may cause membrane remodelling that is able to trigger a similar effect to the one caused by RDK2 depletion, considering possible implications discussed for the PH domain. We know, for instance, that a metalloprotease (TbMSB-B) is responsible for VSG coat detachment (Grandgenett *et al.*, 2007), a fundamental step in the process of bloodstream-to-procyclic differentiation. In addition, CCA and mild protease co-

treatment produces a synergistic effect in differentiation (Sbicego *et al.*, 1999), similar to the one observed here after combination of 8-pCPT-cAMP/CCA successive treatment with RDK2 RNAi.

4.4.6 RDK2 depletion and the *T. brucei* cell cycle

Differentiation to procyclic forms triggered by CCA requires cell cycle-arrested stumpy forms (Rico *et al.*, 2013). To have an effect, RDK2 RNAi may require G1 cells too. After 72 h of induction the procyclin-positive population seems enriched in 1N1K and aberrant cells, while the negative population looks enriched in 2N2K cells (see Figure 4-3C). However, the number of undifferentiated cells currently examined (n=38) limits the reliability of this counting. Indeed, contradicting this datum, DNA staining in immunofluorescence microscopy (n=200) or flow cytometry after 48 h RNAi showed the same cell cycle distribution in differentiated and undifferentiated cells (see Figure 4-3C and D).

Enrichment in 2N2K cells and accumulation of an aberrant population indicated a block in cytokinesis between 36-48 h, coinciding with the first symptoms of growth arrest, see Figure 4-3A. Both cohorts of cells decreased at the 72 h time point, when cell death had clearly occurred, suggesting aberrant cells are those that die first. Surviving EP positive populations were enriched in 1N1K cells, but a remnant of aberrant cells were still present, as cell death continued for another 24h, Figure 4-3C.

Synchronizing the cell cycle, such as with hydroxyurea (Mutomba and Wang, 1996), would help in understanding how the different aspects of the RNAi phenotype come together in the context of the cell cycle, including change of shape, kinetoplast repositioning, EP procyclin expression, cytokinesis arrest and aberrant combinations of nuclear and kinetoplast DNA. In addition, characterising in more detail the aberrant population could also be informative.

In the model we are proposing (Figure 4-14), when CCA induces procyclic formation from stumpy* forms, it also triggers a progressive loss of RDK2, as observed by quantitative proteomics (Dejung *et al.*, 2016). As can be inferred from the effect that RNAi has in differentiation, RDK2 depletion would cause EP

procyclin expression, kinetoplast repositioning and cell elongation. If RDK2 is RNAi-depleted without rearrangement of its interaction network triggered by CCA treatment (and not as a cause of it), an aberrant differentiation would happen that would rapidly render the cells inviable.

4.4.7 RDK2 RNAi phenotype

RDK2 transcript depletion has been reported from 24 h post RNAi induction. The differentiation phenotype was evident from 36 h, with cell elongation, kinetoplast repositioning and EP procycling expression. Growth arrest happens between 36 h and 48 h and is associated with an eventual block in cytokinesis, which develops into aberrant cells that most likely end up dying. After 72 h of induction, cell numbers dropped 10 fold.

Consistent with what has been reported in the literature at a transcriptomic and at a proteomic level (Kabani *et al.*, 2009; Queiroz *et al.*, 2009; Dejung *et al.*, 2016), RDK2 is no longer needed after differentiation is triggered. Differentiation is, most likely, irreversible. No substantial difference was observed between cultures whether tetracycline was maintained or removed (see Figure 4-4). The different attempts made to grow RNAi-differentiated cells as real insect-infective parasites failed to acquire standard densities of procyclic form cultures (~10^7 parasites ml^{-1}). The cells that resulted from these experiments died slowly over the course of a week in fresh SDM79, but maintained population numbers for at least one week if cultured in conditioned media. Nonetheless, in these conditions the cells were still unable to grow. The reasons why conditioned media maintains procyclic forms generated by RDK2-depletion alive and not the fresh media requires further investigation. The same media killed uninduced bloodstream forms in less than 24 h.

The plastic-adhesion phenotype in conditioned medium is also an interesting feature. The fact that it was only observed in plates might be explained if instead of death, cells were attached to the bottom of the standing up flasks, precluding their quantification. This phenotype was long ago reported for epimastigotes, which require attachment for division to the tsetse salivary glands; it has also been observed *in vitro* for *T. vivax*, *T. cruzi* and *T. congolense* (Vickerman, 1969; Figueiredo, Rosa and Soares, 2000; Sakurai, Sugimoto and

Inoue, 2008). As it happens with proventricular forms although they do not attach until reaching the salivary glands, RNAi RDK2 cells could be maintained alive but they were not dividing. *T. brucei* epimastigotes, however have been always challenging to culture (Wallbanks, Molyneux and Dirie, 1989).

It would be interesting to test stage-specific markers of tsetse fly-infective forms in immunofluorescence for better characterisation. Early procyclics, the first stage upon differentiation from bloodstream trypomastigotes, express GPEET procyclin (Vassella *et al.*, 2000). This early stage shows social motility *in vitro* and is present in the midgut before colonization of the ectoperitrophic space (Imhof *et al.*, 2014). Other interesting marker would be BARP, a stage-specific surface antigen of salivary gland forms, that would tell if the plastic adhesion phenotype is closer to them than to early procyclics. Finally, levels of expression of RBP6 (Tb927.3.2930), might be worth checking as overexpression of this RNA binding protein permited developing *in vitro* all stages of the parasite´s life cycle in the tsetse fly (Kolev *et al.*, 2012).

Full understanding of the RDK2 RNAi differentiation phenotype might necessitate using pleomorphic parasites. These are able to produce real stumpy forms in the context of an *in vivo* infection, which renders cells less virulent for the host (Turner, Aslam and Dye, 1995). Some of the natural mechanisms in the process of differentiation may be modified in monomorphs beyond their incapacity to monitor their own population progressing to stumpy in response to high densities.

4.4.8 Final conclusion

This chapter has proven RDK2 to be an active protein kinase, where one particular double mutation of the reported phosphoserines ($RDK2^{S195A/S197E}$) interfered its capacity for phosphorylation. Mixed results attempting complementation of RNAi failed to provide molecular validation of the phenotype. A conditional null mutant or a specific chemical inhibitor could dispel doubts over the cause-effect relationship between RDK2 depletion and the observed phenotype.

The RNAi effect has branded RDK2 as a key regulator in the intricate signalling networks controlling life cycle progression of *Trypanosoma brucei* bloodstream form cells. RDK's presence in the slender trypomastigote form seems to function as a molecular break for acquisition of many features of the insect-infective stage of the life cycle. When released, resulting parasites look like procyclic forms but are unable to survive and divide in their standard culturing conditions. It is interesting how, despite still not getting dense, these cells survive for over a week in conditioned SDM79. In this environment, they acquire a plastic adhesive phenotype, which is more characteristic of later stages in the insect infection.

Overexpression of RDK2 blocks the most exploited method of differentiation for monomorphic parasites *in vitro*. Our results may suggest RDK2 has the capacity to inhibit 8-pCPT-cAMP's ability to produce stumpy* forms. However, we need to use stumpy-specific markers such as PAD to confirm this. Our battery of flow cytometry experimental conditions give indications that RDK2 inactivation and depletion may be a downstream effect of the CCA-TbPPT1 inhibition axis, and that its inhibition alone leads to a frustrated differentiation.

A lot more research is left to be done in order to explore RDK2's role in the cell. Localizing the protein along the differentiation programme, and finding its interaction partners, are critical next steps. Understanding the metabolic implications of the RDK2 phenotype, would be also very interesting. Until a more convincing validation is achieved of the impact that RDK2 kinase activity has on the phenotype we cannot prove that a kinase inhibitor would have value as a drug target. This is the object of a more thorough discussion in the next chapter.

Chapter 5 Mechanism-based phenotypic screen for RDK2 inhibitors

5.1 Introduction

5.1.1 Rationale

As described in the previous chapter and by Jones et al. (2014), RDK2 has a proven protein kinase activity and is essential for survival of *Trypanosoma brucei* bloodstream form *in vitro* and *in vivo*. Uniquely, RNAi knockdown of RDK2 also results in spontaneous differentiation to the insect-infective procyclic form 36 h prior to cell death. We reasoned that this unique phenotype should be amenable to a mechanism based phenotypic screen using protein kinase inhibitor libraries, where inhibition of RDK2 would be detected via the expression of EP procyclin, a marker for differentiation. In the previous chapter an antibody targeting this surface antigen was used extensively in immunofluorescence assay and flow cytometry. These approaches could be adapted to a screening set up using high content cell imaging, measuring fluorescence in a plate reader, or using plate-based flow cytometry. Protein kinase inhibitors causing both cell death and differentiation would be perfect candidates to test against the recombinant protein to confirm on target mechanism of action. An RDK2 inhibitor would be both a useful research tool (permitting pathway dissection) and a potential starting point for drug development.

5.1.2 Anachronistic differentiation to procyclic form as a therapeutic strategy

Natural and synchronous differentiation to procyclic forms happens more readily starting from cell cycle arrested stumpy forms (Rico *et al.*, 2013). They are the key stage for transmission to the tsetse fly, while slender replicative forms are the relevant form of the parasite during establishment of the mammalian infection. Slender forms are still able to differentiate when subjected to CCA, mild acidic conditions or protease digestion treatments but they do so asynchronously and in a lower proportion than stumpy forms (Rico *et al.*, 2013). The parental line used to create our RNAi mutants (2T1) is monomorphic as it

lost its ability to differentiate into the stumpy form. Since they are closer to the life cycle stage that is critical for the infection, it makes differentiation and death triggered by RDK2 depletion more relevant for therapeutics. Cell death caused by RDK2 inactivation would be enhanced by the loss of VSG, concomitant with EP procyclin expression, resulting in parasite clearance mediated by the innate and adaptive immune responses (Schwede *et al.*, 2015).

Interestingly, the concept of triggering anachronistic differentiation to procyclic form in the bloodstream as a chemotherapeutic strategy has been proposed before at the Swiss Tropical and Public Health Institute (Swiss TPH) (Wenzler *et al.*, 2016). Another monomorphic transgenic *T. brucei* line (GUSone), introducing β-glucuronidase (GUS) as a reporter under control of the EP procyclin promoter, was screened against a collection of compounds. These compounds were selected for their structural similarity with citrate and cis-aconitate (CCA), their capacity to act as iron chelators (mimicking CCA function), or their known value as MAPK inhibitors. MAPK5 has been shown to be involved in slender-to-stumpy differentiation (Domenicali Pfister *et al.*, 2006), so the authors reasoned that other MAPK might be involved in further progression of the life cycle. That screen was conducted in parallel to an alamar blue cell viability assay. As a result, 28 compounds were found with activity in the nanomolar range and all of them were capable of killing the parasites *in vitro*. The possibility of EP promoter activation being a stress response upon exposure to trypanocidal drugs was discarded as only ~9% of the lethal compounds produced GUS activation. Still unpublished, the same screening pipeline has been applied to the 1.8 million compound collection of GlaxoSmithKline (GSK) selecting 302 compounds able to induce GUS expression in the lower micromolar range (Tanja Wenzler personal communication).

5.1.3 High throughput screening campaigns targeting protein kinases of *Trypanosoma brucei*

PKs of *T. brucei* have been shown to be druggable (Urbaniak *et al.*, 2012). Being a PK, the prospect of identifying an RDK2 inhibitor with drug-like properties is promising. ATP analogues unable to transfer the phosphate have been exploited to compete with endogenous ATP for the active site causing PK inhibition (type I inhibitors), although allosteric inhibitors (type II-IV) -which do not directly block

ATP binding- can work with higher potencies as type I must compete with the relatively high intracellular concentrations of ATP (1-10 mM) (Gribble *et al.*, 2000). The PK domain is often accompanied by other structural features, which permit the design of molecules to achieve specific targeting of the active site in a particular PK. Conservation of the kinase domain through evolution has not hampered identification of inhibitors able to act with specificity, although it is more crucial to find parasite *versus* host specificity than intra-parasite specificity (Merritt *et al.*, 2014).

Knowledge and historic data available about chemical structures active against homologue PKs in other organisms (mainly humans) have encouraged design of drug repurposing strategies (Pollastri and Campbell, 2011). The proof of concept in *T. brucei* was a study testing 8 compounds in pre-clinical or clinical stages targeting mammalian phosphoinositide-3-kinase (PI3K) or target of rapamycin (TOR) for inhibition (Diaz-Gonzalez *et al.*, 2011), two enzymes related to cellular growth and proliferation. There, a compound (NVP-BEZ235) had sub-nanomolar potency against the parasite both *in vitro* and *in vivo*. However, the exact mode of action in the parasite has not been identified yet. The same approach, where compounds with known activity against human targets are tested and optimized for efficacy/specificity based in SAR studies has been applied to Aurora kinase inhibitors (Jetton *et al.*, 2009; Li, Umeyama and Wang, 2009; Ochiana *et al.*, 2013), and protein-tyrosine kinase inhibitors (Katiyar *et al.*, 2013; Patel *et al.*, 2013; Behera, Thomas and Mensa-Wilmot, 2014).

Later a larger screening campaign tested 2979 compounds with structural analogy, >70% Tanimoto similarity (Bajusz, Rácz and Héberger, 2015), to the mentioned PI3K/TOR inhibitors; the protein kinase inhibitor set (367 compounds) published by GSK (Dranchak *et al.*, 2013); and another 39098 compounds in their collection (Diaz *et al.*, 2014). There, 242 compounds inhibited cell growth and 137 resulted in lethality for the parasite with 100-fold selectivity when compared to human HepG2 cells. All of them had potencies in the sub-micromolar range of concentration.

In addition, another recent publication showed how the entire 1.8 million GSK compound collection was screened for cidal activity against *Leishmania donovani*, *Trypanosoma cruzi* and *Trypanosoma brucei* (Peña *et al.*, 2015). For

each of the parasites a primary screen was conducted followed by a confirmatory replicate with the anti-kinetoplastid hits, an orthogonal assay for validation, an *in silico* filtration process based on drug-like physicochemical properties, a dose-response titration to select those with sub-micromolar potency, and an acute cytotoxicity test against human cells. After this campaign, an anti-*T. brucei* collection of compounds was selected, the "HAT box", containing 192 lead-like chemicals with trypanocidal activity in the nanomolar range and drug-like physicochemical properties (including potential central nervous system penetration). It is encouraging that >90 compounds contained there, had a history of sub-micromolar potency as protein kinase inhibitors in the records of the company.

The Dundee focused kinase inhibitor collection (3885 compounds), has been also used in a phenotypic screen against *T. b. rhodesiense* identifying 121 compounds causing >50% inhibition at 5 μM. These hits were screened against *T. b. brucei* finding three compounds that shared a chemical scaffold and had submicromolar potency and selectivity against human cells. Optimization of these compounds to achieve enhanced metabolic and physicochemical properties led to two promising series of compounds: 1H-imidazo[4,5-b]pyrazin-2(3H)-ones, with cytostatic properties; and 1H-pyrazolo[3,4-b]pyridines, with cytocidal properties (Woodland *et al.*, 2015).

5.2 Research aims

The aim of the experiments presented in this chapter was to identify inhibitors of RDK2, able to provide a chemical validation for the RNAi phenotype: differentiation to procyclic-like forms followed by cell death at 37^{o}C in HMI11. A specific RDK2 inhibitor would have value as a tool to dissect RDK2 signalling without the limitations of RNAi in transcriptionally arrested cells. This would enable work with pleomorphic parasites, able to produce stumpy forms. Such a compound would be also a lead-like chemical with potential for drug discovery. A major output for the project would be the identification of a molecular target for one or more hits within the screening campaigns described in the introduction.

5.3 Results

5.3.1 Compound selection

We reasoned that some RDK2 inhibitors were likely to be found among the sets of compounds described in the introduction to this chapter. They were either potential kinase inhibitors, or agents causing *T. brucei* death/differentiation. Interestingly many of them would have already been selected for their cidal activity in the lower micromolar and nanomolar range, parasite specificity against human cells, and filtered for their drug-like physicochemical properties. We initiated collaboration with researchers at Swiss TPH and GSK in order to test their compounds using our assay. After signing a material transfer agreement, we obtained 518 compounds including 188 members of the HAT box, where many had a history of protein kinase inhibition (Peña *et al.*, 2015); 45 kinase inhibitor chemotypes (Diaz *et al.*, 2014); and 283 GUS inducing compounds (still unpublished), selected in the same screen format as described in the introduction above (Wenzler *et al.*, 2016). 20 μl of each compound were shipped in two 384-well storage plates at 1 mM concentration in DMSO.

5.3.2 High content cell-based screen optimization

We have already shown in the previous chapter that EP procyclin expression is a surface marker of procyclic differentiation that can be traced with a specific antibody conjugated directly or indirectly to a fluorophore (Richardson *et al.*, 1988). As RDK2 RNAi phenotype involved bloodstream-to-procyclic form differentiation before killing the cells, we reasoned that such antibody could be exploited as a reporter in a primary screen for compounds inhibiting RDK2. If able to reproduce the RNAi phenotype, such compound would trigger EP procyclin expression before causing cell death.

5.3.2.1 Fluorescence detection on a plate reader

At first, we considered to analyze the EP-linked fluorescence signal in a high content cell-based screen using a plate reader. *Trypanosoma brucei* bloodstream forms reach the stationary phase of growth around 10^6 parasites x ml^{-1}. It cannot reach concentrations over 4 x 10^6 cells x ml^{-1} and concentration increases 10 times every 24 h when they are in the logarithmic phase of growth, Figure 4-2.

The use of a 384-well format, where the maximum volume capacity is 40 μl, was discarded as the reduced amount of cells required to seed a 48 h-long incubation in such platform (400 cells/well, at 10^4 cells/ml) could risk uneven distribution of cells across the plate. 96-well plates would provide a more robust distribution and was an acceptable setting as we were to screen only 518 compounds.

The protocol we used in the previous chapter for flow cytometry and IFA entailed the use of two antibodies (a mouse IgG anti-EP procyclin, and an antimouse IgG1 allophycocyanin-conjugated), and a fixation step with paraformaldehyde (PFA). We adapted the protocol to use with an automated liquid handler (Beckman BioMek FXp©), available at the Scottish Bioscreening Facility. After testing different makes of plates, with diverse well shapes and types of plastic, we chose 96 well V-bottom sterile Cellstar® (Grainer Bio-one, Cat. Num: 651180). V-shaped bottom enabled a better recovery of cells after centrifugations in the staining protocol than round or flat bottoms. They concentrated the parasite pellet and reduced the volume left in the plate ("dead volume") after supernatant removal. Growth curves performed in these plates with 2T1 parasites confirmed that parasite cell density increased at the same rate as they did in flasks or larger plates, Figure 5-2C. After staining, cells were transferred to black 96-well plates with flat bottom before measuring in a plate reader.

A test run was conducted using RDK2 RNAi line (sTL629), tet-induced and uninduced, plating the cells 48 h before fluorescence analysis (6 replicates), Figure 5-1A. Unfortunately, overlapping of differentiation and death phenotypes quenched the fluorescence signal. As the number of cells in the Tet-induced wells had decreased in comparison with the uninduced (see RDK2 RNAi phenotype in the previous chapter, Figure 4-2), the EP procyclin-associated fluorescence signal could not be distinguished from the background (Figure 5-1A). To confirm this, the same experiment was performed, incubating the cells in separate flasks for 48 h, counting them, and plating equal numbers before proceeding with antibody staining and measurement in the plate reader, Figure 5-1B. As expected, a significant increase in fluorescence associated to EP procyclin expression could be detected in the Tet-induced wells.

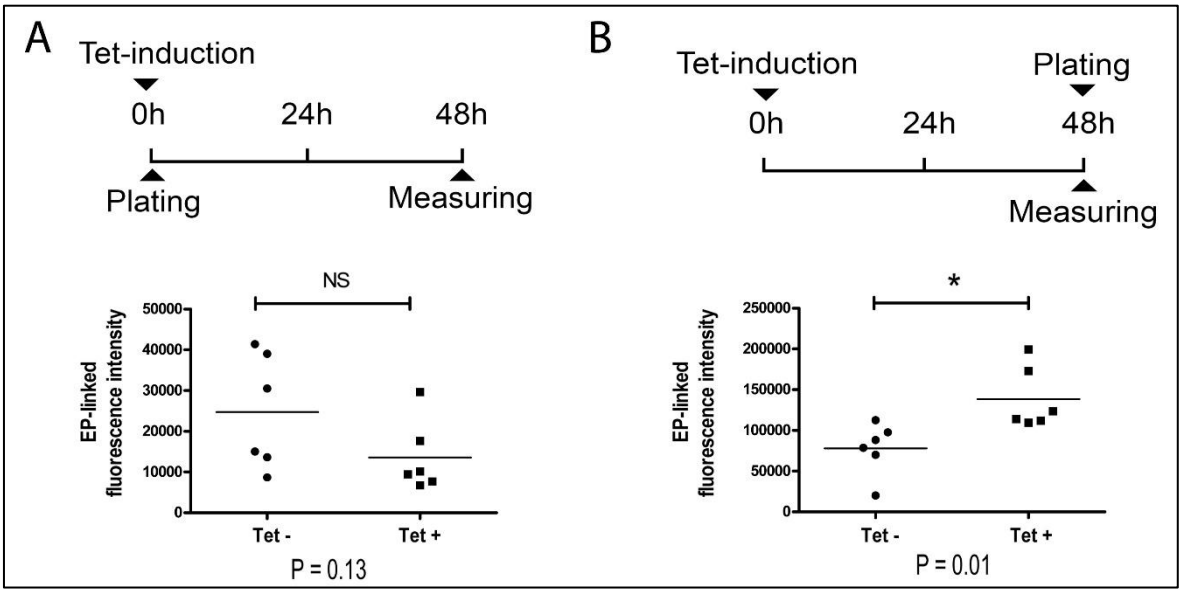


Figure 5-1. Differentiation phenotype caused by RDK2 RNAi as observed in a plate reader using an anti EP procyclin antibody.
Cell death quenched fluorescence intensity in RNAi-differentiated cells. **A.** Plate reader results of endpoint intensity for indirect immunofluorescence (EP-Alexa488), uninduced (Tet-, n=6), and 48 h induced (Tet+, n=6). Cells plated at 10^4 parasites x ml^{-1} 48 h before reading. **B.** Same as A, but plating equal number of Tet+ and Tet- cells just before reading. Comparisons made with paired T tests (* = $P<0.05$, and ns= $P>0.05$).

5.3.2.2 Fluorescence detection on a flow cytometer

In the previous chapter, we used flow cytometry to assess differentiation via detection of fluorescence associated to the anti EP-procyclin antibody. In flow cytometry, fluorescence intensity is assessed at a single-cell level and for a fixed number of recorded events. We reasoned that this would circumvent limitations described for the plate reader when differentiation and cell death phenotypes overlapped. If as many cells as were necessary to match the set number of recorded events remained alive in the wells, we could still establish meaningful comparisons. In addition, EP procyclin positive results would be expressed as a percentage of the population. It would make different wells comparable regardless the number of recorded events. Availability of a flow cytometer equipped with a 96 well plate sampler (Miltenyi Biotec MACSQuant®) at the Wellcome Trust Centre for Molecular Parasitology permitted the adaptation of the EP procyclin expression assays shown in the previous chapter to a 96-well plate format.

Incubation time

We assumed that the phenotype caused by protein kinase inhibition would be detectable sooner than the phenotype caused by RNAi. A cell permeable compound would block RDK2 activity almost immediately, whilst RNAi-mediated transcript depletion would only prevent production of new protein. It would require depletion of available cellular stocks before producing a phenotypic effect. As a reference, EP procyclin is detectable 6 h after treatment in monomorphic cells differentiated *in vitro* with cis-aconitate (Bütikofer *et al.*, 1999). A compromise between detectable levels of EP procyclin and sufficient numbers of viable cells needs to be met. Since RNAi-triggered EP procyclin expression was traceable from 36 h onwards and cell death started after 48 h, the chosen incubation time was 24 h.

Culture volume and layout

T. brucei culture conducted in chapter 4 was carried out always in volumes equal or larger than 5 ml. For this screen, assay volumes were reduced to 100 μl in order to minimize the amount of compound consumed, and maintaining concentration of their solvent (DMSO) under 1%. Previous tests performed in our lab showed that 1% DMSO caused no harm to the parasites. 100 μl instead of 200 μl left also extra room for the 5% CO_2 atmosphere.

A schematic representation of the plate preparation is shown in Figure 5-2A, including the lay out, Figure 5-2B. 2T1 bloodstream forms were used both as the negative control (8 wells, incubated with 1% DMSO only) and the screening subject (incubated with the compounds, also at 1% DMSO). The positive control was RDK2 RNAi line (sTL629) induced 48 h prior to flow cytometry analysis and plated only for the last 24 h. 8 wells were seeded (at the same time than 2T1 cells) at $1x10^6$ parasites x ml^{-1} because they experience growth arrest between 24-48 h of RNAi induction (chapter 4). The automated flow cytometer (MACSQuant®) takes samples in a column-by-column basis, moving from A-H within each of them and progressing through the plate from left to right,Figure 5-2A. The bottom row of each plate (except for wells H1 and H12) was filled with 100 μl of flow cytometry cleaning solution. This enabled washing the equipment circuit once every 7 samples. These arrangements left room for 70

compounds to be tested per plate and gave us capacity to run 2 plates per day, with an average sampling time of 4h per plate.

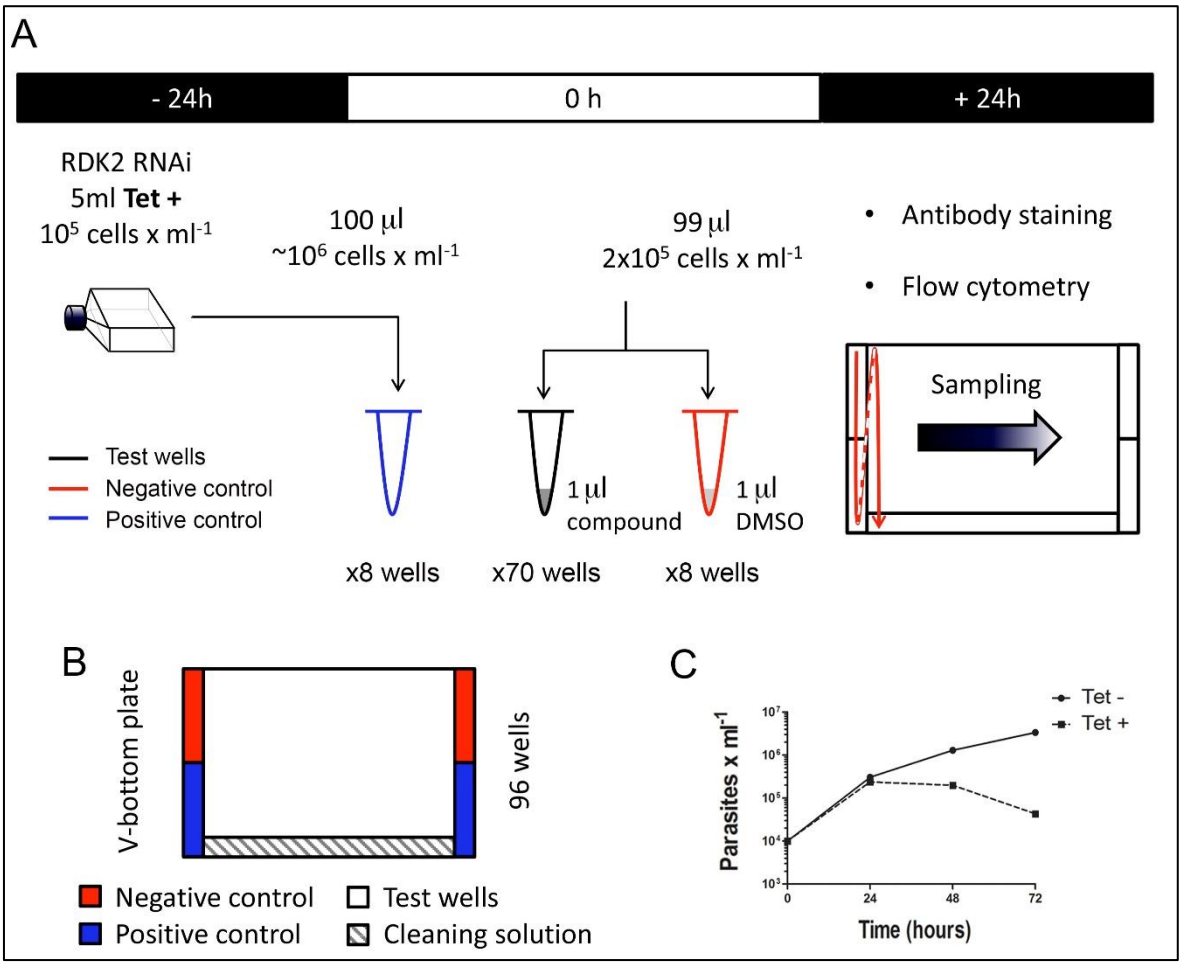


Figure 5-2. Flow cytometry screen plate processing strategy.
A. Schematic representation of the primary screen time frame including: induction of the positive controls 24 h prior to plate layout; compound stamping and plating of cells (0 h); and antibody staining and flow cytometry screen after 24 h incubation. Sampling direction highlighted. **B.** Plate layout. **C.** Growth curve of the RNAi line (Tet+/Tet-) in one of the 96 well V-bottom sterile Cellstar® plates used in the screen.

In *T. brucei* 2T1 (or uninduced derived lines) cell numbers grow 10 times every 24 h in HMI11 at 37°C, also in the 96-well V-bottom plates used for this screen. They reached the stationary phase at around $2x10^6$ parasites x ml^{-1}, Figure 5-2C. With this in mind, we seeded our assay plates at $2x10^5$ parasites x ml^{-1} 24 h prior to flow cytometry analysis. In this way, we could maximize the number of cells per well at the sampling time point: 100 µl at $2x10^6$ parasites x ml^{-1}, i.e. $2x10^5$ cells in total.

Antibody optimisation

Antibodies are expensive consumables. Besides reducing the volume needed per well, optimizing the staining protocol was critical. The immuno-detection strategy used for flow cytometry in the previous chapter was indirect. The use of one fixation step, two antibody incubations and profuse washing entailed high risk of losing cells. The primary antibody was a mouse IgG anti-EP procyclin (Richardson *et al.*, 1988), and the secondary was an anti-mouse IgG1 made in goat and conjugated to the fluorophore allophycocyanin (APC). APC can be excited at 594 and 633 nm having a fluorescence emission peak at 660 nm.

In Figure 5-3, we show different antibody conditions (A-D) that were tested using: 2T1 bloodstream forms, as a negative control for EP procyclin expression; Lister 427 wild type procyclic forms, as a positive control for procyclic expression; differentiated 2T1 cells upon incubation 24 h with 100 μM pCPT-cAMP and 24 h with 3 mM citrate-cis-aconitate (CCA); RDK2 RNAi cells differentiated by 48h induction with 1 μM tetracycline. There, we show that using fluorescein isothiocyanate (FITC) or equivalents such as Alexa Fluor®488 (490 nm excitation/525 nm emission), produced a reduced signal:background ratio when compared to APC. This may happen due to parasite autofluorescence, which has been observed in the green channel of detection using immunofluorescence microscopy upon fixation. It is unfortunate as a primary anti EP procyclin antibody directly conjugated to FITC was commercially available and could have reduced the number of incubation and washing steps. The concentration of this antibody required to obtain a useful signal:background window was too high (1:100), Figure 5-3.

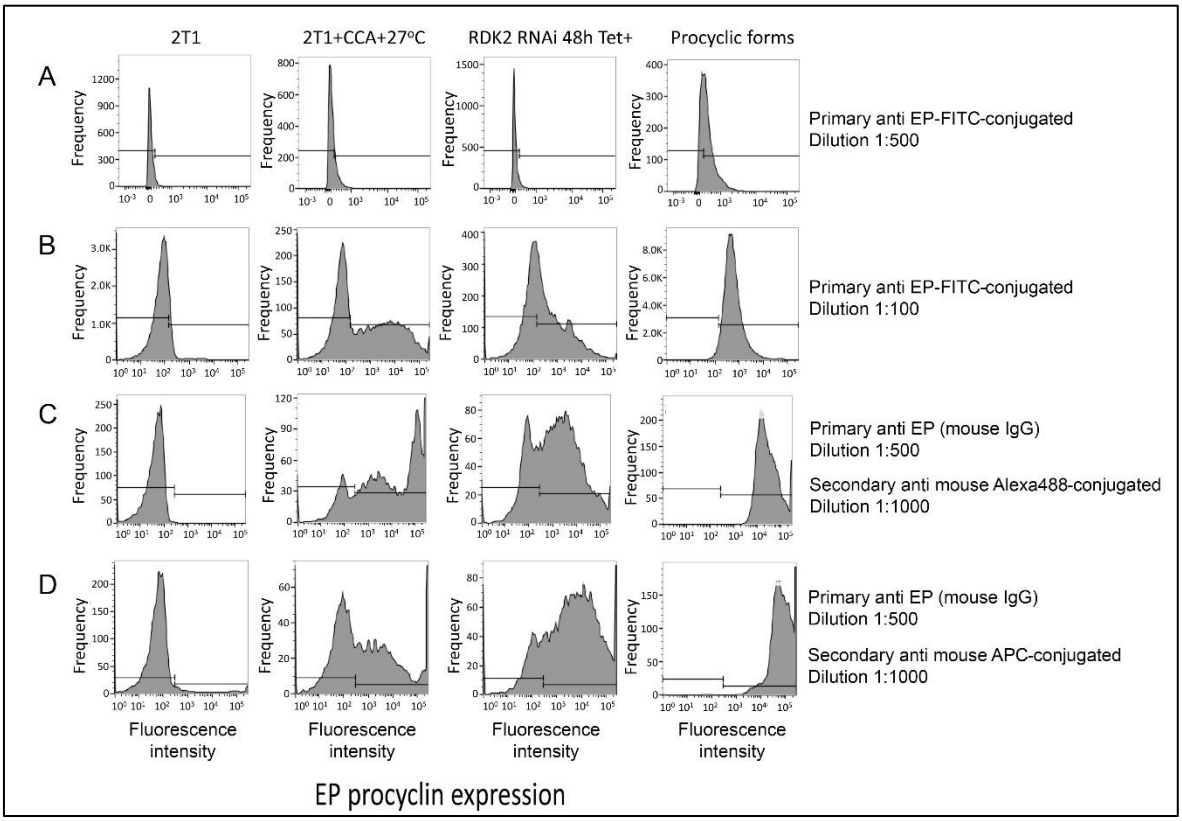


Figure 5-3. EP procyclin antibody optimization.
Histograms showing number of events detected in flow cytometry (frequency) *versus* fluorescence intensity associated to EP procyclin expression. Data analysed with four different antibody staining conditions (**A-D**, annotated in the right hand side); and for four different type of cells: wild type 2T1 bloodstream forms; 2T1s differentiated to procyclic forms with pCPT-cAMP + citrate/cis-aconitate (CCA) at 27ºC; RDK2 targeting RNAi line induced for 48h; and Lister 427 procyclic forms.

We diminished the antibody volume used in both incubations, from 100 μl to 50 μl per well at the staining dilution (1:500 for the primary, and 1:1000 for the secondary). And we reduced from 3 to 2 the number of washes after, including a previous large-dilution step adding 150 μl of PBS over the 50 μl of antibody.

Cell fixation

Large-scale experiments described in the previous chapter, relied on 2% paraformaldehyde treatment in order to fix the cells prior to flow cytometry analysis. A challenge encountered during protocol optimization was cell clumping and adhesion to the walls. These problems caused abundant loss of cells during transfer of the final parasite suspension into the flow cytometry

tubes by filtration through a Nitex mesh membrane. This step is recommended as cell clumps can block the microfluidics system in the flow cytometer or hinder identification of events associated to individual cells. The problem of plastic adhesion was solved first by fixing at the end, after the antibody staining, with cells already in the destination tubes. But still it involved cell clumping that would entail cell loss during the filtration process, or would risk obstructing the equipment. Given the small cell numbers we handled in the 96 well setup, we decided to miss the fixation step by keeping cells alive on ice and cold during the whole screening process.

Once optimized, this system ensured that it was possible to record 6000 parasite-associated events in flow cytometry for the 2T1 negative controls. At the growth rates observed in Figure 5-2C, these wells would contain ~10^5 cells 24 h after seeding them at 2 x 10^5 parasites ml^{-1}. This means ~90% excess, providing a margin to record 6000 events also in wells treated with trypanocidal compounds.

5.3.3 Medium throughput screening workflow

After completing the assay development, we reasoned that many resources generated in the previous chapter would be useful in a screen for RDK2 inhibitors. In addition to the adapted flow cytometry assay -which would be the basis for our primary screen- other tools might be applied in subsequent validation steps: the RDK2 overexpressor line, the RDK2 active recombinant protein and the α-casein phosphorylation assay. The planned workflow after the assay development phase is represented in Figure 5-4.

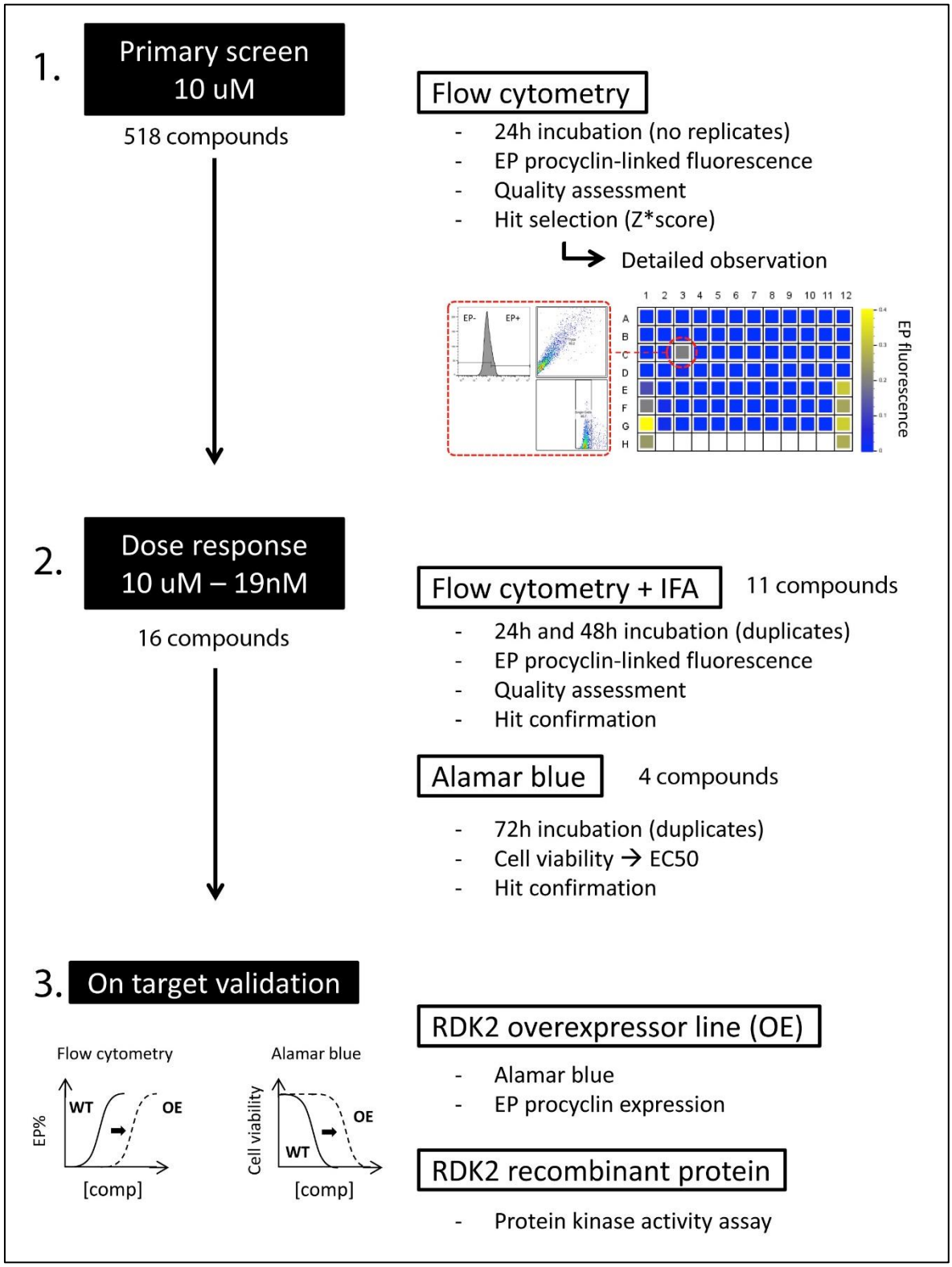


Figure 5-4. Schematic design and summary of the screening for RDK2 inhibitors.

