



**Evaluation of Glycogen Synthase  
Kinase 3 as a drug target in African  
trypanosomes**

Raffaella Grimaldi

PhD Thesis

December 2014

Supervisor: Professor Alan H. Fairlamb

University of Dundee

To my daughter Gaia

## Table of Contents

List of abbreviations.....	VIII
Acknowledgements.....	XI
Declaration.....	XII
Abstract.....	XIII
Chapter 1 Introduction.....	1
1.1 Human African trypanosomiasis .....	2
1.1.1 Epidemiology and distribution .....	2
1.1.2 Trypanosome cell cycle in vector and in human host.....	4
1.1.3 Trypanosome cell architecture.....	6
1.1.4 Symptoms and diagnosis .....	7
1.2 Chemotherapy of sleeping sickness.....	8
1.2.1 <i>T. b. rhodesiense</i> infection.....	8
1.2.2 <i>T. b. gambiense</i> infection.....	10
1.2.3 New drug candidates for HAT.....	14
1.3 Target assessment and target product profile.....	16
1.4 Protein kinases.....	18
1.4.1 The <i>T. brucei</i> kinome .....	19
1.4.2 Protein kinases in <i>T. brucei</i> as drug targets .....	21
1.5 Glycogen synthase kinase: role in diseases and inhibitors development .....	25
1.6 Glycogen synthase kinase as a drug target in kinetoplastida .....	30
1.7 Aims of the project .....	34
Chapter 2 Materials and methods .....	35
2.1 Materials.....	36
2.2 General molecular biology .....	36
2.2.1 Isolation of genomic DNA from bloodstream form <i>T. brucei</i> .....	36
2.2.2 DNA amplification .....	37
2.2.3 Agarose gel electrophoresis and extraction of DNA .....	37
2.2.4 Cloning of PCR products.....	38
2.2.5 Transformation of competent <i>E. coli</i> cell lines.....	38
2.2.6 Purification and digestion of plasmid DNA .....	39
2.2.7 DNA sequencing .....	39
2.2.8 Ligation .....	40
2.2.9 Southern analysis of genomic DNA .....	40
2.3 Recombinant protein expression and purification .....	41
2.3.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) .....	41
2.3.2 Protein quantification .....	42
2.3.3 Recombinant protein expression .....	42
2.3.4 Immobilised metal ion affinity chromatography .....	43

## II

2.3.5	TEV cleavage of hexahistidine-tag.....	44
2.3.6	Anion-exchange chromatography.....	45
2.3.7	Gel filtration chromatography .....	45
2.3.8	Protein characterization by mass spectrometry .....	46
2.3.9	Protein characterization by Western blotting.....	46
2.4	Enzymatic and kinetic studies .....	47
2.4.1	<i>TbGSK3</i> biochemical characterization .....	47
2.4.2	Determination of the <i>TbGSK3</i> short $K_m$ for the substrates .....	47
2.4.3	Determination of mechanism of inhibition .....	48
2.4.4	Surface Plasmon Resonance .....	48
2.4.5	Isothermal Titration calorimetry.....	50
2.5	Screening .....	50
2.5.1	Compound library.....	50
2.5.2	Screening workflow.....	50
2.5.3	Primary screening assay – KinaseGlo format.....	51
2.5.4	Potency screening assay – Flashplate assay .....	52
2.5.5	Screening batch – pass criteria .....	53
2.5.6	Mammalian kinase profiling.....	53
2.5.7	Data analysis.....	54
2.5.8	Proliferation assay of MRC5 and BSF <i>T. brucei</i> cells.....	54
2.6	<i>T. brucei</i> cell culture and genetic manipulation.....	55
2.6.1	Bloodstream form <i>T. brucei</i> cell culture.....	55
2.6.2	Cell density.....	55
2.6.3	Generation of <i>TbGSK3</i> short single knock-outs cell lines.....	56
2.6.4	Generation of transgenic <i>T. brucei</i> cell lines .....	56
2.6.5	Cloning of genetic modified <i>T. brucei</i> lines .....	57
2.6.6	Growth inhibition studies of <i>T. brucei</i> .....	58
2.6.7	Cell lysis for Western blotting.....	59
2.7	Chemical-proteomics.....	59
2.7.1	Kinobeads method .....	59
2.8	Molecular modelling.....	60
2.8.1	Homology modelling.....	60
2.8.2	Generation of binding modes .....	60
Chapter 3	Results.....	61
3.1	Identification of small molecules inhibitors .....	62
3.1.1	Biochemical characterization of <i>TbGSK3</i> .....	62
3.1.2	Primary Screen of <i>TbGSK3</i> short.....	63
3.1.3	Potency Screen of <i>TbGSK3</i> short.....	64
3.2	Hit validation .....	65
3.2.1	<i>TbGSK3</i> short series classification .....	65
3.2.2	GSK3 07 series: correlation, specificity and selectivity .....	66
3.2.3	GSK3 07 series: mode of inhibition and $K_i$ characterization .....	68

### III

3.2.4	GSK3 09 series: correlation, specificity and selectivity .....	69
3.2.5	GSK3 09 series: Mode of inhibition and $K_i$ determination.....	71
3.2.6	Mode of binding of series GSK3 07 and GSK3 09 .....	71
3.2.7	DDD00085893 is a cidal inhibitor of <i>T. brucei</i> .....	73
3.3	Target identification by chemical proteomic approach.....	74
3.3.1	Kinobeads approach .....	74
3.4	Protein expression and purification .....	77
3.4.1	Protein expression and characterization of BAP-tagged <i>TbGSK3</i> short .....	77
3.4.2	Protein expression and characterization of <i>TbGSK3</i> long.....	78
3.4.3	Western blot analysis of recombinant <i>TbGSK3</i> short and <i>TbGSK3</i> long.....	81
3.4.4	Protein expression and purification of <i>TbGSK3</i> short inactive mutant .....	82
Chapter 4	Chemical-genetic validation of <i>TbGSK3</i> short .....	84
4.1	Chemical validation using genetic approach .....	85
4.1.1	Rationale of the chemical-genetic approach.....	85
4.1.2	Null mutants of <i>TbGSK3</i> short have a lethal phenotype .....	86
4.1.3	Generation of conditional cell lines and tuning of the ectopic expression .....	87
4.1.4	Over-expression of <i>TbGSK3</i> short is toxic for cell viability .....	89
4.1.5	Expression of <i>TbGSK3</i> short and long in BSF <i>T. b. brucei</i> .....	89
4.1.6	Chemical validation of <i>TbGSK3</i> short as drug target.....	90
4.1.7	Over-expression of inactive <i>TbGSK3</i> short in bloodstream form .....	92
4.1.8	Over-expression of <i>TbGSK3</i> long in bloodstream form <i>T. brucei</i> .....	93
4.2	Genetic validation of <i>TbGSK3</i> short .....	95
4.2.1	Generation of conditional double knock-out using pLew100v5.....	95
4.2.2	Generation of conditional double knock-out using pLew100.....	96
Chapter 5	Discussion .....	98
5.1	Key findings .....	99
5.2	<i>TbGSK3</i> short is a druggable target in African trypanosomes .....	100
5.3	<i>TbGSK3</i> short is an essential gene .....	104
5.4	Alternative strategies of inhibition to obtain selectivity and efficacy.....	105
5.4.1	Allosteric inhibitors .....	106
5.4.2	Irreversible inhibition .....	108
5.4.3	Substrate inhibitors.....	110
5.4.4	Alternative screening strategies.....	112
5.5	Generation of a chemical tool to dissect new biology .....	113
5.6	Conclusions .....	114
	List of references.....	116
	Appendix.....	139

## List of Figures

	Facing page
Figure 1.1 Classification of the endemic countries for HAT according to the reported cases in 2009.....	2
Figure 1.2 Comparison between reported number of cases of <i>T. b. gambiense</i> and screen of the population.....	2
Figure 1.3 Number of cases reported of <i>T. b. gambiense</i> and <i>T. b. rhodesiense</i> in the period 2000-2009.....	3
Figure 1.4 Drug treatments for the second stage of <i>T. b. gambiense</i> infection in the period 2001-2010.....	3
Figure 1.5 Life cycle of African trypanosome in the vector and in the human host.....	5
Figure 1.6 Trypanosome cell architecture.....	6
Figure 1.7 Generic structure of a protein kinase.....	18
Figure 1.8 Amino acid sequence alignment of <i>HsGSK3<math>\beta</math></i> , <i>HsGSK3<math>\alpha</math></i> , <i>TbGSK3short</i> , and <i>TbGSK3long</i> .....	30
Figure 2.1 pET-15b_TEV_BAP cloning/expression region.....	42
Figure 2.2 KinaseGlo assay principle.....	51
Figure 2.3 Schematic representing cloning steps in the construction of knock-out cassettes.....	56
Figure 3.1 Determination of the kinetic parameters for <i>TbGSK3 short</i> .....	62
Figure 3.2 GW8510 standard inhibitor IC <sub>50</sub> determination.....	64
Figure 3.3 Primary screen.....	65
Figure 3.4 Potency screen.....	66
Figure 3.5 Compound series identified from focused screening and testing of literature compounds.....	66
Figure 3.6 Correlation plot for GSK3 07 compounds.....	67
Figure 3.7 Mechanism of inhibition was determined for (A) DDD00088338 and (B) DDD00088340.....	68
Figure 3.8 Correlation plot between inhibition of <i>TbGSK3 short</i> and inhibition of <i>T. b. brucei</i> cell growth for GSK3 09 series.....	70
Figure 3.9 Mechanism of inhibition of DDD00085893 towards ATP.....	71
Figure 3.10 Mechanism of inhibition of DDD00085893 towards GSP2 substrate .....	71
Figure 3.11 Superposition of the <i>HsGSK3</i> crystal structure 1R0E with the homology model of <i>TbGSK3 short</i> .....	72
Figure 3.12 Proposed binding modes for (A) DDD00065658, (B) DDD00085893, and (C) DDD00101234 in the homology model of <i>TbGSK3 short</i> .....	73
Figure 3.13 DDD00085893 inhibition of <i>TbGSK3 short</i> activity and rate of killing of BSF <i>T. b. brucei</i> cell proliferation.....	73
Figure 3.14 Proteomic profiling of DDD00085893 in BSF <i>T. b. brucei</i> and MRC5 lysates by kinobeads competition assay.....	75
Figure 3.15 ClustalW alignment of <i>TbGSK3 short</i> and <i>TbGSK3 long</i> .....	77

Figure 3.16 Expression and purification of recombinant <i>TbGSK3</i> short tagged at the N-terminus with His6-BAP tag.....	77
Figure 3.17 Size-exclusion purification of recombinant <i>TbGSK3</i> short tagged at the N-terminus with BAP tag.....	77
Figure 3.18 Characterization of recombinant <i>TbGSK3</i> short tagged at the N-terminus with BAP peptide.....	78
Figure 3.19 Expression and purification of recombinant <i>TbGSK3</i> long tagged at the N-terminus with His6-BAP tag.....	79
Figure 3.20 Expression and purification of recombinant <i>TbGSK3</i> long tagged at the N-terminus with BAP peptide in <i>E. coli</i> .....	80
Figure 3.21 Isothermal titration calorimetry of DDD00085893 with <i>TbGSK3</i> short and <i>TbGSK3</i> long.....	80
Figure 3.22 Western blot analysis of recombinant BAP- <i>TbGSK3</i> short and long.....	81
Figure 3.23 Expression and purification of HIS6-BAP- <i>TbGSK3</i> (K49A) in <i>E. coli</i> and western blot analysis.....	82
Figure 3.24 Purification of recombinant HIS6-BAP- <i>TbGSK3</i> (K49A) in <i>E. coli</i> .....	82
Figure 3.25 Purification of recombinant HIS6-BAP- <i>TbGSK3</i> (K49M) in <i>E. coli</i> .....	83
Figure 4.1 Generation of <i>TbGSK3</i> single knock-out (SKO) cell lines.....	86
Figure 4.2 Strategy for inducible expression of an ectopic copy.....	87
Figure 4.3 Genotypic analyses of conditional cell lines.....	88
Figure 4.4 Cumulative growth analyses of wild-type and genetically modified <i>TbGSK3</i> short cell lines.....	89
Figure 4.5 Western blot analysis of BSF <i>T. b. brucei</i> .....	90
Figure 4.6 Targeting of <i>TbGSK3</i> short by DDD00085893 in bloodstream form <i>T. b. brucei</i> .....	90
Figure 4.7 Western blot analysis of mutant cell lines for <i>TbGSK3</i> short.....	91
Figure 4.8 Genotypic analysis of conditional over-expressing <i>TbGSK3</i> (K49A) kinase dead cell lines.....	92
Figure 4.9 Cumulative growth of <i>T. brucei</i> WT and conditional over-expressing <i>TbGSK3</i> (K49A) kinase dead cell lines.....	92
Figure 4.10 DDD00085893 sensitivity towards conditional over-expressing <i>TbGSK3</i> (K49A) kinase dead cell lines.....	92
Figure 4.11 Western Blot analysis of <i>T. brucei</i> conditional over-expressing <i>TbGSK3</i> (K49A) kinase dead mutant cell lines.....	93
Figure 4.12 DDD00085893 sensitivity towards conditional over-expressing <i>TbGSK3</i> long cell line.....	94
Figure 4.13 Western Blot analysis of <i>T. brucei</i> conditional over-expressing <i>TbGSK3</i> long cell lines.....	94
Figure 4.14 Southern blot analysis of conditional double knock-out cell lines obtained using the conditional over-expressor pLew100v5 plasmid.....	95
Figure 4.15 Cumulative growth of <i>T. brucei</i> WT and conditional null mutant cell line.....	95

Figure 4.16 DDD00085893 sensitivity towards conditional null mutant (cDKO) cell line	96
Figure 4.17 Southern blot analysis of pLew100 conditional WT (cWT) and conditional single knock-out (cSKO) cell lines	96
Figure 4.18 Cumulative growth of <i>T. brucei</i> WT, conditional WT and conditional single knock-out cell lines obtained using pLew100 as rescue vector	96
Figure 4.19 DDD00085893 sensitivity towards pLew100 conditional WT and conditional single knock-out cell lines	96
Figure 4.20 Southern blot analysis of cDKO cell lines obtained using pLew100 as rescue vector	97

## List of Tables

	Facing page
Table 1.1 Drugs used for the treatment of human African trypanosomiasis	9
Table 1.2 Classification of PKs by groups in human and in <i>Trypanosoma brucei</i>	19
Table 1.3 <i>T. b. brucei</i> protein kinases genetically and chemically investigated	24
Table 2.1 Primers used for recombinant protein expression in <i>E. coli</i>	37
Table 2.2 Primers used for generation of knock-outs and rescue constructs	38
Table 2.3 Protein expression and purification conditions	43
Table 2.4 Primary and secondary antibodies used for Western blotting analysis	46
Table 2.5 Lists of mutant cell lines generated in BSF <i>T. brucei</i>	57
Table 3.1 Determination of $K_m$ for ATP and GSP2	62
Table 3.2 <i>TbGSK3</i> short assay conditions and screening statistics	64
Table 3.3 Ranking of the kinase set compound potencies	66
Table 3.4 GSK3 07 Selection of early hits with activities against <i>TbGSK3</i> short, <i>HsGSK3<math>\beta</math></i> , <i>HsCDK2</i> , BSF <i>T. b. brucei</i> and MRC5 cells	67
Table 3.5 GSK3 07 Selection of late hits with activities against <i>TbGSK3</i> short, <i>HsGSK3<math>\beta</math></i> , <i>HsCDK2/cyclin A</i> , BSF <i>T. b. brucei</i> and MRC5 cells	67
Table 3.6 Mammalian kinase selectivity panel for series GSK3 07	68
Table 3.7 GSK3 09 Selection of R1 substituted amynopyrazoles with activities against <i>TbGSK3</i> short, <i>HsGSK3<math>\beta</math></i> , <i>HsCDK2/cyclin A</i> , BSF <i>T. b. brucei</i> and MRC5 cells	69
Table 3.8 Mammalian kinase selectivity panel for series GSK3 09	69
Table 3.9 GSK3 09 Selection of R2 substituted amynopyrazoles with activities against <i>TbGSK3</i> short, <i>HsGSK3<math>\beta</math></i> , <i>HsCDK2/cyclin A</i> , BSF <i>T. b. brucei</i> and MRC5 cells	70
Table 3.10 Differences in the ATP binding pockets of <i>TbGSK3</i> , <i>HsGSK3<math>\beta</math></i> and <i>HsCDK2</i>	72
Table 3.11 Determination of DDD00085893 $IC_{50}$ , for <i>TbGSK3</i> short and $EC_{50}$ , $EC_{90}$ , and $EC_{99}$ for BSF <i>T. brucei</i>	74
Table 3.12 Kinobeads competition assay for the BSF <i>T. b. brucei</i> and MRC5 targets of DDD00085893	76
Table 4.1 Sensitivity to DDD00085893 of WT, SKO, <i>TbGSK3</i> conditional over-expressing (cWT <sup>OE</sup> ) and conditional SKO (cSKO <sup>OE</sup> ) cell lines	90



## VII

Table 4.2 Sensitivity of DDD00085893 towards cell lines conditional over-expressing <i>TbGSK3(K49A)</i> kinase dead mutant.....	92
Table 4.3 Sensitivity of DDD00085893 towards conditional WT (cWT) and conditional SKO (cSKO) cell lines.....	96

**List of abbreviations**

AdoMetC	Adenosylmethionine decarboxylase
AQP	Aquaglyceroporin
BAP	Biotin acceptor peptide
BSF	Bloodstream form
CATT	Card agglutination test
CNS	Central nervous system
CSF	Cerebrospinal fluid
DDU	Drug Discovery Unit
DMSO	Dimethyl sulfoxide
DFMO	D,L- $\alpha$ -difluoromethylornithine or eflornihitine
DNDi	Drug for Neglected Disease initiative
ePKs	Eukaryotic protein kinases
GSK3	Glycogen Synthase Kinase 3
HAPT1	High affinity pentamidine transporter 1
HAT	Human African trypanosomiasis
His <sub>6</sub>	Hexahistidine tag
HRP	Horseradish peroxidase
<i>Hs</i> GSK3	Human Glycogen Synthase kinase 3
HYG	Hygromycin phosphotransferase
IEX	Ion exchange chromatography
IM	Intramuscular
IMAC	Metal-ion affinity chromatography
IPI	International protein index
ISG75	Invariant surface glycoprotein 75

iTRAQ	Isobaric tags for relative and absolute quantitation
IV	Intravenous
LB	Luria Bertani
LDL	Low-density lipoprotein
LPT1	Low-affinity pentamidine transporter 1
MBP	Maltose binding protein
Mel T	Melarsoprol–trypanothione adduct
MRPA	Multidrug resistance protein A
MSF	Médecins Sans Frontières
MTS	Medium throughput screening
MW	Molecular weight
NECT	Nifurtimox-eflornithine combination therapy
NEO	G418-neomicin
NGO	Non-governmental organization
NTR	Nitroreductase
ODC	Ornithine decarboxylase
ORF	Open reading frame
PAC	Puromycin acetyltransferase
PAD	Proteins associated with differentiation
PARP	Procyclic acidic repetitive protein
PCF	Procyclic form
PFR	Paraflagellar rod
QC	Quality control
S427	Strain 427
S927	Strain 927

SAR	Structure-activity relationship
SD	Standard deviation
SEC	Size exclusion chromatography
SEM	Standard error of the mean
SIF	Stumpy derived factor
SN	Signal to noise
SNP	Single nucleotide polymorphism
SPR	Surface plasmon resonance
<i>T. b.</i>	<i>Trypanosoma brucei</i>
<i>T. brucei</i>	<i>Trypanosoma brucei</i>
T7RNAP	T7 RNA polymerase
<i>TbGSK3</i>	<i>T. brucei</i> Glycogen Synthase kinase 3
Tet	Tetracycline
TetR	Tetracycline repressor protein (TetR)
TEV	Tobacco Etch Virus
TPP	Target Product Profile
Trityps	<i>Trypanosoma brucei</i> , <i>Trypanosoma cruzi</i> , and <i>Leishmania major</i>
UTR	Untranslated region
VSG	Variant surface glycoproteins
WHO	World Health Organization
WT	Wild-type
Z'	Zeta prime

## **Acknowledgments**

I am deeply grateful to Professor Alan Fairlamb and Professor Julie Frearson for their supervision and for the opportunity offered me to enrol in a part-time PhD program. I am also grateful to my boss Dr David Gray and to Dr Anthony Hope for their support.

I am grateful to the colleagues that have contributed to my work, in particular, Dr Michael Urbaniak and the chemists Dr Andrew Woodland, Dr Robert Urich, and Dr Torsten Luksch.

I have to say a big thank to Dr Natasha Sienkiewicz for teaching me molecular biology and cell biology and sharing the office in the AHF lab. A big thank also to Dr Scott Cameron who has introduced me to protein production and purification. I am also very grateful to all the previous and present members of the AHF lab, BCDD and the DDU, every single one has taught me something.

A big thank you to all the support staff and technical services for hugely contributing to make the College of Life Sciences a great place to work!

Last but not least I have to thank my Scottish-Turkish family here in Dundee and my family in Italy for all their love and patience.

This work was supported by grants to Professor Julie Frearson from the Wellcome Trust.

**Declaration**

I hereby certify that this thesis is of my own composition and is based upon the results of my own work, carried under the supervision of Professor Alan H. Fairlamb in the College of Life Sciences, University of Dundee. Work other than my own is clearly specified in the text by reference to the relevant researchers or to their publications. I have consulted all references cited herein, unless otherwise stated. No part of this thesis has been previously submitted for a higher degree.



Raffaella Grimaldi

I certify that Raffaella Grimaldi has performed the research described in this thesis under my supervision and has fulfilled the conditions of the relevant ordinance and regulations of the University of Dundee, and that she is qualified to submit this thesis for the degree of Doctor of Philosophy.



Professor Alan H. Fairlamb

Wellcome Principal Research Fellow

## Abstract

Human African Trypanosomiasis (HAT), caused by *Trypanosoma brucei* subspecies, is one of the most neglected diseases: available treatments are old, toxic, and difficult to administer; they are not efficacious against all parasite species or disease stages and drug resistance is an increasing problem.

Protein kinases are well validated drug targets for a variety of human diseases with many inhibitors under development or in the clinic. The *T. brucei* kinome has been annotated and there is evidence of essentiality of some of the members of this family. This thesis aims at evaluating the essentiality of Glycogen Synthase Kinase 3 (*TbGSK3* short; Tb927.10.13780) and chemically validating it as a potential drug target in *T. brucei*.

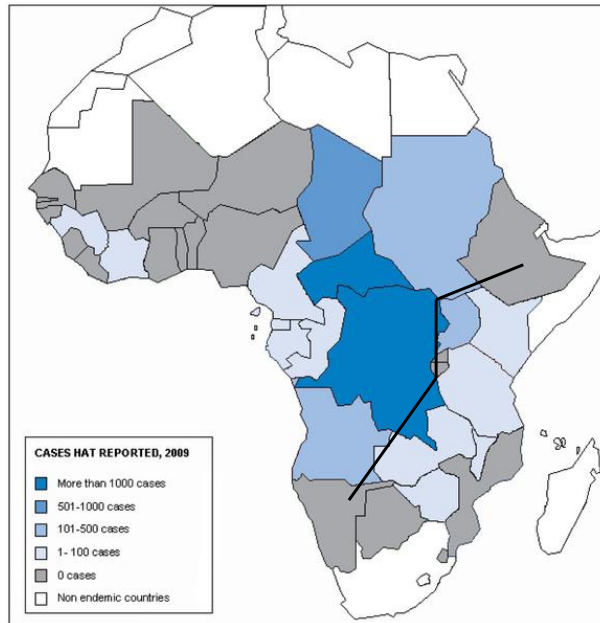
*TbGSK3* recombinant protein was biochemically characterised and screened against a focussed kinase library using the KinaseGlo assay method. Further repurchase and synthesis of novel compounds yielded 10 validated chemical series against *TbGSK3* short. In particular two series showed anti-proliferative activity against the parasite. GSK3 07 series was further investigated by the Drug Discovery Unit with a phenotypic approach for its off-target effects, and GSK3 09 series was further validated to act “on target”. The latter series showed a good correlation between biochemical potency and cellular efficacy. Using a combination of chemical and genetic approaches *TbGSK3* short was demonstrated to be specifically targeted by a GSK3 09 tool molecule in *T. brucei* lysates. Furthermore, the *in vitro* efficacy in trypanosomes could be reverted by target over-expression. Further validation of its activity “on target” was given by its ability to modulate the cell toxicity caused by *TbGSK3* short over-expression.

The genetic validation of *TbGSK3* short by generation of conditional null mutants was not possible due to the tight regulation of the protein levels and the cell toxicity associated with protein over-expression.

The validated *TbGSK3* short chemical tool could be used to elucidate the functions of *TbGSK3* short in *T. brucei*, identify its substrates and increase the chance to solve the crystal structure of this enzyme for the design of novel inhibitors with different mechanism of inhibition and/or increased selectivity.

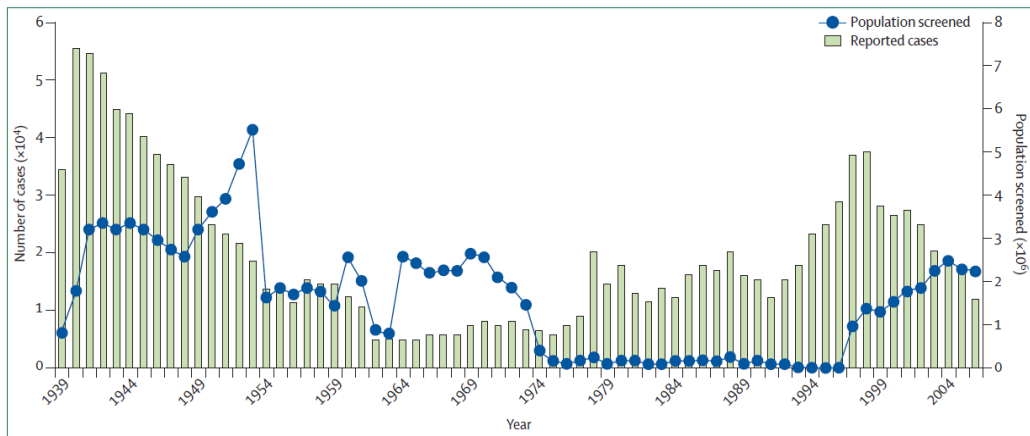
# **Chapter 1 Introduction**





**Figure 1.1 Classification of the endemic countries for HAT according to the reported cases in 2009**

The dark line separates the Western endemic countries for *T. b. gambiense* from the Eastern endemic countries for *T. b. rhodesiense* (figure adapted from Simarro *et al.*, 2011).



**Figure 1.2 Comparison between reported number of cases of *T. b. gambiense* and screen of the population**

Number of cases are reported as green columns and the population screened as blue circles over the period 1939–2004 in Africa (Figure from Brun *et al.*, 2010).

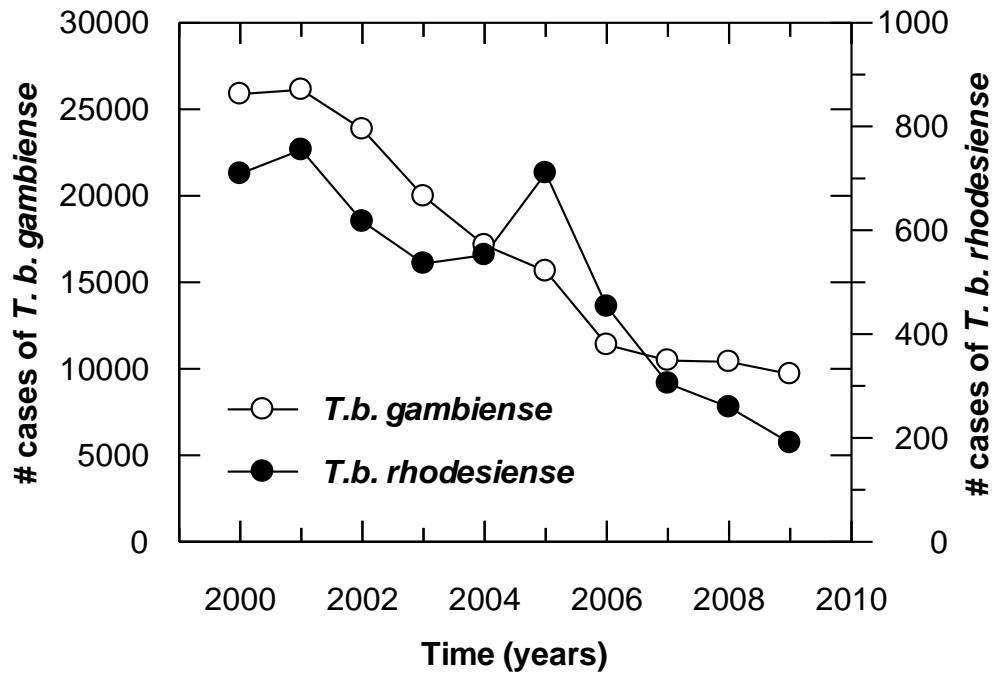
## 1.1 Human African trypanosomiasis

Human African trypanosomiasis (HAT, or also known as sleeping sickness) is one of the most neglected parasitic diseases. If untreated, this disease is invariably lethal. Endemic in Sub-Saharan Africa, it puts at risk the lives of approximately 70 million of people distributed over a territory that covers 36 countries (Simarro *et al.*, 2010). It is caused by infection with the unicellular parasite *Trypanosoma brucei* (*T. brucei* or *T.b.*) and transmitted by the bite of tse-tse fly of the genus *Glossina*.

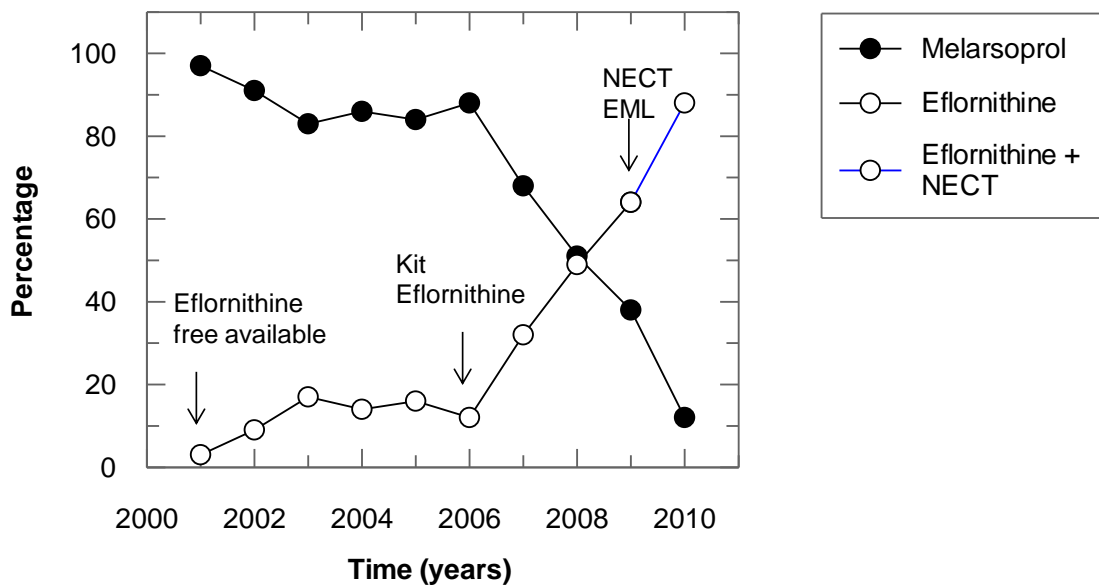
There are three sub-species of *T. brucei* causing Trypanosomiasis in humans and animals: *T. b. gambiense*, which causes the human chronic form of the disease in West and Central Africa and represents 95% of the cases; *T. b. rhodesiense*, which is responsible for the human acute form in East Africa in the remaining 5% of the cases; and *T. b. brucei*, which infects primarily cattle causing the animal form of the disease, called Nagana (Brun *et al.*, 2010) (**Figure 1.1**).

### 1.1.1 Epidemiology and distribution

Human African Trypanosomiasis reached an estimated 800,000 deaths between 1896 and 1906; between 1920 and 1940 there was a second epidemic, but at that time the colonial control program almost achieved the elimination of the disease by active screening and treatment of the population. After independence from colonialism, the political instability associated with wars and the lack of surveillance under the local health-systems caused the spread of a new epidemic in 1990 (Brun *et al.*, 2010). The correlation between active control of the disease and reduction of number of cases is well established, as well as the reappearance when the surveillance is reduced (Brun *et al.*, 2010) (**Figure 1.2**).



**Figure 1.3** Number of cases reported of *T. b. gambiense* and *T. b. rhodesiense* in the period 2000 – 2009 (Figure adapted from Simarro *et al.*, 2010).



**Figure 1.4** Drug treatments for the second stage of *T. b. gambiense* infection in the period 2001-2010 NECT-EML: introduction of the nifurtimox-eflornithine combination therapy (NECT) in WHO list of the essential medicines (EML) (Figure adapted from Simarro *et al.*, 2012)

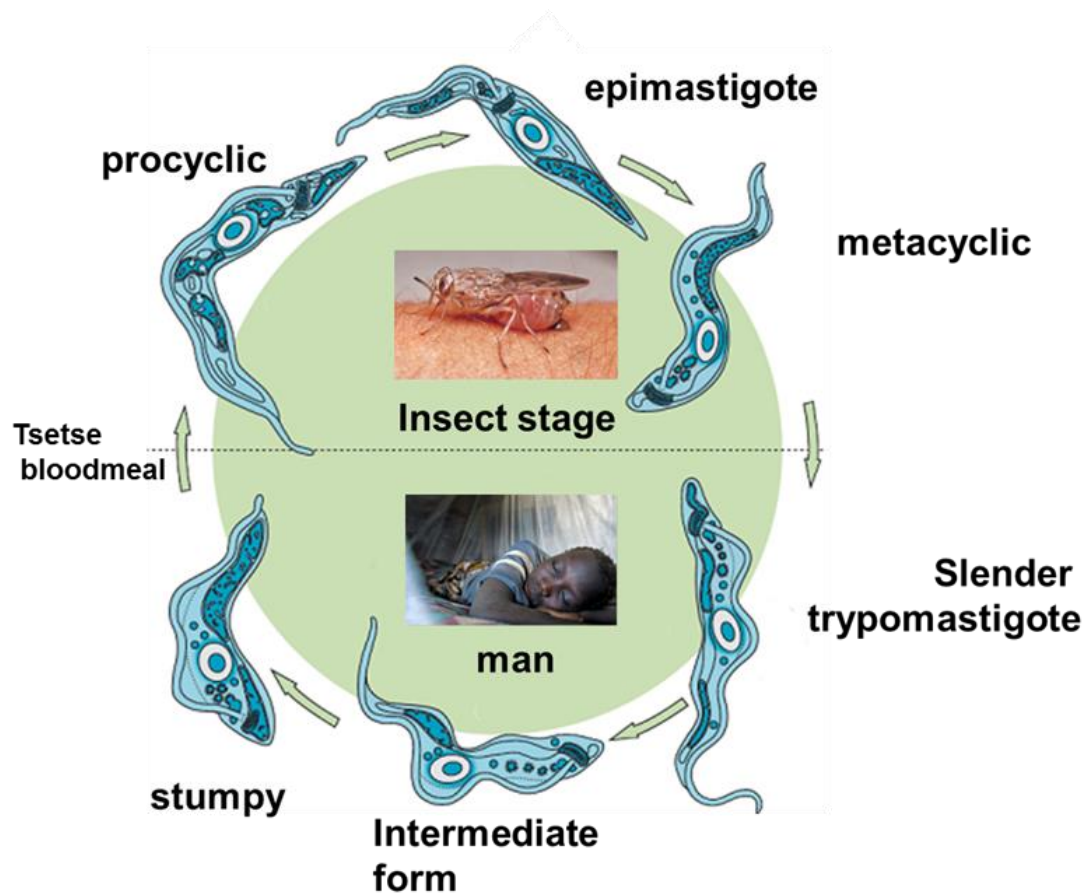
In 1997 a resolution from the World Health Organization (WHO) advocated interventions for HAT, increased the funding and the control activities, including active screen of the population and access to treatment tools, which resulted in a reported number of cases below 10,000 in 2009 (Simarro *et al.*, 2010) (**Figure 1.3**). Despite the efforts made to compile an atlas of HAT, which associates foci to number of cases, it is likely that at least two thirds of the cases are not reported. HAT is endemic in 36 African countries, but in 2009, nineteen disease-endemic countries reported no cases (Simarro *et al.*, 2010). The Central African Republic and the Democratic Republic of the Congo are the most affected countries, reporting more than 1,000 cases annually; Chad more than 500 cases; and Angola, Sudan and Uganda more than 100 cases. The remaining eleven countries have reported less than 100 cases (Simarro *et al.*, 2010) (**Figure 1.1**). The number of cases reported annually is only a fraction of the real number of infected people and the WHO estimate of the number of cases in 2006 was between 50,000 -70,000 (WHO, 2010).

HAT is considered a neglected disease because it affects poor people living in rural areas in Sub-Saharan Africa. Furthermore, the drugs available to cure it are toxic, difficult to administer in poor settings and inefficacious for all forms and stages of the disease. Also resistance in the field is becoming an increasing problem, making the already poor treatments available even less efficacious (Fairlamb, 2003). Yet, there is very little economic incentive in developing new drugs as the disease affects marginalised people with very little political voice. This creates a very vicious cycle between poverty and disease aggravated by the fact that the animal form of the disease limits the availability of meat and milk and reduces the income of these already poor households. The FAO (Food and Agriculture Organization) estimates the annual loss in income associated to HAT to be around US\$ 1.5 billion (WHO, 2010).

Although the numbers of cases have reduced in the recent years due to the joint effort of the WHO, non-governmental organizations (NGOs) such as Médecins Sans Frontières (MSF) and local governments together with the introduction of a new combined therapy (NECT) as first line treatment included in the WHO list of Essential Medicines (**Figure 1.4**), this disease still threatens the economic and human development of a large part of the African continent (Simarro *et al.*, 2012). Reducing the attention and focus now could cause as in the past an increase in number of cases; on the other hand the reduction in number of cases opens the opportunity for the elimination of the disease, considering also that two new molecules have recently reached clinical development (section 1.2.3). However, the path to the clinical use of new molecules is extremely hard and characterised by high rate of failure (Brown and Superti-Furga *et al.*, 2003); therefore it will be a mistake at this time to stop the development of new drugs for HAT.

### **1.1.2 Trypanosome cell cycle in vector and in human host**

*Trypanosoma brucei* belongs to the order of Kinetoplastida. Species of this order are unicellular flagellated protozoa, characterised by a single large mitochondrion, which contains compacted DNA called the “kinetoplast” (Stuart *et al.*, 2008). Trypanosomes are transmitted by the bite of flies of the genus *Glossina*, where both male and female flies are obligate blood-feeders and can transmit the disease. Around 33 species of tse tse flies exist that can be sub-divided in three groups (*Morsitans*, *Palpalis* and *Fusca*) according to the habitat in which they live (savannah, river and forest) (Gooding and Krafsur, 2005). Climate and demographic changes have caused a redistribution of the vectors with consequent shift of the foci of the disease from savannah to forest areas and in some cases foci have been found in populated urban centres (Malvy and Chappuis, 2011).

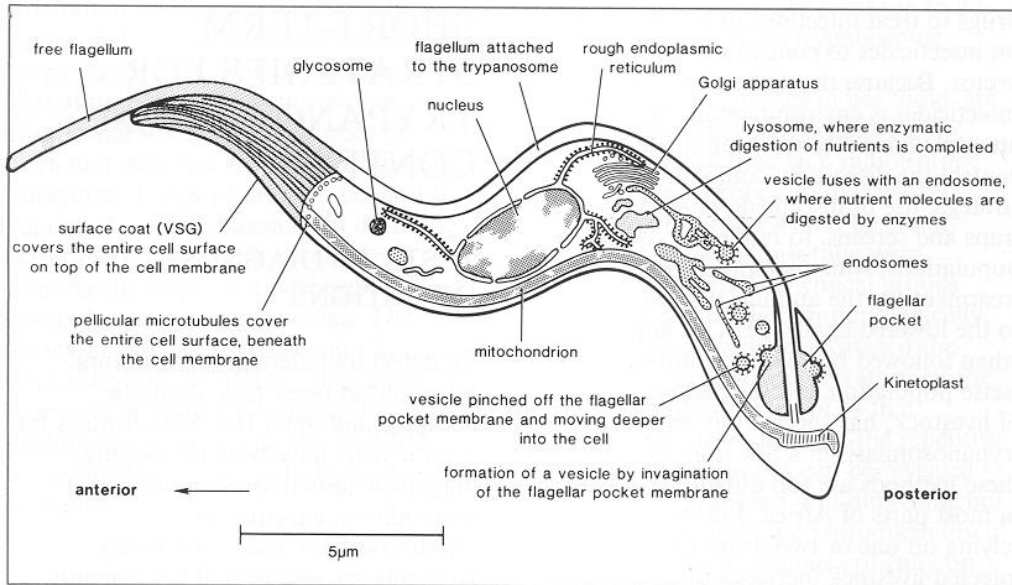


**Figure 1.5 Life cycle of African trypanosome in the vector and in the human host**  
 (Figure adapted from Vickerman, 1985 and Barrett *et al.*, 2007; pictures from WHO website, [www.who.int](http://www.who.int)).

The two major changes between the mammalian and the insect developmental cycle of African trypanosomes are at the level of mitochondrion and the surface membrane (**Figure 1.5**). In mammals the trypanosomes rely on glycolysis for ATP production because of the availability of glucose in the host fluids and are coated with variant surface glycoproteins (VSG) that protect the parasite from the host immune system by antigenic variation (Vickerman, 1985; Pays, 2005; Horn and McCulloch, 2010). Only one VSG gene is expressed at a time from among 10-20 telomeric expression sites, in addition to around 1000 silent VSGs and VSG pseudogenes in subtelomeric regions (Horn and McCulloch, 2010). This remarkable capacity for antigenic variation makes the development of a vaccine for HAT difficult, relying exclusively on control activities and chemotherapeutics for surveillance of the disease.

During a blood meal the fly deposits infective metacyclic trypanosomes in the dermal connective tissue, where they divide to form a chancre, and then move to the lymph nodes and into the bloodstream. Parasites can then multiply by mitosis as slender trypomastigotes. The parasitemia fluctuates according to the host immune response directed against the VSG on the plasma membrane of trypanosome, with increases in IgG and IgM and reduction in parasite numbers till the antigenic variation makes the immune response inadequate (Vickerman, 1985, Paulnock and Freeman, 2010). As parasite numbers increase, a parasite-derived soluble factor, stumpy derived factor (SDF) promotes cell cycle arrest in G0/G1 and generation of stumpy forms (Fenn and Matthews, 2007). The non-dividing stumpy forms are taken-up by the flies during the blood meal and they continue their life cycle in the vector (Vickerman, 1985).

In the midgut of the vector the stumpy transforms into an elongated parasite, the proliferative procyclic form (Vickerman, 1985). *In vitro*, the transition from stumpy to procyclic can be obtained by a reduction in temperature and exposure to citrate/cis-aconitate, both compatible with a drop in temperature during the night when the fly



**Figure 1.6 Trypanosome cell architecture**

(Figure from ILRI website (International Livestock Research Institute),  
[www.ilri.org](http://www.ilri.org))



tends to feed and with the citrate concentration in the fly midgut (Fenn and Matthews, 2007). A family of transporters called PAD (proteins associated with differentiation), expressed on the surface of the stumpy form, is required to sense the citrate/cis-aconitate signal (Dean *et al.*, 2009). Once in the vector the insect form of the parasite exchanges the variable antigen coat for procyclins and relies upon proline metabolism and oxidative phosphorylation as principal sources of energy (Fenn and Matthews, 2007; Vickerman, 1985). The differentiation of the insect form of the parasite into an infective form takes between 3 and 5 weeks, during which the parasites undergo complex steps of differentiation and migrate to the salivary glands. Once in the salivary glands the procyclics transform into epimastigotes, which are characterised by a prenuclear kinetoplast. Finally, epimastigotes develop into the infective metacyclic forms, which are injected in the host during the blood meal. The metacyclic forms regain some of the characteristics of the forms infecting humans with a reduction in size of the mitochondrion and the re-appearance of variable antigen coat (Vickerman, 1985) (**Figure 1.5**).

### **1.1.3 Trypanosome cell architecture**

The trypanosome cell has an elongated and polarized form (**Figure 1.6**) (Matthews, 2005). During the life cycle the morphology of the cells changes with respect to the position of the flagellum, nucleus and kinetoplast, size of mitochondrion, efficiency of endocytosis, and protein expression of the parasite coat. In the slender trypomastigote and stumpy form the kinetoplast is positioned at the posterior end of the parasite and is connected through the mitochondrial membrane to the basal body of the flagellum. At the posterior end is also located the flagellar pocket that is the only site of exo- and endocytosis. The antigenic variation, occurring in the mammalian form of the parasite, requires that the full VSG coat is renewed every 12 mins in order to assure

evasion of the immune response (Overath and Englster, 2004). In the bloodstream trypomastigote the mitochondrion is an elongated organelle positioned all long the cell body from the posterior to the anterior end. The absence of an active oxidative respiration in bloodstream forms explains the lack of internal cristae. The mobility of trypanosomes depends upon a single flagellum that is anchored to the basal body at the posterior end of the cell and is attached to the cell body along its length. The flagellum has a typical “9+2” microtubule axoneme, and is connected to the cell through the paraflagellar rod (PFR) and a set of four microtubules that constitute the flagellum attachment zone (FAZ) (Vaughan and Gull, 2003). During proliferation, the new flagellum is replicated beside the old one, starting from the duplication of the basal body and kinetoplast. Once the duplication of the new flagellum is completed the separation starts from the anterior tip. The process of duplication and separation requires coordination of the replication and regulation of positioning of single organelles (flagellar pocket, flagellum, mitochondrion, kinetoplast and nucleus). The major changes in the insect procyclic and epimastigote cells is the positioning of the kinetoplast to a more anterior position that results in a progressively longer anterior flagellum that might help attachment to the salivary gland; the replacement of the VSG coat with the procyclin coat; the increase of mitochondrial activity and the reduction of the endocytic rate (Matthews, 2005).

#### **1.1.4 Symptoms and diagnosis**

After infection the disease evolves in two different stages. The initial haemolymphatic stage, when the parasite is present in the bloodstream and lymphatic system, is characterised by generic symptoms: fever, headache, lymphadenopathies, and oedema and in some cases splenomegaly or hepatomegaly (Malvy and Chappuis, 2011). The disease progresses to the late meningo-encephalitic stage when the parasites

penetrate the central nervous system (CNS). The late stage of the disease is characterised by far more severe symptoms: disturbed sleep patterns, confusion, tremor, general motor weakness, hemiparesis, abnormal movements and neuro-psychiatric disorders. In the terminal phase of the disease the patient is in an unconscious state and if untreated the disease usually results in death (Malvy and Chappuis, 2011). The progression of the disease caused by *T. b. gambiense* is chronic and takes on average three years. In contrast *T. b. rhodesiense* causes an acute form of the disease that progress rapidly to the late stage in weeks or months (Brun *et al.*, 2010). The diagnostic test available for *T. b. gambiense* is the card agglutination test (CATT) (Magnus *et al.*, 1978), a serological test with 87-98% sensitivity and 93-95% specificity (Brun *et al.*, 2010). A positive test has to be corroborated by a parasitological confirmation of the presence of the trypanosomes in lymph node aspirate and blood. The CATT test is sensitive only against *T. b. gambiense* and there are no available serological tests for *T. b. rhodesiense*. Thus, the diagnosis relies on the clinical symptoms and on the parasitological confirmation, which is easier for *T. b. rhodesiense* because of the higher number of circulating parasites (Brun *et al.*, 2010; Malvy and Chappuis, 2011).

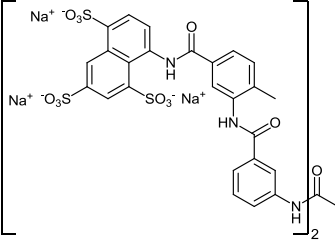
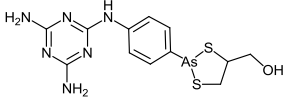
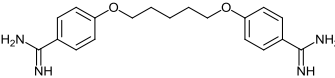
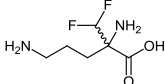
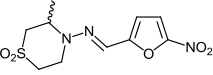
Because treatment differs between stages of disease, the staging of the disease is compulsory after parasitological confirmation. The staging is determined by detection of trypanosomes and white blood cells in the cerebrospinal fluid (CSF) (>5 per  $\mu$ l) (Brun *et al.*, 2010; Malvy and Chappuis, 2011).

## **1.2 Chemotherapy of sleeping sickness**

### **1.2.1 *T. b. rhodesiense* infection**

**Suramin** (introduced in the early 1920s), a polyanionic sulphonated naphthylamine, is used to treat first-stage *T. b. rhodesiense* infection. This molecule

**Table 1.1 Drugs used for the treatment of human African trypanosomiasis**

Causitive agent	Disease stage	Drug (date introduced)	Chemical structure	Treatment regimen adult doses
<i>T. b. rhodesiense</i>	First stage	Suramin (1920s)		IV test dose 4-5 mg/kg on day 1 then 20 mg/kg IV weekly for 5 weeks
	Second stage	Melarsoprol (1949)		slow IV infusion 2.2 mg/kg/day for 10 days
<i>T. b. gambiense</i>	First stage	Pentamidine (1941)		IM or slow IV infusion 4 mg/kg/day for 7 days
	Second stage	Eflornithine (1990)		IV infusion 100 mg/kg every 6 hours for 14 days
	Second stage	Nifurtimox-eflornithine combination therapy, NECT (2009)		oral nifurtimox 3x5 mg/kg/day for 10 days + IV infusion 200 mg/kg every 12 h for 7 days

IV intravenous; IM intramuscular.

(Table adapted from Malvy and Chappuis, 2011)

does not penetrate the CNS, being highly charged, so it cannot be used for the second-stage of the disease (Fairlamb, 2003). The recommended dose regimen consists of an initial test dose of 4-5 mg/kg intravenous injection (IV) for the first day, and, in the absence of an acute reaction, the drug can be administered at the dose of 20 mg/kg IV weekly for 5 weeks (Malvy and Chappuis, 2011) (**Table 1.1**). The side effects are nephrotoxicity, peripheral neuropathy and bone marrow toxicity; generally their severity is mild and they revert once the treatment is completed. The mechanism of action of this drug is not completely understood, but the fact that resistance in the field has been rarely reported despite being used for many decades suggests that this drug might inhibit multiple targets (Barrett *et al.*, 2003; Fries and Fairlamb, 2003). It has also been suggested that suramin mode of action might involve inhibition of glycolytic enzymes (Fairlamb and Bowman, 1980) and the synergistic action with eflornithine (DFMO) might implicate a role in polyamine metabolism (Clarkson *et al.*, 1984). It is known that suramin enters the trypanosomes by endocytosis through the flagellar pocket, this process is receptor-mediated due to the tight binding to low-density lipoprotein (LDL) (Voogd *et al.*, 1993). Recently Alsford and colleagues (2012) have elucidated how suramin uptake is mediated by an invariant surface glycoprotein (ISG75) that is specifically expressed in bloodstream form. Then ISG75 trafficking delivers the drug to the lysosome where it is liberated by the action of lysosomal proteases and possibly delivered to the cytosol through a lysosomal transmembrane protein (major facilitator superfamily transporter). They also linked spermidine and N-acetylglucosamine biosynthesis to the mode of action of suramin.

**Melarsoprol** (introduced in 1949), a melaminophenyl-based organic arsenical, remains the only therapeutic option for the second-stage of *T. b. rhodesiense* infection despite the extreme toxicity. The dose regimen is 2.2 mg/kg/day IV for 10 days (Malvy and Chappuis, 2011) (**Table 1.1**). Five to 10% of patients develop reactive

encephalopathy, which causes the death of half of them. The active metabolite is melarsen oxide, which is rapidly converted in the blood (Keiser *et al.*, 2000). Melarsen oxide forms with trypanothione (bis-glutathionyl-spermidine) a stable and reversible adduct called MeIT, which depletes the intracellular pool of trypanothione and acts as competitive inhibitor of trypanothione reductase (Fairlamb *et al.*, 1989), an enzyme essential in maintaining the correct thiol-redox balance of trypanosomes. However, it has been argued that arsenicals interact more tightly with other thiols, including lipoic acid (Fairlamb *et al.*, 1992) and that when lysis occurs due to arsenicals only a small fraction of trypanothione is conjugated with the melarsen oxide, making unclear if trypanothione and the enzymes involved in its metabolism are the real targets for this drug (Denise and Barrett, 2001).

Melarsen oxide is transported in the cells by the P2 aminopurine transporter (Carter and Fairlamb, 1993), and mechanisms of resistance both *in vitro* and in the field are associated to loss of the P2 transporter (Barrett *et al.*, 2007). An ATP-binding cassette (ABC) transporter, MRPA (multidrug resistance protein A) is responsible for the efflux of MeIT from the cells, and over-expression of this transporter increases resistance to melarsoprol (Shahi *et al.*, 2002). Alsford and colleagues (2012) also found that aquaglyceroporins (AQPs) were responsible for the cross-resistance of melarsoprol and pentamidine. In particular *TbAQP2* has been further characterised as a high affinity melarsoprol and pentamidine transporter, previously reported as high affinity pentamidine transporter (HAPT1), and as a major determinant of cross-resistance to these drugs (Baker *et al.*, 2012, Munday *et al.*, 2014).

### 1.2.2 *T. b. gambiense* infection

**Pentamidine** (introduced in 1941), an aromatic diamidine, has been used as first-line treatment for the haemolymphatic stage of *T. b. gambiense* infection for over 60 years. Being positively charged at physiological pH this molecule has poor oral

availability and no CNS penetration (Fairlamb, 2003). The use of this drug is thus limited to the treatment of first stage of the disease, for which it is administered by intramuscular (IM) or IV injection with a schedule of 4 mg/kg/day per 7 days (Malvy and Chappuis, 2011) (**Table 1.1**). Toxicity includes pain at the injection site, nephrotoxicity, leucopenia and hypoglycaemia (Malvy and Chappuis, 2011).

The mechanism of action of pentamidine is not fully understood. Being a dication, pentamidine binds negatively charged molecules, as DNA, including the mitochondrial kinetoplast. However, the disruption of the kinetoplast DNA cannot account for all action of the drug as parasites that have lost the kinetoplast (a condition called dyskinetoplastidy) retain viability (Fairlamb, 2003). Other hypothetical mechanisms include effects on trans-splicing and RNA editing in trypanosomes; in fact pentamidine has been reported to inhibit group I intron catalytic activity in *Candida albicans* and *Pneumocystis carinii* (Zhang *et al.*, 2002; Liu *et al.*, 1994), but this hypothesis has not been further investigated. Also the enzyme S-adenosylmethionine decarboxylase (AdoMetC) was proposed as a target, due to the enzyme inhibition *in vitro*, but no changes in the levels of putrescine and spermidine were observed in *T. brucei* exposed to pentamidine *in vivo* (Berger *et al.*, 1993) and no changes in sensitivity to the drug were observed in null mutants and in cells overexpressing AdoMetC in *Leishmania donovani* (Roberts *et al.*, 2002). It is known that the drug is concentrated in the trypanosomes by the P2 amino-purine transporter and by two pentamidine transporters, a low-capacity high-affinity pentamidine transporter (HAPT1) and a high-capacity low-affinity pentamidine transporter (LAPT1) (Carter *et al.*, 1995; de Koning, 2001). Pentamidine reaches millimolar concentrations in the cell (Carter *et al.*, 1995), so that its toxicity could be caused by multiple cellular targets. Recently it has been found that plasma membrane H<sup>+</sup>-ATPases may also be required for the pentamidine uptake (Alsford *et al.*, 2012, Baker *et al.*, 2012) and the high-affinity

pentamidine transporter (HAPT1) has been genetically identified to be *TbAQP2* and responsible for the cross-resistance of pentamidine and melarsoprol.

**Eflornithine** (1990), or D,L- $\alpha$ -difluoromethylornithine (DFMO), is an analogue of the amino acid ornithine and acts as an irreversible inhibitor of ornithine decarboxylase (ODC), the first enzyme in the biosynthesis of polyamines (Barrett *et al.*, 2007). It was initially developed as anticancer agent, but it has never been registered for this indication. DFMO is equi-potent against the human and parasite enzyme, but the specificity is achieved by the slower turnover of the parasite enzyme (Ghoda *et al.*, 1990; Heby *et al.*, 2003). Inhibition of the enzyme causes an accumulation of S-adenosyl methionine (Yarlett and Bacchi, 1988) and depletion of polyamines and trypanothione (Fairlamb *et al.*, 1987) with consequent inappropriate methylation and increase in oxidative stress. The trypanosomes stop growing and transform into a stumpy-like form that is incapable of antigenic variation. A fully functional host immune system is then required to clear the parasites (Bitonti *et al.*, 1986). Unfortunately the *T. b. rhodesiense* ODC is naturally resistant to the drug, probably due to a faster turnover of the enzyme (Iten *et al.*, 1997). The transport of the drug in the parasite is carrier-mediated and in the laboratory it has been demonstrated that resistance to eflornithine can develop by loss of the amino acid transporter TbAAT6 (Vincent *et al.*, 2010; Baker *et al.*, 2011; Schumann Burkard *et al.*, 2011).

The drug is active against the *T. b. gambiense* infection, but the cost and complexity of the administration does not make its use feasible for the first-stage of the disease. The regimen for the single therapy with eflornithine consists of 100 mg/kg every 6 hours for 14 days by IV infusion with a total of 56 injections for a full treatment (Malvy and Chappuis, 2011) (**Table 1.1**). The main adverse reactions include fever, headache, hypertension, peripheral neuropathy, tremor and gastrointestinal problems, including diarrhea (Brun *et al.*, 2011). The complexity of the treatment in terms of



logistics and costs (the kit for one treatment weighs 20 kg at a cost of €554 including the transport to the rural health centres (Simarro *et al.*, 2012) prompted a search for alternative therapies.

DNDi (Drug for Neglected Disease initiative, a not-for-profit product development partnership) and Médecins Sans Frontières' efforts to shorten the eflornithine-based treatment have been successful with the inclusion of the **nifurtimox-eflornithine combination therapy (NECT)** in 2009 in the WHO list of Essential Medicines (**Figure 1.4**). A multicenter clinical trial has demonstrated that NECT was not inferior compared with the standard eflornithine treatment (Priotto *et al.*, 2009). NECT regimen schedule combines eflornithine by IV at 200 mg/kg every 12 h for 7 days (for a total of 14 IV injections rather than the 56 of the single therapy) with nifurtimox being given orally at 5 mg/kg three times a day for 10 days (**Table 1.1**). Nifurtimox is a nitrofurantoin derivative, already in use for the treatment of acute Chagas disease, caused by *Trypanosoma cruzi*. Previously, nifurtimox was given on compassionate grounds to second-stage patients resistant to melarsoprol, but because of its limited efficacy (30-80%), it was never granted registration as a monotherapy for HAT. Nifurtimox acts as a pro-drug and requires the activation of a specific parasite nitroreductase (NTR) to produce an unsaturated open-chain nitrile that is equally cytotoxic towards mammalian and parasite cells (Hall *et al.*, 2011). Nifurtimox resistance can be quickly generated in the laboratory by loss of a single NTR allele either through selection of drug resistant clones or targeted gene deletion (Wilkinson *et al.*, 2008; Sokolova *et al.*, 2010, Alsford *et al.*, 2012). The only significant adverse events reported for NECT are a higher incidence of tremors, gastrointestinal nausea and anorexia (Priotto *et al.*, 2009).

### 1.2.3 New drug candidates for HAT

A major advancement towards a better cure for HAT is the current clinical development of two new molecules.

**Fexinidazole**, a 2-substituted 5-nitroimidazole, was identified as a promising candidate for the treatment of HAT by DNDi through a compound mining approach aimed at the re-discovery of potential anti-trypanosomal drugs from the already well-known bioactive class of nitroimidazoles (Torreele *et al.*, 2010). The re-utilization of this forgotten molecule (Winkelmann and Raether, 1978; Jennings and Urquhart, 1983) has the potential to represent a major breakthrough in the treatment for HAT. Fexinidazole is the first drug to enter clinical trials for stage two HAT in the last 30 years and it is currently in clinical phase II/III study aimed at assessing safety and efficacy compared to NECT in patients with late stage HAT (<http://clinicaltrials.gov/>). If successfully developed as an oral drug, it will have the invaluable advantage that it can be used in both stage I and stage II of the disease, simplifying both the case management and the diagnostic process, and abolishing the need for a lumbar puncture. Fexinidazole has many of the desirable attributes defined by the Target Product Profile (TPP) for HAT: it is effective against both stages of the disease caused by both *T. b. gambiense* and *T. b. rhodesiense* and a short course of oral treatment (<7 days) is curative in murine models of acute and chronic disease (Torreele *et al.*, 2010). In addition fexinidazole has the potential to achieve cure in the acute model with a single high dose (Torreele *et al.*, 2010). Fexinidazole is quickly metabolised into the sulfoxide and sulfone metabolites, both having similar trypanocidal activity (Sokolova *et al.*, 2010). It is probable that the sustained *in vivo* activity of this drug in mice, rats and dogs is due to the combined exposure of the three chemical species having different half-lives: 1-3 h for fexinidazole, 2-7 h for the sulfoxide, and up to 24 h for the sulfone after oral administration (Torreele *et al.*, 2010). The mechanism of action of this drug seems

to be similar to other nitrodrugs and is mediated by the activation by NTR to generate a cytotoxic species (Hall *et al.*, 2011). Resistance to nitrodrugs can be easily generated in the laboratory (Sokolova *et al.*, 2010): cell lines resistant to nifurtimox show increased resistance to fexinidazole, raising the risk that resistance to this drug could develop even before its approval for clinical use. In order to reduce the risk of resistance, it would be necessary to evaluate a combination therapy of fexinidazole with existing drugs or new drugs that might become available in the future (Kaiser *et al.*, 2011).

**SCYX-7158**, a benzoxaborole derivative, represents the latest hope to develop a new treatment for HAT. This molecule was discovered as part of collaborative Drug Discovery program funded by DNDi and performed in collaboration with Scynexis and Pace University. This chemical series was initially synthesised by Anacor Pharmaceuticals and its activity against *T. brucei* was determined by phenotypic screening performed at the Sandler Center. Scynexis further developed this series until the identification of SCYX-7158 (Nare *et al.*, 2010; Jacobs *et al.*, 2011). This compound is active *in vitro* against both *T. b. gambiense* and *T. b. rhodesiense*, it is also efficacious in both stage 1 and stage 2 murine HAT models. The *in vivo* pharmacokinetic studies have demonstrated high bioavailability after oral administration, long half-life (24 h), high brain exposure and quick distribution to CSF to achieve therapeutically relevant concentrations (Jacobs *et al.*, 2011). Up to date the mechanism of action of the oxaborole compounds is not known, but the rapid trypanocidal effect (12 h above the minimum inhibitory concentration) suggests that oxaboroles are retained within the parasites or they exert irreversible effects on the potential target(s) during this limited time frame (Nare *et al.*, 2010).

Based on its promising characteristics SCYX-7158 has been selected as a preclinical candidate and in 2012 has entered phase I clinical studies on healthy

volunteers from Sub-Saharan Africa to determine safety, tolerability, pharmacokinetics and pharmacodynamics.

### **1.3 Target assessment and target product profile**

The lack of efficacious drugs for all subspecies and stages of the disease urges research for drugs with novel mechanisms of action. The clinical development of fexinidazole and SCYX-7158 gives hope for an oral treatment for both stages of the disease and both subspecies, and for the simplification of the diagnosis by elimination of the disease staging. Unfortunately, the attrition rate of the clinical development process is quite high, with only one in 10 molecules completing the clinical development process (Brown and Superti-Furga *et al.*, 2003).

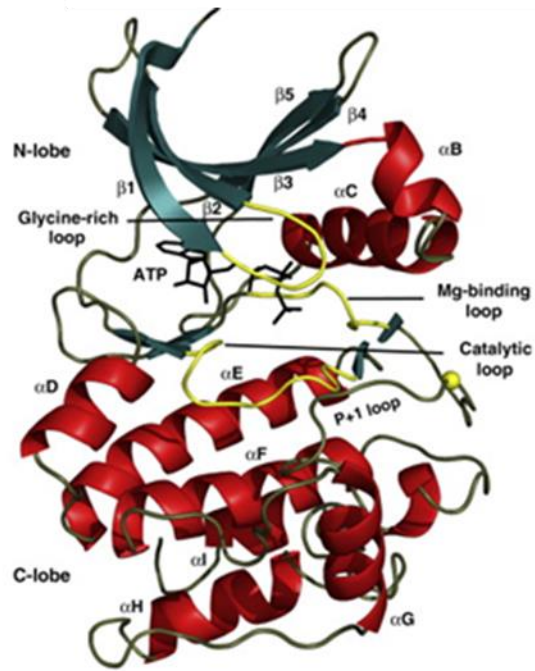
New drugs can be developed either using a phenotypic approach or by target-driven screening campaigns (Gilbert, 2013). Although the first approach selects molecules active in cells and overcomes problems related to drug import and efflux, the lack of knowledge of the mechanism of action requires target deconvolution and the chemical development of the active compounds cannot be guided by structural information. The second approach requires the identification of validated drug targets as starting point for a screening campaign. This approach has the advantage to be target-directed and the availability of structural knowledge or preliminary chemical validation in other species can speed up the process of drug discovery, but poses the challenge to translate the biochemical potency into efficacy against cells. Further, many compounds active in protein-based assays are inactive in whole cells owing to their physicochemical properties (Nwaka and Hudson, 2006; Renslo and McKerrow, 2006).

The process of drug target identification and validation at the Drug Discovery Unit in Dundee (DDU) is based upon a traffic-light system. Target assessment is performed against a series of parameters, such as genetic and chemical validation,

druggability, assay feasibility, toxicity, structural information and resistance potential as illustrated in detail by Frearson and colleagues (2007). The target assessment process has to be linked to the target-product profile (TPP) that defines the therapeutic requirements for a pre-clinical candidate molecule (Frearson *et al.*, 2007; Wyatt *et al.*, 2011). The biological and physico-chemical properties are linked to the specific disease setting and the need of the affected populations. In the case of sleeping sickness an oral or parenteral treatment for the late stage of both *gambiense* and *rhodesiense* infection is among the priorities.

The sequencing of the genome of the Trityps, *Trypanosoma brucei* (Berriman *et al.*, 2005; Jackson *et al.*, 2010), *Leishmania major* (Ivens *et al.*, 2005) and *Trypanosoma cruzi* (El-sayed *et al.*, 2005) has undiscovered a plethora of potential drug targets. In order to rationalize and prioritize these targets various initiatives have been developed aimed at collating all available information and ranking the targets using similar parameters as those used for the target assessment in the DDU (Agüero *et al.*, 2008).

The DDU has a balanced pipeline, including drug targets unique to the parasite for which the biology is well understood, but with little knowledge about their druggability, such as trypanothione reductase (Patterson *et al.*, 2011), trypanothione synthetase (Torrie *et al.*, 2009), and *Trypanosome brucei* pteridine reductase 1 (Spinks *et al.*, 2011); drug targets that have been well validated in other organisms both genetically and chemically are also included, but the presence of orthologues poses the problem of toxicity, which could be overcome by exploiting the structural differences between the mammalian and parasite target (examples of such targets are *Trypanosoma brucei* N-myristoyltransferase (Frearson *et al.*, 2010) and *T. brucei* cdc2-related protein kinase CRK3 (Cleghorn *et al.*, 2011)). In combination the DDU adopts a phenotypic strategy aimed at the identification of bioactive molecules against the whole organism,



**Figure 1.7 Generic structure of a protein kinase**

Protein kinases have a small N-lobe containing five  $\beta$  sheets and one  $\alpha$  helix called  $\alpha$ C-helix. The bigger C-lobe contains 6  $\alpha$ -helices. The ATP binds in the cleft between the two lobes. The glycine-rich loop and the autophosphorylation loop are indicated in yellow (Figure from Taylor and Kornev, 2011).

postponing to a later stage the identification of the molecular target (Smith *et al.*, 2011, De Rycker *et al.*, 2013).