The original design involved two primary 96-well plate based flow cytometry screens at two different concentrations (10 μM and 1 μM). Unfortunately, time

restrictions limited our schedule to a single 10 μM campaign. Selected hits were subject of a confirmatory flow cytometry replicate but with some variations. Rather than a single 10 μM repeat, serial dilution was applied to analyse the dose dependent response after incubation of compounds for 24 h and 48 h. Each of these assays was carried out in duplicate.

In order to provide further validation of the EP procyclin signal, one of the replicates at each incubation time was analysed in an orthogonal immunofluorescence assay performed with a high content imaging system (GE IN Cell 2000®). Hoechst 33342 was added to stain DNA and facilitate focusing on the cells.

Finally, EC50s were calculated upon 72 h of compound treatment using alamar blue assay (Räz *et al.*, 1997). Time restriction did not allow us to pursue this but in the original design, EC50 values would have been calculated also for compounds in an RDK2-overexpressor line. If higher doses were required to produce the same inhibitory effect observed in 2T1 parasites, compounds would be tested for their capacity to inhibit recombinant RDK2 α-casein kinase activity.

5.3.4 Stock plate preprocessing

Stock plates were delivered to us with 20 μl of compounds at 1 mM, in a layout including column 6 filled with DMSO and column 18 empty. Such distribution required DMSO disposal and reallocation of certain compounds to ease stamping them in the final screening plates, Figure 5-5A.

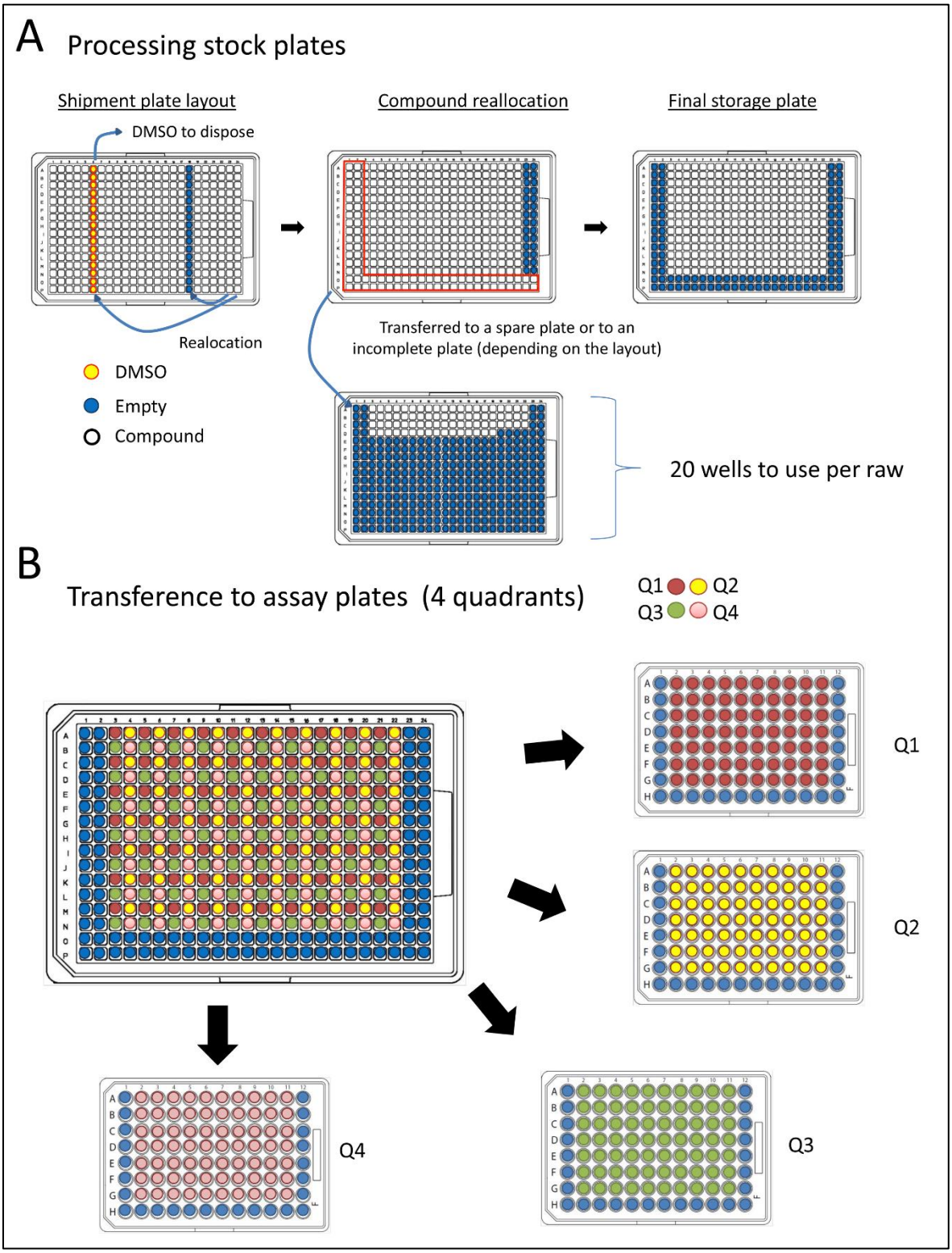


Figure 5-5. Schematic representation of the plate processing.
Two 384 GSK plates were shipped containing 20 µl of compound at 1 mM. Column 6 contained DMSO and column 18 was empty. Plates were rearranged **(A)** to permit direct transfer by quadrants **(B)** to the assay 96 well plates.

A 96-tip head available in our liquid handler equipment enabled transfer of a complete 384-well plate to four 96-well plates by quadrants. To obtain the

desired layout at the screening plates, columns 1, 2, 23 and 24 and rows O and P, needed to be empty. A schematic representation of this processing is displayed in Figure 5-5B. 'Mother' 384-well plates generated after these manoeuvres were labelled *GSKV2_P1* and *GSKV2_P2*. Derived 'daughter' 96-well plates stamped for the screen were labelled *GSKV2_P*X*_Q*Y. In this code, 'X' (1 or 2) refers for the 'mother' plate, and 'Y' (1-4) refers to the quadrant stamped.

5.3.5 Primary screen

Data acquisition

We had the capacity to screen two 96-well plates per day. 24 h before flow cytometry, we stamped 1 μl of each compound at the original 1 mM concentration in the 'test' wells of both assay plates, Figure 5-2A.

MACSQuant® flow cytometer software was used during the data acquisition process to display events comparing the area of signals detected with the forward and side scatters. We could then gate events corresponding to the trypanosome population and, within such gate, select for single-cell events by representation of height versus width for the forward scatter signal. Displaying single-cell events in a histogram for APC fluorescence intensity, permitted a third gate to be set quantifying the EP positive population. The negative controls (2T1 cells) were used as a reference to establish the threshold. The instrument was primed to record between 6000 and 10000 events with these settings, Figure 5-6.

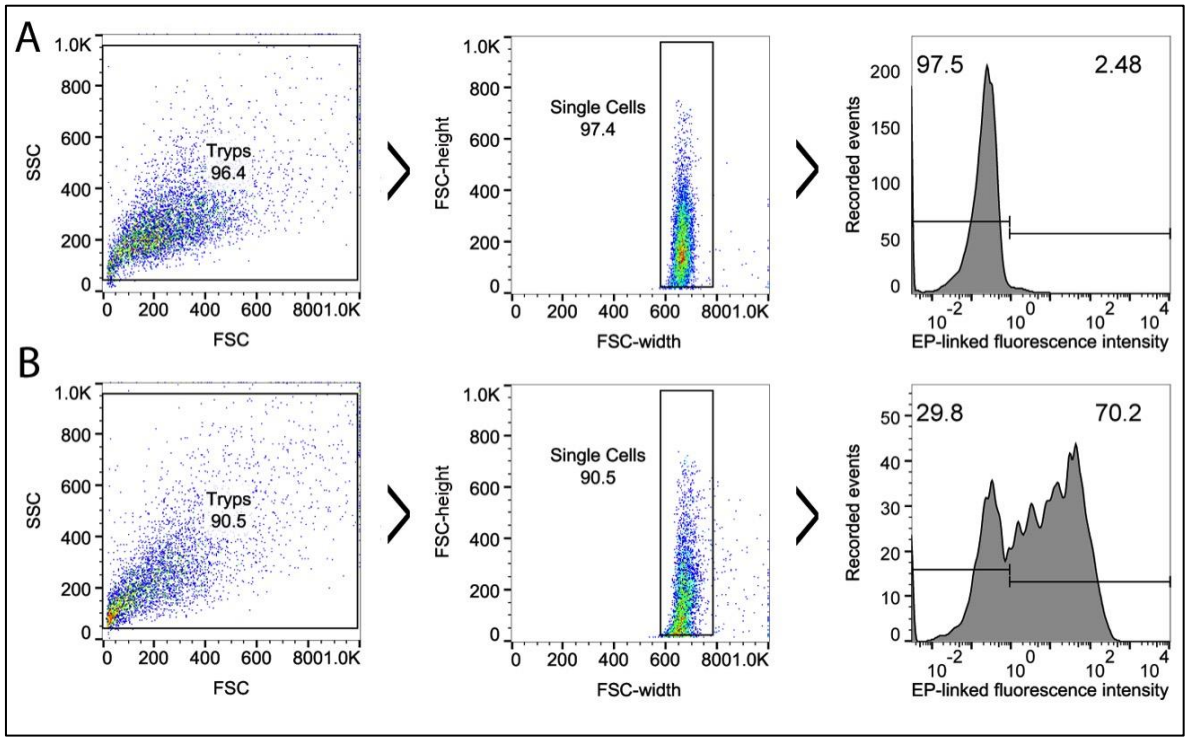


Figure 5-6. Gating strategy used for acquisition and analysis of the cellular events by flow cytometry.
Exemplified with uninduced **(A)** and induced **(B)** RDK2 RNAi line. From left to right: *Trypanosoma brucei* population was selected by area comparing forward and side scatter; selection of single cell events by comparison of the width and height of the forward scatter signal; and selection of the EP procyclin positive population in a histogram for fluorescence intensity of the excited fluorochrome APC, indirectly bound to EP procyclin. The threshold to gate the EP positive population was based on the negative reference, the uninduced sample.

Quality control

In every high throughput screening campaign, the assay read out is susceptible of variations natural to the biological system; human or instrumental errors, which can be systematic or occasional; and physicochemical properties of the compounds themselves. According to the literature, in order to assess reliability of the results statistic quality control approaches (Shun *et al.*, 2011) need to be undertaken. All analysed parameters indicated poor quality in the first three plates of this screen, Table 5-1.

Table 5-1. Primary screen quality control. Summary of results.

Plate ID	Normal distribution	Z´factor	%CV + ctrl	%CV - crtl	S:B	Median test % EP+	Z-score > 20
P1_Q1	Yes	0.13	33.36	27.56	15.64	3.30	0
P1_Q2	No	0.33	14.90	50.00	9.59	4.60	0
P1_Q3	No	0.36	21.07	38.11	141.08	0.38	4
P1_Q4	No	0.54	14.76	53.85	101.08	0.54	1
P2_Q1	No	0.65	11.26	26.88	111.72	1.05	4
P2_Q2	No	0.48	17.18	22.09	128.60	1.19	7
P2_Q3	No	0.88	3.50	39.67	96.69	1.41	0
P2_Q4	No	0.93	2.06	59.35	144.94	0.86	0

%CV, coefficient of variation; S:B, Signal/Background ratio; ctrl, control. Formulas depicted in Figure 5-7 and Figure 5-10.

Z´-factor is the most widely accepted coefficient used to assess robustness of a high throughput screen (Zhang, Chung and Oldenburg, 1999), Figure 5-7. On the one hand, it collects the dynamic range for the signal window: the absolute value for the difference between the medians of the positive and negative controls. On the other hand, it measures a margin of variation based on the intrinsic fluctuation of the positive and negative signals, represented by 3 times their median absolute deviation (MAD). Values < 0.5 are considered poor quality (Shun *et al.*, 2011). Plotted in the same graph, we present the coefficient of variation (%CV) of the positive and negative controls independently. %CV is a normalizing parameter defined as the percentage of the median of the controls that represents their MAD. Such normalization permits a more meaningful comparison between MADs calculated in different plates. %CVs < 20 are considered good quality values which can be said for the positive controls in most of our screen plates but not for the negative.

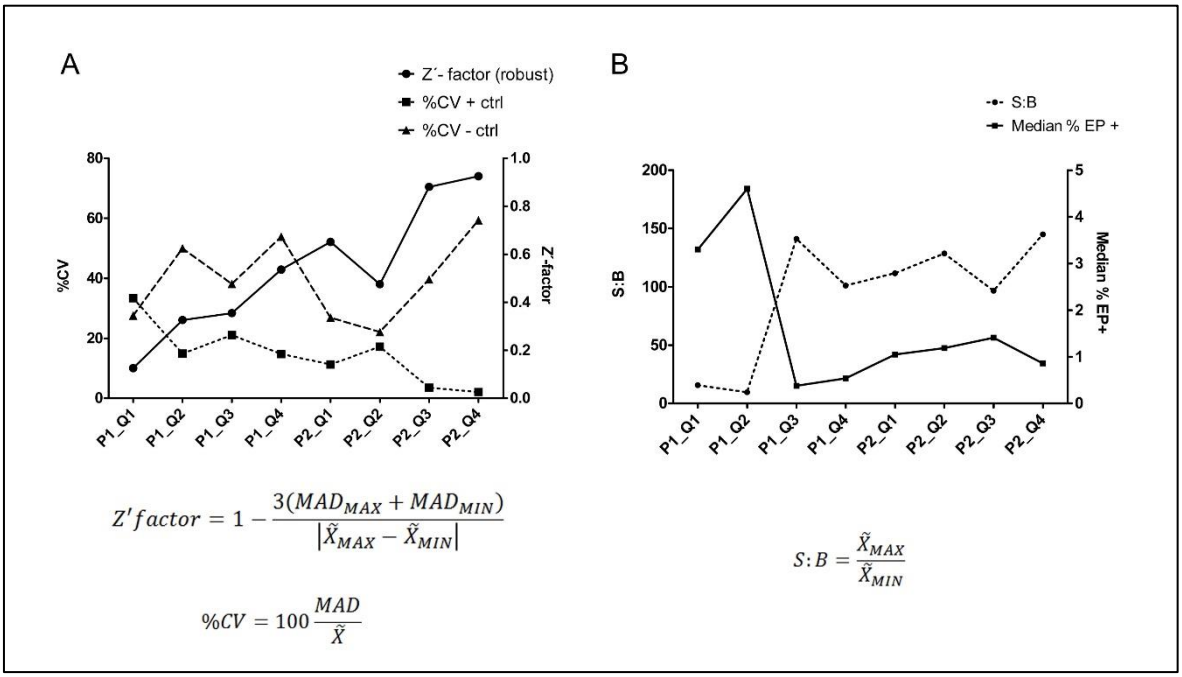


Figure 5-7. Primary screen quality control.
A. Coefficient of variation (%CV) of the negative and the positive references plotted per plate in the 10 μM screening campaign versus Z´factors. **B.** Signal/Background ratios (S:B) plotted against median of compound wells per plate. Formulas depicted under the graphs. MAD=Median absolute deviation; MAX= Positive control; MIN=Negative control; $\tilde{X}$=Median.

Another key feature of a robust assay is having a good separation between the median signals obtained for both the positive and negative controls. *Signal/background ratio (S:B)*, dividing median signal of the positive controls by the median signal of the negative controls, showed a narrow detection window in the first two plates which widens for the rest. Accordingly, *median %EP positive population of the "test" wells* (calculated per plate) followed the opposite trend highlighting a problem with the negative control in the first two plates.

Depiction of *number of events recorded* (and considered for analysis) across the complete screening campaign showed some inter-plate variability. However, for most samples at least 5000 events were recorded. Those with lower recorded events were likely due to cell death, Figure 5-8. *Heat maps for %EP positive population* (Figure 5-9) permitted observation of a systematic pattern in some plates, where positive controls sampled at the end (column 12, wells E-H) showed a less intense fluorescence, possibly due to progressive signal

attenuation. Processing each plate in the flow cytometer took around 4 h. Light exposure along this period may be a cause for photobleaching that could explain a difference in intensity between positive controls in columns 1 and 12.

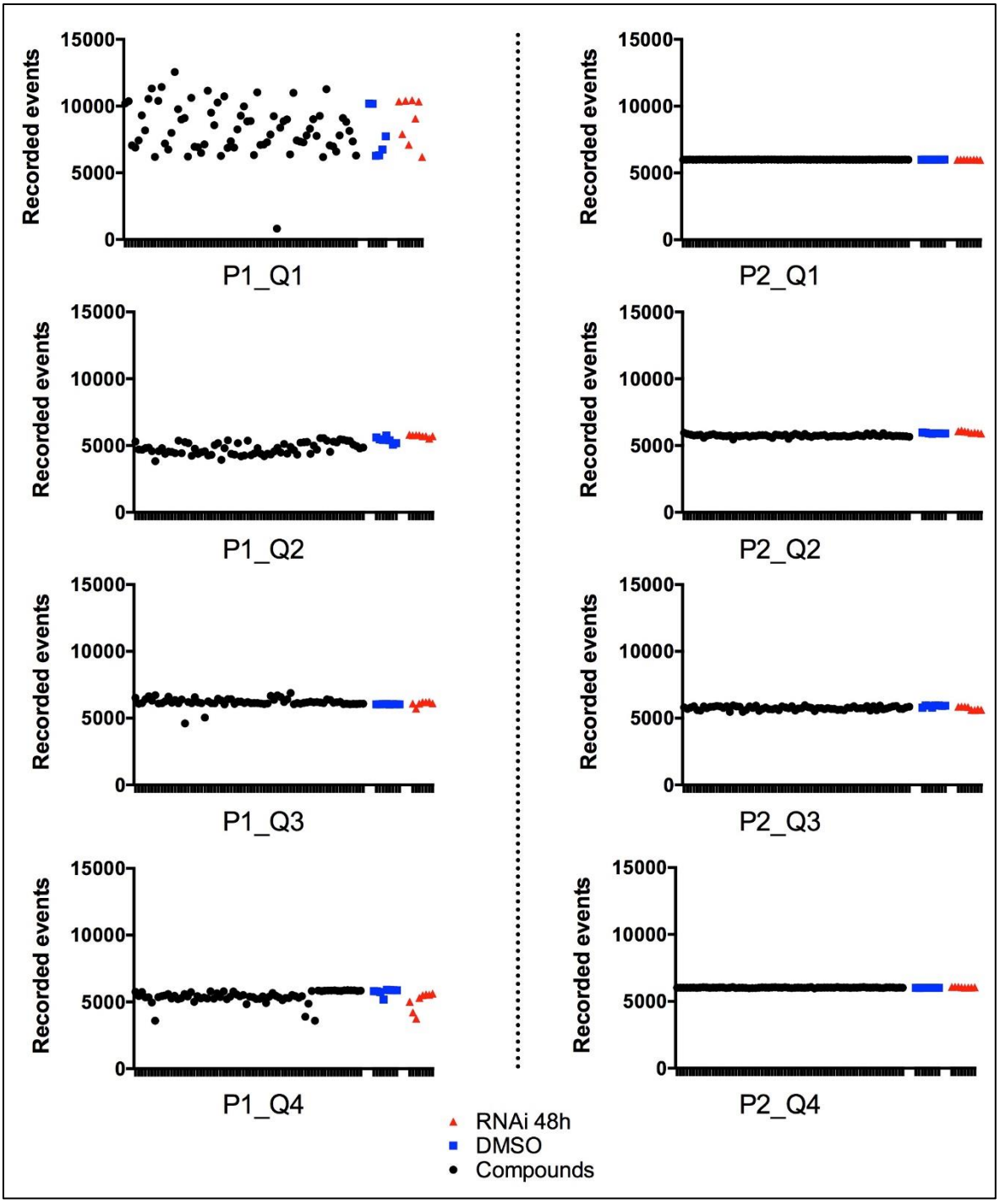


Figure 5-8. Number of flow cytometry events collected per well in primary screen plates

Hit selection

Data analysis was carried out in Flow Jo. We reproduced gating strategy applied in the acquisition process, and used the histogram of EP-linked fluorescence intensity (APC emission) in the negative controls as a reference to establish a threshold identifying the EP positive population. Percentage of parental gate statistics were extracted from all samples in this population. Heat maps for these values on plate views (Figure 5-9) and data tables were generated (Table S6).

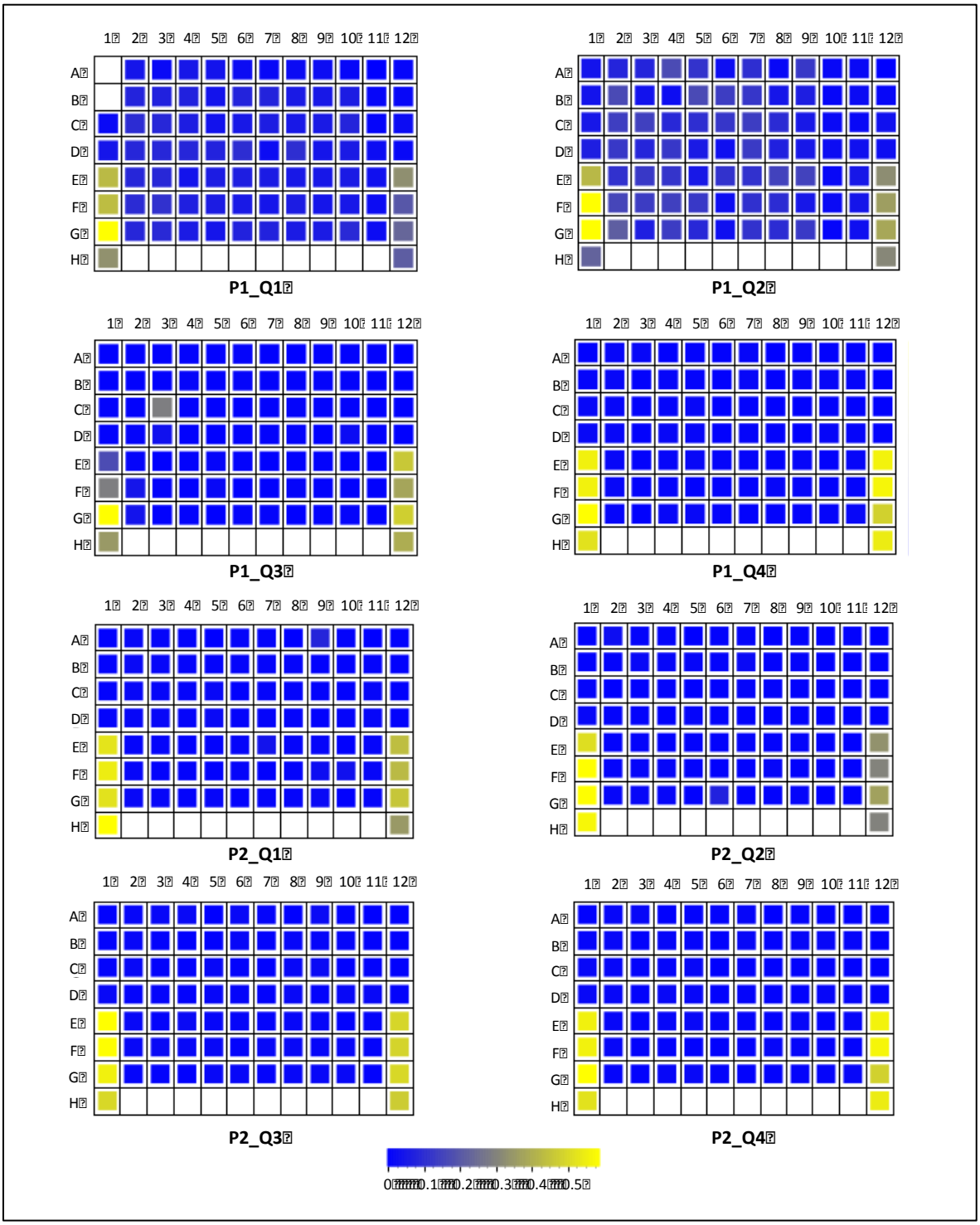


Figure 5-9. Heat maps of %EP positive population.
Normalized respect to the maximum signal in the plate.

Several analytic methods are commonly used in high throughput screening to perform hit selection (Shun *et al.*, 2011). "Z-score" is widely accepted in campaigns without replicates, as in our case (Zhang, 2011). It makes the assumption that the screen signal shows equal variability across the plate, so the

variability calculated for the negative controls (8 per plate) can be extrapolated to any other well. In the standard version, when the dataset has a normal distribution, it is defined as the distance of every data point from the mean, divided by the standard deviation of the negative control.

However, three different normality tests performed with Graph Pad Prism 6® showed that, with the exception of P1_Q1, none of the plates in our screen show a normal distribution. This is the most common observation in high throughput screening and, when this happens, a more robust version of this parameter, Z*-score, is recommended. It uses "median" instead of "mean" and "median of absolute deviation" (MAD) instead of "standard deviation" for the negative controls. MAD is the median of the absolute deviation of every data point to the median for the complete dataset.

In practice, although the measured values are not normally distributed in most of high throughput screens, true hits behave still like outliers (Zhang, 2011). Based in their clear separation (Figure 5-10), we considered "hits" to follow up in our screen those compounds presenting a Z*-score > 20.

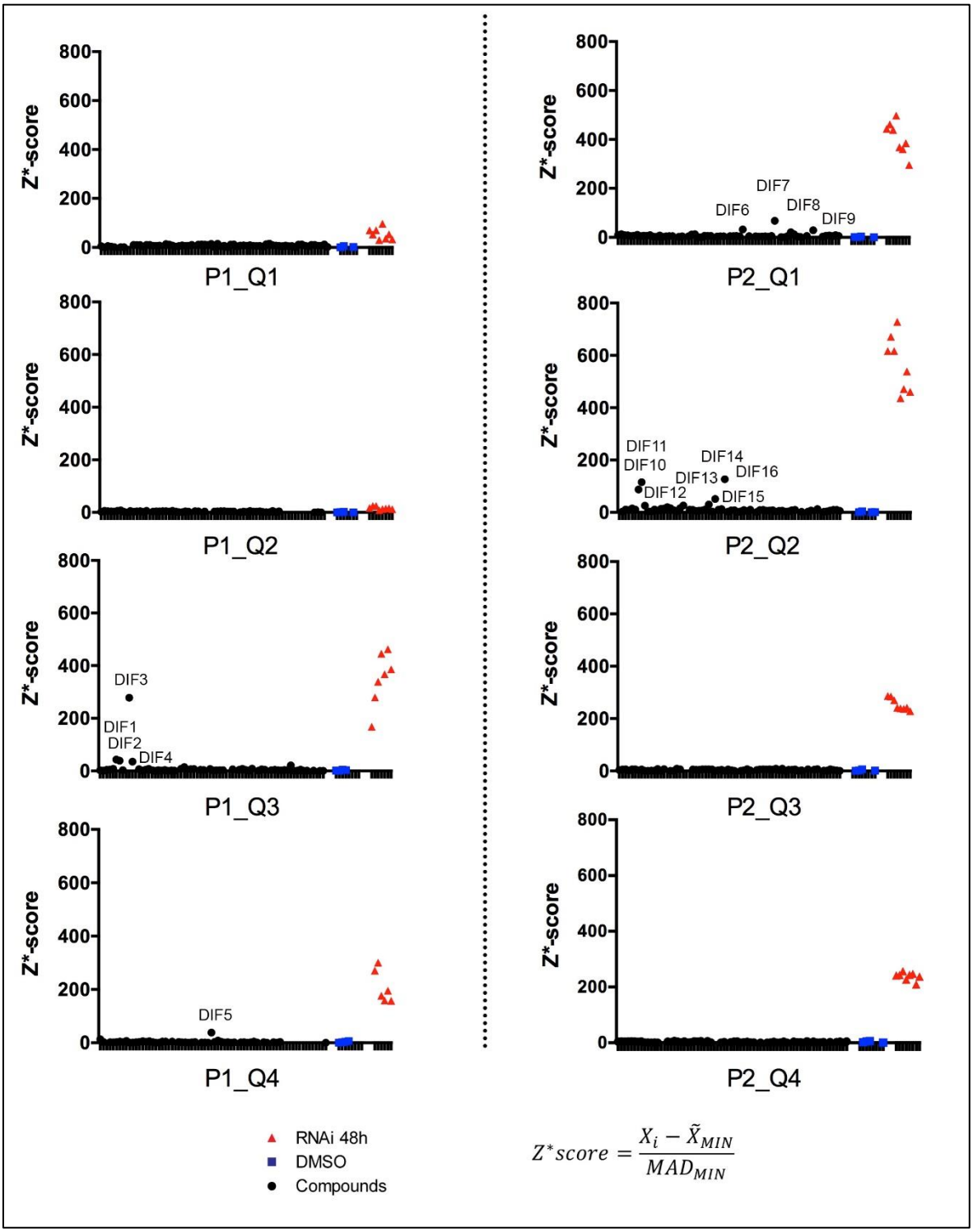


Figure 5-10. Z*score distribution of %EP positive values in the primary screen plates.
Highlighted the 16 "DIF" compounds considered for further analysis. Formula depicted under the graphs. MAD=Median absolute deviation; MIN=Negative control; X_i=%EP positive of a particular well; $\tilde{X}$=Median.

Under these premises, 16 compounds were considered for further analysis, Table S6. As they resulted from a differentiation assay they were called DIF1-DIF16.

The set included 5 HATbox compounds (Peña *et al.*, 2015), 7 protein kinase inhibitor chemotypes (Diaz *et al.*, 2014), and 4 GUS compounds, linked to procyclic form differentiation through the screening pipeline as described in (Wenzler *et al.*, 2016).

5.3.6 Validation

After considering the quality check summary in Table 5-1, none of the compounds selected came from plates P1_Q1 or P1_Q2, which were regarded as poor quality based on their Z´factor and the narrow detection window indicated by their reduced S:B ratio. Four were detected in P1_Q3 that was still low quality according to the Z´score. However the S:B ratio was a lot larger in that plate and the event number distribution was comparatively more homogeneous Figure 5-8. We decided to include these compounds (DIF1-4) in the set for follow up experiments.

Analysing the flow cytometry data

Analysing in detail the flow cytometry data for the selected 16 “DIF” compounds, Figure 5-11 and Figure 5-12, most of the histograms showed a low shoulder protruding into the %EP positive population. The exceptions were DIF3, where the whole population was synchronously shifted towards the higher fluorescence intensities, with 19.8% of events falling within the EP positive area; and DIF7 where the whole peak became wider with 8.89% of events displaying within the EP positive gate.

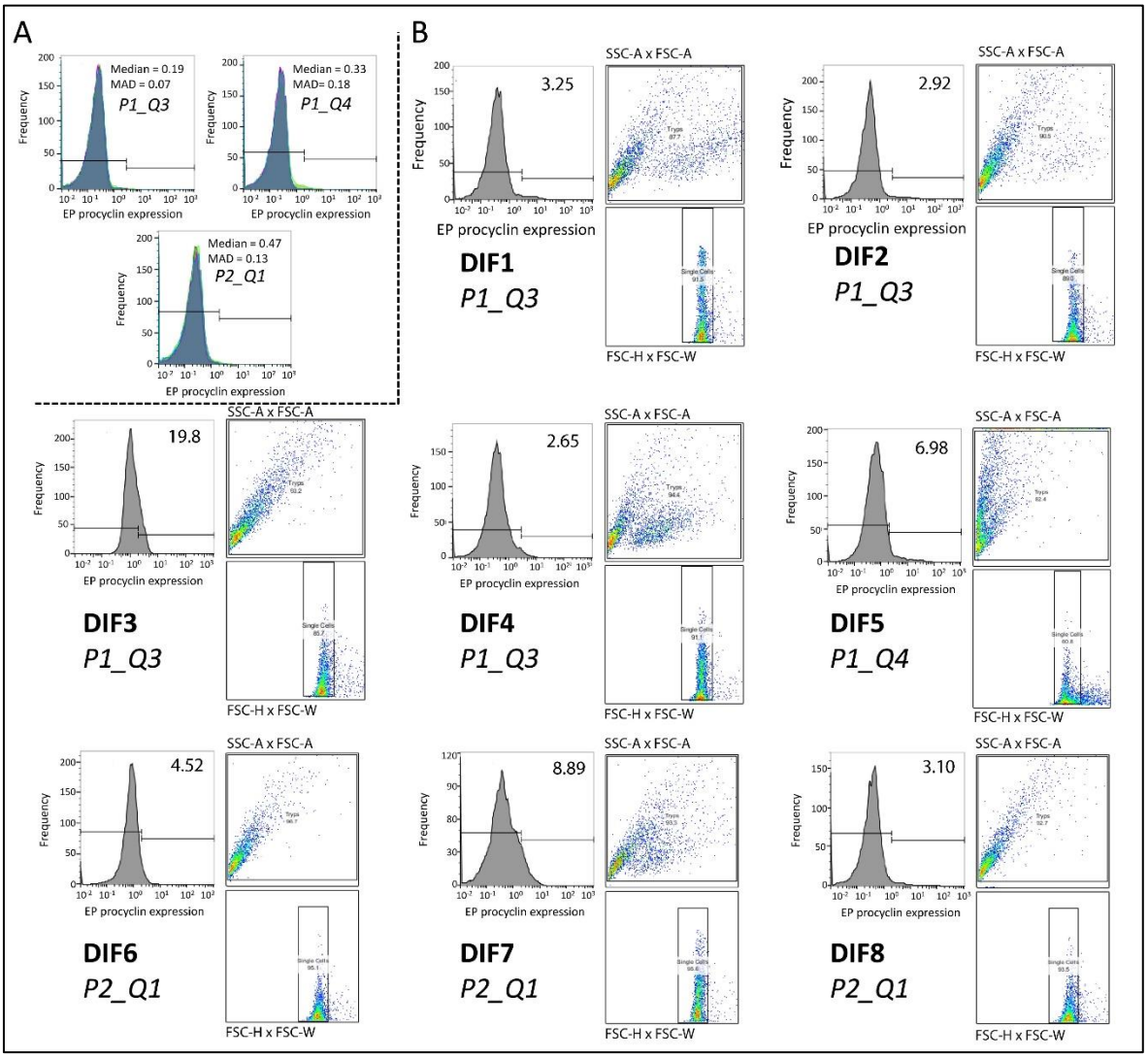


Figure 5-11. Detailed analysis of the flow cytometry results for DIF1-8. **A.** Superimposed EP procyclin expression histograms for the negative controls in each plate with a compound depicted in this figure. Their median EP procyclin positive population, and their median absolute deviation (MAD) are inscribed in the plots. **B.** Histograms of EP procyclin expression for the DIF compounds. EP positive % of the population is inscribed in the upper-right corner of the plots. Next to it, gatting strategy leading to that population (same as Figure 5-6): trypanosomes (SSC-A x FSC-A), and single cells (FSC-H x FSC-W).

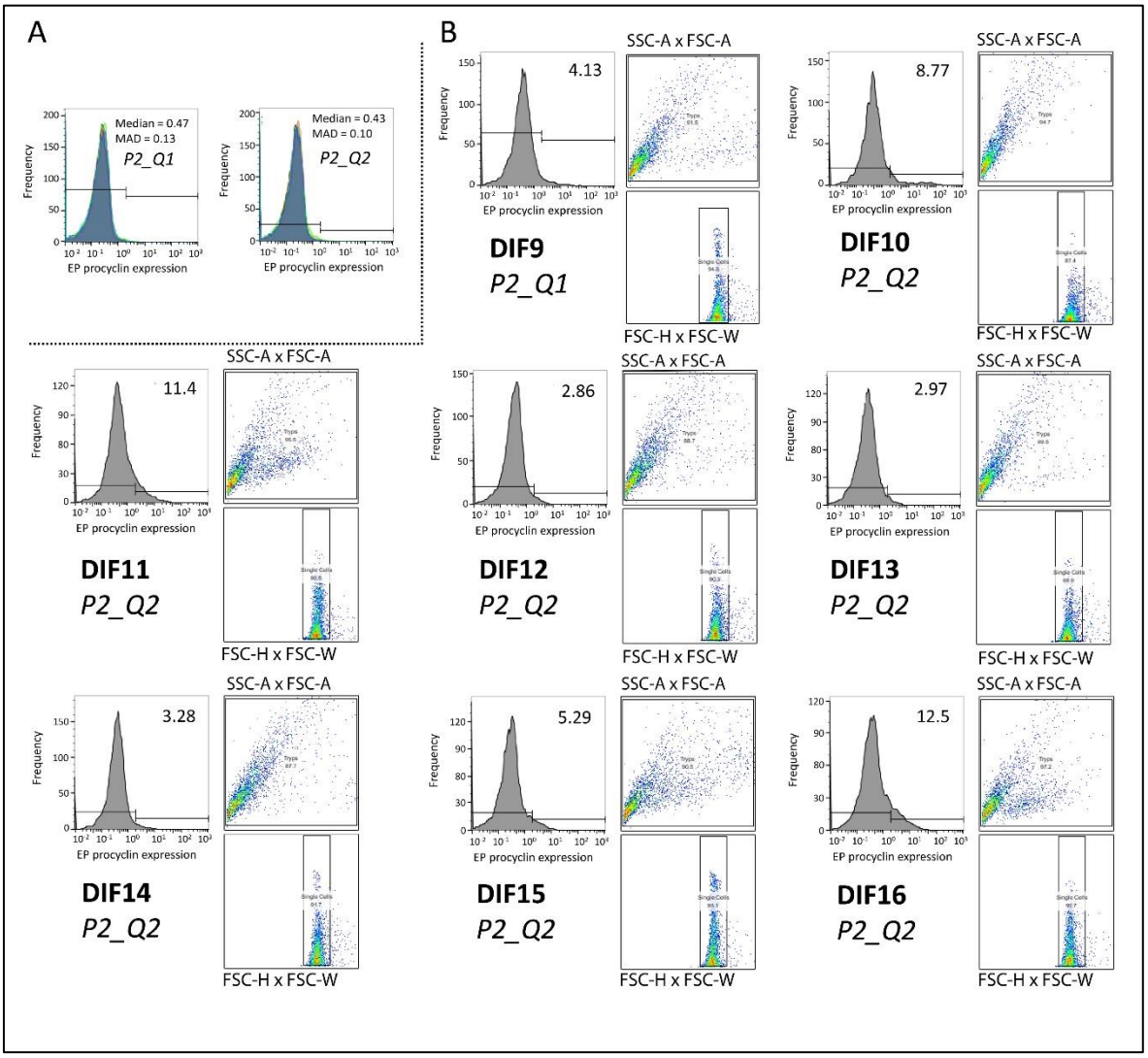


Figure 5-12. Detailed analysis of flow cytometry results for DIF9-16. **A.** Superimposed EP procyclin expression histograms for the negative controls in each plate with a compound depicted in this figure. Their median EP procyclin positive population, and their median absolute deviation (MAD) are inscribed in the plots. **B.** Histograms of EP procyclin expression for the DIF compounds. EP positive % of the population is inscribed in the upper-right corner of the plots. Next to it, gatting strategy leading to that population (same as Figure 5-6): trypanosomes (SSC-A x FSC-A), and single cells (FSC-H x FSC-W).

In flow cytometry, the forward scatter area signal is proportional to the cell size, while the side scatter area signal is proportional to the granularity (i.e. surface complexity). Analysing the forward versus side scatter dot plot, seven compounds (DIF1, 4, 5, 7, 11, 15 and 16) presented two differentiated clusters of cellular events, Figure 5-13. One tended to expand throughout the upper left quadrant, indicative of smaller cells but more complex in their surface granularity; and the other localized more towards the bottom right, indicative of larger but less complex cells. Independent analysis of these populations did not

show a common trend but always involved enrichment of one of the two sets in the EP positive population. Such enrichment was particularly remarkable in DIF1, DIF7 and DIF15 for the shorter cells with a more granular surface. The other subset displayed an almost null EP procyclin expression, Figure 5-13.

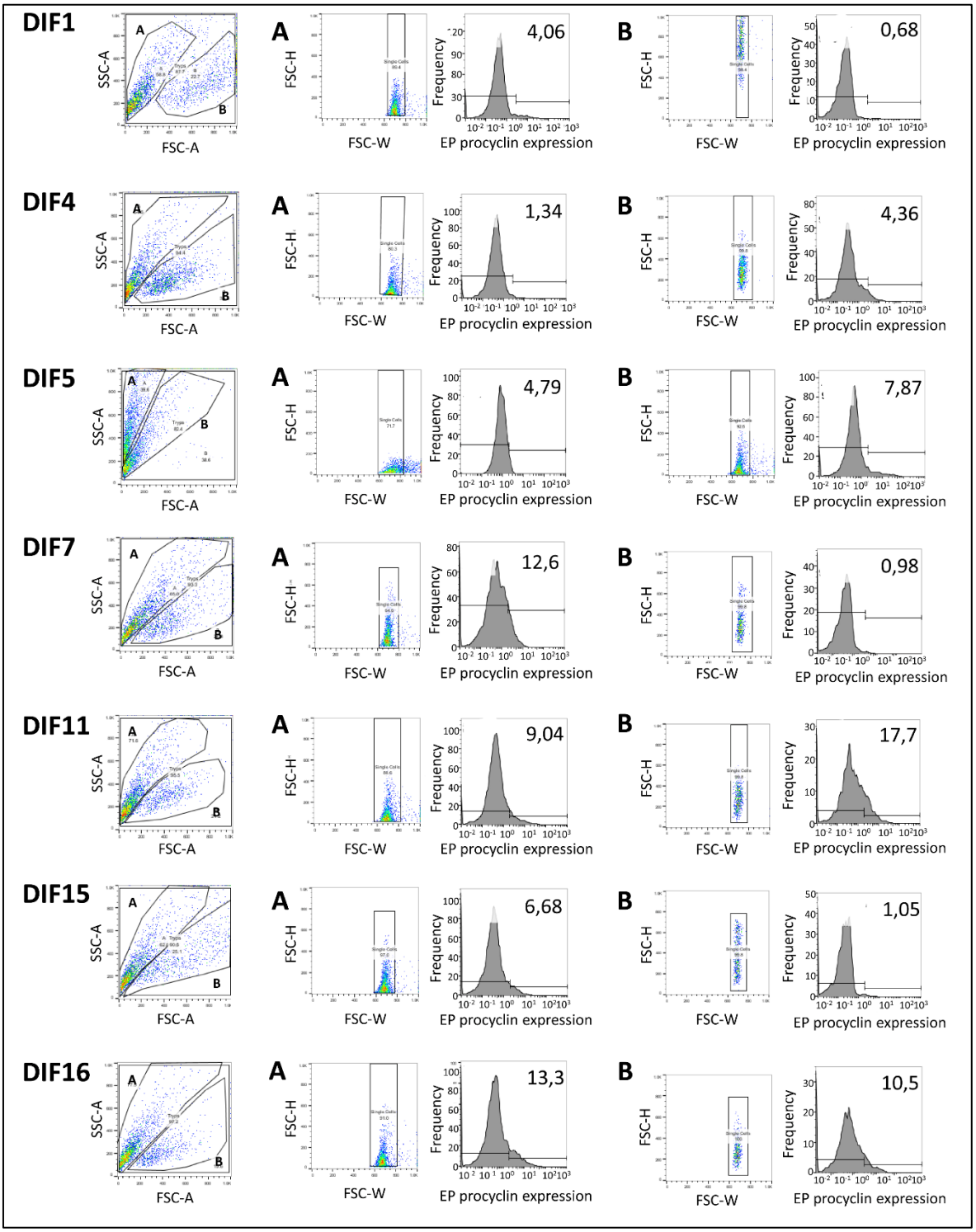


Figure 5-13. Seven DIF compounds that showed in the scatter plots two differentiated clusters of cells, in terms of size and granularity.

Discrete analysis of the % of EP procyclin positive cells within each of these populations for all of them. A = shorter and more granular events. B = larger and less granular events.

Dose-response validation curves

In order to confirm the primary screen results, the flow cytometry assay was repeated but applying nine steps of serial dilution (1:2) to the 16 selected compounds. It was carried out after two different incubation times (24h and 48h), doing everything in duplicate. Plates were set up with positive and negative controls arranged in the same layout used for the primary screen. Compound dilutions were carried out across every single row. Unlike in the primary screen, we also used the bottom line of the plate (H2-H11) for a compound test allowing 8 molecules to be screened per plate. Cleaning was performed every 16 wells in a separate tube by stopping the system and resuming after washing. For the 24 h incubation we proceeded as described for the primary screen. For the 48 h incubation, 2T1 cells were seeded at 5 x 10^4 parasites ml^{-1}, and positive controls (RDK2 RNAi line) at 10^5 parasites ml^{-1}, inducing with tetracycline immediately before plating.

Plates labelled as “A” contained DIF1-5, and DIF10-12. Plates named “B” contained DIF13-16, and DIF6-9. Results can be found in Tables S7 and S8.

For each time point, after antibody staining, one of the replicates was transferred to a black 384-well clear bottom plate, centrifuged to bring parasites to the bottom, and imaged in a high content microscopy system (GE IN Cell 2000®). Four images per well were captured, using three different channels: bright field (parasite shape), Hoechst 33342 (DNA staining), and APC (EP procyclin expression). After imaging, cells were transferred back to a 96-well V-bottom plate for cell cytometry.

Quality assessment was conducted as for the primary screen, Table 5-2. None of the plates had a normal distribution. All of them had acceptable Z´factors (> 0.6) except for the two replicates of “B_48h”, which had Z´ scores of 0.41 and 0.31. A comparison of the median for the whole set of tested compounds was not used as a quality control this time, because all compounds were expected to

produce a signal to some extent and neither the positive nor the negative reference were meant to behave the same. We also need to take into consideration that certain controls were removed from a few plates during the plate processing, due to problems with the liquid handler and the flow cytometer setup, see Tables S7 and S8. The detection window (S:B) was comparatively smaller than in the primary screen. The coefficient of variation remained acceptable for the positive controls, %CV<20 (Shun *et al.*, 2011), while too high for the negative controls.

Table 5-2. Quality control of the validation run. Summary of results.

Plate ID	Normal distribution	Z´factor	%CV + ctrl	%CV - crtl	S:B
A_24h_1	No	0.61	10.57	34.57	20.14
A_24h_2	No	0.65	11.09	3.08	24.27
A_48h_1	No	0.82	4.65	29.24	26.05
A_48h_2	No	0.77	6.68	26.63	34.84
B_24h_1	No	0.87	2.50	51.83	29.27
B_24h_2	No	0.84	4.56	41.59	57.92
B_48h_1	No	0.41	6.71	50.36	5.37
B_48h_2	No	0.31	19.01	26.09	12.23

%CV, coefficient of variation; S:B, Signal/Background ratio; ctrl, control. Formulas depicted in Figure 5-7 and Figure 5-10.

The flow cytometer was set to collect 100 μl of sample per well recording a fixed number of events within the gate described previously (Figure 5-6). As can be observed in the Tables S7 and S8, and Figure 5-14, the number of events/well programmed to be recorded in the individual plates was different (6000, 3500 or 2000). Due to problems during the antibody staining and the transfer between plates, the number of events recorded was also reduced for the positive and negative reference in some plates. This is not analytically relevant for the differentiation assay as EP positive values are normalized in percentages. However, many of the 16 DIF compounds caused cell death besides EP procyclin expression, implicating a reduction in the number of events that the system could possibly record within the 100 μl sample.

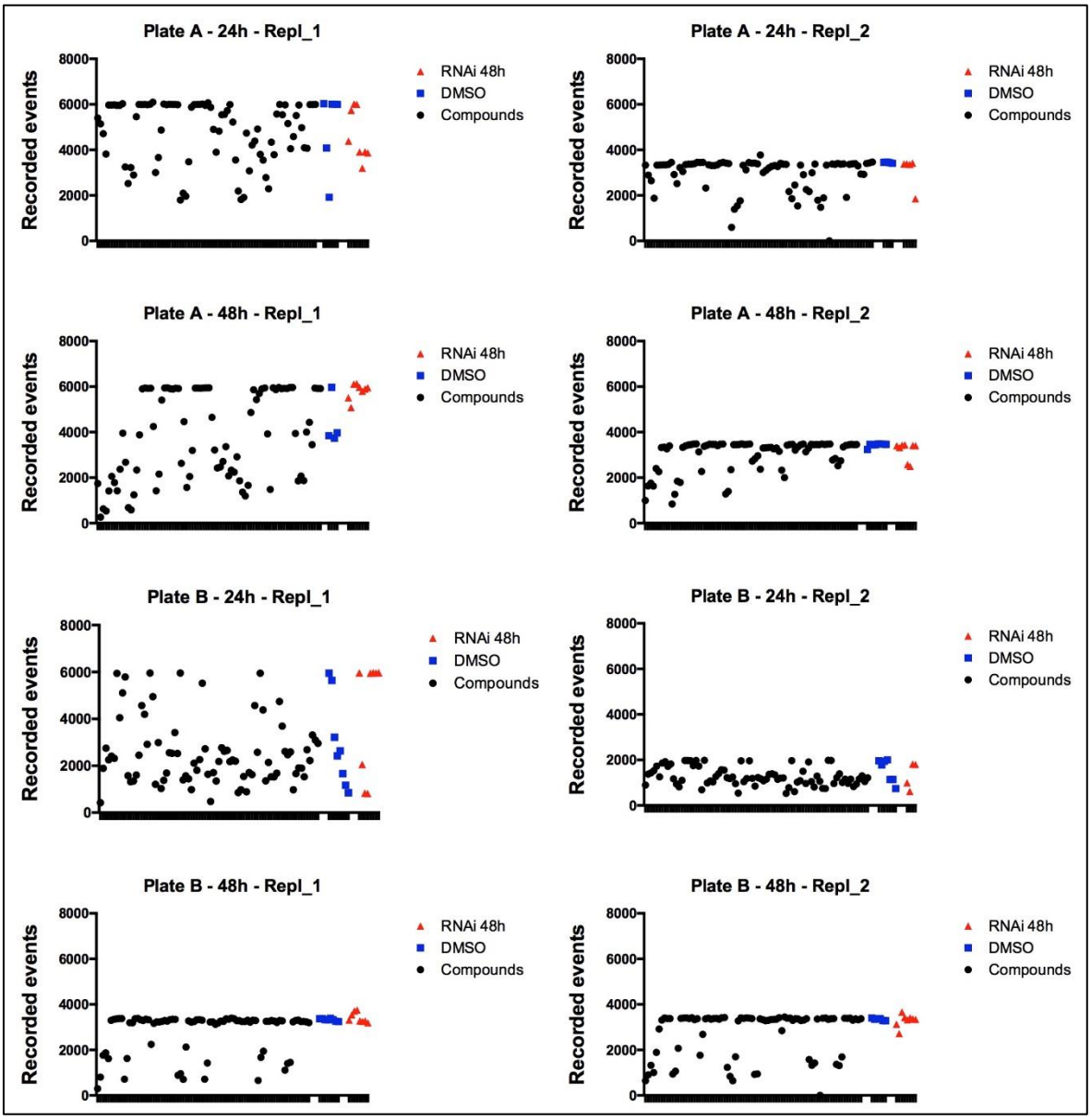


Figure 5-14. Recorded events distribution in the validation flow cytometry run.

In order to establish meaningful comparisons between replicates, the recorded events for each well were normalized against the number set to be counted per well in that particular run. Average of normalized counts was calculated between the 2 replicates analysed per condition and results plotted together with Z*-scores in Figure 5-15 to Figure 5-18.

Data analysis was carried out as for the primary screen, calculating Z*-scores to detect significant EP-linked fluorescence standing up over the background. When the threshold considered for significant results in the primary screen (Z*-score >

20) was applied to the validation test, only compounds in plates "A" reproduced an EP procyclin positive signal: DIF1-5 and DIF 10-12. However, poor quality indicated by Z´factors observed for the "B_48h" replicates did not allow reliability, so they were discarded. Although always under that threshold, for some compounds assayed in these plates %EP positive population increased with compound concentration, although it is not visible at the axis scale used in the figures. It was the case for DIF9 (after 24 h and 48 h), DIF13 and DIF14 (both after 48 h incubation), Table S8. The 24 h differentiation signal observed in the primary screen was not reproduced for DIF6, DIF7, DIF8, DIF15 and DIF16.

As expected, the trend was a reduction in recorded events coincident with the highest Z*-score at the highest compound concentrations Figure 5-15 to Figure 5-18. However, the peak of EP procyclin was observed at concentrations lower than 10 μM for DIF1, DIF9 and DIF10 (both after 24h and 48h incubation), and for DIF12 and DIF13 (after 48h incubation). This may be due to an overlap between cell death and differentiation. Supporting this hypothesis, curves of normalized recorded events and %EP positive population along the gradient of compound concentration mirrored each other in all cases except DIF5.

Compounds screened in plates "A", except DIF3, showed maximum Z*-scores between 2-4 times higher upon 24 h than after 48 h incubation. For compounds screened in "B" showing a proportional dose-response relation with the %EP positive population (although with Z*<20), 24 h scores are always 2 times higher than 48 h.

High content microscopy analysis for the parasites treated with the 8 compounds showing a Z* score>20, confirmed cell loss and death phenotype in most cases. Positive signals were only identified in dead parasites, and could be just antibody aggregates. Figure 5-19 shows an extract of these images at the different concentrations where they were first observed. However, weakness of the signal even in some of the positive controls makes these results inconclusive. IFA analysis on individual parasites might provide more robust data.

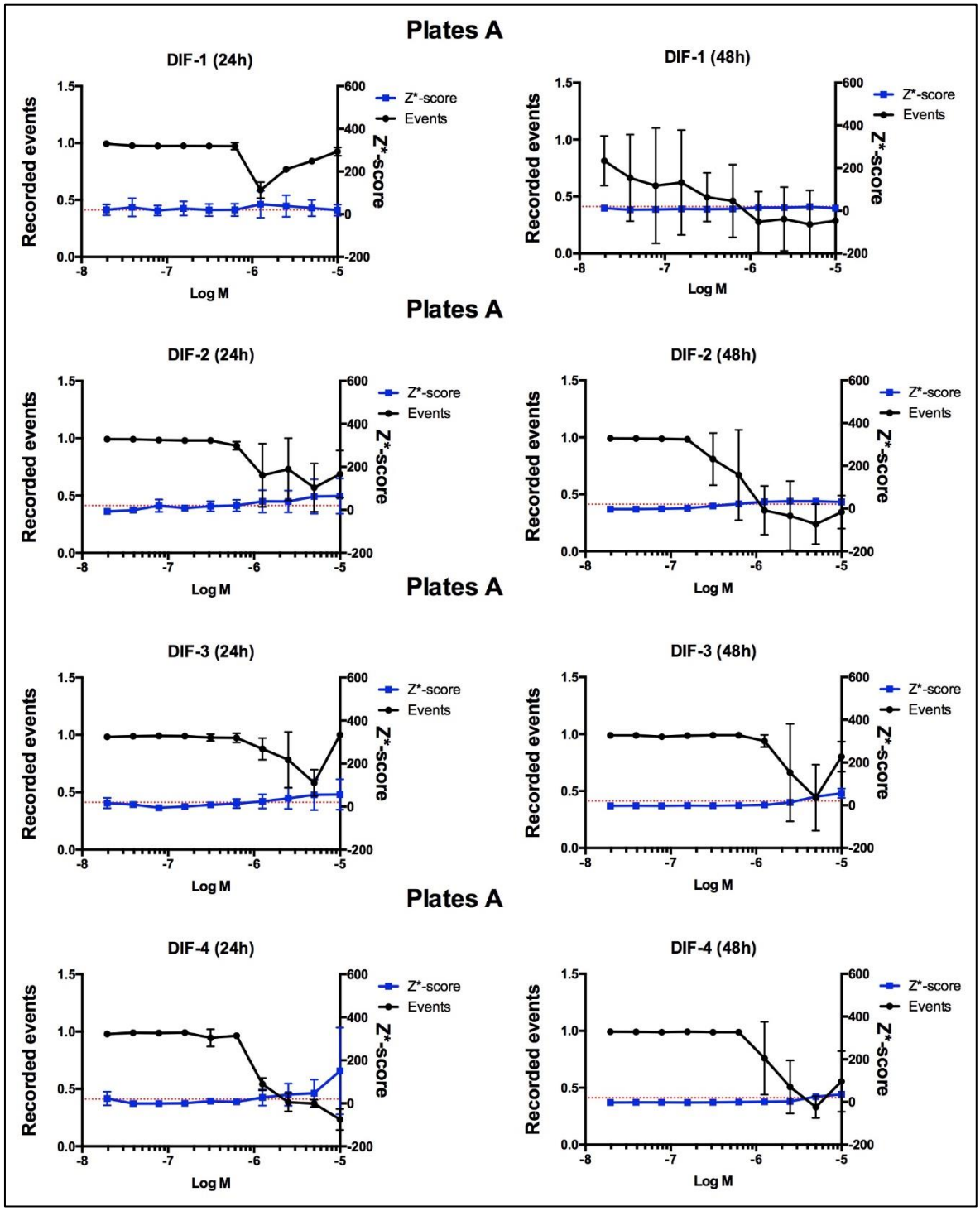


Figure 5-15. Dose-response curves depicting average (n=2) Z* scores and average normalized recorded events for the validation run of compounds DIF1-4.

Red dotted line representing the Z*score=20 threshold. Plate "A" or "B", refers to the validation run where the data can be found, see Tables S7 and S8.

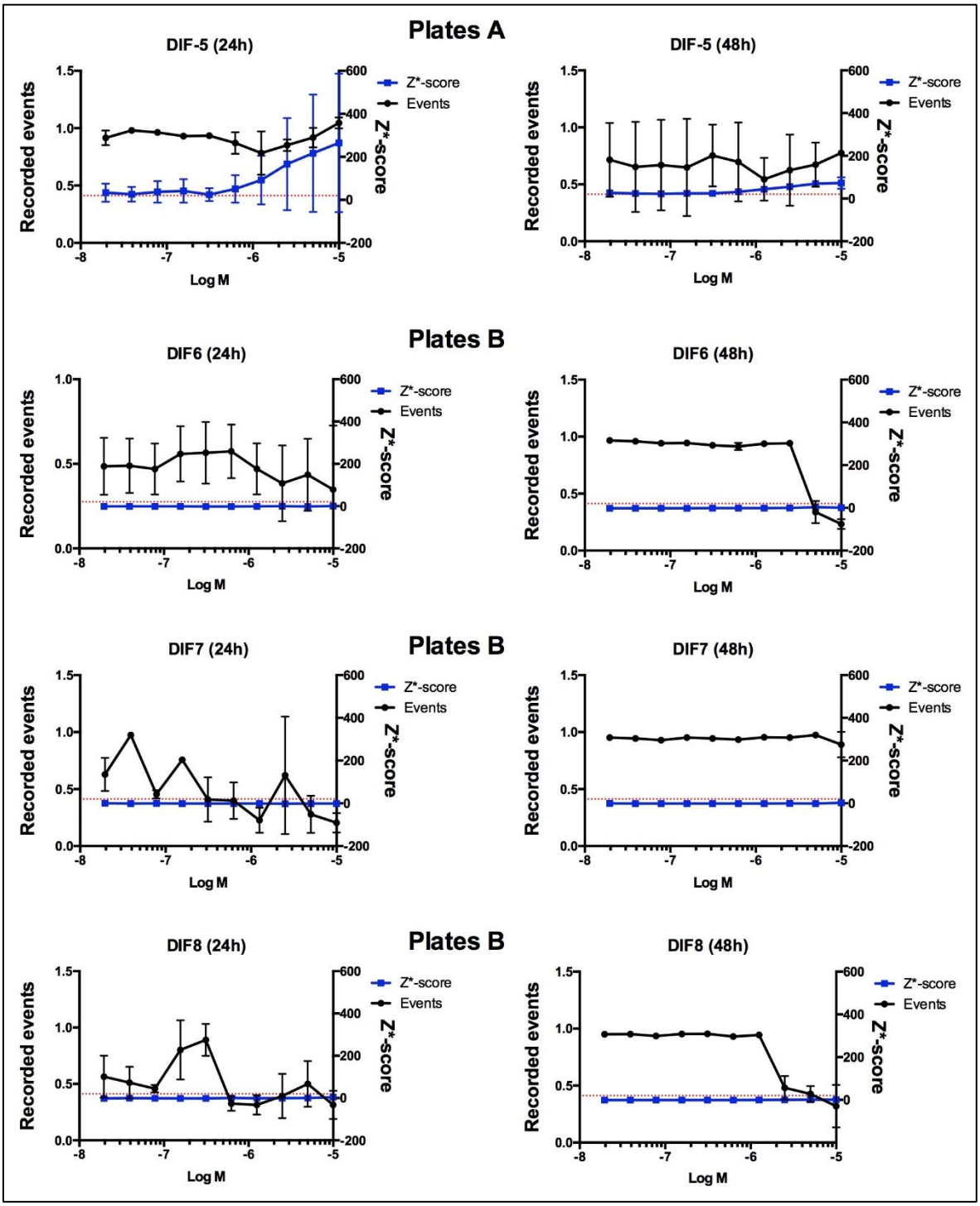


Figure 5-16. Dose-response curves depicting average (n=2) Z*scores and average normalized recorded events for the validation run of compounds DIF5-8.

Red dotted line representing the Z*score=20 threshold. Plate "A" or "B", refers to the validation run where the data can be found, see Tables S7 and S8.

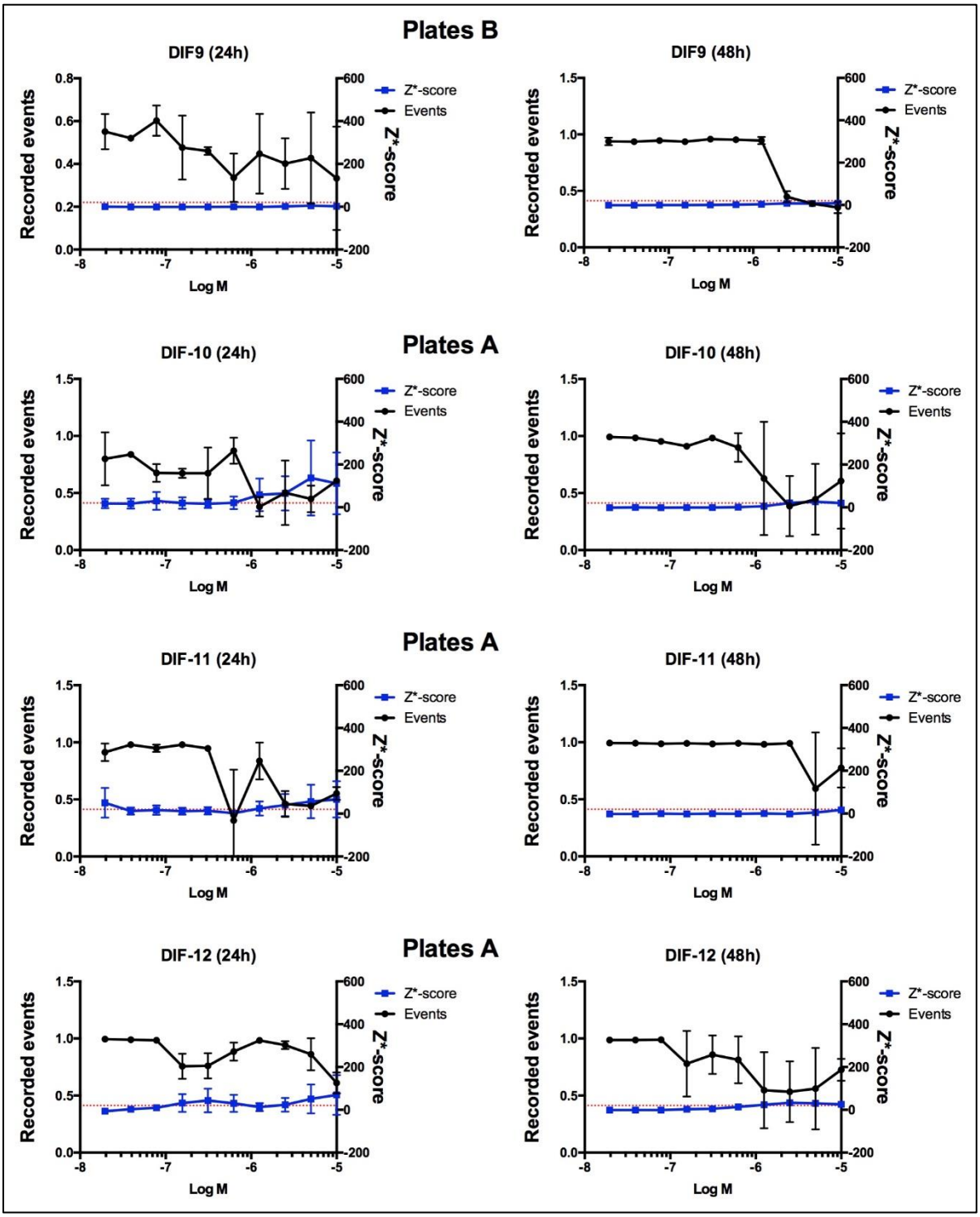


Figure 5-17. Dose-response curves depicting average (n=2) Z*scores and average normalized recorded events for the validation run of compounds DIF9-12.
Red dotted line representing the Z*score=20 threshold. Plate "A" or "B", refers to the validation run where the data can be found, see Tables S7 and S8.

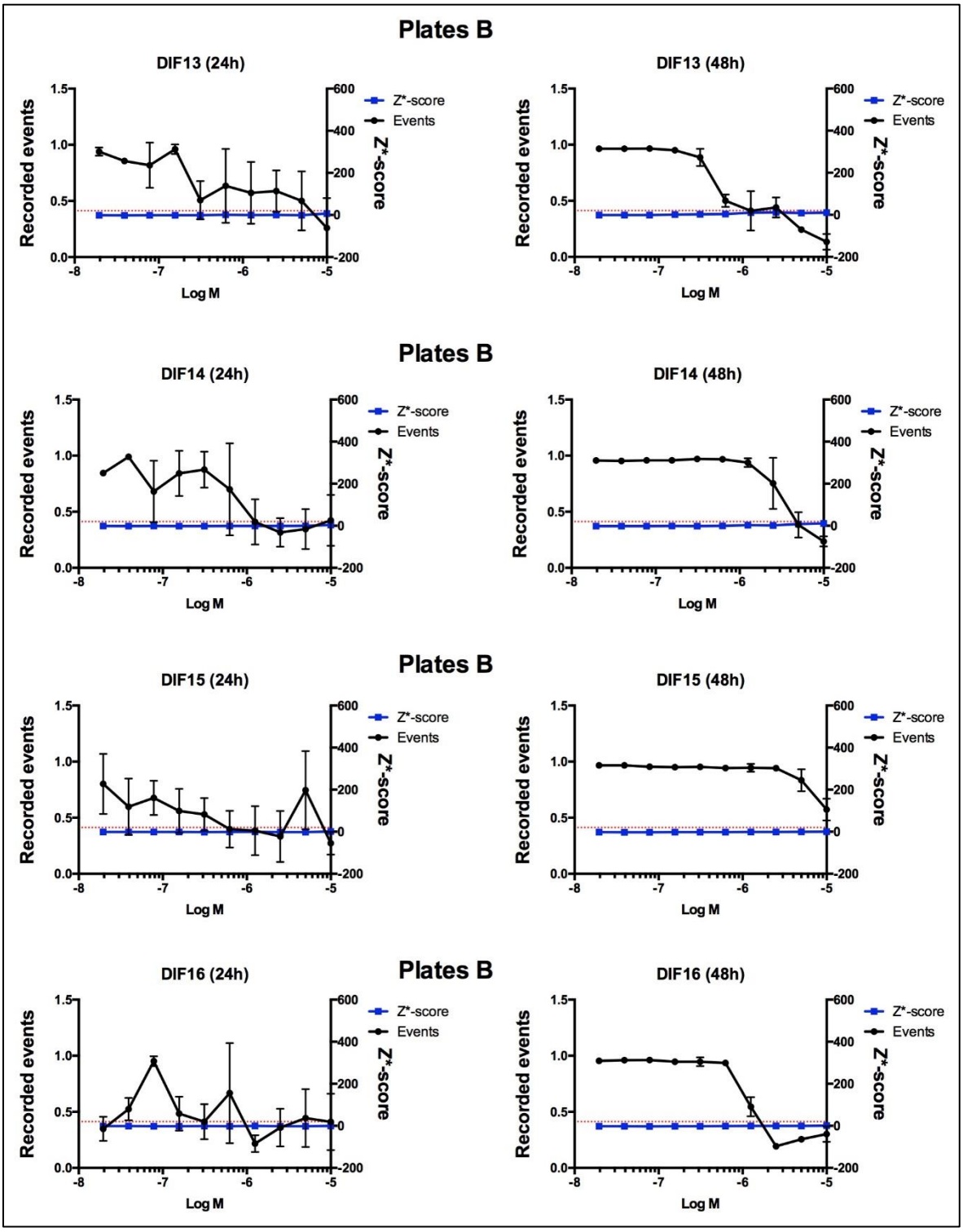


Figure 5-18. Dose-response curves depicting average (n=2) Z*scores and average normalized recorded events for the validation run of compounds DIF13-16.

Red dotted line representing the Z*score=20 threshold. Plate "A" or "B", refers to the validation run where the data can be found, see Tables S7 and S8.

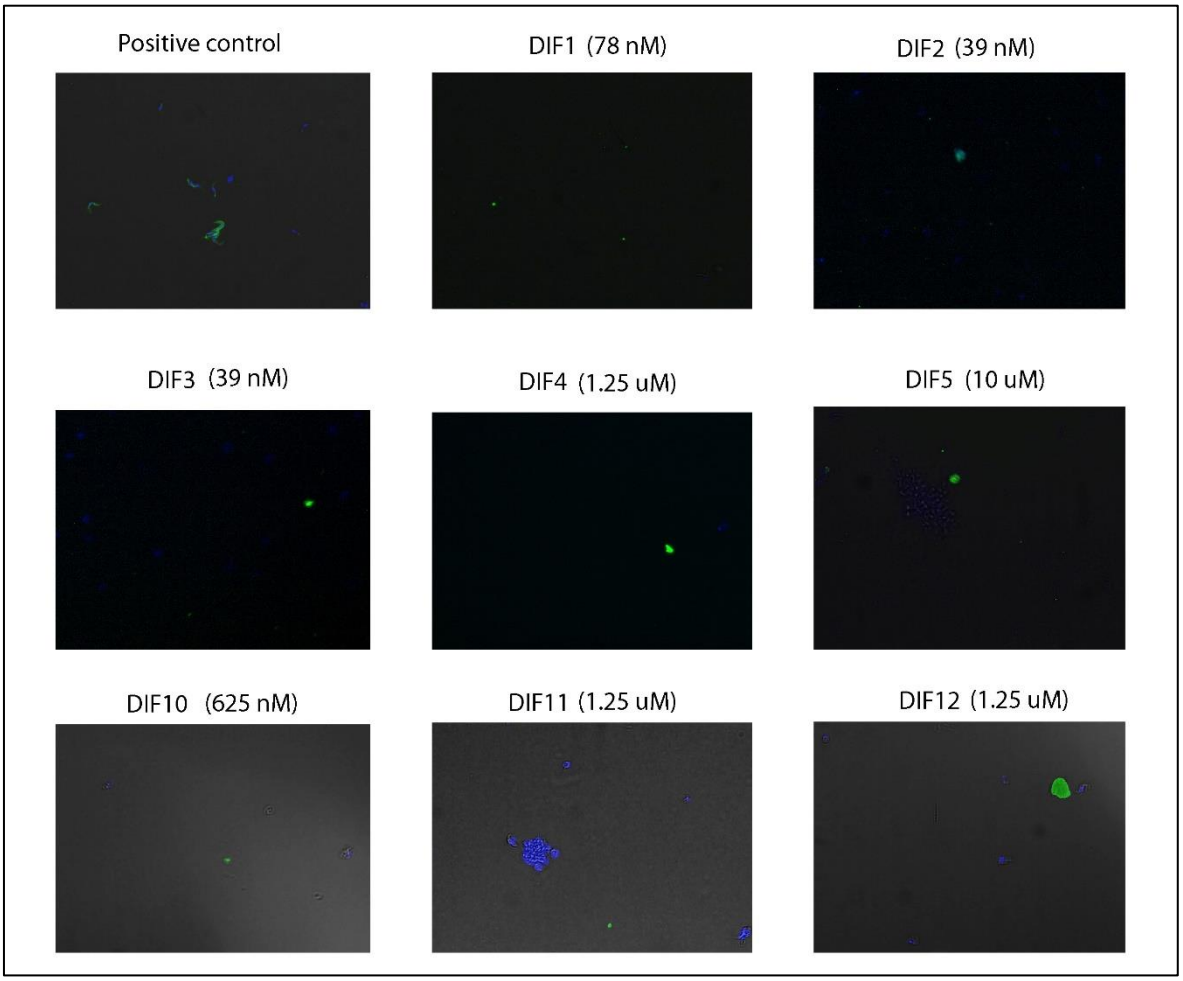


Figure 5-19. In cell fluorescence microscopy images showing positive APC signals for the 8 compounds with a validated signal in the dose-response curve.
Selection of the incubation dilutions showing positive signals after 24 h incubation. Green = EP procyclic-linked APC fluorescence; Blue = DNA-linked Hoechst 33342 fluorescence.

Alamar blue cell viability assays

To calculate EC50 values, Alamar blue cell viability assays were conducted after 72 h of compound incubation, adding the substrate during the last 24 h of growth (Räz *et al.*, 1997). Two technical replicates were carried out per compound with eleven steps of dilution starting at 5 μM. Fluorescence results were fitted to a logarithmic dose-response curve conducting non-linear regression in GraphPad Prism 5 and EC50 values were calculated when possible.

All compounds were trypanocydal except for DIF11 and DIF12, Figure 5-20 and Figure 5-21. Two other compounds did not arrest growth completely at the top 5

μM concentration: DIF7 and DIF14. When Alamar blue results were compared to cell death informed by reduction in events recorded by flow cytometry (Figure 5-15-Figure 5-18), excellent overlap was found with cell loss detected in the other 12 compounds after 48 h incubation. The only exceptions were the two found to kill everything after 72h, DIF1 and DIF5. The fixed number of events was almost fully countable in the 24h and 48h flow cytometry dose response-screen.

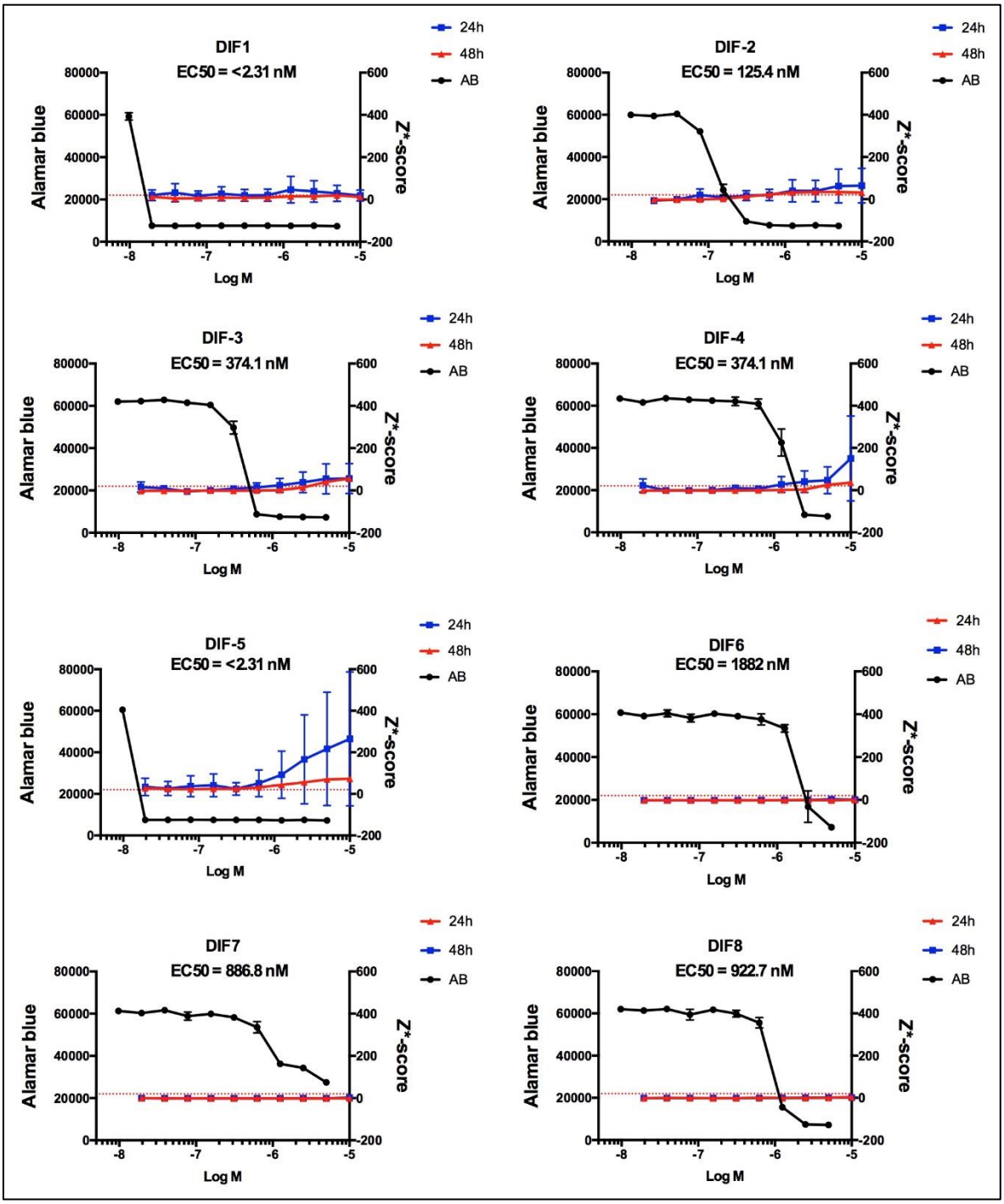


Figure 5-20. Alamar blue (AB) fluorescence intensity results overlaid to Z*-scores of %EP positive population after 24 h and 48 h incubation with compounds DIF1-8.