## 1.4 Protein kinases

Eukaryotic protein kinases (ePKs) catalyse a chemical reaction that requires the binding of ATP (or GTP) as a complex with  $Mg^{2+}$  or  $Mn^{2+}$ , the binding of the protein substrate and transfer of the  $\gamma$ -phosphate from the phosphate donor to the hydroxyl residue (serine, threonine or tyrosine) of the protein substrate (Hanks and Hunter, 1995). The catalytic domain of protein kinases consists of twelve subdomains conserved across the major groups and containing characteristic conserved residues (**Figure 1.7**). The twelve subdomains are organised in two lobes: a smaller amino-terminal lobe (N-lobe) composed of five  $\beta$  sheets and one  $\alpha$  helix and a larger carboxy-terminal lobe (C-lobe) composed of six  $\alpha$ -helical domains (Hans and Hunter, 1995; Huse and Kuriyan, 2002). The two lobes are connected by a hinge linker. The amino-terminal lobe is involved in the binding of the ATP and the carboxy-terminal lobe in the binding of the peptide substrate and initiation of the chemical reaction, with the cleft between the two domains being responsible for the catalysis. ATP is bound in this cleft between the two lobes, in an area described by a conserved loop (P loop). The P loop is a highly flexible glycine-rich area also containing an aromatic residue: the glycine residues allow the coordination of the phosphates of the ATP via backbone interactions and the aromatic residue participates in the phosphate transfer. The activation loop, generally 20-30 amino acids in length, is centrally located and allows the binding of the peptide acceptor in proximity of the  $\gamma$ -phosphate. Phosphorylation of the activation loop allows the stabilization of the kinase in an active conformation. The position and number of phosphorylation sites varies among ePKs (Hans and Hunter, 1995; Huse and Kuriyan, 2002, Taylor and Kornev, 2011) (**Figure 1.7**).

**Table 1.2 Classification of PKs by groups in human and in *Trypanosoma brucei***

Protein kinase group	<i>Hs</i> kinases	<i>T. brucei</i> kinases			
	Manning <i>et al.</i> , 2002	Parsons <i>et al.</i> , 2005	Nett <i>et al.</i> , 2009a	Urbaniak <i>et al.</i> , 2012	Jones <i>et al.</i> , 2014
AGC	63	12	22	18	14
CAMK	74	14	28	33	14
CK1	12	5	5	5	5
CMGC	61	42	47	47	42
RGC	5	0	0	0	0
Other	83	39	40	38	42
STE	47	25	28	30	25
TK	90	0	0	0	0
TKL	43	0	0	0	0
unique	0	19	0	0	28
atypical	40	20	12	11	20
Total	518	176	182	182	190

Protein kinases classification as defined by Miranda-Saavedra and Burton (2007): **AGC** (including cyclic-nucleotide and calcium-phospholipid dependent kinases, ribosomal S6-phosphorylating kinases, G protein coupled kinases, and close relatives); **CAMK** (calcium/calmodulin-regulated kinases); **CK1** (casein kinases I, and close relatives); **CMGC** (cyclin-dependent kinases, mitogen activated protein kinases, glycogen synthase 3 kinases, casein kinases II, CDK-like kinases); **RGC** (receptor guanylate cyclases); **Other** (mixed collection of kinases that could not be classified easily into the previous families); **STE** (protein kinases involved in MAP kinase cascades, homologs of yeast sterile kinases); **TK** (tyrosine kinases); and **TKL** (tyrosine-kinase like, which are, in fact, serine/threonine kinases); **atypical kinases** (including alpha, PIKK [phosphatidylinositol 3' kinase-related kinases], PDHK [pyruvate dehydrogenase kinases], and RIO [right open reading frame kinases]).



Eukaryotic PKs have been classified into major groups (Hanks and Hunter, 1995; Manning *et al.*, 2002; Miranda-Saavedra and Barton, 2007) as described in more detail in **Table 1.2**. ePKs play an important role in eukaryotic organisms (such as *Caenorhabditis elegans* (Plowman *et al.*, 1999), *Drosophila melanogaster* (Rubin *et al.*, 2000), and *Homo sapiens* (Manning *et al.*, 2002) including cell cycle regulation, differentiation, apoptosis, cell shape and organization, transcription, and metabolism (Cohen, 2000). Although for years protein phosphorylation was believed to be a regulation mechanism present only in higher eukaryotes, several bacterial homologues have been identified in recent years (Kennelly, 2003). Mutation and deregulation of protein kinases are responsible for human diseases, such as cancer and auto-inflammatory diseases. The development of new molecules targeting this family has demonstrated that ePKs have structural features that make them druggable (Hopkins and Groom, 2002), resulting in the introduction into the clinic of blockbusters such as Gleevec and Iressa (Druker *et al.*, 1996; Barker *et al.*, 2001). A total of 22 kinase inhibitors have been approved as drugs for the treatment of various cancers and more than 300 compounds are under current clinical development (<http://clinicaltrials.gov/>). The economic potential to exploit this family to develop drugs for “profitable” diseases such as cancer, inflammation, diabetes and Alzheimer’s disease, has resulted in an exponential increase in structural and chemical knowledge (Noble *et al.*, 2004).

#### **1.4.1 The *T. brucei* kinome**

By comparative analysis of the kinome of the three trypanosomatids, *Trypanosoma brucei*, *Trypanosoma cruzi*, and *Leishmania major*, Parsons and colleagues (2005) have initially found that orthologues are conserved across the three species and their kinome consists of 176, 190 and 199 PKs, respectively. The *T. brucei* kinome is around one third of the human one (518 PKs) and it is approximately 2% of

the parasite genome (~7500 proteins in total), well above the number of PKs for other unicellular parasites, such as *Plasmodium falciparum* (Leroy and Doerig, 2008; Parsons *et al.*, 2005) (**Table 1.2**). Further studies have refined the initial classification done by Parsons and colleagues and up to date the *T. brucei* kinome consists of a total of 190 members: 170 ePKs (including 12 predicted pseudo-kinases) and 20 atypical PKs (Nett *et al.*, 2009a; Urbaniak *et al.*, 2012a; Jones *et al.*, 2014) (**Table 1.2**). Parsons and colleagues (2005) also identified 19 unique PKs in *Trypanosoma brucei*, which are the least conserved among the trypanosomatids and their divergence might be a potential source of selective drug targets, although they have been reassigned to ePKs groups by other studies (Nett *et al.*, 2009a; Urbaniak *et al.*, 2012a).

The most striking difference among the parasite and human kinome is the absence of kinases belonging to the receptor-linked TK and TKL groups. Nevertheless tyrosine phosphorylation in *T. brucei* has been reported and attributed to the action of atypical tyrosine kinases such as Wee1, and dual specificity kinases (DYRKs, CLKs, and STE7) (Parsons *et al.*, 1991; Nett *et al.*, 2009b; Urbaniak *et al.*, 2013). Moreover certain groups and families of ePKs are over-represented, such as CMGC, STE and NEK kinases. Another unique feature is the relative scarcity of transmembrane domains, to date only few kinases have been reported to be associated to membranes; eukaryotic initiation factor 2 alpha kinase at the membrane of the flagellar pocket (Moraes *et al.*, 2007), LDK at the membrane of lipid droplets (Flaspohler *et al.*, 2010) and recently RDK1 (Jones *et al.*, 2014; see section 1.4.2). Finally approximately 8% of the ePKs were reported to be pseudokinases, catalytically inactive, lacking some of the conserved residues indispensable for the enzymatic activity (Parsons *et al.*, 2005).

Phosphorylation has to play a key role in the parasite biology if protein kinases represent such large proportion of the genome. This has been recently confirmed by a global SILAC phosphoproteomic analysis performed in both procyclic form (PCF) and

bloodstream form (BSF) of *T. b. brucei* that has identified just above 10,000 phosphorylation sites on 2,551 proteins (Urbaniak *et al.*, 2013). PKs not only represent 2% of the entire genome, but also one third of the expressed proteins seem to be regulated by phosphorylation. This is particularly important considering that *T. brucei* lacks of transcription factors and regulation of gene expression is done through mRNA processing (Clayton and Shapira, 2007) and post-translational modification (Urbaniak *et al.*, 2013). The fact that phosphorylation as well as protein expression is differential in the two life stages of the parasite just adds a further degree of complexity to the understanding of the role of protein phosphorylation in *T. brucei* and in the kinetoplastids.

#### **1.4.2 Protein kinases in *T. brucei* as drug targets**

Protein kinases were initially proposed as drug targets in trypanosomes and *Leishmania* (Naula *et al.*, 2005), due to their relative abundance in the trypanomatids genome and the potential role played in important functions such as regulation of the cell cycle and cytokinesis (Naula *et al.*, 2005; Hammarton *et al.*, 2007a; Hammarton, 2007b). Indeed multiple kinases have now been studied and found to play a role in cell cycle regulation; these studies have also elucidated the differences in cell cycle between the two life stages of the parasite (Li, 2012; Farr and Gull, 2012, Tu *et al.*, 2005). As expected cyclins and the CDK-related kinases (CRKs) are among the key regulators of the cell cycle: CRK1/cyc2 regulates DNA replication and G1/S transition, whilst CRK3/cyc6 controls G2/M progression (Tu and Wang, 2004; Tu *et al.*, 2005). AUK1 together with TLK1 plays a role in mitosis and spindle assembly (Tu *et al.*, 2006, Li *et al.*, 2006; Li *et al.*, 2007; Li *et al.*, 2008b), whereas MAPK6, PLK, PK50 and PK53 have been characterised as main regulators of cytokinesis (Wei and Li, 2014; Hammarton *et al.*, 2007c; Li *et al.*, 2010; Ma *et al.*, 2008). Akiyoshi and Gull (2014)

have recently identified 19 unconventional kinetochore proteins in *Trypanosoma brucei*, four of them are protein kinases: CLK1, CLK2, PK6 and a protein previously annotated as putative protein kinase (named KK10, KK19, KKT3 and KKT2, respectively, after the discovery of their function). These proteins are all responsible for DNA segregation.

Cell replication in *T. brucei* requires also a coordinated duplication of its organelles: CRK9 seems to play a role in kinetoplast segregation in both PCF and BSF (Gourguechon *et al.*, 2009; Jones *et al.*, 2014); PLK is required for bilobe duplication and Golgi biogenesis as well as basal body duplication and flagellum elongation (de Graffenried *et al.*, 2008; Ikeda and de Graffenried, 2012); and Vps34, a lipid kinase, segregates the Golgi complex after duplication (Hall *et al.*, 2006). Jones and colleagues (2014) have performed a kinome wide RNAi study looking for cell viability in BSF and confirmed previously published findings, but also further characterised the function of known essential kinases as GSK3 short and Casein kinase 1.2 (Ojo *et al.*, 2008; Urbaniak, 2009), and identified 15 novel cell cycle regulators.

Protein kinases in *T. brucei* also play important roles in cell differentiation: ZFK (Vassella *et al.*, 2001), mTOR4 (Barquilla *et al.*, 2012) and MAPK5 (Domenicali Pfister *et al.*, 2006) have been characterized as negative regulators of the differentiation from the proliferative trypomastigote to the quiescent stumpy induced by SIF in the human host. More recently an RNAi study has identified positive regulators of the quorum signalling and not surprisingly found among others three protein kinases (FlaK, NEK, and YAK) responsible for the signal transduction pathway that induces the cell-cycle arrest in G0/G1 characteristic of stumpy forms (Mony *et al.*, 2014). The aforementioned RNAi study from Jones and colleagues (2014) has also identified two new kinases acting as repressors of differentiation from bloodstream to procyclic form, the STE-11 like RDK1, which works in concert with the tyrosine phosphatase PTP1 and localizes at the membrane, and the RDK2 belonging to the NEK family, which works

independently from other triggers. Finally MKK1, a MAP Kinase Kinase, is the first kinase associated with the transmission of the disease from insect to human (Morand *et al.*, 2012).

Protein kinases in *T. brucei* are also important for cell proliferation: *TbTOR1* regulates protein synthesis and nucleolar structure (Barquilla *et al.*, 2008); *TbTOR2* controls cell polarization and endocytosis (Barquilla *et al.*, 2009); eukaryotic initiation factor 2 alpha kinase is associated to the flagellar pocket and is involved in sensing and translational control (Moraes *et al.*, 2007) and CRK12 plays a role in endocytosis (Monnerat *et al.*, 2013). Very few physiological substrates of protein kinases have been identified compared to the abundance of phosphorylated proteins and sites in both life stages. The *T. brucei* phosphoproteome is enriched in proteins involved in functions such as mRNA binding and flagellum motility for which very little is known of the role played by protein kinases (Urbaniak *et al.*, 2013). In the phosphoproteomic study, it also emerged that the MAPK kinase pathway is activated in both life cycle stages (Urbaniak *et al.*, 2013). It is known that in promastigote *Leishmania mexicana* MPK9 regulates the flagellum biogenesis and maintenance (Bengs *et al.*, 2005), thus more work is necessary to investigate whether MAPK kinases control the flagellum biogenesis and functions also in *T. brucei* (Rotureau *et al.*, 2009).

In addition to single gene studies, various RNAi studies have confirmed that protein kinases represent a source of drug targets in *T. brucei*. A small RNAi screen has attempted to genetically validate 30 uncharacterized protein kinases in *T. brucei* and identified MPK2 and CRK12 as important genes for cell proliferation (Mackey *et al.*, 2011). A target sequencing RNAi screen against the *T. brucei* genome has found that protein kinases are over-represented among genes responsible for loss-of-fitness in bloodstream form cells and differentiation (Alsford *et al.*, 2011). The already mentioned RNAi study of the whole *T. brucei* kinome has identified a comprehensive list of

**Table 1.3 *T. brucei* protein kinases genetically and chemically investigated**

<b>Gene Product</b>	<b>Function</b>	<b>Genetic validation</b>	<b>References</b>
ZFK (Tb927.11.9270)	Differentiation from BSF to stumpy form, but not essential for BSF growth	gene knock-out	Vassella <i>et al.</i> , 2001
Nuclear DBF-2-related (NDR) kinases PK50 and PK53 (Tb927.10.4940 and Tb927.7.5770)	PK50 and PK53 are essential regulators of cytokinesis. <b>Chemical validation:</b> HTS against molecular targets PK50 and PK53 resulted in potent inhibition in the biochemical assay but low activity in cellular assay. Only PK53 inhibitor could bind the protein target in lysate.	RNAi	Hammarton <i>et al.</i> , 2005; Ma <i>et al.</i> , 2010; DDU project ; Urbaniak <i>et al.</i> , 2012a
PKAC1/PKAC2 (Tb927.9.11100 and Tb927.9.11030)	Growth defect in BSF, cytokinesis	RNAi, gene knockout	Kramer <i>et al.</i> , 2004
LDK (Tb927.11.8940)	Mild growth defect in BSF, lipid droplets biogenesis		Flaspohler <i>et al.</i> , 2010
Casein Kinase 1.2 (Tb927.5.800)	Growth inhibition in BSF. Role in kinetoplast division and cytokinesis <b>Chemical validation:</b> purvalanol B and imidazopyridines bind native <i>Lmex</i> CK1.2 and <i>Lmaj</i> CK1.2, respectively. Lapatinib, canertinib, AEEE788 bind CK1.2 in <i>T. brucei</i> lysate.	RNAi, gene knock-out attempted	Urbaniak <i>et al.</i> ; 2009; Jones <i>et al.</i> , 2014, Knockaert <i>et al.</i> , 2000; Allocco <i>et al.</i> , 2006, Katiyar <i>et al.</i> , 2013
CRK1 (Tb927.10.1070)	Growth defect in BSF and PCF Interaction with <i>cyc2</i> , <i>cyc4</i> , <i>cyc5</i> and <i>cyc7</i> . Role in DNA replication and in G1/S transition <b>Chemical validation:</b> Diaminopyrimidines	RNAi	Tu and Wang, 2004; Tu and Wang, 2005; Li <i>et al.</i> , 2012; Mercer <i>et al.</i> , 2011; Jones <i>et al.</i> , 2014
CRK3 (Tb927.10.4990)	Growth defect, interaction with <i>cyc6</i> , block in G2/M <b>Chemical validation:</b> <i>Lmex</i> CRK3/ <i>cyc6</i> nanomolar inhibitors were identified by HTS, but they did not translate in potent inhibitors <i>in vitro</i> against either <i>T. brucei</i> or <i>Leishmania</i>	RNAi	Tu and Wang, 2004; Cleghorn <i>et al.</i> , 2011

Gene Product	Function	Genetic validation	References
(Tb927.6.1780)	<b>Chemical validation</b> in both BSF and PCF		
MAPK5 (Tb927.6.4220)	Role in differentiation from BSF to stumpy form	gene knock-out in PCF	Domenicali Pfister <i>et al.</i> , 2006
MAPK6 (MPK2) (Tb927.10.5140)	Furrow ingression and cytokinesis completion in PCF. Rapid cell death, essential for cytokinesis initiation in BSF. <b>Chemical validation:</b> Diaminopyrimidines	RNAi	Wei and Li, 2014; Jones <i>et al.</i> , 2014, Mercer <i>et al.</i> , 2011, Mackey <i>et al.</i> , 2011
MAPK14 (Tb927.3.690)	Cell death; essential in cytokinesis in BSF. <b>Chemical validation:</b> Diaminopyrimidines	RNAi	Jones <i>et al.</i> , 2014, Mercer <i>et al.</i> , 2011
MAPK9 (Tb927.10.14800)	<b>Chemical validation:</b> Diaminopyrimidines		Mercer <i>et al.</i> , 2011
ECK1 (Tb927.11.16790)	Essential in both PCF and BSF. In procyclic C-terminal truncation phenotype: abnormal cell proliferation	Expression of truncated <i>TbECK1</i>	Ellis <i>et al.</i> , 2004
KKT2 (Tb927.11.10520)	Unconventional kinetochore protein: segregation of both megabase chromosomes and minichromosomes.	RNAi	Akiyoshi and Gull, 2014
KKT3 (Tb927.9.10920)	Unconventional kinetochore protein: arrested cell growth in BSF, normal cell cycle.	RNAi	Akiyoshi and Gull, 2014; Jones <i>et al.</i> , 2014
FlaK (Tb927.2.2720)	Signal transduction component of the quorum signalling	RNAi	Mony <i>et al.</i> , 2014
Aurora Kinase I (AUK1) (Tb927.11.8220)	Divergent function: chromosome segregation and cytokinesis initiation in PCF; cytokinesis in BSF without a control on mitosis or organelle duplication. <b>Chemical validation:</b> VX-680 and Hesperidin replicate AUK1 RNAi phenotype. Danusertib inhibits AUK1 and is efficacious in cells and animal model	RNAi	Tu <i>et al.</i> , 2006 ; Li and Wang, 2006; Li <i>et al.</i> , 2008a; Jetton <i>et al.</i> , 2009; Li <i>et al.</i> , 2009; Hu <i>et al.</i> , 2014; Ochiana <i>et al.</i> , 2013.

Gene Product	Function Chemical validation	Genetic validation	References
CRK6 (Tb927.11.1180)	No defect <b>Chemical validation:</b> Diaminopyrimidines	RNAi	Tu and Wang, 2004; Mercer <i>et al.</i> , 2011
CRK9 (Tb927.2.4510)	Mitosis and cytokinesis initiation in PF and BSF Role in kinetoplast segregation	RNAi	Gourguechon and Wang, 2009; Jones <i>et al.</i> , 2014
CRK12 (Tb927.11.12310)	CRK12:cyc9 complex in both PCF and BSF Growth inhibition in BSF, role in endocytosis, no effect on cell cycle	RNAi, gene knock-out	Mackey <i>et al.</i> , 2011; Monnerat <i>et al.</i> , 2013; Jones <i>et al.</i> , 2014; Merritt and Stuart, 2013
CLK1 and CLK2 (Tb927.11.12410 and Tb927.11.12420)	Growth defect in BSF. Responsible for missegregation in anaphase cells <b>Chemical validation:</b> hypothemycin, a covalent inhibitor of CDXG kinase, targets native CLK1	RNAi	Akiyoshi and Gull, 2014; Nishino <i>et al.</i> , 2013; Jones <i>et al.</i> , 2014
DYRK/YAK (Tb927.10.15020)	Signal transduction component of the quorum signalling	RNAi	Mony <i>et al.</i> , 2014
<i>Tb</i> GSK3 short and <i>Tb</i> GSK3 long (Tb927.10.13780 and Tb927.7.2420)	<i>Tb</i> GSK3 short RNAi causes growth defects in BSF and differentiation. <i>Tb</i> GSK3 long RNAi causes gain of fitness in procyclic stage. <b>Chemical validation:</b> Non selective/non-specific inhibitors of <i>Hs</i> GSK3 cause growth defects in <i>T. brucei</i> and <i>Leishmania</i> . HTS vs Merck and Pfizer GSK3 focused libraries. Target in cell of hypothemycin, canertinib, lapatinib, and DDD00085893	RNAi	Ojo <i>et al.</i> , 2008; Oduor <i>et al.</i> , 2011; Urbaniak <i>et al.</i> , 2012a; Nishino <i>et al.</i> , 2013; Katiyar <i>et al.</i> , 2013; Jones <i>et al.</i> , 2014.
KFR1 (Tb927.10.7780)	Proliferation of BSF induced by interferon- $\gamma$		Hua and Wang, 1997
MAPK2 (Tb927.10.16030)	Role in differentiation from BSF to PCF	gene knock-out	Muller <i>et al.</i> , 2002
MAPK4	Confers resistance to temperature stress, but not essential	gene knock-out	Guttinger <i>et al.</i> , 2007



Gene Product	Function Chemical validation	Genetic validation	References
<i>T. brucei</i> Tousled like Kinase 1 ( <i>TbTLK1</i> ) (Tb927.4.5180)	TLK1 cooperates with Aurora in the regulation of the spindle assembly and chromosome segregation <b>Chemical validation:</b> lapatinib, canertinib, AEEE788	RNAi	Li <i>et al.</i> , 2007; Li <i>et al.</i> , 2008b; Katiyar <i>et al.</i> , 2013, Jones <i>et al.</i> , 2014
Casein kinase 2 ( <i>TbCK2</i> )	Localized at the nucleolus, interacts with the nucleolar proteins Nopp44/46 and NOG1. Growth arrest in BSF. Role in cytokinesis	RNAi	Park <i>et al.</i> , 2002; Jensen <i>et al.</i> , 2006; Jones <i>et al.</i> , 2014
NEK12.1/ RDK2 (Tb927.4.5310)	Repressor of differentiation in BSF	RNAi	Jones <i>et al.</i> , 2014
NEK17 (Tb927.10.5950)	Signal transduction component of the quorum signalling	RNAi	Mony <i>et al.</i> , 2014
<i>TbNRKC</i> kinase (Tb927.10.460)	Basal body and cytokinesis	RNAi and overexpression mutants	Pradel <i>et al.</i> , 2006
Polo-like kinase ( <i>TbPLK</i> ) (Tb927.7.6310)	Basal body replication and kinetoplast migration, cytokinesis initiation. Localized at the flagellum is responsible for flagellum attachment elongation. <b>Chemical validation:</b> GW843682X causes block of cytokinesis. Analogue-sensitive approach to specifically target PLK in cells.	RNAi	Graham <i>et al.</i> , 1998; Kumar and Wang, 2006; Hammarton <i>et al.</i> , 2007c, de Graffenried <i>et al.</i> , 2008 ; Li <i>et al.</i> , 2010; Sun and Wang, 2011; Yu <i>et al.</i> , 2012; Ikeda <i>et al.</i> , 2012; Lozano-Núñez <i>et al.</i> , 2013
RDK1 (Tb927.11.14070)	Repressor of differentiation in BSF	RNAi	Jones <i>et al.</i> , 2014
<i>TbTOR1</i> ( <i>T. brucei</i> Target of Rapamycin) (Tb927.10.8420)	Regulation of cell growth by regulating cell cycle, protein synthesis and nucleolus structure <b>Chemical validation:</b> NVP-BEZ235 is a nanomolar mTOR/PI3K inhibitor, with efficacy indication in an acute mouse infection model.	RNAi	Barquilla <i>et al.</i> , 2008, de Jesus <i>et al.</i> , 2010, Diaz-Gonzalez <i>et al.</i> , 2011.

<b>Gene Product</b>	<b>Function</b>	<b>Genetic validation</b>	<b>References</b>
<i>Tb</i> TOR2 (Tb927.4.420)	Cell polarization, endocytosis and cytokinesis <b>Chemical validation:</b> TOR complex 2 is sensitive to rapamycin	RNAi	Barquilla <i>et al.</i> , 2009
<i>Tb</i> TOR4 (Tb927.1.1930)	Negative regulator of differentiation	RNAi	Barquilla <i>et al.</i> , 2012
<i>Tb</i> Vps34 (Tb927.8.6210)	Lipid kinase. Segregation of the Golgi complex	RNAi	Hall <i>et al.</i> , 2006
Eukaryotic initiation factor 2 alpha kinase (eIF2alpha kinase) (Tb927.4.2500)	Associated to the membrane of the flagellar pocket, involved in sensing and translational control	NA	Moraes <i>et al.</i> , 2007
MKK1	Role in transmission from insect to human Not essential in BSF and in PCF	Knock-out	Morand <i>et al.</i> , 2012; Jensen <i>et al.</i> , 2011

essential protein kinases, cell cycle and differentiation regulators (Jones *et al.*, 2014). These studies have a certain degree of agreement among them and the previous literature, but still the proposed essential targets will require further chemical and genetic validation (Jones *et al.*, 2014, Frearson *et al.*, 2007).

In recent years more information has been collated regarding which protein kinase targets might be essential in *T. brucei*. The majority of the genetic studies are performed by RNAi and over-expression with tetracycline inducible systems as gene knock-outs remain generally labour intensive and tricky for proteins that are finely regulated (Merritt and Stuart, 2013). The divergences among RNAi studies derive from the differential levels of gene knock-down achieved using different strains and constructs and off-target effects (degradation of mRNA of genes that were not intentionally targeted), making gene validation by RNAi not always conclusive (Merritt and Stuart, 2013).

In some cases chemical validation has been attempted using known human protein kinase inhibitors that are not selective or specific for *T. brucei* (Ojo *et al.*, 2008; Jetton *et al.*, 2009; Diaz-Gonzalez *et al.*, 2011). High throughput screening campaigns have been performed only for a handful of protein kinases (*Tb*GSK3 short, PK50, PK53, *Tb*PLK), mainly due to the lack of active recombinant proteins (Ojo *et al.*, 2008; Oduor *et al.*, 2011; Urbaniak *et al.*, 2012a; DDU unpublished data).

Although a detailed description of all the *T. brucei* kinases is beyond the scope of this thesis, a list of the best characterized kinases, their function and whether have been validated by genetic or chemical means is reported in **Table 1.3**.

Protein kinases are potential drug targets for parasitic diseases. On the one hand, they offer the comparative advantage of piggybacking on the extensive knowledge already existing on human protein kinase inhibitors. On the other hand targeting enzymes present in the human host poses the challenge to develop selective inhibitors

against the parasite ortholog. However, the fact that most of the parasite kinases share less than 60% identity in the catalytic domain with respect to the corresponding human kinases suggests that it is possible to develop selective kinase inhibitors that specifically target the parasite enzyme (Naula *et al.*, 2005). Furthermore the availability of structural information relative to the human kinases combined with computational modelling is a powerful tool for the design of selective inhibitors exploiting key differences in the amino acid sequences.

This thesis assesses the short form of glycogen synthase kinase 3 (GSK3) as a drug target in *Trypanosoma brucei* using a combined chemical and genetic approach (see section 1.7).

## **1.5 Glycogen synthase kinase: role in diseases and inhibitor development**

In humans the two isoforms of glycogen synthase kinase 3 (GSK3) are called GSK3 $\alpha$  and GSK3 $\beta$ , these two enzymes have molecular weight of 51 kDa and 47 kDa, respectively. They are very similar in their kinase domain (98% identity), but differ in their amino- and carboxyl-terminal and functional activity (Doble and Woodget, 2003). Yet they are not interchangeable as demonstrated by the inability of GSK3 $\alpha$  to rescue the embryonic lethality of GSK3 $\beta$  knock-out (Hoeflich *et al.*, 2000). There are two variants of HsGSK3 $\beta$ , the minor splice variant HsGSK3- $\beta$ 2 has an insert of 13 amino acids within the kinase domain. The insertion seems to be responsible for a differential recognition of scaffold proteins and also a differential localization (Mukai *et al.*, 2002).

The crystal structure of HsGSK3 $\beta$  has been solved shedding light on the mechanisms of activation, substrate recognition and inhibition (Dajani *et al.*, 2001).

GSK3 is constitutively active, and its regulation occurs either by inactivation by upstream signals or by changes in substrate recognition.

The activation loop of GSK3 is phosphorylated at Tyrosine 216 and Tyrosine 279 in GSK3 $\beta$  and GSK3 $\alpha$ , respectively. The tyrosine phosphorylation is not required for kinase activity, but facilitates the phosphorylation of the substrate and increases protein stability (Dajani *et al.*, 2001; Cole *et al.*, 2004). The activation loop phosphorylation has been shown to be an autophosphorylation event (Cole *et al.*, 2004), which happens after translation requiring the involvement of a chaperone protein (Lochhead *et al.*, 2006).

GSK3 preferentially phosphorylates substrates that are pre-phosphorylated at the priming site at the C-terminus in its consensus sequence Ser/Thr-X-X-X-(phospho)Ser/Thr. The priming site is believed to interact with a positively charged pocket (Arg96, Arg180 and Lys205) that positions the protein substrate in the right orientation for the subsequent phosphorylation. Un-primed substrates can also be recognised and phosphorylated by GSK3, as they often have a negatively charged residue that mimics the primed site (Doble and Woodget, 2003).

GSK3 can be regulated by phosphorylation on regulatory Ser9 and Ser21 residues, respectively in GSK3 $\beta$  and GSK3 $\alpha$ . The upstream kinases protein kinase B (also called AKT), MAPK-activated protein kinase 1 (also called p90RSK), p70 ribosomal S6 kinase, cAMP-activated protein kinase and protein kinase C (PKC) are responsible for phosphorylation of Ser9/21 causing reduction in kinase activity and consequently reduction in phosphorylation of downstream substrates (Doble and Woodget, 2003, Cross *et al.*, 1995; Cohen *et al.*, 1997). The phosphorylation on Ser9/21 mimics the primed substrate, by occupying the site for the substrate in a competitive manner. High concentrations of primed substrate can compete for this intermolecular inhibition (Doble and Woodget, 2003).

GSK3 can also be regulated by interaction with docking proteins that allow the substrate to make contacts with the priming kinase. Specific combinations of scaffolding proteins and substrates (e.g. axin is the docking protein for casein kinase 1 and presenilin is the docking protein for protein kinase A) determine different functions of GSK3 in the cell (Meijer *et al.*, 2004).

Apart from glycogen synthase, GSK3 has a plethora of substrates in the cell in various compartments: metabolic proteins such as ATP citrate lyase, acetyl-CoA carboxylase, cyclin D1 and E, the eukaryotic initiation factor eIF2B, axin, hexokinase, and presenilin; structural proteins such as tau, neurofilaments, dynein, kinesin light chain, and microtubule associated proteins (MAPIB); and transcription factors such  $\beta$ -catenin, CREB, Myc, NF $\kappa$ B, HSF-1, Notch, p53, and HIF-1. This list is not exhaustive and the functions and phosphorylation effects on various substrates have been reviewed by Doble and Woodget (2003). Thus it is not surprising that GSK3 in mammalian cells is involved in a wide spectrum of cellular processes, including cell proliferation, differentiation, embryonic development, metabolism, transcription, translation, cytoskeletal regulation, intracellular vesicular transport, cell cycle progression, circadian rhythm regulation and apoptosis (Frame and Cohen, 2001).

GSK3 plays an important role in two major pathways, the Wnt/ $\beta$ -catenin pathway, essential during embryonic development, and the Hedgehog pathway, involved in cell fate determination and morphology (Doble and Woodget, 2003). Protein stability seems to be one of the mechanisms through which GSK3 regulates cellular processes (Xu *et al.*, 2009). The phosphorylation of target proteins by GSK3 induces ubiquitination and proteolysis, in the case of cytoplasmic levels of  $\beta$ -catenin in the wnt pathway, or proteolytic cleavage of the substrate, in the case of transcription activator Ci (*Cubitus interruptus*) in the Hedgehog pathway. GSK3 also plays an inhibitory role in cell cycle progression and cell proliferation as phosphorylation of both

cyclin D1 and cyclin E induces their rapid degradation (Kang *et al.*, 2008). GSK3 activity is elevated in cells quiescent and in G1 phase, but is lowered when cells progress in S phase. GSK3 may also play a role in DNA repair after UV irradiation by inducing degradation of cdc25A, and allowing cells more time for DNA repair (Kang *et al.*, 2008). Over-expression of GSK3 and the consequent increase in its activity induce apoptosis in cells undergoing cellular stress (Cross *et al.*, 2001; Pap and Cooper, 1998). It looks like GSK3 is a master regulator, able to integrate both positive and negative cell stimuli through regulation of the stability and turnover of the protein targets.

There are several diseases that could benefit from GSK3 inhibitors based upon its role in different cellular processes. Lithium is an inhibitor of GSK3 used for the treatment of bipolar disorder and mood stabilization despite the potency in cells is only in the millimolar range. Lithium has a dual mechanism of inhibition: it inhibits GSK3 directly through competition with the  $Mg^{2+}$  coordinating the ATP and indirectly by increasing the phosphorylation of the inhibitory serine in the N-terminus (Jope, 2003). GSK3 has also been linked to other disorders of the nervous system, such as Alzheimer's disease caused by hyperphosphorylation of tau and formation of neurofibrillary tangles. Furthermore, elevated GSK-3 activity can cause overproduction of  $\beta$ -amyloid and the senile plaques observed in the Alzheimer patients (Martinez *et al.*, 2011). GSK3 inhibition has been also considered as therapeutic target for neuronal cell death, Parkinson's disease, Huntington's disease, transmissible spongiform encephalopathies and regulation of the circadian-clock (Meijer *et al.*, 2004). As the inhibition of GSK3 by protein kinase B is insulin mediated and results in activation of glycogen synthase with consequent increase in glucose utilization, pharmacological inhibition of GSK3 has also been considered for the treatment of type 2 diabetes. The possible role of GSK3 inhibition in cancer treatment is unclear: on one hand inhibition of GSK3 can cause activation of wnt signalling and accumulation of  $\beta$ -catenin,

responsible for many colorectal cancers and stabilization of cyclin D1 and C and Myc-1 all involved in tumorigenesis; on the other hand dual CDK-GSK3 inhibitors could synergistically induce apoptosis in prostate cancer through a mechanism that involves enhancement of TRAIL (TNF-related apoptosis-inducing ligand) apoptotic activity. It has also been reported that GSK3 plays a role in the mitotic spindle and GSK3 inhibition is responsible for arrest in prometaphase (Meijer *et al.*, 2004). The use of GSK3 in cancer therapy therefore requires the careful design of inhibitors that do not interact with the wnt pathway or are specifically directed towards certain cellular compartments (Meijer *et al.*, 2004).

Additional therapeutic areas for GSK3 have been identified in the treatment of osteoporosis and vascular calcification. These conditions could be treated by the activation and inhibition of the Wnt/ $\beta$ -catenin signalling pathway, respectively. Although activators of GSK3 could also play a role in the treatment of cardiac hypertrophy and cancer, no direct activators of GSK3 are under development (Takahashi-Yanaga *et al.*, 2013)

More than 30 GSK3 inhibitors have been reported (Meijer *et al.*, 2004); some of them have been co-crystallized with GSK3 $\beta$  in the ATP pocket (Bertrand *et al.*, 2003). The majority of GSK3 inhibitors are not selective and generally GSK3 inhibition is associated with CDK2 inhibition as well, as the two enzymes are very closely related in their kinase domain (Bain *et al.*, 2007; Meijer *et al.*, 2004).

Tideglusib, belonging to the chemical family of the thiadiazolidindiones (TDZDs), is the only GSK3 inhibitor currently in clinical development. Initially it was erroneously characterized as a non-competitive ATP inhibitor, but it has been recently demonstrated that it is in reality an irreversible inhibitor of GSK3 (Martinez *et al.*, 2002a; Dominguez *et al.*, 2012). Disappointingly the results for the completed phase II clinical trial for the treatment of Alzheimer's disease and progressive supranuclear palsy



```

HsGSK3_beta 1 MSG.....RPRTTFAES.....CKPVQQPSAFGSMKVS.....DKDGSKVTTVVATPGGGPDRPQEVSYTD 58
HsGSK3_alpha 1 MSGGGPSGGGPGGSGRARTSSFAEPGGGGGGGGGGGGPGGSASGPGGTGGGKASVGMGGGVGASSSGGGPGGSGGGGSGGPGAGTSFPPPGVKLGRD SGKVTTVVATLGQGPERSQEVAYTD 121
TbGSK3_short 1.....MSLNLTDAADDRSYKEMEKYTV 22
TbGSK3_long 1.....MSERILPSTLRGV TNG.....QKEVTASVGERVPLLPR.....RFSARPQGNQEAQERTAVKCEQVRYAI 60

HsGSK3_beta 59 TKVIGNGSFGVVYQAKLCDSGELVAIKKVLQDKRFKNRELQIMR.....KLDHCNIVRLRYFFYSSG--EK 122
HsGSK3_alpha 122 IKVIGNGSFGVVYQARLAETRELVAIKKVLQDKRFKNRELQIMR.....KLDHCNIVRLRYFFYSSG--EK 185
TbGSK3_short 23 ERVAGQGTFGTVQLARDKSTGSLVAIKKVIQDPRFKNRELQIMQ-HLA.....RLRHPNIVMLKNYFYTVGGEGR 91
TbGSK3_long 61 QEVIGRGAFGEVSSAEVVGTRDLVAIKRVIHDGRLRQRELTLMR-DHLGPNTQQGGVSSLDVGNVGVAHATSVTGSNGEEANGTTAIDGLESWNMPTIVPYHPCVVKLLDHFASD...P 176

HsGSK3_beta 123 KDEVYLNLLDYVPETVYRVARHYSR-AKQTLPIYVKLYMYQLFRSLAYIHS-FGI-CHRDIKPQNL LLDPD TAVLKLCDFGSAKQLVVRG.....EPNVSYICSRYYRAPELIFG-AT 232
HsGSK3_alpha 186 KDELYLNLLVLEYVPETVYRVARHFTK-AKLTIPILYVKVYMYQLFRSLAYIHS-QGV-CHRDIKPQNL LLDPD TAVLKLCDFGSAKQLVVRG.....EPNVSYICSRYYRAPELIFG-AT 295
TbGSK3_short 92 RNDVYLNVMFVVPETLHRTCRNYR-RMTNPLILVKVFMFQLLSIACLHIVINI-CHRDIKPHNVLVDEQTGELKLCDFGSAKRLAAD.....EPNVAYICSRYYRAPELIFG-NQ 203
TbGSK3_long 177 SGVQYLFMMDYIPLDVRRLHHMFLRQREQQMPIILVKVIMFQLARALAF L HAR-GI-CHRDVVKPNNILVDQEGVVKLCDFGSAKQKMQAVGGEGPREKNVPIFSRYRAPEL L L G-SQ 293

HsGSK3_beta 233 DYTSSIDVWSAGCVLAELL LGQPI FPGDS-GVDQLVEIKVLGTP TREQIREMNPNT.....EFKFPQIKAHPWTKVFRPR...TPPE 312
HsGSK3_alpha 296 DYTSSIDVWSAGCVLAELL LGQPI FPGDS-GVDQLVEIKVLGTP TREQIREMNPNT.....EFKFPQIKAHPWTKVFKSR...TPPE 375
TbGSK3_short 204 FYTTAVDIWSVGCIFAEMLLGEPIFCGEN-TSGQLREIVKILGKPTKEELHKLNGSST.....EINAN-AKATPWENVFKQP...LP AE 282
TbGSK3_long 294 YHFHVDMMWAFGCVLAELL CGKVLFKGSSSTMDQLVEIKVLGKPSERELFALNPQSAGSALIRTWGDSHNASQLSPTPSGPLPSSNSANADYMQRRSAPRVKSLLLWVEVLPPN...TSQA 411

HsGSK3_beta 313 AIALCSRLLEYTP TARTLPLEACAH SFFDEL RDPNVKLPNGRDTPA-LFNFTTQELSSNPPLATILIPPHARIQAAASTPTNATAASDANTGDRGQTNNAASASASNST... 420
HsGSK3_alpha 376 AIALCSSLLEYTPSSRLSPLEACAH SFFDEL RCLGTL PNNRPLP-LFNFSAGELSIQPSLNAI LIPPHLRSPAGTTTTLTPSSQAL TETPTSSDWQST DATPTLTNS... 483
TbGSK3_short 283 VYDLGKIKFYVPDQRITPLDALCH PFFNELREPTTKLPSGNPLPAHL YQFTPDEV EAMTEAQREYLLKK..... 352
TbGSK3_long 412 ALSLI EQLLRYTPEERL TSAEVL EHVFFDEL FSDDARLPNGAPL PASMFOV TREAEILPPWLLERMAAAEGVAKGRELNQSATAPENAI..... 501

```

**Figure 1.8** Amino acid sequence alignment of *HsGSK3 $\beta$* , *HsGSK3 $\alpha$* , *TbGSK3short*, and *TbGSK3long*

The background colour varies from the least conserved residues (white) to the most conserved residues (dark blue).

have revealed that although well tolerated by the patients one year after administration, this molecule has failed to show any clinical efficacy (Tolosa *et al.*, 2014).

## 1.6 Glycogen synthase kinase as a drug target in kinetoplastida

Glycogen synthase kinase 3 is a conserved protein kinase across eukaryotes and has been proposed as a drug target against *Toxoplasma gondii* (Qin *et al.*, 1998), *Plasmodium falciparum* (Droucheau *et al.*, 2004), *Trypanosoma brucei* (Ojo *et al.*, 2008), and *Leishmania donovani* (Xingi *et al.*, 2009).

In *T. brucei* there are two isoforms of glycogen synthase kinase 3, *TbGSK3* short (Tb927.10.13780) and long (Tb927.7.2420), respectively 40 kDa and 55 kDa. Both isoforms are more similar to *HsGSK3 $\beta$*  than *HsGSK3 $\alpha$*  with an identity of 40.9% and 33.1% for *TbGSK3* short and *TbGSK3* long to *HsGSK3 $\beta$* , respectively. The two parasite isoforms share an identity of 30.6% between them (Ojo *et al.*, 2008). The identity is more over the kinase domain, while the amino-terminal and carboxyl-terminal domains are not very well conserved (**Figure 1.8**).

A phosphoproteomic study performed by Nett and colleagues (2009a) in bloodstream form cells has shown that *TbGSK3* short is phosphorylated in the activation loop at Tyr187, suggesting that the mechanisms of activation might be conserved between mammals and parasites. It is unclear whether the identified phosphorylation on Ser2 might mimic the mechanism of inhibition of GSK3 performed by the phosphorylated Ser9 in the N-terminus of *HsGSK3 $\beta$* . Recently Urbaniak and colleagues (2013), performing a global phosphoproteomic study in procyclic and bloodstream form, have added two additional phosphorylation sites for *TbGSK3* short on T6 and S41 and determined a differential protein expression in the two life stages (higher in bloodstream than procyclic form). The phosphorylation on the activation loop was also identified in the procyclic form only for *TbGSK3* long at Tyr277 (Nett *et al.*,

2009b). Tyrosine phosphorylation on GSK3 activation loop has also been found by phosphoproteomic studies in *Leishmania infantum* (Hem *et al.*, 2010) and *Trypanosoma cruzi* (Nakayasu *et al.*, 2009; Marchini *et al.*, 2011).

The substrate recognition patterns (S/TXXXphosphoS/T) are conserved considering that recombinant *TbGSK3* short (Ojo *et al.*, 2008) and *LdGSK3* short (Xingi *et al.*, 2009) could trans-phosphorylate a known primed GSK3 substrate based upon the GSK3 phosphorylation sites of mammalian glycogen kinase (Meijer *et al.*, 2004). In epimastigote *T. cruzi*, around 7% of the identified phosphoproteins were compatible with the GSK3 consensus sequence (Nakayasu *et al.*, 2009) confirming also in kinetoplastids the relevant role played by GSK3, as one of the kinases with most substrates in the cell (Linding *et al.*, 2007).

The role of *TbGSK3* is unknown. *TbGSK3* short RNAi has been associated with defects in growth in both bloodstream form *T. brucei* and in differentiation (Ojo *et al.*, 2008; Alsford *et al.*, 2011). On the other hand *TbGSK3* long RNAi did not show significant effect on growth in bloodstream form, but was responsible for a gain-of-fitness in procyclic indicating a differential role for this isoform of the enzyme (Ojo *et al.*, 2008; Alsford *et al.*, 2011). Recently Jones and colleagues (2014) have shown that *TbGSK3* short knockdown causes cell arrest with a delay in mitosis and/or defects in cytokinesis. These findings are supported by a study performed on *Leishmania donovani* that has established that *LdGSK3* short is involved in cell cycle control and apoptosis-like death. The same study has also determined that in promastigote (insect form) cells the enzyme is cytosolic during logarithmic growth, but localises to the nucleus and flagellum at the stationary-phase, whereas in axenic amastigotes (model system used to study the intracellular amastigotes infecting the human macrophages) the cytosolic localisation is more punctate (Xingi *et al.*, 2009). In *Plasmodium falciparum* the expression of GSK3 (*PfGSK3*) is predominant in the trophozoite stage; here the

protein is rapidly exported to the erythrocyte cytoplasm where it associates with vesicle-like structures called Maurer's clefts suggesting a function in protein trafficking (Droucheau *et al.*, 2004). An interesting speculative role for this kinase in *P. falciparum* is the regulation of the parasite cycle that occurs in a circadian-dependent manner (always multiple of 24 hours) (Droucheau *et al.*, 2004).

Up to date no crystal structure of *Tb*GSK3 short has been solved, but it has been possible to build homology models using the solved structures for *Hs*GSK3 $\beta$  (Ojo *et al.*, 2008; Oduor *et al.*, 2011; Woodland *et al.*, 2013) and *Lmaj*GSK3 (Ojo *et al.*, 2011). These computational studies have identified key differences in the parasite ATP binding site that could be exploited to develop selective inhibitors.

Non-specific inhibitors of *Hs*CDK2 and *Hs*GSK3 have shown cellular efficacy with some correlation with the potency *in vitro* against recombinant *Tb*GSK3 providing a preliminary indication of chemical inhibition of this enzyme in cells (Ojo *et al.*, 2008). Furthermore a focused screen of Pfizer proprietary compounds targeting *Hs*GSK3 $\beta$  has identified two compounds with 7-fold selectivity for *Tb*GSK3 short over *Hs*GSK3 $\beta$  and this achieved selectivity has been rationalised by a key difference in the gatekeeper residue of the active site (*Tb*M101/*Hs*L132) (Oduor *et al.*, 2011). This selectivity is in agreement with the bioinformatics analysis performed by Osolodkin and colleagues (2011) that has proposed to target the gatekeeper residue in order to achieve selectivity for GSK3 protozoal kinases. The substitution of the gatekeeper residue (Leu132Met) affects the volume of the binding site, and is present not only in *Trypanosoma*, but also in *Plasmodium*, *Leishmania* and *Toxoplasma* (Osolodkin *et al.*, 2011). The methionine gatekeeper is longer and more flexible than the leucine explaining the selectivity towards the human enzyme of the majority of known GSK3 inhibitors. On the one hand the longer lateral chain can obstruct inhibitor binding due to steric clash, but on the other hand smaller inhibitors may be able to make hydrogen bonds with the sulphur of

the methionine thereby preferentially inhibiting the protozoan kinases (Osolodkin *et al.*, 2011).

Unfortunately the selective compounds identified by Oduor and colleagues (2011) are quite promiscuous inhibitors of mammalian kinases with associated cellular toxicity and are unlikely to specifically target *TbGSK3* in cells.

Interestingly two distinct chemoproteomic studies, aimed at identifying the cellular targets of lapatinib, a kinase tyrosine inhibitor, (Katiyar *et al.*, 2013) and of hypothemycin, a covalent inhibitor of kinase containing a CDXG motif (Cys-Asp-Xaa-Gly) (Nishino *et al.*, 2013), have identified *TbGSK3* short among their targets *in vivo* in *T. brucei*. Both lapatinib and hypothemycin kill *T. b. brucei* in proliferation assays with potency of 1.5  $\mu$ M and 170 nM, respectively. As *TbGSK3* is not the only target identified, the cellular activity cannot be completely attributed to this enzyme. Nonetheless, these studies provide further evidence that this protein kinase can be targeted in the cells.

None of the compounds studied up to date are specific enough as inhibitors of *TbGSK3* short to be used as chemical probes to interrogate the still unknown biological functions of this kinase in *T. brucei* or selective enough towards *HsGSK3 $\beta$*  to be considered as starting points for medicinal chemistry optimization as anti-trypanosomal drugs.

## 1.7 Aims of the project

The objective of this project is to evaluate with a combination of chemical and genetic approaches Glycogen Synthase Kinase 3 (*TbGSK3*) as potential drug target in *Trypanosoma brucei*.

The specific goals of the project are to:

- (1) perform a biochemical characterization of the molecular target;
- (2) identify small molecule inhibitors from a medium throughput screening (MTS) campaign against the molecular target;
- (3) identify any correlations between inhibition of the enzymatic activity ( $IC_{50}$ ) with respect to inhibition of trypanosome proliferation ( $EC_{50}$ );
- (4) determine the mode of inhibition and binding kinetics of identified inhibitors (determination of  $IC_{50}$ ,  $K_i$ ,  $K_d$ ,  $k_{on}$  and  $k_{off}$ );
- (5) investigate potential off-target mode(s) of action and identify the molecular targets responsible for the efficacy *in vitro*;
- (6) establish whether or not the mode of action is on target against the whole parasite by over-expression or under-expression studies looking for shifts in inhibitor potency.
- (7) examine essentiality by gene knock-out.

## **Chapter 2 Materials and methods**

## 2.1 Materials

The kinase focused set was collated by the Computational Group at the University of Dundee (Brenk *et al.*, 2008). Further synthesis and chemical characterisation of inhibitors of *Tb*GSK3 short were performed by Dr Andrew Woodland and Dr Laura Cleghorn for series GSK3 07 (Woodland *et al.*, 2013) and by Dr Robert Urich for series GSK3 09 (Urich *et al.*, 2014). All the reagents were bought from Sigma Aldrich if not otherwise stated. All the solutions were prepared with milliQ water. All the reagents were bought at the highest grade of purity and stored according to the manufacturer's recommendations.

## 2.2 General molecular biology

### 2.2.1 Isolation of genomic DNA from bloodstream form *T. brucei*

Genomic DNA (gDNA) was isolated from bloodstream form (BSF) *T. brucei* *brucei* strain 427 Lister (S427) (~ 100 ml of  $1-2 \times 10^6$  cell  $\text{ml}^{-1}$ ). Cells were harvested by centrifugation (800 x g, 10 min at room temperature) and resuspended in 450  $\mu\text{l}$  of lysis buffer (10 mM Tris-HCl, pH 8.0; 25 mM EDTA; 100 mM NaCl; 0.5% (w/v) SDS, 0.1  $\text{mg ml}^{-1}$  proteinase K). The resuspended cells was incubated overnight at 56 °C. DNA was extracted with Tris-buffered, pH 8.0 phenol/chloroform/isoamyl alcohol (PCI, 25:24:1). The organic phase was separated from the aqueous one containing the DNA by two repeated cycles of extraction and centrifugation at 16,000 x g for 1 min. A final extraction was done with chloroform/isoamyl alcohol (CI, 24:1). Finally the extracted DNA was precipitated with 70% ethanol and let to dry prior to resuspension in 10 mM Tris-HCl, pH 8.5 and stored at 4 °C.



**Table 2.1 Primers used for recombinant protein expression in *E. coli***

Primer use	Name	Primer sequence (5'-3')	Restriction Endonucleases
Cloning in expression vectors			
1	<i>Tb</i> GSK3short_ORF_s	<u>g</u> cgcc <u>t</u> cgagATGTCGCTCAACCTTACCGATGC	<i>Xho</i> I
2	<i>Tb</i> GSK3short_ORF_as	<u>g</u> cgcc <u>g</u> gatccTTACTTCTTCAGCAGATACTC	<i>Bam</i> HI
3	<i>Tb</i> GSK3long_ORF_s	<u>g</u> cgcc <u>t</u> cgagATGAGTGAGCGGATTTTGCCGTCG	<i>Xho</i> I
4	<i>Tb</i> GSK3long_ORF_as	<u>g</u> cgcc <u>g</u> gatccTTATATCGCATTCTCCGGC	<i>Bam</i> HI
5	<i>Tb</i> GSK3short_K49A_s	AAGCACGGGGTCACTAGTAGCAATT <b>G</b> CAAAGGTGATACAAGATCCG	Mutation of
6	<i>Tb</i> GSK3short_K49A_as	CGGATCTTGTATCACCTT <b>TG</b> CAATTGCTACTAGTGACCCCGTGCTT	K49A
7	<i>Tb</i> GSK3short_K49M_s	CGGGGTCCTAGTAGCAATT <b>TG</b> AAGGTGATACAAG	Mutation of
8	<i>Tb</i> GSK3short_K49M_as	CTTGTATCACCT <b>CA</b> TAATTGCTACTAGTGACCCCG	K49M

Upper case letters refer to nucleotides corresponding to gene sequences in *T. brucei*; lower case refers to additional sequences used in generating constructs. The restriction sites are underlined. Bold case letters correspond to the mutation sites.

### **2.2.2 DNA amplification**

DNA was amplified by the polymerase chain reaction (PCR) using Platinum PCR SuperMix High Fidelity (Invitrogen) with proofreading ability according to the manufacturer's guidelines. A 50  $\mu$ l reaction contained approximately 100 pmol of gene-specific forward and reverse primers (**Table 2.1** and **Table 2.2**) and approximately 1-200 ng of genomic DNA or plasmid template. All PCR reactions were carried out in 200  $\mu$ l tubes (Thermo Scientific) in a Thermo-Hybrid MSB 0.2G thermal cycler. Generally the tubes were incubated at 94 °C for 2 min to denature the template and activate the enzyme. PCR amplification was generally performed in 30 cycles of denaturation (94 °C for 30 s), annealing (30 s 55-68 °C depending on the melting temperature of the primers) and extension (1 min at 68 °C per kb). A final step of extension for 10 min at 72 °C was included. PCR reaction products were stored at 4 °C.

### **2.2.3 Agarose gel electrophoresis and extraction of DNA**

DNA size and purity were analysed by agarose gel electrophoresis. Gels were prepared by dissolving 0.8% (w/v) agarose in TAE buffer (40 mM Tris-acetate, pH 8.0, 1 mM EDTA) containing 0.1  $\mu$ g ml<sup>-1</sup> ethidium bromide for DNA staining. Separation was performed in TAE buffer at 80 V. DNA was visualised using an UVP transilluminator. One Kb Plus DNA Ladder (Invitrogen) was used as standard for size determination of DNA bands. If required, the DNA band of interest was excised from the gel and purified with the QIAquick Gel Extraction Kit (Qiagen) as per manufacturer's guidelines.

**Table 2.2 Primers used for generation of knock-outs and rescue constructs**

Primer use	Name	Primer sequence (5'-3')	Restiction Endonucleases
Knock-outs			
9	<i>TbGSK3short_5UTR_s</i>	ataagaat <u>gcgccgc</u> GAGTGAACAAACACCTCCAAG	<i>NoI</i>
10	<i>TbGSK3short_5UTR_as</i>	<u>gtttaaact</u> acggaccgtcaagcttGTAGTGAATAGTGCATTTTG	<i>HindIII/PmeI</i>
11	<i>TbGSK3short_3UTR_s</i>	gacggtccgtaag <u>tttaacggatc</u> CATGGTCGTGTTTTAAGTTGTG	<i>PmeI/BamHI</i>
12	<i>TbGSK3short_3UTR_as</i>	ataagtaag <u>cgccgc</u> CGTGCCACACCTACTTGCTTC	<i>NoI</i>
Silent Mutation of <i>HindIII</i> site in <i>TbGSK3</i> short			
13	<i>TbGSK3short (T960C)_s</i>	GGGAGCCAACAACGAAGCTCCCCAGTGGTA	Silent mutation of <i>HindIII</i>
14	<i>TbGSK3short (T960C)_as</i>	TACCACTGGGGAGCTTCGTTGTTGGCTCCC	
Ectopic copy			
15	<i>TbGSK3short_ORF_s</i>	gcgcaagcttcATGTCGCTCAACCTTACCGATGC	<i>HindIII</i>
16	<i>TbGSK3short_ORF_s</i>	gcgcatATGTCGCTCAACCTTACCGATGC	<i>NdeI</i>
17	<i>TbGSK3short_ORF_as</i>	gcgcgatccTTTACTTCTTCAGCAGATACTC	<i>BamHI</i>
18	<i>TbGSK3short_ORF_as</i>	gcgcttaattaaCTTCTTCAGCAGATACTC	<i>PacI</i>
Silent Mutation of <i>HindIII</i> site in <i>TbGSK3</i> long			
19	<i>TbGSK3long (T501C)_s</i>	CCGTGCGTTGTGAAGCTCCTGGATCATTCTTCGC	Silent mutation of <i>HindIII</i>
20	<i>TbGSK3long (T501C)_as</i>	GCGAAGAAATGATCCAGGAGCTTCACAACGCACGG	
Ectopic copy			
21	<i>TbGSK3long_ORF_s</i>	gcgcaagcttcATGAGTGAGCGGATTTTGCCGTCG	<i>HindIII</i>
22	<i>TbGSK3long_ORF_as</i>	gcgcgatccTTATATCGCATTCTCCGGC	<i>BamHI</i>
23	<i>TbGSK3long_ORF_as</i>	gcgcttaattaaTATCGCATTCTCCGGC	<i>PacI</i>

Upper case letters refer to nucleotides corresponding to gene sequences in *T. brucei*; lower case refers to additional sequences used in generating constructs. The restriction sites are underlined. Bold case letters correspond to the mutation sites.

#### **2.2.4 Cloning of PCR products**

Following PCR (section 2.2.2.) and, when necessary, gel extraction (section 2.2.3.), blunt-end PCR products were usually sub-cloned into pCR-BluntII-TOPO vector (Invitrogen). Ligations were carried out for 30 min on ice, following the manufacturer's guidelines. Ligated TOPO vectors were immediately transformed into 50 µl TOP10 chemically competent cells, as described in section 2.2.5. Alternatively when the PCR products were obtained by primers with overhanging regions they were digested with the restriction endonucleases corresponding to their cloning sites (**Table 2.1** and **Table 2.2.**) and ligated in the vector according to section 2.2.8.

#### **2.2.5 Transformation of competent *E. coli* cell lines**

Transformation of plasmids was achieved by heat shock of the following competent cells: *Escherichia coli* strains TOP10 (Invitrogen) were used for cloning and sequencing, JM109 (Stratagene) for DNA amplification, XL10-Gold Ultracompetent (Stratagene) for ligation of expression plasmids, BL21 Star (DE3) (Invitrogen) for protein expression of *TbGSK3* short and long, and ArcticExpress (DE3) RIL (Stratagene) for expression of insoluble proteins at low temperature. Typically, 50 ng of plasmid DNA was added to 10 - 50 µl competent cells and incubated on ice for 10 min. Cells were then subjected to heat shock according to the manufacturer's conditions in a water bath at 42 °C, followed by a further incubation on ice for 2 min. SOC medium (250 µl, containing 0.5% (w/v) yeast extract, 2% (w/v) tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 20 mM MgSO<sub>4</sub> and 20 mM glucose) was added and the resulting culture was incubated at 37 °C for 1 h. Cells were then plated out on Luria Bertiani (LB, 0.5% (w/v) yeast extract, 1% (w/v) tryptone and 1% (w/v) NaCl, pH 7.5) agar plates supplemented with appropriate antibiotics. Plates were incubated at 37 °C overnight and colonies were picked and grown at 37 °C with agitation at 200 rpm in 10 ml LB

medium containing appropriate antibiotics. All the LB media and plates were prepared by the media kitchen of the University of Dundee.

#### **2.2.6 Purification and digestion of plasmid DNA**

Transformed cultures were incubated overnight at 37 °C with agitation at 200 rpm. Cells were harvested by centrifugation (3,800 x g, 15 min, room temperature) and plasmid DNA was purified using a QIAprep kit (Qiagen) according to the manufacturer's instructions. Plasmids were analysed by digestion with appropriate restriction enzymes. Typically, 10 units of each restriction enzyme (New England Biolabs, Promega or Fermentas) and 1 µl of the recommended 10-fold buffer were added to 200 ng plasmid DNA, with autoclaved milliQ water added to a final volume of 10 µl. The total amount of restriction enzymes was kept below 10% for both single and double digests. Reactions were incubated at 37°C in a water bath for 2 h and then analysed by agarose-gel electrophoresis (section 2.2.3.). For preparative purposes, DNA was digested in a larger volume (up to 100 µl) maintaining the equivalent ratio between DNA, restriction enzymes and buffer, the reactions were incubated overnight at 37 °C. If digested DNA was used for ligation, the restriction enzymes were inactivated according to the manufacturer's indications. Linearised plasmids were treated with Antarctic phosphatase (New England Biolabs) according to the manufacturer's indication and phosphatase inactivated for 5 min at 65 °C. Both plasmids and inserts were purified by gel extraction (section 2.2.3.) prior to ligation.

#### **2.2.7 DNA sequencing**

DNA sequencing was performed by DNA Sequencing and Services at the University of Dundee (<http://www.dnaseq.co.uk>), using Applied Biosystems Big-Dye

Terminator chemistry version 3.1 on an Applied Biosystems 3730 automated capillary DNA sequencer. Plasmid DNA samples were submitted at concentrations of 16 – 20 ng  $\mu\text{l}^{-1}$ .

### **2.2.8 Ligation**

Ligation reactions were carried out with gel-purified restriction fragments (section 2.2.6). Purified target vector (~50 ng) and insert were typically combined in molar ratios of 1:1, 1:3 and 1:5, respectively, in a 10- $\mu\text{l}$  reaction. To control for background vector self-ligation, 50 ng of vector was incubated in the absence of insert. Ligations were carried out in the presence of 1 U T4 DNA ligase (Promega or Roche) at RT for 5 h or at 4 °C overnight. Following incubation, the T4 DNA ligase was inactivated by heating at 65 °C for 10 min. The ligation reaction (1  $\mu\text{l}$ ) was transformed in 30 – 50  $\mu\text{l}$  XL10-Gold Ultracompetent cells (Stratagene) as in section 2.2.5. Single colonies of the transformed cells were used for plasmid purification (section 2.2.6.), and the accuracy of the assembled constructs was verified by restriction digest and DNA sequencing as previously described (section 2.2.6. and 2.2.7.).

### **2.2.9 Southern analysis of genomic DNA**

Genomic DNA isolated from wild-type and genetically modified *T. brucei* was analysed by Southern blot (Southern, 1975). DNA (~5  $\mu\text{g}$ ) was digested overnight at 37 °C in the presence of the restriction enzymes (New England Biolabs). The digested DNA was resolved on a 0.8% (w/v) agarose gel (section 2.2.3.) and depurinated in 0.25 M HCl for 10 min. Following equilibration in 0.4 M NaOH, DNA was transferred to a positively charged nylon membrane (GE Healthcare) by the downward capillary transfer method (Sambrook and Russell, 2001). When the transfer was complete, the membrane

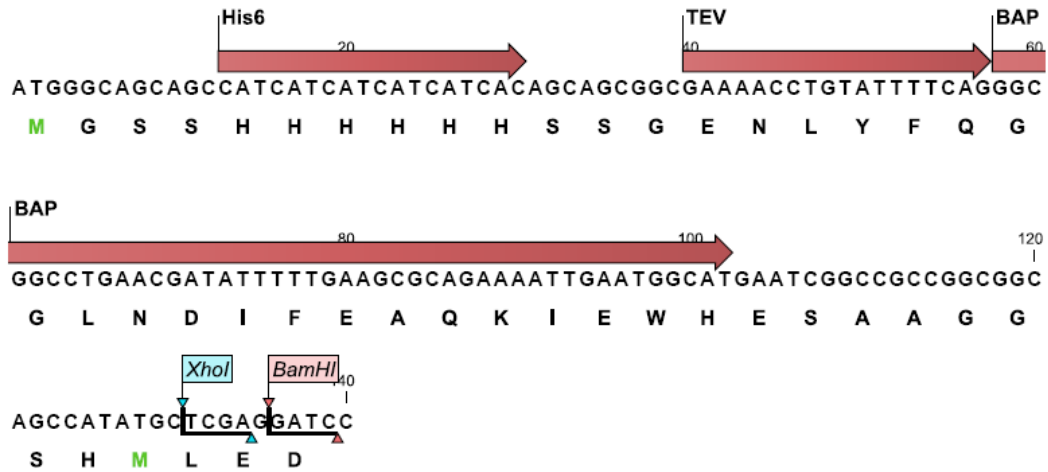
was rinsed in 5 X SSC buffer (supplied as 20 X buffer by Invitrogen) for 25 min and DNA was cross-linked to the membrane by UV irradiation at 1200  $\mu\text{J cm}^{-2}$ . The membrane was pre-hybridised at 42 °C for 2 h in DIG Easy Hyb (Roche) prior to overnight hybridisation at 42 °C with the denatured digoxigenin-dUTP-labelled DNA probe. Probes were prepared in advance from the open reading frame (ORF) of *TbGSK3* short or its 5'UTR, using PCR DIG Probe Synthesis Kit (Roche), following the manufacturer's guidelines.

Washing of the membrane post hybridisation, as well as the immunological detection of the probe, were performed with the DIG Wash and Block Buffer Set and the CSPD ready-to-use reagent (Roche), according to the manufacturer's guidelines. The processed membrane was exposed to Amersham Hyperfilm ECL (GE Healthcare) for 2 min. To strip the membrane for subsequent hybridisations, the membrane was rinsed in autoclaved ultrapure MilliQ water for 1 min and washed twice for 15 min at 37 °C with 0.2 M NaOH, 0.1% SDS. The stripped membrane was neutralised in 2 X SSC buffer and either stored at 4 °C or pre-hybridised and processed further, as described.

## **2.3 Recombinant protein expression and purification**

### **2.3.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)**

SDS-PAGE was used to determine the size and the purity of recombinant proteins (Laemmli, 1970). Equal volume of 2 X Laemmli buffer (Bio-Rad) containing 0.5% (v/v)  $\beta$ -mercaptoethanol (BME) was added to each sample and heated for 10 min at 70 °C prior to analysis. Electrophoretic separation of the bands was carried out using NuPAGE Novex 4-12% Bis-Tris mini Gels (Invitrogen) and the XCell SureLock Mini-cell apparatus at 200 V for approximately 50 min using MOPS SDS (Invitrogen) as



**Figure 2.1 pET-15b\_TEV\_BAP cloning/expression region**

Starting Methionine (M), linker (GSS), His<sub>6</sub> tag (HHHHHH), TEV protease recognition motif (ENLYFQG, the cleavage occurs between the Q and G residues), Biotin Acceptor Peptide (BAP) (GGLNDIFEAQKIEWH) and cloning sites (*XhoI* and *BamHI*).



running buffer. Protein bands were stained with Coomassie Blue (0.25 % (w/v) coomassie Brilliant Blue R250, 40% (v/v) methanol and 10% (v/v) acetic acid) and destained with 20% (v/v) methanol and 10% (v/v) acetic acid. Precision Plus (All blue, Bio-Rad) protein standard was used to determine the relative molecular mass ( $M_r$ ) of the resolved bands.

### 2.3.2 Protein quantification

Recombinant proteins and protein content from lysates were quantified using the Bio-Rad protein assay based on the Bradford method (Bradford, 1976) according to the manufacturer's instructions using BSA as standard. Absorbance at 595 nm was measured in a WPA Biowave II spectrophotometer. Recombinant proteins were also quantified using the calculated extinction coefficient (ProtParm) at 280 nm using the WPA Biowave II spectrophotometer.

### 2.3.3 Recombinant protein expression

The plasmid pET-15b-TEV-BAP was a modification of pET-15b (Novagen) where the sequence encoding the thrombin cleavage site was replaced with the Tobacco Etch Virus (TEV) protease recognition motif and the Biotin Acceptor Peptide sequence was also inserted (**Figure 2.1**). This plasmid was a generous gift of Ms Sharon Shepherd (University of Dundee). This expression plasmid allows the purification of the recombinant protein by metal-ion affinity chromatography using the Hexahistidine tag ( $\text{His}_6$ ), the cleavage by TEV protease that can be easily produced in house and the specific biotinylation at the N-terminus by the biotin-protein ligase BirA (Predonzani *et al.*, 2008). Expression constructs were transformed into the expression strains of *E. coli* indicated in **Table 2.3** as described in section 2.2.5. Single colonies were cultured as

**Table 2.3 Protein expression and purification conditions**

Protein	No	Expression conditions	Nickel affinity chromatography binding buffer <sup>a</sup>	Purification method <sup>b</sup>
<i>Tb</i> GSK3 short	1	BL21 Star (DE3) Auto-induction medium + Amp, 22 °C, 20 h	20 mM Hepes, pH 7.4; 0.5 M NaCl; 1 mM TCEP; 0.05% (w/v) CHAPS; 5% (v/v) glycerol; 30 mM imidazole	IMAC TEV cleavage IMAC SEC
<i>Tb</i> GSK3 long	2	BL21 Star (DE3) Auto-induction medium + Amp, 22 °C, 20 h	20 mM Hepes, pH 7.4; 0.5 M NaCl; 1 mM TCEP; 0.05% CHAPS; 5% (v/v) glycerol; 30 mM imidazole	IMAC TEV cleavage IMAC
<i>Tb</i> GSK3(K49A) short	3	BL21 Star (DE3) Auto-induction medium + Amp, 22 °C, 20 h; 37°C, 20 h	20 mM Hepes, pH 7.4; 20-500 mM NaCl; 1 mM TCEP; 0.05-0.5% CHAPS; 20-30 mM imidazole	IMAC: low level of expression
<i>Tb</i> GSK3(K49A) short	4	ArcticExpress (DE3) RIL Auto-induction medium + Amp, 13 °C, 48 h	20 mM Hepes, pH 7.4; 0.5 M NaCl; 1 mM TCEP; 0.05% CHAPS; 30 mM imidazole	IMAC IEX IMAC after ATP/MgCl <sub>2</sub> incubation
<i>Tb</i> GSK3(K49M) short	5	ArcticExpress (DE3) RIL Auto-induction medium + Amp, 13 °C, 48 h	20 mM Hepes, pH 7.4; 0.5 M NaCl; 1 mM TCEP; 0.05% CHAPS; 30 mM imidazole	IMAC IMAC with washes with ATP/MgCl <sub>2</sub> /denaturated protein from <i>E. coli</i> .

<sup>a</sup>Lysis buffers were obtained supplementing binding buffers with 500 μM activated sodium-orthovanadate, 500 μM sodium fluoride, 500 μM sodium-β-glycerophosphate, 500 μM sodium-pyrophosphate; EDTA-free protease inhibitor tablet (Roche) and DNase I (Roche).

<sup>b</sup> Metal-ion affinity chromatography (IMAC), size-exclusion chromatography (SEC), ion-exchange chromatography (IEX).

previously described (section 2.2.5.). Cultures were harvested by centrifugation (3800 x g, 15 min, RT) and resuspended in fresh LB containing 100  $\mu\text{g ml}^{-1}$  ampicillin. The cultures were then used to inoculate 1 L of pre-warmed auto-induction medium with ampicillin. BL21 Star (DE3) cells in auto-induction medium were incubated with agitation at 200 rpm for 3 h at 30 °C, followed by 24 h incubation at the selected temperatures (**Table 2.3**). ArcticExpress (DE3) RIL cells in auto-induction medium were incubated with agitation at 200 rpm for 6 h at 30 °C, followed by 48 h incubation at 13 °C. Following incubation, cells were harvested by centrifugation (4,500 x g, 30 min, 4 °C). Cell pellets were resuspended in lysis buffer (**Table 2.3**), either processed or flash frozen in liquid nitrogen and stored at -80 °C.

#### **2.3.4 Immobilised metal ion affinity chromatography**

Hexahistidine-tagged recombinant proteins were purified by immobilised metal ion affinity chromatography (IMAC). Cells were lysed using a Continuous One-Shot Cell Disruptor (Constant Systems) under a 30 kpsi pressure. Lysates were clarified by centrifugation (20,000 x g, 30 min, 4 °C) and supernatants were passed through a 0.22  $\mu\text{m}$  PES membrane filter (Helena Biosciences). A 5 ml HiTrap chelating HP column (GE Healthcare) was pre-equilibrated using a binding buffer identical to the composition of lysis buffer (10 X column volume) before the filtrate was passed through the column. The column was washed with 20 ml of binding buffer and connected to an AKTA FPLC purifier. Bound proteins were eluted in 2 ml fractions from the column using a 0-100% gradient of 500 mM imidazole in the binding buffer at a flow rate of 5 ml  $\text{min}^{-1}$ . Eluates were analysed by SDS-PAGE (section 2.3.1.). Fractions (2 ml) containing the highest concentrations of recombinant proteins were pooled together and dialysed (4 °C for 16 h) against binding buffer in absence of

imidazole in a 10-kDa molecular weight cut-off slide-A-lyzer cassette (Pierce Biotechnology) to remove imidazole.

Purified *TbGSK3* short and long were further cleaved of their hexahistidine tag (section 2.3.5). Further purification of contaminating proteins co-purifying with *TbGSK3*(K49A) and *TbGSK3*(K49M) short kinase dead mutants was attempted by anion exchange chromatography (section 2.3.6.) and further metal ion affinity chromatography after either incubation with the binding buffer supplemented with 1 mM ATP, 1 mM MgCl<sub>2</sub> or extensive washes with binding buffer containing 5 mM ATP, 10 mM MgCl<sub>2</sub> and 0.1 mg ml<sup>-1</sup> denaturated proteins from *E. coli* (Rial *et al.*, 2002).

### **2.3.5 TEV cleavage of hexahistidine-tag**

Hexahistidine tags were removed from recombinant *TbGSK3* short and long by TEV cleavage following nickel column affinity chromatography. Recombinant *TbGSK3* short and long were incubated with His<sub>6</sub>-tagged TEV protease (prepared by Mr Adam Roberts, University of Dundee), with a ratio of 10:1 (w/w) recombinant enzyme versus TEV protease at 4°C for 16 h, during dialysis in binding buffer in absence of imidazole. Following incubation, reactions were passed through a 1-ml HiTrap chelating HP column to remove free tags and the tagged proteins, which were bound to the column. The flowthrough containing the cleaved proteins was collected. *TbGSK3* long was dialysed in the storage buffer (20 mM Hepes, pH 7.4; 150 mM NaCl; 1 mM DTT), concentrated by centrifugation (3,800 x g, 4°C) using vivaspin 20 concentrators (Sartorius Stedim Biotech) with 10 kDa MWCO and polyethersulfone membranes. Concentrated *TbGSK3* long was quantified (section 2.3.2) and flash frozen in small aliquots following the addition of 10% (v/v) glycerol. The aliquots were stored at -80 °C. *TbGSK3* short was further purified by gel filtration chromatography (section 2.3.7).

### 2.3.6 Anion-exchange chromatography

Prior to anion-exchange chromatography pooled fractions of cleaved recombinant protein were desalted using PD-10 columns (GE Healthcare) according to manufacturer's guidelines and concentrated using a vivaspin 20 concentrator (Sartorius Stedim Biotech) (10 Kda MWCO). A 5-ml HiTrap Q HP column (GE Healthcare) was pre-equilibrated with 10 X column volume of the appropriate binding buffer. Desalted proteins were loaded manually onto the column using a syringe. The column was washed with 20 ml binding buffer and connected to the AKTA FPLC purifier. Bound proteins were eluted in 2 ml fractions from the column using a 0-100% gradient of 0.5 M NaCl in the binding buffer at a flow rate of 5 ml min<sup>-1</sup>. Eluates were analysed by SDS-PAGE and fractions containing the highest concentrations of pure recombinant proteins were pooled.

### 2.3.7 Gel filtration chromatography

Cleaved *Tb*GSK3 short was dialysed overnight at 4 °C in the gel filtration buffer (20 mM Hepes, pH 7.4; 150 mM NaCl; 1 mM DTT; 10% (v/v) glycerol) and separation by gel filtration was performed using HiLoad 26/60 Superdex 200 prep grade column. Column equilibration, automated sample loading and protein separation were carried out according to the manufacturer's guidelines. Fractions (2 ml) were collected and analysed by SDS-PAGE (section 2.3.1). Following each run, gel filtration standards (Bio-Rad) were applied to the column and resolved under the same conditions as the sample. To determine the molecular mass ( $M_r$ ) of the resolved proteins, a calibration curve was prepared by plotting log  $M_r$  for each gel filtration standard as a function of its elution volume. *Tb*GSK3 short was concentrated by centrifugation (3,800 x g, 4°C) using a vivaspin 20 concentrator (Sartorius Stedim Biotech) with 10 kDa MWCO.

**Table 2.4 Primary and secondary antibodies used for Western blotting analysis**

<b>Primary antibody</b>	<b>Immunogen</b>	<b>Dilution</b>	<b>Secondary antibody</b>	<b>Dilution</b>
Mouse anti-phospho-GSK3 (Tyr279/Tyr216) mAb, clone 5G-2F (Upstate, Millipore)	c-KQLLHGEPNVS[pY]ICSRI 203-219 from Drosophila GSK3/shaggy enzyme	1:375 or 1:500	Goat anti-mouse IgG, HRP-conjugated (Bethyl)	1:5,000
Mouse anti-GSK3 $\alpha/\beta$ mAb, clone 4G-1E (Upstate, Millipore)	c-KQLLHGEPNVSYICSRI 203-219 Drosophila GSK3/shaggy enzyme	1:375 or 1:500	Goat anti-mouse IgG, HRP-conjugated (Bethyl)	1:5,000
Rabbit anti-GSK3 $\alpha/\beta$ mAb (Upstate, Millipore)	Recombinant full length human GSK3 $\beta$	1:1000	Goat anti-rabbit IgG, HRP-conjugated (BioRAD)	1:3,000

Concentrated *TbGSK3* short was quantified (section 2.3.2.) and flash frozen in small aliquots. The aliquots were stored at -80 °C.

### **2.3.8 Protein characterization by mass spectrometry**

Identification of proteins by mass fingerprinting analysis was performed by the FingerPrints Proteomics Facility at the University of Dundee. Samples were submitted to the facility resolved on a NuPAGE Novex 4-12% Bis-Tris gel and stained with Coomassie blue (section 2.3.1). Protein bands of interest were excised from the gel and trypsin-digested, prior to analysis by nano-LC coupled to electrospray ionisation tandem MS (ESI-MS-MS), using the 4000 QTRAP (Applied Biosystems) tandem MS system.

### **2.3.9 Protein characterization by Western blotting**

Prior to analysis by Western blotting, unstained SDS-PAGE gels (section 2.3.1.) were equilibrated in Towbin transfer buffer (25 mM Tris, 192 mM glycine, 20% (v/v) methanol; Towbin *et al.*, 1979). Proteins were transferred onto Protran nitrocellulose membrane (Whatman) using Trans-Blot SD semi-dry electrophoretic transfer cell (Bio-Rad) at 25 V for 20 min. Membranes were incubated for at least 1 h with 5% (w/v) dry milk in PBS containing 0.5% (v/v) Tween 20 to block non-specific binding sites. Primary and secondary (horseradish peroxidase (HRP)-conjugated) antibodies (**Table 2.4**) were sequentially incubated with the blot for 1 h each in PBS containing 1% (w/v) dry milk and 0.5% (v/v) Tween 20, with an intermediate washing step (3 x 10 min) in antibody dilution buffer. Finally, membranes were washed (3 x 10 min) in buffer without milk. Washed blots were incubated with Amersham ECL or ECL Plus detection reagent (GE Healthcare) and exposed to Hyperfilm ECL (GE Healthcare), following the manufacturer's protocol.

## 2.4 Enzymatic and kinetic studies

### 2.4.1 *Tb*GSK3 biochemical characterization

*Tb*GSK3 short was sub-cloned and fused to an N-terminal maltose binding protein which was subsequently expressed and purified as previously reported in literature (Ojo *et al.*, 2008). For biochemical characterization of *Tb*GSK3 short, the kinase assay buffer (25 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 5 mM DTT, 0.02% (w/v) CHAPS, 2 U ml<sup>-1</sup> Heparin) and mix of ATP/[γ-<sup>33</sup>P]-ATP and GSP2 substrate (Biotin-C<sub>6</sub>-YRRAAVPPSPSLSAHSSPHQ[pS]EDEEE) (Pepceuticals) were used in a radiometric P-81 cellulose filterplate assay (Whatman).

### 2.4.2 Determination of the *Tb*GSK3 short $K_m$ for the substrates

The Michaelis-Menten constants ( $K_m$ ) for the ATP and the peptide substrate GSP2 were determined by varying the concentrations of both substrates in a matrix experiment. To determine bireactant kinetic parameters, the equation:

$$v = \frac{V_{\max} [A][B]}{\alpha K_a K_b + \alpha K_b [A] + \alpha K_a [B] + [A][B]} \quad (\text{equation 1})$$

was used, where  $v$  represents the measured velocity,  $V_{\max}$  is the maximum velocity,  $[A]$  and  $[B]$  and  $K_a$  and  $K_b$  are the concentrations and Michaelis Menten constants, respectively, of substrates A and B, and  $\alpha$  is the co-operativity factor between the two substrates (Segel, 1993). The  $\alpha$  parameter was fitted either as free or fixed to 1 and the best fit was assessed using the Akaike information criterion (Burnham and Anderson, 2002). Graphs and analysis were performed using SigmaPlot 10.0.



### 2.4.3 Determination of mechanism of inhibition

To establish mode of inhibition, rates were determined at 10 inhibitor concentrations with 4 varied concentrations of one substrate with a saturating concentration of the other. Each data set was individually fitted to the Michaelis-Menten equation, and the resulting Lineweaver-Burk plots were examined for diagnostic patterns for competitive (equation 2), mixed (equation 3), or non-competitive inhibition (equation 4).

$$v = \frac{V_{\max} \cdot [S]}{K_m \left(1 + \frac{[I]}{K_i}\right) + [S]}$$

Competitive inhibition (equation 2)

$$v = \frac{V_{\max} \cdot [S]}{K_m \left(1 + \frac{[I]}{K_i}\right) + \left(1 + \frac{[I]}{K_i'}\right) [S]}$$

Mixed inhibition (equation 3)

$$v = \frac{V_{\max} \cdot [S] \cdot \frac{1}{1 + [I]/K_i}}{K_m + [S]}$$

Non-competitive inhibition (equation 4)

Where  $v$  represents the measured velocity,  $V_{\max}$  is the maximum velocity,  $[S]$  is the concentration of substrate S and  $K_m$  is the Michaelis-Menten constant, and  $K_i$  and  $K_i'$  are the constants of inhibition (Segel, 1993). Graphs and analysis were done using Grafit 6.0.

### 2.4.4 Surface Plasmon Resonance

Amine-coupling reagents (*N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide [EDC], *N*-hydroxy-succinimide [NHS], and ethanolamine HCl) were purchased from Biacore AB.

The Running Buffer used in all SPR experiments was 25 mM Tris-HCl (pH 7.5), 150 mM NaCl, 10 mM MgCl<sub>2</sub> and 5 mM DTT plus 1% DMSO, and degassed and sterile-filtered through 20 µm filters by vacuum filtration apparatus (Millipore).

Interaction analyses were performed using Biacore T100 optical biosensor equipped with research-grade CM5 sensor chips (Biacore AB). Biacore instrument was primed with HBS-N buffer (Biacore) after removing maintenance chip and docking CM5 sensor chip. To prepare the sensor chip for interaction analysis, the dextran layer coating the CM5 chip surface was preconditioned with two consecutive injections for each of 100 mM HCl, 50 mM NaOH, and 0.5 % SDS. Aliquots of these preconditioning solutions were injected over all four flow cells for 6 s at a flow rate of 100 µL min<sup>-1</sup>. Data were always collected at the highest collection rate.

Coupling conditions were first optimised by preparing 1 µL of *Tb*GSK3 short in 50 µL of Immobilisation Buffer (Biacore) at two different pHs (5.5 and 4.5). The ligand solutions were injected over sensor chip surface using a contact time of 2 min at a flow rate of 10 µL min<sup>-1</sup>. Subsequently, a short pulse of 50 mM NaOH was injected over the surface at flow rate 100 µL min<sup>-1</sup> to remove last traces of the electrostatically bound ligand.

Using a flow rate of 10 µL min<sup>-1</sup>, the sensor chip's four flow cells were activated at 25 °C with a 7 min injection of a 1:1 mixture of EDC and NHS. For amine coupling, 1 µL of GSK-3 (concentration = 0.6 mg mL<sup>-1</sup>) along with 2 µL of 10 mM ATP and 100 µL acetate (pH 5.5) was injected at a flow rate of 10 µL min<sup>-1</sup> for 7 min. Finally, 1M ethanolamine was injected for 7 min to deactivate the surface. To test for analyte (ATP) binding, a threefold dilution series of six concentrations with a starting concentration of 500 µM was prepared and injected at flow rate of 10 µL min<sup>-1</sup> at 20 °C. One flow cell was used as a reference to subtract nonspecific binding, drift and bulk refractive index.

### **2.4.5 Isothermal Titration calorimetry**

Both *TbGSK3* short and *TbGSK3* long were extensively dialyzed against buffer containing 25 mM Tris-HCl (pH 7.5), 10 mM MgCl<sub>2</sub>, 1 mM DTT, 0.02% CHAPS and 1% DMSO. Experiments were carried out at 20 °C with a VP-ITC titration calorimeter (MicroCal Inc.) with 1 μM protein in the cell and 30 μM DDD00085893 in the syringe. Each titration experiment consisted of a 5 μl injection followed by 29 injections of 10 μl. Data were fitted to a single binding site model using Origin (version 7.0 MicroCal).

## **2.5 Screening**

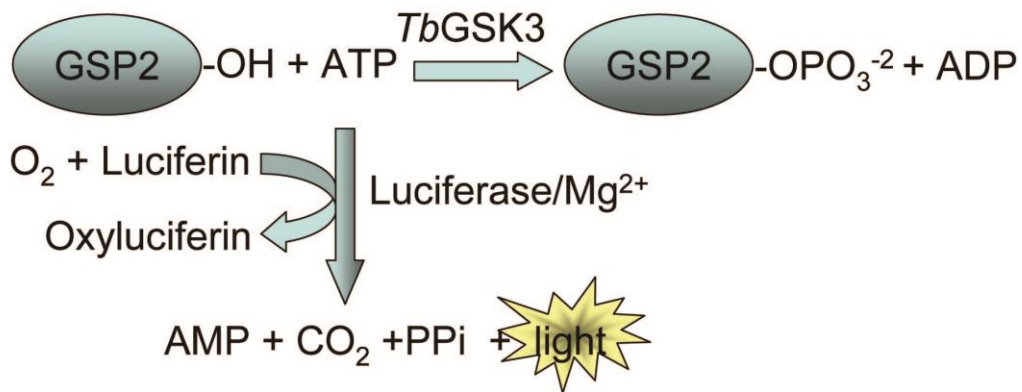
### **2.5.1 Compound library**

The kinase focused set library was collated from commercial sources. At the time of screening, the library included 4110 compounds, with purity confirmed to be higher than 90%. The criteria used for the library selection were previously described in literature (Brenk *et al.*, 2008). Briefly, the selection was based upon literature and patent searches of already known kinase scaffolds, which were then used to assemble the collection of commercially available inhibitors. The virtual set was filtered in order to eliminate non drug-like compounds, reactive or toxic groups. The final library contained 113 scaffolds that were represented with a minimum of 1 compound and a maximum of 50 examples per scaffold.

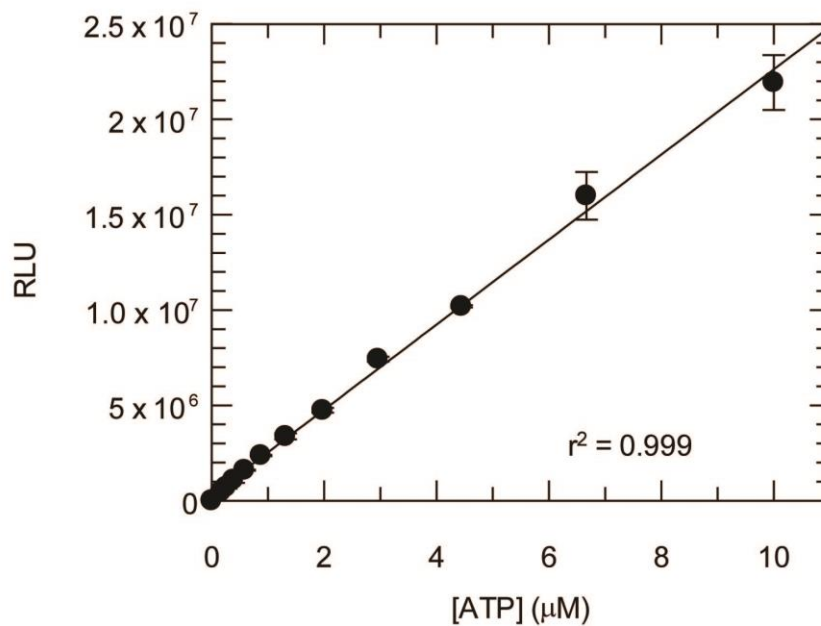
### **2.5.2 Screening workflow**

The kinase library was screened in a single point assay at 25 μM in a KinaseGlo format (Promega). Compounds showing a percentage of inhibition higher than 3-times the standard deviation (SD) of the high controls were cherry picked and re-tested in

A



B



**Figure 2.2 KinaseGlo assay principle**

(A) KinaseGlo is a luminescent-based assay that detects the amount of ATP left un-reacted at the end of the kinase reaction. The emission of light is due to the oxidation of the substrate luciferin to oxyluciferin by the addition of Luciferase enzyme in the presence of  $\text{Mg}^{2+}$ . (B) Linear regression of the titration curve of the ATP from 0 to 10  $\mu\text{M}$  up to 170 nM, linear regression and regression coefficient are shown.

double point at 25  $\mu\text{M}$ . The reconfirmed hits were tested in 10-data points curves from 30  $\mu\text{M}$  to 1 nM in two independent determinations in a Flashplate assay (PerkinElmer). The material from the top concentration of the curve was analysed by LC-MS in order to confirm structure and purity of the hits. Where possible, validated hits were repurchased or re-synthesised in house for the final validation. Only fully validated hits were further investigated. Data from the screening campaign were analysed in collaboration with a medicinal chemist and a computational chemist in the DDU (Dr Andrew Woodland and Dr Torsten Luksch, respectively)

### 2.5.3 Primary screening assay – KinaseGlo format

For the primary screening of the focussed kinase inhibitor library a 384-well KinaseGlo (Promega) luminescence-based assay previously described was used (Ojo *et al.*, 2008). The luminescence is inversely related to kinase activity. Unreacted ATP is used as substrate by Ultra-Glo Luciferase when Kinase-Glo Reagent is added to stop the kinase reaction, products of the reactions are the mono-oxygenation of luciferin and the generation of light (**Figure 2.2 A**). The ATP titration curve done in *TbGSK3* kinase buffer confirmed that the light emission was linear up to 10  $\mu\text{M}$  ATP (**Figure 2.2 B**). Luminescence is generally less prone to interference from library compounds than fluorescence-based assays, and in this particular assay format inhibitors of the Luciferase enzyme are not detected as false positives.

The reactions contained 7.5 nM *TbGSK3* short, 3.2  $\mu\text{M}$  GSP2 substrate peptide (Biotin-C<sub>6</sub>-YRRAAVPPSPSLSAHSSPHQ[pS]EDEEEE) (Pepceuticals), 1  $\mu\text{M}$  ATP and 25  $\mu\text{M}$  test inhibitor compound in optimised kinase assay buffer in a final volume of 30  $\mu\text{l}$ . DMSO and an 8-point titration of GW8510 (Sigma) from 1  $\mu\text{M}$  to 50 pM were included as controls. Reactions were incubated at room temperature for 1 h and stopped

by the addition of 30  $\mu$ l KinaseGlo reagent. Plates were then sealed and the signal was left to stabilize for 1 h in the dark before luminescence was determined using a TopCount NXT HTS counter (PerkinElmer).

#### 2.5.4 Potency screening assay – Flashplate assay

For hit validation and all subsequent compound potency determinations, a radiometric 96-well Flashplate assay (PerkinElmer) was adopted, with a final volume of 50  $\mu$ l per reaction. Compounds were solubilised in DMSO at a top concentration of 3 mM and serially diluted to achieve 10-point titration of final assay concentrations from 30  $\mu$ M to 0.3 nM with a final DMSO concentration of 1% (v/v). The reaction mixtures contained 1  $\mu$ M biotinylated GSP2 substrate (Biotin-C<sub>6</sub>-YRRAAVPPSPSLSAHSSPHQ[pS]EDEEEE), 1  $\mu$ M ATP, 3.7 KBq/well [ $\gamma$ -<sup>33</sup>P]-ATP and 2.5 nM *Tb*GSK3 short in the *Tb*GSK3 kinase assay buffer (25 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 5 mM DTT, 0.02% (w/v) CHAPS, 2 U ml<sup>-1</sup> Heparin). Inhibitors were screened for selectivity assessment also against *Hs*GSK3  $\beta$  and *Hs*CDK2/Cyclin A.

For *Hs*GSK3  $\beta$  assay the reaction mixes contained 1  $\mu$ M biotinylated GSP2 substrate, 2  $\mu$ M ATP, 7.4 KBq/well [ $\gamma$ -<sup>33</sup>P]-ATP and 15 nM *Hs*GSK3  $\beta$  in the *Tb*GSK3 kinase assay buffer. For *Hs*CDK2/cyclin A assay the reaction mixtures contained 1 mM CDK5 biotinylated peptide substrate (Biotin-C<sub>6</sub>-PKTPKKAKKL), 1  $\mu$ M ATP, 7.4 KBq/well [ $\gamma$ -<sup>33</sup>P]-ATP and 2 nM *Hs*CDK2/cyclin A in the kinase assay buffer (50 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 2 mM DTT, 100 mM NaCl, 0.2 mM EGTA, 0.02% (v/v) Brij35).

Reactions were incubated for a further 60 min at room temperature on a shaker before being terminated by the addition of 25 mM EDTA. Reaction mixtures were

transferred to 96-well streptavidin-coated Flashplates (PerkinElmer) and incubated for 60 min at room temperature on an orbital shaker. Plates were washed twice with PBS, 0.1% (v/v) Tween 20 using a microplate washer (Wellwash AC Thermo LabSystems), sealed and read using a TopCount NXT HTS counter (PerkinElmer).

### 2.5.5 Screening batch – pass criteria

During screening quality control (QC) plates were added as follows: 2 QC plates for every 6 assay plates for the primary screen and re-test screen; 2 QC plates for every 12 assay plates for the potency screen. Each assay plate had also internal QC represented by high (un-inhibited reaction) and low controls (no enzyme) and a standard inhibitor curve. Analysis of the QC data enabled the acceptance of a batch of plates based upon three main parameters: standard inhibitor IC<sub>50</sub> in a range of acceptance, Z'-factor (Z') value  $\geq 0.6$  (equation 5), coefficient of variation (CV)  $\leq 10\%$ .

$$Z' = 1 - \frac{(3\sigma_{High} + 3\sigma_{Low})}{|\mu_{High} - \mu_{Low}|} \quad (\text{equation 5})$$

Where Z' represents the Z'-factor,  $\sigma_{High}$  and  $\sigma_{Low}$  are the standard deviations of the high and low controls respectively, and  $\mu_{High}$  and  $\mu_{Low}$  are the average of the high and low controls respectively (Zhang *et al.*, 1999).

### 2.5.6 Mammalian kinase profiling

Selected compounds were screened against a panel of mammalian kinases routinely run by the Division of Signal Transduction Therapy (DSTT) at the University of Dundee in duplicate point at 10  $\mu\text{M}$ . Enzymes included in the panel and assay conditions are reported in literature (Bain *et al.*, 2007). All biochemical assays are run

below the  $K_m$  for the ATP for each enzyme allowing comparison of inhibition across the panel.

### 2.5.7 Data analysis

IC<sub>50</sub> values were determined using a 4-parameter equation in XLFit 4.2 model 205 (equation 6):

$$y = \frac{Top - bottom}{1 + \left( \frac{IC_{50}}{[I]} \right)^{Hillslope}} + bottom \quad (\text{equation 6})$$

The inhibition was background corrected and normalized for the un-inhibited reaction, bottom and top values were left free to float, but could be fixed at 0 and 100% inhibition respectively, if needed. In equation 6, y is the percentage of inhibition at inhibitor concentration [I]. IC<sub>50</sub> is defined as the concentration of inhibitor for which the activity of the enzyme is inhibited of 50% and Hill slope is a measure of the stoichiometry of the reaction. For 1:1 interaction between inhibitor and ligand a Hill slope factor of 1 is expected (Copeland, 2005).

### 2.5.8 Proliferation assay of MRC5 and BSF *T. brucei* cells

The routine screen of hits identified against BSF *T. brucei* and MRC5 (human lung fibroblast) was performed by the Drug Discovery Unit, and in particular Mr Iain Collie, Ms Bhavya Rao and Ms Irene Hallyburton. Mr Daniel James assisted with the data management and analysis. Measurement of inhibition of the proliferation of MRC5 cells was performed using a modification of cell viability assay previously described (Raz *et al.*, 1997). Compounds (50 μM to 0.5 nM, 0.5% DMSO final concentration) were incubated with 2 x 10<sup>3</sup> cells per well in a final volume of 200 μl in the appropriate culture medium (MEM with 10% foetal bovine serum) in clear 96 well plates (Greiner).



Plates were incubated at 37 °C in the presence of 5% CO<sub>2</sub> for 69 h. Resazurin was then added to a final concentration of 50 µM, and plates were incubated as above for a further 4 h before being read on a BioTek flx800 fluorescent plate reader. Measurement of inhibition of the proliferation of BSF *T. brucei brucei* strain 427 lister and analysis of the data was done accordingly to section 2.6.6.

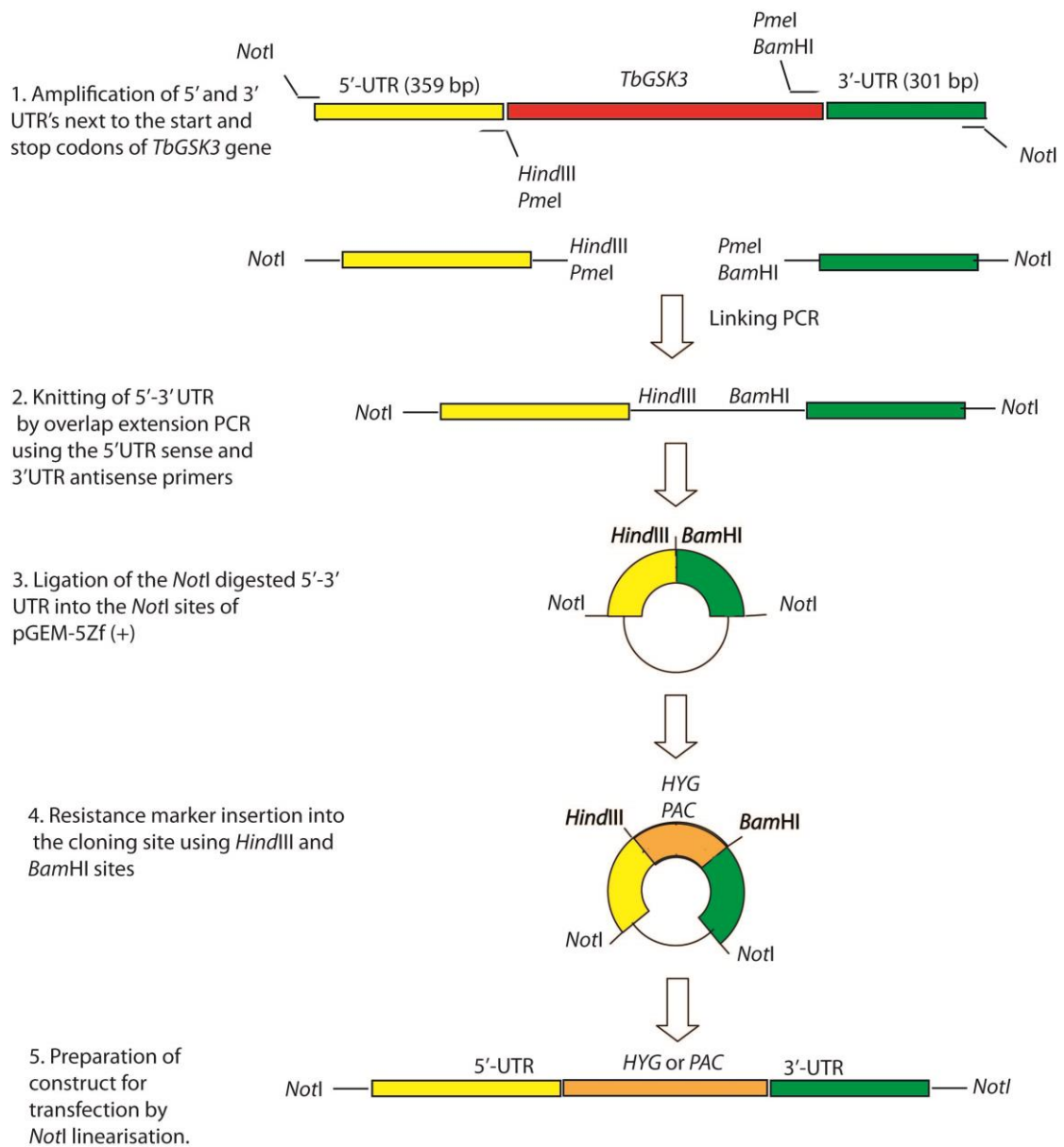
## **2.6 *T. brucei* cell culture and genetic manipulation**

### **2.6.1 Bloodstream form *T. brucei* cell culture**

BSF *T. b. brucei* strain Lister 427 (S427), previously genetically modified to express T7 RNA polymerase (T7RNAP) and the tetracycline repressor protein (TetR) (Wirtz *et al.*, 1999) was used as the parental cell line referred hereafter as wild-type (WT). This cell line is under G418-neomycin (*NEO*) selection and allows inducible expression of ectopic genes under the control of the T7 promoter and tetracycline operator. Wild-type and genetically modified *T. brucei* cell lines generated during this study were routinely diluted every time they reached a concentration of approximately  $2 \times 10^6$  cells ml<sup>-1</sup> in fresh modified HMI9-T medium (HMI-9 containing 56 mM thioglycerol in place of 200 mM 2-mercaptoethanol) supplemented with 2.5 µg ml<sup>-1</sup> of G418 to sustain the drug selection and the required drugs for the genetic modified cell lines. Cells were routinely cultured at 37 °C with 5% CO<sub>2</sub>. HMI-9T medium was routinely prepared by the Drug Discovery Unit, University of Dundee.

### **2.6.2 Cell density**

BSF *T. brucei* cells were counted manually using a Neubauer hemocytometer chamber under a light microscope (Zeiss). Alternatively, cells were counted using a CASY cell counter Model TT (Sharfe) according to the manufacturer's guidelines.



**Figure 2.3 Schematic representing cloning steps in the construction of knock-out cassettes**

### 2.6.3 Generation of *Tb*GSK3 short single knock-out cell lines

Single knock-outs (SKO) of *Tb*GSK3 short (*Tb*927.10.13780) were generated by homologous recombination (**Figure 2.3** and section 4.1.2). The replacement cassettes were generated by designing primers for the 3' and 5' untranslated regions (UTR) immediately before and after the open reading frame (ORF). Using the Tritryps.org genome browser application, the 5'-UTR region upstream and 3'-UTR region downstream of the codifying region of *Tb*GSK3 short were selected with a size of 359 bp and 301 bp, respectively. The 5'-UTR and 3'-UTR were PCR amplified from *T. brucei* S427 genomic DNA (section 2.2.1.) using *Tb*GSK3short\_5UTR\_s and *Tb*GSK3short\_5UTR\_as, *Tb*GSK3short\_3UTR\_s and *Tb*GSK3short\_3UTR\_as coupled primers, respectively (**Table 2.2**). The 5'-3'UTR region was subsequently knitted by overlap extension PCR (Roper *et al.*, 2002) of the two UTRs using the *Tb*GSK3short\_5UTR\_s and *Tb*GSK3short\_3UTR\_as primers and introduction of the linking region *Hind*III, *Pme*I and *Bam*HI. The resulting PCR product was inserted using the cloning *Not*I sites in a pGEM-5Zf (+) vector (Promega). Either the resistance genes hygromycin phosphotransferase (*HYG*) or puromycin acetyltransferase (*PAC*) were subsequently inserted into the cloning sites *Hind*III and *Bam*HI. DNA sequencing confirmed the accuracy of the final transfection constructs. Genetically modified cell lines were named accordingly the nomenclature defined by Clayton and colleagues (1998) (**Table 2.5**).

### 2.6.4 Generation of transgenic *T. brucei* cell lines

Wild-type BSF *T. brucei* cells were transfected using the Amaxa Nucleofector system (Lonza), according to Burkard (2007). Briefly, plasmid DNA was linearised by

**Table 2.5 Lists of mutant cell lines generated in BSF *T. brucei***

<b>Genotype<sup>a</sup></b>	<b>Restriction sites<sup>b</sup></b>	<b>Markers</b>	<b>Abbreviations<sup>c</sup></b>
<i>TetR NEO GSK3short /GSK3short</i>		<i>NEO</i>	WT
<b>pGEM-5Zf(+): knock-outs</b>			
<i>Δgsk3short::PAC/GSK3short</i>	<i>HindIII/BamHI</i>	<i>NEO, PAC</i>	SKO (PAC)
<i>Δgsk3short::HYG/GSK3short</i>	<i>HindIII/BamHI</i>	<i>NEO, HYG</i>	SKO (HYG)
<b>pLew82: leaky over-expression</b>			
<i>GSK3short<sup>Ti</sup></i>	<i>HindIII/BamHI</i>	<i>NEO, BLE</i>	WT <sup>OE</sup> <i>GSK3short<sup>Ti</sup></i>
<i>GSK3short<sup>Ti</sup> Δgsk3short::PAC/GSK3short</i>	<i>HindIII/BamHI</i>	<i>NEO, PAC, BLE</i>	SKO <sup>OE</sup> <i>GSK3short<sup>Ti</sup></i>
<b>pLew100v5: conditional over-expression</b>			
<i>GSK3short<sup>Ti</sup></i>	<i>HindIII/BamHI</i>	<i>NEO, BLE</i>	cWT <sup>OE</sup> <i>GSK3short<sup>Ti</sup></i>
<i>GSK3short<sup>Ti</sup> Δgsk3short::PAC/GSK3short</i>	<i>HindIII/BamHI</i>	<i>NEO, PAC, BLE</i>	cSKO <sup>OE</sup> <i>GSK3short<sup>Ti</sup></i>
<i>GSK3short<sup>Ti</sup> Δgsk3short::HYG/GSK3short</i>	<i>HindIII/BamHI</i>	<i>NEO, HYG, BLE</i>	cSKO <sup>OE</sup> <i>GSK3short<sup>Ti</sup></i>
<i>GSK3short<sup>Ti</sup> Δgsk3short::PAC/Δgsk3short::HYG</i>	<i>HindIII/BamHI</i>	<i>NEO, PAC, BLE, HYG</i>	cDKO <sup>OE</sup> <i>GSK3short<sup>Ti</sup></i>
<i>GSK3(K49A)short<sup>Ti</sup></i>	<i>HindIII/BamHI</i>	<i>NEO, BLE</i>	cWT <sup>OE</sup> <i>GSK3(K49A)short<sup>Ti</sup></i>
<i>GSK3(K49A)short<sup>Ti</sup> Δgsk3short::PAC/GSK3short</i>	<i>HindIII/BamHI</i>	<i>NEO, PAC, BLE</i>	cSKO <sup>OE</sup> <i>GSK3(K49A)short<sup>Ti</sup></i>
<i>TbGSK3long<sup>Ti</sup></i>	<i>HindIII/BamHI</i>	<i>NEO, BLE</i>	cWT <sup>OE</sup> <i>GSK3long<sup>Ti</sup></i>
<b>pLew100: conditional expression</b>			
<i>GSK3short<sup>Ti</sup></i>	<i>NdeI/BamHI</i>	<i>NEO, BSD</i>	cWT <i>GSK3short<sup>Ti</sup></i>
<i>GSK3short<sup>Ti</sup> Δgsk3short::PAC/GSK3short</i>	<i>NdeI/BamHI</i>	<i>NEO, BSD, PAC</i>	cSKO <i>GSK3short<sup>Ti</sup></i>
<i>GSK3short<sup>Ti</sup> Δgsk3short::HYG/GSK3short</i>	<i>NdeI/BamHI</i>	<i>NEO, BSD, HYG</i>	cSKO <i>GSK3short<sup>Ti</sup></i>
<i>GSK3short<sup>Ti</sup> Δgsk3short::PAC/Δgsk3short::HYG</i>	<i>NdeI/BamHI</i>	<i>NEO, BSD, PAC, HYG</i>	cDKO <i>GSK3short<sup>Ti</sup></i>

<sup>a</sup> Genotype is defined according Clayton and colleagues (1998) rules for genetic nomenclature for *Trypanosoma* and *Leishmania*.

<sup>b</sup> The restriction sites refer to the cloning sites for the plasmid of interest. All constructs were linearised by *NotI*.

<sup>c</sup> The abbreviations are defined upon the degree of gene over-expression and control of tetracycline induction.

*NotI* digestion, purified by precipitation in 70% ethanol (v/v) containing 80 mM sodium acetate, and resuspended in autoclaved MilliQ water to a final concentration of  $1 \mu\text{g} \mu\text{l}^{-1}$ . For a single transfection  $10 \mu\text{g}$  of DNA was mixed with  $4 \times 10^7$  cells (logarithmic phase of growth) and  $100 \mu\text{l}$  of Amaxa Nucleofactor system (Lonza). The electroporation was performed using the program X-001 of the Nucleofector II electroporator and cells were subsequently incubated in pre-warmed medium overnight (section 2.6.1). Approximately 16 h later, the selection drugs were added ( $5 \mu\text{g} \text{ml}^{-1}$  puromycin (*PAC*), hygromycin (*HYG*), phleomycin (*BLE*), or blasticidin (*BSD*)) and cultures were diluted 25- and 625-times and incubated in 24-well plate for 4-5 days. Transfected cells with tetracycline inducible expression vectors were cultured in medium supplemented with  $1 \mu\text{g} \text{ml}^{-1}$  tetracycline added daily in order to induce the exogenous protein expression. **Table 2.5** reports the genotype, the cloning sites, the markers used for selection and the abbreviated name of all generated BSF *T. brucei* cell lines.

### 2.6.5 Cloning of genetic modified *T. brucei* lines

All mutant *T. brucei* cell lines were cloned by limiting dilution. After drug selection and two passages in fresh medium, parental cell lines were diluted to  $1.5 \text{ cells} \text{ml}^{-1}$  and plated in 96-well plate ( $200 \mu\text{l}$  per well) in order to obtain  $\frac{1}{3}$  parasite per well in a medium supplemented with the relevant drug for selection and tetracycline was added daily where required. After 1 week, dividing cells were present generally in 30% of the wells, between 5 to 10 clones were selected for further analysis. Clonal cell lines were also stabulated at a density of  $1 \times 10^6 \text{ cell} \text{ml}^{-1}$  in HMI9-T medium containing 10 % (v/v) glycerol in liquid nitrogen.

### 2.6.6 Growth inhibition studies of *T. brucei*

The *in vitro* sensitivity of inhibitors to BSF *T. brucei* was measured at an effective compound concentration which inhibits cell growth by 50% (EC<sub>50</sub>). One hundred and ninety-nine µl of a cell suspension (2.5-5 [x 10<sup>3</sup>] cells per ml based upon cell growth regeneration time) were plated in 96-well plates. Inhibitors were diluted in 100% DMSO in a 10-point dilution curve (1:2 dilutions) and 1µl inhibitor was added to each well by a manual multichannel pipette. Each plate had 8 wells with high controls (DMSO only) and 8 wells with low controls (only medium). Generally pentamidine or melarsoprol were added as a standard control. After incubation for 69 h at 37 °C, 50 µM resazurin was added and the plates were incubated at 37 °C for additional 4 h. Fluorescence due to the formation of resorufin was measured at λ<sub>excitation</sub> of 528 nm and λ<sub>emission</sub> of 590 nm by a BioTek flx800 fluorescent plate reader.

Background corrected data and normalized with respect to the un-inhibited growth were analysed with a four- parameter non-linear regression using GraFit 6.0:

$$y = \frac{\text{range}}{1 + \left( \frac{[I]}{\text{EC}_{50}} \right)^{\text{Hillslope}}} + \text{bottom} \quad (\text{equation 7})$$

In equation 7 *y* is the percentage of inhibition at the corresponding inhibitor concentration [*I*], range is the difference between the high and low controls, bottom is the background and EC<sub>50</sub> is the effective dose of inhibitor required to inhibit the cell growth by 50%. Data are reported as the weighted mean ± SD of three independent determinations done in triplicate.

The weighted mean EC<sub>50</sub> and Hill slope coefficient were subsequently used to calculate the EC<sub>99</sub> (concentration of inhibitor required to kill 99% of cells):

$$\text{EC}_{99} = \text{EC}_{50} \times 99^{1/\text{Hill slope}} \quad (\text{equation 8})$$

### **2.6.7 Cell lysis for Western blotting**

Cells were pelleted at 800 g at room temperature for 10 min. The cell lysate was prepared as reported by Nett and colleagues (2009b). Briefly, cells were washed three times in ice-cold PBS prior to lysis with ice-cold RIPA (10 mM Tris-HCL, pH7.5; 1 mM Sodium- $\beta$ -glycerophosphate, 1 mM sodium pyrophosphate; 1 mM sodium fluoride, 5 mM EDTA, 0.5% NP40, 0.2% sodium deoxycholate, 0.2% SDS, EDTA-free protease inhibitor tablet (Roche), 100  $\mu$ M activated sodium-orthovanadate) at a ratio of  $1 \times 10^9$  cells per 1 ml of buffer. The lysate was sonicated in ice three times for 15 s at 80% power (Jencons) and centrifuged at 14,000 rpm for 20 min at 4 °C. The protein concentration of the supernatant was measured using the Bradford protein assay kit (Bio-Rad). Equal amount of protein lysate ( $\sim 10 \mu$ g) or cells number ( $\sim 1 \times 10^7$ ) were loaded per lane.

## **2.7 Chemical-proteomics**

### **2.7.1 Kinobeads method**

This method relies on the use of kinobeads immobilizing a mixture of seven protein kinase inhibitors that interact with a very broad spectrum of kinases and purine-binding proteins (Bantscheff *et al.*, 2007). The incubation of cell lysate with increasing concentration of an unmodified inhibitor results in competition between the inhibitor target and the kinobeads. This competition will decrease the iTRAQ (isobaric Tags for Relative and Absolute Quantitation) ion sensitivity for peptides derived from the target protein. The advantage of this method is that the inhibitor treatment of the cell lysate is done before the incubation with the kinobeads, so that is possible by competition to identify the inhibitor's targets and to estimate their  $IC_{50}$  values (White, 2007). More details regarding the methods are reported by Urbaniak and colleagues (2012). The chemical-proteomic validation was performed in collaboration with Cellzome using hit

compounds identified by the target-based screening campaigns I performed for *Tb*GSK3 short, PK50 and PK53.

## **2.8 Molecular modelling**

The homology modelling of *Tb*GSK3 short and the generation of binding modes of GSK3 07 and GSK3 09 compounds were performed by Dr Torsten Luksch and Dr Robert Urich, University of Dundee.

### **2.8.1 Homology modelling**

The homology model for *Tb*GSK3 short (Tb927.10.1378) was built using the human GSK3 beta crystal structures as templates (Protein Data Bank accession no. 1R0E). *Hs*GSK3 $\beta$  provided a template for 91% of the *Tb*GSK3 sequence (amino acids 20-348). ClustalW was used to generate the sequence alignments which were subsequently used for building the homology models with Modeller9-2 (Sali and Blundell, 1993).

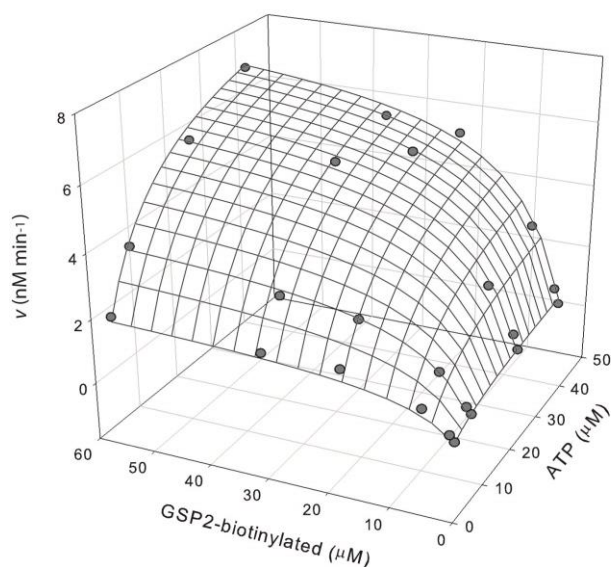
### **2.8.2 Generation of binding modes**

An energy-minimised conformation of DDD00065658 was built using Sybyl (Tripos Inc., Sybyl, St. Louis, MO (USA) 2004). This conformation was manually placed in the binding pocket of the *Tb*GSK3 short homology model with Moloc (Gerber and Müller, 1995) and subsequently minimised with the Moloc MAB force field. During the minimisation process the ligand was considered as fully flexible whereas all protein amino acids were kept rigid.

The putative binding modes of DDD00085893 and DDD00101234 were built in the same manner as described for DDD00065658.



## **Chapter 3 Results**



**Figure 3.1 Determination of the kinetic parameters for *Tb*GSK3 short**

The  $K_m$  for the ATP and the substrate (GSP2) were determined in a matrix experiment where the initial velocities ( $v$ ) were determined as nM of product formed per minute. The grid represents the predicted velocities fitted with a Random Mechanism with alpha equal to 1.

**Table 3.1 Determination of  $K_m$  for ATP and GSP2**

Substrate	MBP- <i>Tb</i> GSK3 short		His <sub>6</sub> - <i>Tb</i> GSK3 short
	$K_m$ ( $\mu\text{M}$ )	Published $K_m^{\text{app}}$ ( $\mu\text{M}$ ) (Ojo et al., 2008) <sup>a</sup>	Published $K_m$ ( $\mu\text{M}$ ) (Ojo et al., 2011) <sup>b</sup>
GSP2	$8.4 \pm 1.3^c$	2.4	$6.6 \pm 3.6$
ATP	$11.0 \pm 1.8^c$	4.5	$9.4 \pm 2.0$

<sup>a</sup>The  $K_m^{\text{app}}$  values for ATP and GSP2 were measured with a filterplate assay, it is not reported if the kinetic parameters were measured at saturating concentration of both substrates.

<sup>b</sup>The  $K_m$  values for ATP and GSP2 were measured with a filterplate assay by varying the concentrations of both substrates.

<sup>c</sup>The standard deviation reported is the error associated to the fitting by multi-regression analysis. The results reported are representative of one experiment, but the kinetic parameters were confirmed by multiple experiments.

### 3.1 Identification of small molecule inhibitors

#### 3.1.1 Biochemical characterization of *Tb*GSK3

The short form of the recombinant *T. brucei* Glycogen Synthase kinase 3 (*Tb*GSK3 short, *Tb*927.10.1378) was provided by the University of Washington as part of a collaboration with the Drug Discovery Unit (DDU) with the aim of discovering novel inhibitors. The recombinant enzyme was tagged at the N-terminus with a maltose binding protein (MBP), which simplified the purification procedure and enhanced the solubility of the protein. This enzyme had been previously used to screen a small set of *Hs*GSK3 inhibitors revealing some preliminary indication of correlation between enzyme potency and inhibition of *T. brucei* growth (Ojo *et al.*, 2008).

The substrate specificity of *Tb*GSK3 short shows similarity to the human isoform, *Hs*GSK3 $\beta$ , as it can phosphorylate primed substrates and recognises the same consensus sequence as the human enzyme: S/TXXXphosphoS/T (Dajani *et al.*, 2001). The substrate used for the biochemical assay was the peptide substrate GSP2 (biotin-C<sub>6</sub>-YRRAAVPPSPSLSAHSSPHQ[phosphoS]EDEEE). I determined the  $K_m$  values for the ATP and the peptide substrate in a matrix experiment by varying the concentrations of both substrates while measuring the rate of the reaction as nmole of product formed per minute. This experimental design permits the determination of the kinetic parameters, and also tests for co-operativity between substrates (**Figure 3.1**). The  $K_m$  for ATP and GSP2 were  $8.4 \pm 1.3 \mu\text{M}$  and  $11.0 \pm 1.8 \mu\text{M}$  respectively, with no evidence of co-operativity. These values are of the same order of magnitude as those previously published for the same recombinant MBP-*Tb*GSK3 short used here and a His<sub>6</sub>-tagged *Tb*GSK3 short (Ojo *et al.*, 2008; Ojo *et al.*, 2011) (**Table 3.1**). The lack of co-operativity between the two substrates indicates that the concentration of one substrate

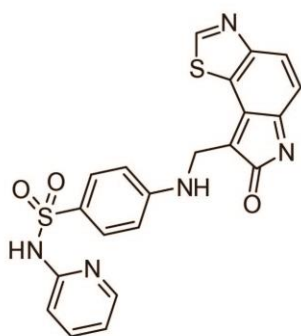
does not affect the  $K_m$  of the other one, allowing more flexibility in the set up of biochemical assays suitable for hit identification (Copeland, 2003; Yang *et al.*, 2009).

### 3.1.2 Primary Screen of *Tb*GSK3 short

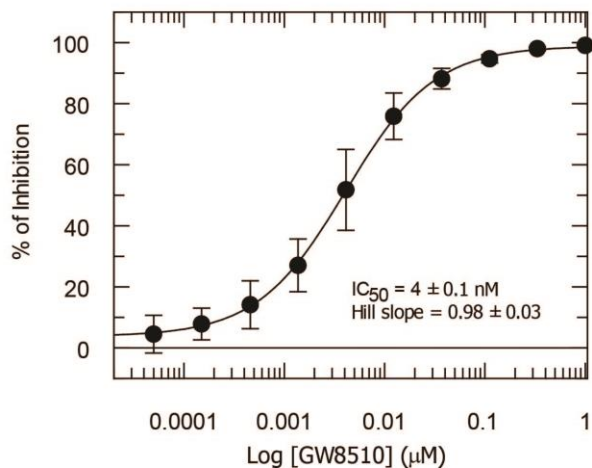
A focused kinase library of 4,110 compounds was assembled at the DDU based upon literature and patent review that had identified 113 scaffolds able to bind into the ATP pocket of kinases; for each scaffold up to 50 examples were purchased that fulfilled the criteria of lead-like properties and did not have unwanted functional groups (Brenk *et al.*, 2008). This set is biased towards the identification of ATP-competitive inhibitors, as it had been collated using already commercial available kinase inhibitor scaffolds (Brenk *et al.*, 2008). The primary screening assay I set-up was aimed to increase the sensitivity towards ATP-competitive inhibitors, i.e. the ATP concentration in the assay was fixed at 1  $\mu$ M, 10-fold below the  $K_m$  for the ATP of 11  $\mu$ M (Copeland, 2003).

KinaseGlo is a luminescent-based assay that detects the amount of ATP left unreacted at the end of the kinase reaction. The detection is attributed to the emission of light from oxidation of the substrate luciferin to oxyluciferin by the addition of Luciferase enzyme in the presence of  $Mg^{2+}$  (**Figure 2.1 A**). This assay format has many advantages: first, it does not require radiolabeled ATP or modified substrates, and, second, it is a homogeneous assay that does not require separation steps, hence it can be easily automated in microplate assays to increase the assay throughput. One disadvantage of this format is its reliance on the measurement of the concentration of the unreacted substrate, hence in order to obtain a reasonable signal to noise ratio (SN) it is necessary to ensure that the consumption of substrate is higher than 10%. This has implications for the accuracy of the potency of the identified hits, in particular with

A



B



**Figure 3.2 GW8510 standard inhibitor  $IC_{50}$  determination**

(A) GW8510 structure. (B) GW8510 prototypical inhibition curve (data are  $IC_{50} \pm SD$ , Hill slope  $\pm SD$ , of 27 independent determinations). GW8510 was used as positive control in all assay plates and its potency used as measure of the assay reproducibility.

**Table 3.2 *Tb*GSK3 short assay conditions and screening statistics**

	KinaseGlo assay	Flashplate assay
<i>Tb</i> GSK3 short	7.5 nM	2.5 nM
GSP2	3.2 μM	1 μM
ATP	1 μM ( $<K_m$ )	1 μM ( $<K_m$ )
Z <sup>a</sup>	0.61 ± 0.02 (n = 15) <sup>b</sup>	0.80 ± 0.01 (n = 34)
GW8510 $IC_{50}$	10 ± 0.05 nM (n = 15)	6.0 ± 0.7 nM (n = 34)

<sup>a</sup>Z' is a measure of the robustness of the assay (Zhang *et al.*, 1999).

<sup>b</sup>Values are reported as mean ± SEM and number of independent determinations (n).

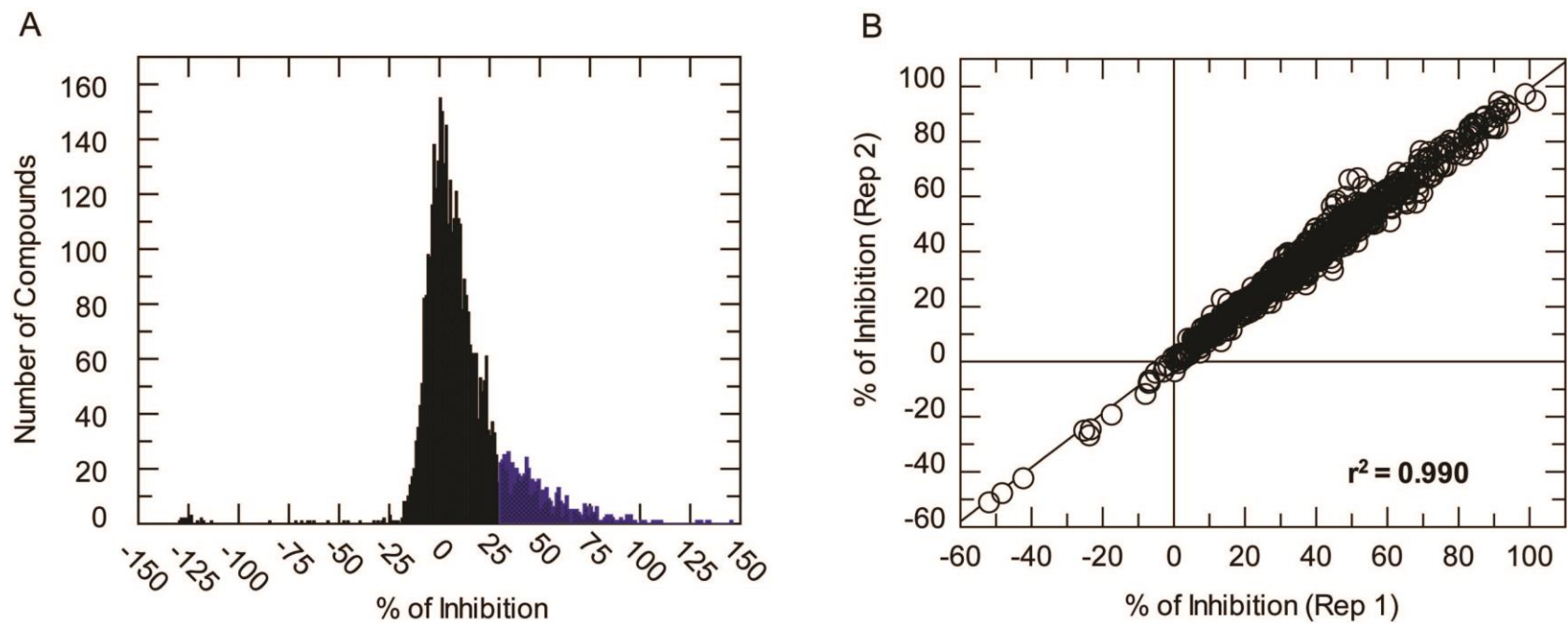
respect to the weaker inhibitors that could be missed as a consequence of deviation of enzyme linearity (Copeland, 2005).

In the primary screen, known inhibitor GW8510 (Ojo *et al.*, 2008) was included as a standard control and used as an indicator of the reproducibility of the assay conditions across all plates (**Figure 3.2** A and B). I screened the focused kinase set in 384-well assay format as a single point at 25  $\mu\text{M}$  generating a robust set of data with a  $Z'$  parameter above 0.6 (Zhang *et al.*, 1999) (**Table 3.2**).

From this primary screen, 567 compounds with percentage of inhibition higher than 30% were subsequently re-tested in a duplicate point screen, which gave 517 reconfirmed compounds with inhibition values higher than 30%, representing a hit rate of 12.8% (**Figure 3.3** A and B). The high hit rate was due to the fact that the library was biased towards kinase inhibitors and that the sensitivity of the assay was enhanced towards ATP-competitive inhibitors by the concentration of ATP in the assay set below its  $K_m$ .

### 3.1.3 Potency Screen of *Tb*GSK3 short

For hit validation and all subsequent compound potency determinations, I developed a radiometric 96-well Flashplate assay (PerkinElmer). Although the KinaseGlo format brings many advantages for screening of chemical libraries, its reliance upon ATP consumption means it requires a level of substrate consumption higher than 10% to achieve an acceptable signal window and hence it is not a suitable assay for accurate  $\text{IC}_{50}$  determinations. The alternative Flashplate assay measures the product formation relying on the transfer of  $^{33}\text{P}$ -labeled phosphate from ATP to the biotinylated substrate. The substrate is selectively bound to the streptavidinated plates, the unreacted  $[\gamma\text{-}^{33}\text{P}]\text{-ATP}$  removed by a series of washes and the emission of beta radiation is detected by addition of scintillation liquid.



**Figure 3.3 Primary screen**

(A) Distribution of the percentage of inhibition of focussed kinase set. Hits (highlighted in blue) were selected using as threshold 3 standard deviation units from the average of high controls ( $\geq 30\%$  inhibition). (B) Correlation plot of the 567 selected hits re-tested in duplicate.

The assay was performed at a concentration below the  $K_m$  for ATP of the enzyme, so that the  $K_i^{app}$  for ATP-competitive inhibitors approximates the measured  $IC_{50}$  value, aiding the assessment of selectivity (Cheng and Prusoff, 1973). I carried out potency evaluations (10-point curves) in duplicate for the 100 most potent compounds. As in the primary screen, the potency assay format yielded highly robust data ( $Z' = 0.80 \pm 0.08$ , **Table 3.2**). The compounds exhibited a range of potencies for *TbGSK3* short (**Table 3.3**), which were highly reproducible (**Figure 3.4**), with 15 compounds having  $IC_{50}$  values  $<1 \mu\text{M}$ . The structure and purities of the hits taken into potency assessment were subsequently confirmed by LCMS.

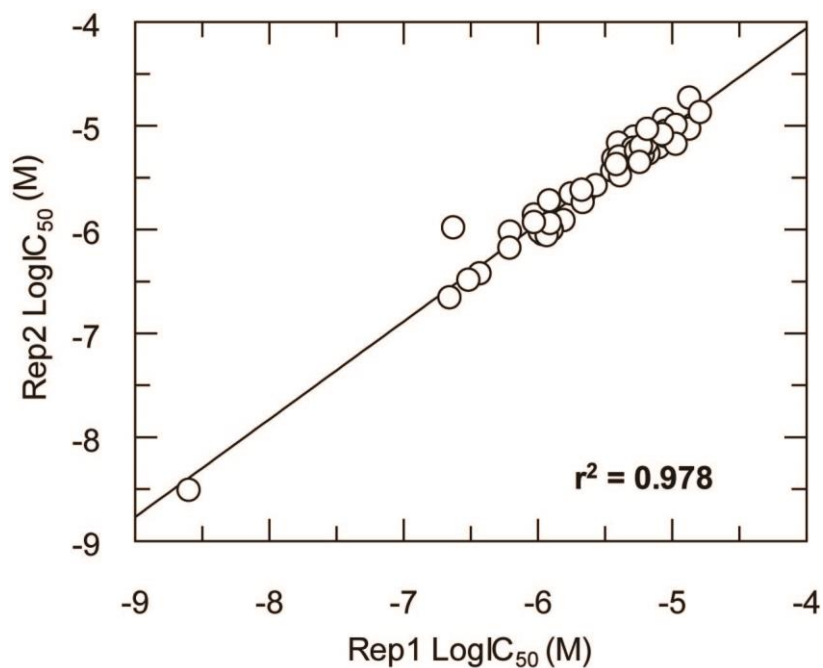
## 3.2 Hit validation

### 3.2.1 *TbGSK3* short series classification

The confirmed hits were classified by Dr Andrew Woodland into series based upon structure similarity and the most potent compound of each series was considered as its representative. From the screening of the recombinant *TbGSK3* short eight compound series were identified (GSK3 01-08), and additional examples of two inhibitors series, GSK3 09 and GSK3 10 were synthesised by Dr Robert Urich based on a combination of in-house expertise and literature data (Wyatt *et al.*, 2008). Both GSK3 09 and GSK3 10 series had highly potent inhibitors against *TbGSK3* short (**Figure 3.5**).