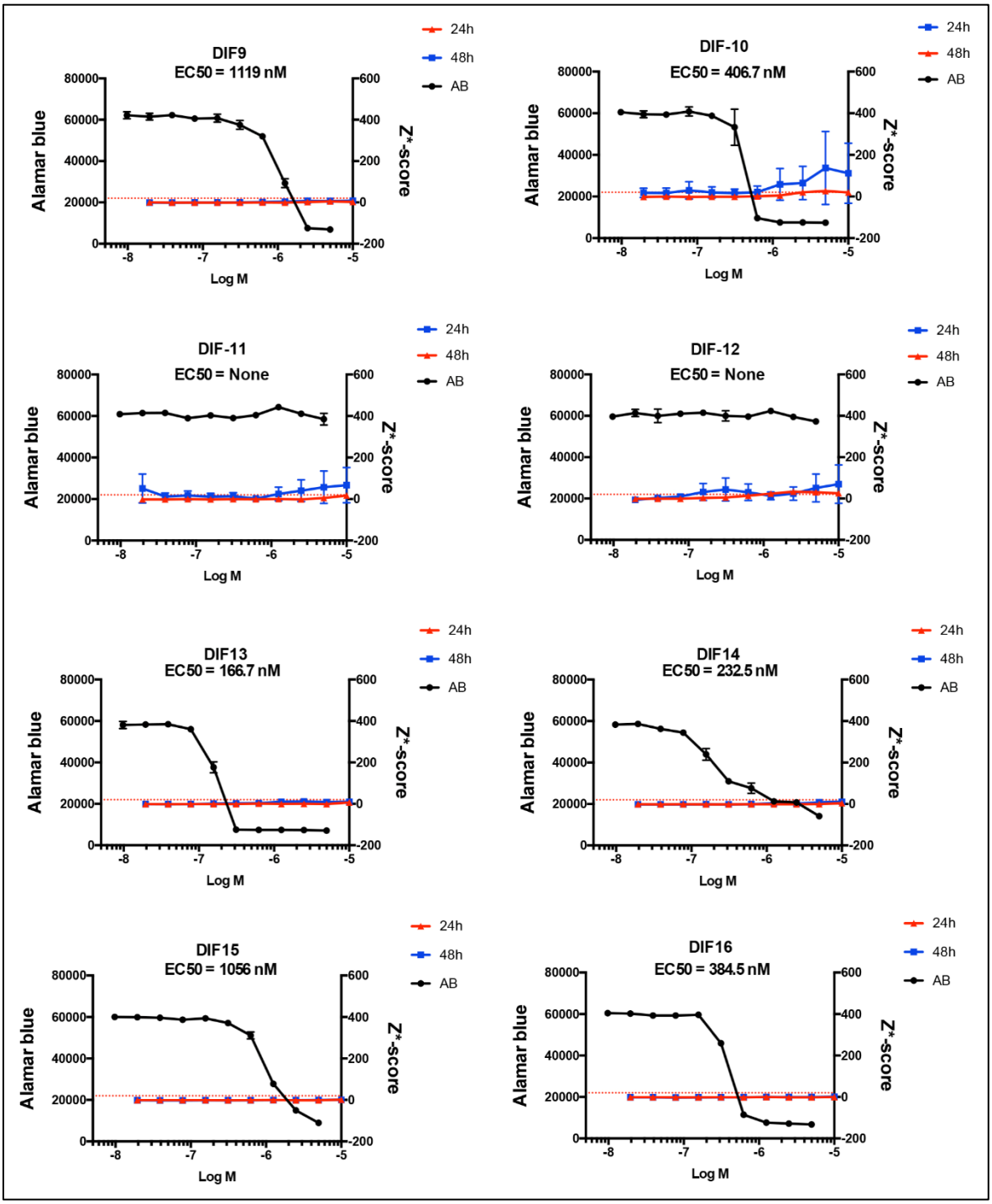


Figure 5-21. Alamar blue (AB) fluorescence intensity results overlaid to Z*-scores of %EP positive population after 24 h and 48 h incubation with compounds DIF9-16.

After all, 8 compounds were validated as expressing EP procyclin with Z*-scores>20, although individualized immunofluorescence microscopy confirmation would be recommended. Two can be discarded as RDK2 inhibitors as they do not

kill the parasite and two kill the parasite in the 72h range even at the lower concentration assayed here, Table 5-3.

Table 5-3. Summary of screening results.

Selected hits	Assay plate	Collection	1ry (10 μM)		Val. Av. - 24h		Val. Av. - 48h		EC50 (72h AB)
			%EP +	Z*score	%EP +	Z*score	%EP +	Z*score	
DIF-1	P1_Q3	HATbox	3,23	43,19	4,88	19,25	10,60	12,14	<2.31 nM
DIF-2	P1_Q3	HATbox	2,92	38,79	11,75	64,39	23,50	30,78	125.4 nM
DIF-3	P1_Q3	HATbox	19,80	278,23	10,80	56,02	40,05	56,36	374.1 nM
DIF-4	P1_Q3	HATbox	2,65	34,96	19,55	150,22	26,20	35,69	374.1 nM
DIF-5	P1_Q4	HATbox	6,98	38,03	45,45	265,01	51,55	73,32	<2.31 nM
DIF-6	P2_Q1	GUS	4,52	32,44	0,64	0,05	14,00	0,94	1882 nM
DIF-7	P2_Q1	NEU	8,89	67,40	0,45	-0,91	11,96	2,18	886.8 nM
DIF-8	P2_Q1	NEU	3,10	21,08	1,71	3,35	7,64	1,10	922.7 nM
DIF-9	P2_Q1	GUS	4,13	29,32	1,21	2,20	23,20	8,00	1119 nM
DIF-10	P2_Q2	GUS	8,77	87,79	17,50	111,75	15,30	19,55	406.7 nM
DIF-11	P2_Q2	NEU	11,40	115,47	12,45	67,30	13,65	16,88	None
DIF-12	P2_Q2	GUS	2,86	25,58	10,75	69,40	19,70	25,41	None
DIF-13	P2_Q2	NEU	2,97	26,74	2,66	6,79	33,95	10,01	166.7 nM
DIF-14	P2_Q2	NEU	3,28	30,00	1,88	3,37	36,85	10,74	232.5 nM
DIF-15	P2_Q2	NEU	5,29	51,16	1,00	1,29	10,08	0,56	1056 nM
DIF-16	P2_Q2	NEU	12,50	127,05	0,58	-0,16	11,80	0,55	384.4 nM

In bold, best fit candidates to serve as RDK2 inhibitors. In red, Z*scores>20. Val. Av.= Validation average at 10 μM. Compound collections: HATbox (Peña *et al.*, 2015); NEU (Diaz *et al.*, 2014); and GUS, same screening pipeline as (Wenzler *et al.*, 2016).

5.4 Discussion

Here we present the initial steps towards the first cell-based, mechanism-directed phenotypic screen for drug candidates performed with *Trypanosoma brucei* to our knowledge. The aim was looking for compounds causing cell death preceded by EP procyclin expression, i.e. the phenotypic properties of the parasite upon RDK2 RNAi. The work presented in the previous chapter proved that such phenotype might be due, at least partiality, to inhibition of its kinase activity. RDK2 is a divergent NEK kinase that is important for life-stage maintenance, thus is essential for the parasite both *in vivo* and *in vitro* (here, and Jones et al. 2014). In a second phase, hits rising from this screen need to be tested under RDK2 overexpression conditions, in order to see whether cells become more resistant to the compound; and in the recombinant protein, to see whether they are RDK2 inhibitors *ex-vivo*.

Phenotypic screens have been proven more useful than *ex-vivo* target-based screens in drug discovery (Swinney, 2013). Compounds are identified directly with the physicochemical properties required to act upon the living parasite (solving problems such as cell uptake, or cell efflux), and the desired cytocidal activity. Good drugs may kill the parasite throughout interaction with more than one biological target, and they can be overlooked in target-based screens (Gilbert, 2013). In the case of infectious diseases caused by protozoans, the key aspect is killing the parasite with selectivity over the host cells rather than inhibiting a very precise target (Merritt *et al.*, 2014). On the other hand, knowing the mode of action of a prospective drug helps refining its molecular properties in order to reduce toxicity towards the host, increasing potency against the infectious agent, and predicting possible resistance mechanisms. With this approach we attempted, at once, identification of molecules with a particular global effect *in vivo,* whose exact mechanism of action can be traced up to a particular target, using the resources generated in the previous chapter. We took advantage both of the detailed description we made of *T. brucei* RDK2 phenotype, and the knowledge accumulated about the drug-like properties of the selected collection of compounds.

After our 10 μM flow cytometry primary screen, we identified 16 compounds out of 518 (3.03% hit rate) that produced an EP procyclin positive signal above the threshold (Z*score >20). As we incubated the compounds with bloodstream trypanosomes, this may be indicative of differentiation to procyclic forms. Among these compounds, such phenotype was reproducible for 7 molecules (8 if we consider DIF1, that was just under the threshold, Z*=19.25). 6 of them also showed trypanocidal activity with EC50s in the submicromolar range, Table 5-3. EP procyclin expression and cell death combined, make these compounds candidate RDK2 inhibitors. For the 6 compounds, the two effects are dose-dependent after 24 and 48 h incubation, an indication of non-stochastic cause-effect relationship, Figure 5-15-Figure 5-18. Z*score after 48 h incubation is at least 1 fold smaller for all of them, maybe as a consequence of parasite death. The only exception is DIF3, whose Z* score remains constant in the two analysed time points.

These 518 compounds were selected from phenotypic screens where they had proven their capacity to kill *T. brucei*, and their drug-like physicochemical properties. They are mainly divided in four blocks: trypanocidal compounds without records of kinase inhibition, HATbox-A (Peña *et al.*, 2015); trypanocidal compounds with a history of human kinase inhibition, HATbox-B (Peña *et al.*, 2015); trypanocidal kinase inhibitor chemotypes, NEU (Diaz *et al.*, 2014); and compounds selected in a phenotypic screen for activators of the promoter of EP procyclin (GUS), not necessarily trypanocidal (Wenzler *et al.*, 2016).

5.4.1 Five trypanocidal compounds from the HATbox producing an EP procyclin-positive signal

DIF1-5 belongs to the HATbox, a selection of compounds able to kill *T. brucei* with drug-like physicochemical properties including central nervous system penetration (Peña *et al.*, 2015). They have been also assayed in human cell lines to prove selectivity against the parasite.

DIF1 (MW=327.384) and DIF5 (MW=344.327) were identified as TCMDC-143364 (or N3447-30-41), and TCMDC-143638 (or U20127/42/3) respectively at the HATbox screen (Peña *et al.*, 2015). They were the 2 most potent compounds of our selected hits as they killed everything after 72 h incubation. The lowest concentration tested in our Alamar blue assay, 2.31 nM, was enough to kill them. These numbers are close to the EC50s calculated in the HATbox report after 48h incubation: 12-60 nM, for DIF1; and 4-20 nM for DIF5. They showed a differentiation signal, particularly strong for DIF5 (Z*=265 after 24 h, and Z*=73.32 after 48 h), but it may not come from viable parasites according to the high content imaging data. In the primary flow cytometry screen, at 10 μM, both showed a big cluster of events in the upper-left corner of the FSC x SSC dot plot, which suggests small-sized cells with high granularity, maybe representing dying parasites.

DIF2 (MW=369.464) was identified as TCMDC-143357 (or N3426-35-101) in the HATbox screen, where it had an EC50 (after 48 h) of 125-158 nM, matching the 125.4 nM EC50 at 72 h found in our screen. Z*scores for differentiation were of 64.39 at 24 h, and 30.78 at 48 h.

DIF3 (MW=281.215), also known as TCMDC-143457 (or ST/1949505) in the HATbox, is the only showing a stable EP procyclin signal (Z*=56.02-56.36) not decreasing from 24 to 48 h incubation. The EC50 we report after 72 h incubation is 374.1 nM, about half the 630 nM reported at 48 h by Peña et al. (2015). In the primary flow cytometry screen, the peak of DIF3 events in the EP procyclin fluorescence intensity histogram was uniformly shifted toward the positive gate, being the most active hit of the screen with 20% of positive cells.

The last, DIF4 (MW=425.322), is TCMDC-143642 (or JE98054-138P2) in the HATbox and shows more than one third decrease in EP procyclin positive signal from 24 to 48 h, and our detected EC50 after 72 h incubation (374.1 nM) is about half what Peña et al. (2015) reported at 48 h (794 nM-1 μM). DIF4 was also divided in two clear clusters of different size and granularity in the FSC x SSC analysis at the primary screen, suggesting abundance of aberrant cells.

Unfortunately, we have no data about which of the HATbox compounds showed kinase inhibition but we know that almost half of them did. They reported more than 90 PK inhibitors out of 188 compounds in the set. They were clustered in >29 chemotypes, including common hinge binding motifs such as pyrimidines, pyrrolopyrimidines, azaindoles, and benzimidazoles. Peña et al. (2015) did a phylogenetic prediction of possible *T. brucei* targets for the HATbox based on their orthology with human protein kinases against which at least one compound had proven activity. RDK2 was not predicted maybe because it is NEK12.2 (Jones et al. 2014), a quite divergent PK within a relatively expanded family in *T. brucei* respect to other eukaryotes. However, NEK2 was among the human kinases targeted by at least one HATbox compound. NEK2 is a mitotic regulator that has been explored as a drug target in anticancer therapy (Frett *et al.*, 2014). Several compounds have been found with potency and selectivity against this human PK (Whelligan *et al.*, 2010; Henise and Taunton, 2011; Solanki *et al.*, 2011; Innocenti *et al.*, 2012). If any of the 5 compounds we identified in our screen is a human NEK2-targeting inhibitor, we could analyse their structure-activity relationship to optimize potency and specificity. A crystal structure of RDK2 (still not available) would be useful at this respect. However, other compounds in HAT box-A, not having typical type I PK inhibitor profiles (ATP-competitors) may still work as allosteric inhibitors. We have already mentioned

in this thesis that these compounds tend to have more potency as they have no need to compete with physiological ATP.

5.4.2 One GUS compound producing an EP procyclin-positive signal with trypanocidal activity

DIF10 (MW= 482,949) was the only hit of our screen coming from the 283 "GUS" compounds, selected for their capacity to activate the EP procyclin promoter in a screen (Wenzler et al. in preparation) following the same procedures reported by Wenzler et al. (2015), but with a larger collection of compounds. In our screen, DIF10 had an EP procyclin signal of Z*=111.75 after 24 h incubation, that was reduced to Z*=19.55 after 48 h. It had an EC50 of 406.7 nM after 72h. Unfortunately we do not have structural or functional information about GUS compounds enabling useful comparisons with other hits.

Low hit-rate for this set (1 out of 283 molecules) was unexpected, as the selection criteria that generated this library followed a similar rationale to that of our anti-RDK2 screen. We detected EP procyclin expression directly with a specific antibody while they did it indirectly, by introduction of β-glucuronidase (*GUS*) as a reporter in the EP procyclin locus. When expressed, GUS activity over a substrate (4-methylumbelliferyl-β-D-glucuronide), could be measured by changes in absorbance at 405 nm that were proportional to EP procyclin promoter activation (Sbicego *et al.*, 1999).

5.4.3 Cell death limits our detection capacity

EC50s of compounds identified in the GUS screen were not disclosed to us but we know that 61 compounds had their peak of activity at concentrations under 1.25 μM; while 236 had it at concentrations over 1.25 μM, all measured after 43 h incubation (Tanja Wenzler, personal communication). DIF10 belonged to the last group but we cannot know how many we overlooked due to an excessive concentration of the compound that killed the parasites quenching the EP procyclin-positive signal. Wenzler et al. (2015) published a smaller library of compounds using the same GUS approach, and they detected a majority of cases where overlapping of cell death with EP locus expression activation impeded

detecting the differentiation-associated signal. This may be also the case at the 10 μM concentration assayed here.

We showed in a plate reader how the increasing background noise associated with normal growth in the negative control population quenched the signal emitted by the positive population, whose growth was arrested or which decreased in numbers, Figure 5-1. Although single-cell analysis carried out by flow cytometry solved this numeric problem by estimating the proportion of the total parasites that produced a positive signal per well, active compounds can still be overlooked if the signal disappear with the dead parasites.

One possibility is that at the 24 and 48 h time points used in the experiment, differentiated parasites were already dead, having lost their shape and having started to disintegrate. Sub-lethal concentrations of compound may not have induced differentiation, since the inhibitor could be less abundant than the natural substrate (potentially ATP). Under this hypothesis, parasite death would be the last step of the differentiation mechanism. This is supported by the fact that being dose-responsive, the EP-linked signal is always 2-4 times higher after 24 than after 48 h incubation. At the same time point, the more compound used the higher the signal; but, for the same concentration, the signal fades as time goes by, Figure 5-20 and Figure 5-21.

High content immunofluorescence microscopy on cells from the primary screen did not suggest that the signals detected in the screen come from viable parasites, Figure 5-19. They could be either dead parasites or debris. Slides for immunofluorescence microscopy need to be done after compound treatment for validation.

As we said for DIF1, DIF4 and DIF5 when the flow cytometry scatter plots (SSC-A x FSC-A) were analysed, trypanosome events spread from the bottom left quadrant drawing a “comet-like” profile (Figure 5-11 and Figure 5-12). This is due to the elongated shape of the trypanosome cell body and the different possible orientations of the parasites when crossing the detector. Interestingly, treatment with 7 of the 16 hit compounds separated the parasites into two different clusters. The EP positive fraction in each of these subgroups was not too different for cells treated with 4 of these molecules (DIF4, DIF5, DIF11 and

DIF16). However for the other 3, the EP positive population was enriched in the subset containing the smaller sized and granular events (DIF1, DIF7, and DIF15), see Figure 5-13. As the high content imaging suggested, this may also support the idea of EP positive population being small dying parasites. As we suggested before, large scale IFA would provide appropriate validation for this observation.

As discussed in chapter 4, an antibody raised against RDK2 or endogenous labelling of the RDK2 gene is needed to understand the protein behaviour during RNAi-mediated depletion. The current lack of detailed information about the time when RDK2 RNAi phenotype occurred in relation to the moment when RDK2 disappeared from the parasite, made it difficult to estimate the best time frame to carry out the experiment. We aimed for the 24 h incubation since RDK2 RNAi phenotype became evident from around 36 h. The 48 h time point used for validation was supported by assays with GUS activity performed after 40 h incubation (Wenzler *et al.*, 2016). But upon CCA treatment, for instance, cells express EP procyclin since 6 h post induction (Rico *et al.*, 2013).

An active compound should be expected to have an equal effect over all parasites in the well unless for some reason the compound acted only in certain subsets of the population. An example would be if the molecule targets proteins that function only in certain stages of the cell cycle or the life cycle. We theorized in the previous chapter that RDK2 might have different roles in different developmental stages, so this would most likely happen with prospective RDK2 inhibitors.

Another possibility, since the flow cytometry screening was performed in live cells, is that parasites internalized and degraded antibodies by means of their flagellar pocket through hydrodynamic flow contributing to signal fading (Engstler *et al.*, 2007). Interestingly, procyclic forms have 10-fold lower endocytosis that bloodstream forms, therefore it may be lower in differentiating cells (Field and Carrington, 2004).

5.4.4 Two non trypanocidal compounds produced an EP procyclin-positive signal

A second GUS compound, DIF12 or HL209352-009A1 (MW=299.338), was found to produce an EP procyclin-positive signal, however it did not kill cells after 72 h at 10 μM. A Z*score=69.4 was detected after 24 h, that reduced to Z*=25.41 after 48 h incubation. The pictures obtained in high content imaging, did not seem to come from viable parasites either.

The last compound identified with a reproducible EP procyclin-positive signal, although unable to kill the parasite after 72h was DIF11 (MW= 293,322), the only representative of the NEU set (NEU-0001221 or CA200460-075A9). This collection is the product of the largest kinase-targeted screen ever directed against *T. brucei* (Diaz *et al.*, 2014).

Data mining and statistical analysis on the "NEU" results showed that compounds inhibiting a particular set of human kinases are more likely going to have inhibitory effects over *T. brucei* (Amata *et al.*, 2016). This has been called "preferred lead repurposing" and selecting compounds with that criteria showed an enhanced hit rate when compared with screening of non-filtered PK inhibitor chemotypes (Amata *et al.*, 2016). Although we do not know how many of the compounds tested in our screen were potential NEK inhibitors, among the "preferred" leads there were human NEK8 inhibitors that may have activity also against RDK2. At least the PKIS collection contained several of these.

Seven other "NEU" compounds were identified among the top 16 hits of our screen (Table 5-3), however their positive signal could not be reproduced in the validation step. Some showed a dose-response signal under the threshold though, indicative of a cause-effect relationship with the compound that upon concentration and time-exposure adjustment may lead to positive results.

5.4.5 Differentiation is not a stress response against a trypanolytic agent.

Differentiation is a natural process in *T. brucei* bloodstream forms when it comes to escape environmental stress encountered in the tsetse midgut (Rico *et*

al., 2013). But EP positive signals observed in the screen cannot be always a response to stress caused by exposure to a trypanocydal drug.

5 out of the 6 ultimate hits belong to the HATbox (Peña *et al.*, 2015) and they are well profiled trypanocidals, however the presence of at least another 184 compounds from the same set in our screen, discount parasite death as a cause for EP procyclin expression. Most are negative in our screen while reportedly killing the parasite in the nanomolar range (Diaz *et al.*, 2014; Peña *et al.*, 2015). Such killing capacity is evidenced here by the drop in numbers of the recorded events observed in different plates across the primary screen, Figure 5-8, and DIC images of high content microscopy.

Our collaborators at the Swiss TPH have reported something similar recently. In the screen they published with the GUS assay (Wenzler *et al.*, 2016), all molecules producing GUS activity were trypanocidal while just 9% of trypanocidal compounds caused GUS activation.

5.4.6 Follow up

The next step would be independent validation in IFA and continuing our screening plan (Figure 5-4), comparing EC50s of the parental cell line with the RDK2 overexpressor that would tell us if we are possibly on target. If EC50s were higher for the overexpressor, we would look for confirmation in the kinase assay with RDK2 recombinant protein using α-casein as a substrate (see previous chapter). The second round of primary screen at 1 μM, could show more hits.

Other libraries that would be worth analysing in this pipeline are the Dundee protein kinase inhibitor collection (Woodland *et al.*, 2015) and the complete GSK PKIS (Drewry, Willson and Zuercher, 2014). We had a material transfer agreement for the last one but time pressure did not allow us to access them. In addition, it may be good to test some NEK inhibitor libraries developed for cancer therapy (Frett *et al.*, 2014) and try to optimize hits if successful based on SAR analysis.

This preliminary screen has been useful as a starting point for developing a more robust pipeline. If optimized it could still produce more reliable hits. It would at

least validate at a protein expression level hits from the Swiss TPH screen (Wenzler *et al.*, 2016). If compounds showing both phenotypes (differentiation and death) are confirmed in validation steps with both the RDK2 overexpressor and protein kinase assay with the RDK2 recombinant protein, we would have identified a drug target for a compound in the screened sets, which would have excellent prospects in terms of drug discovery.

Such compound would also become an excellent research tool able to provide chemical validation of the RDK2 RNAi phenotype and help out dissecting the intricate signalling pathways that rule the balance between life-stage maintenance and cell cycle progression in *Trypanosoma brucei*.

Chapter 6 General discussion

In this thesis I applied RITseq to identify protein kinases that are essential for infection and survival of *Trypanosoma brucei* in the mammalian bloodstream. The value of this information is significant as protein kinases are known to be susceptible to small molecule inhibition and are therefore exploitable as drug targets. Within the 49 protein kinases whose depletion caused a loss of fitness *in vivo*, we uncovered a number involved in withstanding stress conditions related with adaptation to that particular environment. I validated some involved in osmotic shock and serum exposure resistance. Among them, putative members of PI3K/TOR and MAPK pathways have been highlighted, however the way they interact to each other and with other factors that may be also essential for *in vitro* survival needs to be investigated. Exploring further these non-canonical virulence factors and their contribution towards withstanding the initial environmental shock that involves the transference from the tsetse's salivary glands to the mammalian bloodstream is an exciting perspective.

Many of these 49 genes, including the most divergent when compared to other eukaryotes, have orthologues in other kinetoplastid parasites, such as *T. cruzi* and *Leishmania sp*. Since RNAi technology is not available for these parasites (Lye *et al.*, 2010) and molecular genetic tools permitting gene deletion are so far more laborious and not high throughput friendly (Merritt *et al.*, 2014), exploring if proteins with drug target value identified in high throughput screens of *T. brucei* have orthologues with equivalent crucial functions in these organisms may be a promising strategy.

For further analysis, I focused in RDK2, a promising drug target that was essential both *in vivo* and *in vitro* and had very interesting biological implications. Inducible downregulation of RDK2 in the bloodstream form by RNAi caused differentiation to an aberrant procyclic form, which was unable to survive in culture unless placed in conditioned procyclic form media. In that circumstance, they survived but were unable to proliferate while acquiring a plastic adhesion phenotype that recalls proventricular forms within the tsetse fly.

A recently published proteomics dataset profiling CCA-mediated differentiation (Dejung *et al.*, 2016) showed that RDK2 levels increase in the stumpy form, but are reduced dramatically upon development to the procyclic form, being absent in established procyclic form cultures. Monitoring EP procyclin expression as a marker of bloodstream-to-procyclic differentiation, I show that the RNAi effect on differentiation was reduced in all conditions that favour stumpy* formation (Rico *et al.*, 2013), such as 8-pCPT-cAMP treatment and cold shock (may be due to transcriptional silencing), while these conditions ease differentiation mediated by CCA in uninduced cells. CCA treatment, -with limitations in the absence of PAD transporters (Dean *et al.*, 2009), or a more efficient simulation of its downstream effects, -PTP1 inhibition with cell-permeable BZ3 (Szöor *et al.*, 2010), entered in conflict with differentiation by RNAi, reducing the size of the EP procyclin positive population. When CCA and RNAi were applied together in a population pre-treated with pCPT-cAMP to induce stumpy* formation there was a synergistic effect though, suggesting that cells where RNAi had no effect were induced to differentiate by CCA, and cells where CCA had no effect were induced to differentiate by RNAi of RDK2. I propose that RDK2 depletion is a downstream effect of the CCA/PTP1 pathway, and that its premature depletion causes changes in the interactome that reflect interference with the CCA/PTP1 signalling (Szöor *et al.*, 2010), and that render imperfect procyclic forms that cannot survive.

Overexpression of RDK2 blocked differentiation triggered in monomorphic trypomastigotes by preincubation with 8-pCPT-cAMP (inducing stumpy* forms) followed by CCA treatment. However, data may suggest that RDK2 overexpression blocks 8-pCPT-cAMP-mediated stumpy differentiation mainly. Unfortunately, although initial findings suggested it was possible, complementation experiments did not permit reproducibly assigning RNAi phenotype to RDK2 depletion. More clear-cut gene deletion or specific chemical inhibition may be attempted to clarify this point. The originally successful complementation with a recoded active mutant (*rdk2*^{REC}) compared with unsuccessful attempts carried out with the inactive version (*rdk2*$^{REC/K70M}$), suggested that the RNAi effect may be caused by the lack of kinase activity and not by defect of other putative scaffolding functions which disappear when the protein is missing.

The detailed analysis of RDK2 was exploited to design a mechanism-directed phenotypic screen. In contrast to drug repurposing strategies, on-target confirmation should be easier if the predicted target is known. In this type of screen, we intend to attack a particular pathway or target class essential for the parasite that, upon mechanistic disruption, shows a characteristic-traceable phenotype. When confronted with target-based *ex-vivo* assays, phenotypic screens provide hits more readily. Selected molecules are active against the whole cells and detected based on specificity for the parasite compared to host cells, regardless the actual drug target in the objective organism. Effective drugs may have impact over more than one target at once.

In the mechanism-directed phenotypic screen I presented in this thesis, compound libraries with a history of trypanocidal activity (Peña *et al.*, 2015), kinase inhibition (Diaz *et al.*, 2014; Peña *et al.*, 2015), or activation of the EP procyclin promoter (Wenzler *et al.*, 2016), were screened for chemical entities phenocopying the RNAi effect assigned to RDK2 inactivation: triggering EP procyclin expression before (or while) killing the parasite. 6 compounds were identified in a primary screen conducted at 10 μM, a concentration intended to compete with cellular levels of ATP. Application of research tools developed for RDK2 analysis, such as the recombinant protein or the overexpression line, will allow an analysis of whether the 6 compounds were targeting the protein kinase.

There were many aspects of this research that encouraged targeting RDK2 for drug discovery besides its depletion killing *T. brucei* both *in vitro* and *in vivo*. RDK2 is a protein kinase and therefore is potentially druggable (Urbaniak *et al.*, 2012). It is rather divergent and belongs to the NEK family, which is expanded in the parasite compared to the mammalian host. It is involved in a parasite-specific function (life-stage maintenance), what highlights the chances of finding a potential drug capable of lethal parasite-specific inhibition. We proved RDK2 to be an active protein kinase that could use α-casein as a substrate, so a miniaturised assay could be developed to carry out a high throughput screen and initiate a structure-based drug discovery project.

There is an urgent need for new drugs if we want to achieve the WHO 2020 goal of HAT eradication. Since the London Declaration on Neglected Tropical Diseases

(2012) the research community, in an unprecedented display of collaborative and internationalist efforts, has made good progress towards the objective. However, we need to be cautious. Whilst the disease was almost eliminated as a public health issue in the 1970s, the lack of sustained efforts and investment of resources ever since, relaxed surveillance and caused a dramatic re-lapse that has lasted until today (Franco *et al.*, 2014). The current diagnostic and therapeutic tools are really obsolete: suramin (*T. b. rhodesiense*), and pentamidine (*T. b. gambiense*) were developed in the 1920s and 1940s (World Health Organization, 2013) and are still being used today (Streit and Matsumoto, 2016). Stage 2 drugs are also very inconvenient: the current treatment for *T. b. gambiense*, NECT, requires hospitalization and frequent intravenous infusions (Eperon *et al.*, 2014); while the only available treatment for *T. b. rhodesiense*, melarsoprol, kills 5% of the patients (World Health Organization, 2013).

There is hope since 2 oral treatments (fexinidazole and SCYX-7158) entered clinical trials, phase II and III in 2016 (Drugs for Neglected Diseases initiative, 2016). The existing perspective of both drugs targeting *T. b. gambiense* and *T. b. rhodesiense* HAT regardless of the disease stage is very promising. Key strategic advantages would be the possibilities of treating patients without need of hospitalization, and avoidance of the painful lumbar puncture that determines the stage of the disease. However, it is a concern that rising resistant parasites against the current and prospective drugs is fairly easy in the laboratory (Jones *et al.*, 2015; Wyllie *et al.*, 2015). There is a need for a more diverse therapeutic arsenal -with different modes of actions, internalization, or extrusion routes- that would prevent also cross-resistance mechanisms in combination therapies.

In this thesis, I faced the drug discovery process from two complementary angles, which represent the opposite starting points towards identification of an active molecule. On the one hand, I identified potential drug targets in an RNAi screen, which defined protein kinases relevant for the establishment of the *in vivo* infection; and on the other, I identified compounds capable of killing the parasite but showing specific phenotypic features that permitted their linkage with a particular mode of action.

Coupling these two approaches may be a great perspective for drug discovery. First, RITseq screens of genome-wide libraries (or functionally-focused, as the one used here that targets protein kinases) may identify potential drug targets involved in essential biological processes, whose depletion elicit tractable phenotypes prior to causing death to the parasite. Then, compound-screening campaigns may follow in order to select for drugs able to reproduce such a phenotype, and potentially simplifying the often-difficult target deconvolution phase. The more information we have about components of the targeted pathway, the more rationally we can choose the chemotypes integrating the library intended to screen. Focused libraries (e.g. kinase focussed library) have advantages as they are small and specific to target, but disadvantages as they exclude novel chemotypes that are not predicted to be inhibitors. Screening whole compound collections as the 1.2 million from GSK (Peña *et al.*, 2015), also against the recombinant enzyme may be worth doing.

The early evolutionary divergence of *T. brucei* presents many opportunities for this method. Genes involved in divergent specific features of *T. brucei* biology as, for instance, unconventional regulation of the cell cycle, posttranscriptional regulation of gene expression, VSG switching, metabolism, or progression throughout the life-cycle should be accessible to this type of combined screens (Matthews, 2005; Hammarton, 2007; Clayton, 2014). With the same flow cytometry approach that we used to detect EP procyclin expression, we could select for RNAi phenotypes or compounds triggering PAD expression, disrupting the VSG coat, retaining the cell cycle in a particular stage, or dysregulating expression of the splice leader, for instance, prior to killing the parasite. We could design RITseq screens coupled to fluorescence–activated cell sorting to select for parasites able (or unable) to express these markers, and we could look for potentially druggable targets within these subpopulations showing a reduced fitness respect to uninduced controls. A mechanism-based phenotypic screen would follow trying to replicate the phenotype. We would need ideally a detectable variable highly specific for interruption of a pathway known to be essential that we can measure through a biochemical reaction, an antibody, or by generation of the library into a cell line, where a reporter element is endogenously labelled.

Besides potential drug target identification, RITseq has helped detecting proteins whose depletion generates resistance to trypanocidal molecules already identified in phenotypic screens (Alsford *et al.*, 2011, 2012). Such hits can be factors involved in mode of action, as transporters or activators whose depletion increases resistance of the parasite to the drug. Knowing these factors can assist greatly rational modelling of the drug candidate based on structure-activity relationship. When applied to resistant lines generated in the lab against a particular molecule, RITseq can theoretically identify the compound target if, for instance, the mechanism increasing tolerance to the compound was upregulation of the target itself.

Although the importance of protein phosphorylation observed in other eukaryotes is well conserved in *T. brucei*, many protein kinases are phosphorylated under physiological conditions, indicating the presence of intricate phosphorylation networks. Just a few mechanisms, however, have been described (Nett *et al.*, 2009). The pTL library, particularly in the pooled format, can be a great resource to explore the myriad of processes where protein kinases are implicated.

List of References

Van Den Abbeele, J., Claes, Y., van Bockstaele, D., Le Ray, D. and Coosemans, M. (1999) '*Trypanosoma brucei spp*. development in the tsetse fly: characterization of the post-mesocyclic stages in the foregut and proboscis.', *Parasitology*, 118(5), pp. 469–78. doi: 10.1017/S0031182099004217.

Abraham, R. T. (2004) 'PI 3-kinase related kinases: "Big" players in stress-induced signaling pathways', *DNA Repair*, 3(8–9), pp. 883–887. doi: 10.1016/j.dnarep.2004.04.002.

Acosta-Serrano, a, Vassella, E., Liniger, M., Kunz Renggli, C., Brun, R., Roditi, I. and Englund, P. T. (2001) 'The surface coat of procyclic *Trypanosoma brucei*: programmed expression and proteolytic cleavage of procyclin in the tsetse fly.', *Proceedings of the National Academy of Sciences of the United States of America*, 98(4), pp. 1513–8. doi: 10.1073/pnas.041611698.

Adams, J. a. (2001) 'Kinetic and catalytic mechanisms of protein kinases', *Chemical Reviews*, 101(858), pp. 2271–2290. doi: 10.1021/cr000230w.

Agron, P. G., Reed, S. L. and Engel, J. N. (2005) 'An essential, putative MEK kinase of *Leishmania major*', *Molecular and Biochemical Parasitology*, 142(1), pp. 121–125. doi: 10.1016/j.molbiopara.2005.03.007.

Agwuh, K. N. and MacGowan, A. (2006) 'Pharmacokinetics and pharmacodynamics of the tetracyclines including glycylcyclines', *Journal of Antimicrobial Chemotherapy*, pp. 256–265. doi: 10.1093/jac/dkl224.

Akiyoshi, B. and Gull, K. (2014) 'Discovery of unconventional kinetochores in kinetoplastids.', *Cell*. Elsevier, 156(6), pp. 1247–58. doi: 10.1016/j.cell.2014.01.049.

Aksoy, S., Gibson, W. C. and Lehane, M. J. (2003) 'Interactions between tsetse and trypanosomes with implications for the control of trypanosomiasis.', *Advances in parasitology*, 53, pp. 1–83. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14587696.

Alsford, S., Currier, R. B., Guerra-Assunção, J. A., Clark, T. G. and Horn, D. (2014) 'Cathepsin-L Can Resist Lysis by Human Serum in *Trypanosoma brucei brucei*', *PLoS Pathogens*, 10(5). doi: 10.1371/journal.ppat.1004130.

Alsford, S., Eckert, S., Baker, N., Glover, L., Sanchez-flores, A., Leung, K. F., Turner, D. J., Field, M. C., Berriman, M. and Horn, D. (2012) 'High-throughput decoding of anti-trypanosomal drug efficacy and resistance', *Nature*, 482(7384), pp. 232–236. doi: 10.1038/nature10771.High-throughput.

Alsford, S. and Horn, D. (2008) 'Single-locus targeting constructs for reliable regulated RNAi and transgene expression in *Trypanosoma brucei*.', *Molecular and biochemical parasitology*, 161(1), pp. 76–9. doi: 10.1016/j.molbiopara.2008.05.006.

Alsford, S., Kawahara, T., Glover, L. and Horn, D. (2005) 'Tagging a *T. brucei* RRNA locus improves stable transfection efficiency and circumvents inducible expression position effects.', *Molecular and biochemical parasitology*, 144(2), pp. 142–8. doi: 10.1016/j.molbiopara.2005.08.009.

Alsford, S., Turner, D. J., Obado, S. O., Sanchez-flores, A., Glover, L., Berriman, M., Hertz-fowler, C. and Horn, D. (2011) 'High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome', *Genome research*, 21, pp. 915–924. doi: 10.1101/gr.115089.110.Freely.

Altschul, S. F., Madden, T. L., Schäffer, A. a., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) 'Gapped BLAST and PSI-BLAST: A new generation of protein database search programs', *Nucleic Acids Research*, 25(17), pp. 3389–3402. doi: 10.1093/nar/25.17.3389.

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) 'Basic local alignment search tool.', *Journal of Molecular Biology*, 215(3), pp. 403–410. doi: 10.1016/S0022-2836(05)80360-2.