Where possible the hit compounds were repurchased and re-tested. In addition a programme of hit expansion, via purchasing of analogues or investigation of the presence of similar scaffolds in the DDU large diversity set (Brenk *et al.*, 2008) was conducted to expand on the preliminary structure-activity relationship (SAR). All the GSK3 series were assessed for chemical properties, evidence of an initial SAR in the biochemical assay, correlation of the activity against *TbGSK3* short versus the *T. b.*



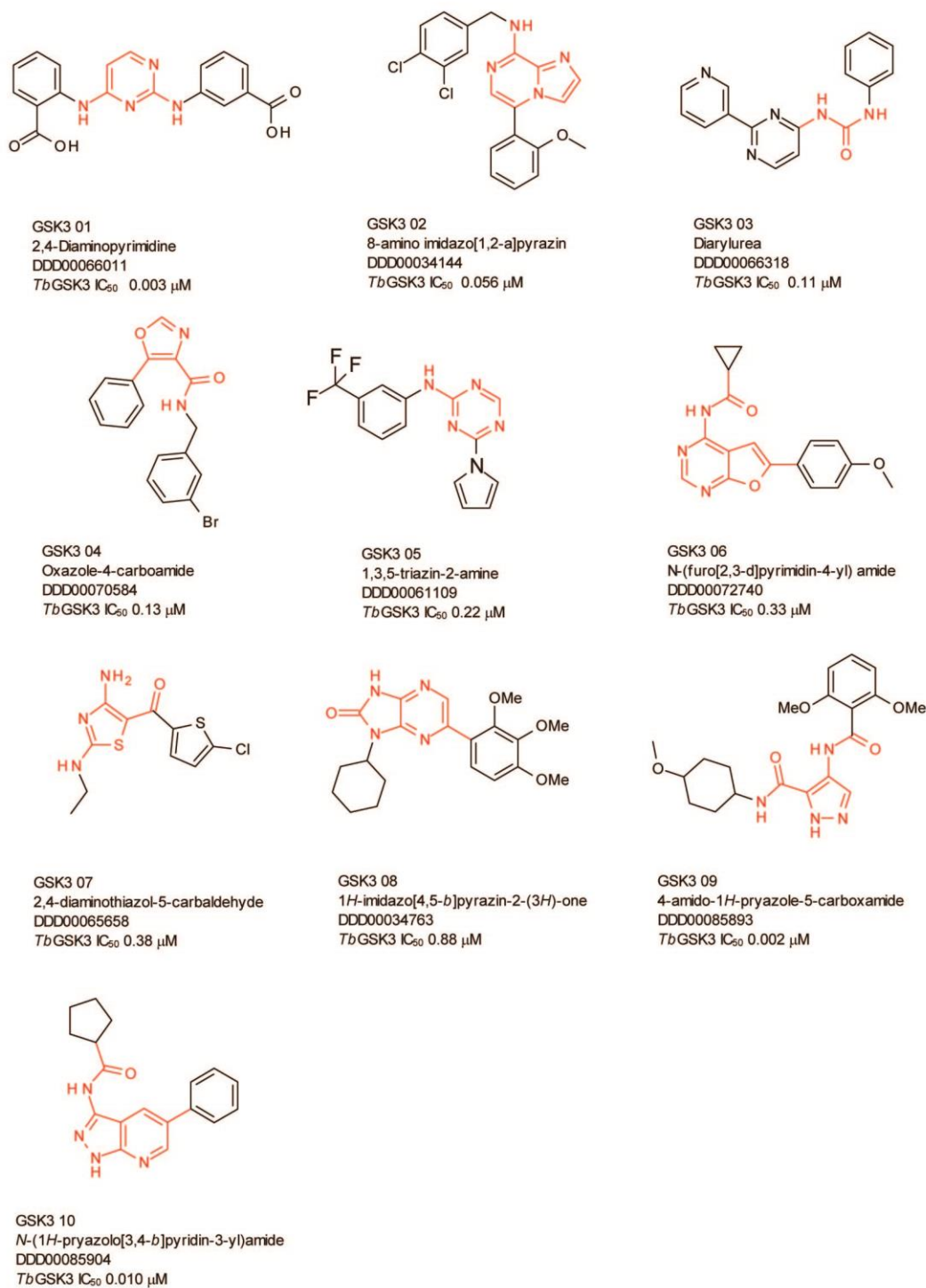


**Figure 3.4 Potency screen**

The 100 most potent hits were tested in 10-data point curve in 2 independent determinations, the correlation plot between the  $\text{LogIC}_{50}$  of the replicates is reported.

**Table 3.3 Ranking of the kinase set compound potencies**

$\text{IC}_{50}$ ( $\mu\text{M}$ )	Number of Compounds
<1	15
1-10	75
10-20	10
20-30	-



**Figure 3.5 Compound series identified from focused screening and testing of literature compounds**

Most active examples identified from the screen are shown with the core scaffold highlighted in red.

*brucei* proliferation assay, toxicity against MRC5 proliferation assay, and selectivity against a panel of mammalian recombinant kinases. In particular, two series, GSK3 07 and GSK3 09, were subjected to further studies (Woodland *et al.*, 2013, Urich *et al.*, 2014).

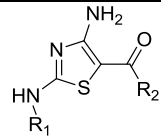

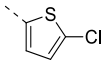

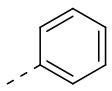
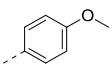
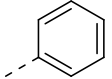
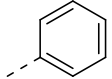
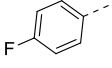
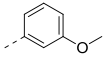
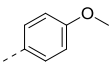
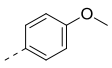
### 3.2.2 GSK3 07 series: correlation, specificity and selectivity

GSK3 07 is a series of 2,4-diaminothiazol-5-carbaldehydes (**Figure 3.5**) for which only two examples were present in the initial DDU focused kinase set. The most potent compound, DDD00065658 had an IC<sub>50</sub> value of 0.38 μM. This compound showed a high ligand efficiency of 0.52 kcal mol<sup>-1</sup> <sup>a</sup> and hence was deemed to be an interesting starting point for hit validation (Hopkins *et al.*, 2004). Using a combination of diversity and docking studies (section 3.2.7) 21 compounds were selected for testing from the DDU general screening and commercial sources. The IC<sub>50</sub> values were determined for *Tb*GSK3 short, *Hs*GSK3 β and *Hs*CDK2/cyclin A as well as EC<sub>50</sub> values for bloodstream form (BSF) *T. b. brucei* and MRC5 cell lines. The data for some key examples of GSK3 07 are summarised in **Table 3.4**. The hit expansion was successful with compound DDD00086390 (IC<sub>50</sub> value of 0.03 μM) which was over 10-fold more active than the initial hit DDD00065658 (IC<sub>50</sub> value of 0.38 μM). DDD00086390 showed also a greater activity for *Hs*GSK3 (IC<sub>50</sub> value of <10 nM) than for *Tb*GSK3 short which seemed to be a general trend for the GSK3 inhibitors from this series. GSK3 07 compounds generally showed good levels of selectivity for *T. b. brucei* versus MRC5 cells, giving preliminary indication that inhibition of *Hs*GSK3 could be tolerated

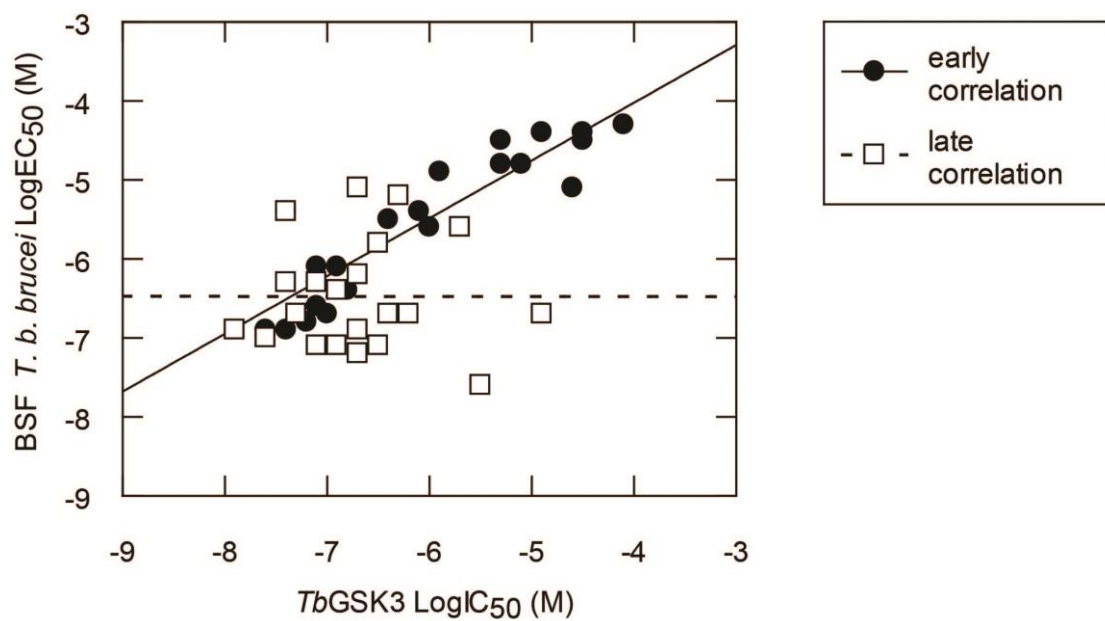
---

<sup>a</sup> Ligand efficiency is a measure of the affinity related to the MW: Ligand efficiency = [-RT Ln (IC<sub>50</sub>)] / N<sub>non-hydrogen atoms</sub>

**Table 3.4 GSK3 07 Selection of early hits with activities against *Tb*GSK3 short, *Hs*GSK3 $\beta$ , *Hs*CDK2, BSF *T. b. brucei* and MRC5 cells**

	$R_1$	$R_2$	<i>Tb</i> GSK3 short IC <sub>50</sub> ( $\mu$ M)	<i>Hs</i> GSK3 $\beta$ IC <sub>50</sub> ( $\mu$ M)	<i>Hs</i> CDK2 IC <sub>50</sub> ( $\mu$ M)	BSF <i>T. b. brucei</i> EC <sub>50</sub> ( $\mu$ M)	MRC5 EC <sub>50</sub> ( $\mu$ M)
DDD00065658			0.38	<0.010	0.15	2.6	>15
DDD00066736	Me		>30	ND	ND	38	>50
DDD00066742	Me		>30	ND	ND	29	>50
DDD00063206		Me	1.1	ND	ND	13	26
DDD00086390			0.03	<0.010	0.27	0.16	13
DDD00086393			0.04	<0.010	0.045	0.13	>50
DDD00086398			0.1	<0.010	0.72	0.7	4

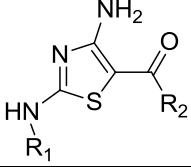
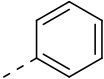
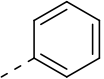
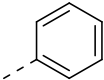
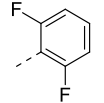
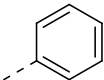
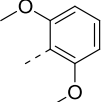
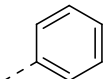
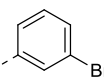
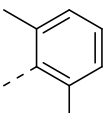
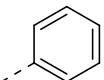
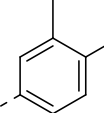
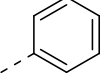
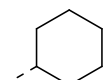
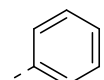
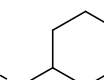
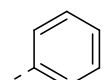
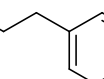
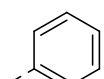
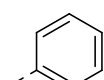

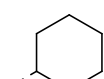

The number of determinations was two or higher. ND: not determined.



**Figure 3.6 Correlation plot for GSK3 07 compounds**

Correlation plot between inhibition of *TbGSK3* short and inhibition of bloodstream form *T. b. brucei* cell growth for the initial set of compounds (solid line) and for the later set (dashed line).

**Table 3.5 GSK3 07 Selection of late hits with activities against *Tb*GSK3 short, *Hs*GSK3 $\beta$ , *Hs*CDK2/cyclin A, BSF *T. b. brucei* and MRC5 cells**

	R <sub>1</sub>	R <sub>2</sub>	<i>Tb</i> GSK3 short IC <sub>50</sub> (μM)	<i>Hs</i> GSK3 $\beta$ IC <sub>50</sub> (μM)	<i>Hs</i> CDK2/ cyclin A %I At 1 μM	BSF <i>T. b. brucei</i> EC <sub>50</sub> (μM)	MRC5 EC <sub>50</sub> (μM)
DDD00086390			0.03	0.007	91%	0.2	13
DDD00088340			0.02	<0.01	97%	0.12	1.6
DDD00088327			0.4	2.5	31%	0.4	20
DDD00088328			0.1	<0.01	90%	0.2	>50
DDD00088331			0.2	0.13	17%	9	35
DDD00088337			0.06	<0.01	84%	0.18	38
DDD00088338			0.19	<0.01	91%	0.13	11
DDD00088336			0.7	<0.01	30%	0.4	41
DDD00088334			0.5	<0.01	65%	5.7	38
DDD00088342			3.4	0.5	21%	0.04	20
DDD00090123			12	ND	ND	0.30	>50

The number of determinations was two or higher. ND: not determined.

to some extent by these cells. It was also observed that it was possible to achieve selectivity over *Hs*CDK2, a desirable feature considering the similarity between these enzymes and potential toxicity issues related to inhibition of *Hs*CDK2 (Meijer *et al.*, 2004) (**Table 3.4**).

For the early compounds investigated from this series cell efficacy (BSF *T. brucei* LogEC<sub>50</sub>) correlated with the enzyme potency (*Tb*GSK3 LogIC<sub>50</sub>) (correlation coefficient equal to 0.90) with a 5-fold drop off between the *Tb*GSK3 short and *T. brucei* activities (**Figure 3.6**). For ATP-competitive inhibitors a drop-off of activity is expected between the IC<sub>50</sub> values measured in a biochemical assay, where the concentration of ATP is below its *K<sub>m</sub>*, and the EC<sub>50</sub> values measured in the cellular assay, where the physiological concentration of ATP is in the millimolar range (Cheng and Prusoff, 1973; Knight and Shokat, 2005). Considering that in the case of *Tb*GSK3 short the *K<sub>m</sub>* for ATP determined is 11 μM, the expected drop-off between biochemical and cellular assay should be around 100 fold (Knight and Shokat, 2005). Indeed, other factors such as protein binding or the requirement for a high level of inhibition of the enzyme to achieve a phenotypic effect, could actually result in a drop off greater than 100-fold. The 5-fold difference between potency against the enzyme and the cell activity for this series was inconsistent with an exclusive ATP-competitive mechanism of inhibition of *Tb*GSK3 short as the direct cause of the cell activity.

Further optimisation within this series was designed to improve inhibitor potency. The newly synthesised compounds were also tested against *Tb*GSK3 short, *Hs*GSK3, *Hs*CDK2/cyclin A, BSF *T. b. brucei* and MRC5 cells (**Table 3.5**). Compound DDD00088327 showed 6-fold selectivity towards *Tb*GSK3 short (IC<sub>50</sub> 0.4 μM) with respect to *Hs*GSK3 β (IC<sub>50</sub> 2.5 μM), giving an indication that it was possible to obtain compounds more potent towards the parasite isoform of the enzyme. The most significant development was the diverging correlation between *Tb*GSK3 potency and *T.*

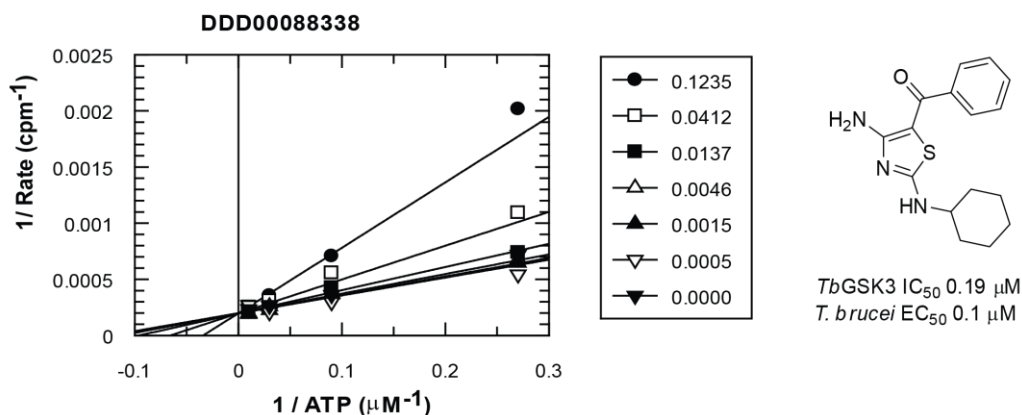
**Table 3.6 Mammalian kinase selectivity panel for series GSK3 07**

Enzyme	DDD000086390	DDD000086405	DDD000086406	DDD000086411	DDD000088338	DDD000088340	DDD000090123		Enzyme	DDD000086390	DDD000086405	DDD000086406	DDD000086411	DDD000088338	DDD000088340	DDD000090123
MKK1	0	74	84	94	68	96	94		AMPK	41	74	78	85	96	67	101
ERK1	59	91	66	127	20	72	73		MARK3	34	93	59	89	100	25	112
ERK2	26	58	58	59	16	68	95		BRSK2	42	79	53	62	93	55	95
JNK1	36	67	54	103	60	75	98		MELK	34	57	79	42	70	47	89
JNK2	48	65	69	88	39	81	90		CK1	14	37	91	91	82	13	74
p38 $\alpha$ MAPK	14	16	48	47	24	8	102		CK2	107	108	99	102	114	111	118
p38 $\beta$ MAPK	47	55	44	74	22	12	82		DYRK1A	12	14	60	71	25	10	88
p38 $\gamma$ MAPK	53	75	81	47	71	77	89		DYRK2	25	26	60	67	39	79	90
p38 $\sigma$ MAPK	56	87	63	65	67	62	102		DYRK3	51	69	81	89	57	83	96
ERK8	9	24	19	36	18	14	43		NEK2a	68	88	79	89	93	92	93
RSK1	24	50	48	63	67	29	89		NEK6	60	87	91	102	113	97	100
RSK2	47	82	78	104	85	38	104		IKKb	55	64	67	116	70	82	99
PDK1	64	96	82	94	106	65	108		PIM1	71	80	74	81	94	79	90
PKB $\alpha$	89	95	85	94	85	84	90		PIM2	80	88	76	88	94	91	109
PKB $\beta$	70	88	88	86	73	85	83		PIM3	32	69	68	48	69	69	82
SGK1	80	85	89	100	81	98	91		SRPK1	26	46	51	45	93	6	92
S6K1	53	55	39	91	78	57	86		MST1	42	94	64	92	97	40	104
PKA	77	86	75	73	97	90	106		EF2K	84	79	83	93	89	95	111
ROCK 2	42	72	56	97	97	55	107		HIPK2	3	35	31	37	64	16	79
PRK2	33	75	62	90	98	57	95		PAK4	32	66	73	61	88	47	88
PKC $\alpha$	70	85	73	84	101	64	100		PAK5	38	80	72	75	103	66	110
PKC zeta	33	68	50	85	83	83	104		PAK6	52	82	94	91	93	87	98
PKD1	29	40	23	53	35	46	85		SRC	42	91	63	85	99	47	99
MSK1	73	82	74	99	122	87	102		LCK	53	86	87	93	86	49	85
MNK1	71	75	70	102	105	110	105		CSK	57	87	84	79	95	90	98
MNK2	70	82	87	99	94	99	96		FGF-R1	39	62	59	99	97	64	87
MAPKAP-K2	98	100	99	104	96	99	102		IRR	16	22	22	32	25	24	93
PRAK	72	74	88	82	96	63	103		EPH A2	77	99	82	106	110	71	106
CAMKKb	15	52	35	98	63	12	83		MST4	65	90	69	79	76	57	90
CAMK1	34	58	62	52	86	55	109		SYK	45	84	81	86	91	37	111
SmMLCK	22	47	51	59	44	29	74		YES1	25	63	31	76	114	26	109
PHK	74	90	92	105	105	88	112		IKK $\epsilon$	35	81	48	85	88	84	108
CHK1	24	91	75	110	96	86	80		TBK1	53	98	80	91	88	92	95
CHK2	27	75	49	90	61	30	53		IGF-1R	41	90	144	78	80	80	118
GSK3 $\beta$	0	12	2	3	1	4	15		VEG-FR	30	77	55	86	57	17	82
CDK2-Cyclin A	2	34	17	48	5	2	13		BTK	60	86	86	98	75	68	79
PLK1	65	84	87	120	82	86	93		IR-HIS	63	85	88	103	75	88	93
Aurora B	59	91	76	72	75	24	107		EPH-B3	78	95	92	105	88	59	94

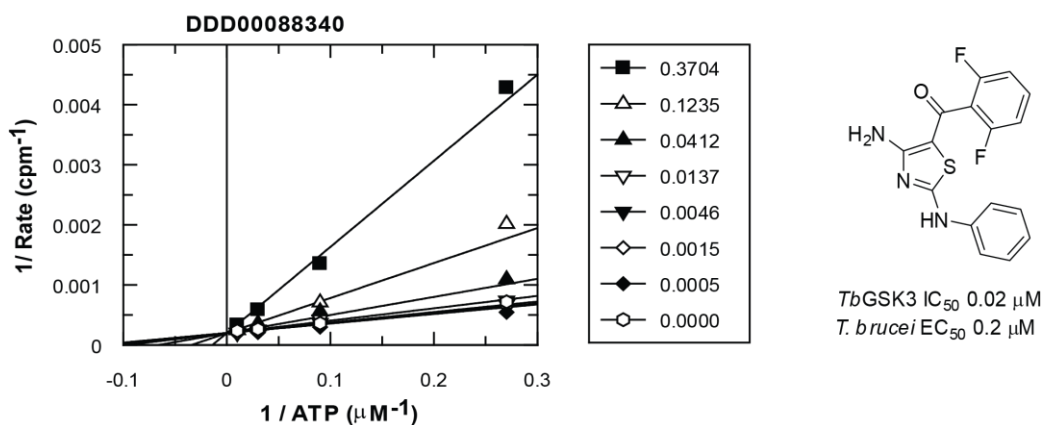
Compounds were tested in duplicate at 10  $\mu$ M. Data shown as percentage of uninhibited activity. The ranking of inhibition is shown as a heat map: red below 30% residual activity; orange 31-50% residual activity; yellow 51-70% residual activity; white above 70% residual activity. The kinase profiling was performed by the Division of Signal Transduction Therapy (DSTT) at the University of Dundee.



A



B



**Figure 3.7 Mechanism of inhibition was determined for (A) DDD00088338 and (B) DDD00088340**

Rates (cpm in 20 min) were determined at the reported inhibitor concentrations in μM with 4 varied concentrations of ATP at saturating concentration of the other substrate. The resulting Lineweaver-Burke plots were examined for diagnostic patterns for competitive inhibition and globally fitted to the equation for competitive inhibition.

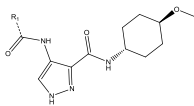
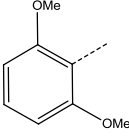
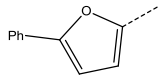
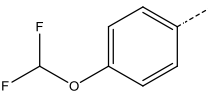
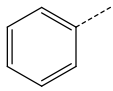
*brucei* anti-proliferative activity, as it is shown by the lack of correlation for these latter compounds (**Figure 3.6**). The unrelated SAR between enzyme activity and cell efficacy *in vitro* is particularly evident for compounds DDD00088342 and DDD00090123, which have greater potency against *T. brucei* cells ( $EC_{50}$  values of 40 nM and 300 nM respectively) than against *TbGSK3* short ( $IC_{50}$  values of 3.4  $\mu$ M and 12  $\mu$ M respectively).

Finally, the potential toxicity of these compounds was also assessed by testing a select group against a panel of mammalian kinases performed by DSTT at the University of Dundee (Bain *et al.*, 2003; Bain *et al.*, 2007). Initial compounds (DDD00086390, DDD00086405, DDD00086406 and DDD00089411) and late compounds (DDD00088338, DDD0008340 and DDD0009213) were profiled at 10  $\mu$ M in duplicate (**Table 3.6**). The profiling results illustrate that these compounds were potent inhibitors of the human *HsGSK3 $\beta$*  (percentage of inhibition of at least 85%) and also consistently hit *HsCDK2/cyclin A*, but selectivity against *HsCDK2/cyclin A* could be achieved. The compounds had a diverse ranking of toxicity against MRC5 cells, with the less promiscuous compounds generally being inactive against this cell model.

### 3.2.3 GSK3 07 series: mode of inhibition and $K_i$ characterization

I performed studies of the mechanism of inhibition of DDD00088338 and DDD00088340, selected compounds from GSK3 07 series, which confirmed that they are ATP-competitive inhibitors of *TbGSK3* short (**Figure 3.7**). The resulting  $K_i$  values ( $0.25 \pm 0.03$   $\mu$ M and  $0.05 \pm 0.01$  for DDD00088338 and DDD00088340, respectively) correlated well with the determined  $IC_{50}$  values in the biochemical assay (**Table 3.6**). This was expected, considering that the concentration of ATP in the assay is below the  $K_m$  for the ATP, the  $IC_{50}$  values are a good approximation of  $K_i$ . These compounds had

**Table 3.7 GSK3 09 Selection of R<sub>1</sub> substituted aminopyrazoles with activities against *Tb*GSK3 short, *Hs*GSK3β, *Hs*CDK2/cyclin A, BSF *T. b. brucei* and MRC5 cells**

	R <sub>1</sub>	<i>Tb</i> GSK3 short IC <sub>50</sub> (μM)	<i>Hs</i> GSK3β IC <sub>50</sub> (μM)	<i>Hs</i> CDK2/ cyclin A IC <sub>50</sub> (μM)	BSF <i>T. b. brucei</i> EC <sub>50</sub> (μM)	MRC5 EC <sub>50</sub> (μM)
DDD00085893		0.002	< 0.005	0.019	0.52	31
DDD00085905		0.007	< 0.005	0.010	0.32	0.13
DDD00085896		0.016	< 0.005	0.269	1.09	2.5
DDD00085887		0.024	< 0.005	0.038	1.31	0.76

The number of determinations was two or higher.

**Table 3.8 Mammalian kinase selectivity panel for series GSK3 09**

Enzyme	DDD00085887	DDD00085893	DDD00085905	DDD00101234
MKK1	54	83	17	76
ERK1	75	100	33	93
ERK2	18	51	2	92
JNK1	31	77	8	71
JNK2	51	82	15	63
P38 $\alpha$ MAPK	81	80	71	81
P38 $\beta$ MAPK	74	95	91	90
P38 $\gamma$ MAPK	60	72	24	95
P38 $\sigma$ MAPK	57	85	38	72
ERK8	14	27	7	47
RSK1	74	70	54	77
RSK2	92	82	93	78
PDK1	87	87	86	94
PKB $\alpha$	96	85	86	95
PKB $\beta$	45	98	84	80
SGK1	89	92	81	95
S6K1	100	100	58	94
PKA	81	74	76	91
ROCK2	81	74	72	88
PRK2	95	83	65	77
PKC $\alpha$	49	90	48	91
PKC zeta	57	83	60	96
PKD1	46	59	59	72
MSK1	80	86	66	91
MNK1	76	81	82	103
MNK2	100	63	77	97
MAPKAP-K2	92	33	90	13
PRAK	52	92	65	84
CAMKK $\beta$	34	50	27	41
CAMK1	36	76	46	87
SmMLCK	25	65	23	55
PHK	89	63	45	77
CHK1	100	84	94	83
CHK2	100	75	51	59
GSK3 $\beta$	33	0	0	10
CDK2	1	9	1	38
PLK1	100	96	99	95
Aurora B	46	87	47	91
AMPK	46	88	49	104

Enzyme	DDD00085887	DDD00085893	DDD00085905	DDD00101234
MARK3	84	93	83	74
BRSK2	55	77	60	83
MELK	32	100	13	96
CK1	55	88	25	89
CK2	74	100	21	81
DYRK1A	6	84	1	79
DYRK2	3	63	1	28
DYRK3	24	100	2	80
NEK2a	100	99	69	95
NEK6	100	84	70	82
IKK $\beta$	54	100	36	86
PIM1	46	100	13	92
PIM2	90	97	82	85
PIM3	19	94	1	96
SRPK1	90	78	72	83
MST2	72	81	69	84
EF2K	99	85	92	79
HIPK2	14	100	4	93
PAK4	68	74	27	39
PAK5	68	75	47	51
PAK6	97	90	71	77
Src	98	94	85	91
Lck	76	90	67	87
CSK	85	77	73	91
FGF-R1	99	95	72	116
IRR	39	87	26	75
EPHA2	100	100	100	99
MST4	46	49	30	80
SYK	78	100	70	84
YES1	100	79	45	103
IKK $\epsilon$	62	85	46	87
TBK1	63	96	73	91
IGF-1R	96	100	15	79
VEG-FR	52	90	29	93
BTK	100	100	100	91
IR-HIS	84	99	50	81
EPH-B3	60	100	72	88
IKK	68	90	71	ND
MINK1	ND	ND	ND	18

Compounds were tested in duplicate at 10  $\mu$ M and data shown as percentage of uninhibited activity. The ranking of inhibition is shown as a heat map: red below 30% residual activity; orange 31-50% residual activity; yellow 51-70% residual activity; white above 70% residual activity. The kinase profiling was performed by the Division of Signal Transduction Therapy (DSTT) at the University of Dundee.

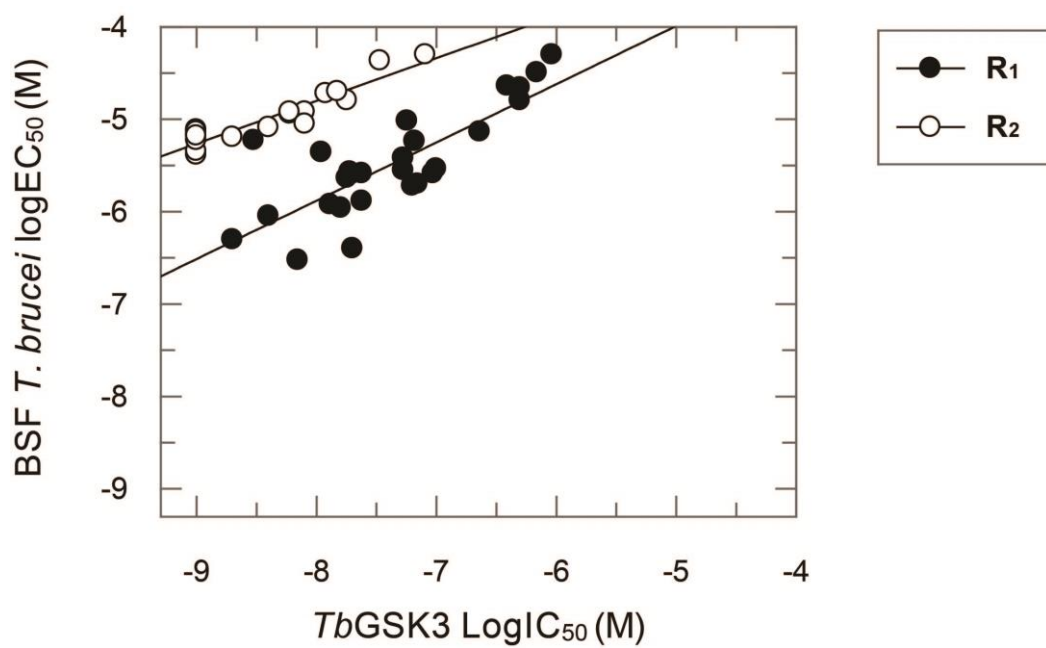
ND: not determined.

a different ratio between efficacy in cells and biochemical potency (the ratio *T. brucei*  $EC_{50} / K_i$  varied from 0.4 for DDD00088338 to 4.3 for DDD00088340) which was unexpected for inhibitors with the same mechanism of inhibition. The divergent ratios of cellular and enzymatic activities, despite the ATP-competitive mechanism of inhibition towards the molecular target *TbGSK3* short, give additional hints that these compounds have a common pharmacophore shared with other unknown targets in *T. brucei*.

### 3.2.4 GSK3 09 series: correlation, specificity and selectivity

GSK3 09 is a series of 4-amino-1*H*-pyrazole-5-carboxamide (**Figure 3.5**). These compounds were synthesised based on the literature, by varying residues to explore the relationship between *TbGSK3* short, *HsCDK2* and *HsGSK3 $\beta$*  activity (Wyatt *et al.*, 2008). Initially 19 examples were tested and some of them were found to be nanomolar inhibitors of *TbGSK3* short, such as compounds DDD00085893, DDD00085905, DDD00085896 and DDD00085887 with  $IC_{50}$  values ranging from 2 to 24 nM (**Table 3.7**). Interestingly, the most potent compound DDD00085893 was active against *T. b. brucei* cells ( $EC_{50}$  of 0.52  $\mu$ M), and it was not toxic against MRC5 cells ( $EC_{50}$  of 30  $\mu$ M). This compound was profiled against the mammalian panel of kinases at 10  $\mu$ M; only *HsGSK3 $\beta$*  and *HsCDK2/cyclin A* were inhibited more than 90% (**Table 3.8**) with  $IC_{50}$  values of 1 nM and 19 nM, respectively, indicating that in MRC5 cells *HsGSK3* inhibition did not cause cell toxicity. All the compounds from this series obtained by replacement of **R<sub>1</sub>** were not selective against *HsGSK3 $\beta$* , but selectivity could be achieved against *HsCDK2/cyclin A*.

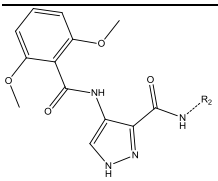
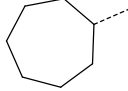
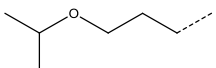

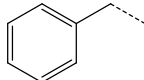
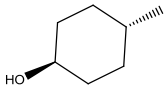
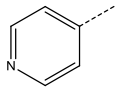
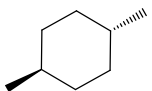
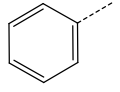
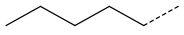
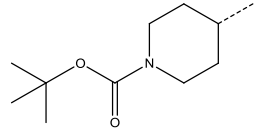
The other two compounds (DDD00085887 and DDD00085905) profiled against the mammalian kinase panel were more promiscuous and this was reflected by a higher

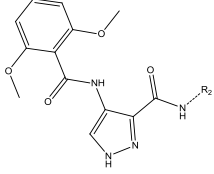
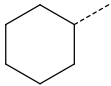
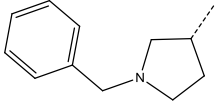
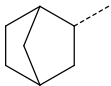
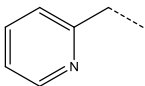

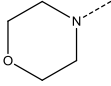
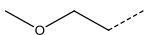
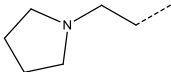
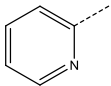
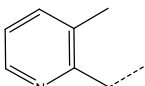
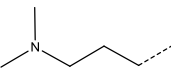
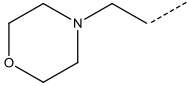


**Figure 3.8 Correlation plot between inhibition of *TbGSK3* short and inhibition of *T. brucei* cell growth for GSK3 09 series**

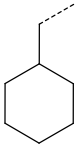
The compounds substituted in position R1 are indicated with close circles and the compounds substituted in position R2 are indicated with open circles.

**Table 3.9 GSK3 09 Selection of R<sub>2</sub> substituted aminopyrazoles with activities against *Tb*GSK3 short, *Hs*GSK3β, *Hs*CDK2/cyclin A, BSF *T. b. brucei* and MRC5 cells**

		R <sub>2</sub>	<i>Tb</i> GSK3 short IC <sub>50</sub> (μM)	<i>Hs</i> GSK3β IC <sub>50</sub> (μM)	<i>Hs</i> CDK2/ cyclin A IC <sub>50</sub> (μM)	BSF <i>T. b. brucei</i> EC <sub>50</sub> (μM)	MRC5 EC <sub>50</sub> (μM)
DDD00101238			0.001	0.02	2.04	4.51	42.08
DDD00101236			0.001	ND	> 10	6.59	> 50
DDD00101223			0.008	0.08	2.4	11.64	> 50
DDD00101234			0.001	0.33	> 10	5.92	50
DDD00101250			0.002	0.07	1.56	6.42	33.68
DDD00101247			0.004	0.12	1.31	8.15	> 50
DDD00101237			< 0.001	ND	> 4.76	6.69	45.28
DDD00101224			0.006	0.07	4.3	11.50	> 50
DDD00101246			0.001	0.1	4.8	7.29	> 50
DDD00101242			0.006	0.14	4.67	12.22	> 50

		<b>R<sub>2</sub></b>	<b>TbGSK3 short IC<sub>50</sub> (μM)</b>	<b>HsGSK3β IC<sub>50</sub> (μM)</b>	<b>HsCDK2/ cyclin A IC<sub>50</sub> (μM)</b>	<b>BSF <i>T. b. brucei</i> EC<sub>50</sub> (μM)</b>	<b>MRC5 EC<sub>50</sub> (μM)</b>
DDD00101248			0.001	0.05	1.16	4.09	35.43
DDD00101235			0.008	0.08	> 10	8.90	37.82
DDD00101220			0.018	ND	> 10	15.58	> 50
DDD00101240			0.015	0.32	> 10	20.22	> 50
DDD00101228			0.012	0.22	> 10	18.84	> 50
DDD00101229			0.081	0.45	> 10	50	> 50
DDD00101239			0.034	0.30	> 10	42.75	> 50
DDD00101221			0.316	0.66	> 10	> 50	> 50
DDD00101230			0.318	0.94	> 10	> 50	> 50
DDD00101232			0.142	0.87	> 10	> 50	> 50
DDD00101241			0.332	0.47	> 10	> 50	> 50
DDD00101243			0.054	0.63	> 10	> 50	> 50



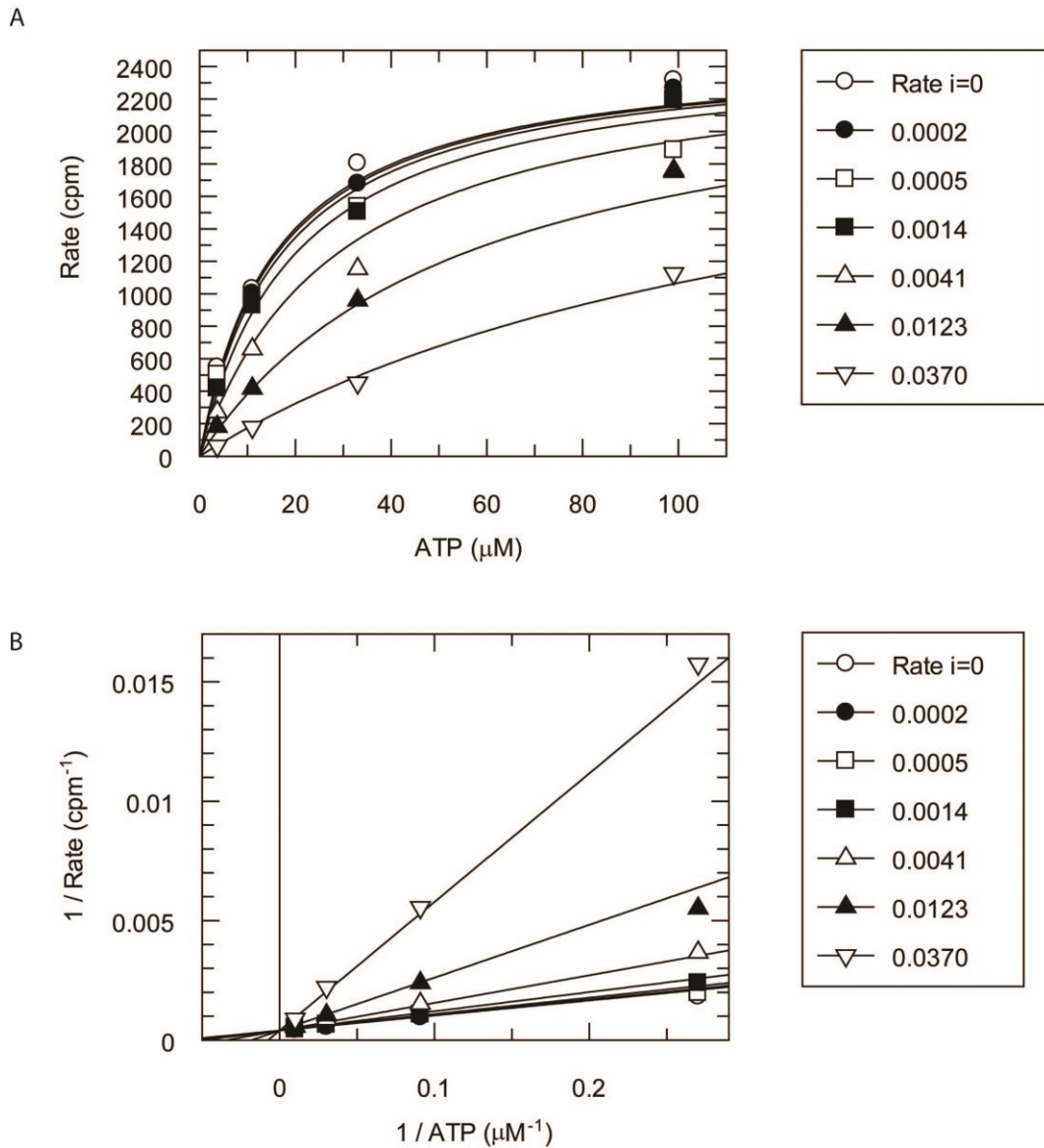
	<b>R<sub>2</sub></b>	<b><i>Tb</i>GSK3 short IC<sub>50</sub> (μM)</b>	<b><i>Hs</i>GSK3β IC<sub>50</sub> (μM)</b>	<b><i>Hs</i>CDK2/ cyclin A IC<sub>50</sub> (μM)</b>	<b>BSF <i>T. b.</i> <i>brucei</i> EC<sub>50</sub> (μM)</b>	<b>MRC5 EC<sub>50</sub> (μM)</b>
DDD00107470		0.001	ND	ND	7.66	>50

The number of determinations was two or higher. ND: not determined.

toxicity against MRC5 cells with  $EC_{50}$  values of 0.76  $\mu$ M and 0.13  $\mu$ M, respectively (**Table 3.7** and **Table 3.8**). GSK3 09 series overall had a very good correlation between inhibition of the recombinant *TbGSK3* short and efficacy *in vitro* versus BSF *T. b. brucei* cells (correlation coefficient of 0.80) giving an indication that this series could exert its action “on target” against *TbGSK3* short (**Figure 3.8**). Moreover, the drop-off between enzyme inhibition and cellular efficacy was  $\sim$  100-fold. This was consistent with the expected ATP-competitive mode of inhibition (as discussed in section 3.2.2).

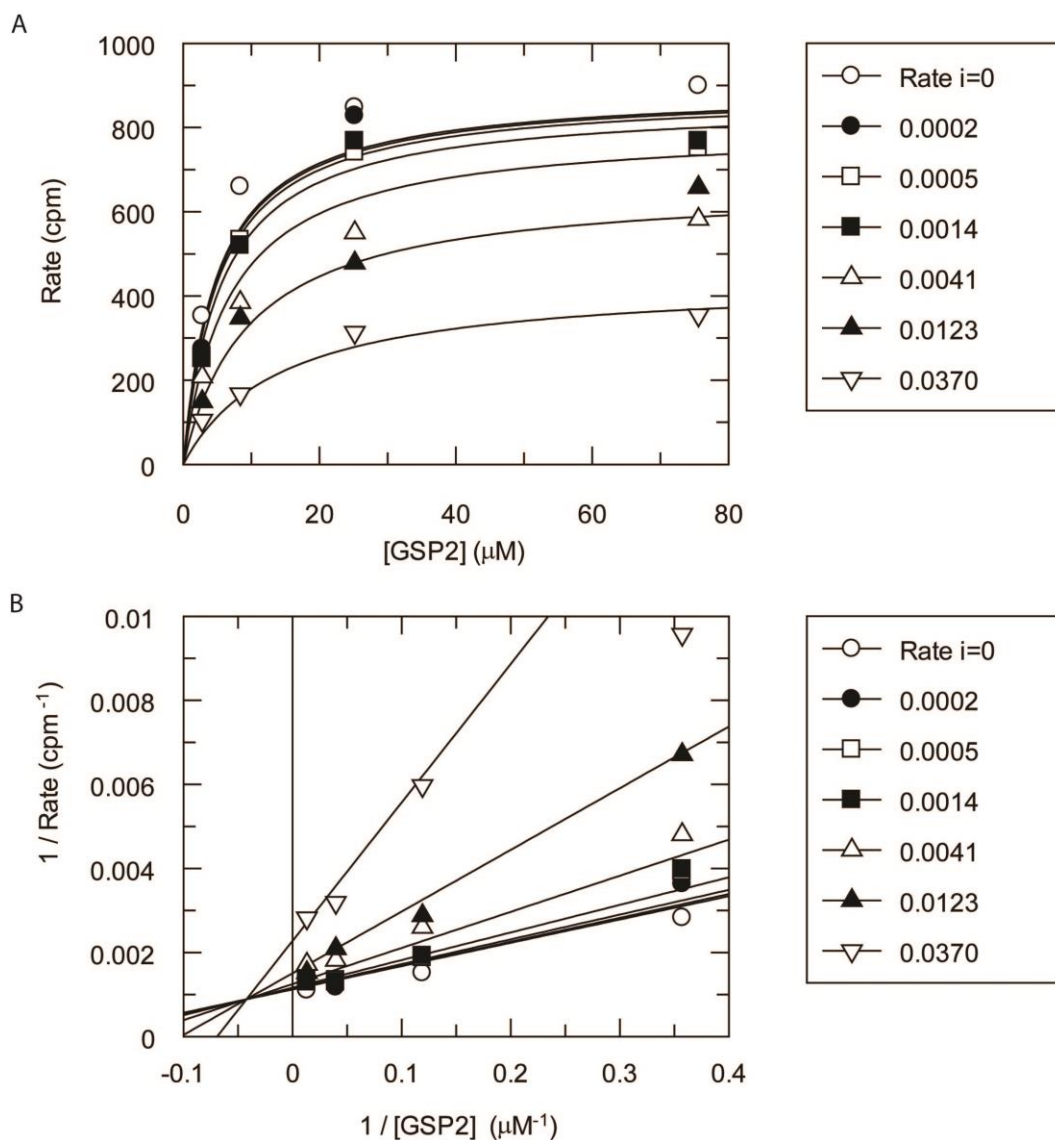
Further optimization within the series was aimed at exploring whether it was possible to obtain selectivity towards *HsGSK3 $\beta$*  retaining potency against *TbGSK3* short. As DDD00085893 was the most promising inhibitor (potent against *TbGSK3* short in the biochemical assay, blocking proliferation of BSF *T. b. brucei* in culture, and not toxic against MRC5 cells) the 2,6-dimethoxybenzamide group was retained in position **R<sub>1</sub>** and an array of substituents was made in position **R<sub>2</sub>** (**Table 3.9**). The majority of variations on this position retained the activity against *TbGSK3* short in the nanomolar range. The **R<sub>2</sub>**-substituted compounds were tested in proliferation assay against BSF *T. b. brucei* and MRC5 cells. As for the **R<sub>1</sub>**-substituted analogues, also for the **R<sub>2</sub>**-substitutes there was a good correlation (correlation coefficient of 0.94) between the enzymatic inhibition of *TbGSK3* short and cell efficacy in BSF *T. b. brucei* culture with a 1000-fold drop in potency between the biochemical and the cellular assay (**Figure 3.8**).

Interestingly, compared to the variations in position **R<sub>1</sub>**, a majority of **R<sub>2</sub>** analogues showed selectivity over both *HsCDK2/cyclin A* and *HsGSK3 $\beta$*  up to reaching more than 10,000-fold selectivity towards *HsCDK2* and more than 300-fold selectivity towards *HsGSK3 $\beta$* . Compound DDD00101234, substituted with a benzyl in position **R<sub>2</sub>** is the most selective *TbGSK3* short inhibitor known so far (**Table 3.9**) (Ojo *et al.*, 2008; Oduor *et al.*, 2011). This compound not only is a potent *TbGSK3* inhibitor selective



**Figure 3.9 Mechanism of inhibition of DDD00085893 towards ATP**

Rates (cpm in 20 min) were determined at the reported inhibitor concentrations ( $\mu\text{M}$ ) with 4 varied concentrations of ATP at the  $K_m^{\text{app}}$  concentration of the other substrate ( $8.4 \mu\text{M}$ ) and  $2.5 \text{ nM}$  of *TbGSK3* short. The direct plot (**A**) and the derived Lineweaver-Burke plot (**B**) were examined for diagnostic patterns before globally fitting the data to the equation for competitive inhibition. The  $K_i$  value of DDD00085893 is  $5.7 \pm 0.1 \text{ nM}$ .



**Figure 3.10 Mechanism of inhibition of DDD00085893 towards GSP2 substrate**

Rates (cpm in 20 min) were determined at the reported inhibitor concentrations ( $\mu\text{M}$ ) with 4 varied concentrations of GSP2 substrate at the  $K_m^{\text{app}}$  concentration of the ATP (11  $\mu\text{M}$ ) and 2.5 nM of *Tb*GSK3 short. The direct plot (**A**) and the derived Lineweaver-Burke plot (**B**) were examined for diagnostic patterns before globally fitting the data to the equation for mixed inhibition. The calculated  $K_i$  and  $K_i'$  value of DDD00085893 are  $7.5 \pm 3.1$  nM and  $35.9 \pm 11.9$  nM, respectively.

towards the human isoform, but also retains efficacy against the BSF *T. b. brucei* in culture ( $EC_{50}$  of 5.9  $\mu$ M) and is not toxic for MRC5 cells ( $EC_{50}$  of 50  $\mu$ M) (**Table 3.9**).

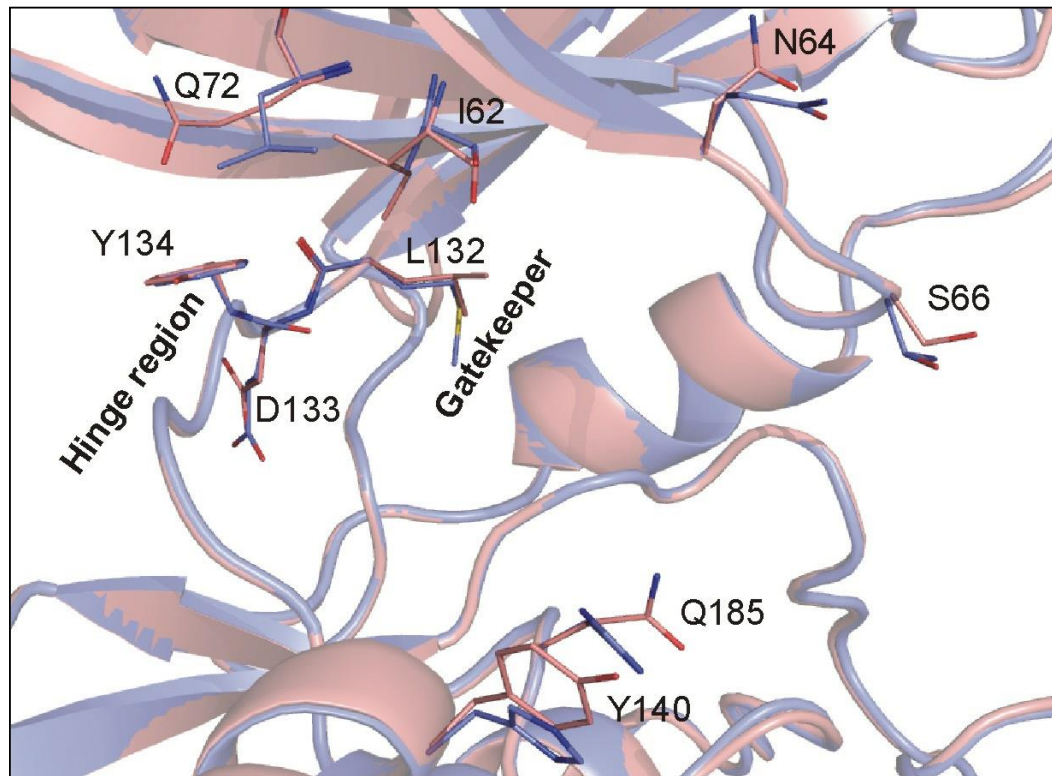
DDD00101234 was also profiled against a panel of mammalian kinases. This compound inhibited the activity of only three PKs more than 80%: GSK3 $\beta$ , MAPKAP-K2 and MINK1, confirming that is a selective kinase inhibitor (**Table 3.8**).

### 3.2.5 GSK3 09 series: Mode of inhibition and $K_i$ determination

I determined the  $K_i$  of DDD00085893 proving that this compound acts through competitive inhibition with respect to ATP (**Figure 3.9**) and mixed inhibition towards the peptide substrate GSP2 (**Figure 3.10**). Increasing the concentration of DDD00085893 resulted in an increase in the  $K_m^{app}$  for ATP, but the  $V_{max}$  remained unchanged. On the other hand, varying the concentration of inhibitor caused changes in both  $V_{max}$  and  $K_m^{app}$  for the peptide substrate GSP2. The calculated  $K_i$  of DDD00085893 ( $5.7 \pm 0.1$  nM) correlates as expected with the  $IC_{50}$  in the biochemical assay, taking into consideration that the  $IC_{50}$  value was determined at an ATP concentration below its  $K_m$ .

### 3.2.6 Mode of binding of series GSK3 07 and GSK3 09

The homology model for *Tb*GSK3 short was built by Dr Torsten Luksch using as template the solved crystal structure of *Hs*GSK3 $\beta$  (pdb code 1R0E). These two proteins shared a high degree of identity (52%) in their kinase domains, but a closer look at their binding pocket revealed that inhibitor selectivity could be achieved exploiting the difference in nine residues in the binding site (**Figure 3.11** and **Table 3.10**). *Tb*GSK3 short is also very similar to *Hs*CDK2, although in this case there are differences in some key residues in the kinase domain that could be exploited to achieve selectivity (**Table 3.10**). These findings were in agreement with the independent



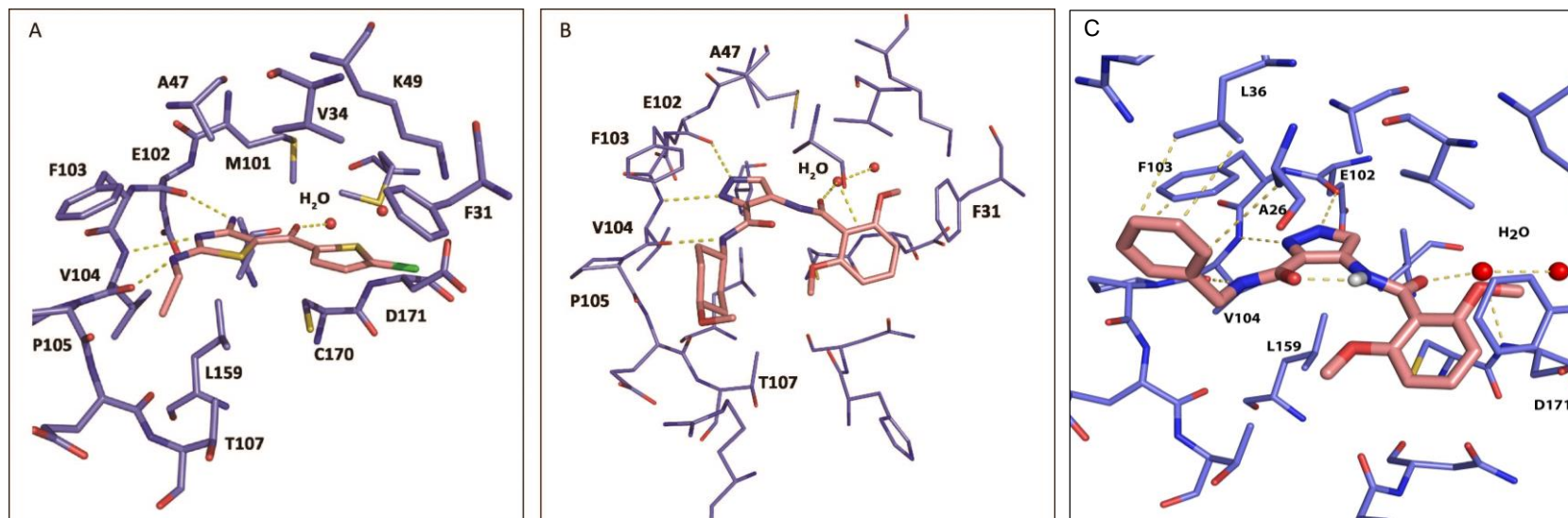
**Figure 3.11 Superposition of the *Hs*GSK3 crystal structure 1R0E with the homology model of *Tb*GSK3 short**

The *Hs*GSK3 $\beta$  structure is in pink, the *Tb*GSK3 short homology model in blue. Both structures are shown in cartoon representation. Amino acids that are different in the binding pocket of *Tb*GSK3 short in comparison to *Hs*GSK3 $\beta$  are represented as sticks. Residue numbers are for *Hs*GSK3 $\beta$ . The homology modelling was performed by Dr Torsten Luksch (figure from Woodland *et al.*, 2013).

**Table 3.10 Differences in the ATP binding pockets of *Tb*GSK3, *Hs*GSK3 $\beta$  and *Hs*CDK2**

<i>Tb</i> GSK3	<i>Hs</i> GSK3 $\beta$	<i>Hs</i> CDK2
V25	V61	<b>K9</b>
A26	<b>I62</b>	<b>I10</b>
G27	G63	G11
Q28	<b>N64</b>	<b>E12</b>
G29	G65	G13
T30	<b>S66</b>	T14
F31	F67	<b>Y15</b>
V34	V70	V18
L36	<b>Q72</b>	<b>K20</b>
A47	A83	A31
K49	K85	K33
E61	E97	<b>I52</b>
M65	M101	<b>L55</b>
V77	V110	V64
M101	<b>L132</b>	<b>F80</b>
E102	<b>D133</b>	E81
F103	<b>Y134</b>	F82
I04	V135	<b>L83</b>
P105	P136	<b>H84</b>
E106	E137	<b>Q85</b>
T107	T138	<b>D86</b>
H109	<b>Y140</b>	<b>K88</b>
R110	R141	<b>K89</b>
K154	K183	K129
H156	<b>Q185</b>	<b>Q131</b>
N157	N186	N132
L159	L188	L134
C170	C199	<b>A144</b>
D171	D200	D145

Amino acids of *Hs*GSK3 $\beta$  or *Hs*CDK2 which differ in *Tb*GSK3 short are shown in bold (table from Urich *et al.*, 2014).



**Figure 3.12 Proposed binding modes for (A) DDD00065658, (B) DDD00085893, and (C) DDD00101234 in the homology model of *TbGSK3* short**

Both, ligands and protein are represented in sticks and colour coded by atom types. Ligand carbon atoms are shown in pink and protein carbon atoms in light blue. Hydrogen bonds are shown as yellow dotted lines. The binding modes were generated by Dr Torsten Luksch and Dr Robert Urich (figures adapted from Woodland *et al.*, 2013 and Urich *et al.*, 2014).

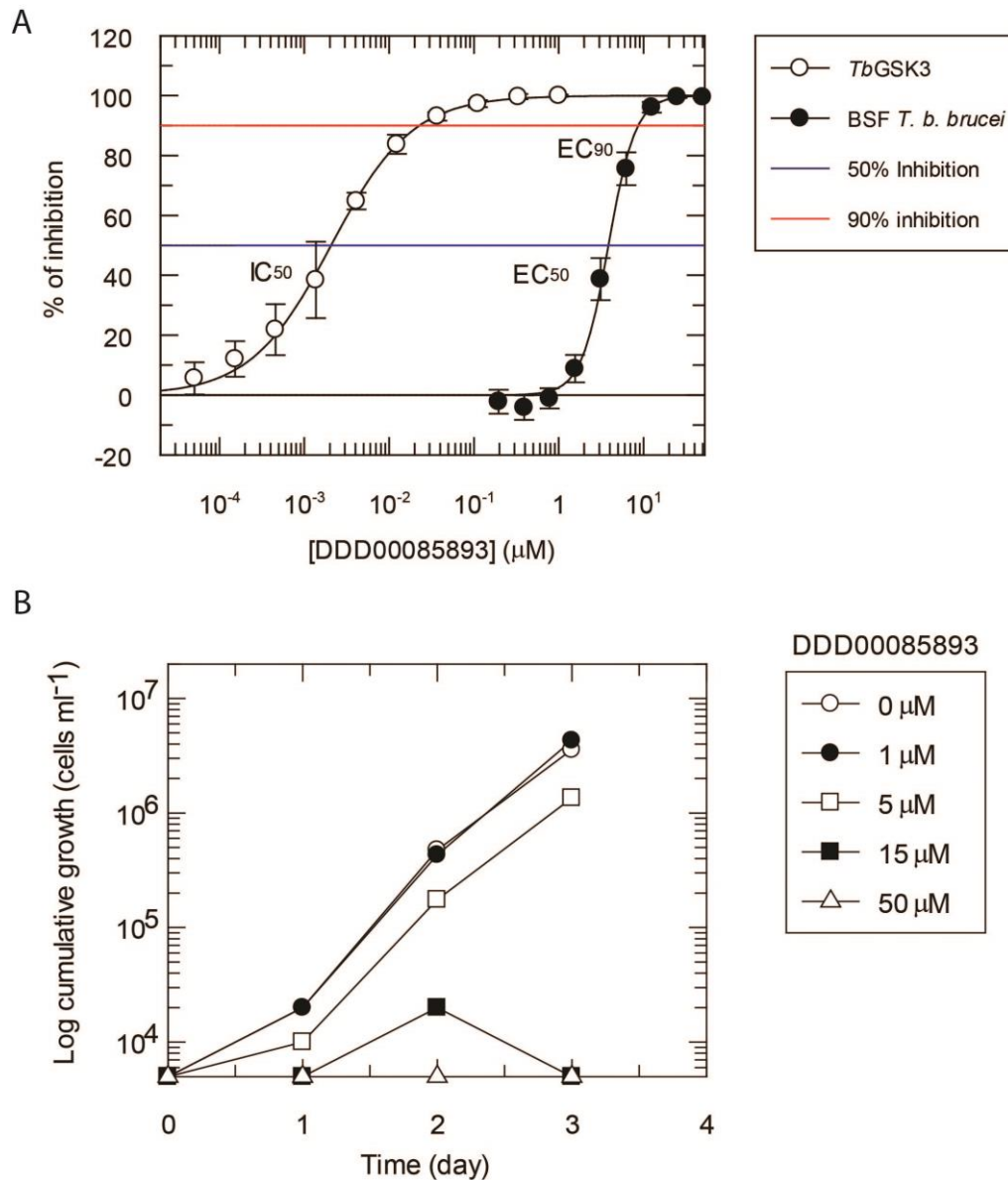


observations of Ojo and co-workers (2008; 2011). This homology model was used for the docking of the identified inhibitors from both series GSK3 07 and GSK3 09 which were predicted to bind in the ATP pocket (**Figure 3.12** A, B and C).

In the binding mode of the GSK3 07 initial hit, DDD00065658, the 2, 4-diaminothiazole group has a kinase hinge binding recognition motif allowing hydrogen bonds to the backbone of Glu102, Phe103, and Val104 (**Figure 3.12** A). All the other residues DDD00065658 is interacting with are conserved between *Hs*GSK3 $\beta$  and *Tb*GSK3 short, so it is not surprising that all the compounds tested from this series in the kinase panel (**Table 3.6**) are not selective towards *Hs*GSK3 $\beta$  (Woodland *et al.*, 2014).

A potential binding mode was generated also for DDD00085893 and protein-ligand interactions were analysed by Dr Robert Urich (**Figure 3.12** B). Also the 4-amido-5-carboxamide-pyrazole group bears a kinase hinge binding motif and makes hydrogen bonds to the backbone of Glu102, Phe103, and Val104. Notably, the gatekeeper residue differs between *Tb*GSK3 short, *Hs*GSK3 $\beta$ , and *Hs*CDK2 (**Table 3.10**). While *Hs*GSK3 $\beta$  and *Tb*GSK3 short enzymes have Leu132 or Met101, respectively, in this position, in *Hs*CDK2 a Phe80 is present. The gatekeeper region of *Hs*CDK2 is more constrained and less accessible by bulky R<sup>1</sup>-substituents such as the 2,6-dimethoxybenzamide group of DDD00085893. This main difference could explain the selectivity of this compound towards *Hs*CDK2.

The achieved selectivity towards *Hs*GSK3 $\beta$  of the compounds substituted in R<sup>2</sup> could be rationalised again docking the compounds in the homology model for *Tb*GSK3 short. The docking of DDD00101234 suggests that the source of the selectivity is the replacement of Phe103 in a hydrophobic pocket of *Tb*GSK3 short with Tyr134 in *Hs*GSK3 $\beta$ . The phenyl of the N-benzylamide moiety of DDD00101234 can interact with the aromatic side chain of Phe103 in *Tb*GSK3 short, but it is electrostatically



**Figure 3.13 DDD00085893 inhibition of *TbGSK3* short activity and rate of killing of BSF *T. b. brucei* cell proliferation**

(A) The inhibition of the recombinant *TbGSK3* short was measured using a filterplate assay (open circles), whereas cell proliferation inhibition of bloodstream trypanosomes was measured with resazurin assay (close circles) (as reported in Materials and Methods). Fifty percent and 90% of inhibition are represented respectively with a blue and red line. Data points are average of triplicate determinations  $\pm$  SD. (B) BSF *T. b. brucei* cells were cultured in presence of different concentrations of DDD00085893 as reported in the legend. Cell densities were determined daily by manual counting with a haemocytometer. Each data point is average of two determinations.

repulsed by the polar lateral chain of Tyr134 in *HsGSK3 $\beta$*  (Urich *et al.*, 2014). This moiety also has hydrophobic interactions with Leu36 and Ala26 in *TbGSK3* short. Leu36 is replaced by polar amino acids in *HsGSK3 $\beta$*  and *HsCDK2* (Gln72 and Lys20, respectively), further diminishing hydrophobic interactions and increasing the selectivity of DDDD001010234 over *HsGSK3 $\beta$*  and *HsCDK2* (**Figure 3.12 C**).

### 3.2.7 DDD00085893 is a cidal inhibitor of *T. brucei*

DDD00085893 is a potent ATP-competitive inhibitor of the recombinant *TbGSK3* short with an  $IC_{50}$  of 3 nM and  $K_i$  of 5.4 nM. In my assays, this compound also demonstrated a detrimental effect on BSF *T. brucei* proliferation with an  $EC_{50}$  of 3.9  $\mu M^a$ . The shift in potency is in agreement with the ATP-competitive mechanism of inhibition and the differential concentration of ATP in *T. brucei* (mM range) and in the biochemical assay (1  $\mu M$ ) (as discussed in section 3.2.2 and 3.2.4) (**Figure 3.13 A**).

In order to verify whether DDD00085893 had a cidal or a static effect on BSF *T. brucei*, cells were incubated for 72 h with inhibitor concentration ranging from 0 to 50  $\mu M$  (**Figure 3.13 B**). Cells were counted daily and subcultured in fresh medium after removal of the inhibitor. As expected, 50% of growth inhibition was obtained at 5  $\mu M$ , a concentration equivalent to the  $EC_{50}$  value determined with the resazurin assay. After incubation for 24 h in presence of 50  $\mu M$  of DDD00085893, no cells were detectable, but cells washed free of the inhibitor could recover after 8 days. In contrast after

---

<sup>a</sup> This value was 10-times higher with respect to the one determined by the DDU routine screening, the discrepancy could be explained by the fact that: (i) a new batch of inhibitor was synthesised and (ii) in the meantime some medium components were re-placed resulting with faster growth of the BSF *T. brucei*. Both factors could have impacted in the change of potency of DDD00085893, which remained thereafter constant throughout my study.

**Table 3.11 Determination of DDD00085893 IC<sub>50</sub>, for *Tb*GSK3 short and EC<sub>50</sub>, EC<sub>90</sub>, EC<sub>99</sub> and EC<sub>99,9</sub> for BSF *T. brucei***

<b>DDD00085893</b>		<b><i>Tb</i>GSK3 short (nM) (n=4)</b>		
<b>Hill slope</b>	<b>IC<sub>50</sub></b>			
0.9 ± 0.1	3 ± 1			

<b>DDD00085893</b>		<b>BSF <i>T. brucei</i> (µM) (n=4)</b>		
<b>Hill slope</b>	<b>EC<sub>50</sub></b>	<b>EC<sub>90</sub></b>	<b>EC<sub>99</sub></b>	<b>EC<sub>99,9</sub></b>
2.5 ± 0.2	3.9 ± 0.6	9.5 ± 2.1	25.3 ± 7.1	25.4 ± 7.1

Where n is the number of 4 independent experiments.

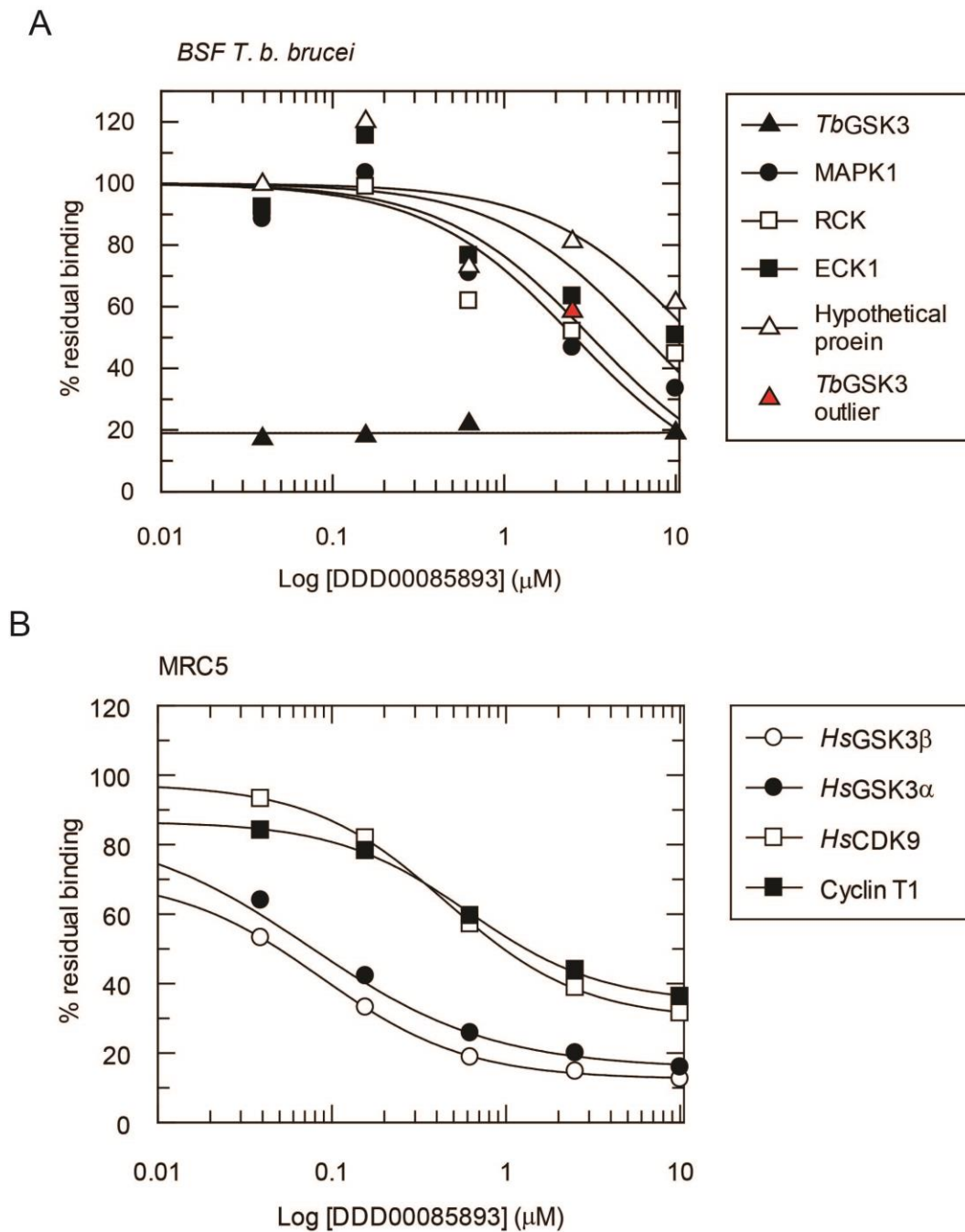
incubation for 48 h with 50  $\mu$ M the trypanosomes did not recover, even after washing off of the inhibitor. Thus, DDD00085893 is a cidal inhibitor of *Trypanosoma brucei* after 48 h incubation at the concentration of 50  $\mu$ M. Overall these findings are consistent with the EC<sub>99</sub> (or minimum inhibitory concentration, MIC) of this inhibitor which is greater than 25  $\mu$ M (**Table 3.11**).

### **3.3 Target identification by chemical proteomic approach**

#### **3.3.1 Kinobeads approach**

From the screening of *Tb*GSK3 short, two chemical series were identified with an effect on the proliferation of BSF *T. brucei*. Intriguingly, series GSK3 07 had a very limited drop-off between cellular and enzyme potency, with some examples clearly having off-target effects displaying higher potency against the cells rather than the enzyme. In contrast, series GSK3 09 had a tighter correlation between cell and enzyme activity in agreement with an ATP-competitive mechanism of inhibition. Consequently this was considered a promising series for the chemical validation of *Tb*GSK3 short as a drug target in *T. brucei*. Although the biochemical assay of the recombinant *Tb*GSK3 short seemed to be a good model for the prediction of the efficacy of the compounds in the cells, the biological complexity in cells is far higher than in a biochemical assay, therefore it is important to prove that the cellular effect is due to the action of the inhibitor on the target enzyme (Patricelli *et al.*, 2011). To achieve this, we decided to investigate the targets of these compound series in *T. brucei* lysate in collaboration with Cellzome AG (Heidelberg, Germany).

A chemical proteomic approach was undertaken for the direct identification of kinase inhibitor targets in cell lysates (Bantscheff *et al.*, 2007) utilising kinobeads and iTRAQ technology (section 2.7.1). Bantscheff and colleagues (2007) tested their



**Figure 3.14 Proteomic profiling of DDD00085893 in BSF *T. b. brucei* and MRC5 lysates by kinobeads competition assay**

Competition binding curves calculated by iTRAQ reporter signals (measured as percentage of residual binding). DDD00085893 was dosed in 5 concentrations ranging from 39 nM to 10 μM in *T. brucei* (A) and in MRC5 cell lysates (B). The curves are the non-linear fits using the 2- and 4-parameter  $IC_{50}$  equation with GraFit.

This work was done in collaboration with Cellzome (Heidelberg) (Urbaniak *et al.*, 2012a).

technology against numerous human and rodent cell lines proving that a broad spectrum of protein kinases interacted with the kinobeads. Kinobeads not only capture protein kinases, but also other ATP- and purine-binding proteins such as chaperones, helicases, ATPases, motor proteins, transporters and metabolic enzymes (Bantscheff *et al.*, 2007). Similarly this technology could be used to identify directly in *T. brucei* lysate the native targets of inhibitors and also to verify potential toxicity directly in MRC5 lysate.

The kinobeads could bind *T. brucei* kinases with a coverage of 57 kinases belonging to AGC (5 examples), CAMK (10), CK1 (1), CMGC (22), other (3), STE (6), NEK (9), and atypical (1) families (Urbaniak *et al.*, 2012a).

Compound DDD00085893 is a potent inhibitor of *TbGSK3 short in vitro*, with a good efficacy against cultured *T. brucei* and good selectivity against mammalian cells (**Table 3.7**). The kinobead profiling of DDD00085893 against *T. b. brucei* cell lysates confirmed the *TbGSK3 short* was specifically targeted by this compound. In the chemical proteomic experiment a total of 171 proteins, 49 of them protein kinases, were detected by affinity enrichment with the kinobeads (Urbaniak *et al.*, 2012a). When incubated with DDD00085893 ranging from 10  $\mu\text{M}$  to 39 nM only five proteins were no longer bound, four of which were protein kinases (**Figure 3.14 A**). The top binder of this inhibitor was *TbGSK3 short* with 17% of residual binding at 39 nM (an  $\text{IC}_{50}$  value could not be extrapolated). The protein kinase hits belonged to the same family of CMGC kinases and in the rank order of affinity were MAPK kinase (Tb427.10.3230) with an  $\text{IC}_{50}$  of 2.6  $\mu\text{M}$ , RCK kinase (Tb427.10.14800) with  $\text{IC}_{50}$  of 3.2  $\mu\text{M}$ , CDK-like kinase ECK1 (Tb11.01.8550) with  $\text{IC}_{50}$  of 6.6  $\mu\text{M}$ , and finally a hypothetical protein (Tb927.4.4060) with an  $\text{IC}_{50}$  of 12.9  $\mu\text{M}$ .

These results confirm that *TbGSK3 short* is the main target of DDD00085893 in *T. b. brucei* cells and that this compound has an affinity in the low nM range for the endogenous protein as well as the recombinant one. The coverage of the *T. b. brucei*

**Table 3.12 Kinobeads competition assay for the BSF *T. b. brucei* and MRC5 targets of DDD00085893**

BSF <i>T. b. brucei</i> targets	GeneDB accession number	IC <sub>50</sub> ± SD (μM) or percentage of inhibition
CMGC/GSK	Tb427.10.13780	17% at 39 nM
CMGC/MAPK	Tb427.10.3230	2.62 ± 0.85
CMGC/RCK	Tb427.10.14800	3.23 ± 1.6
CMGC/ECK1	Tb427.01.8550	6.61 ± 2.9
hypothetical protein	Tb427.04.4060	12.9 ± 6.4
MRC5 targets	IPI accession number	IC <sub>50</sub> ± SD (μM)
<i>Hs</i> GSK3β	IPI00216190.1	0.087 ± 0.03
<i>Hs</i> GSK3α	IPI00292228.1	0.076 ± 0.16
CDK9	IPI00552413.2	0.45 ± 0.02
Cyclin T1	IPI00030247.1	0.61 ± 0.08

These data have been published by Urbaniak *et al.*, 2012a

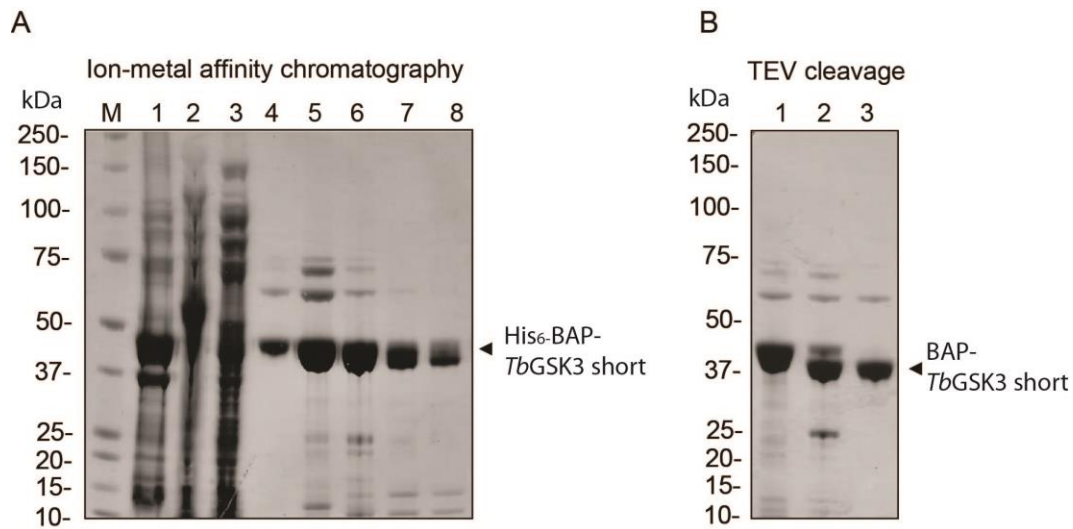




kinome by the kinobeads in this experiment was 27% (considering a total of 182 PKs) and DDD00085893 was not a promiscuous inhibitor of these kinases.

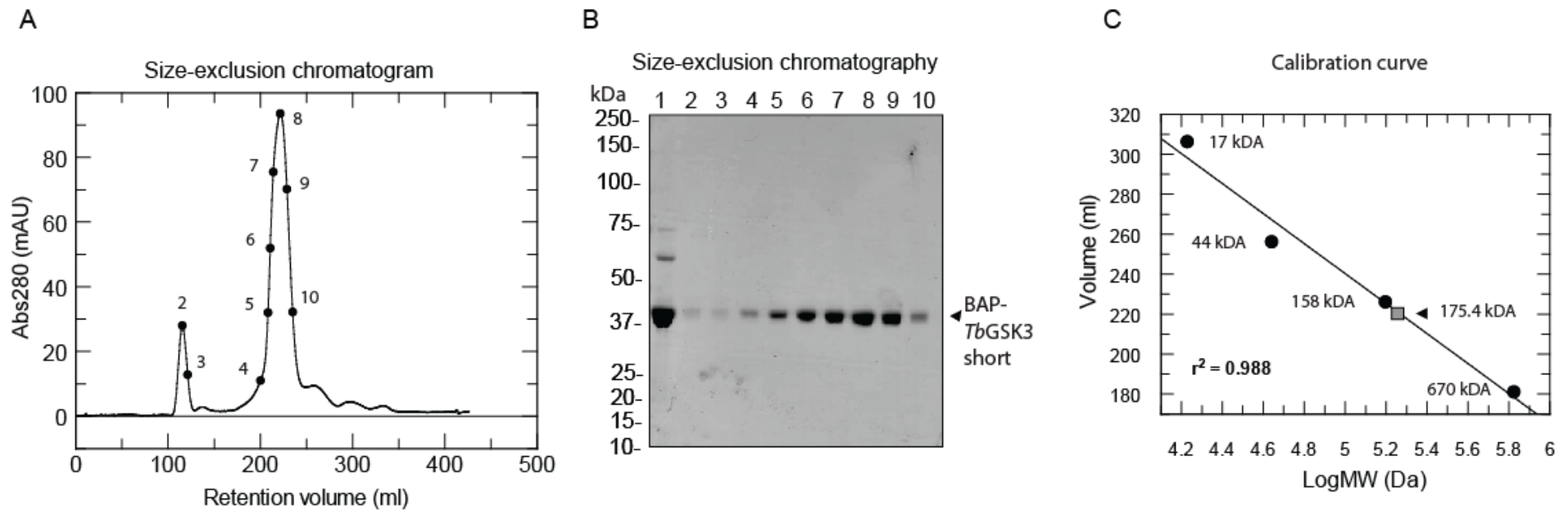
A similar competition experiment was conducted with MRC5 lysates in order to evaluate the selectivity of this compound against the parasite kinome and potential toxic effects against the host kinome. In MRC5 lysates the kinobeads could bind 238 proteins, 125 of which were annotated as protein kinases. Only four were actually displaced by incubation of DDD00085893 ranging from 10  $\mu$ M to 39 nM. The top binders were *HsGSK3 $\beta$*  with an  $IC_{50}$  of 87 nM, other binders were *HsGSK3 $\alpha$*  with an  $IC_{50}$  of 76 nM, CDK9 with an  $IC_{50}$  of 0.45  $\mu$ M and its binding partner cyclin T1 with a comparable affinity (**Figure 3.14 B** and **Table 3.12**). These results indicate that the compound DDD00085893 is a specific inhibitor of *TbGSK3* short, but not completely selective for *T. brucei*.

*T. brucei* has two isoforms annotated in GeneDB, *TbGSK3* long (Tb927.07.2420) and *TbGSK3* short (Tb927.10.1378), which are 40% identical at the amino acid level. In the competition experiment all the identified peptides were unique to *TbGSK3* short with only one peptide sequence in common (LCDFGSAK) between the two isoforms (**Figure 3.15**). The results from the kinobeads experiment prove that DDD00085893 binds the native *TbGSK3* short, although binding to the long isoform cannot be definitely ruled out. To date *TbGSK3* long protein has never been identified in any global proteomic and phosphoproteomic study in bloodstream form *T. b. brucei* (Urbaniak *et al.*, 2012b; Urbaniak *et al.*, 2013).



**Figure 3.16 Expression and purification of recombinant *TbGSK3* short tagged at the N-terminus with His<sub>6</sub>-BAP tag**

Protein sample were run on 4-12% SDS-PAGE gel and stained with Coomassie Blue. **(A)** Purification by Ni-affinity column: (M) marker; (1) insoluble fraction; (2) clarified lysate by centrifugation; (3) flowthrough; (4-8) fractions at higher absorbance of the eluate of the Ni-affinity column. **(B)** TEV protease cleavage and purification by Ni-affinity chromatography: (1) pooled fractions from first Ni-affinity column; (2) over-night reaction with TEV protease; (3) flowthrough by the second Ni-affinity chromatography collecting the cleaved BAP-*TbGSK3* short protein.



**Figure 3.17 Size-exclusion purification of recombinant *TbGSK3* short tagged at the N-terminus with BAP tag**

Protein samples were run on 4-12% SDS-PAGE gel and stained with Coomassie Blue. **(A)** Size exclusion chromatogram and **(B)** corresponding SDS-PAGE of analysed fractions: (1) His<sub>6</sub>-cleaved BAP-*TbGSK3* short pooled fractions; (2-10) labeled fractions in the chromatogram (A) with higher absorbance values. **(C)** Determination of native mass of *TbGSK3* short indicated with a gray square on the calibration curve obtained using as standards thyroglobulin (670 kDa), bovine gamma-globulin (158 kDa), chicken ovalbumin (44 kDa), equine myoglobin (17kDa).

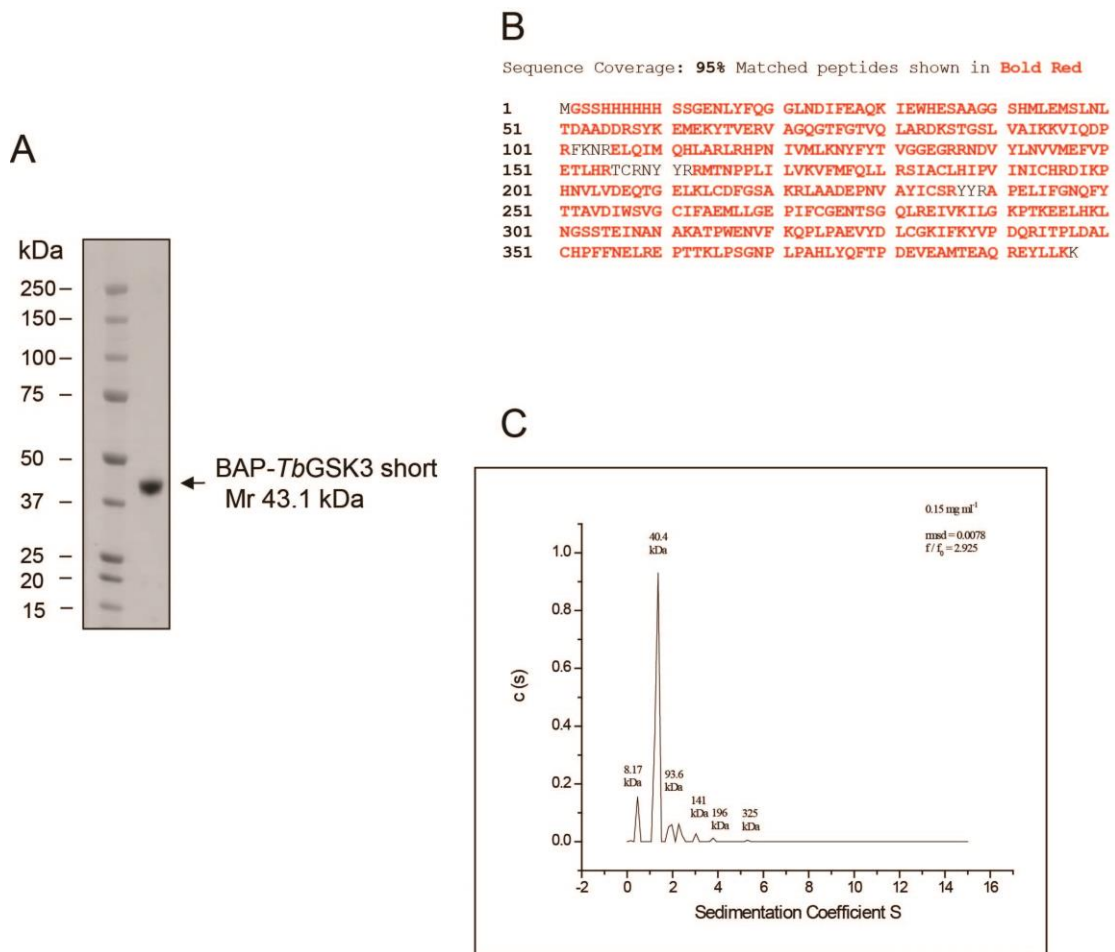
## 3.4 Protein expression and purification

### 3.4.1 Protein expression and characterization of BAP-tagged *Tb*GSK3 short

In order to support further studies of inhibitor characterization and chemical-genetic validation, I expressed recombinant *Tb*GSK3 short protein in house. The protein used for the hit identification (kindly provided by the University of Washington) had at the N-terminus a large MBP tag (45.1 kDa) (Ojo *et al.*, 2008). The attempts I made to determine the binding kinetics of GSK3 07 and 09 compounds using this construct by Surface Plasmon Resonance (SPR) were unsuccessful, due to the fact that the recombinant *Tb*GSK3 tagged with MBP was inactive after immobilization onto the CM5 chip by amino-coupling. This procedure includes a step at pH 5.5 that can cause loss of activity. Furthermore, proteins with bulky tags are not ideal for immobilization for SPR because they decrease the assay sensitivity due to the increased distance between the binding interaction and chip surface.

The Biotin Acceptor Peptide (BAP) tag, which can be site-specifically biotinylated by the biotin-protein ligase BirA (Predonzani *et al.*, 2008), improves the rate of success of SPR analysis by specific biotinylation in a known position that does not interfere with the activity of the protein and the immobilization at the chip surface can happen through the extremely strong interaction of biotin with streptavidin ( $K_d = 10^{15}$  M) (Predonzani *et al.*, 2008).

The *Tb*GSK3 short encoding region was cloned into the modified pET-15b-TEV-BAP vector and the expressed protein was purified by metal ion-affinity chromatography (**Figure 3.16 A**). The N-terminal His<sub>6</sub>-tag was subsequently cleaved by TEV protease and the cleaved protein purified by a second nickel column (**Figure 3.16 B**). The cleaved protein was further purified by size-exclusion chromatography (**Figure 3.17 A and B**). The protein had the expected Mr of 43.1 kDa by SDS-PAGE analysis,



**Figure 3.18 Characterization of recombinant *TbGSK3* short tagged at the N-terminus with BAP peptide**

(A) Protein sample after TEV cleavage and size exclusion purification. Protein sample was run on 4-12% SDS-PAGE gel and stained with Coomassie Blue. (B) Identity confirmation by Mass-fingerprint analysis of the His<sub>6</sub>-BAP tagged *TbGSK3* short performed by the Mass Proteomics Facility at the University of Dundee (Mr Douglas Lamont). (C) Plot of the distribution of sedimentation coefficients generated by analytical ultracentrifugation (AUC) of *TbGSK3* short at 0.15 mg/ml in buffer 20 mM Hepes pH 7.4, 1 mM DTT, 150 mM NaCl, 10% glycerol (Dr Mark Agacan).

but the elution of *TbGSK3* short from the gel filtration column indicated a tetrameric status (**Figure 3.17 C**). The protein identity was confirmed by mass-fingerprint analysis (**Figure 3.18 A and B**), and ultracentrifugation revealed that the protein was primarily monomeric, but aggregates were also present in solution (**Figure 3.18 C**). The detergent present in the purification buffer might be responsible for the increased tendency of the protein to aggregate.

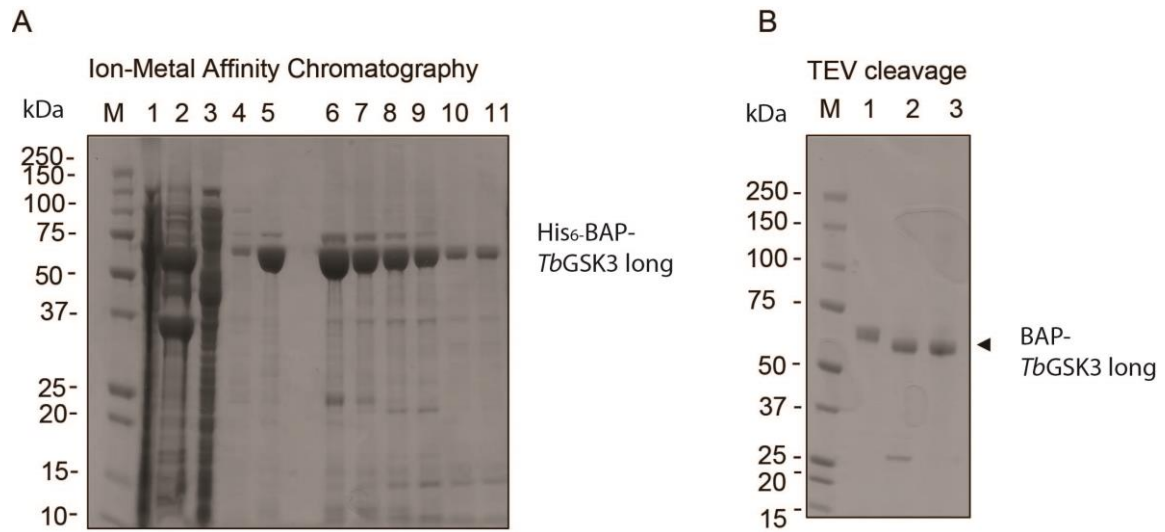
The protein yield for the His<sub>6</sub>-tagged protein was higher than 20 mg per litre of culture, but the cleavage of the His<sub>6</sub>-tag and the following steps of purification gave a final yield of 5 mg per litre.

The BAP-tagged *TbGSK3* short was active with a specific activity of 8.7 nM min<sup>-1</sup> nM<sup>-1</sup> at room temperature (201.4 nM min<sup>-1</sup> mg<sup>-1</sup>, MW 43,129.6 Da) that was ~2-times higher than the MBP-tagged *TbGSK3* preparation received from University of Washington that had an activity of 4 nM min<sup>-1</sup> nM<sup>-1</sup> (46.8 nM min<sup>-1</sup> mg<sup>-1</sup>, MW 85,391.6 Da).

Due to time constraints and limited access to instrumentation this *TbGSK3* short construct was not used for the biophysical characterization of the GSK3 inhibitors by Surface Plasmon Resonance, but it represents a useful reagent for further studies and potentially for undertaking fragment screening in the search for novel inhibitors with potential alternative mechanism of binding and inhibition.

### **3.4.2 Protein expression and characterization of *TbGSK3* long**

*TbGSK3* long protein has never been found expressed in bloodstream form trypanosomes in previous proteomic and phosphoproteomic studies (Urbaniak *et al.*, 2012b; Urbaniak *et al.*, 2013), whereas the peptide corresponding to the autophosphorylation loop of *TbGSK3* long has been identified by a phosphoproteomic analysis conducted in procyclic cells (NVPposphoYIFSR) (Nett *et al.*, 2009b), meaning



**Figure 3.19 Expression and purification of recombinant *TbGSK3* long tagged at the N-terminus with His<sub>6</sub>-BAP tag**

Protein sample were run on 4-12% SDS-PAGE gel and stained with Coomassie Blue.

**(A)** Purification by Ni-affinity column. (M) marker; (1) insoluble fraction; (2) clarified lysate by centrifugation; (3-4) flowthrough; (5-11) fractions at higher absorbance of the eluate of the Ni-affinity column. **(B)** TEV protease cleavage and purification by Ni-affinity chromatography: (1) pooled fractions from first Ni-affinity column; (2) over-night reaction with TEV protease; (3) flowthrough by the second Ni-affinity chromatography collecting the cleaved BAP-*TbGSK3* long protein.



that the protein is expressed in the insect life stage most likely in an active form. Moreover, RNAi studies have shown that this protein kinase is not essential in bloodstream form cells or for differentiation to procyclic form, and its knock-down is beneficial for the proliferation of procyclic cells (Alsford *et al.*, 2011; Jones *et al.*, 2014). This suggests that *TbGSK3* short and *TbGSK3* long have different expression and functions in the different life stages and they cannot compensate for each other.

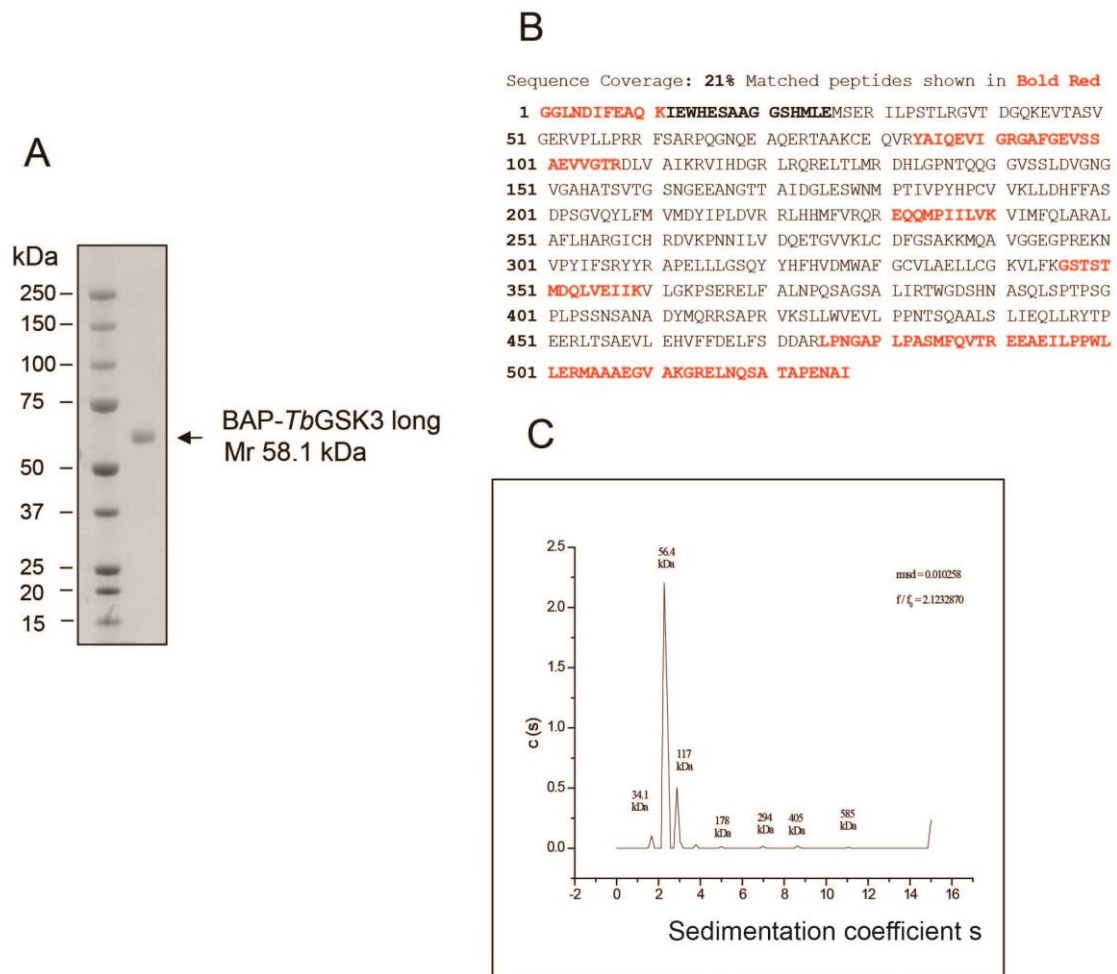
In order to address the question whether DDD00085893 could bind *TbGSK3* long *in vitro* and in cultured cells I expressed recombinant protein in *E. coli* and overexpressed this gene in BSF *T. b. brucei* (see section 4.1.8).

The open reading frame (ORF) region of the long form of *TbGSK3* was cloned from the genomic DNA of strain 427 (S427) using as reference the coding region reported in the genome for the strain 927 (S927) in GeneDB. The consensus sequence of three independent PCR reactions was compared to the DNA sequence reported in the reference S927 genome and 11 single nucleotide polymorphisms (SNPs) were identified between S427 and S927, four of these SNPs translated in four different amino acids (D15N, A51V, V201L and T322S).<sup>a</sup>

The ORF of the long isoform of *TbGSK3* was cloned in the same modified pET-15b-TEV-BAP plasmid used for the protein expression of *TbGSK3* short. The His-tagged form of the recombinant protein was expressed in *E. coli* and purified in a soluble form by metal-affinity chromatography; subsequently the protein was untagged by TEV cleavage and purified by a second nickel ion-affinity column (**Figure 3.19 A** and **B**). The protein had the expected Mr of 58.1 kDa by SDS-PAGE analysis (**Figure**

---

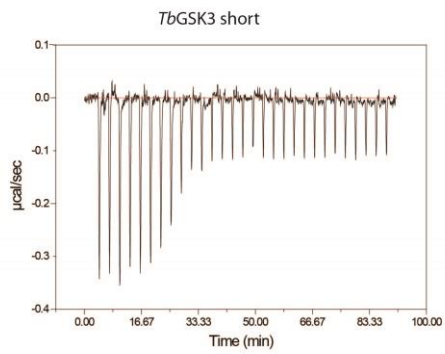
<sup>a</sup> These differences were in agreement with the genome sequence subsequently reported in Trytripsdb.org for *T. brucei* Lister strain 427.



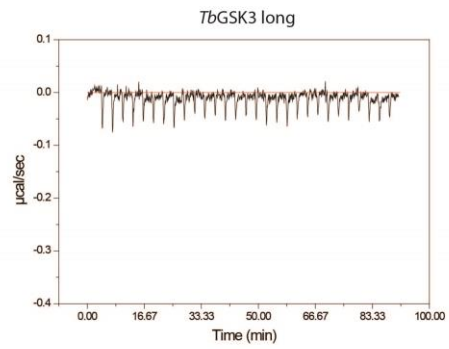
**Figure 3.20 Expression and purification of recombinant *TbGSK3* long tagged at the N-terminus with BAP peptide in *E. coli***

A) Protein sample after TEV cleavage and second nickel ion affinity column. Protein sample was run on 4-12% SDS-PAGE gel and stained with Coomassie Blue. B) Identity confirmation by Mass-fingerprint analysis performed by the Mass Proteomics Facility at the University of Dundee (Mr Douglas Lamont). C) Plot of the distribution of sedimentation coefficients generated by analytical ultracentrifugation (AUC) of *TbGSK3* long at 0.75 mg/ml in buffer 20 mM Hepes pH 7.4, 2 mM DTT, 150 mM NaCl (Dr Mark Agacan).

A



B



**Figure 3.21 Isothermal titration calorimetry of DDD00085893 with *TbGSK3 short* and *TbGSK3 long***

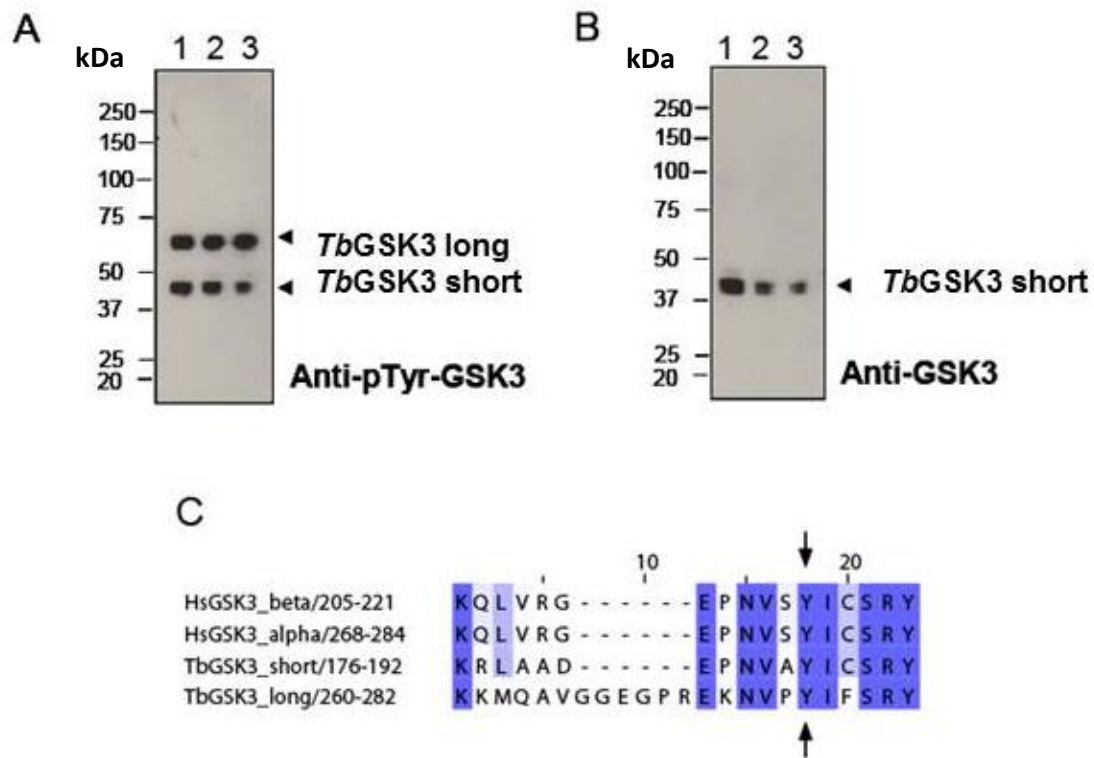
Isothermal titration calorimetry data for a solution of A) 3  $\mu\text{M}$  of *TbGSK3 short* and B) 3  $\mu\text{M}$  of *TbGSK3 long* titrated with 30  $\mu\text{M}$  DDD00085893. The two panels show the raw data.

**3.20** A), the identity was confirmed by mass-fingerprint and verified to be primarily a monomer by ultracentrifugation (**Figure 3.20** B and C).

The protein yield for the His<sub>6</sub>-tagged *TbGSK3* long was of 5 mg per litre of culture, but after the cleavage of the His<sub>6</sub>-tag the final yield was of 1.6 mg per litre.

Using the same peptide substrate previously used for the biochemical assay, I compared the enzymatic activity of *TbGSK3* long BAP-tagged at the N-terminus to the activity of *TbGSK3* short, BAP-tagged as well at the N-terminus. *TbGSK3* long was ~380-fold less active than *TbGSK3* short with a specific activity of 0.023 nM min<sup>-1</sup> nM<sup>-1</sup> (0.4 nM min<sup>-1</sup> mg<sup>-1</sup>, MW 58,126.3 Da). Neither enzyme was capable of autophosphorylation in the absence of substrate, but considering that both proteins cross-reacted with an antibody directed against the phosphorylated tyrosine of the activation loop (see following section), it is plausible that the proteins were already fully phosphorylated on this residue during protein production and no further autophosphorylation could happen. This is in agreement with what has been reported in literature that the autophosphorylation on the activation loop is an intramolecular reaction happening immediately after translation, during the protein folding and assisted by chaperone proteins, such as Hsp90 (Lochhead *et al.*, 2006).

Because of the extremely low enzymatic activity it was not possible to address the question whether DDD00085893 could inhibit *TbGSK3* long *in vitro* with an activity assay, but isothermal titration calorimetry (ITC), which does not require enzymatic activity, gave a preliminary indication that DDD00085893 does not bind *TbGSK3* long in solution (**Figure 3.21** A). On the other hand *TbGSK3* short did bind DDD00085893, but due to buffer mismatch it was not possible to integrate the results and obtain a  $K_d$  value (**Figure 3.21** B). Further experiments are also required to fully characterise the kinetics of binding of DDD00085893 to *TbGSK3* short.



**Figure 3.22 Western blot analysis of recombinant BAP-*TbGSK3* short and long**

Western blot analysis showed that only the p-Tyr-GSK3 antibody directed against the activation loop (C-KQLLHGEPNVS[pY]ICSRY), cross-reacted with the long form of *TbGSK3* protein expressed in *E. coli* (Figure 3.20 A). In contrast, the anti-GSK3 directed versus the same un-phosphorylated peptide did cross-react with the short isoform but not cross-react with the long isoform (Figure 3.20 B).

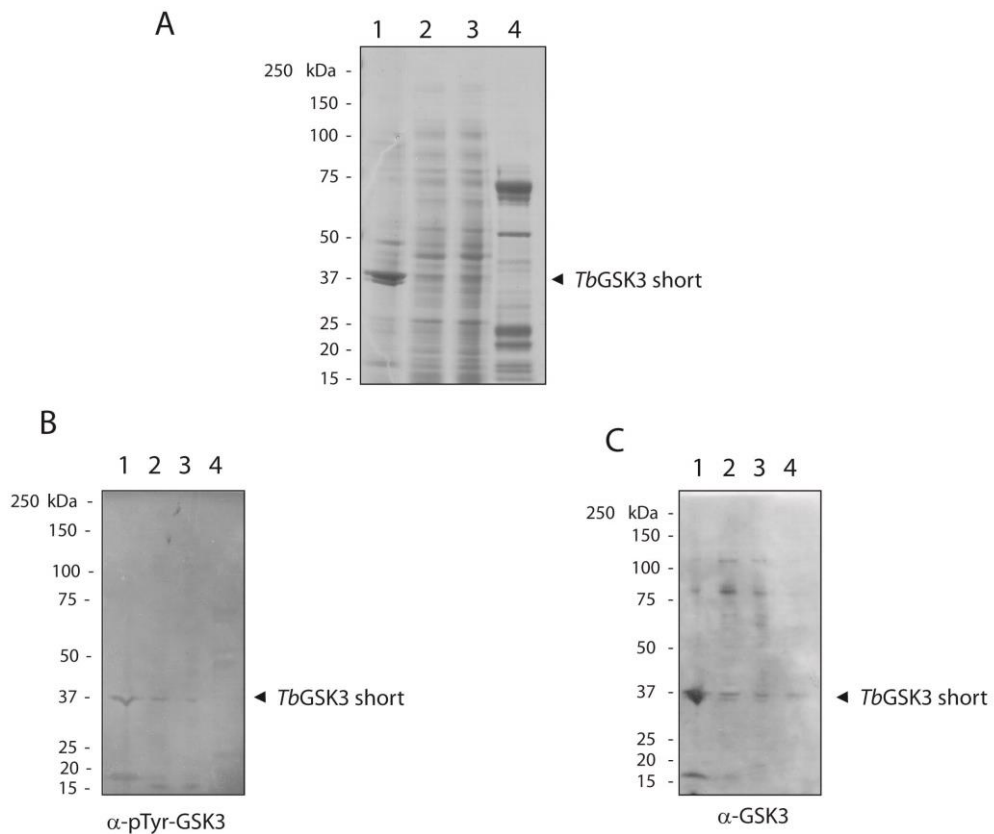
Serial dilutions (1:2 from 0.2  $\mu$ g to 0.05  $\mu$ g) of recombinant *TbGSK3* short and long probed with anti-phosphoTyr-GSK3 (A) and anti-GSK3 antibodies (B). (C) ClustalW alignment of the activation loop of *HsGSK3* and *TbGSK3* enzymes. The arrows indicate Tyr216 in *HsGSK3* $\beta$  and the conserved Tyr across species and isoforms.

### 3.4.3 Western blot analysis of recombinant *TbGSK3* short and *TbGSK3* long

Commercially available antibodies (Millipore) raised against both the phosphorylated and un-phosphorylated form of a synthetic peptide (KQLLHGEPNVS[phosphoY]ICSRY) corresponding to the autophosphorylation loop of the *Drosophila* GSK3/shaggy enzyme were tested for cross-reactivity against recombinant *TbGSK3* short and long. Considering that this region is highly conserved in all species, the antibodies were also expected to recognise the *T. brucei* isoforms of the enzyme.

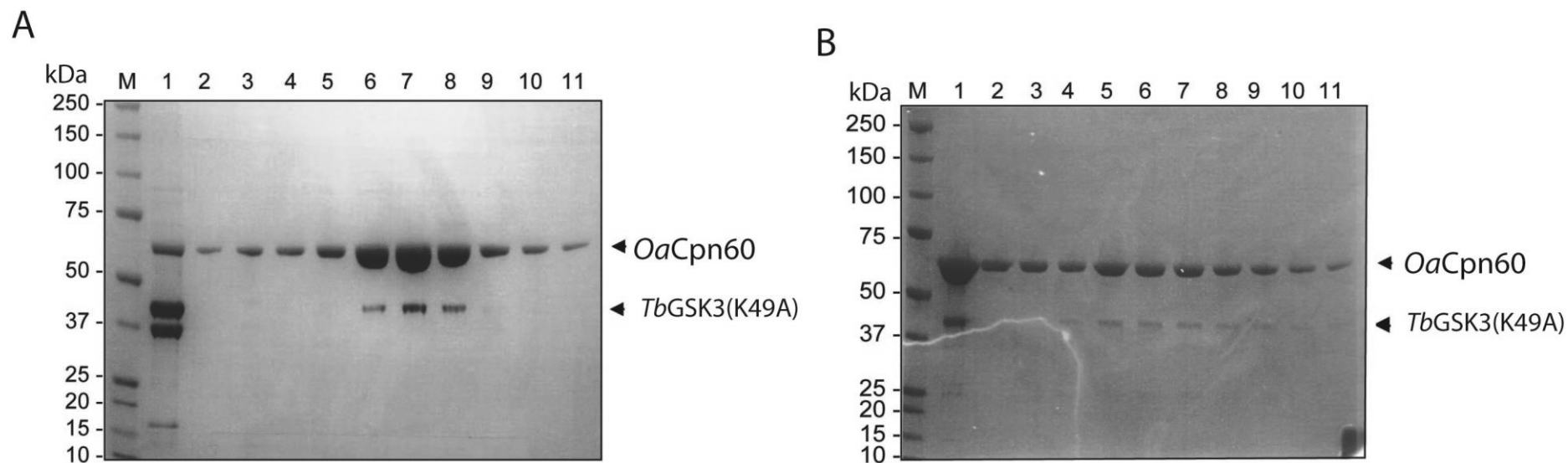
Western blot analysis conducted on the recombinant proteins showed that the phospho-Tyr-GSK3 antibody directed against the activation loop (C-KQLLHGEPNVS[pY]ICSRY), cross-reacted with both the short and long isoforms of *TbGSK3* expressed in *E. coli* (**Figure 3.22 A**). In contrast, the GSK3 antibody directed against the same un-phosphorylated peptide cross-reacted only with the short isoform (**Figure 3.22 B**). The insertion of six amino acids in the antibody recognition sequence of *TbGSK3* long might account for the lack of cross-reactivity with the antibody directed against the unphosphorylated substrate (**Figure 3.22 C**).

Together with the activity data these results indicate that the *TbGSK3* long gene encodes for a protein that can be expressed in *E. coli* in a soluble form, that can autophosphorylate in the activation loop and trans-phosphorylate a primed substrate, although with an extremely low specific activity with respect to the short isoform. Interestingly *TbGSK3* long has in its sequence an Arg residue instead of a catalytically important Lys residue (K86 in *HsGSK3 $\beta$* , R88 in *TbGSK3* long). Usually this mutation is used to make kinase dead mutants and this might explain why the enzyme activity is so low and lack of significant catalytic activity might also explain why this isoform is not essential and not expressed or expressed at undetectable levels in *T. brucei*. It might also be the case that the protein is also less stable and easily degraded because of this



**Figure 3.23 Expression and purification of HIS<sub>6</sub>-BAP-*TbGSK3*(K49A) in *E. coli* and western blot analysis**

(A) *TbGSK3*(K49A) expression and partial purification from BL21(DE3)\* cells cultured in auto-induction media at 22 °C for 24 hours before lysis. Coomassie blue staining of SDS-PAGE: (1) insoluble fraction; (2) clarified bacterial cell lysate by centrifugation (19,000 rpm for 30 min at 4 °C); (3) flowthrough; (4) pooled fractions purified by Ni-affinity chromatography. The same samples were transferred into two nitrocellulose blots and probed with antibodies directed against anti-phosphotyrosine-GSK3 ( $\alpha$ -pTyr-GSK3) (B) and anti-GSK3 ( $\alpha$ -GSK3) (C). Proteins at the expected size of *TbGSK3* short (40.3 kDa) are marked by arrows. The HIS<sub>6</sub>-BAP-tagged *TbGSK3*(K49A) has a molecular weight of 45.3 kDa.



**Figure 3.24 Purification of recombinant HIS<sub>6</sub>-BAP-*TbGSK3(K49A)* in *E. coli***

(A) *TbGSK3(K49A)* expression and partial purification from ArcticExpress (DE3) RIL cells cultured in auto-induction media at 13 °C for 48 h before lysis. Coomassie blue stained SDS-PAGE: (M) marker; (1) insoluble fraction; (2) clarified lysate by centrifugation (19,000 rpm for 30 min at 4 °C) (dil 1/10); (3) flowthrough from the Ni-affinity column; (4) washes with binding buffer; (5-11) fractions at higher absorbance of the eluate of the Ni-affinity column. (B) Second Ni-affinity column after incubation at 4 °C for 1 h in Buffer A supplemented with 1 mM ATP, 1 mM MgCl<sub>2</sub>. Coomassie blue stained SDS-PAGE: (M) marker; (1) pooled fractions after the first Ni-affinity and following purification by ion exchange chromatography; (2) flowthrough from the second Ni-affinity column; (3) washes; (4-11) fractions at higher absorbance of the eluate from the second Ni-affinity column. In both panels the arrows indicate the expected Mr for HIS<sub>6</sub>-BAP-tagged *TbGSK3(K49A)* (45.3 kDa) and for *Oleispira antarctica* Cpn60 (60 kDa).



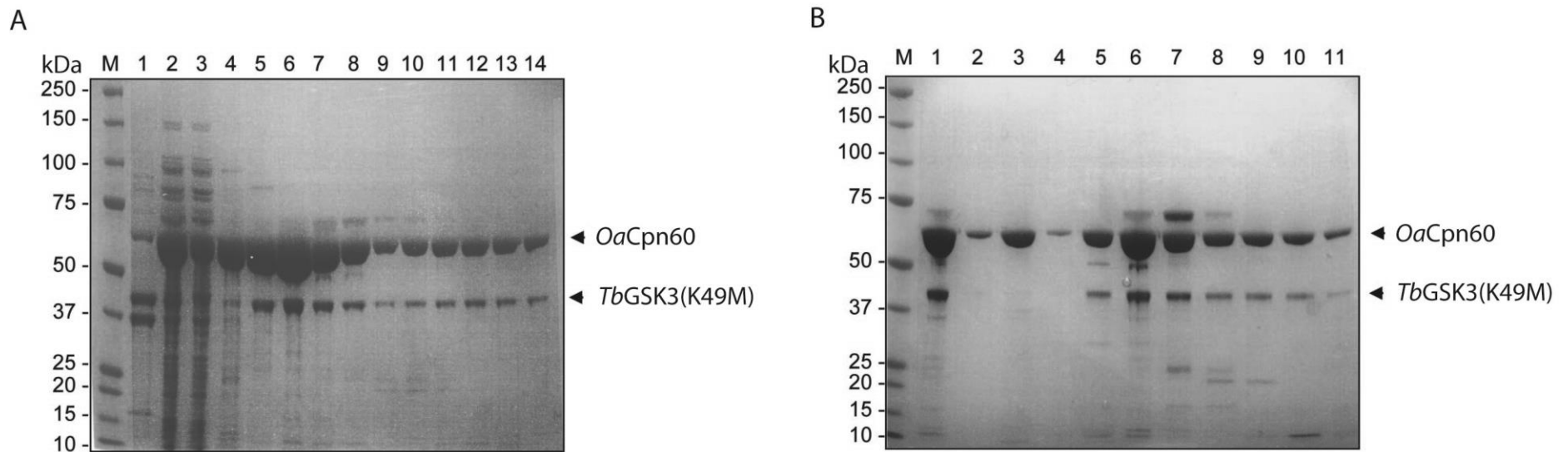
mutation, the reason for this could be that usually autophosphorylation contributes to the stability of the protein, but, given that the protein is almost inactive, a lack of autophosphorylation might cause instability and degradation.

#### **3.4.4 Protein expression and purification of *Tb*GSK3 short inactive mutant**

In order to address the question whether DDD00085893 could bind the inactive form of *Tb*GSK3 short *in vitro* and in cultured cells, kinase dead mutants of *Tb*GSK3 short were expressed in *E. coli* and overexpressed in *T. b. brucei* cultured cells (see section 4.1.7).

Based on the homology model of mammalian GSK3 $\beta$  the catalytic Lys85 corresponding to Lys49 in *Tb*GSK3 short was mutated to Ala (K49A) or Arg (K49R) to obtain kinase-dead mutants (He *et al.*, 1995; Eldar-Finkelman *et al.*, 1996).

The mutated *Tb*GSK3(K49A) short was cloned into the pET-15b-TEV-BAP plasmid (**Figure 2.1**). The plasmid was transformed in BL21(DE3)\* cells for protein expression in auto-induction medium at 22°C for 20 h (Studier, 2005) using similar conditions previously reported for the expression of the N-terminus His<sub>6</sub>-tagged *Tb*GSK3 short (Ojo *et al.*, 2011). The mutant protein was poorly expressed in these conditions and primarily insoluble (**Figure 3.23** lane 1) as demonstrated also by Western analysis (**Figure 3.23** B and C). An increase in protein expression and solubility was obtained using ArcticExpress (DE3) cells due to the co-expression of chaperonins Cpn60 and Cnp10 from *Oleispira antarctica* (Ferrer *et al.*, 2003; Ferrer *et al.*, 2004). Indeed, *Tb*GSK3(K49A) short was co-purified with a contaminant protein at higher molecular weight (~60 kDa), which mass-fingerprint analysis confirmed to be the *Oa*Cpn60. The attempts to separate this co-purified contaminant by anion exchange chromatography (data not shown) and by incubation with the binding buffer supplemented with 1 mM ATP, 1 mM MgCl<sub>2</sub> were unsuccessful (**Figure 3.24** A and B).



**Figure 3.25 Purification of recombinant HIS<sub>6</sub>-BAP-TbGSK3(K49M) in *E. coli***

(A) *TbGSK3(K49M)* expression and partial purification from ArcticExpress (DE3) RIL cells cultured in autoinduction media at 13 °C for 48 h. Coomassie blue stained SDS-PAGE: (M) marker; (1) insoluble fraction; (2) clarified lysate by centrifugation (19,000 rpm for 30 min at 4 °C); (3) flowthrough from the Ni-affinity column; (4) washes with binding buffer supplemented with 10 mM MgCl<sub>2</sub> and 1 mM ATP; (5-14) fractions at higher absorbance of the eluate of the Ni-affinity column. (B) Second Ni-affinity column after extensive washes with binding buffer containing 5 mM ATP, 10 mM MgCl<sub>2</sub> and 0.1 mg/ml denaturated proteins from *E. coli*. Coomassie blue stained SDS-PAGE: (M) marker; (1) pooled fractions from the first Ni-affinity column dialysed overnight in buffer A without imidazole; (2) flowthrough; (3) washes with 5 mM ATP, 10 mM MgCl<sub>2</sub> and 0.1 mg/ml denaturated proteins from *E. coli*; (4-11) fractions at higher absorbance of the eluate from the Ni-ion affinity column. In both panels the arrows indicate the expected Mr for HIS<sub>6</sub>-BAP-tagged *TbGSK3(K49M)* (45.3 kDa) and for *Oleispira antartica* Cpn60 ( ~60 kDa).

Similar results were also obtained when an alternative kinase dead mutant was generated by replacement of Lysine49 with a Methionine, obtaining the kinase dead mutant called *TbGSK3(K49M)* short. This mutant protein also co-purified with *OaCpn60* when expressed from ArcticExpress (DE3) cells at 13 °C for 48 h (**Figure 3.25 A**). Also, in this case attempts to separate the contaminating protein after extensive washes with binding buffer containing 5 mM ATP, 10 mM MgCl<sub>2</sub> and 0.1 mg ml<sup>-1</sup> denaturated proteins from *E. coli* were unsuccessful (Rial *et al.*, 2002) (**Figure 3.25 B**).

In conclusion, the expression of the kinase dead mutants *TbGSK3(K49A)* and *TbGSK3(K49M)* from ArcticExpress(DE3) cells resulted in partially soluble proteins tightly associated with a chaperonin. It is likely that also in cultured BSF *T. b. brucei* these mutants are unstable when over-expressed (see section 4.1.6). Indeed, it is reported that autophosphorylation in the activation loop increases the stability of *HsGSK3β* (Cole *et al.*, 2004) and it is an intramolecular chaperone-dependent event that occurs after translation (Lochhead *et al.*, 2006). This might also explain why the inactive mutant protein would be so tightly bound to the chaperonin protein and failed to progress from this transitional intermediate to the fully active protein.

Due to protein instability it is not possible to confirm whether DDD00085893 is able to bind the kinase dead mutants *TbGSK3(K49A)* and *TbGSK3(K49M)*.

Interestingly a band at the expected molecular weight for *TbGSK3(K49A)* short was observed in the Western analysis also with the anti-phosphoTyr antibody directed against the activation loop (Figure 3.23 B), this is surprising considering that this is an intramolecular autophosphorylation, meaning that this expressed protein although unstable is not completely inactive. Further work would be required to test the activity of these preparations using an in gel kinase assay.

**Chapter 4 Chemical-genetic validation of  
*Tb*GSK3 short**

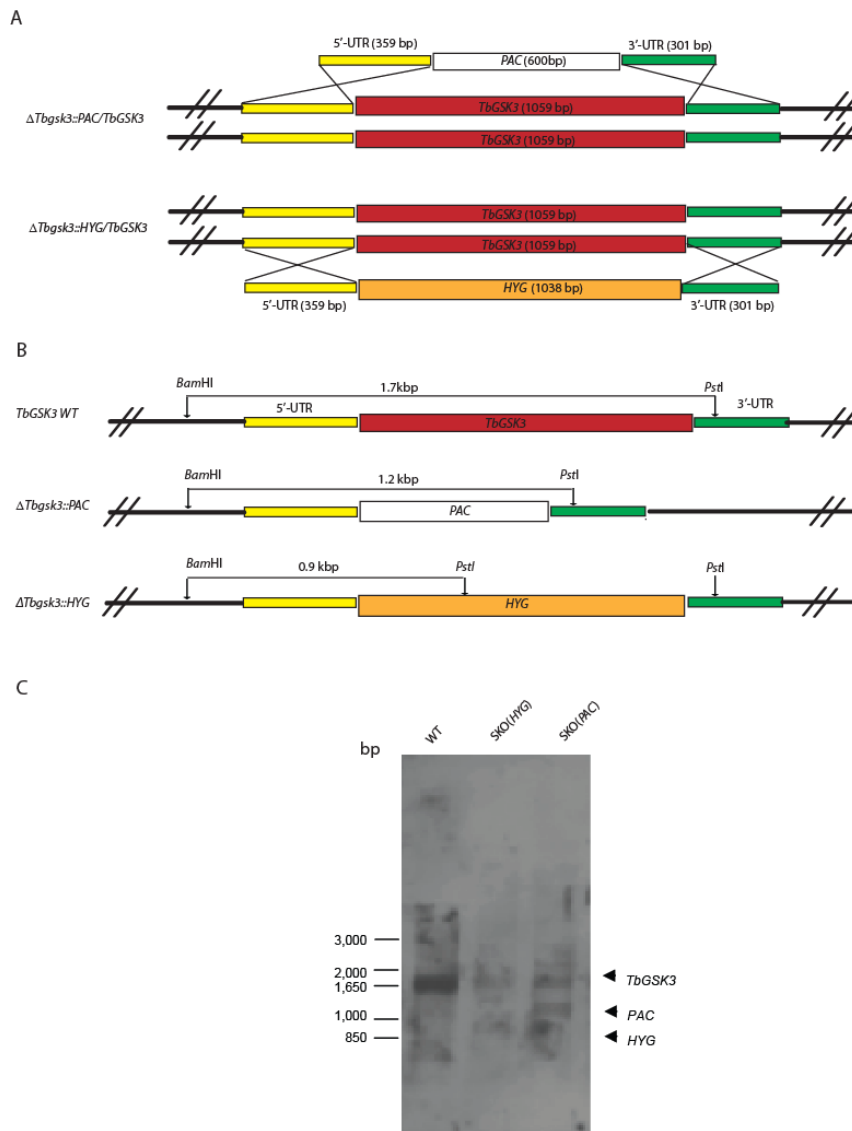
## 4.1 Chemical validation using genetic approach

### 4.1.1 Rationale of the chemical-genetic approach

*TbGSK3* short was initially proposed as a drug target for human African trypanosomiasis by Ojo and colleagues (2008). Phosphoproteomic studies conducted in *T. brucei* revealed that this protein is expressed in both bloodstream (Nett *et al.*, 2009a) and in procyclic trypanosomes (Nett *et al.*, 2009b). In both the parasite life stages the activation loop of *TbGSK3* short was found to be phosphorylated, an indication that this kinase should be enzymatically active considering that in other species this is an auto-phosphorylation event (Eldar-Finkelman *et al.*, 1996; Cole *et al.*, 2004).

RNA interference (RNAi) studies demonstrated that knock-down of *TbGSK3* short causes a growth defect in BSF *T. brucei*, while the knock-down of the long isoform has no effect on cell growth (Ojo *et al.*, 2008). Subsequently Alsford and colleagues (2011) conducted a genome-wide RNA interference target sequencing study directed against both bloodstream and procyclic *T. b. brucei*. Their findings confirmed that RNAi of *TbGSK3* short resulted in inhibition of proliferation and differentiation in bloodstream form only. Recently another RNAi study has confirmed that *TbGSK3* short knock-down in bloodstream form trypanosomes causes arrest of growth and plays a role in mitosis and/or cytokinesis (Jones *et al.*, 2014).

To date this potential drug target has never been genetically validated by generation of knock-outs and the chemical compounds used by others for chemical validation of *TbGSK3* short were not sufficiently specific to clearly attribute their cellular effects to exclusive “on target” activity (Ojo *et al.*, 2008; Oduor *et al.*, 2011; Ojo *et al.*, 2011). The chemical proteomic approach we performed in collaboration with



**Figure 4.1 Generation of *TbGSK3* single knock-out (SKO) cell lines**

(A) Schematic representation of the generation of the *TbGSK3* single knock-out (SKO) cell lines. The replacement of an endogenous allele by incorporation of the puromycin-resistance gene (*PAC*) produced  $\Delta Tbgsk3::PAC/TbGSK3$  cell line (SKO(*PAC*), top) and the replacement with the hygromycin-resistance gene (*HYG*), produced  $\Delta Tbgsk3::HYG/TbGSK3$  cell line (SKO(*HYG*), bottom). (B) Analysis of restriction sites for *Bam*HI/*Pst*I for *TbGSK3/TbGSK3* (WT),  $\Delta Tbgsk3::PAC/TbGSK3$  (SKO(*PAC*)) and  $\Delta Tbgsk3::HYG/TbGSK3$  (SKO(*HYG*)) loci and expected fragment sizes when probed with *TbGSK3* 5'-UTR. (C) Southern analysis of ~5  $\mu$ g of DNA digested with *Bam*HI and *Pst*I from WT, SKO(*PAC*) and SKO(*HYG*) and probed with *TbGSK3* 5'-UTR.

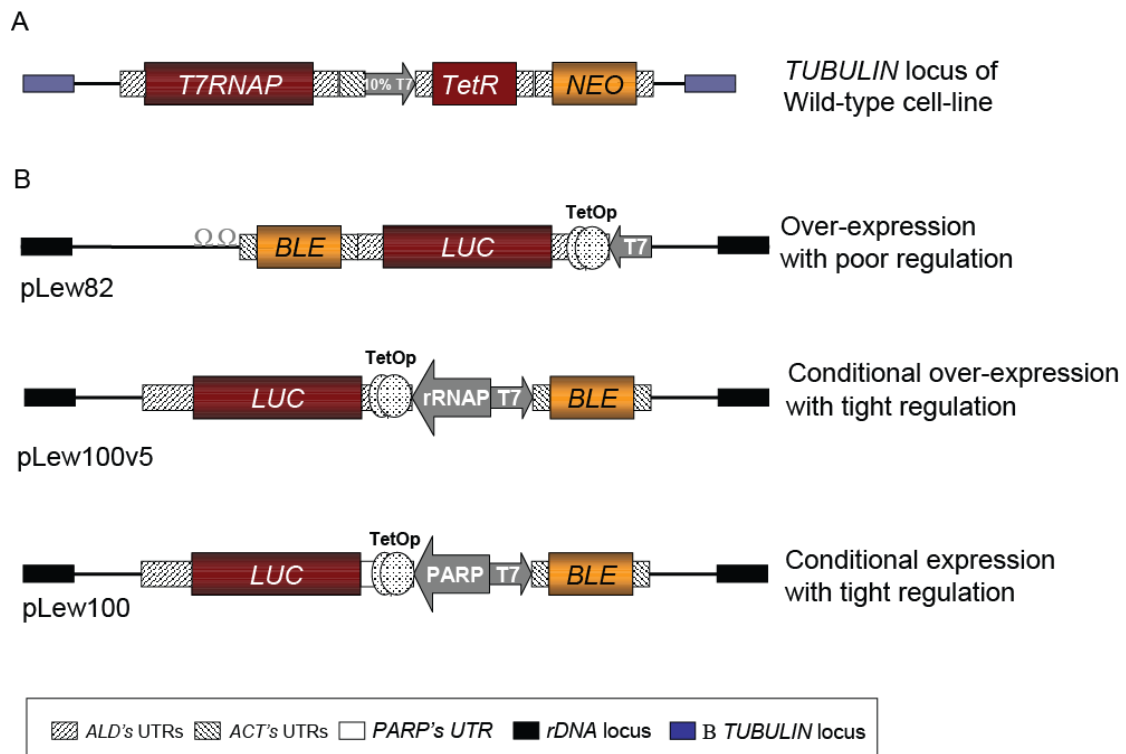
Cellzome demonstrated that DDD00085893 binds to the *TbGSK3* short native protein with high affinity (IC<sub>50</sub> below 40 nM) and is also a specific GSK3 inhibitor for both *T. brucei* and human MRC5 cells (see section 3.3 and Urbaniak *et al.*, 2012a). Next I wanted to assess whether this compound could specifically act “on target” in living cells.

To date the role of *TbGSK3* short has not yet been elucidated, and its physiological substrates in *T. brucei* are unknown; hence it was not possible to validate DDD00085893 directly by inhibiting the phosphorylation of *TbGSK3* short downstream targets. An alternative strategy to demonstrate that DDD00085893 specifically inhibits *TbGSK3* short in cultured trypanosomes is by modulating its protein expression in mutant cell lines. An inhibitor that acts “on target” is more potent against mutant cell lines under-expressing the molecular target and becomes less potent when the target protein is overexpressed (Wyllie *et al.*, 2009), or in other words the potency to a small molecule inhibitor acting “on target” can be modulated by changing the level of protein expression of its molecular target.

#### **4.1.2 Null mutants of *TbGSK3* short have a lethal phenotype**

It was previously demonstrated that *TbGSK3* short gene is present as a single copy per haploid genome by Southern analysis (Ojo *et al.*, 2008).

Single knock-out cell lines (SKO) were obtained by replacement of one *TbGSK3* allele with the puromycin N-acetyl transferase resistance gene (*PAC*) flanked with the 5'- and 3'- untranslated regions (5'-UTR and 3'-UTR) of *TbGSK3* short generating the  $\Delta Tbgsk3::PAC/TbGSK3$  cell line. Likewise it was possible to incorporate the hygromycin phosphotransferase resistance gene (*HYG*) generating the  $\Delta Tbgsk3::HYG/TbGSK3$  cell line. (**Figure 4.1** A and B). The replacement of a single



**Figure 4.2 Strategy for inducible expression of an ectopic copy**

(A) The WT cell line constitutively expresses T7 RNA polymerase and the tetracycline repressor protein (TetR) under G418-neomycin (*NEO*) selection. (B) Inducible expression vectors designed for integration in the ribosomal locus. Figure adapted from Wirtz *et al.*, 1999 and Professor George Cross' website ([http://tryps.rockefeller.edu/trypsru2\\_plasmids.html](http://tryps.rockefeller.edu/trypsru2_plasmids.html)).

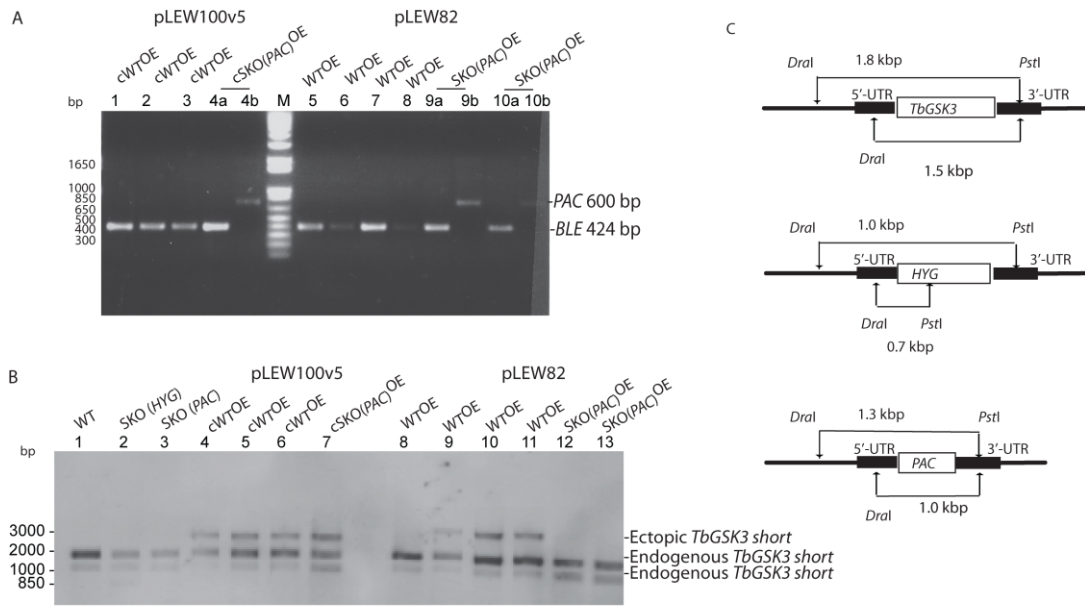


allele with the resistance gene cassettes in the *TbGSK3* locus was confirmed by Southern analysis (**Figure 4.1 C**). Repeated attempts to obtain a null mutant replacing the second allele by homologous recombination with either *PAC* or *HYG* were unsuccessful. Similarly also for other essential genes studied up to date in *T. brucei*, in which there is not a metabolic by-pass or compensation of function by other isoforms and/or enzymes, it has not been possible to generate null mutants (Roper *et al.*, 2002; Sienkiewicz *et al.*, 2008; Urbaniak, 2009; Wyllie *et al.*, 2009). This negative result suggests that *TbGSK3* short is an essential gene in *T. brucei*, although it cannot be considered a definitive proof of its essentiality (Frearson *et al.*, 2007).