Alvarez, V. E., Kosec, G., Sant'Anna, C., Turk, V., Cazzulo, J. J. and Turk, B. (2008) 'Blocking autophagy to prevent parasite differentiation: A possible new strategy for fighting parasitic infections?', *Autophagy*, 4(3), pp. 361–363. doi:

10.1074/jbc.M708474200.

Amata, E., Xi, H., Colmenarejo, G., Gonzalez-Diaz, R., Cordon-Obras, C., Berlanga, M., Manzano, P., Erath, J., Roncal, N. E., Lee, P. J., Leed, S. E., Rodriguez, A., Sciotti, R. J., Navarro, M. and Pollastri, M. P. (2016) 'Identification of "preferred" Human Kinase Inhibitors for Sleeping Sickness Lead Discovery. Are Some Kinases Better than Others for Inhibitor Repurposing?', *ACS Infectious Diseases*, 2(3), pp. 180–186. doi: 10.1021/acsinfecdis.5b00136.

Amiguet-Vercher, A., Pérez-Morga, D., Pays, A., Poelvoorde, P., Van Xong, H., Tebabi, P., Vanhamme, L. and Pays, E. (2004) 'Loss of the mono-allelic control of the VSG expression sites during the development of *Trypanosoma brucei* in the bloodstream', *Molecular Microbiology*, 51(6), pp. 1577–1588. doi: 10.1111/j.1365-2958.2003.03937.x.

Anding, a L. and Baehrecke, E. H. (2015) 'Vps15 is required for stress induced and developmentally triggered autophagy and salivary gland protein secretion in *Drosophila*.', *Cell death and differentiation*. Nature Publishing Group, 22(3), pp. 457–64. doi: 10.1038/cdd.2014.174.

Araki, Y., Ku, W. C., Akioka, M., May, A. I., Hayashi, Y., Arisaka, F., Ishihama, Y. and Ohsumi, Y. (2013) 'Atg38 is required for autophagy-specific phosphatidylinositol 3-kinase complex integrity', *Journal of Cell Biology*, 203(2), pp. 299–313. doi: 10.1083/jcb.201304123.

Archer, S. K. (2009) 'Standard culture medium allows clonal dilution of *Trypanosoma brucei* procyclic cells after auto-conditioning', *Molecular and Biochemical Parasitology*, 164(1), pp. 100–103. doi: 10.1016/j.molbiopara.2008.11.003.

Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., Depledge, D. P., Fischer, S., Gajria, B., Gao, X., Gardner, M. J., Gingle, a., Grant, G., Harb, O. S., Heiges, M., Hertz-Fowler, C., Houston, R., Innamorato, F., Iodice, J., Kissinger, J. C., Kraemer, E., Li, W., Logan, F. J., Miller, J. a., Mitra, S., Myler, P. J., Nayak, V., Pennington, C., Phan, I., Pinney, D. F., Ramasamy, G., Rogers, M. B., Roos, D. S., Ross, C., Sivam, D., Smith, D.

F., Srinivasamoorthy, G., Stoeckert, C. J., Subramanian, S., Thibodeau, R., Tivey, a., Treatman, C., Velarde, G. and Wang, H. (2010) 'TriTrypDB: a functional genomic resource for the Trypanosomatidae', *Nucleic Acids Research*, 38(Database), pp. D457–D462. doi: 10.1093/nar/gkp851.

Atayde, V. D., Tschudi, C. and Ullu, E. (2011) 'The emerging world of small silencing RNAs in protozoan parasites', *Trends in Parasitology*. Elsevier Ltd, 27(7), pp. 321–327. doi: 10.1016/j.pt.2011.03.002.

Badjatia, N., Park, S. H., Ambrósio, D. L., Kirkham, J. K. and Günzl, A. (2016) 'Cyclin-Dependent Kinase CRK9, Required for Spliced Leader trans Splicing of Pre-mRNA in Trypanosomes, Functions in a Complex with a New L-Type Cyclin and a Kinetoplastid-Specific Protein', *PLOS Pathogens*, 12(3), p. e1005498. doi: 10.1371/journal.ppat.1005498.

Bähler, J. and Nurse, P. (2001) 'Fission yeast Pom1p kinase activity is cell cycle regulated and essential for cellular symmetry during growth and division.', *The EMBO journal*, 20(5), pp. 1064–73. doi: 10.1093/emboj/20.5.1064.

Bajusz, D., Rácz, A. and Héberger, K. (2015) 'Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations?', *Journal of Cheminformatics*, 7(1), p. 20. doi: 10.1186/s13321-015-0069-3.

Baker, N., Alsford, S. and Horn, D. (2011) 'Genome-wide RNAi screens in African trypanosomes identify the nifurtimox activator NTR and the eflornithine transporter AAT6.', *Molecular and biochemical parasitology*. Elsevier B.V., 176, pp. 55–7. doi: 10.1016/j.molbiopara.2010.11.010.

Bantscheff, M., Eberhard, D., Abraham, Y., Bastuck, S., Boesche, M., Hobson, S., Mathieson, T., Perrin, J., Raida, M., Rau, C., Reader, V., Sweetman, G., Bauer, A., Bouwmeester, T., Hopf, C., Kruse, U., Neubauer, G., Ramsden, N., Rick, J., Kuster, B. and Drewes, G. (2007) 'Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors', *Nature biotechnology*, 25(9), pp. 1035–1044. doi: 10.1038/nbt1328.

Barnes, R. L., Shi, H., Kolev, N. G., Tschudi, C. and Ullu, E. (2012) 'Comparative

genomics reveals two novel RNAi factors in *Trypanosoma brucei* and provides insight into the core machinery', *PLoS Pathogens*, 8(5). doi: 10.1371/journal.ppat.1002678.

Barouch-bentov, R. and Sauer, K. (2012) 'Mechanisms of Drug-Resistance in Kinases', *Expert Opinion on Investigational Drugs*, 20(2), pp. 153–208. doi: 10.1517/13543784.2011.546344.Mechanisms.

Barquilla, A., Crespo, J. L. and Navarro, M. (2008) 'Rapamycin inhibits trypanosome cell growth by preventing TOR complex 2 formation', *Proceedings of the National Academy of Sciences*, 105(38), pp. 14579–14584. doi: 10.1073/pnas.0802668105.

Barquilla, A. and Navarro, M. (2009a) 'Trypanosome TOR as a major regulator of cell growth and autophagy', *Autophagy*, 5(2), pp. 256–258. doi: 10.4161/auto.5.2.7591.

Barquilla, A. and Navarro, M. (2009b) 'Trypanosome TOR complex 2 functions in cytokinesis', *Cell Cycle*, 8(5), pp. 697–699. doi: 10.4161/cc.8.5.7808.

Barquilla, A., Saldivia, M., Diaz, R., Bart, J.-M., Vidal, I., Calvo, E., Hall, M. N. and Navarro, M. (2012) 'Third target of rapamycin complex negatively regulates development of quiescence in *Trypanosoma brucei*.', *Proceedings of the National Academy of Sciences of the United States of America*, 109(36), pp. 14399–404. doi: 10.1073/pnas.1210465109.

Barrett, M. P. (1999) 'The fall and rise of sleeping sickness.', *Lancet (London, England)*, 353(9159), pp. 1113–4. doi: 10.1016/S0140-6736(98)00416-4.

Basnet, H., Su, X. B., Tan, Y., Meisenhelder, J., Merkurjev, D., Ohgi, K. A., Hunter, T., Pillus, L. and Rosenfeld, M. G. (2014) 'Tyrosine phosphorylation of histone H2A by CK2 regulates transcriptional elongation', *Nature*, 516(7530), pp. 267–271. doi: 10.1038/nature13736.

Batram, C., Jones, N. G., Janzen, C. J., Markert, S. M. and Engstler, M. (2014) 'Expression site attenuation mechanistically links antigenic variation and

development in *Trypanosoma brucei*', *eLife*, 3(3:e02324), pp. 1–18. doi: 10.7554/eLife.02324.

Behera, R., Thomas, S. M. and Mensa-Wilmot, K. (2014) 'New chemical scaffolds for human African trypanosomiasis lead discovery from a screen of tyrosine kinase inhibitor drugs', *Antimicrobial Agents and Chemotherapy*, 58(4), pp. 2202–2210. doi: 10.1128/AAC.01691-13.

Ben-David, Y., Letwin, K., Tannock, L., Bernstein, a and Pawson, T. (1991) 'A mammalian protein kinase with potential for serine/threonine and tyrosine phosphorylation is related to cell cycle regulators.', *The EMBO journal*, 10(2), pp. 317–25. PMID: 1825055

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., Böhme, U., Hannick, L., Aslett, M. a, Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U. C. M., Arrowsmith, C., Atkin, R. J., Barron, A. J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T.-J., Churcher, C., Clark, L. N., Corton, C. H., Cronin, A., Davies, R. M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M. C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B. R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A. X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., Macleod, A., Mooney, P. J., Moule, S., Martin, D. M. a, Morgan, G. W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C. S., Peterson, J., Quail, M. a, Rabbinowitsch, E., Rajandream, M.-A., Reitter, C., Salzberg, S. L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A. J., Tallon, L., Turner, C. M. R., Tait, A., Tivey, A. R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M. D., Embley, T. M., Gull, K., Ullu, E., Barry, J. D., Fairlamb, A. H., Opperdoes, F., Barrell, B. G., Donelson, J. E., Hall, N., Fraser, C. M., Melville, S. E. and El-Sayed, N. M. (2005) 'The genome of the African trypanosome *Trypanosoma brucei*.', *Science (New York, N.Y.)*, 309(5733), pp. 416–22. doi: 10.1126/science.1112642.

Besteiro, S., Williams, R. A. M., Morrison, L. S., Coombs, G. H. and Mottram, J. C. (2006) 'Endosome sorting and autophagy are essential for differentiation and

virulence of *Leishmania major*', *Journal of Biological Chemistry*, 281(16), pp. 11384–11396. doi: 10.1074/jbc.M512307200.

Biéler, S., Waltenberger, H., Barrett, M. P., McCulloch, R., Mottram, J. C., Carrington, M., Schwaeble, W., McKerrow, J., Phillips, M. A., Michels, P. A., Büscher, P., Sanchez, J.-C., Bishop, R., Robinson, D. R., Bangs, J., Ferguson, M., Nerima, B., Albertini, A., Michel, G., Radwandska, M. and Ndung'u, J. M. (2016) 'Evaluation of Antigens for Development of a Serological Test for Human African Trypanosomiasis.', *PloS one*, 11(12), p. e0168074. doi: 10.1371/journal.pone.0168074.

Bierhoff, H., Dundr, M., Michels, A. A. and Grummt, I. (2008) 'Phosphorylation by casein kinase 2 facilitates rRNA gene transcription by promoting dissociation of TIF-IA from elongating RNA polymerase I', *Molecular and Cellular Biology*, 28(16), pp. 4988–98. doi: 10.1128/MCB.00492-08.

Bishop, A. C., Ubersax, J. a, Petsch, D. T., Matheos, D. P., Gray, N. S., Blethrow, J., Shimizu, E., Tsien, J. Z., Schultz, P. G., Rose, M. D., Wood, J. L., Morgan, D. O. and Shokat, K. M. (2000) 'A chemical switch for inhibitor-sensitive alleles of any protein kinase', *Nature*, 407(6802), pp. 395–401. doi: 10.1038/35030148.

Blenis, J. and Resh, M. D. (1993) 'Subcellular localization specified by protein acylation and phosphorylation.', *Current opinion in cell biology*, 5(6), pp. 984–9. PMID: 8129952.

Bouyer, J., Seck, M. T. and Sall, B. (2013) 'Misleading guidance for decision making on tsetse eradication: Response to Shaw et al. (2013)', *Preventive Veterinary Medicine*, 112(3–4), pp. 443–446. doi: 10.1016/j.prevetmed.2013.05.017.

Bradford, M. M. (1976) 'A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.', *Analytical biochemistry*, 72, pp. 248–54. PMID: 942051.

Brecht, M. and Parsons, M. (1998) 'Changes in polysome profiles accompany

trypanosome development', *Molecular and Biochemical Parasitology*, 97(1–2), pp. 189–198. doi: 10.1016/S0166-6851(98)00151-0.

Brown, J. R., Koretke, K. K., Birkeland, M. L., Sanseau, P. and Patrick, D. R. (2004) 'Evolutionary relationships of Aurora kinases: implications for model organism studies and the development of anti-cancer drugs.', *BMC evolutionary biology*, 4, p. 39. doi: 10.1186/1471-2148-4-39.

Brun, R., Blum, J., Chappuis, F. and Burri, C. (2010) 'Human African trypanosomiasis.', *Lancet*. Elsevier Ltd, 375(January), pp. 148–59. doi: 10.1016/S0140-6736(09)60829-1.

Brun, R., Don, R., Jacobs, R. T., Wang, M. Z. and Barrett, M. P. (2011) 'Development of novel drugs for human African trypanosomiasis.', *Future microbiology*, 6(6), pp. 677–91. doi: 10.2217/fmb.11.44.

Brun, R. and Schonenberger, M. (1981) 'Stimulating Effect of Citrate and Cis-Aconitate on the Transformation of *Trypanosoma brucei* Bloodstream Forms to Procyclic Forms in Vitro', *Zeitschrift fur Parasitenkunde*, 66, pp. 17–24.

Brushia, R. J. and Walsh, D. A. (1999) 'Phosphorylase kinase: the complexity of its regulation is reflected in the complexity of its structure.', *Frontiers in bioscience : a journal and virtual library*, 4, pp. D618-41. PMID: 10487978.

Buhrow, S. A., Cohen, S., Garbers, D. L. and Staros, J. V (1983) 'Characterization of the interaction of 5'-p-fluorosulfonylbenzoyl adenosine with the epidermal growth factor receptor/protein kinase in A431 cell membranes.', *The Journal of biological chemistry*, 258(12), pp. 7824–7. PMID: 6305956.

Burbank, L. P. and Stenger, D. (2016) 'A temperature-independent cold-shock protein homolog acts as a virulence factor in *Xylella fastidiosa*.', *Molecular plant-microbe interactions : MPMI*, 1, p. 26808446. doi: 10.1094/MPMI-11-15-0260-R.

Burnett, G. and Kennedy, E. P. (1954) 'The enzymatic phosphorylation of proteins', *Journal of Biological Chemistry*, (211), pp. 969–980.

Bütikofer, P., Vassella, E., Ruepp, S., Boschung, M., Civenni, G., Seebeck, T., Hemphill, A., Mookherjee, N., Pearson, T. W. and Roditi, I. (1999) 'Phosphorylation of a major GPI-anchored surface protein of *Trypanosoma brucei* during transport to the plasma membrane', *Journal of cell science*, (112), pp. 1785–1795.

Cameron, A., Frirdich, E., Huynh, S., Parker, C. T. and Gaynor, E. C. (2012) 'Hyperosmotic Stress Response of *Campylobacter jejuni*', *Journal of Bacteriology*, 194(22), pp. 6116–6130. doi: 10.1128/JB.01409-12.

Capewell, P., Cren-Travaillé, C., Marchesi, F., Johnston, P., Clucas, C., Benson, R. A., Gorman, T.-A., Calvo-Alvarez, E., Crouzols, A., Jouvion, G., Jamonneau, V., Weir, W., Stevenson, M. L., O'Neill, K., Cooper, A., Swar, N. K., Bucheton, B., Ngoyi, D. M., Garside, P., Rotureau, B. and MacLeod, A. (2016) 'The skin is a significant but overlooked anatomical reservoir for vector-borne African trypanosomes', *eLife*, 5. doi: 10.7554/eLife.17716.

Capewell, P., Monk, S., Ivens, A., MacGregor, P., Fenn, K., Walrad, P., Bringaud, F., Smith, T. K. and Matthews, K. R. (2013) 'Regulation of *Trypanosoma brucei* Total and Polysomal mRNA during Development within Its Mammalian Host', *PLoS ONE*, 8(6). doi: 10.1371/journal.pone.0067069.

Carling, D., Mayer, F. V, Sanders, M. J. and Gamblin, S. J. (2011) 'AMP-activated protein kinase: nature's energy sensor.', *Nature chemical biology*, 7(8), pp. 512–518. doi: 10.1038/nchembio.610.

Carmena, M., Ruchaud, S. and Earnshaw, W. C. (2009) 'Making the Auroras glow : regulation of Aurora A and B kinase function by interacting proteins', pp. 796–805. doi: 10.1016/j.ceb.2009.09.008.

Cassola, A., De Gaudenzi, J. G. and Frasch, A. C. (2007) 'Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes.', *Molecular microbiology*, 65(3), pp. 655–70. doi: 10.1111/j.1365-2958.2007.05833.x.

Chan, K. F., Hurst, M. O. and Graves, D. J. (1982) 'Phosphorylase kinase specificity. A comparative study with cAMP-dependent protein kinase on

synthetic peptides and peptide analogs of glycogen synthase and phosphorylase.', *The Journal of biological chemistry*, 257(7), pp. 3655–9. PMID: 6277942.

Checchi, F., Cox, A. P., Chappuis, F., Priotto, G., Chandramohan, D. and Haydon, D. T. (2012) 'Prevalence and under-detection of gambiense human African trypanosomiasis during mass screening sessions in Uganda and Sudan.', *Parasites & vectors*, 5, p. 157. doi: 10.1186/1756-3305-5-157.

Chhajer, R., Bhattacharyya, A., Didwania, N., Shadab, M., Das, N., Palit, P., Vaidya, T. and Ali, N. (2016) '*Leishmania donovani* Aurora kinase: A promising therapeutic target against visceral leishmaniasis.', *Biochimica et biophysica acta*, 1860(9), pp. 1973–88. doi: 10.1016/j.bbagen.2016.06.005.

Clayton, C. E. (2014) 'Networks of gene expression regulation in T*rypanosoma brucei*', *Molecular and biochemical parasitology*, pp. 96–106. doi: 10.1016/j.molbiopara.2014.06.005.

Clayton, C. and Michaeli, S. (2011) '3′ processing in protists', *Wiley Interdisciplinary Reviews: RNA*, 2(2), pp. 247–255. doi: 10.1002/wrna.49.

Cohen, P. (2000) 'The regulation of protein function by multisite phosphorylation--a 25 year update.', *Trends in biochemical sciences*, 25(12), pp. 596–601. doi: 10.1016/S0968-0004(00)01712-6.

Cohen, P. (2002) 'Protein kinases--the major drug targets of the twenty-first century?', *Nature reviews. Drug discovery*, 1(April), pp. 309–315. doi: 10.1038/nrd773.

Cohen, P. and Tcherpakov, M. (2010) 'Will the ubiquitin system furnish as many drug targets as protein kinases?', *Cell*. Elsevier Inc., 143(5), pp. 686–693. doi: 10.1016/j.cell.2010.11.016.

Cole, A., Frame, S. and Cohen, P. (2004) 'Further evidence that the tyrosine phosphorylation of glycogen synthase kinase-3 (GSK3) in mammalian cells is an autophosphorylation event', *Biochem. J*, 377, pp. 249–255. doi:

10.1042/BJ20031259.

Coles, J. A., Myburgh, E., Ritchie, R., Hamilton, A., Rodgers, J., Mottram, J. C., Barrett, M. P. and Brewer, J. M. (2015) 'Intravital imaging of a massive lymphocyte response in the cortical dura of mice after peripheral infection by trypanosomes.', *PLoS neglected tropical diseases*, 9(4), p. e0003714. doi: 10.1371/journal.pntd.0003714.

Cook, P. F., Neville, M. E., Vrana, K. E., Hartl, F. T. and Roskoski, R. (1982) 'Adenosine cyclic 3',5'-monophosphate dependent protein kinase: kinetic mechanism for the bovine skeletal muscle catalytic subunit.', *Biochemistry*, 21(23), pp. 5794–5799. doi: 10.1021/bi00266a011.

Cross, G. A. M., Kim, H. S. and Wickstead, B. (2014) 'Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427', *Molecular and Biochemical Parasitology*, 195(1), pp. 59–73. doi: 10.1016/j.molbiopara.2014.06.004.

D'Assoro, A. B., Haddad, T. and Galanis, E. (2015) 'Aurora-A Kinase as a Promising Therapeutic Target in Cancer.', *Frontiers in oncology*, 5(January), p. 295. doi: 10.3389/fonc.2015.00295.

Damerow, M., Graalfs, F., Güther, M. L. S., Mehlert, A., Izquierdo, L. and Ferguson, M. A. J. (2016) 'A gene of the β3-glycosyltransferase family encodes N-acetylglucosaminyltransferase II function in *Trypanosoma brucei*', *Journal of Biological Chemistry*, 291(26), pp. 13834–13845. doi: 10.1074/jbc.M116.733246.

Daniels, J. P., Gull, K. and Wickstead, B. (2010) 'Cell Biology of the Trypanosome Genome', *Microbiology and Molecular Biology Reviews*, 74(4), pp. 552–569. doi: 10.1128/mmbr.00024-10.

Das, A., Zhang, Q., Palenchar, J. B., Chatterjee, B., Cross, G. A. M. and Bellofatto, V. (2005) 'Trypanosomal TBP Functions with the Multisubunit Transcription Factor tSNAP To Direct Spliced-Leader RNA Gene Expression', *Molecular and Cellular Biology*, 25(16), pp. 7314–7322. doi: 10.1128/MCB.25.16.7314.

Davis, A. T., Wang, H., Zhang, P. and Ahmed, K. (2002) 'Heat shock mediated modulation of protein kinase CK2 in the nuclear matrix', *Journal of Cellular Biochemistry*, 85(3), pp. 583–591. doi: 10.1002/jcb.10158.

Dean, S., Gould, M. K., Dewar, C. E. and Schnaufer, A. C. (2013) 'Single point mutations in ATP synthase compensate for mitochondrial genome loss in trypanosomes.', *Proceedings of the National Academy of Sciences of the United States of America*, 110(36), pp. 14741–6. doi: 10.1073/pnas.1305404110.

Dean, S., Marchetti, R., Kirk, K. and Matthews, K. R. (2009) 'A surface transporter family conveys the trypanosome differentiation signal.', *Nature*. Nature Publishing Group, 459(7244), pp. 213–7. doi: 10.1038/nature07997.

Deborggraeve, S., Koffi, M., Jamonneau, V., Bonsu, F. A., Queyson, R., Simarro, P. P., Herdewijn, P. and B??scher, P. (2008) 'Molecular analysis of archived blood slides reveals an atypical human Trypanosoma infection', *Diagnostic Microbiology and Infectious Disease*, 61(4), pp. 428–433. doi: 10.1016/j.diagmicrobio.2008.03.006.

Deininger, M. and Druker, B. (2003) 'Specific targeted therapy of chronic myelogenous leukemia with imatinib', *Pharmacological reviews*, 55(3), pp. 401–423. doi: 10.1124/pr.55.3.4.401.

Deitsch, K. W., Lukehart, S. a and Stringer, J. R. (2009) 'Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens.', *Nature reviews. Microbiology*. Nature Publishing Group, 7(July), pp. 493–503. doi: 10.1038/nrmicro2145.

Dejung, M., Subota, I., Bucerius, F., Dindar, G., Freiwald, A., Engstler, M., Boshart, M., Butter, F. and Janzen, C. J. (2016) 'Quantitative Proteomics Uncovers Novel Factors Involved in Developmental Differentiation of *Trypanosoma brucei*', *PLoS Pathogens*, 12(2), pp. 1–20. doi: 10.1371/journal.ppat.1005439.

Delespaux, V. and de Koning, H. P. (2007) 'Drugs and drug resistance in African trypanosomiasis', *Drug Resistance Updates*, 10(1–2), pp. 30–50. doi:

10.1016/j.drup.2007.02.004.

Diaz-Gonzalez, R., Kuhlmann, F. M., Galan-Rodriguez, C., da Silva, L. M., Saldivia, M., Karver, C. E., Rodriguez, A., Beverley, S. M., Navarro, M. and Pollastri, M. P. (2011) 'The susceptibility of trypanosomatid pathogens to PI3/mTOR kinase inhibitors affords a new opportunity for drug repurposing', *PLoS Neglected Tropical Diseases*, 5(8), pp. 1–11. doi: 10.1371/journal.pntd.0001297.

Diaz, R., Luengo-Arratta, S. a, Seixas, J. D., Amata, E., Devine, W., Cordon-Obras, C., Rojas-Barros, D. I., Jimenez, E., Ortega, F., Crouch, S., Colmenarejo, G., Fiandor, J. M., Martin, J. J., Berlanga, M., Gonzalez, S., Manzano, P., Navarro, M. and Pollastri, M. P. (2014) 'Identification and Characterization of Hundreds of Potent and Selective Inhibitors of *Trypanosoma brucei* Growth from a Kinase-Targeted Library Screening Campaign.', *PLoS neglected tropical diseases*, 8(10), p. e3253. doi: 10.1371/journal.pntd.0003253.

Dissmeyer, N. and Schnittger, A. (2011) 'Use of phospho-site substitutions to analyze the biological relevance of phosphorylation events in regulatory networks.', *Methods in molecular biology (Clifton, N.J.)*, 779, pp. 93–138. doi: 10.1007/978-1-61779-264-9_6.

Domenicali Pfister, D., Burkard, G., Morand, S., Renggli, C. K., Roditi, I. and Vassella, E. (2006) 'A Mitogen-activated protein kinase controls differentiation of bloodstream forms of Trypanosoma brucei.', *Eukaryotic cell*, 5(7), pp. 1126–35. doi: 10.1128/EC.00094-06.

Domingo-Sananes, M. R., Szoor, B., Ferguson, M. a. J., Urbaniak, M. D. and Matthews, K. R. (2015) 'Molecular control of irreversible bistability during trypanosome developmental commitment', *The Journal of Cell Biology*, 211(2), pp. 455–468. doi: 10.1083/jcb.201506114.

Dow, L. E., Fisher, J., O'Rourke, K. P., Muley, A., Kastenhuber, E. R., Livshits, G., Tschaharganeh, D. F., Socci, N. D. and Lowe, S. W. (2015) 'Inducible in vivo genome editing with CRISPR-Cas9', *Nature Biotechnology*, 33(4), pp. 390–394. doi: 10.1038/nbt.3155.

Dranchak, P., MacArthur, R., Guha, R., Zuercher, W. J., Drewry, D. H., Auld, D. S. and Inglese, J. (2013) 'Profile of the GSK Published Protein Kinase Inhibitor Set Across ATP-Dependent and-Independent Luciferases: Implications for Reporter-Gene Assays', *PLoS ONE*, 8(3), pp. 1–9. doi: 10.1371/journal.pone.0057888.

Dreesen, O., Li, B. and Cross, G. A. M. (2006) 'Telomere structure and function in trypanosomes: a proposal', *Nature Reviews Microbiology*, 5(1), pp. 70–75. doi: 10.1038/nrmicro1577.

Drewry, D. H., Willson, T. M. and Zuercher, W. J. (2014) 'Seeding collaborations to advance kinase science with the GSK Published Kinase Inhibitor Set (PKIS)', *Curr Top Med Chem*, 14(3), pp. 340–342. doi: 10.2174/1568026613666131127160819.

Drugs for Neglected Diseases initiative (2015) *SLEEPING SICKNESS, TARGET PRODUCT PROFILE*. Available at: http://www.dndi.org/diseases-projects/hat/hat-target-product-profile/ (Accessed: 1 December 2016).

Drugs for Neglected Diseases initiative (2016) *Human African Trypanosomiasis. Fact Sheet*.

Durand-Dubief, M. and Bastin, P. (2003) 'TbAGO1, an argonaute protein required for RNA interference, is involved in mitosis and chromosome segregation in *Trypanosoma brucei*.', *BMC biology*, 1, p. 2. doi: 10.1186/1741-7007-1-2.

Durand-Dubief, M., Kohl, L. and Bastin, P. (2003) 'Efficiency and specificity of RNA interference generated by intra- and intermolecular double stranded RNA in *Trypanosoma brucei*', *Molecular and Biochemical Parasitology*, 129(1), pp. 11–21. doi: 10.1016/S0166-6851(03)00071-9.

Duszenko, M., Ginger, M. L., Brennand, A., Gualdr??n-L??pez, M., Colombo, M. I., Coombs, G. H., Coppens, I., Jayabalasingham, B., Langsley, G., De Castro, S. L., Menna-Barreto, R., Mottram, J. C., Navarro, M., Rigden, D. J., Romano, P. S., Stoka, V., Turk, B. and Michels, P. A. M. (2011) 'Autophagy in protists', *Autophagy*, 7(2), pp. 127–158. doi: 10.4161/auto.7.2.13310.

Dyer, N. A., Rose, C., Ejeh, N. O. and Acosta-Serrano, A. (2013) 'Flying tryps: Survival and maturation of trypanosomes in tsetse flies', *Trends in Parasitology*. Elsevier Ltd, 29(4), pp. 188–196. doi: 10.1016/j.pt.2013.02.003.

Edgar, R. C. (2004) 'MUSCLE: multiple sequence alignment with high accuracy and high throughput', *Nucleic Acid Research*, 32(5), pp. 1792–1797. doi: 10.1093/nar/gkh340.

Ekwanzala, M., Pépin, J., Khonde, N., Molisho, S., Bruneel, H. and De Wals, P. (1996) 'In the heart of darkness: Sleeping sickness in Zaire', *Lancet*, 348(9039), pp. 1427–1430. doi: 10.1016/S0140-6736(96)06088-6.

El-Sayed, N. M., Hegde, P., Quackenbush, J., Melville, S. E. and Donelson, J. E. (2000) 'The African trypanosome genome.', *International journal for parasitology*, 30(4), pp. 329–45. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10731558.

El-Sayed, N. M., Myler, P. J., Bartholomeu, D. C., Nilsson, D., Aggarwal, G., Tran, A.-N., Ghedin, E., Worthey, E. a, Delcher, A. L., Blandin, G., Westenberger, S. J., Caler, E., Cerqueira, G. C., Branche, C., Haas, B., Anupama, A., Arner, E., Aslund, L., Attipoe, P., Bontempi, E., Bringaud, F., Burton, P., Cadag, E., Campbell, D. a, Carrington, M., Crabtree, J., Darban, H., da Silveira, J. F., de Jong, P., Edwards, K., Englund, P. T., Fazelina, G., Feldblyum, T., Ferella, M., Frasch, A. C., Gull, K., Horn, D., Hou, L., Huang, Y., Kindlund, E., Klingbeil, M., Kluge, S., Koo, H., Lacerda, D., Levin, M. J., Lorenzi, H., Louie, T., Machado, C. R., McCulloch, R., McKenna, A., Mizuno, Y., Mottram, J. C., Nelson, S., Ochaya, S., Osoegawa, K., Pai, G., Parsons, M., Pentony, M., Pettersson, U., Pop, M., Ramirez, J. L., Rinta, J., Robertson, L., Salzberg, S. L., Sanchez, D. O., Seyler, A., Sharma, R., Shetty, J., Simpson, A. J., Sisk, E., Tammi, M. T., Tarleton, R., Teixeira, S., Van Aken, S., Vogt, C., Ward, P. N., Wickstead, B., Wortman, J., White, O., Fraser, C. M., Stuart, K. D. and Andersson, B. (2005) 'The genome sequence of Trypanosoma cruzi, etiologic agent of Chagas disease.', *Science (New York, N.Y.)*, 309(5733), pp. 409–15. doi: 10.1126/science.1112631.

El-Sayed, N. M., Myler, P. J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal,

G., Caler, E., Renauld, H., Worthey, E. a, Hertz-Fowler, C., Ghedin, E., Peacock, C., Bartholomeu, D. C., Haas, B. J., Tran, A.-N., Wortman, J. R., Alsmark, U. C. M., Angiuoli, S., Anupama, A., Badger, J., Bringaud, F., Cadag, E., Carlton, J. M., Cerqueira, G. C., Creasy, T., Delcher, A. L., Djikeng, A., Embley, T. M., Hauser, C., Ivens, A. C., Kummerfeld, S. K., Pereira-Leal, J. B., Nilsson, D., Peterson, J., Salzberg, S. L., Shallom, J., Silva, J. C., Sundaram, J., Westenberger, S., White, O., Melville, S. E., Donelson, J. E., Andersson, B., Stuart, K. D. and Hall, N. (2005) 'Comparative genomics of trypanosomatid parasitic protozoa.', *Science (New York, N.Y.)*, 309(5733), pp. 404–9. doi: 10.1126/science.1112181.

Elia, A. E. H., Rellos, P., Haire, L. F., Chao, J. W., Ivins, F. J., Hoepker, K., Mohammad, D., Cantley, L. C., Smerdon, S. J. and Yaffe, M. B. (2003) 'The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain', *Cell*, 115(1), pp. 83–95. doi: 10.1016/S0092-8674(03)00725-6.

Ellinger-Ziegelbauer, H., Brown, K., Kelly, K. and Siebenlist, U. (1997) 'Direct activation of the stress-activated protein kinase (SAPK) and extracellular signal-regulated protein kinase (ERK) pathways by an inducible mitogen-activated protein kinase/ERK kinase kinase 3 (MEKK) derivative', *Journal of Biological Chemistry*, 272(5), pp. 2668–2674. doi: 10.1074/jbc.272.5.2668.

Ellis, D. S. and Evans, D. A. (1977) 'Passage of *Trypanosoma brucei rhodesiense* through the peritrophic membrane of Glossina morsitans morsitans.', *Nature*, 267(5614), pp. 834–5. PMID: 895841.

Elphick, L. M., Lee, S. E., Gouverneur, V. and Mann, D. J. (2007) 'Using Chemical Genetics and ATP Analogues To Dissect Protein Kinase Function', *ACS Chemical Biology*, 2(5), pp. 299–314. doi: 10.1021/cb700027u.

Engstler, M., Pfohl, T., Herminghaus, S., Boshart, M., Wiegertjes, G., Heddergott, N. and Overath, P. (2007) 'Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes.', *Cell*, 131(3), pp. 505–15. doi: 10.1016/j.cell.2007.08.046.

Eperon, G., Balasegaram, M., Potet, J., Mowbray, C., Valverde, O. and Chappuis, F. (2014) 'Treatment options for second-stage gambiense human African trypanosomiasis.', *Expert review of anti-infective therapy*, 12(11), pp. 1407–17. doi: 10.1586/14787210.2014.959496.

Erben, E. D., Fadda, A., Lueong, S., Hoheisel, J. D. and Clayton, C. (2014) 'A Genome-Wide Tethering Screen Reveals Novel Potential Post-Transcriptional Regulators in *Trypanosoma brucei*', *PLoS Pathogens*, 10(6). doi: 10.1371/journal.ppat.1004178.

Ersfeld, K. (2011) 'Nuclear architecture, genome and chromatin organisation in *Trypanosoma brucei*', *Research in Microbiology*, 162(6), pp. 626–636. doi: 10.1016/j.resmic.2011.01.014.

Evans, D. A. and Ellis, D. E. (1983) 'Recent observations on the behaviour of certain trypanosomes within their insect hosts', *Advances in Parasitology*, 22, pp. 1–42. doi: Doi 10.1016/S0065-308x(08)60460-1.

Farr, H. and Gull, K. (2012) 'Cytokinesis in trypanosomes', *Cytoskeleton*, pp. 931–941. doi: 10.1002/cm.21074.

Ferguson, M. a, Homans, S. W., Dwek, R. a and Rademacher, T. W. (1988) 'Glycosyl-phosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane.', *Science (New York, N.Y.)*, 239(4841 Pt 1), pp. 753–759. doi: 10.1126/science.3340856.

Ferrante, a and Allison, a C. (1983) 'Alternative pathway activation of complement by African trypanosomes lacking a glycoprotein coat.', *Parasite immunology*, 5(5), pp. 491–8. PMID: 6634218.

Field, M. C. and Carrington, M. (2004) 'Intracellular Membrane Transport Systems in *Trypanosoma brucei*', *Traffic*, 5(12), pp. 905–913. doi: 10.1111/j.1600-0854.2004.00234.x.

Field, M. C. and Carrington, M. (2009) 'The trypanosome flagellar pocket.', *Nature reviews. Microbiology*, 7(11), pp. 775–86. doi: 10.1038/nrmicro2221.

Field, M. C., Menon, A. K. and Cross, G. A. M. (1991) 'Developmental variation of glycosylphosphatidylinositol membrane anchors in *Trypanosoma brucei*: Identification of a candidate biosynthetic precursor of the glycosylphosphatidylinositol anchor of the major procyclic stage surface glycoprotein', *Journal of Biological Chemistry*, 266(13), pp. 8392–8400.

Figueiredo, R. C., Rosa, D. S. and Soares, M. J. (2000) 'Differentiation of *Trypanosoma cruzi* epimastigotes: metacyclogenesis and adhesion to substrate are triggered by nutritional stress.', *The Journal of parasitology*, 86(6), pp. 1213–1218. doi: 10.1645/0022-3395(2000)086[1213:DOTCEM]2.0.CO;2.

Finn, R. D., Bateman, a., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, a., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L. L., Tate, J. and Punta, M. (2014) 'Pfam: the protein families database', *Nucleic Acids Research*, 42(D1), pp. D222–D230. doi: 10.1093/nar/gkt1223.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. and Mello, C. C. (1998) 'Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans.', Nature,* 391(6669), pp. 806–811.

Ford, J. (1969) 'Control of the African trypanosomiases with special reference to land use.', *Bulletin of the World Health Organization*, 40(6), pp. 879–892.

Fotaki, V., Dierssen, M., Alcántara, S., Martínez, S., Martí, E., Casas, C., Visa, J., Soriano, E., Estivill, X. and Arbonés, M. L. (2002) 'Dyrk1A haploinsufficiency affects viability and causes developmental delay and abnormal brain morphology in mice.', *Molecular and cellular biology*, 22(18), pp. 6636–47. doi: 10.1128/MCB.22.18.6636.

Franco, J. R., Simarro, P. P., Diarra, a, Ruiz-Postigo, J. a and Jannin, J. G. (2014) 'The journey towards elimination of gambiense human African trypanosomiasis: not far, nor easy.', *Parasitology*, 141, pp. 748–60. doi: 10.1017/S0031182013002102.

Frett, B., Brown, R. V., Ma, M., Hu, W., Han, H. and Li, H. (2014) 'Therapeutic Melting Pot of Never in Mitosis Gene A Related Kinase 2 (Nek2): A Perspective on

Nek2 as an Oncology Target and Recent Advancements in Nek2 Small Molecule Inhibition', *Journal of Medicinal Chemistry*, 57(14), pp. 5835–5844. doi: 10.1021/jm401719n.

Frevert, U., Movila, A., Nikolskaia, O. V, Raper, J., Mackey, Z. B., Abdulla, M., McKerrow, J. and Grab, D. J. (2012) 'Early invasion of brain parenchyma by African trypanosomes.', *PloS one*, 7(8), p. e43913. doi: 10.1371/journal.pone.0043913.

Fry, A. M., O'Regan, L., Sabir, S. R. and Bayliss, R. (2012) 'Cell cycle regulation by the NEK family of protein kinases', *Journal of Cell Science*, 125(19), pp. 4423–4433. doi: 10.1242/jcs.111195.

Funamoto, S., Milan, K., Meili, R. and Firtel, R. A. (2001) 'Role of Phosphatidylinositol 3′ Kinase and a Downstream Pleckstrin Homology Domain–Containing Protein in Controlling Chemotaxis in *Dictyostelium*', *The Journal of Cell Biology*, 153(4), pp. 795–810. doi: 10.1083/jcb.153.4.795.