#### **4.1.3 Generation of conditional cell lines and tuning of the ectopic expression**

When it is not possible generate a null mutant directly by replacement of both alleles, an established procedure is to insert an ectopic copy of the gene in an alternative locus under tetracycline-inducible regulation (Wirtz *et al.*, 1999). This strategy requires a background cell line, referred as wild-type (WT), which constitutively expresses T7 RNA polymerase (T7RNAP) and the tetracycline repressor protein (TetR) under G418-neomycin (*NEO*) selection (**Figure 4.2 A**) (Wirtz *et al.*, 1999). The ectopic copy is generally inserted in the ribosomal locus, its expression being regulated by a tetracycline-inducible promoter. Selection is possible by the presence of a resistance cassette regulated by a constitutive T7 promoter. Different vectors can induce different level of expression of the ectopic protein based upon their “leaky” nature or tighter regulation, so that the ectopic protein expression can be finely tuned, in particular for toxic genes (Wirtz *et al.*, 1999).

For the generation of *TbGSK3* short over-expressing cell lines, the pLew82 over-expressor vector was initially used as it gives the highest level of expression at the



### Figure 4.3 Genotypic analyses of conditional cell lines

**(A)** Agarose gel of PCR reactions performed with 100 ng genomic DNA from conditional cell lines cW<sup>TOE</sup> *GSK3short*<sup>Ti</sup> (*BLE*, pLew100v5) (lane 1-3), cSKO(PAC)<sup>OE</sup> *GSK3short*<sup>Ti</sup> (*BLE*, pLew100v5)(lane 4a-4b), WT<sup>OE</sup> *GSK3short*<sup>Ti</sup> (*BLE*, pLew82) (lane 5-8), and SKO(PAC)<sup>OE</sup> *GSK3short*<sup>Ti</sup> (*BLE*, pLew82) (lane 9a-10b) to confirm incorporation of resistance genes, *PAC* (600bp) and *BLE* (424 bp). M, marker. **(B)** Southern blot analysis of *Dra*I and *Pst*I digested genomic DNA (~5 µg) from WT (lane 1), *TbGSK3short* SKO (*HYG*) (lane 2), *TbGSK3short* SKO (*PAC*) (lane 3), cW<sup>TOE</sup> *GSK3short*<sup>Ti</sup> (*BLE*, pLew100v5) (lane 4-6), cSKO(PAC)<sup>OE</sup> *GSK3short*<sup>Ti</sup> (*BLE*, pLew100v5)(lane 7), WT<sup>OE</sup> *GSK3short*<sup>Ti</sup> (*BLE*, pLew82) (lane 8-11), and SKO(PAC)<sup>OE</sup> *GSK3short*<sup>Ti</sup> (*BLE*, pLew82) (lane 12-13). Southern blot probed with *TbGSK3 short* ORF. WT shows the expected size of 1.5 kbp (completed digested DNA) and 1.8 kbp (partial digested DNA). The membrane had previously been probed with the 5'-UTR probe and stripped. The expected extra band at 0.7 kbp for the introduction of *HYG* in the *TbGSK3short* SKO (*HYG*) cell line is still visible in the blot (lane 2). **(C)** Map of the restriction sites for *Dra*I and *Pst*I in WT cell lines, *TbGSK3short* SKO (*HYG*) and *TbGSK3short* SKO (*PAC*).

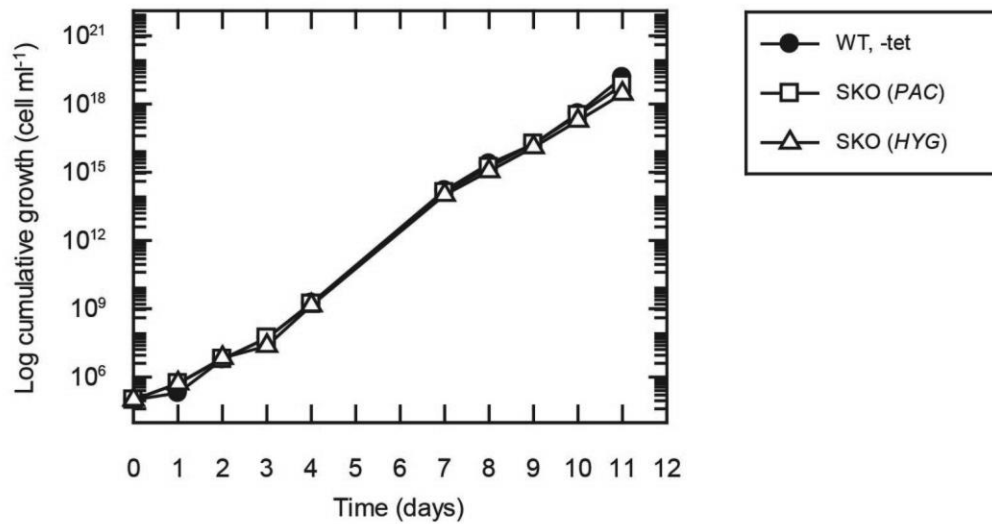
expense of a poor regulation since it is controlled by a leaky tetracycline-inducible T7 promoter. This vector has a single T7 promoter that drives transcription of both the ectopic gene and the drug resistance gene, whereas transcription is terminated downstream by tandem T7 terminators (**Figure 4.2 B**). The unregulated over-expression of *TbGSK3* short by the pLew82 vector caused either cell toxicity or deletion of the ectopic copy by gene rearrangement, although the gene conferring resistance was maintained (lane 5, 9a and 10a in **Figure 4.3 A** and lane 12 and 13 in **Figure 4.3 B**). These results suggested that the level of *TbGSK3* short protein in *T. brucei* is tightly regulated and expression above the physiological levels is as lethal as gene deletion. These findings were in agreement with the fact that over-expression of *HsGSK3 $\beta$*  was able to induce apoptosis in a variety of cell lines (Pap *et al.*, 1998; Beurel and Jope, 2006).

The pLew100 vector is tightly regulated by a PARP (procyclic acidic repetitive protein) promoter with two Tet operators, thus the level of conditional expression of the ectopic copy under induction is comparable to the WT level (**Figure 4.2 B**). Therefore, it is generally used as a rescue construct for generation of conditional null mutants, but it is not suitable for protein over-expression above physiological levels (Wirtz *et al.*, 1999). A modified version of the pLew100 vector, called pLew100v5<sup>a</sup> combines expression above physiological levels with a tighter regulation that can be finely tuned by tetracycline concentration. This vector has a back-to-back dual promoter, a constitutive T7 promoter which drives the drug resistance and a regulated rRNA polymerase (rRNAP) that drives the expression of the ectopic gene (**Figure 4.2 B**). Using pLew100v5 it was possible to introduce an ectopic copy of *TbGSK3* short both

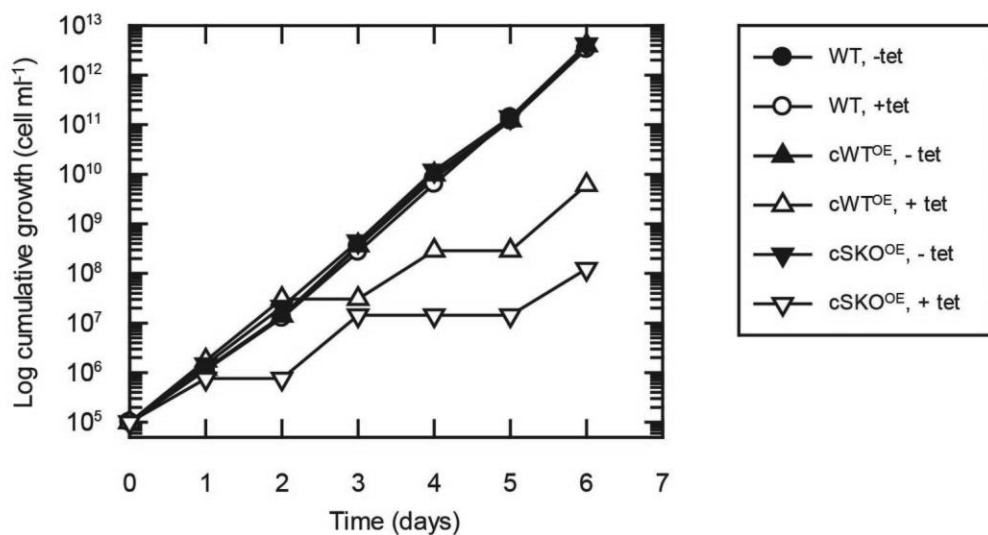
---

<sup>a</sup> [http://tryps.rockefeller.edu/trypsru2\\_plasmids.html](http://tryps.rockefeller.edu/trypsru2_plasmids.html)

A



B



**Figure 4.4 Cumulative growth analyses of wild-type and genetically modified *TbGSK3* short cell lines**

(A) WT growth was compared to the growth of SKO (*PAC*) and SKO (*HYG*) clonal cell lines. (B) WT growth in the absence (closed symbols) and in the presence (open symbols) of 1  $\mu\text{g ml}^{-1}$  of tetracycline (tet) was compared to the cumulative growth of a conditional over-expressing clone (cWT<sup>OE</sup>) and a *TbGSK3* conditional single knock-out clone (cSKO<sup>OE</sup>).

into WT and SKO background cell lines generating conditional over-expressing (cWT<sup>OE</sup>) and conditional single knock-out (cSKO<sup>OE</sup>) cell lines. PCR and southern blot analysis of genomic DNA from these cell lines confirmed that the ectopic copy under phleomycin (*BLE*) selection was introduced and retained (lane 1 to 4a in **Figure 4.3 A** and lane 4 to 7 in **Figure 4.3 B**).<sup>a</sup>

#### 4.1.4 Over-expression of *TbGSK3* short is toxic for cell viability

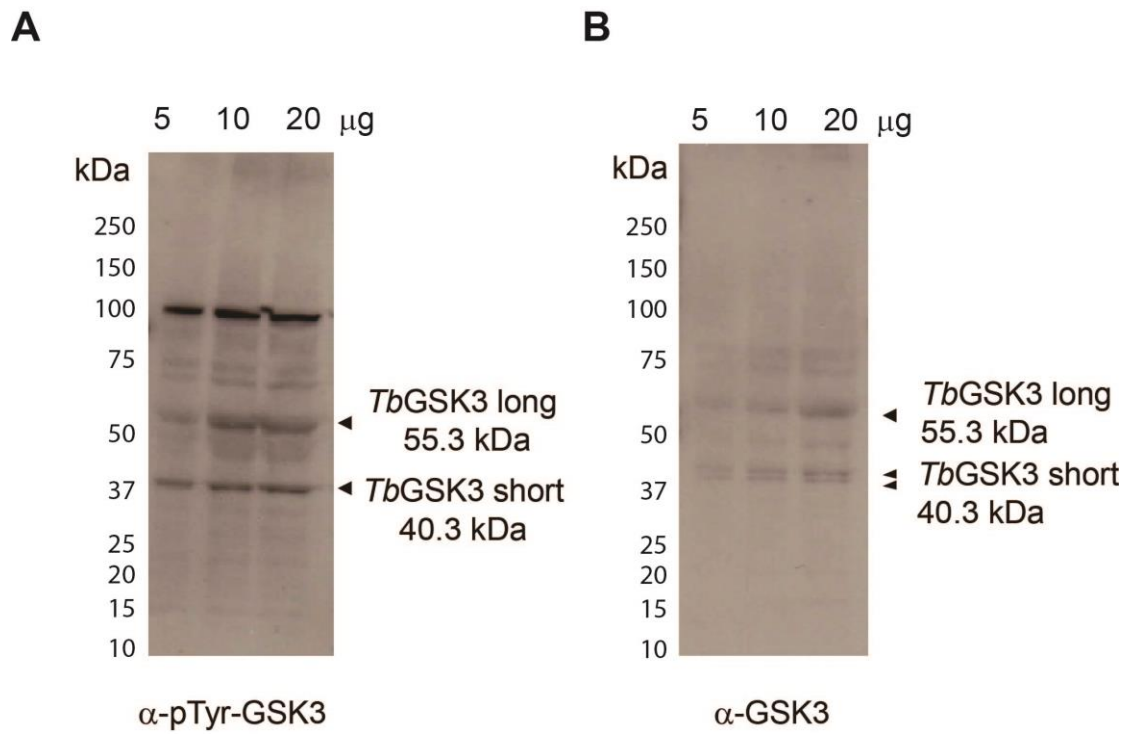
The growth of the mutant cell lines generated using pLew100v5 as conditional over-expressor was compared to the growth of WT cell lines. SKO clones (both SKO (*PAC*) and SKO (*HYG*)) showed comparable growth to the WT (**Figure 4.4 A**). The growth of the conditional over-expressing cell lines in the presence and absence of tetracycline (tet) compared to WT trypanosomes showed growth defects when induced by tetracycline around day 2-3 and an erratic growth profile (**Figure 4.4 B**). The defect in growth observed under tetracycline induction supports the indication that *TbGSK3* short over-expression is toxic for the cells.

#### 4.1.5 Expression of *TbGSK3* short and long in BSF *T. b. brucei*

Bloodstream form *T. b. brucei* lysates were probed with the same anti-phosphotyrosine GSK3 antibody that cross-reacted with recombinant *TbGSK3* short and

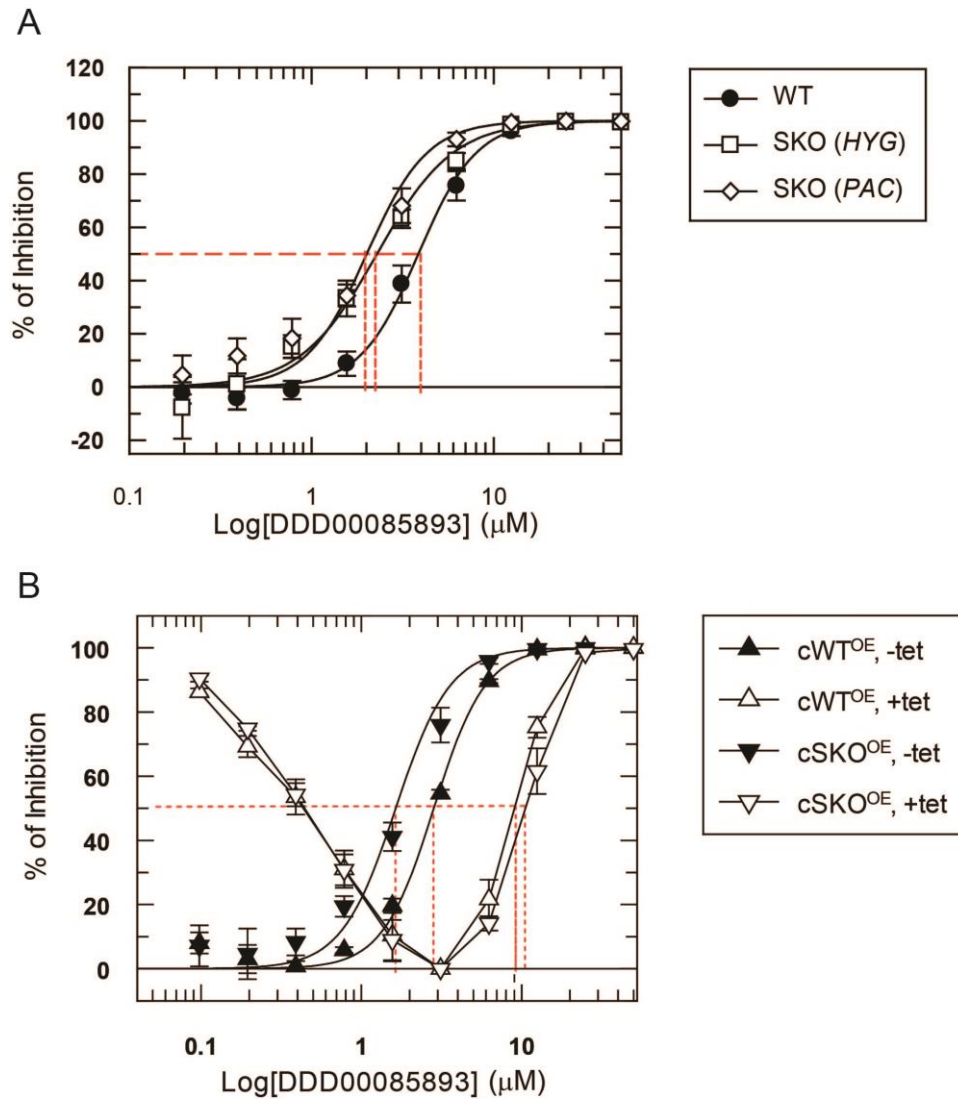
---

<sup>a</sup> The restriction enzymes used in the Southern analysis (*DraI* and *PstI*) gave two bands for the WT when probed with the *TbGSK3* short probe due to the presence of two *DraI* restriction sites in the 5'-UTR region of *TbGSK3* short and partial digestion (see **Figure 4.3 B** and **C**). The partial digestion of the DNA does not compromise the validity of the information obtained regarding the incorporation of the ectopic copy and therefore the blot is shown as evidence of ectopic copy incorporation. To avoid this problem in further Southern analysis a new pair of restriction enzymes was chosen (*BamHI* and *PstI*, see **Figure 4.1B**).



**Figure 4.5 Western blot analysis of BSF *T. b. brucei***

Immunoblots of whole cell extracts (5, 10 and 20  $\mu\text{g}$ ) from BSF *T. b. brucei* were probed with **(A)** anti-phospho-GSK3 ( $\alpha\text{-pTyr-GSK3}$ ) and **(B)** anti-GSK3 ( $\alpha\text{-GSK3}$ ) antibodies in two independent blots. In both panels the arrows indicate the expected Mr for *TbGSK3* short (40.3 kDa) and for *TbGSK3* long (55.3 kDa).



**Figure 4.6 Targeting of *TbGSK3* short by DDD00085893 in bloodstream form *T. b. brucei***

(A)  $\text{EC}_{50}$  values determination against WT (closed circles), SKO (*HYG*) (open squares) and SKO (*PAC*) (open diamonds). (B)  $\text{EC}_{50}$  determination against *TbGSK3* short-overexpressing (cWT<sup>OE</sup>) and conditional SKO (cSKO<sup>OE</sup>) cells non-induced (close upward triangles and close downward triangles, respectively). Tetracycline-induced cWT<sup>OE</sup> (open upward triangles) and cSKO<sup>OE</sup> (open downward triangles) showed cell toxicity that could be rescued by increasing concentration of DDD00085893.  $\text{EC}_{50}$  values are indicated in the graph by the red dotted lines. Data are shown as mean  $\pm$  SD of triplicate measurements.

**Table 4.1 Sensitivity to DDD00085893 of WT, SKO, *Tb*GSK3 conditional over-expressing (cWT<sup>OE</sup>) and conditional SKO (cSKO<sup>OE</sup>) cell lines**

Cell line	EC <sub>50</sub> , μM		Ratio EC <sub>50</sub>	Ratio EC <sub>50</sub>	Hill slope		Ratio Hill slope
	minus Tet	plus Tet	plus Tet / minus Tet	plus Tet / WT	minus Tet	plus Tet	plus Tet / minus Tet
WT	3.9 ± 0.1	3.8 ± 0.1	1.0	1.0	2.5 ± 0.3	1.9 ± 0.6	0.8
SKO ( <i>PAC</i> )	2.1 ± 0.1	2.3 ± 0.1	1.1	**0.5	2.1 ± 0.1	2.0 ± 0.1	1.0
SKO ( <i>HYG</i> )	2.2 ± 0.2	2.3 ± 0.2	1.0	*0.6	1.5 ± 0.1	1.8 ± 0.5	1.2
cWT <sup>OE</sup>	3.5 ± 0.1	9.3 ± 0.8	*2.7	*2.4	2.6 ± 0.1	3.9 ± 0.1	*1.5
cSKO <sup>OE</sup>	1.9 ± 0.3	11.4 ± 0.9	**6.0	*2.9	2.2 ± 0.2	4.0 ± 0.2	*1.8

EC<sub>50</sub> values are weighted means and standard errors (SEM) of at least two independent determinations done in triplicate.

pLew100v5 was used as rescue construct for the conditional over-expression of the ectopic copy under tetracycline control.

The two-tailed P values were calculated using an unpaired *t*-test performed using the weighted mean, the SEM and number of independent determinations. P values below 0.05 (95%) were considered statistically significant (\*), P values below 0.01 (99%) were considered highly statistically significant (\*\*).

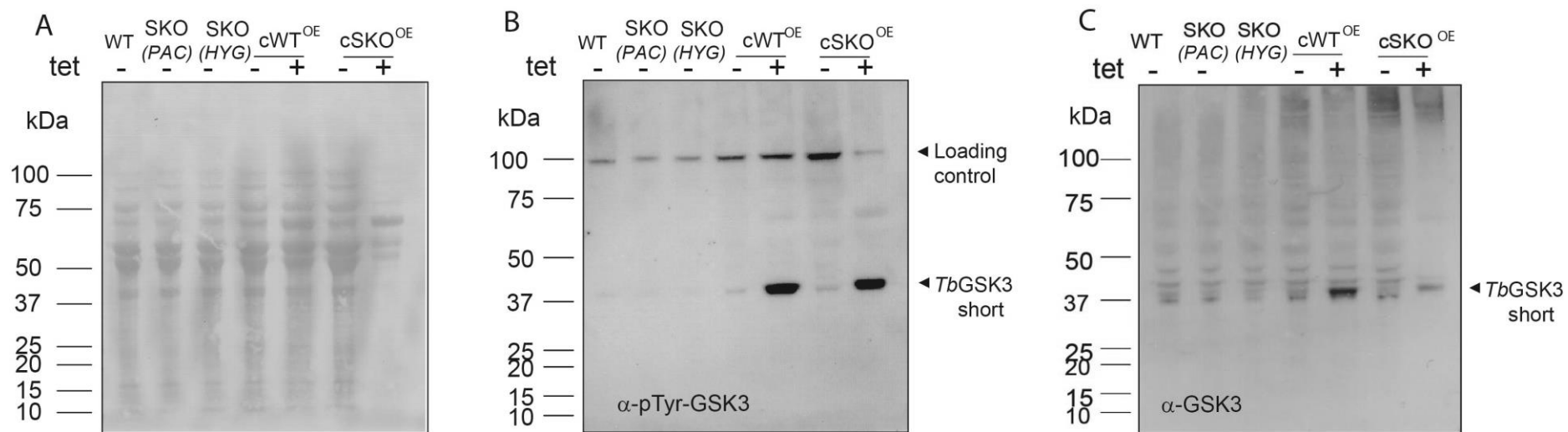


*TbGSK3* long, and the anti-GSK3 antibody that recognised only the recombinant *TbGSK3* short (see section 3.4.3, Figure 3.22 A and B). Bands at the expected molecular mass of ~40 kDa for *TbGSK3* short and ~55 kDa for *TbGSK3* long were visible on both blots, in addition to an intense non-specific band at 100 kDa for the immunoblot probed with the anti-phospho-tyrosine GSK3 antibody (**Figure 4.5** A and B). Two distinct bands were visible at the predicted molecular mass for *TbGSK3* short in the blot probed with the anti-GSK3 antibody, presumably the phosphorylated and unphosphorylated form of the enzyme, and a single band for the blot probed with the anti-phospho-tyrosine GSK3. These results seem to indicate that these antibodies are able to cross-react with not only the recombinant *TbGSK3* short but also the protein in the cell lysate. Considering that the anti-GSK3 antibody could not recognise the recombinant *TbGSK3* long, it is unclear if the band observed at ~50 kDa in both blots is *TbGSK3* long or an aspecific cross-reacting protein.

#### **4.1.6 Chemical validation of *TbGSK3* short as drug target**

To assess whether *TbGSK3* short was specifically targeted by DDD00085893 in trypanosomes, *TbGSK3* short SKO cell lines and over-expressing cell lines (cWT<sup>OE</sup> and cSKO<sup>OE</sup>), obtained using pLew100v5 as conditional over-expressor, were compared with the WT for their relative sensitivity to this inhibitor.

The increase in sensitivity of the SKO cell lines to DDD00085893 correlated well with an expected reduction of the 50% in protein expression due to deletion of one allele. Indeed, after 72 h incubation DDD00085893 EC<sub>50</sub> values were  $2.2 \pm 0.2$  and  $2.1 \pm 0.1$   $\mu\text{M}$  for SKO cell lines (SKO(*HYG*) and SKO(*PAC*), respectively) compared to an EC<sub>50</sub> value of  $3.9 \pm 0.1$   $\mu\text{M}$  for the WT (**Figure 4.6** A, **Table 4.1**). Western analysis demonstrated that *TbGSK3* short protein levels were higher in tetracycline-induced

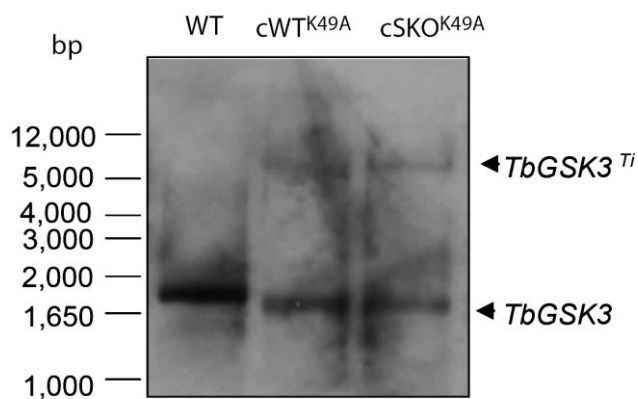


**Figure 4.7 Western blot analysis of mutant cell lines for *TbGSK3* short**

Cells were grown for 72 hours in presence (+) or absence (-) of tetracycline (tet) prior to analysis. **(A)** Ponceau staining of the blot is used as loading control. In each lane were loaded ~10 μg of protein at the exception of tetracycline-induced cSKO for which not enough cells survived to the induction. **(B)** The blot was probed with anti-pTyr-GSK3. A 100 kDa non-specific band was also detected, which was used as an internal loading control. **(C)** The same blot was stripped and blot with anti-GSK3.

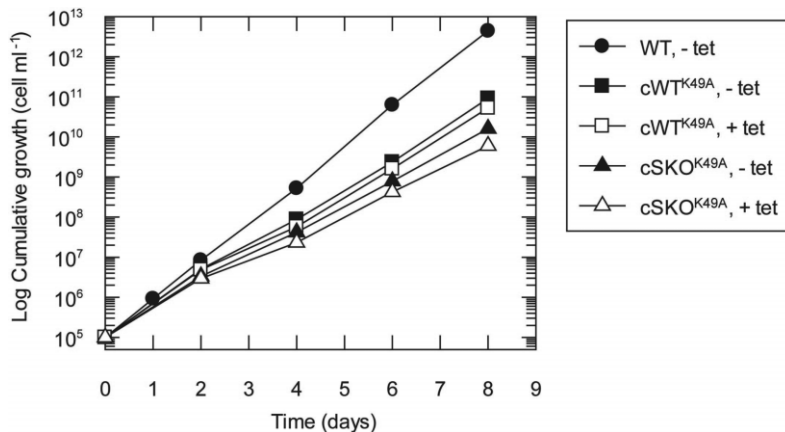
cWT<sup>OE</sup> and cSKO<sup>OE</sup> cell lines than in WT cells (**Figure 4.7 A, B and C**). The non-induced cWT<sup>OE</sup> and cSKO<sup>OE</sup> cells had EC<sub>50</sub> values of  $3.5 \pm 0.1$  and  $1.9 \pm 0.3$   $\mu\text{M}$ , respectively, comparable to the values of their respective parental cell lines (**Figure 4.6 B, Table 4.1**). Conditional over-expression of *TbGSK3 short* affected cell growth causing cell toxicity that could be rescued by increasing concentrations of DDD00085893 up to 3  $\mu\text{M}$ , but higher concentrations of compound resulted in inhibition of growth in itself. The induced cell lines reached maximum growth at 3  $\mu\text{M}$ , and this growth level was used to normalize the inhibitor effect of DDD00085893; the derived inhibition profile had an unusual U shape depicting rescue of growth inhibition below 3  $\mu\text{M}$  and the inhibitory effect above this concentration. The inhibition shifted with respect to the parental cell lines, and in particular tetracycline-induced cWT<sup>OE</sup> and cSKO<sup>OE</sup> had EC<sub>50</sub> value of  $9.3 \pm 0.8$  and  $11.4 \pm 0.9$   $\mu\text{M}$ , respectively, displaying a decrease in sensitivity. These tetracycline-induced cell lines also gave an increase in their Hill-slope values, which increased 1.5 and 1.8 fold, respectively, compared to the parental cell lines (**Figure 4.6 B and Table 4.1**). A steeper Hill-slope can be attributed to off-target effects or to a reduction of the percentage of enzyme inhibition necessary to achieve a cellular effect (Czock and Keller, 2007; Sampah *et al.*, 2011).

In conclusion, an increase in *TbGSK3 short* protein expression and/or activity is detrimental to cell proliferation in bloodstream form *T. b. brucei* and the associated toxicity can be chemically modulated by DDD00085893. Taken together these results provide evidence that DDD00085893 targets *TbGSK3 short* in cultured trypanosomes and chemically validates *TbGSK3 short* as drug target in BSF *T. brucei*.



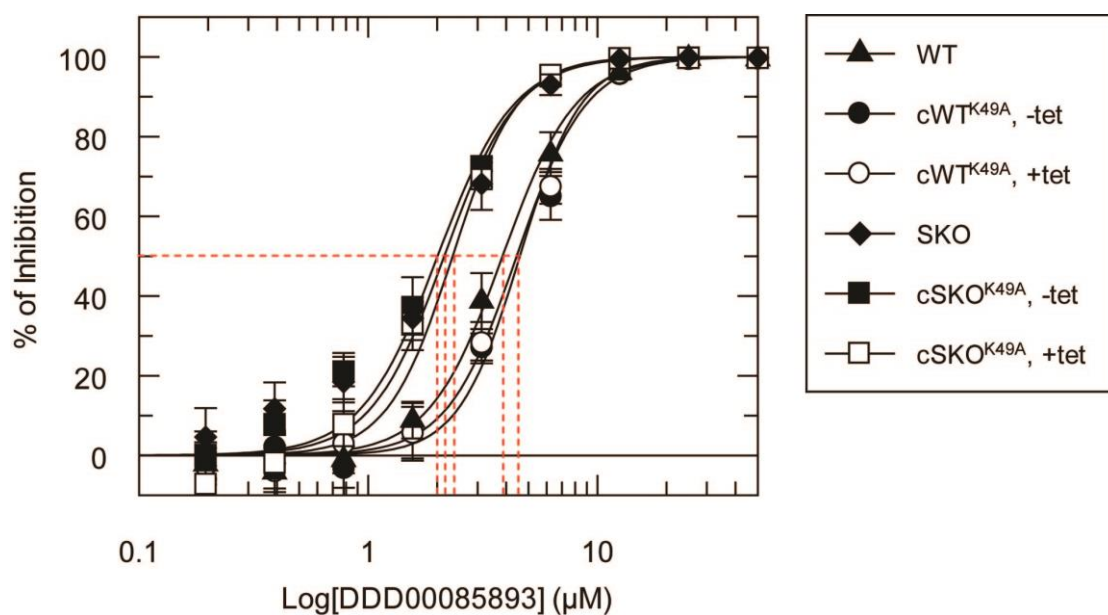
**Figure 4.8 Genotypic analysis of conditional over-expressing *TbGSK3*(K49A) kinase dead cell lines**

Southern blot analysis of conditional over-expressing kinase dead *TbGSK3* short cell lines obtained using pLew100v5 plasmid. Genomic DNA from the *T. b. brucei* 427 wild-type (WT), from *TbGSK3*(K49A) over-expressing cell line (cWT<sup>K49A</sup>) and from *TbGSK3K49A<sup>Ti</sup> ΔTbgsk3::PAC/TbGSK3* cell line (cSKO<sup>K49A</sup>) was digested with *Bam*HI and *Pst*I and probed with the *TbGSK3* ORF probe. The blot shows the incorporation of the ectopic copy.



**Figure 4.9 Cumulative growth of *T. brucei* WT and conditional over-expressing *TbGSK3*(K49A) kinase dead cell lines**

Wild-type (WT) (closed circles); non-induced and tetracycline-induced *TbGSK3*(K49A) kinase conditional over-expressing cell lines (cWT<sup>K49A</sup>) (closed and open squares, respectively); non-induced and tetracycline-induced *TbGSK3*(K49A) kinase dead conditional single knock-out (cSKO<sup>K49A</sup>) (close and open triangles, respectively).



**Figure 4.10 DDD00085893 sensitivity towards conditional over-expressing *Tb*GSK3(K49A) kinase dead cell lines**

Non-induced and tetracycline-induced cWT<sup>K49A</sup> EC<sub>50</sub> values of  $4.4 \pm 0.2$  and  $4.6 \pm 0.2$   $\mu\text{M}$  were determined (closed and open circles, respectively). Non-induced and tetracycline-induced cSKO<sup>K49A</sup> EC<sub>50</sub> values of  $2.1 \pm 0.1$  and  $2.3 \pm 0.2$   $\mu\text{M}$  were determined (open and closed squares, respectively). WT (closed triangle) and SKO (close diamond) inhibition profiles were added as comparison to cWT<sup>K49A</sup> and cSKO<sup>K49A</sup> cell lines. EC<sub>50</sub> values are indicated in the graph by the red dotted lines. Data are shown as mean  $\pm$  SD of triplicate measurements.

**Table 4.2 Sensitivity of DDD00085893 towards cell lines conditional over-expressing *Tb*GSK3(K49A) kinase dead mutant**

DDD00085893 Cell line	EC <sub>50</sub> ± SD, µM		Ratio EC <sub>50</sub>	Ratio EC <sub>50</sub>	Hill slope ± SD		Ratio Hill slope
	minus Tet	plus Tet	plus Tet / minus Tet	plus Tet / WT <sup>a</sup>	minus Tet	plus Tet	plus Tet / minus Tet
WT <sup>a</sup>	3.9 ± 0.1	3.8 ± 0.1	1.0	1.0	2.5 ± 0.3	1.9 ± 0.6	0.8
cWT <sup>K49A</sup>	4.4 ± 0.2	4.6 ± 0.2	1.1	1.2**	2.7 ± 0.2	3.1 ± 0.2	1.1
SKO <sup>a</sup>	2.1 ± 0.1	2.3 ± 0.1	1.1	0.5**	2.1 ± 0.1	2.0 ± 0.1	1.0
cSKO <sup>K49A</sup>	2.1 ± 0.1	2.3 ± 0.2	1.1	0.6*	2.9 ± 0.2	3.3 ± 0.2	1.1

<sup>a</sup> EC<sub>50</sub> value for WT and SKO as reported in Table 4.1.

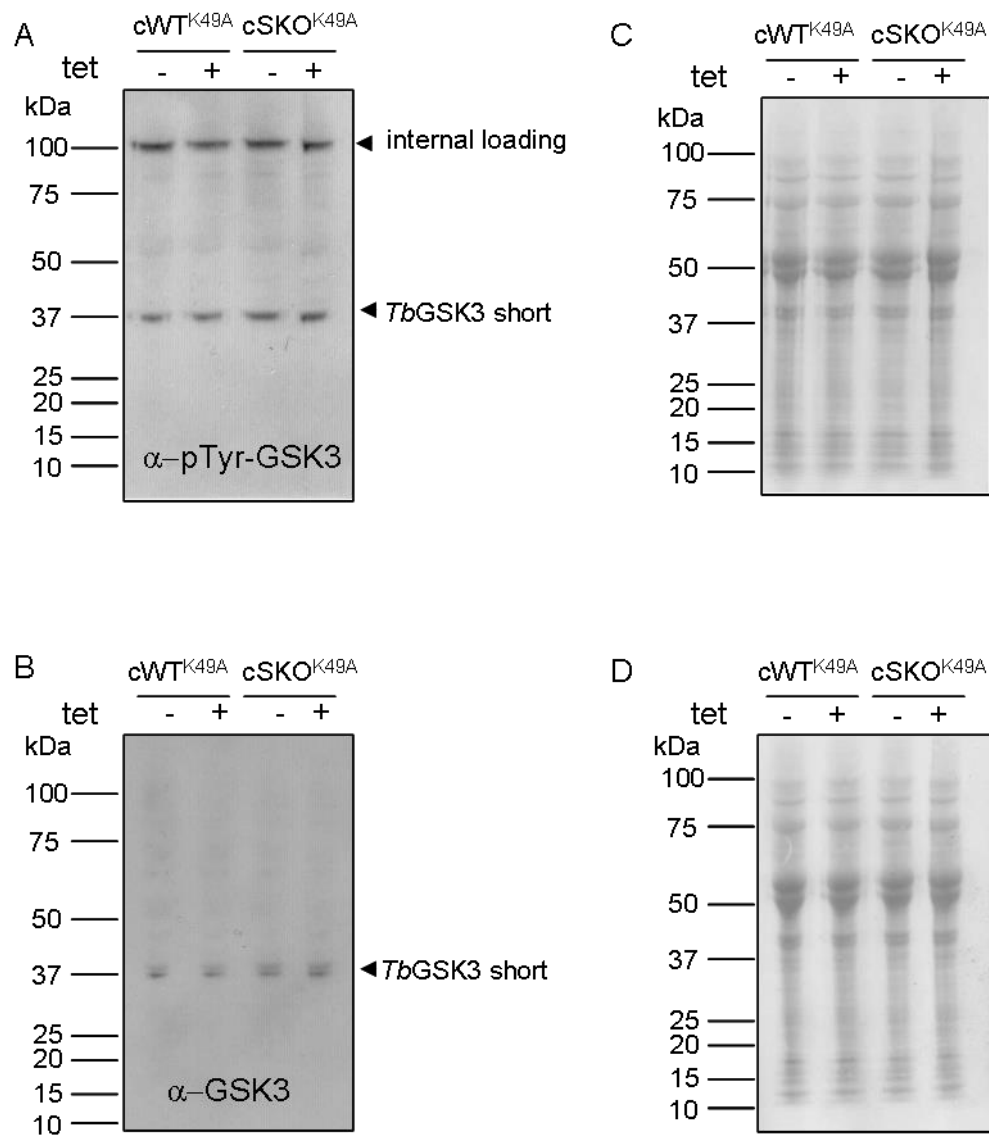
EC<sub>50</sub> values are means and standard errors of triplicate measurements.

The two-tailed P values were calculated using an unpaired *t* test calculated using the weighted mean, the SD and number of replicates. P values <0.05 (95%) were considered statistically significant (\*), P values <0.01 (99%) were considered highly statistically significant (\*\*).

#### 4.1.7 Over-expression of inactive *TbGSK3* short in bloodstream form

Having shown that DDD00085893 is an “on target” inhibitor of *TbGSK3* short, I wanted to address the question as to whether this compound targeted both the active and inactive forms of the enzyme, as well as whether the toxic effects associated with *TbGSK3* over-expression were due to an increase in enzyme activity or an increase in cellular protein concentration. In *HsGSK3 $\beta$*  there is a conserved lysine (Lys) at position 85 that plays an important role in the catalytic activity of the enzyme (Cole *et al.*, 2004). Mutation of this residue to alanine (Ala) results in a catalytically inactive form of the enzyme. This residue is conserved in all orthologues of GSK3 including the kinetoplastids (**Figure 1.8**).

Using this information I generated a kinase dead mutant of *TbGSK3* short by site-direct mutagenesis in which Lys49 was mutated to Ala, namely hereafter *TbGSK3*(K49A) short. Conditional over-expressing cell lines were obtained by transfection of an ectopic copy of this mutant using pLew100v5 vector into both wild-type and single knock-out background cell lines. The incorporation of the ectopic copy in selected clones was confirmed by Southern analysis (**Figure 4.8**). The growth of cWT<sup>K49A</sup> and cSKO<sup>K49A</sup> cells was compared to the growth of their parental cell lines and were found to be slightly slower (**Figure 4.9**). The sensitivity of DDD00085893 to the tetracycline induced and un-induced cWT<sup>K49A</sup> and cSKO<sup>K49A</sup> cells was tested after 72 h exposure and found unchanged with respect to the sensitivity to WT and SKO cell lines. The tetracycline induction had no effect on the sensitivity of DDD00085893 against cWT<sup>K49A</sup> clonal cell lines with EC<sub>50</sub> values of  $4.4 \pm 0.2$  and  $4.6 \pm 0.2$   $\mu$ M for non-induced and tetracycline-induced, respectively. Similarly also cSKO<sup>K49A</sup> cell lines had EC<sub>50</sub> values of  $2.1 \pm 0.1$  and  $2.3 \pm 0.2$   $\mu$ M for non-induced and tetracycline-induced, respectively (**Figure 4.10** and **Table 4.2**).



**Figure 4.11 Western Blot analysis of *T. brucei* conditional over-expressing *TbGSK3(K49A)* kinase dead mutant cell lines**

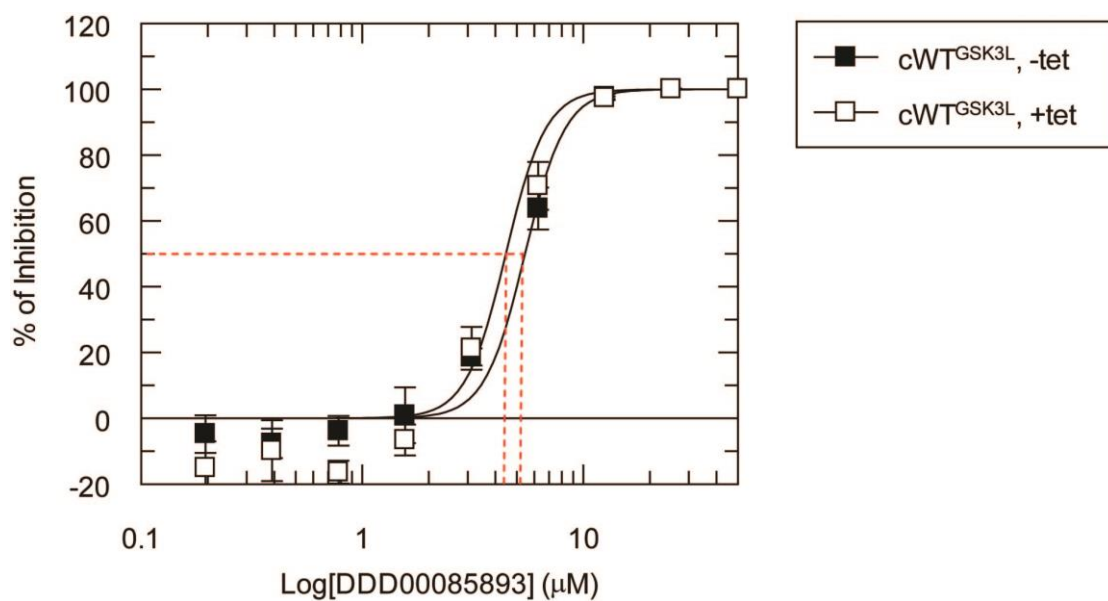
Cells were induced with tetracycline (tet) for 72 h prior to analysis. Blots were probed with antibody directed against phospho-tyrosine-GSK3 ( $\alpha$ -pTyr-GSK3) (A) and anti-GSK3 ( $\alpha$ -GSK3) (B), the corresponding Ponceau staining of the blots is also shown as loading control (C and D). Ten  $\mu$ g of total protein was loaded for each lane.



Previously it had been reported that the phosphorylation of *HsGSK3 $\beta$*  in the activation loop is an autophosphorylation event (Cole *et al.*, 2004), thus it is assumed that the kinase-dead enzyme could not be phosphorylated in the activation loop. Western blot analysis of tetracycline-induced cWT<sup>K49A</sup> and cSKO<sup>K49A</sup> did indeed show no increase in the amount of the phosphorylated form of the enzyme, but also total protein levels did not appear to be increased either (**Figure 4.11**). The fact that total protein was not elevated after tetracycline induction suggests that the kinase-dead either is not over-expressed or this form of the enzyme is unstable and quickly targeted for degradation considering that the wild-type protein was over-expressed despite being toxic (**Figure 4.7**). This hypothesis would be in agreement with Cole and colleagues' (2004) observation that the phosphorylation in the activation loop is not required for protein activity but rather for stability. Further evidence that the kinase dead might be unstable in *T. b. brucei* cells comes from the inability to obtain soluble recombinant protein in *E. coli* not interacting with folding proteins (see section 3.4.4, **Figure 3.24** and **Figure 3.25**). Thus, the failure to demonstrate over-expression of the kinase dead mutant leaves unanswered the question whether the cellular toxicity is caused by the increase in enzyme activity or protein level of *TbGSK3* short.

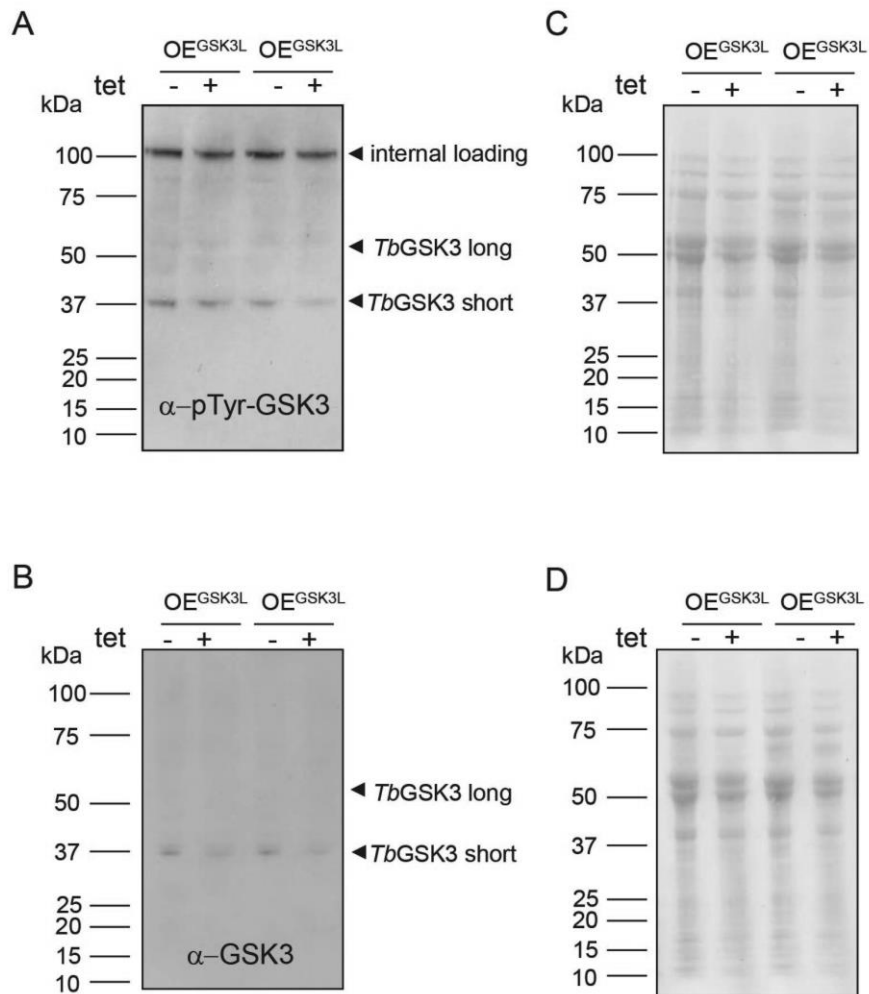
#### **4.1.8 Over-expression of *TbGSK3* long in bloodstream form *T. brucei***

The chemical proteomic approach clearly identified *TbGSK3* short as target of DDD00085893 in *T. brucei* lysate. All the peptides attributed to *TbGSK3* were unique to this isoform at the exclusion of one peptide (LCDFGSAK) that was shared between the short and the long form (Tb427.07.2420) (section 3.3.1, Figure 3.13). The low number of observations for this peptide and the fact that no unique peptides for the long isoform were observed makes it unlikely that the long isoform of *TbGSK3* was among the binders of DDD00085893 in the chemical-proteomic experiment.



**Figure 4.12 DDD00085893 sensitivity towards conditional over-expressing *TbGSK3* long cell line**

EC<sub>50</sub> values of  $4.5 \pm 0.7$  and  $5.4 \pm 1.0$  μM were determined for a non-induced and tetracycline-induced clonal cell line over-expressing the long form of *TbGSK3*. EC<sub>50</sub> values are indicated by the red dotted lines. Data are shown as mean ± SD of triplicate measurements.

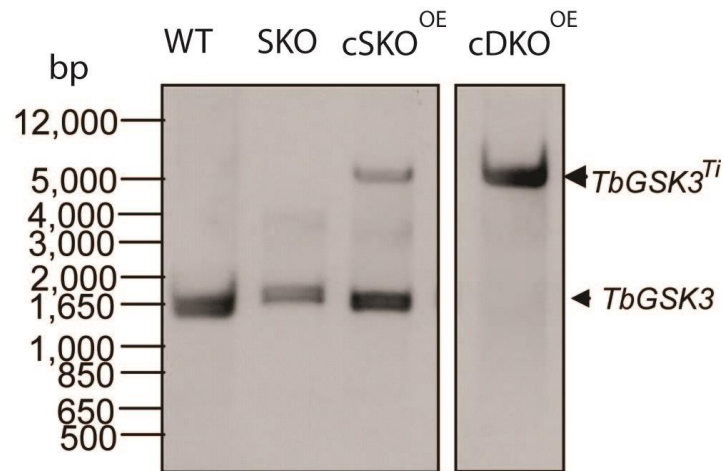


**Figure 4.13 Western Blot analysis of *T. brucei* conditional over-expressing *TbGSK3* long cell lines**

Cells were induced with tetracycline (tet) for 72 h prior of the study. Blots were probed with antibody directed against phospho-tyrosine-GSK3 ( $\alpha$ -pTyr-GSK3) (A) and anti-GSK3 ( $\alpha$ -GSK3) (B), the corresponding Ponceau staining of the blots is also shown as loading control (C and D). Arrows indicate the expected size for *TbGSK3* short and *TbGSK3* long. An unspecific band at 100 kDa is reported as internal loading control. Ten  $\mu$ g of total protein was loaded per each lane.

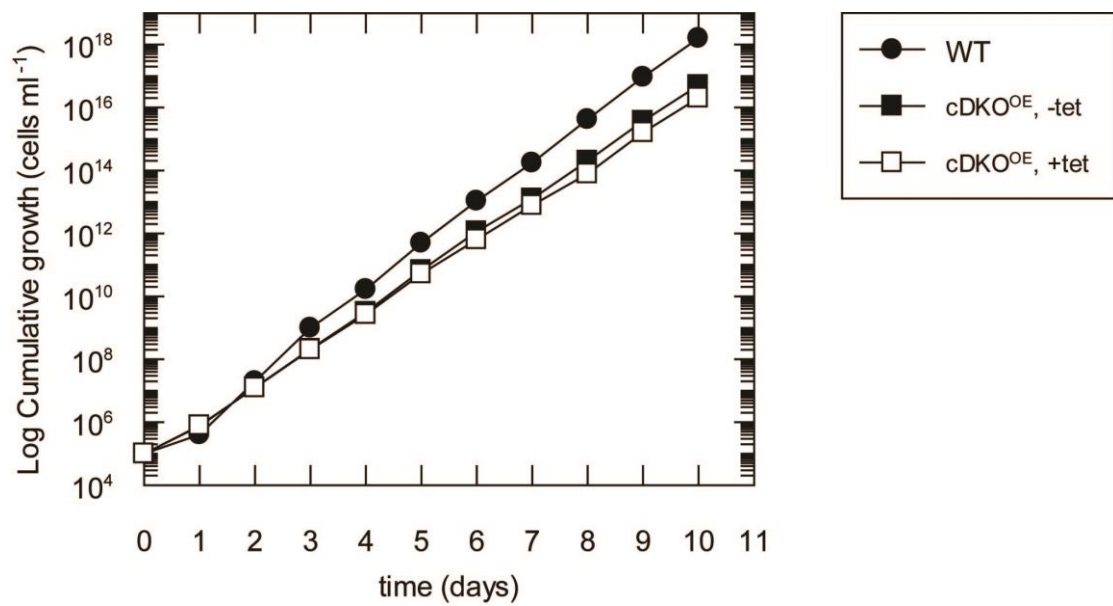
In previous studies, the knock-down of *TbGSK3* long did not affect the growth of BSF *T. brucei* (Ojo *et al.*, 2008, Alsford *et al.*, 2011) and was associated to a gain-of-fitness in procyclic form (Alsford *et al.*, 2011). The phosphoproteomic studies conducted by Nett and colleagues (2009a; 2009b) identified no phosphorylated peptides belonging to the long isoform in the BSF, but in procyclic a unique phosphorylated peptide was identified corresponding to its activation loop (NVPpYIFSR). These data suggest that *TbGSK3* long is expressed in the procyclic stage, probably in an active form, but there is no evidence to date of the expression of this protein in BSF in an active form. Thus, I sought to investigate whether this protein could be over-expressed in BSF *T. brucei* and could be inhibited by DDD00085893.

I cloned the *TbGSK3* long ORF (see section 3.4.2) into the pLew100v5 plasmid for ectopic over-expression in *T. brucei*. The growth of two distinct clonal cell lines resistant to phleomycin (*BLE*) were examined and found comparable to wild-type cells (data not shown). These cell lines had also unaltered sensitivity to DDD00085893 even when induced by tetracycline (**Figure 4.12**). Although a p-Tyr-GSK3 antibody can recognise the recombinant *TbGSK3* long (see section 3.4.4), the protein was not detected in BSF lysates even when the expression was induced with tetracycline (**Figure 4.13**). A lack of expression or reduced stability in BSF would explain why *TbGSK3* long does not play a role in the clinically relevant form of *T. brucei*, and therefore DDD00085893 presumably targets only *TbGSK3* short. Not only Lys85, but also the conserved residue Lys86 plays a role in the kinase activity of GSK3 across species. Interestingly in *TbGSK3* long this residue is mutated to Arg88, making *TbGSK3* long more similar to a kinase dead mutant than an active protein. This might explain the extremely low activity of the recombinant protein (see section 3.4.2) and the associated instability might account for the low level of expression or rapid protein turnover.



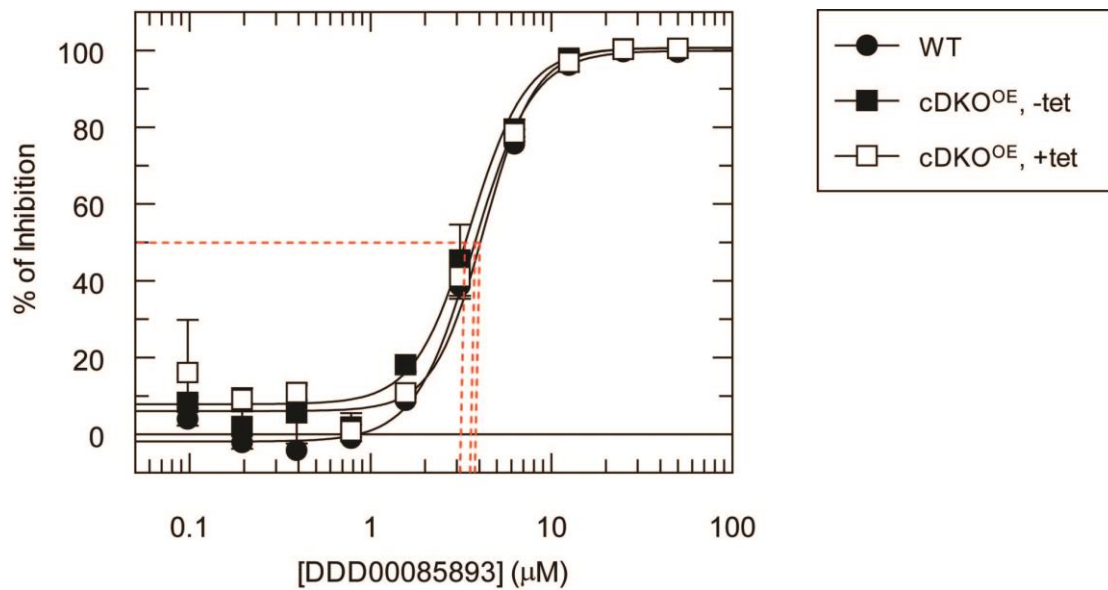
**Figure 4.14 Southern blot analysis of conditional double knock-out cell lines obtained using the conditional over-expressor pLew100v5 plasmid**

Genomic DNA from the wild-type (WT) *T. brucei* S427, from  $\Delta Tbgsk3::PAC/TbGSK3$  cell line (SKO), from  $TbGSK3^{Ti} \Delta Tbgsk3::PAC/TbGSK3$  cell line (cSKO<sup>OE</sup>) and  $TbGSK3^{Ti} \Delta Tbgsk3::PAC/\Delta Tbgsk3::HYG$  cell line (cDKO<sup>OE</sup>) was digested with *Bam*HI and *Pst*I and probed with the *TbGSK3* ORF probe. The arrows show the expected size for *TbGSK3* endogenous copy and for the ectopic  $TbGSK3^{Ti}$ . Five  $\mu$ g of digested DNA was loaded per lane.



**Figure 4.15 Cumulative growth of *T. brucei* WT and conditional null mutant cell line**

Wild-type (WT, closed circles) was compared to the growth of non-induced and tetracycline-induced conditional null mutant cell line (cDKO<sup>OE</sup>, closed and open squares, respectively).



**Figure 4.16 DDD00085893 sensitivity towards conditional null mutant (cDKO) cell line**

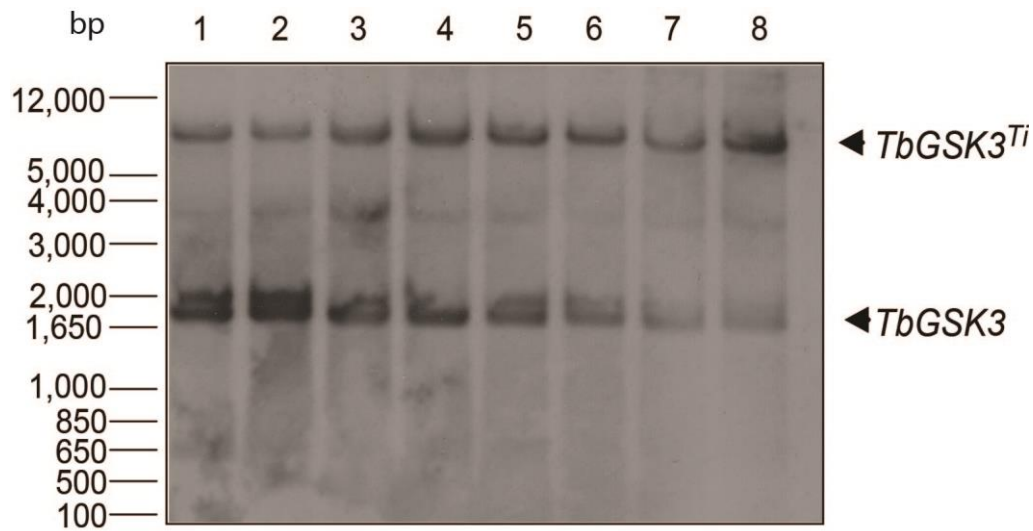
EC<sub>50</sub> values of  $3.6 \pm 0.3$  and  $4.2 \pm 0.3$  μM were determined for non-induced and tetracycline-induced *TbGSK3* cDKO<sup>OE</sup> cell lines (close and open squares respectively). WT (closed circle) inhibition profile was added as comparison to cDKO<sup>OE</sup> cell line. EC<sub>50</sub> values are indicated by the red dotted lines. Data are shown as mean  $\pm$  SD of triplicate measurements.

## 4.2 Genetic validation of *TbGSK3* short

### 4.2.1 Generation of conditional double knock-out using pLew100v5

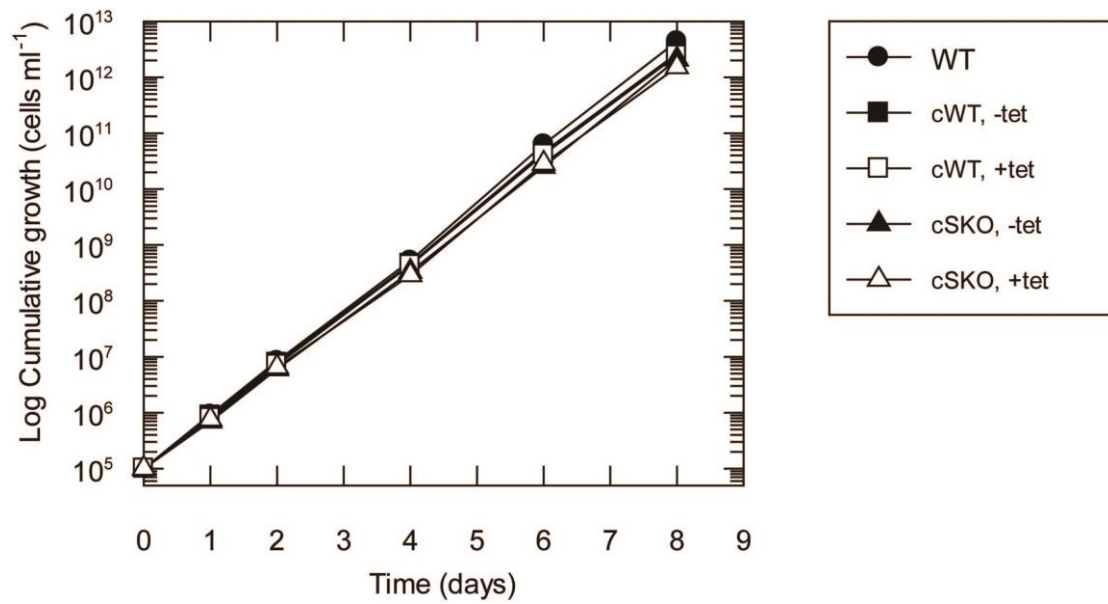
As reported in section 4.1.2 the failure to knock-out both alleles of *TbGSK3* gene suggests a lethal phenotype. Thus, generation of a conditional null mutant was attempted by insertion of an ectopic copy. When pLew100v5 was used as rescue construct, after several attempts it was possible to select a clonal cell line that was resistant to both *PAC* and *HYG*, suggesting that both alleles of the WT gene were replaced. The removal of both copies was confirmed by Southern analysis (**Figure 4.14**). The growth of this conditional null mutant was compared to wild type cells and found to be equivalent both in presence and absence of tetracycline (**Figure 4.15**). The sensitivity to DDD00085893 of the cDKO<sup>OE</sup> cell line was unaltered with EC<sub>50</sub> values of 3.6  $\mu$ M and 4.2  $\mu$ M when un-induced and tetracycline-induced, these values were not significantly different from the EC<sub>50</sub> value of 3.9  $\mu$ M for the WT cell line (**Figure 4.16**). The unaltered sensitivity to DDD00085893 indicated that the level of protein expression was unchanged, even when tetracycline was withdrawn from the culture medium. It is known that conditional null mutants can escape from the tetracycline regulatory machinery in particular when the protein is essential for the parasite survival. Previously it has been reported that reverting cell lines had undergone genetic rearrangement and deleted the *TETR* gene thereby losing the tetracycline control (Roper *et al.*, 2002). Other mechanisms could also be responsible for the lack of tetracycline control and the constitutively expression of *TbGSK3* short, as such as mutations in the *TetR* gene or expression of a non-functional TetR protein (Sienkiewicz *et al.*, 2008). The cause of reversion was not deemed to be part of the remit of the study and was therefore not further investigated.





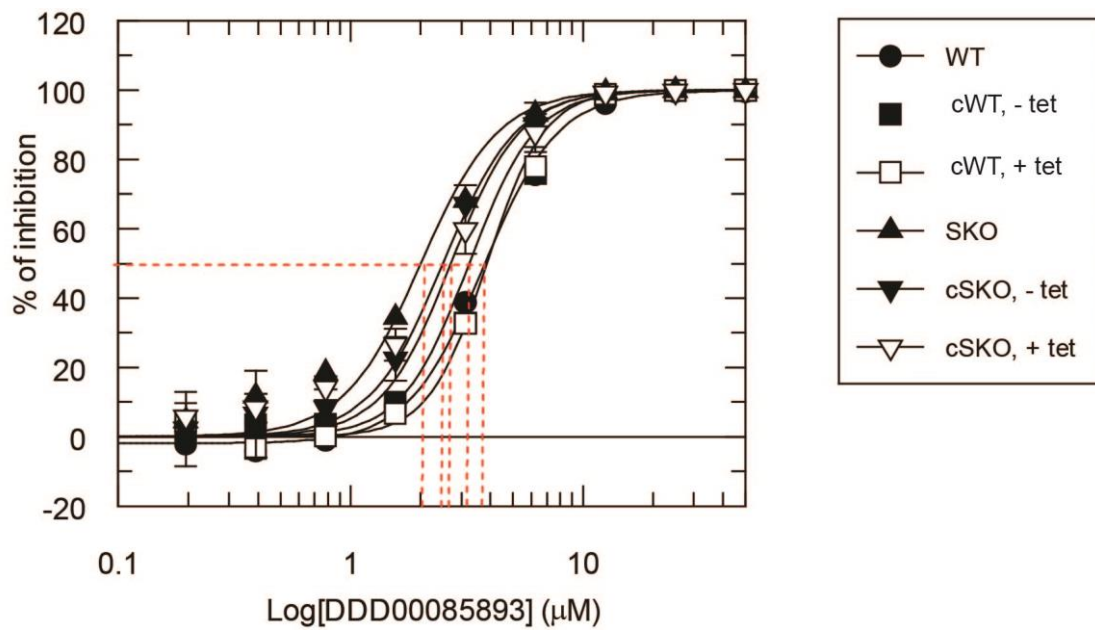
**Figure 4.17 Southern blot analysis of pLew100 conditional WT (cWT) and conditional single knock-out (cSKO) cell lines**

Five  $\mu\text{g}$  of genomic DNA from four different clones of *TbGSK3* conditional WT (lanes 1 to 4) and conditional single knock-out cell lines (lanes 5 to 8) were digested with *Bam*HI and *Pst*I and probed with the ORF of *TbGSK3 short*. Expected size for the endogenous *TbGSK3* and ectopic *TbGSK3<sup>Ti</sup>* are reported on the right hand-side.



**Figure 4.18 Cumulative growth of *T. brucei* WT, conditional WT and conditional single knock-out cell lines obtained using pLew100 as rescue vector**

Wild-type (WT) growth in the absence of tetracycline (tet) was compared to the growth of non-induced and tetracycline-induced pLew100 conditional WT (cWT) and conditional single knock-out (cSKO) cell lines.



**Figure 4.19 DDD00085893 sensitivity towards pLew100 conditional WT and conditional single knock-out cell lines**

EC<sub>50</sub> values of  $2.5 \pm 0.2$  and  $2.7 \pm 0.1$   $\mu\text{M}$  were determined for non-induced and tetracycline-induced *TbGSK3* cSKO cell lines (close and open inverted triangles respectively). EC<sub>50</sub> values of  $3.2 \pm 0.1$  and  $3.8 \pm 0.1$   $\mu\text{M}$  were determined for non-induced and tetracycline-induced *TbGSK3* cWT cell lines (closed and open squares respectively). WT (closed circle) and SKO (close triangle) inhibition profiles were added as comparison. EC<sub>50</sub> values are indicated by the red dotted lines. Data are shown as mean  $\pm$  SD of triplicate measurements.

**Table 4.3 Sensitivity of DDD00085893 towards conditional WT (cWT) and conditional SKO (cSKO) cell lines**

DDD00085893 Cell line	EC <sub>50</sub> , μM		Ratio EC <sub>50</sub>		Hill slope		Ratio Hill slope
	minus Tet	plus Tet	plus Tet / minus Tet	plus Tet / WT	minus Tet	plus Tet	plus Tet / minus Tet
WT <sup>a</sup>	3.9 ± 0.1	3.8 ± 0.1	1.0	1.0	2.5 ± 0.3	1.9 ± 0.6	0.8
SKO <sup>a</sup>	2.1 ± 0.1	2.3 ± 0.1	1.1	0.5**	2.1 ± 0.1	2.0 ± 0.1	1.0
<b>pLew100v5</b>							
cWT <sup>OEa</sup>	3.5 ± 0.1	9.3 ± 0.8	2.7*	2.4*	2.6 ± 0.1	3.9 ± 0.1	1.5*
cSKO <sup>OE a</sup>	1.9 ± 0.3	11.4 ± 0.9	6.0**	2.9*	2.2 ± 0.2	4.0 ± 0.2	1.8*
<b>pLew100</b>							
cWT <sup>b</sup>	3.2 ± 0.1	3.8 ± 0.1	1.2**	1.0	2.7 ± 0.1	2.9 ± 0.2	1.1
cSKO <sup>b</sup>	2.5 ± 0.2	2.7 ± 0.1	1.1	0.7***	3.0 ± 0.1	3.5 ± 0.1	1.2**

<sup>a</sup>EC<sub>50</sub> values are weighted means and standard errors (SEM) of at least two independent experiments done in triplicate.

<sup>b</sup>EC<sub>50</sub> values are means and standard deviations (SD) of triplicate measurements.

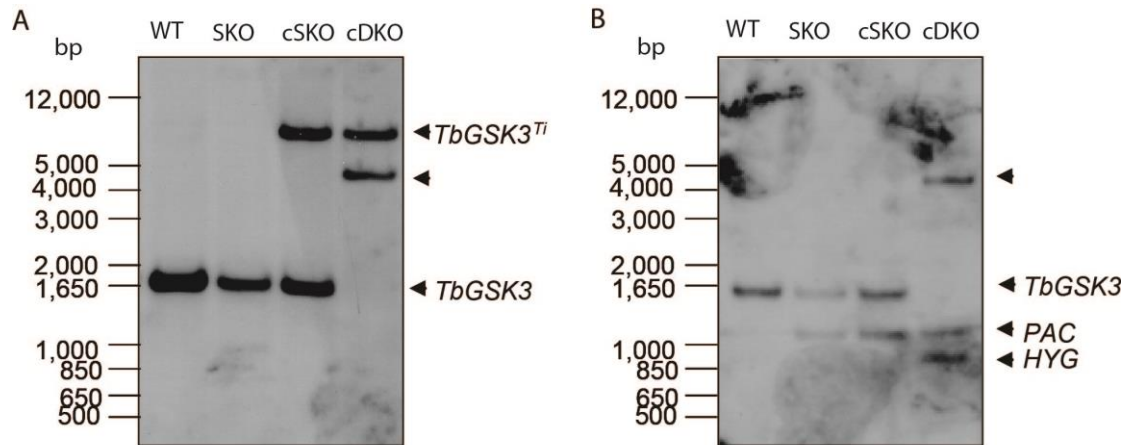
The two-tailed P values were calculated using an unpaired *t* test calculated using the weighted mean, the SD or SEM and number of replicates. P values <0.05 (95%) were considered statistically significant (\*), P values <0.01 (99%) were considered highly statistically significant (\*\*), P values ≤0.001 (99.9%) were considered extremely statistically significant (\*\*\*)

As previously reported (section 4.1.4) over-expression above physiological levels of *TbGSK3* was toxic for the cells, but they seemed to be able to adjust to the increased protein expression after one week of continuously culturing in presence of tetracycline (**Figure 4.4**). The loss of tetracycline control may have occurred at the level of the cSKO<sup>OE</sup> cell lines when continuously induced with tetracycline for one week prior to the final transfection to obtain the cDKO<sup>OE</sup>. Incidentally, when cSKO<sup>OE</sup> cells were cultured in the presence of tetracycline for a shorter period of time (3 days) before the final transfection event, no cells were recovered after drug selection.

#### 4.2.2 Generation of conditional double knock-out using pLew100

An alternative strategy was used to generate a potential cDKO by the use of pLew100, an expression vector that could achieve lower protein expression more comparable to the physiological levels. Using pLew100 as rescue construct it was possible to insert an ectopic copy of *TbGSK3 short* into WT and  $\Delta TbgsK3::PAC/TbGSK3$  (SKO) cells obtaining, respectively,  $TbGSK3^{Ti}$  (cWT) and  $TbGSK3^{Ti} \Delta TbGSK3::PAC/TbGSK3$  (cSKO) cell lines, as confirmed by Southern analysis of clonal lines (**Figure 4.17**). The growth of the  $TbGSK3^{Ti}$  (cWT) and  $TbGSK3^{Ti} \Delta TbgsK3::PAC/TbGSK3$  (cSKO) clonal cell lines was comparable to the background cell lines WT and SKO respectively (**Figure 4.18**), without any toxic side effects due to tetracycline induction. Moreover, there was no shift in sensitivity towards DDD00085893 (**Figure 4.19** and **Table 4.3**), suggesting that the level of protein was not elevated above physiological levels.

The transfection of *HYG* flanked with the UTR's regions of *TbGSK3 short* into a cSKO background ( $GSK3short^{Ti} \Delta gsk3short::PAC/GSK3short$ ) generated cell lines that were resistant to both selectable markers and were selected in clonal cell lines by limiting dilution. Unfortunately, southern analysis showed that these clones had not



**Figure 4.20** Southern blot analysis of cDKO cell lines obtained using pLew100 as rescue vector

Genomic DNA from the wild-type (WT) *T. brucei* 427,  $\Delta Tbgsk3::PAC/TbGSK3$  cell line (SKO), from  $TbGSK3^{Ti} \Delta Tbgsk3::PAC/TbGSK3$  cell line (cSKO) and  $TbGSK3^{Ti} \Delta Tbgsk3::PAC/\Delta Tbgsk3::HYG$  cell line (cDKO) was digested with *Bam*HI and *Pst*I and probed with the *TbGSK3* short ORF probe (A) and *TbGSK3* short 5'UTR probe (B). **Panel A** depicts the incorporation of the ectopic copy in the cSKO and cDKO cell lines, the removal of the endogenous copies and the insertion of an additional copy of *TbGSK3* gene in an alternative locus (band indicated by an arrow at ~4,500 bp). **Panel B** shows the incorporation of puromycin-resistance gene (*PAC*) in the SKO and cSKO cell lines and the further incorporation of hygromycin-resistance gene (*HYG*) at the expected bp size in the cDKO. Also this blot shows the extra band at ~ 4,500 bp. Same amounts of digested DNA (5  $\mu$ g) were loaded into two different gels and transferred into two distinct blots.

only replaced both allelic copies of *TbGSK3* short, but had also undergone gene rearrangement and that the endogenous copy of *TbGSK3* had shifted to an alternative locus (**Figure 4.20**). The observation that the second allelic copy of the gene had translocated to another locus probably indicates that the level of expression of the ectopic copy under pLew100 control did not reach levels that could maintain the parasites' viability once the second allelic copy of the gene was removed.

These data suggest that the gene knock-outs can be particularly difficult to obtain for proteins that have tightly regulated levels of expression or are toxic at concentrations above physiological levels. Under these circumstances alternative methods of gene validation are required. Taken together, these experiments indicate that *TbGSK3* short is an essential druggable gene in BSF *T. brucei* due to the following: (i) knock-outs could not be obtained even using different rescue constructs and (ii) its activity can be modulated by small molecule inhibitors directly in cells.

## **Chapter 5 Discussion**



## 5.1 Key findings

At the beginning of this project *TbGSK3* short had been just proposed as a drug target for Human African trypanosomiasis. Its genetic validation was based upon a mild growth defect by RNAi and chemically validated by using compounds that were not selective or specific for *TbGSK3* short (Ojo *et al.*, 2008). In the previous studies the correlation existing between biochemical and cellular activity was quite loose and the cellular activity could not be completely attributed to *TbGSK3* short, considering the promiscuity of the compounds. Finally, the highest selectivity that could be achieved towards *HsGSK3* was only seven fold (Oduor *et al.*, 2011).

Despite the very limited genetic and chemical validation available for *TbGSK3* short as potential drug target, what made it attractive from a drug discovery point of view was 1) the availability of recombinant active enzyme and biochemical assays for medium throughput screening (Ojo *et al.*, 2008), 2) the pre-existing chemical knowledge on protein kinase inhibitors together with the availability of an in-house focused kinase library of chemical starting points (Wyatt *et al.*, 2008; Brenk *et al.*, 2007), and 3) the availability of crystal structures for *HsGSK3* and *HsCDK2*, the most closely related kinases to *TbGSK3* short in humans.

In addition to the main question whether *TbGSK3* short was a good drug target or not, there were other unanswered questions about 1) the level of enzyme inhibition required to obtain cell efficacy and in particular if *TbGSK3* short inhibition would be cytotoxic or cytostatic; 2) whether it would be possible to achieve higher selectivity between the human and parasite kinases; and 3) what the function(s) of *TbGSK3* short might be in *T. brucei* or other kinetoplastids. The presence of another isoform in the *T. brucei* genome, *TbGSK3* long, also posed the question whether this enzyme which has

been characterised as not essential by RNAi (Ojo *et al.*, 2008; Alsford *et al.*, 2011), would have been able to compensate for the inhibition of *TbGSK3* short.

This project has contributed greatly to the chemical and genetic validation of *TbGSK3* short as drug target determining that:

- 1) *TbGSK3* short is a druggable target;
- 2) *TbGSK3* short is most likely to be an essential gene;
- 3) selectivity can be achieved towards both *HsGSK3* and greatly towards *HsCDK2*;
- 4) a high enzyme inhibition (~99%) is required for cell efficacy, due to the millimolar concentrations of ATP in the trypanosomes, indicating that alternative mechanisms of inhibition other than competitive with respect to ATP would be required to achieve nanomolar cell efficacy by specific inhibition of *TbGSK3* short;
- 5) *TbGSK3* long is most likely to be an inactive enzyme, poorly expressed in BSF *T. brucei* if at all, therefore unable to compensate for *TbGSK3* short functions;
- 6) a specific chemical tool for *TbGSK3* short has been developed and characterised to reveal the function(s) of this protein kinase in *T. brucei*.

## **5.2 *TbGSK3* short is a druggable target in African trypanosomes**

The druggability of a target is a measure of the likelihood of identifying drug-like molecules that can modulate its activity in cells, in animal models and ultimately in human with both adequate efficacy and safety (Frearson *et al.*, 2007).

The biochemical screening carried out with the recombinant MBP-tagged *TbGSK3* short has led to the identification of a few series of small molecule inhibitors (GSK3 01-08) able to inhibit the enzyme in the nanomolar range (Woodland *et al.*,

2013). The previous knowledge about *HsCDK2* inhibitors had contributed to the synthesis of two additional series (GSK3 09-10) that were extremely active against *TbGSK3* short (Wyatt *et al.*, 2008, Urich *et al.*, 2014).

The identification of small molecule inhibitors against a recombinant enzyme is not *per se* an indication that a protein is druggable as the biochemical assay could be a model system unable to replicate the physiological conditions in which the native protein functions in the cell (Copeland *et al.*, 2005, Patricelli *et al.*, 2011). The establishment of a correlation between biochemical inhibition and cellular efficacy is a first step in the chemical validation of a drug target, although correlation does not necessarily mean causation (Hardy and Peet, 2004).

In the case of compounds from GSK3 09 series their nanomolar potency in the enzymatic assay translated into micromolar efficacy in BSF *T. b. brucei*, with high correlation among all the members of the series substituted in different positions. The shift in potency between the biochemical assay and the cellular assay could be explained by the ATP-competitive mechanism of inhibition and the differential ATP concentration in the cells with respect to the biochemical assay (Cheng and Prusoff, 1973; Hofer *et al.*, 1998). In contrast, although compounds from series GSK3 07 were very potent and had a very promising ligand efficacy, they had a minimal drop in cell efficacy that could not be explained by either the biochemical inhibition of *TbGSK3* short or their ATP-competitive mechanism of inhibition. Therefore this series was not considered to act on target anymore and was subsequently pursued phenotypically. This project, which was initially set-up as target-based approach, has evolved into a new project aimed at identifying phenotypically active compounds. These have been subsequently developed into potent leads for both HAT, Nagana and Leishmaniasis (DDU, unpublished work). Currently compounds developed from this series are facing preclinical development selection for the treatment of Leishmaniasis and identification

of their target(s) is ongoing. Preliminary results indicate that the cellular targets are other protein kinases.

Some of the GSK3 09 inhibitors were specific *Tb*GSK3 short inhibitors and selectivity towards both *Hs*CDK2 and *Hs*GSK3 could be achieved, up to 10,000 and 330 fold, respectively, with the most selective *Tb*GSK3 short compound being DDD00101234 so far. This compound has also moderate efficacy in trypanosomes and no toxicity in human MRC5 cells.

The chemical proteomic approach defined *Tb*GSK3 short as the main target for DD00085893 among the protein kinases that could be bound by this compound. Using kinobeads it was possible to screen a third of the *T. brucei* kinome, the biggest panel of *T. b. brucei* kinases available to date with the advantage of testing endogenous proteins rather than recombinant ones. The binding of an inhibitor to a target protein in the *T. b. brucei* lysate does not confirm that this binding is responsible for the efficacy in cells, particularly considering that the cellular ATP is removed from the lysate to reduce competition.

The definitive proof that a compound is acting on target is the modulation of the inhibitor potency directly in cells expressing different levels of protein (Hardy and Peet, 2004). The deletion of a *Tb*GSK3 short copy by gene knock-out increased the sensitivity to DDD00085893 by about 2-fold, whereas the over-expression of an ectopic copy of *Tb*GSK3 short decreased the sensitivity to DDD00085893 of ~3-fold (**Figure 4.6 A and B**). Further evidence that DDD00085893 acts on target comes from the ability of this inhibitor to revert the toxicity associated with over-expression and(or) hyper-activity of this protein. This small molecule inhibitor is able to work as a chemical switch, modulating the activity in the cells to a level that is no longer toxic for cell proliferation. Taken together, all these data provide the best chemical validation of

*TbGSK3* short as druggable target and provide a validated chemical tool for investigating the role of this kinase in *T. brucei*.

Other chemical proteomic approaches have recently identified two bioactive compounds, lapatinib and hypothemycin, which target *TbGSK3* short among other protein kinases in *T. b. brucei* cells (Katiyar *et al.*, 2013, Nishino *et al.*, 2013). These two molecules are not specific inhibitors of *TbGSK3* short as they act on multiple targets, so they cannot be used as tool compounds for interrogating *TbGSK3* short biology. Interestingly these compounds target a combination of proven essential kinases. Together with *TbGSK3* short, hypothemycin targets mainly *TbCLK1*, an essential kinase recently characterised to be an unconventional kinetochore protein (Akiyoshi and Gull, 2014). Lapatinib targets other three additional protein kinases, two of which have been characterised to be essential: *TbTLK1* (Li *et al.*, 2007; Li *et al.*, 2008b) and CK1.2 (Urbaniak *et al.*, 2009). While hypothemycin is quite a promiscuous inhibitor of human kinases targeting a total of 45 kinases with unsurprising associated toxicity (Schirmer *et al.*, 2006), lapatinib is an extremely specific inhibitor of epidermal growth factor receptor (EGFR) (Karaman *et al.*, 2008) already approved for clinical use<sup>a</sup>. These recent findings add further evidence to the druggability of *TbGSK3* short in *T. brucei*.

---

<sup>a</sup> New derivatives of lapatinib have been developed specifically to increase trypanocidal activity and have shown efficacy in animal model (Patel *et al.*, 2013). The phenotype of the lead compound **23** is similar to the phenotype for the knockdown of CK1.2 described by Jones and colleagues (2014).

Proof that *TbGSK3* short can be specifically targeted with small molecule inhibitors in cells (*in vitro*) is only the first step in its validation as drug target for HAT. It still remains to demonstrate the selective toxicity, efficacy and safety of the inhibition of this target with drug-like molecules in animal models (*in vivo*) and ultimately in human (clinical trials).

### 5.3 *TbGSK3* short is an essential gene

*TbGSK3* short is a druggable target in African trypanosomes, but is it an essential target? The fact that genetic knock-outs were not obtained even in presence of inducible expression of an ectopic copy supports the idea that this is an essential target and that there are no alternative pathways for the cell to compensate for the loss of this gene. Attempts to obtain genetic knock-outs of GSK3 have failed not only in *T. brucei*, but also in *Plasmodium falciparum* (Tewari *et al.*, 2010). Unfortunately the tendency not to publish negative results makes it difficult to review all of the genes for which genetic validation has been attempted and has failed.

All the RNAi studies performed in *T. b. brucei* are in agreement that *TbGSK3* short plays a role in BSF trypanosomes (Ojo *et al.*, 2008; Alsford *et al.*, 2011; Jones *et al.*, 2014). It has been also reported that protein down regulation causes cell arrest rather than cell death (Jones *et al.*, 2014), but irreversible cell death can be obtained dosing DDD0085893 at concentration equal to the EC<sub>99</sub> (**Figure 3.13 B**). This result indicates that almost complete enzymatic inhibition is required to get a cidal effect, and also alerts us that the outcomes of RNAi studies might reflect the level of protein knockdown rather than the real static/cidal nature of the target. The use of a target-specific inhibitor for determining the rate of killing and the recovery after wash-off gives a far better indication of the static or cidal nature of a target compared to RNAi

studies, but it requires well validated and specific tool molecules (Jacobs *et al.*, 2011; de Rycker *et al.*, 2012).

#### **5.4 Alternative strategies of inhibition to obtain selectivity and efficacy**

The extremely high enzyme inhibition required to obtain cellular efficacy poses two problems for the development of drug inhibitors of this protein kinase. First, for ATP-competitive inhibitors picomolar potencies are required to reach nanomolar efficacy in cells, and second, although we have demonstrated that selectivity can be obtained towards *HsGSK3 $\beta$* , lack of selectivity and potential toxicity could be still a problem for extremely potent compounds (Cohen and Goedert, 2004; Martinez *et al.*, 2002b). Furthermore it would be advantageous to have only a moderate inhibition of *HsGSK3 $\beta$*  considering the toxicity concerns associated with the inhibition of the human enzyme (hypoglycaemia, cardiotoxicity, neuronal deregulation, nephrotoxicity and potential tumorigenesis) (Eldar-Finkelman and Martinez, 2011).

The development of inhibitors with alternative mechanisms of inhibition could address both the selectivity and efficacy issues. At present, alternative strategies of inhibition have been already considered for *HsGSK3 $\beta$* : allosteric inhibitors (Palomo *et al.*, 2011); irreversible inhibitors (Castro *et al.*, 2008; Perez *et al.*, 2011; Dominguez *et al.*, 2012); substrate competitive inhibitors (Avrahami *et al.*, 2013; Licht-Murava and Eldar-Finkelman, 2011; Eldar-Finkelman *et al.*, 2010); and activators (Takahashi-Yanaga *et al.*, 2003; Takahashi-Yanaga *et al.*, 2006). Moreover, small molecules that regulate cell localization, interaction with scaffolding proteins or inhibitors of the priming kinases could provide alternative mechanisms of inhibition for *HsGSK3* (Meijer *et al.*, 2004).

### 5.4.1 Allosteric inhibitors

Allosteric inhibitors generally act by inducing conformational changes without binding in the ATP pocket. Allosteric modulators should be more selective with less off-target side effects as they bind in sites that are more unique to a specific kinase rather than the well conserved ATP binding site. However, drug resistance could be developed more easily, as mutations that are not directly responsible for the kinase activity are better tolerated than changes in the catalytic residues that tend to be conserved (Johnson, 2009).

Palomo and colleagues (2011) have performed a computational study on *HsGSK3 $\beta$*  and found that there are 7 targetable cavities which are present in the majority of GSK3-ligand crystal structures analysed. Of these seven cavities, only three are known binding sites of GSK3: 1) the ATP binding pocket, 2) the substrate binding site and 3) the axin/fratide binding pocket (Palomo *et al.*, 2011).

The ATP noncompetitive inhibitor manzamine A binds to the enzyme in the substrate binding site constituted by three basic residues, Arg96, Arg180, and Lys205, near the activation loop which is responsible for the binding of the primed substrate. The binding of this inhibitor in the substrate pocket was postulated by computational studies and experimentally confirmed by competition studies with a synthetic primed substrate (Palomo *et al.*, 2011). A similar site seems to be occupied by L803, a peptide reported to bind in the cavity near loop 85-95 (Avrahami *et al.*, 2013). Lastly the quinoline derivative VP0.7 is a reversible allosteric inhibitor of *HsGSK3 $\beta$*  (IC<sub>50</sub> ~3  $\mu$ M), non-competitive towards both ATP and a synthetic primed substrate, which has been docked to bind preferentially in cavity 7 with a resulting conformational change that will impede substrate binding (Palomo *et al.*, 2011). Other “allosteric” cavities have never been explored by small molecule inhibitors, but could be potential targeted by rational design to achieve both efficacy and selectivity. Further computational studies



are required to evaluate if similar cavities are present in *Tb*GSK3 short and whether key differences could be exploited to achieve selectivity towards the human enzyme.

An alternative “allosteric” inhibition could be achieved by targeting the inactive conformation. This approach has some pharmacological advantages compared to the targeting of the active conformation: the inactive configuration is more unique than the active conformation shared among all protein kinases, and more importantly the inactive kinase has a lower affinity for the ATP and an inhibitor specifically targeting this conformation will be less susceptible to the competition of the intracellular ATP (Johnson, 2009; Knight and Shokat, 2005).

Lapatinib, targeting *Tb*GSK3 short in *T. b. brucei* cells (see section 5.2) (Katiyar *et al.*, 2013), is a tyrosine kinase inhibitor known to specifically target the inactive conformation of the Epidermal Growth Factor Receptor (EGFR) (Wood *et al.*, 2004). Another interesting property of lapatinib is its slow off-rate (half-life 300 min) compared to other inhibitors that target the active conformation of EGFR, such as gefitinib and erlotinib, with off-rates of approximately 30 min (Johnson, 2009; Wood *et al.*, 2004). The inhibitor dissociation from the inactive conformation is probably slow as requires conformational changes or alternatively because of the very tight binding. The combination of a slow off-rate and specificity towards the inactive conformation are desirable characteristics for inhibitors of *Tb*GSK3 short. Lapatinib targets *Tb*GSK3 short, thus this suggests that it is possible to target the inactive conformation of this protein kinase.

*Hs*GSK3 $\beta$  is a constitutively active protein. The tyrosine phosphorylation of the activation loop (Tyr216 in *Hs*GSK3 $\beta$ ) is an intramolecular chaperone-dependent event happening immediately after translation (Lochhead *et al.*, 2006). The tyrosine phosphorylation increases the protein stability (Cole *et al.*, 2004) and locks the protein in an active conformation. Once activated, *Hs*GSK3 $\beta$  is an intermolecular Ser/Thr

kinase that is no longer dependent on chaperone activity (Lochhead *et al.*, 2006). Some inhibitors are able to target specifically the activation process (i.e. geldanamycin, radicicol, LiCl, SB216763) and their effect could be monitored using the rabbit reticulocyte lysate translation system (Lochhead *et al.*, 2006). Biochemical assay formats using the constitutionally active recombinant protein would miss inhibitors of the auto-activation of GSK3, so novel assays able to monitor the transient post-translational activation should be developed to find inhibitors with this alternative mechanism of inhibition.

#### **5.4.2 Irreversible inhibition**

Irreversible inhibitors are an attractive alternative to ATP-competitive inhibitors: high inhibitor concentration is not required for efficacy in cells as the inhibitor cannot dissociate from its target. Thus to restore activity requires protein re-synthesis; irreversible inhibition can prevent the development of resistance; with irreversible inhibitors it is possible to obtain an efficacy far higher than the one obtained using reversible inhibitors without prolonged exposure, and rational design can help the development of molecules that specifically target the protein kinase of interest (Rastelli *et al.*, 2008; Liu *et al.*, 2013). Irreversible inhibitors are also powerful tools as mutants of the reactive cysteine can rescue the inhibition phenotype providing in-cell evidence of the mode of action and specificity of inhibition (Liu *et al.*, 2013).

Usually an irreversible inhibitor is specifically targeted to a binding site where it can initially have reversible interactions, then the proximity between a cysteine (or other amino acid side chain) and the reactive moiety of the molecule allows the establishment of the irreversible interaction. In the case of irreversible inhibitors, the potency cannot be monitored by measurement of  $IC_{50}$  or  $K_i$  values, as the time of incubation plays a relevant role. Both the  $K_i$  of the non-covalent initial binding and the rate of inhibition of

the covalent binding should be taken in consideration for the development of irreversible inhibitors (Liu *et al.*, 2013). The turnover of the protein is another important factor to take in consideration, as the protein function is recovered when new protein is synthesised (Liu *et al.*, 2013).

*HsGSK3* is among the human protein kinases that have a reactive cysteine immediately before of the DFG motif that could be targeted by irreversible inhibitors. Indeed this binding site has been exploited for the development of two inhibitors of *HsGSK3 $\beta$* : Tideglusib (Castro *et al.*, 2008; Dominguez *et al.*, 2012) and compound 40, a maleimide-based molecule substituted with a reactive halomethylketo (HMK) moiety (Perez *et al.*, 2011).

Tideglusib is a thiadiazolidinone based compound currently in clinical development for both Alzheimer's disease and progressive supranuclear palsy. It was initially erroneously classified as non-ATP competitive inhibitor (Castro *et al.*, 2008) as irreversible inhibitors often display same pattern of inhibition as non-ATP competitive inhibitors in reciprocal plots, based on the assumption that the binding of the inhibitor to its target is rapid and reversible. Recently it has been characterised to be an irreversible inhibitor by a more in-depth mode of inhibition study (Dominguez *et al.*, 2012). The fact that its mechanism of inhibition was misinterpreted underlines the importance in drug discovery of undertaking rigorous studies of the mode of inhibition rather than relying exclusively on  $IC_{50}$  and  $K_i$  determination, but also determining reversibility of inhibition, time dependency and study of the kinetics of binding. Compound 40, a maleimide derivative substituted with a reactive halomethylketo (HMK) moiety (Perez *et al.*, 2011), was rationally designed with the aim to transform ATP-competitive reversible inhibitors in irreversible inhibitors, by covalently binding the reactive Cys199 with the HMK moiety. This compound was confirmed to bind covalently by Maldi-TOF analysis and its efficacy was demonstrated in cells by the decrease of Tau

phosphorylation. Irreversible inhibitors of GSK3 targeting Cys199 are specific towards CDK kinases as this residue is not conserved.

Hypothemycin is a polyketide natural product that binds irreversibly to 45 human protein kinases containing a specific CDXG motif (Schirmer *et al.*, 2006). In this motif the reactive Cys is close to the conserved Asp residue responsible for the binding to the Mg<sup>2+</sup> complexed with the ATP (Rastrelli *et al.*, 2008). *TbGSK3* short contains this motif and it is among the protein kinases targeted by this compound in *T. brucei* lysates (Nishino *et al.*, 2013). The cell efficacy of this inhibitor was mainly attributed to *TbCLK1* by virtue of the higher inhibition of both the recombinant enzyme and stronger binding of the native protein immunoprecipitated after incubation with hypothemycin. Nonetheless this demonstrates that irreversible inhibition can be achieved for *TbGSK3* short.

### 5.4.3 Substrate inhibitors

Protein kinases have different specificities for protein substrates and are more likely to bind and trans-phosphorylate specific consensus sequences. Competitive inhibitors for the substrate have generally weak interactions and protein-protein interaction inhibitors have been generally disregarded as difficult to develop.

Although both the binding site for the primed substrate and the auto-inhibition performed by the pseudo-substrate (phosphorylated Ser9 at the N-terminus) are well understood from a structural point of view (Dajani *et al.*, 2001; Ilouz *et al.*, 2008; Palomo *et al.*, 2011; Eldar-Finkelman *et al.*, 2010), very few substrate inhibitors have been reported for *HsGSK3β*. GSK3 has an unique substrate recognition motif (S/TXXXphosphoS/T). The substrate binds between the two lobes, in a site delimited by the P-loop and the short loop from Gln89 to Asn95 (Licht-Murava *et al.*, 2011). It establishes interactions with Gln89, Asn95 and Phe93. The 89-95 loop is well

conserved in kinetoplastids, with the exception of *TbGSK3* long where none of the residues essential for substrate recognition are conserved (see **Figure 1.8**). Thus, it is not surprising that the ability to trans-phosphorylate the synthetic primed substrate GSP2 is considerably lower compared to *TbGSK3* short (see section 4.1.8). This loop is involved in substrate recognition, but it does not affect the ability of GSK3 to auto-phosphorylate the Tyr216 in the activation loop (Licht-Murava *et al.*, 2011).

In addition, the phosphorylated priming site interacts with a positively charged pocket constituted by Arg96, Arg180 and Lys205. GSK3 preference for primed substrates has led to the identification of L803-mts: a peptide mimetic inhibitor based upon the recognition sequence of heat shock-factor 1 (HSF-1) (Eldar-Finkelman *et al.*, 2010). This peptide has a weak  $IC_{50}$  of 40  $\mu$ M, but has shown *in vivo* efficacy in diabetes and CNS models of neuroprotection, antidepressive activity and axon morphogenesis (Eldar-Finkelman *et al.*, 2010; Eldar-Finkelman and Martinez, 2011). Modifications of this peptide inhibitor aimed to increase the hydrophobic interactions with Phe93 have brought to the development of the more potent inhibitors L803F and PQ9P (Eldar-Finkelman and Martinez, 2011). In conclusion, although substrate competitive inhibition can be obtained for *HsGSK3 $\beta$* , this is associated with moderate efficacy, but greater selectivity, both characteristics that are considered positive attributes for the therapeutic uses of *HsGSK3 $\beta$*  inhibitors. It would be interesting to investigate whether it would be possible to develop selective substrate competitive inhibitors of *TbGSK3* short, although both parasite and human kinases are very similar in their substrate recognition sites (the only difference is the substitution of one residue in the 85-95 loop).

#### 5.4.4 Alternative screening strategies

The development of *Tb*GSK3 short inhibitors with an alternative mechanism of inhibition rather than ATP-competitive would require different assays and screening strategies. The *Tb*GSK3 short screening campaigns, conducted here in Dundee and elsewhere, were biased towards the identification of ATP-competitive inhibitors because the ATP concentration was set below the  $K_m$  for the ATP and because of the nature of compounds tested (kinase focussed library and known *Hs*GSK3 inhibitors) (Ojo *et al.*, 2008, Oduor *et al.*, 2011). Allosteric inhibitors could be identified by the screening of fragment libraries in assay platforms not biased towards ATP competitors. The identified binders with higher ligand efficiency could be rationally improved using the available homology models built for *Tb*GSK3 short to expand them towards allosteric cavities (Hopkins *et al.*, 2004; Ciulli and Abell, 2007). By adopting biophysical assays, it would be possible to obtain indications on the kinetic properties of the binders (slow binders, longer residence time); properties that could direct in the selection of the best compound for development, rather than rely exclusively on potency ( $IC_{50}$  or  $K_d$ ) (Copeland *et al.*, 2006). The available recombinant BAP-tagged *Tb*GSK3 short protein, suitable for biotinylation, could be used to develop biophysical assays for both surface plasmon resonance and biolayer-interferometry-based platforms. Unfortunately *Tb*GSK3 short has not been crystallized to date, but the new potent ligands from series GS309 could be used for co-crystallization studies. The availability of the enzyme crystal structure could be of great help in the development of fragment derived compounds, as it would give confirmation of their mode of binding and direct their chemical development.

## 5.5 Generation of a chemical tool to dissect new biology

DDD00085893 has all the wanted characteristics of the chemical probe: it is specific enough to correlate the *in vitro* potency to its cellular efficacy, its mechanism of inhibition has been studied and is coherent with its efficacy in cells, and its efficacy has been tested in target over-expressing cell lines (Hardy and Peet, 2004; Cohen, 2010; Frye, 2010; Workman and Collins, 2010). The combined use of both chemical and genetic approaches provides an orthogonal validation and the establishment of causal relationship between the inhibition of the target and in cell effect. As a validated chemical probe for *TbGSK3* short DDD00085893 can be used for interrogating its role(s) and function(s) in *T. brucei* and for the identification of its physiological substrates.

The functions of *TbGSK3* short in *T. brucei* are unknown, but on the basis of *HsGSK3 $\beta$*  roles and what is known on other species, it can be postulated a role in cell cycle control (Gladden and Diehl, 2005); in protein synthesis (Wang *et al.*, 2001); regulation of flagellar assembly and intraflagellar transport (Wilson and Lefebvre, 2004, Nett *et al.*, 2009b; Absalon *et al.*, 2008a; Absalon *et al.*, 2008b); and pro-apoptotic action (Beurel and Jope, 2006; Xingi *et al.*, 2009) among others. *TbGSK3* could also play a role in protein degradation and targeting for ubiquitination, determining protein stability and turnover (Xu *et al.*, 2009).

Preliminary SILAC experiments have been performed by Dr Michael Urbaniak using DDD00085893 (unpublished data) to identify the substrates of *TbGSK3* short in BSF trypanosomes; although protein levels are not affected by sub-lethal doses, multiple sites have both significantly increased and decreased levels of phosphorylation. Target proteins with higher levels of change in phosphorylation, combined with the knowledge of the pathways in which *HsGSK3 $\beta$*  is involved, could be prioritized for

substrate validation in conjunction with pharmacological studies of the phenotype of cells treated with DDD00085893.

## 5.6 Conclusions

This kinase target-based project had two main outcomes. First, the genetic-chemical validation of *Tb*GSK3 short as drug target using compounds from the GSK3 09 series that were synthesized based upon previous literature knowledge (Wyatt *et al.*, 2008). Second, and the identification of a phenotypically active series (GSK3 07) that acts off-target with efficacy in animal models for both HAT and leishmaniasis.

The limited cell efficacy of specific and potent *Tb*GSK3 short inhibitors in trypanosomes and the narrow window of selectivity towards the *Hs*GSK3 $\beta$  caused the de-prioritization of *Tb*GSK3 short as drug target in the DDU pipeline, even though this protein target can be now regarded as fully chemically validated *in vitro* for HAT. However, some important lessons from this project can be learned, in particular considering more protein kinase target-based projects are entering the DDU pipeline as potential drug targets for HAT and leishmaniasis.

Protein kinases as drug targets should be selected according to their essentiality proven in animal models and/or by RNAi. The target proteins that require lower degree of enzymatic inhibition or protein knock-down for a cidal effect in cells should be prioritised for screening (Jones *et al.*, 2014). Furthermore, the screening cascades should not be limited to the identification of ATP-competitive inhibitors and where possible an array of assay platforms and conditions should be tested in order to identify compounds with alternative mechanisms of inhibition or favourable kinetics.

Although specificity is a vital characteristic for a chemical probe, it is not a requirement for a lead molecule. Promiscuous inhibitors towards a pool of essential parasitic protein kinases, but selective towards the human isoforms would be the best



candidates for further chemical development. Until now the limiting step in the execution of protein kinase target-based projects has been the availability of active recombinant protein, and, even when available, recombinant enzymes have not always been able to predict cell efficacy (Urbaniak *et al.*, 2012a). Recently global approaches have informed us about essentiality of protein kinases (Alsford *et al.*, 2011; Jones *et al.*, 2014), level of protein expression and phosphorylation in *T. brucei* (Urbaniak *et al.*, 2012b and Urbaniak *et al.*, 2013). We need now to automate medium throughput systems integrated with proteomics that allow the characterization of phenotypically active compounds directly on cell lysates (Bantscheff *et al.*, 2007; Patricelli *et al.*, 2011), in order to identify potent binders to multiple essential native protein kinases. The development of these platforms will inform whether it is possible to identify potent kinase inhibitors with favourable network pharmacology for the treatment of HAT (Hopkins *et al.*, 2008).

## List of references

- Absalon, S., Blisnick, T., Bonhivers, M., Kohl, L., Cayet, N., Toutirais, G., Buisson, J., Robinson, D. and Bastin, P. (2008a). Flagellum elongation is required for correct structure, orientation and function of the flagellar pocket in *Trypanosoma brucei*. *J. Cell Sci.* *121*, 3704-3716.
- Absalon, S., Blisnick, T., Kohl, L., Toutirais, G., Dore, G., Julkowska, D., Tavenet, A. and Bastin, P. (2008b). Intraflagellar transport and functional analysis of genes required for flagellum formation in trypanosomes. *Mol. Biol. Cell* *19*, 929-944.
- Aguero, F., Al-Lazikani, B., Aslett, M., Berriman, M., Buckner, F.S., Campbell, R.K., Carmona, S., Carruthers, I.M., Chan, A.W., Chen, F., Crowther, G.J., Doyle, M.A., Hertz-Fowler, C., Hopkins, A.L., McAllister, G., Nwaka, S., Overington, J.P., Pain, A., Paolini, G.V., Pieper, U., Ralph, S.A., Riechers, A., Roos, D.S., Sali, A., Shanmugam, D., Suzuki, T., Van Voorhis, W.C. and Verlinde, C.L. (2008). Genomic-scale prioritization of drug targets: the TDR Targets database. *Nat. Rev. Drug Discov.* *7*, 900-907.
- Akiyoshi, B. and Gull, K. (2014). Discovery of unconventional kinetochores in kinetoplastids. *Cell* *156*, 1247-1258.
- Allocco, J.J., Donald, R., Zhong, T., Lee, A., Tang, Y.S., Hendrickson, R.C., Liberator, P. and Nare, B. (2006). Inhibitors of casein kinase 1 block the growth of *Leishmania major* promastigotes *in vitro*. *Int. J. Parasitol.* *36*, 1249-1259.
- 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 antitrypanosomal drug efficacy and resistance. *Nature* *482*, 232-236.
- 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 Res.* *21*, 915-924.
- Avrahami, L., Licht-Murava, A., Eisenstein, M. and Eldar-Finkelman, H. (2013). GSK-3 inhibition: achieving moderate efficacy with high selectivity. *Biochim. Biophys. Acta* *1834*, 1410-1414.
- Bain, J., McLauchlan, H., Elliott, M. and Cohen, P. (2003). The specificities of protein kinase inhibitors: an update. *Biochem. J.* *371*, 199-204.
- Bain, J., Plater, L., Elliott, M., Shpiro, N., Hastie, C.J., McLauchlan, H., Klevernic, I., Arthur, J.S., Alessi, D.R. and Cohen, P. (2007). The selectivity of protein kinase inhibitors: a further update. *Biochem. J.* *408*, 297-315.
- 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. *Mol. Biochem. Parasitol.* *176*, 55-57.

- Baker, N., Glover, L., Munday, J.C., Aguinaga Andres, D., Barrett, M.P., de Koning, H.P. and Horn, D. (2012). Aquaglyceroporin 2 controls susceptibility to melarsoprol and pentamidine in African trypanosomes. *Proc. Natl. Acad. Sci. USA* *109*, 10996-11001.
- 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. *Nat. Biotechnol.* *25*, 1035-1044.
- Barker, A.J., Gibson, K.H., Grundy, W., Godfrey, A.A., Barlow, J.J., Healy, M.P., Woodburn, J.R., Ashton, S.E., Curry, B.J., Scarlett, L., Henthorn, L. and Richards, L. (2001). Studies leading to the identification of ZD1839 (IRESSA): an orally active, selective epidermal growth factor receptor tyrosine kinase inhibitor targeted to the treatment of cancer. *Bioorg. Med. Chem. Lett.* *11*, 1911-1914.
- Barquilla, A., Crespo, J.L. and Navarro, M. (2008). Rapamycin inhibits trypanosome cell growth by preventing TOR complex 2 formation. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 14579-14584.
- Barquilla, A. and Navarro, M. (2009). Trypanosome TOR complex 2 functions in cytokinesis. *Cell Cycle* *8*, 697-699.
- 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*. *Proc. Natl. Acad. Sci. USA.* *109*, 14399-14404.
- Barrett, M.P. (2001). Drug resistance in sleeping sickness. Report of the Scientific Working Group on African trypanosomiasis. 96-111.
- Barrett, M.P., Boykin, D.W., Brun, R. and Tidwell, R.R. (2007). Human African trypanosomiasis: pharmacological re-engagement with a neglected disease. *Br. J. Pharmacol.* *152*, 1155-1171.
- Bengs, F., Scholz, A., Kuhn, D. and Wiese, M. (2005). LmxMPK9, a mitogen-activated protein kinase homologue affects flagellar length in *Leishmania mexicana*. *Mol. Microbiol.* *55*, 1606-1615.
- Berger, B.J., Carter, N.S. and Fairlamb, A.H. (1993). Polyamine and pentamidine metabolism in African trypanosomes. *Acta Trop.* *54*, 215-224.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., Bohme, U., Hannick, L., Aslett, M.A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U.C., Arrowsmith, C., Atkin, R.J., Barron, A.J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T.J., Churcher, C., Clark, L.N., Corton, C.H., Cronin, A., Davies, R.M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M.C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B.R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K.,

Kerhornou, A.X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., Macleod, A., Mooney, P.J., Moule, S., Martin, D.M., Morgan, G.W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C.S., Peterson, J., Quail, M.A., Rabbinowitsch, E., Rajandream, M.A., Reitter, C., Salzberg, S.L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A.J., Tallon, L., Turner, C.M., Tait, A., Tivey, A.R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M.D., Embley, T.M., Gull, K., Ullu, E., Barry, J.D., Fairlamb, A.H., Opperdoes, F., Barrell, B.G., Donelson, J.E., Hall, N., Fraser, C.M., Melville, S.E. and El-Sayed, N.M. (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Science* 309, 416-422.

Bertrand, J.A., Thieffine, S., Vulpetti, A., Cristiani, C., Valsasina, B., Knapp, S., Kalisz, H.M. and Flocco, M. (2003). Structural characterization of the GSK-3beta active site using selective and non-selective ATP-mimetic inhibitors. *J. Mol. Biol.* 333, 393-407.

Beurel, E. and Jope, R.S. (2006). The paradoxical pro- and anti-apoptotic actions of GSK3 in the intrinsic and extrinsic apoptosis signaling pathways. *Progress in Neurobiology* 79, 173-189.

Bitonti, A.J., McCann, P.P. and Sjoerdsma, A. (1986). Necessity of antibody response in the treatment of African trypanosomiasis with alpha-difluoromethylornithine. *Biochem. Pharmacol.* 35, 331-334.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248-254.

Brenk, R., Schipani, A., James, D., Krasowski, A., Gilbert, I.H., Frearson, J. and Wyatt, P.G. (2008). Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. *ChemMedChem* 3, 435-444.

Brown, D. and Superti-Furga, G. (2003). Rediscovering the sweet spot in drug discovery. *Drug Discov. Today* 8, 1067-1077.

Brun, R., Blum, J., Chappuis, F. and Burri, C. (2010). Human African trypanosomiasis. *Lancet* 375, 148-159.

Brun, R., Don, R., Jacobs, R.T., Wang, M.Z. and Barrett, M.P. (2011). Development of novel drugs for human African trypanosomiasis. *Future Microbiol* 6, 677-691.

Burkard, G., Fragoso, C.M. and Roditi, I. (2007). Highly efficient stable transformation of bloodstream forms of *Trypanosoma brucei*. *Mol. Biochem. Parasitol* 153, 220-223.

Burnham, K.P., and Anderson, D. R. (2002). Model Selection and Inference: A Practical Information-Theoretic Approach, 2nd Ed. New York, Springer-Verlag.

Carter, N.S., Berger, B.J. and Fairlamb, A.H. (1995). Uptake of diamidine drugs by the P2 nucleoside transporter in melarsen-sensitive and -resistant *Trypanosoma brucei brucei*. *J. Biol. Chem.* 270, 28153-28157.

Carter, N.S. and Fairlamb, A.H. (1993). Arsenical-resistant trypanosomes lack an unusual adenosine transporter. *Nature* 361, 173-176.

Castro, A., Encinas, A., Gil, C., Brase, S., Porcal, W., Perez, C., Moreno, F.J. and Martinez, A. (2008). Non-ATP competitive glycogen synthase kinase 3beta (GSK-3beta) inhibitors: study of structural requirements for thiadiazolidinone derivatives. *Bioorg. Med. Chem.* 16, 495-510.

Cheng, Y. and Prusoff, W.H. (1973). Relationship between the inhibition constant (K<sub>1</sub>) and the concentration of inhibitor which causes 50 per cent inhibition (I<sub>50</sub>) of an enzymatic reaction. *Biochem. Pharmacol.* 22, 3099-3108.

Ciulli, A. and Abell, C. (2007). Fragment-based approaches to enzyme inhibition. *Curr. Opin. Biotechnol.* 18, 489-496.

Clarkson, A.B., Jr., Bienen, E.J., Bacchi, C.J., McCann, P.P., Nathan, H.C., Hutner, S.H. and Sjoerdsma, A. (1984). New drug combination for experimental late-stage African trypanosomiasis: DL-alpha-difluoromethylornithine (DFMO) with suramin. *Am. J. Trop. Med. Hyg.* 33, 1073-1077.

Clayton, C., Adams, M., Almeida, R., Baltz, T., Barrett, M., Bastien, P., Belli, S., Beverley, S., Biteau, N., Blackwell, J., Blaineau, C., Boshart, M., Bringaud, F., Cross, G., Cruz, A., Degrave, W., Donelson, J., El-Sayed, N., Fu, G., Ersfeld, K., Gibson, W., Gull, K., Ivens, A., Kelly, J., Vanhamme, L. and et al. (1998). Genetic nomenclature for *Trypanosoma* and *Leishmania*. *Mol. Biochem. Parasitol.* 97, 221-224.

Clayton, C. and Shapira, M. (2007). Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Molecular and Biochemical Parasitology* 156, 93-101.

Cleghorn, L.A., Woodland, A., Collie, I.T., Torrie, L.S., Norcross, N., Luksch, T., Mpmhanga, C., Walker, R.G., Mottram, J.C., Brenk, R., Frearson, J.A., Gilbert, I.H. and Wyatt, P.G. (2011). Identification of inhibitors of the *Leishmania* cdc2-related protein kinase CRK3. *ChemMedChem* 6, 2214-2224.

Cohen, P. (2000). The regulation of protein function by multisite phosphorylation - a 25 year update. *Trends Biochem. Sci.* 25, 596-601.

Cohen, P. (2010). Guidelines for the effective use of chemical inhibitors of protein function to understand their roles in cell regulation. *Biochem. J.* 425, 53-54.

Cohen, P., Alessi, D.R. and Cross, D.A. (1997). PDK1, one of the missing links in insulin signal transduction? *FEBS Lett.* 410, 3-10.

Cohen, P. and Goedert, M. (2004). GSK3 inhibitors: development and therapeutic potential. *Nat Rev Drug Discov* 3, 479-487.

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, 249-255.

Copeland, R.A. (2003). Mechanistic considerations in high-throughput screening. *Anal. Biochem.* 320, 1-12.

Copeland, R.A. (2005). Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide for Medicinal Chemists and Pharmacologists. John Wiley & Sons, Inc.

Copeland, R.A., Pompliano, D.L. and Meek, T.D. (2006). Drug-target residence time and its implications for lead optimization. *Nat. Rev. Drug Discov.* 5, 730-739.

Cross, D.A., Culbert, A.A., Chalmers, K.A., Facci, L., Skaper, S.D. and Reith, A.D. (2001). Selective small-molecule inhibitors of glycogen synthase kinase-3 activity protect primary neurones from death. *J. Neurochem.* 77, 94-102.

Cross, D.A.E., Alessi, D.R., Cohen, P., Andjelkovich, M. and Hemmings, B.A. (1995). Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B. *Nature* 378, 785-789.

Czock, D. and Keller, F. (2007). Mechanism-based pharmacokinetic-pharmacodynamic modeling of antimicrobial drug effects. *J. Pharmacokinet. Pharmacodyn.* 34, 727-751.

Dajani, R., Fraser, E., Roe, S.M., Young, N., Good, V., Dale, T.C. and Pearl, L.H. (2001). Crystal structure of glycogen synthase kinase 3 $\beta$ : Structural basis for phosphate-primed substrate specificity and autoinhibition. *Cell* 105, 721-732.

de Graffenried, C.L., Anrather, D., Von Raussendorf, F. and Warren, G. (2013). Polo-like kinase phosphorylation of bilobe-resident TbCentrin2 facilitates flagellar inheritance in *Trypanosoma brucei*. *Mol. Biol. Cell* 24, 1947-1963.

de Graffenried, C.L., Ho, H.H. and Warren, G. (2008). Polo-like kinase is required for Golgi and bilobe biogenesis in *Trypanosoma brucei*. *J. Cell Biol.* 181, 431-438.

de Jesus, T.C., Tonelli, R.R., Nardelli, S.C., da Silva Augusto, L., Motta, M.C., 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*. *J. Biol. Chem.* 285, 24131-24140.

De Koning, H.P. (2001). Uptake of pentamidine in *Trypanosoma brucei brucei* is mediated by three distinct transporters: implications for cross-resistance with arsenicals. *Mol. Pharmacol.* 59, 586-592.

De Rycker, M., Hallyburton, I., Thomas, J., Campbell, L., Wyllie, S., Joshi, D., Cameron, S., Gilbert, I.H., Wyatt, P.G., Frearson, J.A., Fairlamb, A.H. and Gray, D.W. (2013). Comparison of a high-throughput high-content intracellular

*Leishmania donovani* assay with an axenic amastigote assay. *Antimicrob. Agents Chemother.* 57, 2913-2922.

De Rycker, M., O'Neill, S., Joshi, D., Campbell, L., Gray, D.W. and Fairlamb, A.H. (2012). A static-cidal assay for *Trypanosoma brucei* to aid hit prioritisation for progression into drug discovery programmes. *PLoS Negl Trop Dis* 6, e1932.

Dean, S., Marchetti, R., Kirk, K. and Matthews, K.R. (2009). A surface transporter family conveys the trypanosome differentiation signal. *Nature* 459, 213-217.

Denise, H. and Barrett, M.P. (2001). Uptake and mode of action of drugs used against sleeping sickness. *Biochem. Pharmacol.* 61, 1-5.

Diaz-Gonzalez, R., Kuhlmann, F.M., Galan-Rodriguez, C., Madeira da Silva, L., 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 Negl Trop Dis* 5, e1297.

Doble, B.W. and Woodgett, J.R. (2003). GSK-3: tricks of the trade for a multi-tasking kinase. *J. Cell Sci.* 116, 1175-1186.

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*. *Eukaryot. Cell* 5, 1126-1135.

Dominguez, J.M., Fuertes, A., Orozco, L., del Monte-Millan, M., Delgado, E. and Medina, M. (2012). Evidence for irreversible inhibition of glycogen synthase kinase-3beta by tideglusib. *J. Biol. Chem.* 287, 893-904.

Droucheau, E., Primot, A., Thomas, V., Mattei, D., Knockaert, M., Richardson, C., Sallicandro, P., Alano, P., Jafarshad, A., Baratte, B., Kunick, C., Parzy, D., Pearl, L., Doerig, C. and Meijer, L. (2004). *Plasmodium falciparum* glycogen synthase kinase-3: molecular model, expression, intracellular localisation and selective inhibitors. *Biochim. Biophys. Acta* 1697, 181-196.

Druker, B.J., Tamura, S., Buchdunger, E., Ohno, S., Segal, G.M., Fanning, S., Zimmermann, J. and Lydon, N.B. (1996). Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat. Med.* 2, 561-566.

Eldar-Finkelman, H., Argast, G.M., Foord, O., Fischer, E.H. and Krebs, E.G. (1996). Expression and characterization of glycogen synthase kinase-3 mutants and their effect on glycogen synthase activity in intact cells. *Proc. Natl. Acad. Sci. USA* 93, 10228-10233.

Eldar-Finkelman, H., Licht-Murava, A., Pietrovski, S. and Eisenstein, M. (2010). Substrate competitive GSK-3 inhibitors - strategy and implications. *Biochim. Biophys. Acta* 1804, 598-603.

Eldar-Finkelman, H. and Martinez, A. (2011). GSK-3 Inhibitors: Preclinical and Clinical Focus on CNS. *Front Mol Neurosci* 4, 32.

- Ellis, J., Sarkar, M., Hendriks, E. and Matthews, K. (2004). A novel ERK-like, CRK-like protein kinase that modulates growth in *Trypanosoma brucei* via an autoregulatory C-terminal extension. *Mol. Microbiol.* 53, 1487-1499.
- 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., 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* 309, 404-409.
- Fairlamb, A.H. (2003). Chemotherapy of human African trypanosomiasis: current and future prospects. *Trends Parasitol.* 19, 488-494.
- Fairlamb, A.H. and Bowman, I.B. (1980). Uptake of the trypanocidal drug suramin by bloodstream forms of *Trypanosoma brucei* and its effect on respiration and growth rate *in vivo*. *Mol. Biochem. Parasitol.* 1, 315-333.
- Fairlamb, A.H., Henderson, G.B., Bacchi, C.J. and Cerami, A. (1987). *In vivo* effects of difluoromethylornithine on trypanothione and polyamine levels in bloodstream forms of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 24, 185-191.
- Fairlamb, A.H., Henderson, G.B. and Cerami, A. (1989). Trypanothione is the primary target for arsenical drugs against African trypanosomes. *Proc. Natl. Acad. Sci. USA* 86, 2607-2611.
- Fairlamb, A.H., Smith, K. and Hunter, K.J. (1992). The interaction of arsenical drugs with dihydrolipoamide and dihydrolipoamide dehydrogenase from arsenical resistant and sensitive strains of *Trypanosoma brucei brucei*. *Mol. Biochem. Parasitol.* 53, 223-231.
- Farr, H. and Gull, K. (2012). Cytokinesis in trypanosomes. *Cytoskeleton (Hoboken)* 69, 931-941.
- Fenn, K. and Matthews, K.R. (2007). The cell biology of *Trypanosoma brucei* differentiation. *Curr. Opin. Microbiol.* 10, 539-546.
- Ferrer, M., Chernikova, T.N., Timmis, K.N. and Golyshin, P.N. (2004). Expression of a temperature-sensitive esterase in a novel chaperone-based *Escherichia coli* strain. *Appl. Environ. Microbiol.* 70, 4499-4504.
- Ferrer, M., Chernikova, T.N., Yakimov, M.M., Golyshin, P.N. and Timmis, K.N. (2003). Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nat. Biotechnol.* 21, 1266-1267.



Flaspohler, J.A., Jensen, B.C., Saveria, T., Kifer, C.T. and Parsons, M. (2010). A novel protein kinase localized to lipid droplets is required for droplet biogenesis in trypanosomes. *Eukaryot. Cell* 9, 1702-1710.

Frame, S. and Cohen, P. (2001). GSK3 takes centre stage more than 20 years after its discovery. *Biochem. J.* 359, 1-16.

Frearson, J.A., Brand, S., McElroy, S.P., Cleghorn, L.A., Smid, O., Stojanovski, L., Price, H.P., Guther, M.L., Torrie, L.S., Robinson, D.A., Hallyburton, I., Mpmhanga, C.P., Brannigan, J.A., Wilkinson, A.J., Hodgkinson, M., Hui, R., Qiu, W., Raimi, O.G., van Aalten, D.M., Brenk, R., Gilbert, I.H., Read, K.D., Fairlamb, A.H., Ferguson, M.A., Smith, D.F. and Wyatt, P.G. (2010). N-myristoyltransferase inhibitors as new leads to treat sleeping sickness. *Nature* 464, 728-732

Frearson, J.A., Wyatt, P.G., Gilbert, I.H. and Fairlamb, A.H. (2007). Target assessment for antiparasitic drug discovery. *Trends Parasitol.* 23, 589-595.

Fries, D.S. and Fairlamb, A. H. (2003 ). Antiprotozoal Agents. Burger's Medicinal Chemistry and Drug Discovery, Chemotherapeutic Agents. D. J. Abraham. New York Wiley and Sons. 5, 1033-1087.

Frye, S.V. (2010). The art of the chemical probe. *Nat. Chem. Biol.* 6, 159-161.

Gerber, P.R. and Muller, K. (1995). MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J. Comput. Aided Mol. Des.* 9, 251-268.

Ghoda, L., Wang, C.C. and Coffino, P. (1990). Ornithine decarboxylase of African trypanosomes. *Biochem. Soc. Trans.* 18, 739-740.

Gilbert, I.H. (2013). Drug discovery for neglected diseases: molecular target-based and phenotypic approaches. *J. Med. Chem.* 56, 7719-7726.

Gladden, A.B. and Diehl, J.A. (2005). Location, location, location: The role of cyclin D1 nuclear localization in cancer. *J. Cell. Biochem.* 96, 906-913.

Gooding, R.H. and Krafur, E.S. (2005). Tsetse genetics: contributions to biology, systematics, and control of tsetse flies. *Annu. Rev. Entomol.* 50, 101-123.

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

Graham, T.M., Tait, A. and Hide, G. (1998). Characterisation of a polo-like protein kinase gene homologue from an evolutionary divergent eukaryote, *Trypanosoma brucei*. *Gene* 207, 71-77.

Guttinger, A., Schwab, C., Morand, S., Roditi, I. and Vassella, E. (2007). A mitogen-activated protein kinase of *Trypanosoma brucei* confers resistance to temperature stress. *Mol. Biochem. Parasitol.* 153, 203-206.

- Hall, B.S., Bot, C. and Wilkinson, S.R. (2011). Nifurtimox activation by trypanosomal type I nitroreductases generates cytotoxic nitrile metabolites. *J. Biol. Chem.* 286, 13088-13095.
- 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. *J. Biol. Chem.* 281, 27600-27612.
- Hammarton, T.C. (2007a). Cell cycle regulation in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 153, 1-8.
- Hammarton, T.C., Monnerat, S. and Mottram, J.C. (2007b). Cytokinesis in trypanosomatids. *Curr. Opin. Microbiol.* 10, 520-527.
- Hammarton, T.C., Kramer, S., Tetley, L., Boshart, M. and Mottram, J.C. (2007c). *Trypanosoma brucei* Polo-like kinase is essential for basal body duplication, kDNA segregation and cytokinesis. *Mol. Microbiol.* 65, 1229-1248.
- Hammarton, T.C., Lillico, S.G., Welburn, S.C. and Mottram, J.C. (2005). *Trypanosoma brucei* MOB1 is required for accurate and efficient cytokinesis but not for exit from mitosis. *Mol. Microbiol.* 56, 104-116.
- Hanks, S.K. and Hunter, T. (1995). Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* 9, 576-596.
- Hardy, L.W. and Peet, N.P. (2004). The multiple orthogonal tools approach to define molecular causation in the validation of druggable targets. *Drug Discov. Today* 9, 117-126.
- He, X., Saint-Jeannet, J.P., Woodgett, J.R., Varmus, H.E. and Dawid, I.B. (1995). Glycogen synthase kinase-3 and dorsoventral patterning in *Xenopus* embryos. *Nature* 374, 617-622.
- Heby, O., Roberts, S.C. and Ullman, B. (2003). Polyamine biosynthetic enzymes as drug targets in parasitic protozoa. *Biochem. Soc. Trans.* 31, 415-419.
- Hem, S., Gherardini, P.F., Osorio y Fortea, J., Hourdel, V., Morales, M.A., Watanabe, R., Pescher, P., Kuzyk, M.A., Smith, D., Borchers, C.H., Zilberstein, D., Helmer-Citterich, M., Namane, A. and Spath, G.F. (2010). Identification of *Leishmania*-specific protein phosphorylation sites by LC-ESI-MS/MS and comparative genomics analyses. *Proteomics* 10, 3868-3883.
- Hoeflich, K.P., Luo, J., Rubie, E.A., Tsao, M.S., Jin, O. and Woodgett, J.R. (2000). Requirement for glycogen synthase kinase-3 $\beta$  in cell survival and NF- $\kappa$ B activation. *Nature* 406, 86-90.
- Hofer, A., Ekanem, J.T. and Thelander, L. (1998). Allosteric regulation of *Trypanosoma brucei* ribonucleotide reductase studied *in vitro* and *in vivo*. *J. Biol. Chem.* 273, 34098-34104.
- Hopkins, A.L. and Groom, C.R. (2002). The druggable genome. *Nat Rev Drug Discov* 1, 727-730.

Hopkins, A.L. (2008). Network pharmacology: the next paradigm in drug discovery. *Nat. Chem. Biol.* 4, 682-690.

Horn, D. and McCulloch, R. (2010). Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. *Curr. Opin. Microbiol.* 13, 700-705.

Hu, H., Yu, Z., Liu, Y., Wang, T., Wei, Y. and Li, Z. (2014). The Aurora B kinase in *Trypanosoma brucei* undergoes post-translational modifications and is targeted to various subcellular locations through binding to TbCPC1. *Mol. Microbiol.* 91, 256-274.

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*. *J. Biol. Chem.* 272, 10797-10803.

Huse, M. and Kuriyan, J. (2002). The conformational plasticity of protein kinases. *Cell* 109, 275-282.

Ikeda, K.N. and de Graffenried, C.L. (2012). Polo-like kinase is necessary for flagellum inheritance in *Trypanosoma brucei*. *J. Cell Sci.* 125, 3173-3184.

Ilouz, R., Pietrokovski, S., Eisenstein, M. and Eldar-Finkelman, H. (2008). New insights into the autoinhibition mechanism of glycogen synthase kinase-3beta. *J. Mol. Biol.* 383, 999-1007.

Iten, M., Mett, H., Evans, A., Enyaru, J.C., Brun, R. and Kaminsky, R. (1997). Alterations in ornithine decarboxylase characteristics account for tolerance of *Trypanosoma brucei rhodesiense* to D,L-alpha-difluoromethylornithine. *Antimicrob. Agents Chemother.* 41, 1922-1925.

Ivens, A.C., Peacock, C.S., Wortley, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Adlem, E., Aert, R., Anupama, A., Apostolou, Z., Attipoe, P., Bason, N., Bauser, C., Beck, A., Beverley, S.M., Bianchetti, G., Borzym, K., Bothe, G., Bruschi, C.V., Collins, M., Cadag, E., Ciarloni, L., Clayton, C., Coulson, R.M., Cronin, A., Cruz, A.K., Davies, R.M., De Gaudenzi, J., Dobson, D.E., Duesterhoeft, A., Fazelina, G., Fosker, N., Frasch, A.C., Fraser, A., Fuchs, M., Gabel, C., Goble, A., Goffeau, A., Harris, D., Hertz-Fowler, C., Hilbert, H., Horn, D., Huang, Y., Klages, S., Knights, A., Kube, M., Larke, N., Litvin, L., Lord, A., Louie, T., Marra, M., Masuy, D., Matthews, K., Michaeli, S., Mottram, J.C., Muller-Auer, S., Munden, H., Nelson, S., Norbertczak, H., Oliver, K., O'Neil, S., Pentony, M., Pohl, T.M., Price, C., Purnelle, B., Quail, M.A., Rabbinowitsch, E., Reinhardt, R., Rieger, M., Rinta, J., Robben, J., Robertson, L., Ruiz, J.C., Rutter, S., Saunders, D., Schafer, M., Schein, J., Schwartz, D.C., Seeger, K., Seyler, A., Sharp, S., Shin, H., Sivam, D., Squares, R., Squares, S., Tosato, V., Vogt, C., Volckaert, G., Wambutt, R., Warren, T., Wedler, H., Woodward, J., Zhou, S., Zimmermann, W., Smith, D.F., Blackwell, J.M., Stuart, K.D., Barrell, B. and Myler, P.J. (2005). The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309, 436-442.

Jackson, A.P., Sanders, M., Berry, A., McQuillan, J., Aslett, M.A., Quail, M.A., Chukualim, B., Capewell, P., MacLeod, A., Melville, S.E., Gibson, W., Barry,

J.D., Berriman, M. and Hertz-Fowler, C. (2010). The genome sequence of *Trypanosoma brucei* gambiense, causative agent of chronic human african trypanosomiasis. *PLoS Negl Trop Dis* 4, e658.

Jacobs, R.T., Nare, B., Wring, S.A., Orr, M.D., Chen, D., Sligar, J.M., Jenks, M.X., Noe, R.A., Bowling, T.S., Mercer, L.T., Rewerts, C., Gaukel, E., Owens, J., Parham, R., Randolph, R., Beaudet, B., Bacchi, C.J., Yarlett, N., Plattner, J.J., Freund, Y., Ding, C., Akama, T., Zhang, Y.K., Brun, R., Kaiser, M., Scandale, I. and Don, R. (2011). SCYX-7158, an orally-active benzoxaborole for the treatment of stage 2 human African trypanosomiasis. *PLoS Negl Trop Dis* 5, e1151.

Jennings, F.W. and Urquhart, G.M. (1983). The use of the 2 substituted 5-nitroimidazole, Fexinidazole (Hoe 239) in the treatment of chronic *T. brucei* infections in mice. *Z Parasitenkd* 69, 577-581.

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*. *Mol. Biochem. Parasitol.* 151, 28-40.

Jensen, B.C., Kifer, C.T. and Parsons, M. (2011). *Trypanosoma brucei*: Two mitogen activated protein kinase kinases are dispensable for growth and virulence of the bloodstream form. *Exp. Parasitol.* 128, 250-255.

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. *Mol. Microbiol.* 72, 442-458.

Johnson, L.N. (2009). Protein kinase inhibitors: contributions from structure to clinical compounds. *Q. Rev. Biophys.* 42, 1-40.

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 Pathog* 10, e1003886.

Jope, R.S. (2003). Lithium and GSK-3: one inhibitor, two inhibitory actions, multiple outcomes. *Trends Pharmacol. Sci.* 24, 441-443.

Kaiser, M., Bray, M.A., Cal, M., Bourdin Trunz, B., Torreele, E. and Brun, R. (2011). Antitrypanosomal activity of fexinidazole, a new oral nitroimidazole drug candidate for treatment of sleeping sickness. *Antimicrob. Agents Chemother.* 55, 5602-5608.

Kang, T., Wei, Y., Honaker, Y., Yamaguchi, H., Appella, E., Hung, M.C. and Piwnicka-Worms, H. (2008). GSK-3 beta targets Cdc25A for ubiquitin-mediated proteolysis, and GSK-3 beta inactivation correlates with Cdc25A overproduction in human cancers. *Cancer Cell* 13, 36-47.

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. *Nat. Biotechnol.* 26, 127-132.
- 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, e56150.
- Keiser, J., Ericsson, O. and Burri, C. (2000). Investigations of the metabolites of the trypanocidal drug melarsoprol. *Clin. Pharmacol. Ther.* 67, 478-488.
- Kennelly, P.J. (2003). Archaeal protein kinases and protein phosphatases: insights from genomics and biochemistry. *Biochem. J.* 370, 373-389.
- Knight, Z.A. and Shokat, K.M. (2005). Features of selective kinase inhibitors. *Chemistry & Biology* 12, 621-637.
- 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. and Meijer, L. (2000). Intracellular targets of cyclin-dependent kinase inhibitors: identification by affinity chromatography using immobilised inhibitors. *Chem Biol* 7, 411-422.
- Kramer, S. (2004). Characterisation of a PKA-like kinase from *Trypanosoma brucei*. Thesis, Munichen, Ludwig-Maximilians-Universitat.
- Kumar, P. and Wang, C.C. (2006). Dissociation of cytokinesis initiation from mitotic control in a eukaryote. *Eukaryot. Cell* 5, 92-102.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.
- Leroy, D. and Doerig, C. (2008). Drugging the *Plasmodium* kinome: the benefits of academia-industry synergy. *Trends Pharmacol. Sci.* 29, 241-249.
- Li, Z. (2012). Regulation of the cell division cycle in *Trypanosoma brucei*. *Eukaryot. Cell* 11, 1180-1190.
- 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. *J. Cell Sci.* 120, 3883-3894.
- Li, Z., Lee, J.H., Chu, F., Burlingame, A.L., Gunzl, A. and Wang, C.C. (2008a). Identification of a novel chromosomal passenger complex and its unique localization during cytokinesis in *Trypanosoma brucei*. *PLoS One* 3, e2354.
- Li, Z., Umeyama, T., Li, Z. and Wang, C.C. (2010). Polo-like kinase guides cytokinesis in *Trypanosoma brucei* through an indirect means. *Eukaryot. Cell* 9, 705-716.