Garrington, T. P. and Johnson, G. L. (1999) 'Organization and regulation of mitogen-activated protein kinase signaling pathways', *Current Opinion in Cell Biology*, 11(2), pp. 211–218. doi: 10.1016/S0955-0674(99)80028-3.

Gerber, D. A., Souquere-Besse, S., Puvion, F., Dubois, M. F., Bensaude, O. and Cochet, C. (2000) 'Heat-induced relocalization of protein kinase CK2: Implication of CK2 in the context of cellular stress', *Journal of Biological Chemistry*, 275(31), pp. 23919–23926. doi: 10.1074/jbc.M002697200.

Gibbs, C. S. and Zoller, M. J. (1991) 'Rational scanning mutagenesis of a protein kinase identifies functional regions involved in catalysis and substrate interactions.', *The Journal of biological chemistry*, 266(14), pp. 8923–31. PMID: 2026604.

Gibson, W. and Bailey, M. (2003) 'The development of Trypanosoma brucei within the tsetse fly midgut observed using green fluorescent trypanosomes.', *Kinetoplastid Biology and Disease*, 2(1), p. 1. doi: 10.1186/1475-9292-2-1.

Gilbert, I. H. (2013) 'Drug Discovery for Neglected Diseases: Molecular Target-Based and Phenotypic Approaches', *Journal of medicinal chemistry*, 56, pp. 7719–7726. doi: 10.1021/jm400362b.

Gill, A. L., Verdonk, M., Boyle, R. G. and Taylor, R. (2007) 'A comparison of physicochemical property profiles of marketed oral drugs and orally bioavailable anti-cancer protein kinase inhibitors in clinical development.', *Current topics in medicinal chemistry*, 7(14), pp. 1408–22. doi: 10.2174/156802607781696819.

Glover, L., Alsford, S., Baker, N., Turner, D. J., Sanchez-Flores, A., Hutchinson, S., Hertz-Fowler, C., Berriman, M. and Horn, D. (2014) 'Genome-scale RNAi screens for high-throughput phenotyping in bloodstream-form African trypanosomes', *Nature Protocols*, 10(1), pp. 106–133. doi: 10.1038/nprot.2015.005.

Glover, L., Alsford, S. and Horn, D. (2013) 'DNA Break Site at Fragile Subtelomeres Determines Probability and Mechanism of Antigenic Variation in African Trypanosomes', *PLoS Pathogens*, 9(3). doi: 10.1371/journal.ppat.1003260.

Glover, L. and Horn, D. (2009) 'Site-specific DNA double-strand breaks greatly increase stable transformation efficiency in *Trypanosoma brucei*', *Molecular and Biochemical Parasitology*, 166(2), pp. 194–197. doi: 10.1016/j.molbiopara.2009.03.010.

Glover, L., Hutchinson, S., Alsford, S., Mcculloch, R., Field, M. C. and Horn, D. (2013) 'Antigenic variation in African trypanosomes: The importance of chromosomal and nuclear context in VSG expression control', *Cellular Microbiology*, 15(12), pp. 1984–1993. doi: 10.1111/cmi.12215.

Gluenz, E., Wheeler, R. J., Hughes, L. and Vaughan, S. (2015) 'Scanning and three-dimensional electron microscopy methods for the study of *Trypanosoma brucei* and *Leishmania mexicana* flagella BT - Methods in Cell Biology', in *Methods in Cell Biology*, pp. 509–542. doi: 10.1016/bs.mcb.2014.12.011.

Gough, J., Karplus, K., Hughey, R. and Chothia, C. (2001) 'Assignment of

homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure1', *Journal of Molecular Biology*, 313(4), pp. 903–919. doi: 10.1006/jmbi.2001.5080.

Graaf, K. De, Hekerman, P., Spelten, O., Herrmann, A., Packman, L. C., Bu, K., Mu, G. and Becker, W. (2004) 'Characterization of Cyclin L2 , a Novel Cyclin with an Arginine / Serine-rich Domain', 279(6), pp. 4612–4624. doi: 10.1074/jbc.M310794200.

Grab, D. J., Garcia-Garcia, J. C., Nikolskaia, O. V., Kim, Y. V., Brown, A., Pardo, C. A., Zhang, Y., Becker, K. G., Wilson, B. A., de A. Lima, A. P. C., Scharfstein, J. and Dumler, J. S. (2009) 'Protease Activated Receptor Signaling Is Required for African Trypanosome Traversal of Human Brain Microvascular Endothelial Cells', *PLoS Neglected Tropical Diseases*. Edited by J. Raper, 3(7), p. e479. doi: 10.1371/journal.pntd.0000479.

Grab, D. J. and Kennedy, P. G. (2008) 'Traversal of human and animal trypanosomes across the blood-brain barrier', *Journal of Neurovirology*, 14(5), pp. 344–351. doi: 10.1080/13550280802282934.

Grandgenett, P. M., Otsu, K., Wilson, H. R., Wilson, M. E. and Donelson, J. E. (2007) 'A function for a specific zinc metalloprotease of African trypanosomes', *PLoS Pathogens*, 3(10), pp. 1432–1445. doi: 10.1371/journal.ppat.0030150.

Grant, B. D. and Adams, J. A. (1996) 'Pre-steady-state kinetic analysis of cAMP-dependent protein kinase using rapid quench flow techniques', *Biochemistry*, 35(6), pp. 2022–2029. doi: 10.1021/bi952144+.

Greganova, E. and Bütikofer, P. (2012) 'Ethanolamine phosphoglycerol attachment to eEF1A is not essential for normal growth of *Trypanosoma brucei*.', *Scientific Reports*, 2, p. 254. doi: 10.1038/srep00254.

Gribble, F. M., Loussouarn, G., Tucker, S. J., Zhao, C., Nichols, C. G. and Ashcroft, F. M. (2000) 'A novel method for measurement of submembrane ATP concentration', *Journal of Biological Chemistry*, 275(39), pp. 30046–30049. doi: 10.1074/jbc.M001010200.

Griffioen, G., Anghileri, P., Imre, E., Baroni, M. D. and Ruis, H. (2000) 'Nutritional control of nucleocytoplasmic localization of cAMP-dependent protein kinase catalytic and regulatory subunits in *Saccharomyces cerevisiae*', *Journal of Biological Chemistry*, 275(2), pp. 1449–1456. doi: 10.1074/jbc.275.2.1449.

Griffioen, G., Branduardi, P., Ballarini, a, Anghileri, P., Norbeck, J., Baroni, M. D. and Ruis, H. (2001) 'Nucleocytoplasmic distribution of budding yeast protein kinase A regulatory subunit Bcy1 requires Zds1 and is regulated by Yak1-dependent phosphorylation of its targeting domain.', *Molecular and cellular biology*, 21(2), pp. 511–23. doi: 10.1128/MCB.21.2.511-523.2001.

Gruszynski, A. E., van Deursen, F. J., Albareda, M. C., Best, A., Chaudhary, K., Cliffe, L. J., del Rio, L., Dunn, J. D., Ellis, L., Evans, K. J., Figueiredo, J. M., Malmquist, N. a, Omosun, Y., Palenchar, J. B., Prickett, S., Punkosdy, G. a, van Dooren, G., Wang, Q., Menon, A. K., Matthews, K. R. and Bangs, J. D. (2006) 'Regulation of surface coat exchange by differentiating African trypanosomes.', *Molecular and biochemical parasitology*, 147(2), pp. 211–23. doi: 10.1016/j.molbiopara.2006.02.013.

Guimera, J., Casas, C., Estivill, X. and Pritchard, M. (1999) 'Human minibrain homologue (MNBH/DYRK1): characterization, alternative splicing, differential tissue expression, and overexpression in Down syndrome.', *Genomics*, 57(3), pp. 407–418. doi: 10.1006/geno.1999.5775.

Günzl, A., Bruderer, T., Laufer, G., Schimanski, B., Tu, L. C., Chung, H. M., Lee, P. T. and Lee, M. G. S. (2003) 'RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*', *Eukaryotic Cell*, 2(3), pp. 542–551. doi: 10.1128/EC.2.3.542-551.2003.

Günzl, A., Kirkham, J. K., Nguyen, T. N., Badjatia, N. and Park, S. H. (2015) 'Mono-allelic VSG expression by RNA polymerase I in *Trypanosoma brucei*: Expression site control from both ends?', *Gene*, 556(1), pp. 68–73. doi: 10.1016/j.gene.2014.09.047.

Haenni, S., Renggli, C. K., Fragoso, C. M., Oberle, M. and Roditi, I. (2006) 'The procyclin-associated genes of *Trypanosoma brucei* are not essential for cyclical

transmission by tsetse', *Molecular and Biochemical Parasitology*, 150(2), pp. 144–156. doi: 10.1016/j.molbiopara.2006.07.005.

Haenni, S., Studer, E., Burkard, G. S. and Roditi, I. (2009) 'Bidirectional silencing of RNA polymerase I transcription by a strand switch region in *Trypanosoma brucei*', *Nucleic Acids Research*, 37(15), pp. 5007–5018. doi: 10.1093/nar/gkp513.

Haines, L. R. (2013) 'Examining the tsetse teneral phenomenon and permissiveness to trypanosome infection.', *Frontiers in cellular and infection microbiology*, 3(November), p. 84. doi: 10.3389/fcimb.2013.00084.

Hales, E. C., Taub, J. W. and Matherly, L. H. (2014) 'New insights into Notch1 regulation of the PI3K-AKT-mTOR1 signaling axis: Targeted therapy of ©-secretase inhibitor resistant T-cell acute lymphoblastic leukemia', *Cellular Signalling*. Elsevier Inc., 26(1), pp. 149–161. doi: 10.1016/j.cellsig.2013.09.021.

Hall, B. S., Gabernet-Castello, C., Voak, A., Goulding, D., Natesan, S. K. and Field, M. C. (2006) 'TbVps34, the trypanosome orthologue of Vps34, is required for Golgi complex segregation', *Journal of Biological Chemistry*, 281(37), pp. 27600–27612. doi: 10.1074/jbc.M602183200.

Hall, J. P. J., Wang, H. and David Barry, J. (2013) 'Mosaic VSGs and the Scale of *Trypanosoma brucei* Antigenic Variation', *PLoS Pathogens*, 9(7). doi: 10.1371/journal.ppat.1003502.

Hammarton, T. C. (2007) 'Cell cycle regulation in *Trypanosoma brucei*', *Molecular and Biochemical Parasitology*. Elsevier B.V., 153(1), pp. 1–8. doi: 10.1016/j.molbiopara.2007.01.017.

Hammarton, T. C., Clark, J., Douglas, F., Boshart, M. and Mottram, J. C. (2003) 'Stage-specific differences in cell cycle control in *Trypanosoma brucei* revealed by RNA interference of a mitotic cyclin.', *The Journal of biological chemistry*, 278(25), pp. 22877–86. doi: 10.1074/jbc.M300813200.

Hammarton, T. C., Kramer, S., Tetley, L., Boshart, M. and Mottram, J. C. (2007)

'*Trypanosoma brucei* Polo-like kinase is essential for basal body duplication, kDNA segregation and cytokinesis', *Molecular Microbiology*, 65(5), pp. 1229–1248. doi: 10.1111/j.1365-2958.2007.05866.x.

Hammarton, T. C., Monnerat, S. and Mottram, J. C. (2007) 'Cytokinesis in trypanosomatids', *Current Opinion in Microbiology*, 10(6), pp. 520–527. doi: 10.1016/j.mib.2007.10.005.

Han, J., Luby-Phelps, K., Das, B., Shu, X., Xia, Y., Mosteller, R. D., Krishna, U. M., Falck, J. R., White, M. A. and Broek, D. (1998) 'Role of substrates and products of PI 3-kinase in regulating activation of Rac-related guanosine triphosphatases by Vav', *Science*, 279(5350), pp. 558–560. doi: 10.1126/science.279.5350.558.

Hanks, S. K. (2003) 'Genomic analysis of the eukaryotic protein kinase superfamily: a perspective.', *Genome biology*, 4(5), p. 111. doi: 10.1186/gb-2003-4-5-111.

Hanks, S. K. and Hunter, T. (1995) 'Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification.', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 9(8), pp. 576–96. doi: 7768349.

Hassan, P., Fergusson, D., Grant, K. M. and Mottram, J. C. (2001) 'The CRK3 protein kinase is essential for cell cycle progression of *Leishmania mexicana*.', *Molecular and biochemical parasitology*, 113(2), pp. 189–198.

Hathaway, G. M., Zoller, M. J. and Traugh, J. A. (1981) 'Identification of the catalytic subunit of casein kinase II by affinity labeling with 5'-p-fluorosulfonylbenzoyl adenosine.', *The Journal of biological chemistry*, 256(22), pp. 11442–6. PMID: 6946059.

Hengge-Aronis, R., Lange, R., Henneberg, N. and Fischer, D. (1993) 'Osmotic regulation of rpoS-dependent genes in *Escherichia coli*.', *J. Bacteriol.*, 175(1), pp. 259–265. PMID: 8416901

Henise, J. C. and Taunton, J. (2011) 'Irreversible Nek2 kinase inhibitors with cellular activity', *Journal of Medicinal Chemistry*, 54(12), pp. 4133–4146. doi: 10.1021/jm200222m.

Herwaldt, B. L. (2001) 'Laboratory-acquired parasitic infections from accidental exposures', *Clinical Microbiology Reviews*, 14(4), pp. 659–688. doi: 10.1128/CMR.14.3.659-688.2001.

Hixson, C. S. and Krebs, E. G. (1979) 'Affinity labeling of catalytic subunit of bovine heart muscle cyclic AMP-dependent protein kinase by 5'-p-fluorosulfonylbenzoyladenosine.', *The Journal of biological chemistry*, 254(16), pp. 7509–14. PMID: 224031.

Ho, H. H., He, C. Y., de Graffenried, C. L., Murrells, L. J. and Warren, G. (2006) 'Ordered assembly of the duplicating Golgi in Trypanosoma brucei.', *Proceedings of the National Academy of Sciences of the United States of America*, 103(20), pp. 7676–7681. doi: 10.1073/pnas.0602595103.

Holt, H. R., Selby, R., Mumba, C., Napier, G. B. and Guitian, J. (2016) 'Assessment of animal African trypanosomiasis (AAT) vulnerability in cattle-owning communities of sub-Saharan Africa.', *Parasites & vectors*. Parasites & Vectors, 9(1), p. 53. doi: 10.1186/s13071-016-1336-5.

Hopkins, A. L. and Groom, C. R. (2002) 'The druggable genome.', *Nature reviews. Drug discovery*, 1(9), pp. 727–30. doi: 10.1038/nrd892.

Horn, D. (2008) 'Codon usage suggests that translational selection has a major impact on protein expression in trypanosomatids.', *BMC genomics*, 9, p. 2. doi: 10.1186/1471-2164-9-2.

Horwich, M. D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P. and Zamore, P. D. (2007) 'The Drosophila RNA Methyltransferase, DmHen1, Modifies Germline piRNAs and Single-Stranded siRNAs in RISC', *Current Biology*, 17(14), pp. 1265–1272. doi: 10.1016/j.cub.2007.06.030.

Hsia, R., Beals, T. and Boothroyd, J. C. (1996) 'Use of chimeric recombinant

polypeptides to analyse conformational, surface epitopes on trypanosome variant surface glycoproteins.', *Molecular microbiology*, 19(1), pp. 53–63. doi: 10.1046/j.1365-2958.1996.351878.x.

Hua, S. B. and Wang, C. C. (1997) 'Interferon-gamma activation of a mitogen-activated protein kinase, KFR1, in the bloodstream form of *Trypanosoma brucei*.', *The Journal of biological chemistry*, 272(16), pp. 10797–10803. doi: 10.1074/jbc.272.16.10797.

Hughes, K., Wand, M., Foulston, L., Young, R., Harley, K., Terry, S., Ersfeld, K. and Rudenko, G. (2007) 'A novel ISWI is involved in VSG expression site downregulation in African trypanosomes.', *The EMBO Journal*, 26(9), pp. 2400–2410. doi: 10.1038/sj.emboj.7601678.

Hughes, L., Towers, K., Starborg, T., Gull, K. and Vaughan, S. (2013) 'A cell-body groove housing the new flagellum tip suggests an adaptation of cellular morphogenesis for parasitism in the bloodstream form of *Trypanosoma brucei*.', *Journal of cell science*, 126(Pt 24), pp. 5748–57. doi: 10.1242/jcs.139139.

Hunter, T. (1995) 'Protein kinases and phosphatases: The Yin and Yang of protein phosphorylation and signaling', *Cell*, 80(2), pp. 225–236. doi: 10.1016/0092-8674(95)90405-0.

Huson, D. H. and Bryant, D. (2006) 'Application of phylogenetic networks in evolutionary studies', *Molecular Biology and Evolution*, 23(2), pp. 254–267. doi: 10.1093/molbev/msj030.

Ibar, C., Cataldo, V. F., Vasquez-Doorman, C., Olguin, P. and Glavic, A. (2013) '*Drosophila* p53-related protein kinase is required for PI3K/TOR pathway-dependent growth', *Development*, 140(6), pp. 1282–1291. doi: 10.1242/dev.086918.

Imhof, S., Fragoso, C., Hemphill, A., von Schubert, C., Li, D., Legant, W., Betzig, E. and Roditi, I. (2016) 'Flagellar membrane fusion and protein exchange in trypanosomes; a new form of cell-cell communication?', *F1000Research*, 5, p. 682. doi: 10.12688/f1000research.8249.1.

Imhof, S., Kn??sel, S., Gunasekera, K., Vu, X. L. and Roditi, I. (2014) 'Social Motility of African Trypanosomes Is a Property of a Distinct Life-Cycle Stage That Occurs Early in Tsetse Fly Transmission', *PLoS Pathogens*, 10(10). doi: 10.1371/journal.ppat.1004493.

Imhof, S. and Roditi, I. (2015) 'The Social Life of African Trypanosomes', *Trends in Parasitology*, pp. 490–498. doi: 10.1016/j.pt.2015.06.012.

Inglese, J. and Premont, R. T. (1996) 'Lipid modifications of G-protein-coupled receptor kinases', *Biochemical Society Transactions*, 24(3), pp. 714–717. doi: 10.1042/bst0240714.

Innocenti, P., Cheung, K. M. J., Solanki, S., Mas-Droux, C., Rowan, F., Yeoh, S., Boxall, K., Westlake, M., Pickard, L., Hardy, T., Baxter, J. E., Aherne, G. W., Bayliss, R., Fry, A. M. and Hoelder, S. (2012) 'Design of potent and selective hybrid inhibitors of the mitotic kinase nek2: Structure-activity relationship, structural biology, and cellular activity', *Journal of Medicinal Chemistry*, 55(7), pp. 3228–3241. doi: 10.1021/jm201683b.

Inoue, M., Nakamura, Y., Yasuda, K., Yasaka, N., Hara, T., Schnaufer, A., Stuart, K. and Fukuma, T. (2005) 'The 14-3-3 proteins of *Trypanosoma brucei* function in motility, cytokinesis, and cell cycle', *Journal of Biological Chemistry*, 280(14), pp. 14085–14096. doi: 10.1074/jbc.M412336200.

Inoue, M., Okamoto, K., Uemura, H., Yasuda, K., Motohara, Y., Morita, K., Hiromura, M., Reddy, E. P., Fukuma, T. and Horikoshi, N. (2015) 'Identification and characterization of a cell division-regulating kinase AKB1 (associated kinase of *Trypanosoma brucei* 14-3-3) through proteomics study of the Tb14-3-3 binding proteins', *Journal of Biochemistry*, 158(1), pp. 49–60. doi: 10.1093/jb/mvv018.

Jacks, K. A. and Koch, C. A. (2010) 'Differential regulation of mitogen- and stress-activated protein kinase-1 and -2 (MSK1 and MSK2) by CK2 following UV radiation', *Journal of Biological Chemistry*, 285(3), pp. 1661–1670. doi: 10.1074/jbc.M109.083808.

Jackson, A. P. (2007a) 'Evolutionary consequences of a large duplication event in

Trypanosoma brucei: chromosomes 4 and 8 are partial duplicons.', *BMC genomics*, 8, p. 432. doi: 10.1186/1471-2164-8-432.

Jackson, A. P. (2007b) 'Tandem gene arrays in *Trypanosoma brucei*: comparative phylogenomic analysis of duplicate sequence variation.', *BMC evolutionary biology*, 7, p. 54. doi: 10.1186/1471-2148-7-54.

Jackson, A. P. (2014) 'Genome evolution in trypanosomatid parasites.', *Parasitology*, pp. 1–17. doi: 10.1017/S0031182014000894.

Jensen, B. C., Kifer, C. T., Brekken, D. L., Randall, A. C., Wang, Q., Drees, B. L. and Parsons, M. (2007) 'Characterization of protein kinase CK2 from *Trypanosoma brucei*', *Molecular and Biochemical Parasitology*, 151(1), pp. 28–40. doi: 10.1016/j.molbiopara.2006.10.002.

Jensen, B. C., Sivam, D., Kifer, C. T., Myler, P. J. and Parsons, M. (2009) 'Widespread variation in transcript abundance within and across developmental stages of *Trypanosoma brucei*.', *BMC genomics*, 10, p. 482. doi: 10.1186/1471-2164-10-482.

Jensen, R. E. and Englund, P. T. (2012) 'Network news: the replication of kinetoplast DNA', *Annual Review of Microbiology*, 66, pp. 473–491. doi: 10.1146/annurev-micro-092611-150057.

De Jesus, T. C. L., Tonelli, R. R., Nardelli, S. C., Da Silva Augusto, L., Motta, M. C. M., Girard-Dias, W., Miranda, K., Ulrich, P., Jimenez, V., Barquilla, A., Navarro, M., Docampo, R. and Schenkman, S. (2010) 'Target of rapamycin (TOR)-like 1 kinase is involved in the control of polyphosphate levels and acidocalcisome maintenance in *Trypanosoma brucei*', *Journal of Biological Chemistry*, 285(31), pp. 24131–24140. doi: 10.1074/jbc.M110.120212.

Jetton, N., Rothberg, K. G., Hubbard, J. G., Wise, J., Li, Y., Ball, H. L. and Ruben, L. (2009) 'The cell cycle as a therapeutic target against *Trypanosoma brucei*: Hesperadin inhibits Aurora kinase-1 and blocks mitotic progression in bloodstream forms', *Molecular Microbiology*, 72(2), pp. 442–458. doi: 10.1111/j.1365-2958.2009.06657.x.

Jin, T., Zhang, N., Long, Y., Parent, C. A. and Devreotes, P. N. (2000) 'Localization of the G protein betagamma complex in living cells during chemotaxis.', *Science (New York, N.Y.)*, 287(5455), pp. 1034–6. doi: 10.1126/science.287.5455.1034.

Johnson, J. E. and Cornell, R. B. (1999) 'Amphitropic proteins: regulation by reversible membrane interactions (review).', *Molecular membrane biology*, 16(3), pp. 217–35. PMID: 10503244.

Jones, D. C., Foth, B. J., Urbaniak, M. D., Patterson, S., Ong, H. B., Berriman, M. and Fairlamb, A. H. (2015) 'Genomic and Proteomic Studies on the Mode of Action of Oxaboroles against the African Trypanosome.', *PLoS neglected tropical diseases*, 9(12), p. e0004299. doi: 10.1371/journal.pntd.0004299.

Jones, N. G. (2012) *Validating Protein Kinases of Trypanosoma brucei as Drug Targets*. PhD thesis. University of Glasgow.

Jones, N. G., Thomas, E. B., Brown, E., Dickens, N. J., Hammarton, T. C. and Mottram, J. C. (2014) 'Regulators of Trypanosoma brucei cell cycle progression and differentiation identified using a kinome-wide RNAi screen.', *PLoS pathogens*, 10(1), p. e1003886. doi: 10.1371/journal.ppat.1003886.

Jordan, A. (1977) *Systematics, Tsetse: the future for biological methods in integrated control*. Edited by M. Laird. Ottawa, Canada: International Development Research Centre.

Joshi, P. P., Shegokar, V. R., Powar, R. M., Herder, S., Katti, R., Salkar, H. R., Dani, V. S., Bhargava, A., Jannin, J. and Truc, P. (2005) 'Human trypanosomiasis caused by *Trypanosoma evansi* in India: the first case report', *American Journal of Tropical Medicine & Hygiene*, 73(OCTOBER), pp. 491–495.

Kabani, S., Fenn, K., Ross, A., Ivens, A., Smith, T. K., Ghazal, P. and 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, p. 427. doi: 10.1186/1471-2164-10-427.

Kabayo, J. P. (2002) ‘Aiming to eliminate tsetse from Africa’, *Trends in Parasitology*, 18(11), pp. 473–475. doi: 10.1016/S1471-4922(02)02371-1.

Kainz, P. (2000) ‘The PCR plateau phase - towards an understanding of its limitations.’, *Biochimica et biophysica acta*, 1494(1–2), pp. 23–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11072065.

Kamps, M. P. and Sefton, B. M. (1986) ‘Neither arginine nor histidine can carry out the function of lysine-295 in the ATP-binding site of p60src.’, *Molecular and cellular biology*, 6(3), pp. 751–757. doi: 10.1128/MCB.6.3.751.Updated.

Karaman, M. W., Herrgard, S., Treiber, D. K., Gallant, P., Atteridge, C. E., Campbell, B. T., Chan, K. W., Ciceri, P., Davis, M. I., Edeen, P. T., Faraoni, R., Floyd, M., Hunt, J. P., Lockhart, D. J., Milanov, Z. V, Morrison, M. J., Pallares, G., Patel, H. K., Pritchard, S., Wodicka, L. M. and Zarrinkar, P. P. (2008) ‘A quantitative analysis of kinase inhibitor selectivity.’, *Nature biotechnology*, 26(1), pp. 127–132. doi: 10.1038/nbt1358.

Katiyar, S., Kufareva, I., Behera, R., Thomas, S. M., Ogata, Y., Pollastri, M., Abagyan, R. and Mensa-Wilmot, K. (2013) ‘Lapatinib-Binding Protein Kinases in the African Trypanosome: Identification of Cellular Targets for Kinase-Directed Chemical Scaffolds’, *PLoS ONE*, 8(2). doi: 10.1371/journal.pone.0056150.

Katzen, F. (2007) ‘Recombinational Cloning: a Biological Operating System’, *Expert Opinion on Drug Discovery*, 2(4), pp. 571–589. doi: 10.1517/17460441.2.4.571.

Kennedy, P. G. (2013) ‘Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness)’, *The Lancet Neurology*, 12(2), pp. 186–194. doi: 10.1016/S1474-4422(12)70296-X.

Kentrup, H., Becker, W., Heukelbach, J., Wilmes, A., Schürmann, A., Huppertz, C., Kainulainen, H. and Joost, H. G. (1996) ‘Dyrk, a dual specificity protein kinase with unique structural features whose activity is dependent on tyrosine residues between subdomains VII and VIII’, *Journal of Biological Chemistry*, 271(7), pp. 3488–3495. doi: 10.1074/jbc.271.7.3488.

Khare, S., Nagle, A. S., Biggart, A., Lai, Y. H., Liang, F., Davis, L. C., Barnes, S. W., Mathison, C. J. N., Myburgh, E., Gao, M.-Y., Gillespie, J. R., Liu, X., Tan, J. L., Stinson, M., Rivera, I. C., Ballard, J., Yeh, V., Groessl, T., Federe, G., Koh, H. X. Y., Venable, J. D., Bursulaya, B., Shapiro, M., Mishra, P. K., Spraggon, G., Brock, A., Mottram, J. C., Buckner, F. S., Rao, S. P. S., Wen, B. G., Walker, J. R., Tuntland, T., Molteni, V., Glynne, R. J. and Supek, F. (2016) 'Proteasome inhibition for treatment of leishmaniasis, Chagas disease and sleeping sickness', *Nature*. Nature Publishing Group, pp. 1–24. doi: 10.1038/nature19339.

Knight, Z. A. and Shokat, K. M. (2007) 'Chemical Genetics: Where Genetics and Pharmacology Meet', *Cell*, 128(3), pp. 425–430. doi: 10.1016/j.cell.2007.01.021.

Knockaert, M., Gray, N., Damiens, E., Chang, Y. T., Grellier, P., Grant, K., Fergusson, D., Mottram, J., Soete, M., Dubremetz, J. F., Le Roch, K., Doerig, C., Schultz, P. G. and Meijer, L. (2000) 'Intracellular targets of cyclin-dependent kinase inhibitors: Identification by affinity chromatography using immobilised inhibitors', *Chemistry and Biology*, 7(6), pp. 411–422. doi: 10.1016/S1074-5521(00)00124-1.

Kolev, N. G., Franklin, J. B., Carmi, S., Shi, H., Michaeli, S. and Tschudi, C. (2010) 'The transcriptome of the human pathogen *Trypanosoma brucei* at single-nucleotide resolution', *PLoS Pathogens*, 6(9), pp. 1–15. doi: 10.1371/journal.ppat.1001090.

Kolev, N. G., Ramey-Butler, K., Cross, G. a M., Ullu, E. and Tschudi, C. (2012) 'Developmental progression to infectivity in *Trypanosoma brucei* triggered by an RNA-binding protein.', *Science (New York, N.Y.)*, 338(6112), pp. 1352–3. doi: 10.1126/science.1229641.

Kolev, N. G., Tschudi, C. and Ullu, E. (2011) 'RNA interference in protozoan parasites: Achievements and challenges', *Eukaryotic Cell*, 10(9), pp. 1156–1163. doi: 10.1128/EC.05114-11.

Kolev, N. G., Ullu, E. and Tschudi, C. (2014) 'The emerging role of RNA-binding proteins in the life cycle of *Trypanosoma brucei*', *Cellular Microbiology*, pp. 482–489. doi: 10.1111/cmi.12268.

Kristensson, K., Nygård, M., Bertini, G. and Bentivoglio, M. (2010) 'African trypanosome infections of the nervous system: Parasite entry and effects on sleep and synaptic functions', *Progress in Neurobiology*, 91(2), pp. 152–171. doi: 10.1016/j.pneurobio.2009.12.001.

Kristjanson, P. M., Swallow, B. M., Rowlands, G. J., Kruska, R. L. and De Leeuw, P. N. (1999) 'Measuring the costs of African animal trypanosomosis, the potential benefits of control and returns to research', *Agricultural Systems*, 59(1), pp. 79–98. doi: 10.1016/S0308-521X(98)00086-9.

Kroiher, M., Miller, M. A. and Steele, R. E. (2001) 'Deceiving appearances: signaling by "dead" and "fractured" receptor protein-tyrosine kinases.', *BioEssays*, 23(1), pp. 69–76. doi: 10.1002/1521-1878(200101)23:1<69::AID-BIES1009>3.0.CO;2-K.

van der Krol, A. R., Mur, L. A., Beld, M., Mol, J. N. and Stuitje, A. R. (1990) 'Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression.', *The Plant cell*, 2(4), pp. 291–9. doi: 2152117.

Krupa, A., Preethi, G. and Srinivasan, N. (2004) 'Structural modes of stabilization of permissive phosphorylation sites in protein kinases: Distinct strategies in Ser/Thr and Tyr kinases', *Journal of Molecular Biology*, 339(5), pp. 1025–1039. doi: 10.1016/j.jmb.2004.04.043.

Kubi, C., Van Den Abbeele, J., De Deken, R., Marcotty, T., Dorny, P. and Van Den Bossche, P. (2006) 'The effect of starvation on the susceptibility of teneral and non-teneral tsetse flies to trypanosome infection', *Medical and Veterinary Entomology*, 20(4), pp. 388–392. doi: 10.1111/j.1365-2915.2006.00644.x.

Kumar, P. and Wang, C. C. (2006) 'Dissociation of cytokinesis initiation from mitotic control in a eukaryote', *Eukaryotic Cell*, 5(1), pp. 92–102. doi: 10.1128/EC.5.1.92-102.2006.

Lander, N., Li, Z.-H., Niyogi, S. and Docampo, R. (2015) 'CRISPR/Cas9-Induced Disruption of Paraflagellar Rod Protein 1 and 2 Genes in *Trypanosoma cruzi*

Reveals Their Role in Flagellar Attachment', *mBio*, 6(4), pp. e01012-15. doi: 10.1128/mBio.01012-15.

Langmead, B. and Salzberg, S. L. (2012) 'Fast gapped-read alignment with Bowtie 2.', *Nature methods*, 9(4), pp. 357–9. doi: 10.1038/nmeth.1923.

Langousis, G. and Hill, K. L. (2014) 'Motility and more: the flagellum of *Trypanosoma brucei*.', *Nature reviews. Microbiology*, 12(7), pp. 505–18. doi: 10.1038/nrmicro3274.

Lanham, S. and Godfrey, D. (1970) 'Isolation and of Salivarian Other Mammals Trypanosomes from using DEAE-Cellulose', *Experimental pathology*, (28), pp. 521–534.

Larkin, M. a., Blackshields, G., Brown, N. P., Chenna, R., Mcgettigan, P. a., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, a., Lopez, R., Thompson, J. D., Gibson, T. J. and Higgins, D. G. (2007) 'Clustal W and Clustal X version 2.0', *Bioinformatics*, 23(21), pp. 2947–2948. doi: 10.1093/bioinformatics/btm404.

LaVallie, E. R., DiBlasio, E. a, Kovacic, S., Grant, K. L., Schendel, P. F. and McCoy, J. M. (1993) 'A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm.', *Bio/technology (Nature Publishing Company)*, 11(2), pp. 187–193. doi: 10.1038/nbt0293-187.

Laxman, S., Riechers, A., Sadilek, M., Schwede, F. and Beavo, J. a (2006) 'Hydrolysis products of cAMP analogs cause transformation of *Trypanosoma brucei* from slender to stumpy-like forms.', *Proceedings of the National Academy of Sciences of the United States of America*, 103(50), pp. 19194–9. doi: 10.1073/pnas.0608971103.

Lee, J., Han, C. T. and Hur, Y. (2010) 'Overexpression of BrMORN, a novel "membrane occupation and recognition nexus" motif protein gene from Chinese cabbage, promotes vegetative growth and seed production in *Arabidopsis*', *Molecules and Cells*, 29(2), pp. 113–122. doi: 10.1007/s10059-010-0006-2.

Lee, M. G. and Van der Ploeg, L. H. (1990) 'Homologous recombination and

stable transfection in the parasitic protozoan *Trypanosoma brucei*.', *Science (New York, N.Y.)*, 250(4987), pp. 1583–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2177225.

Lemmon, M. A., Ferguson, K. M. and Abrams, C. S. (2002) 'Pleckstrin homology domains and the cytoskeleton', *FEBS Letters*, 513(1), pp. 71–76. doi: 10.1016/S0014-5793(01)03243-4.

Letunic, I., Doerks, T. and Bork, P. (2015) 'SMART: Recent updates, new developments and status in 2015', *Nucleic Acids Research*, 43(D1), pp. D257–D260. doi: 10.1093/nar/gku949.

Li, B. (2015) 'DNA Double-Strand Breaks and Telomeres Play Important Roles in *Trypanosoma brucei* Antigenic Variation', *Eukaryot Cell*, 14, pp. 196–205. doi: 10.1128/EC.00207-14.

Li, F. J. and He, C. Y. (2014) 'Acidocalcisome is required for autophagy in *Trypanosoma brucei*', *Autophagy*, 10(11), pp. 1978–1988. doi: 10.4161/auto.36183.

Li, Z. (2012) 'Regulation of the cell division cycle in *Trypanosoma brucei*', *Eukaryotic Cell*, 11(10), pp. 1180–1190. doi: 10.1128/EC.00145-12.

Li, Z., Gourguechon, S. and Wang, C. C. (2007) 'Tousled-like kinase in a microbial eukaryote regulates spindle assembly and S-phase progression by interacting with Aurora kinase and chromatin assembly factors.', *Journal of cell science*, 120(Pt 21), pp. 3883–3894. doi: 10.1242/jcs.007955.

Li, Z., Umeyama, T. and Wang, C. C. (2009) 'The Aurora kinase in *Trypanosoma brucei* plays distinctive roles in metaphase-anaphase transition and cytokinetic initiation', *PLoS Pathogens*, 5(9). doi: 10.1371/journal.ppat.1000575.

Li, Z. and Wang, C. C. (2006) 'Changing roles of aurora-B kinase in two life cycle stages of *Trypanosoma brucei*', *Eukaryotic Cell*, 5(7), pp. 1026–1035. doi: 10.1128/EC.00129-06.

Lindner, A. K. and Priotto, G. (2010) 'The unknown risk of vertical transmission in sleeping sickness-a literature review.', *PLoS neglected tropical diseases*, 4(12), pp. 1–5. doi: 10.1371/journal.pntd.0000783.

Liscovitch, M. and Cantley, L. C. (1994) 'Lipid second messengers.', *Cell*, 77(3), pp. 329–34. PMID: 8181054.

Liu, B., Liu, Y., Motyka, S. A., Agbo, E. E. C. and Englund, P. T. (2005) 'Fellowship of the rings: The replication of kinetoplast DNA', *Trends in Parasitology*, pp. 363–369. doi: 10.1016/j.pt.2005.06.008.

Liu, Y. and Gray, N. S. (2006) 'Rational design of inhibitors that bind to inactive kinase conformations.', *Nature chemical biology*, 2(7), pp. 358–364. doi: 10.1038/nchembio799.

Liu, Y., Solis, N. V., Heilmann, C. J., Phan, Q. T., Mitchell, A. P., Klis, F. M. and Filler, S. G. (2014) 'Role of retrograde trafficking in stress response, host cell interactions, and virulence of *Candida albicans*', *Eukaryotic Cell*, 13(2), pp. 279–287. doi: 10.1128/EC.00295-13.

Lochhead, P. a., Sibbet, G., Morrice, N. and Cleghon, V. (2005) 'Activation-loop autophosphorylation is mediated by a novel transitional intermediate form of DYRKs', *Cell*, 121(6), pp. 925–936. doi: 10.1016/j.cell.2005.03.034.

Loh, J. T., Torres, V. J. and Cover, T. L. (2007) 'Regulation of *Helicobacter pylori* cagA expression in response to salt', *Cancer Research*, 67(10), pp. 4709–4715. doi: 10.1158/0008-5472.CAN-06-4746.

London Declaration on Neglected Tropical Diseases (2012) *World Health Organization*. doi: 10.1080/03064228708534241.

Lonsdale-Eccles, J. D. and Grab, D. J. (2002) 'Trypanosome hydrolases and the blood-brain barrier', *Trends in Parasitology*, 18(1), pp. 17–19. doi: 10.1016/S1471-4922(01)02120-1.

Lorusso, P. M. and Eder, J. P. (2008) 'Therapeutic potential of novel selective-

spectrum kinase inhibitors in oncology', *Expert opinion on investigational drugs*, 17(7), pp. 1013–28. doi: 10.1517/13543784.17.7.1013.

Lozano-Núñez, A., Ikeda, K. N., Sauer, T. and de Graffenried, C. L. (2013) 'An analogue-sensitive approach identifies basal body rotation and flagellum attachment zone elongation as key functions of PLK in *Trypanosoma brucei*.', *Molecular biology of the cell*, 24(9), pp. 1321–33. doi: 10.1091/mbc.E12-12-0846.

Lucas, K. A., Pitari, G. M., Kazerounian, S., Ruiz-Stewart, I., Park, J., Schulz, S., Chepenik, K. P. and Waldman, S. A. (2000) 'Guanylyl cyclases and signaling by cyclic GMP.', *Pharmacological reviews*, 52(3), pp. 375–414. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10977868.

Lukeš, J., Skalický, T., Týč, J., Votýpka, J. and Yurchenko, V. (2014) 'Evolution of parasitism in kinetoplastid flagellates.', *Molecular and biochemical parasitology*, 195(2), pp. 115–22. doi: 10.1016/j.molbiopara.2014.05.007.

Luo, K., Zhou, P. and Lodish, H. F. (1995) 'The specificity of the transforming growth factor beta receptor kinases determined by a spatially addressable peptide library.', *Proceedings of the National Academy of Sciences of the United States of America*, 92(25), pp. 11761–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8524844.

Lye, L.-F., Owens, K., Shi, H., Murta, S. M. F., Vieira, A. C., Turco, S. J., Tschudi, C., Ullu, E. and Beverley, S. M. (2010) 'Retention and Loss of RNA Interference Pathways in Trypanosomatid Protozoans', *PLoS Pathogens*. Edited by B. Ullman, 6(10), p. e1001161. doi: 10.1371/journal.ppat.1001161.

Ma, A. D. and Abrams, C. S. (1999) 'Pleckstrin homology domains and phospholipid-induced cytoskeletal reorganization.', *Thrombosis and haemostasis*, 82(2), pp. 399–406. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11911883.

Ma, J., Benz, C., Grimaldi, R., Stockdale, C., Wyatt, P., Frearson, J. and Hammarton, T. C. (2010) 'Nuclear DBF-2-related kinases are essential regulators

of cytokinesis in bloodstream stage *Trypanosoma brucei*.', *The Journal of biological chemistry*, 285(20), pp. 15356–68. doi: 10.1074/jbc.M109.074591.

Macek, B., Mijakovic, I., Olsen, J. V, Gnad, F., Kumar, C., Jensen, P. R. and Mann, M. (2007) 'The serine/threonine/tyrosine phosphoproteome of the model bacterium Bacillus subtilis.', *Molecular & cellular proteomics : MCP*, 6(4), pp. 697–707. doi: 10.1074/mcp.M600464-MCP200.

MacGregor, P., Szöőr, B., Savill, N. J. and Matthews, K. R. (2012) 'Trypanosomal immune evasion, chronicity and transmission: an elegant balancing act.', *Nature reviews. Microbiology*, 10(6), pp. 431–8. doi: 10.1038/nrmicro2779.

Mackey, Z. B., Koupparis, K., Nishino, M. and McKerrow, J. H. (2011) 'High-throughput analysis of an RNAi library identifies novel kinase targets in *Trypanosoma brucei*.', *Chemical biology & drug design*, 78(3), pp. 454–63. doi: 10.1111/j.1747-0285.2011.01156.x.

Macleod, E. T., Darby, A. C., Maudlin, I. and Welburn, S. C. (2007) 'Factors affecting trypanosome maturation in tsetse flies', *PLoS ONE*, 2(2). doi: 10.1371/journal.pone.0000239.

Madeira da Silva, L. and Beverley, S. M. (2010) 'Expansion of the target of rapamycin (TOR) kinase family and function in *Leishmania* shows that TOR3 is required for acidocalcisome biogenesis and animal infectivity.', *Proceedings of the National Academy of Sciences of the United States of America*, 107(26), pp. 11965–11970. doi: 10.1073/pnas.1004599107.

Magez, S., Schwegmann, A., Atkinson, R., Claes, F., Drennan, M., De Baetselier, P. and Brombacher, F. (2008) 'The role of B-cells and IgM antibodies in parasitemia, anemia, and VSG switching in *Trypanosoma brucei*-infected mice.', *PLoS pathogens*, 4(8), p. e1000122. doi: 10.1371/journal.ppat.1000122.

Mallick, B., Ghosh, Z. and Chakrabarti, J. (2008) 'MicroRNA switches in *Trypanosoma brucei*.', *Biochemical and biophysical research communications*, 372(3), pp. 459–63. doi: 10.1016/j.bbrc.2008.05.084.

Mancini, P. E. and Patton, C. L. (1981) ‘Cyclic 3’,5’-adenosine monophosphate levels during the developmental cycle of *Trypanosoma brucei brucei* in the rat’, *Molecular and Biochemical Parasitology*, 3, pp. 19–31.

Manning, G., Plowman, G. D., Hunter, T. and Sudarsanam, S. (2002) ‘Evolution of protein kinase signaling from yeast to man.’, *Trends in biochemical sciences*, 27(10), pp. 514–20. PMID: 12368087.

Manning, G., Whyte, D. B., Martinez, R., Hunter, T. and Sudarsanam, S. (2002) ‘The Protein Kinase Complement of the Human Genome’, *Science*, 298(5600), pp. 1912–1934. doi: 10.1126/science.1075762.

Marcello, L. and Barry, J. D. (2007) ‘Analysis of the VSG gene silent archive in Trypanosoma brucei reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure.’, *Genome research*, 17(9), pp. 1344–52. doi: 10.1101/gr.6421207.

Masocha, W., Robertson, B., Rottenberg, M. E., Mhlanga, J., Sorokin, L. and Kristensson, K. (2004) ‘Cerebral vessel laminins and IFN-γ define *Trypanosoma brucei brucei* penetration of the blood-brain barrier’, *Journal of Clinical Investigation*, 114(5), pp. 689–694. doi: 10.1172/JCI200422104.

Matthews, K. R. (2005) ‘The developmental cell biology of *Trypanosoma brucei*’, *Journal of Cell Science*, 118(2), pp. 283–290. doi: 10.1242/jcs.01649.

Matthews, K. R. and Gull, K. (1994) ‘Evidence for an interplay between cell cycle progression and the initiation of differentiation between life cycle forms of African trypanosomes’, *Journal of Cell Biology*, 125(5), pp. 1147–1156. doi: 10.1083/jcb.125.5.1147.

Matthews, K. R., McCulloch, R. and Morrison, L. J. (2015) ‘The within-host dynamics of African trypanosome infections’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1675), p. 20140288. doi: 10.1098/rstb.2014.0288.

Mattick, K. L., Jørgensen, F., Legan, J. D., Cole, M. B., Porter, J., Lappin-Scott,

H. M. and Humphrey, T. J. (2000) ‘Survival and filamentation of *Salmonella enterica* serovar Enteritidis PT4 and *Salmonella enterica* serovar Typhimurium DT104 at low water activity’, 66(4), pp. 1274–1279. doi: 10.1128/AEM.66.4.1274-1279.2000.

Maudlin, I. and Welburn, S. C. (1989) ‘A single trypanosome is sufficient to infect a tsetse fly.’, *Annals of tropical medicine and parasitology*. Maney, 83(4), pp. 431–3. PMID: 6626009.

Mbang-Benet, D.-E., Sterkers, Y., Crobu, L., Sarrazin, A., Bastien, P. and Pagès, M. (2015) ‘RNA interference screen reveals a high proportion of mitochondrial proteins essential for correct cell cycle progress in *Trypanosoma brucei*’, *BMC Genomics*, 16(1), pp. 1–10. doi: 10.1186/s12864-015-1505-5.

McCall, L. I. and Matlashewski, G. (2010) ‘Localization and induction of the A2 virulence factor in *Leishmania*: Evidence that A2 is a stress response protein’, *Molecular Microbiology*, 77(2), pp. 518–530. doi: 10.1111/j.1365-2958.2010.07229.x.

McCulloch, R. and Field, M. C. (2015) ‘Quantitative sequencing confirms VSG diversity as central to immune evasion by *Trypanosoma brucei*’, *Trends in Parasitology*. Elsevier Ltd, 31(8), pp. 346–349. doi: 10.1016/j.pt.2015.05.001.

Meggio, F. and Pinna, L. a (2003) ‘One-thousand-and-one substrates of protein kinase CK2?’, *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 17(3), pp. 349–368. doi: 10.1096/fj.02-0473rev.

Meirelles, G. V., Perez, A. M., de Souza, E. E., Basei, F. L., Papa, P. F., Melo Hanchuk, T. D., Cardoso, V. B. and Kobarg, J. (2014) ‘“Stop Ne(c)king around”: How interactomics contributes to functionally characterize Nek family kinases.’, *World journal of biological chemistry*, 5(2), pp. 141–60. doi: 10.4331/wjbc.v5.i2.141.

Merritt, C., Silva, L. E., Tanner, A. L., Stuart, K. and Pollastri, M. P. (2014) ‘Kinases as druggable targets in trypanosomatid protozoan parasites’, *Chemical*

Reviews, 114(22), pp. 11280–11304. doi: 10.1021/cr500197d.

Merritt, C. and Stuart, K. (2013) 'Identification of essential and non-essential protein kinases by a fusion PCR method for efficient production of transgenic *Trypanosoma brucei*.', *Molecular and biochemical parasitology*. Elsevier B.V., 190(1), pp. 44–9. doi: 10.1016/j.molbiopara.2013.05.002.

Mhlanga, J. D. ., Bentivoglio, M. and Kristensson, K. (1997) 'Neurobiology of Cerebral Malaria and African Sleeping Sickness', *Brain Research Bulletin*, 44(5), pp. 579–589. doi: 10.1016/S0361-9230(97)00309-2.

Michaeli, S. (2011) 'Trans-splicing in trypanosomes: machinery and its impact on the parasite transcriptome.', *Future microbiology*, 6(4), pp. 459–474. doi: 10.2217/fmb.11.20.

Milne, K. G., Güther, M. L. S. and Ferguson, M. A. J. (2001) 'Acyl-CoA binding protein is essential in bloodstream form *Trypanosoma brucei*', *Molecular and Biochemical Parasitology*, 112(2), pp. 301–304. doi: 10.1016/S0166-6851(00)00369-8.

Miranda-Saavedra, D., Stark, M. J. R., Packer, J. C., Vivares, C. P., Doerig, C. and Barton, G. J. (2007) 'The complement of protein kinases of the microsporidium *Encephalitozoon cuniculi* in relation to those of *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*.', *BMC genomics*, 8, p. 309. doi: 10.1186/1471-2164-8-309.

Mitcheson, D. F., Bottrill, A. R., Carr, K., Coxon, C. R., Cano, C., Golding, B. T., Griffin, R. J., Fry, A. M., Doerig, C., Bayliss, R. and Tobin, A. B. (2016) 'A new tool for the chemical genetic investigation of the *Plasmodium falciparum* Pfnek-2 NIMA-related kinase', *Malaria Journal*. BioMed Central, 15(1), p. 535. doi: 10.1186/s12936-016-1580-3.

Mogk, S., Boßelmann, C. M., Mudogo, C. N., Stein, J., Wolburg, H. and Duszenko, M. (2016) 'African trypanosomes and brain infection - the unsolved question.', *Biological reviews of the Cambridge Philosophical Society*. doi: 10.1111/brv.12301.

Moniz, L., Dutt, P., Haider, N. and Stambolic, V. (2011) 'Nek family of kinases in cell cycle, checkpoint control and cancer.', *Cell division*, 6(Figure 1), p. 18. doi: 10.1186/1747-1028-6-18.

Mony, B. M., Macgregor, P., Ivens, A., Rojas, F., Cowton, A., Young, J., Horn, D. and Matthews, K. (2013) 'Genome-wide dissection of the quorum sensing signalling pathway in *Trypanosoma brucei*.', *Nature*. Nature Publishing Group. doi: 10.1038/nature12864.

Mony, B. M. and Matthews, K. R. (2015) 'Assembling the components of the quorum sensing pathway in African trypanosomes', *Molecular Microbiology*, 96(2), pp. 220–232. doi: 10.1111/mmi.12949.

Moore, A. and Richer, M. (2001) 'Re-emergence of epidemic sleeping sickness in southern Sudan', *Tropical Medicine and International Health*, 6(5), pp. 342–347. doi: 10.1046/j.1365-3156.2001.00714.x.

Moore, S., Shrestha, S., Tomlinson, K. W. and Vuong, H. (2012) 'Predicting the effect of climate change on African trypanosomiasis: integrating epidemiology with parasite and vector biology.', *Journal of the Royal Society, Interface / the Royal Society*, 9(70), pp. 817–30. doi: 10.1098/rsif.2011.0654.

Morand, S., Renggli, C. K., Roditi, I. and Vassella, E. (2012) 'MAP kinase kinase 1 (MKK1) is essential for transmission of *Trypanosoma brucei* by *Glossina morsitans*.', *Molecular and biochemical parasitology*. Elsevier B.V., 186(1), pp. 73–6. doi: 10.1016/j.molbiopara.2012.09.001.

Moreira, D., López-García, P. and Vickerman, K. (2004) 'An updated view of kinetoplastid phylogeny using environmental sequences and a closer outgroup: Proposal for a new classification of the class Kinetoplastea', *International Journal of Systematic and Evolutionary Microbiology*, 54(5), pp. 1861–1875. doi: 10.1099/ijs.0.63081-0.

Morris, J. C., Wang, Z., Drew, M. E. and Englund, P. T. (2002) 'Glycolysis modulates trypanosome glycoprotein expression as revealed by an RNAi library.', *The EMBO journal*, 21(17), pp. 4429–38. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=125414&tool=pmcentrez&rendertype=abstract.

Morrison, D. K. (2001) 'KSR: a MAPK scaffold of the Ras pathway?', *Journal of cell science*, 114(Pt 9), pp. 1609–12. PMID: 11309192.

Morrison, L. J., McCulloch, R. and Hall, J. P. J. (2014) 'DNA Recombination Strategies During Antigenic Variation in the African Trypanosome', *Mobile DNA III*, (1), pp. 409–435. doi: 10.1128/microbiolspec.MDNA3-0016-2014.

Mottram, J. C. and Smith, G. (1995) 'A family of trypanosome cdc2-related protein kinases', *Gene*, 162(1), pp. 147–152. doi: 10.1016/0378-1119(95)00350-F.

Mugnier, M. R., Cross, G. A. M. and Papavasiliou, F. N. (2015) 'The in vivo dynamics of antigenic variation in *Trypanosoma brucei*.', *Science (New York, N.Y.)*, 347(6229), pp. 1470–1473. doi: 10.1126/science.aaa4502.

Mugnier, M. R., Stebbins, C. E. and Papavasiliou, F. N. (2016) 'Masters of Disguise: Antigenic Variation and the VSG Coat in *Trypanosoma brucei*', *PLOS Pathogens*. Edited by L. J. Knoll, 12(9), p. e1005784. doi: 10.1371/journal.ppat.1005784.

Mutomba, M. C. and Wang, C. C. (1996) 'Effects of aphidicolin and hydroxyurea on the cell cycle and differentiation of *Trypanosoma brucei* bloodstream forms', *Molecular and Biochemical Parasitology*, 80(1), pp. 89–102. doi: 10.1016/0166-6851(96)02675-8.

Napoli, C. (1990) 'Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans', *THE PLANT CELL ONLINE*, 2(4), pp. 279–289. doi: 10.1105/tpc.2.4.279.

Naula, C., Parsons, M. and Mottram, J. C. (2005) 'Protein kinases as drug targets in trypanosomes and *Leishmania*', *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1754(1–2), pp. 151–159. doi: 10.1016/j.bbapap.2005.08.018.

Nett, I. R. E., Martin, D. M. a, Miranda-Saavedra, D., Lamont, D., Barber, J. D.,

Mehlert, A. and Ferguson, M. a J. (2009) 'The phosphoproteome of bloodstream form *Trypanosoma brucei*, causative agent of African sleeping sickness.', *Molecular & cellular proteomics : MCP*, 8(7), pp. 1527–38. doi: 10.1074/mcp.M800556-MCP200.

Newton, A. C. (1997) 'Regulation of protein kinase C.', *Current opinion in cell biology*, 9(2), pp. 161–7. PMID: 9069266.

Ngô, H., Tschudi, C., Gull, K. and Ullu, E. (1998) 'Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*.', *Proceedings of the National Academy of Sciences of the United States of America*, 95, pp. 14687–92. PMID: 9843950.

Van Nieuwenhove, S., Betu-Ku-Mesu, V. K., Diabakana, P. M., Declercq, J. and Bilenge, C. M. M. (2001) 'Sleeping sickness resurgence in the DRC: The past decade', *Tropical Medicine and International Health*, 6(5), pp. 335–341. doi: 10.1046/j.1365-3156.2001.00731.x.

Nishino, M., Choy, J. W., Gushwa, N. N., Oses-Prieto, J. a., Koupparis, K., Burlingame, A. L., Renslo, A. R., McKerrow, J. H. and Taunton, J. (2013) 'Hypothemicin, a fungal natural product, identifies therapeutic targets in *Trypanosoma brucei*', *eLife*, 2013(2), pp. 1–15. doi: 10.7554/eLife.00712.

Notredame, C., Higgins, D. G. and Heringa, J. (2000) 'T-coffee: a novel method for fast and accurate multiple sequence alignment', *Journal of Molecular Biology*, 302(1), pp. 205–217. doi: 10.1006/jmbi.2000.4042.

Obado, S. O., Bot, C., Nilsson, D., Andersson, B. and Kelly, J. M. (2007) 'Repetitive DNA is associated with centromeric domains in *Trypanosoma brucei* but not *Trypanosoma cruzi*.', *Genome biology*, 8(3), p. R37. doi: 10.1186/gb-2007-8-3-r37.

Oberholzer, M., Langousis, G., Nguyen, H. T., Saada, E. a, Shimogawa, M. M., Jonsson, Z. O., Nguyen, S. M., Wohlschlegel, J. a and Hill, K. L. (2011) 'Independent analysis of the flagellum surface and matrix proteomes provides insight into flagellum signaling in mammalian-infectious *Trypanosoma brucei*.',

Molecular & cellular proteomics : MCP, 10(10), p. M111.010538. doi: 10.1074/mcp.M111.010538.

Oberholzer, M., Marti, G., Baresic, M., Kunz, S., Hemphill, A. and Seebeck, T. (2007) 'The *Trypanosoma brucei* cAMP phosphodiesterases TbrPDEB1 and TbrPDEB2: flagellar enzymes that are essential for parasite virulence', *The FASEB Journal*, 21(3), pp. 720–731. doi: 10.1096/fj.06-6818com.

Oberle, M., Balmer, O., Brun, R. and Roditi, I. (2010) 'Bottlenecks and the maintenance of minor genotypes during the life cycle of *Trypanosoma brucei*', *PLoS Pathogens*, 6(7), pp. 1–8. doi: 10.1371/journal.ppat.1001023.

Ochiana, S. O., Pandarinath, V., Wang, Z., Kapoor, R., Ondrechen, M. J., Ruben, L. and Pollastri, M. P. (2013) 'The human Aurora kinase inhibitor danusertib is a lead compound for anti-trypanosomal drug discovery via target repurposing', *European Journal of Medicinal Chemistry*, 62, pp. 777–784. doi: 10.1016/j.ejmech.2012.07.038.

Oduor, R. O., Ojo, K. K., Williams, G. P., Bertelli, F., Mills, J., Maes, L., Pryde, D. C., Parkinson, T., van Voorhis, W. C. and Holler, T. P. (2011) 'Trypanosoma brucei Glycogen synthase kinase-3, a target for anti-trypanosomal drug development: A public-private partnership to identify novel leads', *PLoS Neglected Tropical Diseases*, 5(4). doi: 10.1371/journal.pntd.0001017.

Ogbadoyi, E., Ersfeld, K., Robinson, D., Sherwin, T. and Gull, K. (2000) 'Architecture of the *Trypanosoma brucei* nucleus during interphase and mitosis.', *Chromosoma*, 108(8), pp. 501–513. doi: 10.1007/s004120050402.

Ojo, K. K., Arakaki, T. L., Napuli, A. J., Inampudi, K. K., Keyloun, K. R., Zhang, L., Hol, W. G. J., Verlinde, C. L. M. J., Ethan, A. and Voorhis, W. C. Van (2011) 'Structure Determination of Glycogen Synthase Kinase-3 from *Leishmania major* and Comparative Inhibitor Structure-Activity Relationships with *Trypanosoma brucei* GSK-3', 176(2), pp. 98–108. doi: 10.1016/j.molbiopara.2010.12.009.Structure.

Ojo, K. K., Gillespie, J. R., Riechers, A. J., Napuli, A. J., Verlinde, C. L. M. J.,

Buckner, F. S., Gelb, M. H., Domostoj, M. M., Wells, S. J., Scheer, A., Wells, T. N. C. and Van Voorhis, W. C. (2008) 'Glycogen synthase kinase 3 is a potential drug target for African trypanosomiasis therapy.', *Antimicrobial agents and chemotherapy*, 52(10), pp. 3710–7. doi: 10.1128/AAC.00364-08.

Olsen, B. B., Svenstrup, T. H. and Guerra, B. (2012) 'Downregulation of protein kinase CK2 induces autophagic cell death through modulation of the mTOR and MAPK signaling pathways in human glioblastoma cells', *International Journal of Oncology*, 41(6), pp. 1967–1976. doi: 10.3892/ijo.2012.1635.

Ooi, C.-P., Rotureau, B., Gribaldo, S., Georgikou, C., Julkowska, D., Blisnick, T., Perrot, S., Subota, I. and Bastin, P. (2015) 'The Flagellar Arginine Kinase in *Trypanosoma brucei* Is Important for Infection in Tsetse Flies', *Plos One*, 10(7), p. e0133676. doi: 10.1371/journal.pone.0133676.

Orth, P., Schnappinger, D., Hillen, W., Saenger, W. and Hinrichs, W. (2000) 'Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system.', *Nature structural biology*, 7(3), pp. 215–9. doi: 10.1038/73324.

Overath, P., Czichos, J. and Haas, C. (1986) 'The effect of citrate/cis-aconitate on oxidative metabolism during transformation of *Trypanosoma brucei*.', *European journal of biochemistry / FEBS*, 160(1), pp. 175–182.

Overath, P. and Engstler, M. (2004) 'Endocytosis, membrane recycling and sorting of GPI-anchored proteins: *Trypanosoma brucei* as a model system', *Molecular Microbiology*, pp. 735–744. doi: 10.1111/j.1365-2958.2004.04224.x.

Pang, K. M., Ishidate, T., Nakamura, K., Shirayama, M., Trzepacz, C., Schubert, C. M., Priess, J. R. and Mello, C. C. (2004) 'The minibrain kinase homolog, mbk-2, is required for spindle positioning and asymmetric cell division in early C. elegans embryos', *Developmental Biology*, 265(1), pp. 127–139. doi: 10.1016/j.ydbio.2003.09.024.

Panova, T. B., Panov, K. I., Russell, J. and Zomerdijk, J. C. B. M. (2006) 'Casein kinase 2 associates with initiation-competent RNA polymerase I and has multiple

roles in ribosomal DNA transcription', *Molecular and Cellular Biology*, 26(16), pp. 5957–5968. doi: 10.1128/MCB.00673-06.

Paquet, C., Castilla, J., Mbulamberi, D., Beaulieu, M. F., Gastellu Etchegorry, M. G. and Moren, a (1995) 'Trypanosomiasis from *Trypanosoma brucei gambiense* in the center of north-west Uganda. Evaluation of 5 years of control (1987-1991).', *Bulletin de la Société de pathologie exotique (1990)*, pp. 38–41. PMID: 7787452.

Park, J. H., Brekken, D. L., Randall, A. C. and Parsons, M. (2002) 'Molecular cloning of *Trypanosoma brucei* CK2 catalytic subunits: The a isoform is nucleolar and phosphorylates the nucleolar protein Nopp44/46', *Molecular and Biochemical Parasitology*, 119(1), pp. 97–106. doi: 10.1016/S0166-6851(01)00407-8.

Parsons, M., Worthey, E. A., Ward, P. N. and Mottram, J. C. (2005) 'Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*', *BMC Genomics*, 6(127), p. (15 September 2005).

Patel, G., Karver, C. E., Behera, R., Guyett, P. J., Sullenberger, C., Edwards, P., Roncal, N. E., Mensa-Wilmot, K. and Pollastri, M. P. (2013) 'Kinase scaffold repurposing for neglected disease drug discovery: discovery of an efficacious, lapatinib-derived lead compound for trypanosomiasis.', *Journal of medicinal chemistry*, 56(10), pp. 3820–32. doi: 10.1021/jm400349k.

Patel, G., Roncal, N. E., Lee, P. J., Leed, S. E., Erath, J., Rodriguez, A., Sciotti, R. J. and Pollastri, M. P. (2014) 'Repurposing human Aurora kinase inhibitors as leads for anti-protozoan drug discovery.', *MedChemComm*, 5(5), pp. 655–658. doi: 10.1039/C4MD00045E.

Paterson, H. F., Savopoulos, J. W., Perisic, O., Cheung, R., Ellis, M. V, Williams, R. L. and Katan, M. (1995) 'Phospholipase C delta 1 requires a pleckstrin homology domain for interaction with the plasma membrane.', *The Biochemical journal*, 312 (Pt 3, pp. 661–6. PMID: 8554502.

Patrick, K. L., Shi, H., Kolev, N. G., Ersfeld, K., Tschudi, C. and Ullu, E. (2009)

‘Distinct and overlapping roles for two Dicer-like proteins in the RNA interference pathways of the ancient eukaryote Trypanosoma brucei.’, *Proceedings of the National Academy of Sciences of the United States of America*, 106(42), pp. 17933–8. doi: 10.1073/pnas.0907766106.

Pawson, T. and Gish, G. D. (1992) ‘SH2 and SH3 domains: from structure to function.’, *Cell*, 71(3), pp. 359–62. PMID: 1423600.

Pawson, T. and Scott, J. D. (1997) ‘Signaling through scaffold, anchoring, and adaptor proteins.’, *Science (New York, N.Y.)*, 278(5346), pp. 2075–80. PMID: 9405336.

Pays, E., Lips, S., Nolan, D., Vanhamme, L. and Pérez-Morga, D. (2001) ‘The VSG expression sites of *Trypanosoma brucei*: Multipurpose tools for the adaptation of the parasite to mammalian hosts’, *Molecular and Biochemical Parasitology*, pp. 1–16. doi: 10.1016/S0166-6851(01)00242-0.

Peacock, L., Bailey, M., Carrington, M. and Gibson, W. (2014) ‘Meiosis and haploid gametes in the pathogen *Trypanosoma brucei*’, *Current Biology*. The Authors, 24(2), pp. 181–186. doi: 10.1016/j.cub.2013.11.044.

Peacock, L., Ferris, V., Bailey, M. and Gibson, W. (2012) ‘The influence of sex and fly species on the development of trypanosomes in tsetse flies’, *PLoS Neglected Tropical Diseases*, 6(2). doi: 10.1371/journal.pntd.0001515.

Peacock, L., Ferris, V., Bailey, M. and Gibson, W. (2014) ‘Mating compatibility in the parasitic protist *Trypanosoma brucei*.’, *Parasites & vectors*, 7(1), p. 78. doi: 10.1186/1756-3305-7-78.

Peacock, L., Ferris, V., Sharma, R., Sunter, J., Bailey, M., Carrington, M. and Gibson, W. (2011) ‘Identification of the meiotic life cycle stage of *Trypanosoma brucei* in the tsetse fly.’, *Proceedings of the National Academy of Sciences of the United States of America*, 108(9), pp. 3671–6. doi: 10.1073/pnas.1019423108.

Pearson, R. B. and Kemp, B. E. (1991) ‘[3] Protein kinase phosphorylation site

sequences and consensus specificity motifs: Tabulations', in Hunter, T. and Sefton, B. M. (eds) *Protein Phosphorylation*. San Diego: Academic Press, pp. 62–81. doi: 10.1016/0076-6879(91)00127-I.

Peña, I., Pilar Manzano, M., Cantizani, J., Kessler, A., Alonso-Padilla, J., Bardera, A. I., Alvarez, E., Colmenarejo, G., Cotillo, I., Roquero, I., de Dios-Anton, F., Barroso, V., Rodriguez, A., Gray, D. W., Navarro, M., Kumar, V., Sherstnev, A., Drewry, D. H., Brown, J. R., Fiandor, J. M. and Julio Martin, J. (2015) 'New Compound Sets Identified from High Throughput Phenotypic Screening Against Three Kinetoplastid Parasites: An Open Resource', *Scientific Reports*, 5, p. 8771. doi: 10.1038/srep08771.

Perdomo, D., Bonhivers, M. and Robinson, D. R. (2016) 'The Trypanosome Flagellar Pocket Collar and Its Ring Forming Protein-TbBILBO1.', *Cells*, 5(1), p. 9. doi: 10.3390/cells5010009.

Pinna, L. a (2003) 'The raison d'etre of constitutively active protein kinases: the lesson of CK2', *Acc Chem Res*, 36(6), pp. 378–384. doi: 10.1021/ar020164f.

Pitcher, J. A., Freedman, N. J. and Lefkowitz, R. J. (1998) 'G protein-coupled receptor kinases.', *Annual review of biochemistry*, 67, pp. 653–92. doi: 10.1146/annurev.biochem.67.1.653.

Pitcher, J. A., Touhara, K., Payne, E. S. and Lefkowitz, R. J. (1995) 'Pleckstrin homology domain-mediated membrane association and activation of the beta-adrenergic receptor kinase requires coordinate interaction with G beta gamma subunits and lipid.', *The Journal of biological chemistry*, 270(20), pp. 11707–10. doi: 10.1074/jbc.270.20.11707.

Ploubidou, a, Robinson, D. R., Docherty, R. C., Ogbadoyi, E. O. and Gull, K. (1999) 'Evidence for novel cell cycle checkpoints in trypanosomes: kinetoplast segregation and cytokinesis in the absence of mitosis.', *Journal of cell science*, 112 (24), pp. 4641–4650.

Pollastri, M. P. and Campbell, R. K. (2011) 'Target repurposing for neglected diseases.', *Future medicinal chemistry*, 3(10), pp. 1307–1315. doi:

10.4155/fmc.11.92.

Proudfoot, C. and McCulloch, R. (2005) 'Distinct roles for two RAD51-related genes in *Trypanosoma brucei* antigenic variation', *Nucleic Acids Research*, 33(21), pp. 6906–6919. doi: 10.1093/nar/gki996.

Queiroz, R., Benz, C., Fellenberg, K., Hoheisel, J. D. and Clayton, C. (2009) 'Transcriptome analysis of differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons.', *BMC genomics*, 10, p. 495. doi: 10.1186/1471-2164-10-495.

Rachidi, N., Taly, J. F., Durieu, E., Leclercq, O., Aulner, N., Prina, E., Pescher, P., Notredame, C., Meijer, L. and Sp??th, G. F. (2014) 'Pharmacological assessment defines *Leishmania donovani* casein kinase 1 as a drug target and reveals important functions in parasite viability and intracellular infection', *Antimicrobial Agents and Chemotherapy*, 58(3), pp. 1501–1515. doi: 10.1128/AAC.02022-13.

Rasmussen, H. (1989) 'The cycling of calcium as an intracellular messenger.', *Scientific American*, 261(4), pp. 66–73. PMID: 2551033.

Räz, B., Iten, M., Grether-Bühler, Y., Kaminsky, R. and Brun, R. (1997) 'The Alamar Blue assay to determine drug sensitivity of African trypanosomes (T.b. rhodesiense and T.b. gambiense) in vitro.', *Acta tropica*, 68(2), pp. 139–47. PMID: 9386789.

Redmond, S., Vadivelu, J. and Field, M. C. (2003) 'RNAit: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*', *Molecular and Biochemical Parasitology*, 128(1), pp. 115–118. doi: 10.1016/S0166-6851(03)00045-8.

Ren, L., Willet, a. H., Roberts-Galbraith, R. H., McDonald, N. a., Feoktistova, a., Chen, J.-S., Huang, H., Guillen, R., Boone, C., Sidhu, S. S., Beckley, J. R. and Gould, K. L. (2014) 'The Cdc15 and Imp2 SH3 domains cooperatively scaffold a network of proteins that redundantly ensure efficient cell division in fission yeast', *Molecular Biology of the Cell*, 26(2), pp. 256–269. doi: 10.1091/mbc.E14-

10-1451.

Reuner, B., Vassella, E., Yutzy, B. and Boshart, M. (1997) ‘Cell density triggers slender to stumpy differentiation of *Trypanosoma brucei* bloodstream forms in culture.’, *Molecular and biochemical parasitology*, 90(1), pp. 269–80. PMID: 9497048.

Reynolds, D., Hofmeister, B. T., Cliffe, L., Alabady, M., Siegel, T. N., Schmitz, R. J. and Sabatini, R. (2016) ‘Histone H3 Variant Regulates RNA Polymerase II Transcription Termination and Dual Strand Transcription of siRNA Loci in *Trypanosoma brucei*’, *PLOS Genetics*. Edited by L. Figueiredo, 12(1), p. e1005758. doi: 10.1371/journal.pgen.1005758.

Richardson, J., Beecroft, R., Tolson, D., Liu, M. and Pearson, T. (1988) ‘Procyclin: an unusual immunodominant glycoprotein surface antigen from the procyclic stage of African trypanosomes’, *Molecular and Biochemical Parasitology*, 31(3), pp. 203–216. doi: 10.1016/0166-6851(88)90150-8.

Rico, E., Rojas, F., Mony, B. M., Szoor, B., Macgregor, P. and Matthews, K. R. (2013) ‘Bloodstream form pre-adaptation to the tsetse fly in *Trypanosoma brucei*.’, *Frontiers in cellular and infection microbiology*, 3(November), p. 78. doi: 10.3389/fcimb.2013.00078.

Robays, J., Bilengue, M. M. C., Van Der Stuyft, P. and Boelaert, M. (2004) ‘The effectiveness of active population screening and treatment for sleeping sickness control in the Democratic Republic of Congo’, *Tropical Medicine and International Health*, 9(5), pp. 542–550. doi: 10.1111/j.1365-3156.2004.01240.x.

Robays, J., Ebeja Kadima, a, Lutumba, P., Miaka mia Bilenge, C., Kande Betu Ku Mesu, V., De Deken, R., Makabuza, J., Deguerry, M., Van der Stuyft, P. and Boelaert, M. (2004) ‘Human African trypanosomiasis amongst urban residents in Kinshasa: a case-control study.’, *Tropical medicine & International Health*, 9(8), pp. 869–75. doi: 10.1111/j.1365-3156.2004.01266.x.

Rocha, G., Martins, A., Gama, G., Brandão, F. and Atouguia, J. (2004) ‘Possible cases of sexual and congenital transmission of sleeping sickness’, *The Lancet*,

363(9404), p. 247. doi: 10.1016/S0140-6736(03)15345-7.

Rodgers, J. (2010) ‘Trypanosomiasis and the brain.’, *Parasitology*, 137(14), pp. 1995–2006. doi: 10.1017/S0031182009991806.

Rodriguez, M. M., Ron, D., Touhara, K., Chen, C. H. and Mochly-Rosen, D. (1999) ‘RACK1, a protein kinase C anchoring protein, coordinates the binding of activated protein kinase C and select pleckstrin homology domains in vitro’, *Biochemistry*, 38(42), pp. 13787–13794. doi: 10.1021/bi991055k.

Rojas-Benítez, D., Ibar, C. and Glavic, Á. (2013) ‘The Drosophila EKC/KEOPS complex: roles in protein synthesis homeostasis and animal growth.’, *Fly*, 7(3), pp. 168–172. doi: 10.4161/fly.25227.

Rolin, S., Hancocq-Quertier, J., Paturiaux-Hanocq, F., Nolan, D. P. and Pays, E. (1998) ‘Mild acid stress as a differentiation trigger in *Trypanosoma brucei*.’, *Molecular and biochemical parasitology*, 93(2), pp. 251–62. PMID: 9662709.

Rolin, S., Paindavoine, P., Hanocq-Quertier, J., Hanocq, F., Claes, Y., Le Ray, D., Overath, P. and Pays, E. (1993) ‘Transient adenylate cyclase activation accompanies differentiation of *Trypanosoma brucei* from bloodstream to procyclic forms’, *Molecular and Biochemical Parasitology*, 61(1), pp. 115–125. doi: 10.1016/0166-6851(93)90164-S.

Rose, C., Belmonte, R., Armstrong, S. D., Molyneux, G., Haines, L. R., Lehane, M. J., Wastling, J. and Acosta-Serrano, A. (2014) ‘An Investigation into the Protein Composition of the Teneral *Glossina morsitans morsitans* Peritrophic Matrix’, *PLoS Neglected Tropical Diseases*, 8(4). doi: 10.1371/journal.pntd.0002691.

Rossomando, A., Wu, J., Weber, M. J. and Sturgill, T. W. (1992) ‘The phorbol ester-dependent activator of the mitogen-activated protein kinase p42mapk is a kinase with specificity for the threonine and tyrosine regulatory sites.’, *Proceedings of the National Academy of Sciences of the United States of America*, 89(12), pp. 5221–5. PMID: 1319055.

Roux, K. J., Kim, D. I., Raida, M. and Burke, B. (2012) 'A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells', *Journal of Cell Biology*, 196(6), pp. 801–810. doi: 10.1083/jcb.201112098.

Sakurai, T., Sugimoto, C. and Inoue, N. (2008) 'Identification and molecular characterization of a novel stage-specific surface protein of *Trypanosoma congolense* epimastigotes', *Molecular and Biochemical Parasitology*, 161(1), pp. 1–11. doi: 10.1016/j.molbiopara.2008.05.003.

Saldivia, M., Barquilla, A., Bart, J.-M., Diaz-González, R., Hall, M. N. and Navarro, M. (2013) 'Target of rapamycin (TOR) kinase in *Trypanosoma brucei*: an extended family.', *Biochemical Society transactions*, 41(4), pp. 934–8. doi: 10.1042/BST20130052.

Saldivia, M., Ceballos-Pérez, G., Bart, J.-M. and Navarro, M. (2016) 'The AMPKα1 Pathway Positively Regulates the Developmental Transition from Proliferation to Quiescence in *Trypanosoma brucei*', *Cell Reports*, 17(3), pp. 660–670. doi: 10.1016/j.celrep.2016.09.041.

Salminen A, Kauppinen A and Kaarniranta K (2016) 'AMPK/Snf1 signaling regulates histone acetylation: Impact on gene expression and epigenetic functions.', *Cellular signaling*. Elsevier Inc., S0898-6568(16), pp. 30063–8. doi: 10.1016/j.cellsig.2016.03.009.

Sanchez-Casalongue, M. E., Lee, J., Diamond, A., Shuldiner, S., Moir, R. D. and Willis, I. M. (2015) 'Differential Phosphorylation of a Regulatory Subunit of Protein Kinase CK2 by TOR Complex 1 Signaling and the Cdc-like Kinase Kns1', *Journal of Biological Chemistry*, 1(11), p. jbc.M114.626523. doi: 10.1074/jbc.M114.626523.

Savioli, L. and Daumerie, D. (2012) 'Accelerating Work To Overcome the Global Impact of Neglected Tropical Diseases: a Roadmap for Implementation', *WHO*, (Strategic and Technical Advisory Group for Neglected Tropical Diseases).

Sbicego, S., Vassella, E., Kurath, U., Blum, B. and 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(2), pp. 311–22. PMID: 10593184.

Schimanski, B., Nguyen, T. N. and Gu, A. (2005) 'Characterization of a Multisubunit Transcription Factor Complex Essential for Spliced-Leader RNA Gene Transcription in *Trypanosoma brucei*', *Molecular and Cellular Biology*, 25(16), pp. 7303–7313. doi: 10.1128/MCB.25.16.7303.

Schmidt, H. (1983) 'The pathogenesis of trypanosomiasis of the CNS. Studies on parasitological and neurohistological findings in trypanosoma rhodesiense infected vervet monkeys', *Virchows Archiv.A, Pathological anatomy and histopathology*, 399(3), pp. 333–343.

Schnaufer, A., Domingo, G. J. and Stuart, K. (2002) 'Natural and induced dyskinetoplastic trypanosomatids: how to live without mitochondrial DNA.', *International journal for parasitology*, 32(9), pp. 1071–84. PMID: 12117490.

Schumann Burkard, G., Jutzi, P. and Roditi, I. (2011) 'Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters.', *Molecular and biochemical parasitology*. Elsevier B.V., 175(1), pp. 91–4. doi: 10.1016/j.molbiopara.2010.09.002.

Schwede, A., Macleod, O. J. S., MacGregor, P. and Carrington, M. (2015) 'How Does the VSG Coat of Bloodstream Form African Trypanosomes Interact with External Proteins?', *PLoS Pathogens*, 11(12), pp. 1–18. doi: 10.1371/journal.ppat.1005259.

Scott, J. D. and Pawson, T. (2000) 'Cell communication: the inside story.', *Scientific American*, 282(6), pp. 72–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10862426.

Seixas, J. D., Luengo-Arratta, S. A., Diaz, R., Saldivia, M., Rojas-Barros, D. I., Manzano, P., Gonzalez, S., Berlanga, M., Smith, T. K., Navarro, M. and Pollastri, M. P. (2014) 'Establishment of a structure-activity relationship of 1 H-imidazo[4,5-c]quinoline-based kinase inhibitor NVP-BEZ235 as a lead for African sleeping sickness', *Journal of Medicinal Chemistry*, 57(11), pp. 4834–4848. doi:

10.1021/jm500361r.

Servant, G., Weiner, O. D., Herzmark, P., Balla, T., Sedat, J. W. and Bourne, H. R. (2016) 'Polarization of Chemoattractant Receptor Signaling During Neutrophil Chemotaxis', 287(5455), pp. 1037–1040.

Sharma, R., Gluenz, E., Peacock, L., Gibson, W., Gull, K. and Carrington, M. (2009) 'The heart of darkness: growth and form of *Trypanosoma brucei* in the tsetse fly', *Trends in Parasitology*, pp. 517–524. doi: 10.1016/j.pt.2009.08.001.

Sharma, R., Peacock, L., Gluenz, E., Gull, K., Gibson, W. and Carrington, M. (2008) 'Asymmetric cell division as a route to reduction in cell length and change in cell morphology in trypanosomes.', *Protist*, 159(1), pp. 137–51. doi: 10.1016/j.protis.2007.07.004.

Shaw, A. P. M., Torr, S. J., Waiswa, C., Cecchi, G., Wint, G. R. W., Mattioli, R. C. and Robinson, T. P. (2013a) 'Estimating the costs of tsetse control options: An example for Uganda', *Preventive Veterinary Medicine*. Elsevier B.V., 110(3–4), pp. 290–303. doi: 10.1016/j.prevetmed.2012.12.014.

Shaw, A. P. M., Torr, S. J., Waiswa, C., Cecchi, G., Wint, G. R. W., Mattioli, R. C. and Robinson, T. P. (2013b) 'Reply to the letter to the editor by Bouyer et al. (2013)', *Preventive Veterinary Medicine*, 112(3–4), pp. 447–449. doi: 10.1016/j.prevetmed.2013.08.001.

Sheader, K., Vaughan, S., Minchin, J., Hughes, K., Gull, K. and Rudenko, G. (2005) 'Variant surface glycoprotein RNA interference triggers a precytokinesis cell cycle arrest in African trypanosomes.', *Proceedings of the National Academy of Sciences of the United States of America*, 102(24), pp. 8716–21. doi: 10.1073/pnas.0501886102.

Shenolikar, S. (1988) 'Protein phosphorylation: hormones, drugs, and bioregulation.', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 2(12), pp. 2753–64. PMID: 2842213.

Shi, H., Djikeng, A., Mark, T., Wirtz, E., Tschudi, C. and Ullu, E. (2000) 'Genetic

interference in *Trypanosoma brucei* by heritable and inducible double-stranded RNA', *RNA*, 6, pp. 1069–1076.

Shi, H., Djikeng, A., Tschudi, C. and Ullu, E. (2004) 'Argonaute protein in the early divergent eukaryote *Trypanosoma brucei*: control of small interfering RNA accumulation and retroposon transcript abundance.', *Molecular and cellular biology*, 24(1), pp. 420–7. doi: 10.1128/MCB.24.1.420.

Shi, H., Tschudi, C. and Ullu, E. (2006) 'An unusual Dicer-like1 protein fuels the RNA interference pathway in *Trypanosoma brucei*.', *Rna*, 12(12), pp. 2063–2072. doi: 10.1261/rna.246906.

Shi, Y., Han, G., Wu, H., Ye, K., Tian, Z., Wang, J., Shi, H., Ye, M., Zou, H. and Huo, K. (2009) 'Casein kinase 2 interacts with human mitogen- and stress-activated protein kinase MSK1 and phosphorylates it at Multiple sites', *BMB Reports*, 42(12), pp. 840–845. doi: 10.5483/BMBRep.2009.42.12.840.

Shun, T. Y., Lazo, J. S., Sharlow, E. R. and Johnston, P. A. (2011) 'Identifying Actives from HTS Data Sets Practical Approaches for the Selection of an Appropriate HTS Data-Processing Method and Quality Control Review', *Journal of Biomolecular Screening*, 16(1), pp. 1–14. doi: 10.1177/1087057110389039.

Sicheri, F. and Kuriyan, J. (1997) 'Structures of Src-family tyrosine kinases', *Current Opinion in Structural Biology*, 7(6), pp. 777–785. doi: 10.1016/S0959-440X(97)80146-7.

Sidik, S. M., Huet, D., Ganesan, S. M., Huynh, M.-H., Wang, T., Nasamu, A. S., Thiru, P., Saeij, J. P. J., Carruthers, V. B., Niles, J. C. and Lourido, S. (2016) 'A Genome-wide CRISPR Screen in *Toxoplasma* Identifies Essential Apicomplexan Genes', *Cell*. Elsevier Inc., 166(6), p. 1423–1435.e12. doi: 10.1016/j.cell.2016.08.019.

Siegel, T. N., Hekstra, D. R. and Cross, G. A. M. (2008) 'Analysis of the *Trypanosoma brucei* cell cycle by quantitative DAPI imaging', *Molecular and Biochemical Parasitology*, 160(2), pp. 171–174. doi: 10.1016/j.molbiopara.2008.04.004.

Siegel, T. N., Hekstra, D. R., Kemp, L. E., Figueiredo, L. M., Lowell, J. E., Fenyo, D., Wang, X., Dewell, S. and Cross, G. A. M. (2009) 'Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*', *Genes & Development*, 23(9), pp. 1063–1076. doi: 10.1101/gad.1790409.

Siegel, T. N., Hekstra, D. R., Wang, X., Dewell, S. and Cross, G. a M. (2010) 'Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites.', *Nucleic acids research*, 38(15), pp. 4946–4957. doi: 10.1093/nar/gkq237.

Simarro, P. P., Cecchi, G., Franco, J. R., Paone, M., Diarra, A., Priotto, G., Mattioli, R. C. and Jannin, J. G. (2015) 'Monitoring the Progress towards the Elimination of Gambiense Human African Trypanosomiasis.', *PLoS neglected tropical diseases*, 9(6), p. e0003785. doi: 10.1371/journal.pntd.0003785.

Simarro, P. P., Franco, J. R., Ndongo, P., Nguema, E., Louis, F. J. and Jannin, J. (2006) 'The elimination of *Trypanosoma brucei gambiense* sleeping sickness in the focus of Luba, Bioko Island, Equatorial Guinea', *Tropical Medicine and International Health*, 11(5), pp. 636–646. doi: 10.1111/j.1365-3156.2006.01624.x.

Simon, M. N., Pelegrini, O., Veron, M. and Kay, R. R. (1992) 'Mutation of protein kinase A causes heterochronic development of *Dictyostelium*.', *Nature*, 356(6365), pp. 171–2. doi: 10.1038/356171a0.

Simpson, A. G. B., Stevens, J. R. and Lukeš, J. (2006) 'The evolution and diversity of kinetoplastid flagellates', *Trends in Parasitology*, pp. 168–174. doi: 10.1016/j.pt.2006.02.006.

Simpson, L., Aphasizhev, R., Gao, G. and Kang, X. (2004) 'Mitochondrial proteins and complexes in *Leishmania* and *Trypanosoma* involved in U-insertion/deletion RNA editing', *Rna*, 10(2), pp. 159–170. doi: 10.1261/rna.5170704.RNA.

Siomi, H. and Siomi, M. C. (2009) 'On the road to reading the RNA-interference code', *Nature*, 457(January), pp. 396–404. doi: 10.1038/nature07754.

Siomi, M. C., Sato, K., Pezic, D. and Aravin, A. a (2011) 'PIWI-interacting small RNAs: the vanguard of genome defence.', *Nature reviews. Molecular cell biology*. Nature Publishing Group, 12(4), pp. 246–258. doi: 10.1038/nrm3089.

Sleator, R. D. and Hill, C. (2002) 'Bacterial osmoadaptation: The role of osmolytes in bacterial stress and virulence', *FEMS Microbiology Reviews*, 26(1), pp. 49–71. doi: 10.1016/S0168-6445(01)00071-7.

Smith, D. H., Pepin, J. and Stich, A. H. R. (1998) 'Human African trypanosomiasis : an emerging public health crisis', 54(2), pp. 341–355.

Smith, T. K., Vasileva, N., Gluenz, E., Terry, S., Portman, N., Kramer, S., Carrington, M., Michaeli, S., Gull, K. and Rudenko, G. (2009) 'Blocking variant surface glycoprotein synthesis in *Trypanosoma brucei* triggers a general arrest in translation initiation', *PLoS ONE*, 4(10). doi: 10.1371/journal.pone.0007532.

Soisson, S. M., Nimnual, A. S., Uy, M., Bar-Sagi, D. and Kuriyan, J. (1998) 'Crystal structure of the Dbl and pleckstrin homology domains from the human Son of sevenless protein', *Cell*, 95(2), pp. 259–268. doi: 10.1016/S0092-8674(00)81756-0.

Solanki, S., Innocenti, P., Mas-Droux, C., Boxall, K., Barillari, C., Van Montfort, R. L. M., Aherne, G. W., Bayliss, R. and Hoelder, S. (2011) 'Benzimidazole inhibitors induce a DFG-out conformation of never in mitosis gene a-related kinase 2 (Nek2) without binding to the back pocket and reveal a nonlinear structure-activity relationship', *Journal of Medicinal Chemistry*, 54(6), pp. 1626–1639. doi: 10.1021/jm1011726.

Sollelis, L., Ghorbal, M., Macpherson, C. R., Martins, R. M., Kuk, N., Crobu, L., Bastien, P., Scherf, A., Lopez-Rubio, J. J. and Sterkers, Y. (2015) 'First efficient CRISPR-Cas9-mediated genome editing in *Leishmania* parasites', *Cellular Microbiology*, 17(10), pp. 1405–1412. doi: 10.1111/cmi.12456.

Songyang, Z., Carraway, K. L., Eck, M. J., Harrison, S. C., Feldman, R. A., Mohammadi, M., Schlessinger, J., Hubbard, S. R., Smith, D. P. and Eng, C. (1995) 'Catalytic specificity of protein-tyrosine kinases is critical for selective

signalling.', *Nature*, 373(6514), pp. 536–9. doi: 10.1038/373536a0.

Srinivasan, M., Mehta, P., Yu, Y., Prugar, E., Koonin, E. V, Karzai, a W. and Sternglanz, R. (2011) 'The highly conserved KEOPS/EKC complex is essential for a universal tRNA modification, t6A.', *The EMBO journal*. Nature Publishing Group, 30(5), pp. 873–81. doi: 10.1038/emboj.2010.343.

Stanghellini, A. and Josenando, T. (2001) 'The situation of sleeping sickness in Angola: A calamity', *Tropical Medicine and International Health*, 6(5), pp. 330–334. doi: 10.1046/j.1365-3156.2001.00724.x.

Stanne, T. M., Kushwaha, M., Wand, M., Taylor, J. E. and Rudenko, G. (2011) 'TbISWI regulates multiple polymerase I (Pol I)-transcribed loci and is present at Pol II transcription boundaries in *Trypanosoma brucei*', *Eukaryotic Cell*, 10(7), pp. 964–976. doi: 10.1128/EC.05048-11.

Steiger, R. F. (1973) 'On the ultrastructure of *Trypanosoma* (Trypanozoon) *brucei* in the course of its life cycle and some related aspects.', *Acta tropica*, 30(1), pp. 64–168. PMID: 4144959.

Sternberg, J. M., Gierliński, M., Biéler, S., Ferguson, M. A. J. and Ndung'u, J. M. (2014) 'Evaluation of the diagnostic accuracy of prototype rapid tests for human African trypanosomiasis.', *PLoS neglected tropical diseases*, 8(12), p. e3373. doi: 10.1371/journal.pntd.0003373.

Stortz, J. A. (2016) *Identification and characterisation of novel Trypanosoma brucei protein kinases involved in repair of cellular damage*. PhD thesis. University of Glasgow.

Streit, J. A. and Matsumoto, E. (2016) 'African Trypanosomiasis', *New England Journal of Medicine*, 375(24), pp. 2380–2380. doi: 10.1056/NEJMicm1604333.

Stuart, K., Brun, R., Croft, S., Fairlamb, A., Gürtler, R. E., McKerrow, J., Reed, S. and Tarleton, R. (2008) 'Kinetoplastids: related protozoan pathogens, different diseases.', *The Journal of clinical investigation*, 118(4), pp. 1301–10. doi: 10.1172/JCI33945.

Subramaniam, C., Veazey, P., Redmond, S., Hayes-Sinclair, J., Chambers, E., Carrington, M., Gull, K., Matthews, K., Horn, D. and Field, M. C. (2006) 'Chromosome-wide analysis of gene function by RNA interference in the african trypanosome.', *Eukaryotic cell*, 5(9), pp. 1539–49. doi: 10.1128/EC.00141-06.

Subramanya, S. and Mensa-Wilmot, K. (2010) 'Diacylglycerol-stimulated endocytosis of transferrin in trypanosomatids is dependent on tyrosine kinase activity', *PLoS ONE*, 5(1), pp. 1–9. doi: 10.1371/journal.pone.0008538.

Sullivan, L., Wall, S. J., Carrington, M. and Ferguson, M. A. J. (2013) 'Proteomic Selection of Immunodiagnostic Antigens for Human African Trypanosomiasis and Generation of a Prototype Lateral Flow Immunodiagnostic Device', *PLoS Neglected Tropical Diseases*. Edited by C. Tschudi, 7(2), p. e2087. doi: 10.1371/journal.pntd.0002087.

Sunter, J. D. and Gull, K. (2016) 'The Flagellum Attachment Zone: "The Cellular Ruler" of Trypanosome Morphology', *Trends in Parasitology*, pp. 309–324. doi: 10.1016/j.pt.2015.12.010.

Sutherland, C. S., Yukich, J., Goeree, R. and Tediosi, F. (2015) 'A Literature Review of Economic Evaluations for a Neglected Tropical Disease: Human African Trypanosomiasis ("Sleeping Sickness")', *PLOS Neglected Tropical Diseases*. Edited by J. M. Ndung'u, 9(2), p. e0003397. doi: 10.1371/journal.pntd.0003397.

Swinney, D. C. (2013) 'Phenotypic vs. Target-Based Drug Discovery for First-in-Class Medicines', *Clin Pharmacol Ther*, 93(4), pp. 299–301. doi: 10.1038/clpt.2012.236.

Szöőr, B., Dyer, N. a, Ruberto, I., Acosta-Serrano, A. and Matthews, K. R. (2013) 'Independent Pathways Can Transduce the Life-Cycle Differentiation Signal in *Trypanosoma brucei*.', *PLoS pathogens*, 9(10), p. e1003689. doi: 10.1371/journal.ppat.1003689.

Szöor, B., Ruberto, I., Burchmore, R. and Matthews, K. R. (2010) 'A novel phosphatase cascade regulates differentiation in *Trypanosoma brucei* via a glycosomal signaling pathway.', *Genes & development*, 24(12), pp. 1306–16. doi:

10.1101/gad.570310.

Szöor, B., Wilson, J., McElhinney, H., Tabernero, L. and Matthews, K. R. (2006) 'Protein tyrosine phosphatase TbPTP1: A molecular switch controlling life cycle differentiation in trypanosomes.', *The Journal of cell biology*, 175(2), pp. 293–303. doi: 10.1083/jcb.200605090.

Taylor, S. S. and Radzio-Andzelm, E. (1994) 'Three protein kinase structures define a common motif.', *Structure (London, England : 1993)*, 2(5), pp. 345–55. PMID: 8081750.

Taylor, S. S., Shaw, A., Hu, J., Meharena, H. S. and Kornev, A. (2013) 'Pseudokinases from a structural perspective', *Biochemical Society Transactions*, 41(4), pp. 981–986. doi: 10.1042/BST20130120.

Tejedor, F., Zhu, X. R., Kaltenbach, E., Ackermann, A., Baumann, A., Canal, I., Heisenberg, M., Fischbach, K. F. and Pongs, O. (1995) 'minibrain: A new protein kinase family involved in postembryonic neurogenesis in Drosophila', *Neuron*, 14(2), pp. 287–301. doi: 10.1016/0896-6273(95)90286-4.

Tetley, L. and Vickerman, K. (1985) 'Differentiation in *Trypanosoma brucei*: host-parasite cell junctions and their persistence during acquisition of the variable antigen coat.', *Journal of cell science*, 74, pp. 1–19. PMID: 4030903.

Thuita, J. K., Kagira, J. M., Mwangangi, D., Matovu, E., Turner, C. M. R. and Masiga, D. (2008) '*Trypanosoma brucei rhodesiense* transmitted by a single tsetse fly bite in vervet monkeys as a model of human African trypanosomiasis', *PLoS Neglected Tropical Diseases*, 2(5). doi: 10.1371/journal.pntd.0000238.

Tiengwe, C., Marques, C. A. and McCulloch, R. (2014) 'Nuclear DNA replication initiation in kinetoplastid parasites: New insights into an ancient process', *Trends in Parasitology*, pp. 27–36. doi: 10.1016/j.pt.2013.10.009.

Touhara, K., Inglese, J., Pitcher, J. A., Shaw, G. and Lefkowitz, R. J. (1994) 'Binding of G protein beta gamma-subunits to pleckstrin homology domains', *J Biol Chem*, 269(14), pp. 10217–10220. PMID: 8144601.

Trindade, S., Rijo-Ferreira, F., Carvalho, T., Pinto-Neves, D., Guegan, F., Aresta-Branco, F., Bento, F., Young, S. A., Pinto, A., Van Den Abbeele, J., Ribeiro, R. M., Dias, S., Smith, T. K. and Figueiredo, L. M. (2016) 'T*rypanosoma brucei* Parasites Occupy and Functionally Adapt to the Adipose Tissue in Mice', *Cell Host & Microbe*, pp. 837–848. doi: 10.1016/j.chom.2016.05.002.

Truc, P., Jamonneau, V., N'Guessan, P., N'Dri, L., Diallo, P. B. and Cuny, G. (1996) '*Trypanosoma brucei ssp. and T congolense*: mixed human infection in Côte d'Ivoire.', *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 92(5), pp. 537–8. PMID: 9861372.

Tu, X. and Wang, C. C. (2004) 'The Involvement of Two cdc2-related Kinases (CRKs) in *Trypanosoma brucei* Cell Cycle Regulation and the Distinctive Stage-specific Phenotypes Caused by CRK3 Depletion', *Journal of Biological Chemistry*, 279(19), pp. 20519–20528. doi: 10.1074/jbc.M312862200.

Turner, C. M., Aslam, N. and Dye, C. (1995) 'Replication, differentiation, growth and the virulence of *Trypanosoma brucei* infections.', *Parasitology*, 111 (Pt 3, pp. 289–300. PMID: 7567097.

Turner, C. M. R. (1990) 'The use of experimental artefacts in African trypanosome research', *Parasitology Today*, 6(1), pp. 14–17. doi: 10.1016/0169-4758(90)90385-H.

Tyler, K. M., Higgs, P. G., Matthews, K. R. and Gull, K. (2001) 'Limitation of *Trypanosoma brucei* parasitaemia results from density-dependent parasite differentiation and parasite killing by the host immune response.', *Proceedings. Biological sciences*, 268(1482), pp. 2235–43. doi: 10.1098/rspb.2001.1794.

Ullal, P., McDonald, N. A., Chen, J. S., Presti, L. Lo, Roberts-Galbraith, R. H., Gould, K. L. and Martin, S. G. (2015) 'The DYRK-family kinase Pom1 phosphorylates the F-BAR protein Cdc15 to prevent division at cell poles', *Journal of Cell Biology*, 211(3), pp. 653–668. doi: 10.1083/jcb.201504073.

Ullu, E., Tschudi, C. and Chakraborty, T. (2004) 'RNA interference in protozoan parasites', *Cellular Microbiology*, 6(6), pp. 509–519. doi: 10.1111/j.1462-

5822.2004.00399.x.

Urbaniak, M. D. (2009) ‘Casein kinase 1 isoform 2 is essential for bloodstream form *Trypanosoma brucei*.’, *Molecular and biochemical parasitology*, 166(2), pp. 183–5. doi: 10.1016/j.molbiopara.2009.03.001.

Urbaniak, M. D., Mathieson, T., Bantscheff, M., Eberhard, D., Grimaldi, R., Miranda-Saavedra, D., Wyatt, P., Ferguson, M. a J., Frearson, J. and Drewes, G. (2012) ‘Chemical proteomic analysis reveals the drugability of the kinome of *Trypanosoma brucei*.’, *ACS chemical biology*, 7(11), pp. 1858–65. doi: 10.1021/cb300326z.

Urwyler, S., Studer, E., Renggli, C. K. and Roditi, I. (2007) ‘A family of stage-specific alanine-rich proteins on the surface of epimastigote forms of *Trypanosoma brucei*’, *Molecular Microbiology*, 63(1), pp. 218–228. doi: 10.1111/j.1365-2958.2006.05492.x.

Urwyler, S., Vassella, E., Van Den Abbeele, J., Renggli, C. K., Blundell, P., Barry, J. D. and Roditi, I. (2005) ‘Expression of procyclin mRNAs during cyclical transmission of *Trypanosoma brucei*’, *PLoS Pathogens*, 1(3), pp. 0225–0231. doi: 10.1371/journal.ppat.0010022.

Usacheva, A., Tian, X., Sandoval, R., Salvi, D., Levy, D. and Colamonici, O. R. (2003) ‘The WD motif-containing protein RACK-1 functions as a scaffold protein within the type I IFN receptor-signaling complex.’, *Journal of immunology (Baltimore, Md. : 1950)*, 171(6), pp. 2989–2994. doi: 10.4049/jimmunol.171.6.2989.

Valenciano, A. L., Knudsen, G. M. and Mackey, Z. B. (2016) ‘Extracellular-signal Regulated Kinase 8 of *Trypanosoma brucei* Uniquely Phosphorylates its Proliferating Cell Nuclear Antigen Homolog and Reveals Exploitable Properties’, *Cell Cycle*. Taylor & Francis, 15(20), pp. 00–00. doi: 10.1080/15384101.2016.1222340.

Valenciano, A. L., Ramsey, A. C., Santos, W. L. and Mackey, Z. B. (2016) ‘Discovery and antiparasitic activity of AZ960 as a *Trypanosoma brucei* ERK8

inhibitor', *Bioorganic & Medicinal Chemistry*, 24(19), pp. 4647–4651. doi: 10.1016/j.bmc.2016.07.069.

Vanhollebeke, B., De Muylder, G., Nielsen, M. J., Pays, A., Tebabi, P., Dieu, M., Raes, M., Moestrup, S. K. and Pays, E. (2008) 'A haptoglobin-hemoglobin receptor conveys innate immunity to *Trypanosoma brucei* in humans.', *Science (New York, N.Y.)*, 320(5876), pp. 677–81. doi: 10.1126/science.1156296.

Vassella, E., Van Den Abbeele, J., Bütikofer, P., Renggli, C. K., Furger, A., Brun, R. and 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 and Development*, 14(5), pp. 615–626. doi: 10.1101/gad.14.5.615.

Vassella, E., Krämer, R., Turner, C. M., Wankell, M., Modes, C., van den Bogaard, M. and Boshart, M. (2001) 'Deletion of a novel protein kinase with PX and FYVE-related domains increases the rate of differentiation of *Trypanosoma brucei*.', *Molecular microbiology*, 41(1), pp. 33–46. PMID: 11454198.

Vassella, E., Probst, M., Studer, E., Renggli, C. K. and Roditi, I. (2004) 'Expression of a Major Surface Protein of *Trypanosoma brucei* Insect Forms Is Controlled by the Activity of Mitochondrial Enzymes', 15(September), pp. 3986–3993. doi: 10.1091/mbc.E04.

Vassella, E., Reuner, B., Yutzy, B. and Boshart, M. (1997) 'Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway.', *Journal of cell science*, 110 (Pt 2, pp. 2661–71. PMID: 9427384.

Vaughan, S. and Gull, K. (2008) 'The structural mechanics of cell division in *Trypanosoma brucei*.', *Biochemical Society transactions*, 36(Pt 3), pp. 421–424. doi: 10.1042/BST0360421.

Vickerman, K. (1969) 'On the surface coat and flagellar adhesion in trypanosomes.', *Journal of cell science*, 5, pp. 163–193.

Vickerman, K. (1985) 'Developmental cycles and biology of pathogenic trypanosomes.', *British medical bulletin*, 41(2), pp. 105–14. PMID: 3928017.

Voorheis, H. P. and Martin, B. R. (1980) '"Swell dialysis" demonstrates that adenylate cyclase in *Trypanosoma brucei* is regulated by calcium ions', *Eur.J.Biochem.*, 113(1), pp. 223–227. PMID: 6257514.

Wager, T. T., Chandrasekaran, R. Y., Hou, X., Troutman, M. D., Verhoest, P. R., Villalobos, A. and Will, Y. (2010) 'Defining desirable central nervous system drug space through the alignment of molecular properties, in vitro ADME, and safety attributes', *ACS Chemical Neuroscience*, 1(6), pp. 420–434. doi: 10.1021/cn100007x.

Wager, T. T., Hou, X., Verhoest, P. R. and Villalobos, A. (2010) 'Moving beyond rules: The development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties', *ACS Chemical Neuroscience*, 1(6), pp. 435–449. doi: 10.1021/cn100008c.

Walker, R. G., Thomson, G., Malone, K., Nowicki, M. W., Brown, E., Blake, D. G., Turner, N. J., Walkinshaw, M. D., Grant, K. M. and Mottram, J. C. (2011) 'High throughput screens yield small molecule inhibitors of leishmania CRK3:CYC6 cyclin-dependent kinase', *PLoS Neglected Tropical Diseases*, 5(4). doi: 10.1371/journal.pntd.0001033.

Wallbanks, K. R., Molyneux, D. H. and Dirie, M. F. (1989) 'Chitin derivatives as novel substrates for *Trypanosoma brucei brucei* attachment *in vitro*.', *Acta tropica*, 46(1), pp. 63–8. PMID: 2566263.

Walshe, D. P., Lehane, M. J. and Haines, L. R. (2011) 'Post eclosion age predicts the prevalence of midgut trypanosome infections in *Glossina*', *PLoS ONE*, 6(11). doi: 10.1371/journal.pone.0026984.

Ward, P., Equinet, L., Packer, J. and Doerig, C. (2004) 'Protein kinases of the human malaria parasite *Plasmodium falciparum*: the kinome of a divergent eukaryote.', *BMC genomics*, 5, p. 79. doi: 10.1186/1471-2164-5-79.

Weinmann, H. and Metternich, R. (2005) 'Drug discovery process for kinase inhibitors', *ChemBioChem*, 6(3), pp. 455–459. doi: 10.1002/cbic.200500034.

Weinmaster, G., Zoller, M. J. and Pawson, T. (1986) 'A lysine in the ATP-binding site of P130gag-fps is essential for protein-tyrosine kinase activity.', *The EMBO journal*, 5(1), pp. 69–76. PMID: 3007119.

Wenzler, T., Schumann Burkard, G., S Schmidt, R., Mäser, P., Bergner, A., Roditi, I. and Brun, R. (2016) 'A new approach to chemotherapy: drug-induced differentiation kills African trypanosomes.', *Scientific reports*, 6(October 2015), p. 22451. doi: 10.1038/srep22451.

Wheeler, R. J., Gluenz, E. and Gull, K. (2013) 'The limits on trypanosomatid morphological diversity', *PLoS ONE*, 8(11). doi: 10.1371/journal.pone.0079581.

Whelligan, D. K., Solanki, S., Taylor, D., Thomson, D. W., Cheung, K. M. J., Boxall, K., Mas-Droux, C., Barillari, C., Burns, S., Grummitt, C. G., Collins, I., Van Montfort, R. L. M., Aherne, G. W., Bayliss, R. and Hoelder, S. (2010) 'Aminopyrazine inhibitors binding to an unusual inactive conformation of the mitotic kinase Nek2: SAR and structural characterization', *Journal of Medicinal Chemistry*, 53(21), pp. 7682–7698. doi: 10.1021/jm1008727.

Wickstead, B., Ersfeld, K. and Gull, K. (2004) 'The small chromosomes of *Trypanosoma brucei* involved in antigenic variation are constructed around repetitive palindromes.', *Genome research*, 14(6), pp. 1014–24. doi: 10.1101/gr.2227704.

Widmann, C., Gibson, S., Jarpe, M. B. and Johnson, G. L. (1999) 'Mitogen-activated protein kinase: conservation of a three-kinase module from yeast to human.', *Physiological reviews*, 79(1), pp. 143–180.

Wiese, M. (1998) 'A mitogen-activated protein (MAP) kinase homologue of *Leishmania mexicana* is essential for parasite survival in the infected host', *EMBO Journal*, 17(9), pp. 2619–2628. doi: 10.1093/emboj/17.9.2619.

Wilkinson, S. R. and Kelly, J. M. (2009) 'Trypanocidal drugs: mechanisms,

resistance and new targets.', *Expert reviews in molecular medicine*, 11(October), p. e31. doi: 10.1017/S1462399409001252.

Wirtz, E. and Clayton, C. (1995) 'Inducible gene expression in trypanosomes mediated by a prokaryotic repressor.', *Science*, 268(5214), pp. 1179–83. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7761835.

Wirtz, E., Leal, S., Ochatt, C. and Cross, G. a (1999) 'A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*.', *Molecular and biochemical parasitology*, 99(1), pp. 89–101. PMID: 10215027.

Wolburg, H., Mogk, S., Acker, S., Frey, C., Meinert, M., Schönfeld, C., Lazarus, M., Urade, Y., Kubata, B. K. and Duszenko, M. (2012) 'Late Stage Infection in Sleeping Sickness', *PLoS ONE*. Edited by S. N. Moreno, 7(3), p. e34304. doi: 10.1371/journal.pone.0034304.

Woodland, A., Grimaldi, R., Luksch, T., Cleghorn, L. A. T., Ojo, K. K., VanVoorhis, W. C., Brenk, R., Frearson, J. A., Gilbert, I. H. and Wyatt, P. G. (2013) 'From on-target to off-target activity: Identification and optimisation of *Trypanosoma brucei* GSK3 inhibitors and their characterisation as anti-*Trypanosoma brucei* drug discovery lead molecules', *ChemMedChem*, 8(7), pp. 1127–1137. doi: 10.1002/cmdc.201300072.

Woodland, A., Thompson, S., Cleghorn, L. a T., Norcross, N., Rycker, M. De, Grimaldi, R., Hallyburton, I., Rao, B., Norval, S., Stojanovski, L., Brun, R., Kaiser, M., Frearson, J. a, Gray, D. W., Wyatt, P. G., Read, K. D. and Gilbert, I. H. (2015) 'Discovery of Inhibitors of *Trypanosoma brucei* by Phenotypic Screening of a Focused Protein Kinase Library', pp. 1–13. doi: 10.1002/cmdc.201500300.

Woodward, R. and Gull, K. (1990) 'Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*.', *Journal of cell science*, 95, pp. 49–57.

World Health Organization (2013) 'Control and surveillance of human African

trypanosomiasis.', *World Health Organization technical report series*, (984), pp. 1–237.

Worthylake, D. K., Rossman, K. L. and Sondek, J. (2000) 'Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1.', *Nature*, 408(6813), pp. 682–688. doi: 10.1038/35047014.

Wright, J. R., Siegel, T. N. and Cross, G. A. M. (2010) 'Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in *Trypanosoma brucei*', *Molecular and Biochemical Parasitology*, 172(2), pp. 141–144. doi: 10.1016/j.molbiopara.2010.03.013.

Wurst, M., Seliger, B., Jha, B. A., Klein, C., Queiroz, R. and Clayton, C. (2012) 'Expression of the RNA recognition motif protein RBP10 promotes a bloodstream-form transcript pattern in Trypanosoma brucei', *Molecular Microbiology*, 83(5), pp. 1048–1063. doi: 10.1111/j.1365-2958.2012.07988.x.

Wyllie, S., Foth, B. J., Kelner, a., Sokolova, a. Y., Berriman, M. and Fairlamb, a. H. (2015) 'Nitroheterocyclic drug resistance mechanisms in *Trypanosoma brucei*', *Journal of Antimicrobial Chemotherapy*, 71(3), pp. 1–10. doi: 10.1093/jac/dkv376.

Xingi, E., Smirlis, D., Myrianthopoulos, V., Magiatis, P., Grant, K. M., Meijer, L., Mikros, E., Skaltsounis, A.-L. and Soteriadou, K. (2009) '6-Br-5methylindirubin-3'oxime (5-Me-6-BIO) targeting the leishmanial glycogen synthase kinase-3 (GSK-3) short form affects cell-cycle progression and induces apoptosis-like death: exploitation of GSK-3 for treating leishmaniasis.', *International journal for parasitology*. Australian Society for Parasitology Inc., 39(12), pp. 1289–303. doi: 10.1016/j.ijpara.2009.04.005.

Van Xong, H., Vanhamme, L., Chamekh, M., Chimfwembe, C. E., Van Den Abbeele, J., Pays, A., Van Melrvenne, N., Hamers, R., De Baetselier, P. and Pays, E. (1998) 'A VSG expression site-associated gene confers resistance to human serum in *Trypanosoma rhodesiense*', *Cell*, 95(6), pp. 839–846. doi: 10.1016/S0092-8674(00)81706-7.

Yamane, K. and Kinsella, T. J. (2005) 'CK2 Inhibits Apoptosis and Changes Its Cellular Localization Following Ionizing Radiation CK2 Inhibits Apoptosis and Changes Its Cellular Localization Following Ionizing Radiation', (10), pp. 4362–4367.

Yao, L., Janmey, P., Frigeri, L. G., Han, W., Fujita, J., Kawakami, Y., Apgar, J. R. and Kawakami, T. (1999) 'Pleckstrin homology domains interact with filamentous actin', *Journal of Biological Chemistry*, 274(28), pp. 19752–19761. doi: 10.1074/jbc.274.28.19752.

Yao, L., Suzuki, H., Ozawa, K., Deng, J., Lehel, C., Fukamachi, H., Anderson, W. B., Kawakami, Y. and Kawakami, T. (1997) 'Interactions between protein kinase C and pleckstrin homology domains. Inhibition by phosphatidylinositol 4,5-bisphosphate and phorbol 12-myristate 13-acetate', *J Biol Chem*, 272(20), pp. 13033–13039. doi: 10.1074/jbc.272.20.13033.

Zaman, S., Lippman, S. I., Zhao, X. and Broach, J. R. (2008) 'How *Saccharomyces* responds to nutrients.', *Annual review of genetics*, 42, pp. 27–81. doi: 10.1146/annurev.genet.41.110306.130206.

Zervas, C. G. and Brown, N. H. (2002) 'Integrin adhesion: when is a kinase a kinase?', *Current biology : CB*, 12(10), pp. R350-1. PMID: 12015134.

Zhang, W. W. and Matlashewski, G. (1997) 'Loss of virulence in *Leishmania donovani* deficient in an amastigote-specific protein, A2.', *Proceedings of the National Academy of Sciences of the United States of America*, 94(August), pp. 8807–8811. doi: 10.1073/pnas.94.16.8807.

Zhang, W. W. and Matlashewski, G. (2015) 'CRISPR-Cas9-mediated genome editing in *Leishmania donovani*', *mBio*, 6(4), pp. 1–14. doi: 10.1128/mBio.00861-15.

Zhang, X. D. (2011) 'Illustration of SSMD, z score, SSMD*, z* score, and t statistic for hit selection in RNAi high-throughput screens.', *Journal of biomolecular screening*, 16(7), pp. 775–85. doi: 10.1177/1087057111405851.

Zhang, Chung and Oldenburg (1999) ‘A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays.’, *Journal of biomolecular screening*, 4(2), pp. 67–73. PMID: 10838414.

Zhi, H., Tang, L., Xia, Y. and Zhang, J. (2013) ‘Ssk1p-independent activation of Ssk2p plays an important role in the osmotic stress response in *Saccharomyces cerevisiae*: alternative activation of Ssk2p in osmotic stress.’, *PloS one*, 8(2), p. e54867. doi: 10.1371/journal.pone.0054867.

Zhong, X. Y., Ding, J. H., Adams, J. A., Ghosh, G. and Fu, X. D. (2009) ‘Regulation of SR protein phosphorylation and alternative splicing by modulating kinetic interactions of SRPK1 with molecular chaperones’, *Genes and Development*, 23(4), pp. 482–495. doi: 10.1101/gad.1752109.

Zhou, Q., Hu, H. and Li, Z. (2016) ‘An EF-hand-containing protein in Trypanosoma brucei regulates cytokinesis initiation by maintaining the stability of the cytokinesis initiation factor CIF1’, *Journal of Biological Chemistry*, (7), p. jbc.M116.726133. doi: 10.1074/jbc.M116.726133.

Zhou, Z., Qiu, J., Liu, W., Zhou, Y., Plocinik, R. M., Li, H., Hu, Q., Ghosh, G., Adams, J. A., Rosenfeld, M. G. and Fu, X. D. (2012) ‘The Akt-SRPK-SR Axis Constitutes a Major Pathway in Transducing EGF Signaling to Regulate Alternative Splicing in the Nucleus’, *Molecular Cell*, 47(3), pp. 422–433. doi: 10.1016/j.molcel.2012.05.014.

Ziegelbauer, K. and Overath, P. (1993) ‘Organization of two invariant surface glycoproteins in the surface coat of *Trypanosoma brucei*’, *Infection and Immunity*, 61(11), pp. 4540–4545.

Zylbersztejn, A. M. B., de Morais, C. G. V., Lima, A. K. C., Souza, J. E. de O., Lopes, A. H., Da-Silva, S. A. G., Silva-Neto, M. A. C. and Dutra, P. M. L. (2015) ‘CK2 Secreted by *Leishmania braziliensis* Mediates Macrophage Association Invasion: A Comparative Study between Virulent and Avirulent Promastigotes’, *BioMed Research International*, 2015, pp. 1–11. doi: 10.1155/2015/167323.