- Li, Z., Umeyama, T. and Wang, C.C. (2008b). The chromosomal passenger complex and a mitotic kinesin interact with the Tousled-like kinase in trypanosomes to regulate mitosis and cytokinesis. *PLoS one* 3, e3814.
- 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 Pathog* 5, e1000575.
- Li, Z. and Wang, C.C. (2006). Changing roles of aurora-B kinase in two life cycle stages of *Trypanosoma brucei*. *Eukaryot. Cell* 5, 1026-1035.
- Licht-Murava, A. and Eldar-Finkelman, H. (2012). Exploiting substrate recognition for selective inhibition of protein kinases. *Curr. Pharm. Des.* 18, 2914-2920.
- Licht-Murava, A., Plotkin, B., Eisenstein, M. and Eldar-Finkelman, H. (2011). Elucidating substrate and inhibitor binding sites on the surface of GSK-3 $\beta$  and the refinement of a competitive inhibitor. *J. Mol. Biol.* 408, 366-378.
- Linding, R., Jensen, L.J., Ostheimer, G.J., van Vugt, M.A., Jorgensen, C., Miron, I.M., Diella, F., Colwill, K., Taylor, L., Elder, K., Metalnikov, P., Nguyen, V., Pasculescu, A., Jin, J., Park, J.G., Samson, L.D., Woodgett, J.R., Russell, R.B., Bork, P., Yaffe, M.B. and Pawson, T. (2007). Systematic discovery of *in vivo* phosphorylation networks. *Cell* 129, 1415-1426.
- Liu, Q., Sabnis, Y., Zhao, Z., Zhang, T., Buhrlage, S.J., Jones, L.H. and Gray, N.S. (2013). Developing irreversible inhibitors of the protein kinase cysteinome. *Chem. Biol.* 20, 146-159.
- Liu, Y., Tidwell, R.R. and Leibowitz, M.J. (1994). Inhibition of *in vitro* splicing of a group I intron of *Pneumocystis carinii*. *J. Eukaryot. Microbiol.* 41, 31-38.
- Lochhead, P.A., Kinstrie, R., Sibbet, G., Rawjee, T., Morrice, N. and Cleghon, V. (2006). A Chaperone-Dependent GSK3 Transitional Intermediate Mediates Activation-Loop Autophosphorylation. *Mol. Cell* 24, 627-633.
- Lozano-Nunez, 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*. *Mol. Biol. Cell* 24, 1321-1333.
- 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*. *J. Biol. Chem.* 285, 15356-15368.
- 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*. *Chem Biol Drug Des* 78, 454-463.

- Magnus, E., Vervoort, T. and Van Meirvenne, N. (1978). A card-agglutination test with stained trypanosomes (C.A.T.T.) for the serological diagnosis of *T. b. gambiense* trypanosomiasis. *Ann Soc Belg Med Trop* 58, 169-176.
- Malvy, D. and Chappuis, F. (2011). Sleeping sickness. *Clin Microbiol Infect* 17, 986-995.
- Manning, G., Whyte, D.B., Martinez, R., Hunter, T. and Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science* 298, 1912-1934.
- Martinez, A., Alonso, M., Castro, A., Perez, C. and Moreno, F.J. (2002a). First non-ATP competitive glycogen synthase kinase 3 beta (GSK-3beta) inhibitors: thiadiazolidinones (TDZD) as potential drugs for the treatment of Alzheimer's disease. *J. Med. Chem.* 45, 1292-1299.
- Martinez, A., Castro, A., Dorronsoro, I. and Alonso, M. (2002b). Glycogen synthase kinase 3 (GSK-3) inhibitors as new promising drugs for diabetes, neurodegeneration, cancer, and inflammation. *Med. Res. Rev.* 22, 373-384.
- Martinez, A., Gil, C. and Perez, D.I. (2011). Glycogen synthase kinase 3 inhibitors in the next horizon for Alzheimer's disease treatment. *Int J Alzheimers Dis* 2011, 280502.
- Matthews, K.R. (2005). The developmental cell biology of *Trypanosoma brucei*. *J. Cell Sci.* 118, 283-290.
- Meijer, L., Flajolet, M. and Greengard, P. (2004). Pharmacological inhibitors of glycogen synthase kinase 3. *Trends Pharmacol. Sci.* 25, 471-480.
- Mercer, L., Bowling, T., Perales, J., Freeman, J., Nguyen, T., Bacchi, C., Yarlett, N., Don, R., Jacobs, R. and Nare, B. (2011). 2,4-Diaminopyrimidines as potent inhibitors of *Trypanosoma brucei* and identification of molecular targets by a chemical proteomics approach. *PLoS Negl. Trop. Dis.* 5, e956.
- 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*. *Mol. Biochem. Parasitol.* 190, 44-49.
- Miranda-Saavedra, D. and Barton, G.J. (2007). Classification and functional annotation of eukaryotic protein kinases. *Proteins* 68, 893-914.
- Monnerat, S., Almeida Costa, C.I., Forkert, A.C., Benz, C., Hamilton, A., Tetley, L., Burchmore, R., Novo, C., Mottram, J.C. and Hammarton, T.C. (2013). Identification and Functional Characterisation of CRK12:CYC9, a Novel Cyclin-Dependent Kinase (CDK)-Cyclin Complex in. *PLoS One* 8, e67327.
- Mony, B.M., MacGregor, P., Ivens, A., Rojas, F., Cowton, A., Young, J., Horn, D. and Matthews, K. (2014). Genome-wide dissection of the quorum sensing signalling pathway in *Trypanosoma brucei*. *Nature* 505, 681-685.
- Moraes, M.C., Jesus, T.C., Hashimoto, N.N., Dey, M., Schwartz, K.J., Alves, V.S., Avila, C.C., Bangs, J.D., Dever, T.E., Schenkman, S. and Castilho, B.A.

- (2007). Novel membrane-bound eIF2alpha kinase in the flagellar pocket of *Trypanosoma brucei*. *Eukaryot. Cell* 6, 1979-1991.
- 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*. *Mol. Biochem. Parasitol.* 186, 73-76.
- Mukai, F., Ishiguro, K., Sano, Y. and Fujita, S.C. (2002). Alternative splicing isoform of tau protein kinase I/glycogen synthase kinase 3beta. *J. Neurochem.* 81, 1073-1083.
- Muller, I.B., Domenicali-Pfister, D., Roditi, I. and Vassella, E. (2002). Stage-specific requirement of a mitogen-activated protein kinase by *Trypanosoma brucei*. *Mol. Biol. Cell* 13, 3787-3799.
- Munday, J.C., Eze, A.A., Baker, N., Glover, L., Clucas, C., Aguinaga Andres, D., Natto, M.J., Teku, I.A., McDonald, J., Lee, R.S., Graf, F.E., Ludin, P., Burchmore, R.J., Turner, C.M., Tait, A., MacLeod, A., Maser, P., Barrett, M.P., Horn, D. and De Koning, H.P. (2014). *Trypanosoma brucei* aquaglyceroporin 2 is a high-affinity transporter for pentamidine and melaminophenyl arsenic drugs and the main genetic determinant of resistance to these drugs. *J. Antimicrob. Chemother.* 69, 651-663.
- Nakayasu, E.S., Gaynor, M.R., Sobreira, T.J., Ross, J.A. and Almeida, I.C. (2009). Phosphoproteomic analysis of the human pathogen *Trypanosoma cruzi* at the epimastigote stage. *Proteomics* 9, 3489-3506.
- Nare, B., Wring, S., Bacchi, C., Beaudet, B., Bowling, T., Brun, R., Chen, D., Ding, C., Freund, Y., Gaukel, E., Hussain, A., Jarnagin, K., Jenks, M., Kaiser, M., Mercer, L., Mejia, E., Noe, A., Orr, M., Parham, R., Plattner, J., Randolph, R., Rattendi, D., Rewerts, C., Sligar, J., Yarlett, N., Don, R. and Jacobs, R. (2010). Discovery of novel orally bioavailable oxaborole 6-carboxamides that demonstrate cure in a murine model of late-stage central nervous system african trypanosomiasis. *Antimicrob. Agents Chemother.* 54, 4379-4388.
- Naula, C., Parsons, M. and Mottram, J.C. (2005). Protein kinases as drug targets in trypanosomes and *Leishmania*. *Biochim. Biophys. Acta* 1754, 151-159.
- Nett, I.R., Davidson, L., Lamont, D. and Ferguson, M.A. (2009b). Identification and specific localization of tyrosine-phosphorylated proteins in *Trypanosoma brucei*. *Eukaryot. Cell* 8, 617-626.
- Nett, I.R., Martin, D.M., Miranda-Saavedra, D., Lamont, D., Barber, J.D., Mehlert, A. and Ferguson, M.A. (2009a). The phosphoproteome of bloodstream form *Trypanosoma brucei*, causative agent of African sleeping sickness. *Mol. Cell. Proteomics* 8, 1527-1538.
- Nishino, M., Choy, J.W., Gushwa, N.N., Osés-Prieto, J.A., Koupparis, K., Burlingame, A.L., Renslo, A.R., McKerrow, J.H. and Taunton, J. (2013). Hypothemycin, a fungal natural product, identifies therapeutic targets in *Trypanosoma brucei*. *Elife* 2, e00712.



Noble, M.E., Endicott, J.A. and Johnson, L.N. (2004). Protein kinase inhibitors: insights into drug design from structure. *Science* 303, 1800-1805.

Nwaka, S. and Hudson, A. (2006). Innovative lead discovery strategies for tropical diseases. *Nat Rev Drug Discov* 5, 941-955.

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. *Eur. J. Med. Chem.* 62, 777-784.

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 Negl. Trop. Dis.* 5, e1017.

Ojo, K.K., Gillespie, J.R., Riechers, A.J., Napuli, A.J., Verlinde, C.L., Buckner, F.S., Gelb, M.H., Domostoj, M.M., Wells, S.J., Scheer, A., Wells, T.N. and Van Voorhis, W.C. (2008). Glycogen synthase kinase 3 is a potential drug target for African trypanosomiasis therapy. *Antimicrob. Agents Chemother.* 52, 3710-3717.

Ojo, K.K., Arakaki, T.L., Napuli, A.J., Inampudi, K.K., Keyloun, K.R., Zhang, L., Hol, W.G., Verlinde, C.L., Merritt, E.A. and Van Voorhis, W.C. (2011). Structure determination of glycogen synthase kinase-3 from *Leishmania major* and comparative inhibitor structure-activity relationships with *Trypanosoma brucei* GSK-3. *Mol. Biochem. Parasitol.* 176, 98-108.

Osolodkin, D.I., Zakharevich, N.V., Palyulin, V.A., Danilenko, V.N. and Zefirov, N.S. (2011). Bioinformatic analysis of glycogen synthase kinase 3: human versus parasite kinases. *Parasitology* 138, 725-735.

Overath, P. and Engstler, M. (2004). Endocytosis, membrane recycling and sorting of GPI-anchored proteins: *Trypanosoma brucei* as a model system. *Mol. Microbiol.* 53, 735-744.

Palomo, V., Soteras, I., Perez, D.I., Perez, C., Gil, C., Campillo, N.E. and Martinez, A. (2011). Exploring the binding sites of glycogen synthase kinase 3. Identification and characterization of allosteric modulation cavities. *J. Med. Chem.* 54, 8461-8470.

Pap, M. and Cooper, G.M. (1998). Role of glycogen synthase kinase-3 in the phosphatidylinositol 3-kinase/Akt cell survival pathway. *J. Biol. Chem.* 273, 19929-19932.

Park, J.H., Brekken, D.L., Randall, A.C. and Parsons, M. (2002). Molecular cloning of *Trypanosoma brucei* CK2 catalytic subunits: the alpha isoform is nucleolar and phosphorylates the nucleolar protein Nopp44/46. *Mol. Biochem. Parasitol.* 119, 97-106.

Parsons, M., Valentine, M., Deans, J., Schieven, G.L. and Ledbetter, J.A. (1991). Distinct patterns of tyrosine phosphorylation during the life cycle of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 45, 241-248.

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.

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. *J. Med. Chem.* 56, 3820-3832.

Patterson, S., Alphey, M.S., Jones, D.C., Shanks, E.J., Street, I.P., Frearson, J.A., Wyatt, P.G., Gilbert, I.H. and Fairlamb, A.H. (2011). Dihydroquinazolines as a novel class of *Trypanosoma brucei* trypanothione reductase inhibitors: discovery, synthesis, and characterization of their binding mode by protein crystallography. *J. Med. Chem.* 54, 6514-6530.

Patricelli, Matthew P., Nomanbhoy, Tyzoon K., Wu, J., Brown, H., Zhou, D., Zhang, J., Jagannathan, S., Aban, A., Okerberg, E., Herring, C., Nordin, B., Weissig, H., Yang, Q., Lee, J.-D., Gray, Nathanael S. and Kozarich, John W. (2011). In Situ Kinase Profiling Reveals Functionally Relevant Properties of Native Kinases. *Chemistry & Biology* 18, 699-710.

Paulnock, D.M., Freeman, B.E. and Mansfield, J.M. (2010). Modulation of innate immunity by African trypanosomes. *Parasitology* 137, 2051-2063.

Pays, E. (2005). Regulation of antigen gene expression in *Trypanosoma brucei*. *Trends Parasitol.* 21, 517-520.

Perez, D.I., Palomo, V., Perez, C., Gil, C., Dans, P.D., Luque, F.J., Conde, S. and Martinez, A. (2011). Switching reversibility to irreversibility in glycogen synthase kinase 3 inhibitors: clues for specific design of new compounds. *J. Med. Chem.* 54, 4042-4056.

Plowman, G.D., Sudarsanam, S., Bingham, J., Whyte, D. and Hunter, T. (1999). The protein kinases of *Caenorhabditis elegans*: a model for signal transduction in multicellular organisms. *Proc. Natl. Acad. Sci. USA* 96, 13603-13610.

Pradel, L.C., Bonhivers, M., Landrein, N. and Robinson, D.R. (2006). NIMA-related kinase TbNRKC is involved in basal body separation in *Trypanosoma brucei*. *J. Cell Sci.* 119, 1852-1863.

Predonzani, A., Arnoldi, F., Lopez-Requena, A. and Burrone, O.R. (2008). *In vivo* site-specific biotinylation of proteins within the secretory pathway using a single vector system. *BMC Biotechnol.* 8, 41.

Priotto, G., Kasparian, S., Mutombo, W., Ngouama, D., Ghorashian, S., Arnold, U., Ghabri, S., Baudin, E., Buard, V., Kazadi-Kyanza, S., Ilunga, M., Mutangala, W., Pohlig, G., Schmid, C., Karunakara, U., Torreele, E. and Kande, V. (2009).

Nifurtimox-eflornithine combination therapy for second-stage African *Trypanosoma brucei gambiense* trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. *Lancet* 374, 56-64.

Qin, C.L., Tang, J. and Kim, K. (1998). Cloning and *in vitro* expression of TPK3, a *Toxoplasma gondii* homologue of shaggy/glycogen synthase kinase-3 kinases. *Mol. Biochem. Parasitol.* 93, 273-283.

Rastelli, G., Rosenfeld, R., Reid, R. and Santi, D.V. (2008). Molecular modeling and crystal structure of ERK2-hypothemycin complexes. *J. Struct. Biol.* 164, 18-23.

Raz, B., Iten, M., Grether-Buhler, 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 Trop.* 68, 139-147.

Renslo, A.R. and McKerrow, J.H. (2006). Drug discovery and development for neglected parasitic diseases. *Nat. Chem. Biol.* 2, 701-710.

Rial, D.V. and Ceccarelli, E.A. (2002). Removal of DnaK contamination during fusion protein purifications. *Protein Expr. Purif.* 25, 503-507.

Roberts, S.C., Scott, J., Gasteier, J.E., Jiang, Y., Brooks, B., Jardim, A., Carter, N.S., Heby, O. and Ullman, B. (2002). S-adenosylmethionine decarboxylase from *Leishmania donovani*. Molecular, genetic, and biochemical characterization of null mutants and overproducers. *J. Biol. Chem.* 277, 5902-5909.

Roper, J.R., Guther, M.L., Milne, K.G. and Ferguson, M.A. (2002). Galactose metabolism is essential for the African sleeping sickness parasite *Trypanosoma brucei*. *Proc. Natl. Acad. Sci. USA* 99, 5884-5889.

Rotureau, B., Morales, M.A., Bastin, P. and Spath, G.F. (2009). The flagellum-mitogen-activated protein kinase connection in Trypanosomatids: a key sensory role in parasite signalling and development? *Cell. Microbiol.* 11, 710-718.

Rubin, G.M., Yandell, M.D., Wortman, J.R., Gabor Miklos, G.L., Nelson, C.R., Hariharan, I.K., Fortini, M.E., Li, P.W., Apweiler, R., Fleischmann, W., Cherry, J.M., Henikoff, S., Skupski, M.P., Misra, S., Ashburner, M., Birney, E., Boguski, M.S., Brody, T., Brokstein, P., Celniker, S.E., Chervitz, S.A., Coates, D., Cravchik, A., Gabrielian, A., Galle, R.F., Gelbart, W.M., George, R.A., Goldstein, L.S., Gong, F., Guan, P., Harris, N.L., Hay, B.A., Hoskins, R.A., Li, J., Li, Z., Hynes, R.O., Jones, S.J., Kuehl, P.M., Lemaitre, B., Littleton, J.T., Morrison, D.K., Mungall, C., O'Farrell, P.H., Pickeral, O.K., Shue, C., Vossball, L.B., Zhang, J., Zhao, Q., Zheng, X.H. and Lewis, S. (2000). Comparative genomics of the eukaryotes. *Science* 287, 2204-2215.

Sali, A. and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779-815.

Sambrook, J. and Russell, D.W. (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sampah, M.E., Shen, L., Jilek, B.L. and Siliciano, R.F. (2011). Dose-response curve slope is a missing dimension in the analysis of HIV-1 drug resistance. *Proc. Natl. Acad. Sci. U. S. A.* 108, 7613-7618.

Schirmer, A., Kennedy, J., Murli, S., Reid, R. and Santi, D.V. (2006). Targeted covalent inactivation of protein kinases by resorcylic acid lactone polyketides. *Proc. Natl. Acad. Sci. USA* 103, 4234-4239.

Schumann Burkard, G., Jutzi, P. and Roditi, I. (2011). Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters. *Mol. Biochem. Parasitol.* 175, 91-94.

Segel, I.H. (1993). *Enzyme Kinetics*. New York: John Wiley & Sons, Inc.

Shahi, S.K., Krauth-Siegel, R.L. and Clayton, C.E. (2002). Overexpression of the putative thiol conjugate transporter TbMRPA causes melarsoprol resistance in *Trypanosoma brucei*. *Mol. Microbiol.* 43, 1129-1138.

Sienkiewicz, N., Jaroslowski, S., Wyllie, S. and Fairlamb, A.H. (2008). Chemical and genetic validation of dihydrofolate reductase-thymidylate synthase as a drug target in African trypanosomes. *Mol. Microbiol.* 69, 520-533.

Simarro, P.P., Cecchi, G., Paone, M., Franco, J.R., Diarra, A., Ruiz, J.A., Fevre, E.M., Courtin, F., Mattioli, R.C. and Jannin, J.G. (2010). The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. *Int J Health Geogr* 9, 57.

Simarro, P.P., Diarra, A., Ruiz Postigo, J.A., Franco, J.R. and Jannin, J.G. (2011). The human African trypanosomiasis control and surveillance programme of the World Health Organization 2000-2009: the way forward. *PLoS neglected tropical diseases* 5, e1007.

Simarro, P.P., Franco, J., Diarra, A., Postigo, J.A. and Jannin, J. (2012). Update on field use of the available drugs for the chemotherapy of human African trypanosomiasis. *Parasitology* 139, 842-846.

Smith, V.C., Cleghorn, L.A., Woodland, A., Spinks, D., Hallyburton, I., Collie, I.T., Mok, N.Y., Norval, S., Brenk, R., Fairlamb, A.H., Frearson, J.A., Read, K.D., Gilbert, I.H. and Wyatt, P.G. (2011). Optimisation of the anti-*Trypanosoma brucei* activity of the opioid agonist U50488. *ChemMedChem* 6, 1832-1840.

Sokolova, A.Y., Wyllie, S., Patterson, S., Oza, S.L., Read, K.D. and Fairlamb, A.H. (2010). Cross-resistance to nitro drugs and implications for treatment of human African trypanosomiasis. *Antimicrob. Agents Chemother.* 54, 2893-2900.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98, 503-517.

Spinks, D., Ong, H.B., Mpamhanga, C.P., Shanks, E.J., Robinson, D.A., Collie, I.T., Read, K.D., Frearson, J.A., Wyatt, P.G., Brenk, R., Fairlamb, A.H. and Gilbert, I.H. (2011). Design, synthesis and biological evaluation of novel

inhibitors of *Trypanosoma brucei* pteridine reductase 1. *ChemMedChem* 6, 302-308.

Stuart, K., Brun, R., Croft, S., Fairlamb, A., Gurtler, R.E., McKerrow, J., Reed, S. and Tarleton, R. (2008). Kinetoplastids: related protozoan pathogens, different diseases. *The Journal of clinical investigation* 118, 1301-1310.

Sun, L. and Wang, C.C. (2011). The structural basis of localizing polo-like kinase to the flagellum attachment zone in *Trypanosoma brucei*. *PloS one* 6, e27303.

Takahashi-Yanaga, F. (2013). Activator or inhibitor? GSK-3 as a new drug target. *Biochem. Pharmacol.* 86, 191-199.

Takahashi-Yanaga, F., Mori, J., Matsuzaki, E., Watanabe, Y., Hirata, M., Miwa, Y., Morimoto, S. and Sasaguri, T. (2006). Involvement of GSK-3 $\beta$  and DYRK1B in differentiation-inducing factor-3-induced phosphorylation of cyclin D1 in HeLa cells. *J. Biol. Chem.* 281, 38489-38497.

Takahashi-Yanaga, F., Taba, Y., Miwa, Y., Kubohara, Y., Watanabe, Y., Hirata, M., Morimoto, S. and Sasaguri, T. (2003). Dictyostelium differentiation-inducing factor-3 activates glycogen synthase kinase-3 $\beta$  and degrades cyclin D1 in mammalian cells. *J. Biol. Chem.* 278, 9663-9670.

Taylor, S.S. and Kornev, A.P. (2011). Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem. Sci.* 36, 65-77.

Tewari, R., Straschil, U., Bateman, A., Bohme, U., Cherevach, I., Gong, P., Pain, A. and Billker, O. (2010). The systematic functional analysis of *Plasmodium* protein kinases identifies essential regulators of mosquito transmission. *Cell Host Microbe* 8, 377-387.

Tolosa, E., Litvan, I., Hoglinger, G.U., Burn, D., Lees, A., Andres, M.V., Gomez-Carrillo, B., Leon, T. and Del Ser, T. (2014). A phase 2 trial of the GSK-3 inhibitor tideglusib in progressive supranuclear palsy. *Movement disorders* 29, 470-478.

Torreele, E., Bourdin Trunz, B., Tweats, D., Kaiser, M., Brun, R., Mazue, G., Bray, M.A. and Pecoul, B. (2010). Fexinidazole - a new oral nitroimidazole drug candidate entering clinical development for the treatment of sleeping sickness. *PLoS Negl. Trop. Dis.* 4, e923.

Torrie, L.S., Wyllie, S., Spinks, D., Oza, S.L., Thompson, S., Harrison, J.R., Gilbert, I.H., Wyatt, P.G., Fairlamb, A.H. and Frearson, J.A. (2009). Chemical validation of trypanothione synthetase: a potential drug target for human trypanosomiasis. *J. Biol. Chem.* 284, 36137-36145.

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U. S. A.* 76, 4350-4354.

- Tu, X., Kumar, P., Li, Z. and Wang, C.C. (2006). An aurora kinase homologue is involved in regulating both mitosis and cytokinesis in *Trypanosoma brucei*. *J. Biol. Chem.* 281, 9677-9687.
- Tu, X., Mancuso, J., Cande, W.Z. and Wang, C.C. (2005). Distinct cytoskeletal modulation and regulation of G1-S transition in the two life stages of *Trypanosoma brucei*. *J. Cell Sci.* 118, 4353-4364.
- 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. *J. Biol. Chem.* 279, 20519-20528.
- Tu, X. and Wang, C.C. (2005). Pairwise knockdowns of cdc2-related kinases (CRKs) in *Trypanosoma brucei* identified the CRKs for G1/S and G2/M transitions and demonstrated distinctive cytokinetic regulations between two developmental stages of the organism. *Eukaryot. Cell* 4, 755-764.
- Urbaniak, M.D. (2009). Casein kinase 1 isoform 2 is essential for bloodstream form *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 166, 183-185.
- Urbaniak, M.D., Mathieson, T., Bantscheff, M., Eberhard, D., Grimaldi, R., Miranda-Saavedra, D., Wyatt, P., Ferguson, M.A., Frearson, J. and Drewes, G. (2012a). Chemical proteomic analysis reveals the drugability of the kinome of *Trypanosoma brucei*. *ACS Chem Biol* 7, 1858-1865.
- Urbaniak, M.D., Guther, M.L. and Ferguson, M.A. (2012b). Comparative SILAC proteomic analysis of *Trypanosoma brucei* bloodstream and procyclic lifecycle stages. *PLoS One* 7, e36619.
- Urbaniak, M.D., Martin, D.M. and Ferguson, M.A. (2013). Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of *Trypanosoma brucei*. *J. Proteome Res.* 12, 2233-2244.
- Urich, R., Grimaldi, R., Luksch, T., Frearson, J.A., Brenk, R. and Wyatt, P.G. (2014). The design and synthesis of potent and selective inhibitors of *Trypanosoma brucei* Glycogen Synthase Kinase 3 for the treatment of human African trypanosomiasis. *J. Med. Chem.* 57, 7536-7549.
- Vassella, E., Kramer, 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*. *Mol. Microbiol.* 41, 33-46.
- Vaughan, S. and Gull, K. (2003). The trypanosome flagellum. *J. Cell Sci.* 116, 757-759.
- Vickerman, K. (1985). Developmental cycles and biology of pathogenic trypanosomes. *Br. Med. Bull.* 41, 105-114.
- Vincent, I.M., Creek, D., Watson, D.G., Kamleh, M.A., Woods, D.J., Wong, P.E., Burchmore, R.J. and Barrett, M.P. (2010). A molecular mechanism for eflornithine resistance in African trypanosomes. *PLoS Pathog* 6, e1001204.

Voogd, T.E., Vansterkenburg, E.L., Wilting, J. and Janssen, L.H. (1993). Recent research on the biological activity of suramin. *Pharmacol. Rev.* 45, 177-203.

Wang, X., Paulin, F.E.M., Campbell, L.E., Gomez, E., O'Brien, K., Morrice, N. and Proud, C.G. (2001). Eukaryotic initiation factor 2B: identification of multiple phosphorylation sites in the  $\epsilon$ -subunit and their functions *in vivo*. *EMBO J.* 20, 4349-4359.

Wei, Y. and Li, Z. (2014). Distinct roles of a mitogen-activated protein kinase in cytokinesis between different life cycle forms of *Trypanosoma brucei*. *Eukaryot. Cell* 13, 110-118.

White, F.M. (2007). On the iTRAQ of kinase inhibitors. *Nat. Biotechnol.* 25, 994-996.

WHO (2010). Working to overcome the global impact of neglected tropical diseases: first WHO report on neglected tropical diseases. Geneva: World Health Organization.

Wilkinson, S.R., Taylor, M.C., Horn, D., Kelly, J.M. and Cheeseman, I. (2008). A mechanism for cross-resistance to nifurtimox and benznidazole in trypanosomes. *Proc. Natl. Acad. Sci. USA* 105, 5022-5027.

Wilson, N.F. and Lefebvre, P.A. (2004). Regulation of flagellar assembly by glycogen synthase kinase 3 in *Chlamydomonas reinhardtii*. *Eukaryot. Cell* 3, 1307-1319.

Winkelmann, E. and Raether, W. (1978). Chemotherapeutically active nitro compounds. 4. 5-Nitroimidazoles (Part III). *Arzneimittelforschung* 28, 739-749.

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*. *Mol. Biochem. Parasitol.* 99, 89-101.

Wood, E.R., Truesdale, A.T., McDonald, O.B., Yuan, D., Hassell, A., Dickerson, S.H., Ellis, B., Pennisi, C., Horne, E., Lackey, K., Alligood, K.J., Rusnak, D.W., Gilmer, T.M. and Shewchuk, L. (2004). A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. *Cancer Res.* 64, 6652-6659.

Woodland, A., Grimaldi, R., Luksch, T., Cleghorn, L.A., Ojo, K.K., Van Voorhis, 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, 1127-1137.

Workman, P. and Collins, I. (2010). Probing the probes: fitness factors for small molecule tools. *Chem Biol* 17, 561-577.

- Wyatt, P.G., Gilbert, I.H., Read, K.D. and Fairlamb, A.H. (2011). Target validation: linking target and chemical properties to desired product profile. *Curr. Top. Med. Chem.* 11, 1275-1283.
- Wyatt, P.G., Woodhead, A.J., Berdini, V., Boulstridge, J.A., Carr, M.G., Cross, D.M., Davis, D.J., Devine, L.A., Early, T.R., Feltell, R.E., Lewis, E.J., McMenamin, R.L., Navarro, E.F., O'Brien, M.A., O'Reilly, M., Reule, M., Saxty, G., Seavers, L.C., Smith, D.M., Squires, M.S., Trewartha, G., Walker, M.T. and Woolford, A.J. (2008). Identification of N-(4-piperidiny)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (AT7519), a novel cyclin dependent kinase inhibitor using fragment-based X-ray crystallography and structure based drug design. *J. Med. Chem.* 51, 4986-4999.
- Wyllie, S., Oza, S.L., Patterson, S., Spinks, D., Thompson, S. and Fairlamb, A.H. (2009). Dissecting the essentiality of the bifunctional trypanothione synthetase-amidase in *Trypanosoma brucei* using chemical and genetic methods. *Mol. Microbiol.* 74, 529-540.
- 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. *Int. J. Parasitol.* 39, 1289-1303.
- Xu, C., Kim, N.G. and Gumbiner, B.M. (2009). Regulation of protein stability by GSK3 mediated phosphorylation. *Cell Cycle* 8, 4032-4039.
- Yang, J., Copeland, R.A. and Lai, Z. (2009). Defining balanced conditions for inhibitor screening assays that target bisubstrate enzymes. *J. Biomol. Screen.* 14, 111-120.
- Yarlett, N. and Bacchi, C.J. (1988). Effect of DL-alpha-difluoromethylornithine on methionine cycle intermediates in *Trypanosoma brucei brucei*. *Mol. Biochem. Parasitol.* 27, 1-10.
- Yu, Z., Liu, Y. and Li, Z. (2012). Structure-function relationship of the Polo-like kinase in *Trypanosoma brucei*. *J. Cell Sci.* 125, 1519-1530.
- Zhang, J.H., Chung, T.D. and Oldenburg, K.R. (1999). A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* 4, 67-73.
- Zhang, Y., Li, Z., Pilch, D.S. and Leibowitz, M.J. (2002). Pentamidine inhibits catalytic activity of group I intron Ca.LSU by altering RNA folding. *Nucleic Acids Res.* 30, 2961-2971.



## Appendix

Urbaniak, M.D., Mathieson, T., Bantscheff, M., Eberhard, D., Grimaldi, R., Miranda-Saavedra, D., Wyatt, P., Ferguson, M.A., Frearson, J. and Drewes, G. (2012). Chemical proteomic analysis reveals the drugability of the kinome of *Trypanosoma brucei*. *ACS Chem. Biol.* 7, 1858-1865.

Urich, R., Grimaldi, R., Luksch, T., Frearson, J.A., Brenk, R. and Wyatt, P.G. (2014). The design and synthesis of potent and selective inhibitors of *Trypanosoma brucei* Glycogen Synthase Kinase 3 for the treatment of human African trypanosomiasis. *J. Med. Chem.* 57, 7536-7549.

Woodland, A., Grimaldi, R., Luksch, T., Cleghorn, L.A., Ojo, K.K., Van Voorhis, 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, 1127-1137.

# Chemical Proteomic Analysis Reveals the Drugability of the Kinome of *Trypanosoma brucei*

Michael D. Urbaniak,<sup>\*,†</sup> Toby Mathieson,<sup>‡</sup> Marcus Bantscheff,<sup>‡</sup> Dirk Eberhard,<sup>‡</sup> Raffaella Grimaldi,<sup>†</sup> Diego Miranda-Saavedra,<sup>§</sup> Paul Wyatt,<sup>†</sup> Michael A. J. Ferguson,<sup>†</sup> Julie Frearson,<sup>||</sup> and Gerard Drewes<sup>\*,‡</sup>

<sup>†</sup>Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, U.K.

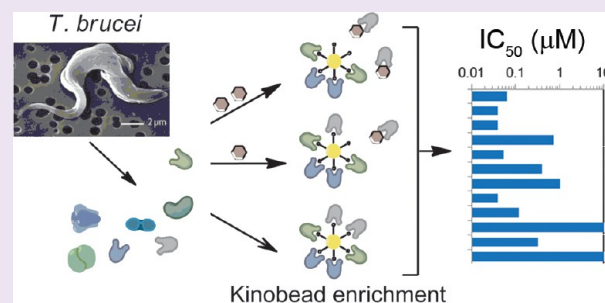
<sup>‡</sup>Cellzome AG, Meyerhofstrasse 1, D-69117 Heidelberg, Germany

<sup>§</sup>World Premier International Immunology Frontier Research Centre, Osaka University, 3-1 Yamadaoka, Suita 565-0871, Osaka, Japan

<sup>||</sup>BioFocus, Chesterford Park, Saffron Walden, Essex CB10 1XL, U.K.

## S Supporting Information

**ABSTRACT:** The protozoan parasite *Trypanosoma brucei* is the causative agent of African sleeping sickness, and there is an urgent unmet need for improved treatments. Parasite protein kinases are attractive drug targets, provided that the host and parasite kinomes are sufficiently divergent to allow specific inhibition to be achieved. Current drug discovery efforts are hampered by the fact that comprehensive assay panels for parasite targets have not yet been developed. Here, we employ a kinase-focused chemoproteomics strategy that enables the simultaneous profiling of kinase inhibitor potencies against more than 50 endogenously expressed *T. brucei* kinases in parasite cell extracts. The data reveal that *T. brucei* kinases are sensitive to typical kinase inhibitors with nanomolar potency and demonstrate the potential for the development of species-specific inhibitors.



The protozoan parasite *Trypanosoma brucei* is transmitted by the bite of an infected Tsetse fly and causes African sleeping sickness, which is also known as Human African Trypanosomiasis (HAT). The disease is invariably fatal if left untreated and results in upward of 10,000 deaths each year in sub-Saharan Africa.<sup>1</sup> *T. brucei* has a complex digenetic lifecycle between the insect vector and mammalian host, and the ability to adapt to these environments is essential to its survival and virulence. During early stages of infection the clinically relevant bloodstream form of the parasite proliferates in the blood and lymph of the human host and then in the second stage enters the cerebrospinal fluid and brain, resulting in coma and death. Current treatments are expensive, toxic, and difficult to administer, leaving an urgent unmet need for improved therapeutic agents.<sup>2</sup>

Protein kinases play key roles in the control of growth and cell signaling and are a major target of the pharmaceutical industry. Parasite protein kinases have been proposed as attractive targets for drug discovery as such efforts can “piggy-back” on the extensive knowledge of the development of inhibitors against human protein kinases.<sup>3</sup> In the case of *T. brucei*, bioinformatic analysis of the genome has identified 176–182 putative protein kinases on the basis of sequence similarity, the majority of which can be placed within well-recognized kinase groups (Supplementary Table S1).<sup>4,5</sup> Efforts to

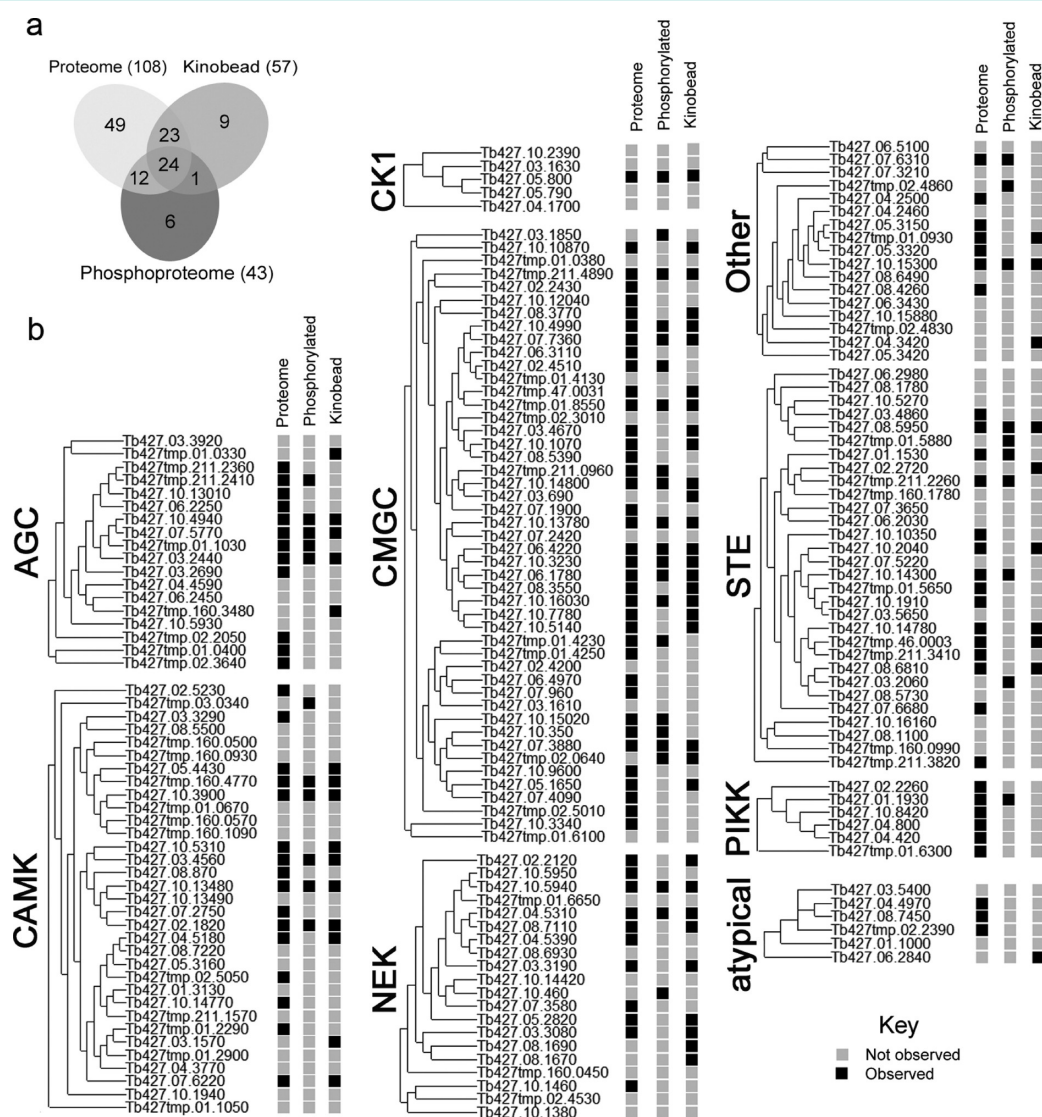
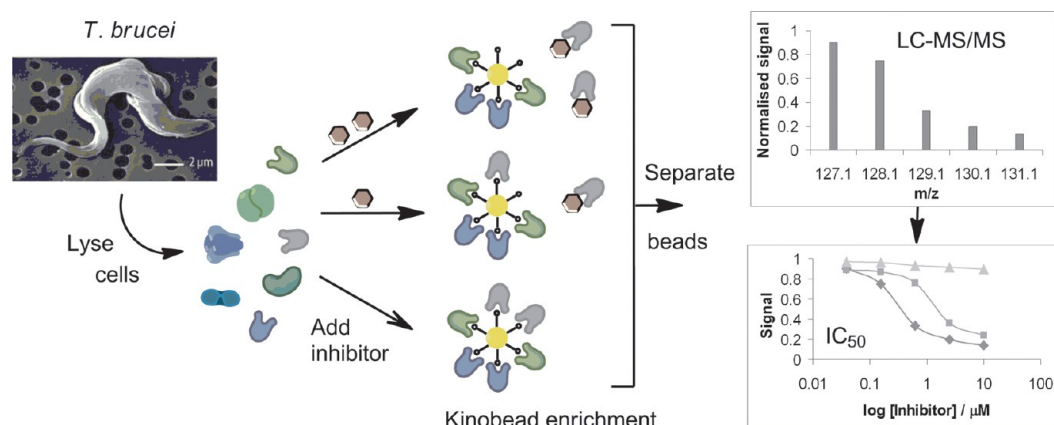
determine the detailed biological role of *T. brucei* protein kinases are ongoing, although knock-down by RNA interference has provided evidence of the essential nature of a significant number of protein kinases.<sup>6</sup> However, the rationale to develop drugs to target the *T. brucei* kinome poses a conundrum: if mammalian and parasite protein kinases are sufficiently similar to be identified and classified on the basis of sequence similarity and are inhibited by typical inhibitors, will parasite kinase inhibitors lack host–parasite specificity? Conversely, if the kinases are sufficiently different that host–parasite specificity can be readily obtained, will they be inhibited by typical inhibitors of mammalian kinases? In other words, we need to consider the similarity of the chemical space that parasite and mammalian protein kinase inhibitors occupy, rather than the similarity in protein kinase sequence.

One way to probe the inhibitor chemical space is to profile inhibitor activity against both the mammalian and parasite kinomes. Such profiling is often achieved using *in vitro* activity assays against a panel of recombinant protein kinases,<sup>7</sup> but there is no such panel available for the *T. brucei* kinome; indeed only a handful of active *T. brucei* kinases have been

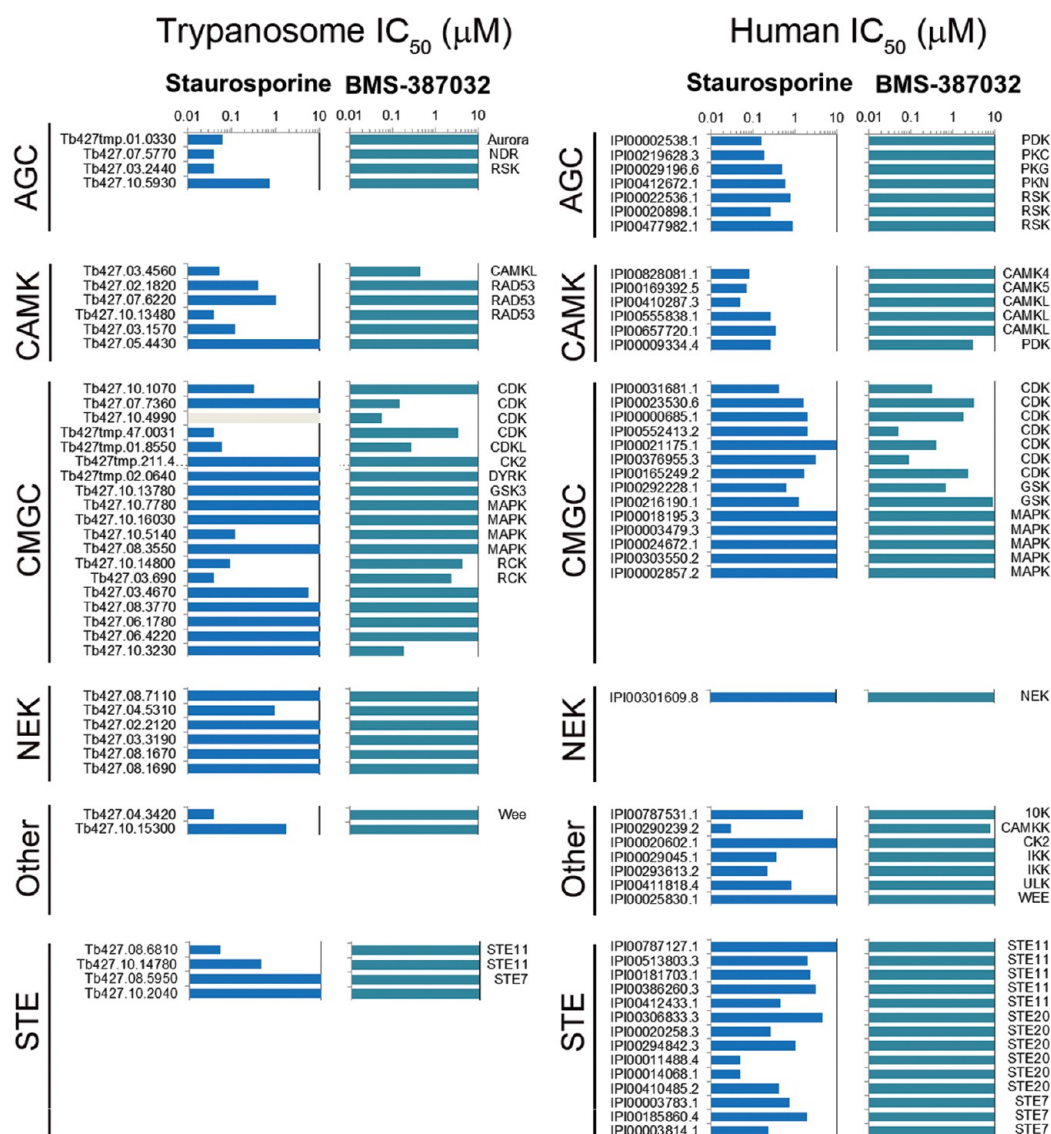
Received: July 3, 2012

Accepted: August 21, 2012

Published: August 21, 2012



**Figure 2.** Profiling the kinome expressed in bloodstream form *T. brucei* using complementary mass spectrometry-based observations. (a) Venn diagram summarizing overlapping protein kinase observations. (b) Details of protein kinases observed (black square), divided by kinase group classification according to the similarity of their catalytic domains. Proteome: detection at natural abundance. Phosphorylated: enrichment of phosphorylated peptides.<sup>5</sup> Kinobeads: enriched by immobilized mixed kinase-inhibitors.



**Figure 3.** Chemical proteomics profiling of Staurosporine and BMS-387032 against the trypanosome and human kinome. Horizontal bars represent  $IC_{50}$  value calculated from the isobaric reporter signals, gray bars indicate where binding was not quantified. Lysates were incubated with varying concentrations of compounds prior to incubation with kinobeads, and the bound fraction was quantified by tandem mass spectrometry. Data for TK, TKL, atypical, and lipid kinase are not shown for clarity; the full data for the trypanosome profile can be found in Supplementary Table S5, and the CLL cell data has been reported previously.<sup>17</sup>

recombinantly expressed as active enzymes.<sup>8–10</sup> A recent advance in kinase inhibitor profiling uses a chemical proteomic methodology that captures a substantial portion of the expressed kinome (and related proteins) contained in cell lysates on a mixed kinase-inhibitor matrix known as kinobeads.<sup>11,12</sup> Addition of a kinase inhibitor to the cell lysates enables it to bind to its specific target(s), occupying the binding sites and preventing binding to the kinobeads, whereas the binding of nontargeted kinases and other proteins are unaffected. Incubation of the lysate with varying concentrations of the inhibitor and subsequent analysis of the kinobead-bound subproteome by quantitative mass spectrometry allows inhibition curves to be generated for each protein observed (Figure 1). We reasoned that this methodology should be species independent, provided that the kinobeads are sufficiently promiscuous to capture a sizable portion of the parasite kinome.

Here, we present the results of our efforts to establish kinobead chemoproteomics profiling in *T. brucei* and estimate the coverage of the parasite kinome. Our strategy enabled us to access more than 50 parasite kinases for inhibitor profiling, which by far exceeds any currently available enzyme panels. We report the profile of the mammalian kinase inhibitors staurosporine and BMS-387032 and several early hit compounds identified as parasite protein kinase inhibitors by the Drug Discovery Unit at the University of Dundee.

## RESULTS AND DISCUSSION

**Coverage of the *T. brucei* Kinome.** Kinobeads consist of immobilized analogues of a variety of ATP competitive kinase inhibitors that show relatively promiscuous binding to the mammalian kinases, but their ability to bind kinases from more divergent organisms has not been examined. To establish to what extent *T. brucei* protein kinases were able to bind to the kinobeads, we examined the subproteome enriched from cell

lysates of the clinically relevant bloodstream form of *T. brucei*. Initially, we compared four versions of kinobeads that differ in the identity of the immobilized inhibitors (see Methods). In these experiments, we observed the enrichment of a total of a 57 protein kinases (Supplementary Table S2). The enrichment of a significant number of trypanosome protein kinases by these promiscuous mammalian kinase inhibitors suggests that the ATP binding pocket architecture is broadly conserved between the two species.

Estimating the portion of the trypanosome kinome captured is not straightforward, as the bloodstream form represents just one of the multiple lifecycle stages of *T. brucei* and may not express every kinase encoded by the genome. Before we could determine the proportion of the bloodstream form kinome that was captured by kinobeads, we needed to estimate how many protein kinases were present in the cell lysates. To achieve this, we took advantage of the differing bias of two orthogonal proteomic techniques. By analyzing the total proteome contained in the non-enriched cell lysate we identified the most abundant 3248 proteins, which included 90 protein kinases (Supplementary Table S3). This data shows good overlap with a recent comparative SILAC proteomic study of the bloodstream and procyclic form *T. brucei*, which identified 65 protein kinases,<sup>13</sup> including 18 not seen in this study. A separate phosphoproteomic study by Nett *et al.*,<sup>5</sup> which used strong cation exchange and TiO<sub>2</sub> chromatography to enrich for phosphopeptides, identified 43 phosphorylated protein kinases in bloodstream form *T. brucei*. Comparison of the protein kinases identified by these three orthogonal proteomics techniques revealed overlapping and complementary coverage of the bloodstream form kinome (Figure 2 and Supplementary Table S4), providing experimental observation of a total of 124 protein kinases out of the predicted 182 (68%). The kinases that bind to kinobeads are not significantly biased toward abundant (Supplementary Figure S1) or phosphorylated protein kinases or any particular kinase group. Mapping the kinases that bind to kinobeads onto the phylogeny of the *T. brucei* kinome shows that kinobead enrichment appears to be independent of the degree of sequence homology (Figure 2). While it is likely that the coverage of this observable bloodstream form kinome is not complete, it is in reasonable agreement with transcriptome studies that suggest 25% of the genome is differentially expressed between bloodstream and procyclic form *T. brucei* cells.<sup>14</sup> The kinobead-enriched subproteome contains 46% of the observed bloodstream form kinome (31% of the predicted genome), comparable to the coverage obtained from analysis of human cell lysates (52% of the predicted genome).

**Profiling of Known Kinase Inhibitors.** Kinobead-based profiling enables access to a sizable fraction of the expressed trypanosome kinome, which can be used to determine the potency and selectivity of kinase inhibitors in cell extracts by means of a multiplexed competition binding assay. The kinobeads version producing the best coverage of the *T. brucei* kinome was used to determine the kinase inhibition profile of two well-studied kinase inhibitors: Staurosporine and BMS-387032. The binding of these inhibitors to their cellular targets was quantified by mass spectrometry using isobaric tags for relative and absolute quantification (iTRAQ)<sup>15,16</sup> for 51 protein kinases and 67 other kinobead-binding proteins (Figure 3 and Supplementary Table S5). Staurosporine, a natural product, is a prototypical ATP-competitive pan-kinase inhibitor that binds to many protein kinases with high affinity and little selectivity.<sup>11</sup> In

a previous study we reported kinobead profiling of staurosporine in primary chronic lymphocytic leukemia cells<sup>17</sup> and demonstrated that more than a third of the observed human kinome (41/112) displayed submicromolar IC<sub>50</sub> values (Figure 3). This pan-kinase activity was retained against the *T. brucei* kinome, with more than a third of the observed trypanosome kinome (18/44) displaying IC<sub>50</sub> values <1 μM, including 10 with IC<sub>50</sub> values below 100 nM (Figure 3).

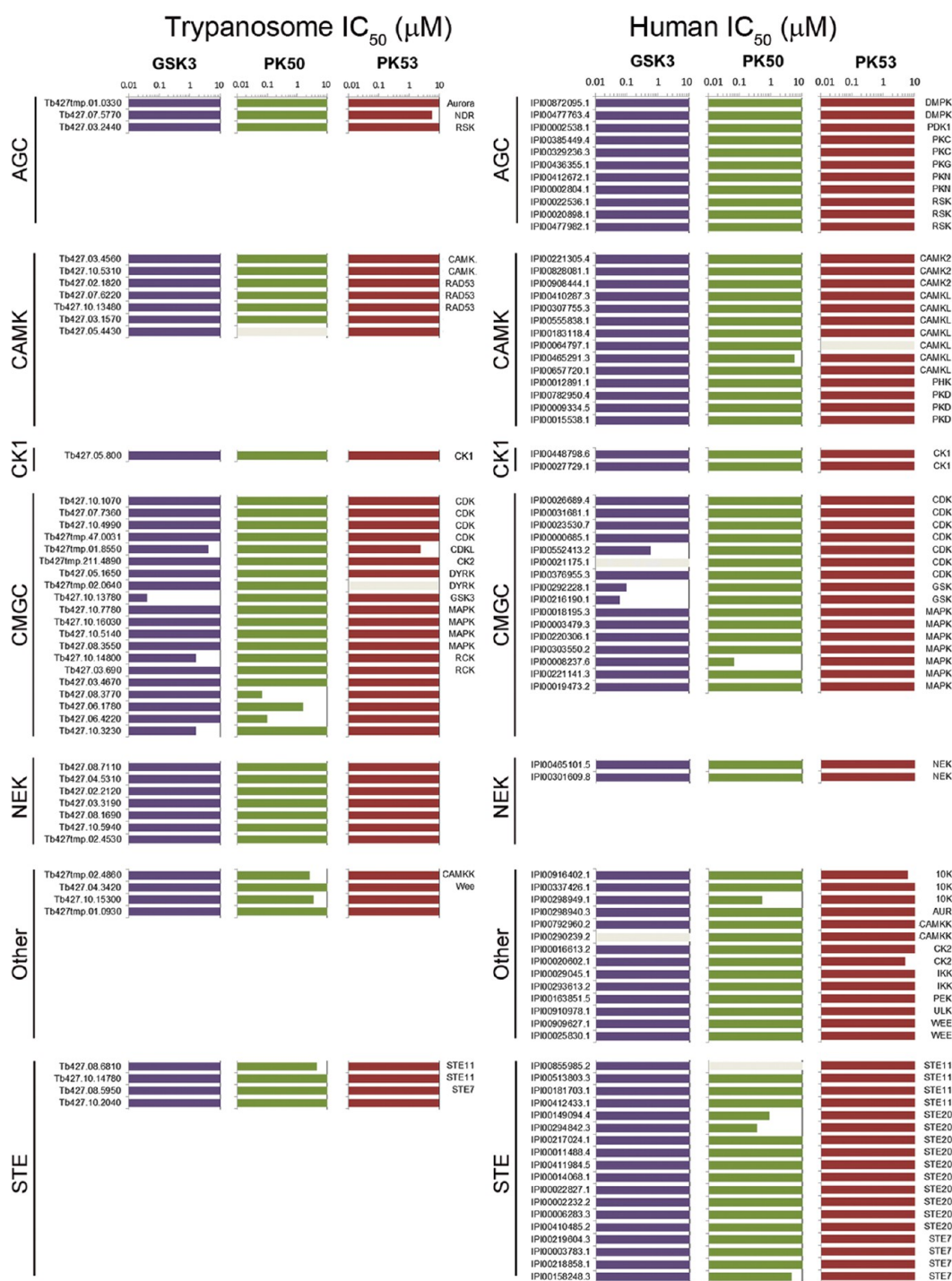
The second kinase inhibitor to be profiled, BMS-387032, an established pan-cyclin-dependent kinase (CDK) inhibitor, was selected because *T. brucei* have a relatively expanded CMGC group including many putative CDK or CDK-like (CDKL) family members.<sup>4,5</sup> Kinobead profiling of BMS-387032 against primary chronic lymphocytic leukemia cells showed that all of the seven observed CDKs were inhibited, with CDK2, CDK9, CRK7, and PCSTAIRE2 displaying submicromolar IC<sub>50</sub> values (Figure 3).<sup>17</sup> BMS-387032 retained the ability to inhibit the majority of the observed trypanosome CDKs, including targeting the CDK2-related kinases CRK2 and CRK3 with submicromolar potency (IC<sub>50</sub> of 148 and 57 nM, respectively), although no inhibition of CRK1 was observed (Figure 3). In addition, the compound selectivity was slightly broader than just the CDKs, with two additional kinases (CMGC and CAMK) also inhibited with IC<sub>50</sub> values below 500 nM, which may reflect the divergence of the trypanosome kinome. These data provide the first molecular evidence that the trypanosome kinome is sensitive to typical mammalian kinase inhibitors with nanomolar potencies and suggest that other standard kinase inhibitor scaffolds may retain substantial activity.

**Profiling Trypanosome Kinase Inhibitors.** The Drug Discovery Unit at the University of Dundee has conducted a number of screening campaigns to help identify preclinical candidates for the treatment of African sleeping sickness, including target-based screens against the *T. brucei* protein kinases Glycogen Synthase Kinase 3 (GSK3, Tb427.10.13780) and the Nuclear DBF-2-related (NDR) kinases PK50 (Tb427.10.4940) and PK53 (Tb427.07.5770).<sup>8</sup> Details of the compound screening and optimization will be reported elsewhere. To further probe the chemical space that mammalian and trypanosome kinase inhibitors occupy and demonstrate the utility of chemical proteomics in antiparasitic drug discovery, we profiled hits selected from these screens (Supplementary Tables S6 and S7). The compounds selected all show nanomolar potency against their respective molecular target *in vitro* and variable efficacy against cultured trypanosome and human hepatocyte (MRC5) cells (Table 1).

**Table 1. Trypanosome Kinase Inhibitors**

	GSK3	PK50	PK53
compound ID	DDD85893	DDD34425	DDD88213
<i>T. brucei</i> enzyme IC <sub>50</sub> (μM) <sup>a</sup>	<0.002	0.013 ± 0.006	0.73 ± 0.14
<i>T. brucei</i> kinobead IC <sub>50</sub> (μM) <sup>b</sup>	<0.039	not observed	5.7
<i>T. brucei</i> EC <sub>50</sub> (μM) <sup>c</sup>	1.3 ± 1.2	0.86 ± 0.52	45 ± 4
<i>H. sapiens</i> MRC5 EC <sub>50</sub> (μM) <sup>c</sup>	28 ± 9.7	>50	>50
<i>H. sapiens</i> kinobead IC <sub>50</sub> (μM) <sup>b</sup>	<0.06	not observed	not observed

<sup>a</sup>Enzyme IC<sub>50</sub> values were determined using a plate-based radiometric assay. <sup>b</sup>Kinobead IC<sub>50</sub> is described in Methods. <sup>c</sup>Cell EC<sub>50</sub> values were determined using an Alamar blue assay.<sup>34</sup>



**Figure 4.** Chemical proteomics profiling of trypanosome kinase inhibitors against the kinome of bloodstream form *T. brucei*. Horizontal bars represent  $IC_{50}$  value calculated from the isobaric reporter signals; gray bars indicate where binding was not quantified. Lysates were incubated with varying concentrations of compounds prior to incubation with kinobeads, and the bound fraction was quantified by tandem mass spectrometry. Data for TK, TKL, atypical, and lipid kinase are not shown for clarity; the full data can be found in Supplementary Tables S6 and S7.

The compound DDD85893 was identified as a potent inhibitor of *T. brucei* GSK3 (TbGSK3) *in vitro*, with good efficacy against cultured *T. brucei* and good selectivity against cultured human cells. The kinobead profiling of DDD85893 against *T. brucei* cell lysates confirmed that TbGSK3 was inhibited with nanomolar potency, with three other CMGC kinases inhibited at micromolar level (Figure 4). The compound also showed a very clean profile against human MRC5 cell lysates, with only human GSK3 $\alpha$ , GSK3 $\beta$ , and

CDK9 inhibited with nanomolar potency. These data show that the compound DDD85893 has excellent selectivity for GSK3 and limited other CMGC members but does not display any species specificity.

The second compound to be profiled (DDD34425) was identified as a potent inhibitor of *T. brucei* PK50 *in vitro*, with good efficacy against cultured *T. brucei* and good selectivity against cultured human cells. Unfortunately, the expected target PK50 was not among the kinases that bound to the kinobeads,

suggesting that the ATP binding site contains features that are not recognized by the set of standard ligands immobilized on the beads. Indeed, none of the four human NDR kinases (NDR1, NDR2, LATS1, LATS2) were among the kinases in MRC5 lysates that bound to the kinobeads. However, the kinobead profile of DDD34425 against *T. brucei* cell lysates revealed that MAPK10 and MAPK5 were inhibited with nanomolar potency, with three other kinases inhibited at micromolar level (Figure 4). As *T. brucei* bloodstream form MAPK5 null mutants grow normally *in vitro*<sup>18</sup> and little is known about MAPK10,<sup>19</sup> it is unclear what effect, if any, their inhibition may contribute to the observed trypanocidal effect of DDD34425. The kinobead profile against MRC5 cell lysates showed a broader specificity, with the MAPK NKL and TKLs RIPK2 and ALK2 inhibited with nanomolar potency. Six additional kinases from four different groups were inhibited at the micromolar level. These data revealed that the compound DDD34425 lacks specificity against the observable trypanosome kinome and significant poly pharmacology against the observable human kinome, including a number of tyrosine-specific kinases that are absent from *T. brucei*.

The compound DDD88213 was identified as a submicromolar inhibitor of *T. brucei* PK53 *in vitro* but lacked efficacy against both cultured *T. brucei* and human cells. The kinobeads profile of DDD88213 against *T. brucei* cell lysates revealed that PK53 was inhibited with an  $IC_{50}$  of 5.7  $\mu$ M, with CK2 $\alpha$ 2 also inhibited with similar potency (Figure 4). The 10-fold drop in potency between the enzyme  $IC_{50}$  and the kinobeads binding assay is not significantly different from that seen with other inhibitors and may reflect differences between purified proteins and cell extracts. Moreover, the potency determined with kinobeads is in line with the weak activity of the compound in the Alamar blue assay. The kinobead profile against MRC5 cell lysates showed that none of the observable human kinome was significantly inhibited by DDD88213. These data show that while this compound does appear to be specifically targeting the desired trypanosome kinase and appears to have little effect on the human kinome, it lacks sufficient potency to achieve the desired trypanocidal effect and would require further optimization.

The final compound to be profiled was recently identified as a potent inhibitor of the *Leishmania major* cyclin dependent kinase 2-related kinase 3 (LmCRK3) (compound 33, Cleghorn *et al.*<sup>20</sup>) but lacked efficacy against cultured *L. major* or the related parasite *T. brucei*.<sup>20</sup> As active recombinant *T. brucei* CRK3 (TbCRK3) is not available for *in vitro* screening, we attempted to use kinobead profiling to investigate whether the lack of cellular potency against *T. brucei* was due to lack of potency against TbCRK3. The kinobead profile of the compound revealed that none of the observed *T. brucei* kinases were significantly inhibited, including TbCRK3, and neither were any mammalian kinase (Supplementary Tables S6 and S7). The complete lack of inhibition suggests that the native state of parasite CRK3s is distinct from the recombinant form used for *in vitro* screening. While these differences may be due to the absence of the associated cyclin, a recent screen against LmCRK3-cyclin 6 also resulted in compounds that lacked efficacy against cultured *L. major*.<sup>21</sup>

**Summary and Conclusions.** We set out to examine the similarity of the chemical space that parasite and mammalian protein kinase inhibitors occupy using a recently developed chemical proteomics approach to profile kinase inhibitors. The data presented here represent the first molecular evidence that

typical ATP-competitive inhibitors can retain low nanomolar potency against *T. brucei* protein kinases. The inhibition profile of the compounds does not seem to map directly between the two species, suggesting that it may be possible to exploit these differences to obtain host/parasite specificity. It is also possible that the differences observed are an artifact due to the limited coverage of the kinome achieved by kinobead enrichment. However, the lack of bias in the enriched kinome suggested that this is unlikely.

Our data suggest that phenotypic screening of known kinase inhibitors against *T. brucei* is likely to identify potent compounds but also show that it is inappropriate to infer the molecular target on the basis of the inhibition profile established in mammalian systems. The chemoproteomics approach presented to profile potential kinase inhibitors simultaneously covers as much as half the observed bloodstream form kinome, representing a 10-fold increase in the selection of active parasite kinases currently available for drug discovery. Development of kinobeads tailored to the trypanosome kinome, for instance, by immobilizing novel inhibitors identified through phenotypic screening, is an attractive approach to extend the coverage of the parasite kinome. Importantly, the approach presented here is species-independent and can be applied to any clinically relevant pathogen for which a genome sequence is available.

## METHODS

**Reagents and Drugs.** All reagents were purchased from Sigma unless otherwise noted. Staurosporine was purchased from IRIS Biotech, and BMS-387032 was custom synthesized by Park Place Research.

**Preparation of Bloodstream Form *T. brucei* Cell Lysates.** Bloodstream form *T. brucei brucei* variant 117 (MITat1.4) was purified from infected rodent blood over DE52 cellulose as described previously.<sup>22</sup> The cells were centrifuged at 800  $\times$  g for 10 min at 4  $^{\circ}$ C and resuspended at  $1 \times 10^9$  cells/mL in ice-cold buffer 1 containing protease and phosphatase inhibitors (0.1 mM TLCK, 1  $\mu$ g/mL Leupeptin, 1  $\mu$ g/mL aprotinin, 1 mM PMSF, 1 mM benzamide, Phosphatase Inhibitor Cocktail II (Roche)), and hypotonic lysis was allowed to proceed for 10 min on ice. An equal volume of ice-cold buffer 2 (100 mM Tris pH 7.5, 10% glycerol, 300 mM NaCl, 50 mM NaF, 3 mM MgCl<sub>2</sub>, 0.2 mM Na<sub>2</sub>VO<sub>4</sub>, 1.6% Igepal-CA630, 2 mM DTT, 0.1 mM TLCK) was added, and the lysate was centrifuged at 145,000  $\times$  g for 1 h at 4  $^{\circ}$ C. The BCA assay (Pierce) was used to determine the total protein content in the supernatant, and the concentration was adjusted to 5 mg/mL. Aliquots were frozen in liquid nitrogen and stored at  $-80$   $^{\circ}$ C prior to use.

**Preparation of *H. sapiens* MRC5 Cell Lysates.** MRC5 cells were grown at 37  $^{\circ}$ C, 5% CO<sub>2</sub> in MEM supplemented with 10% FCS. Cell lysates were prepared washing the cells briefly in PBS, incubating with an equal volume of buffers 1 and 2 for 15 min, and then processing the crude lysate as described above.

**Proteomic Analysis.** The total proteome was determined in duplicate by fractionating 25  $\mu$ g of the bloodstream form *T. brucei* cell lysate by SDS-PAGE and pixilation into 24 bands, followed by in-gel reductive alkylation and tryptic digest. Samples were analyzed by liquid chromatography – tandem mass spectrometry on a Eksigent 1D+ HPLC system coupled to a LTQ-Orbitrap mass spectrometer (Thermo scientific). MS spectra were searched using Mascot (Matrix Science) against a nonredundant, in-house compiled database of *Trypanosoma brucei* 927 and 427 strains obtained from TriTrypDB 3.0<sup>23</sup> with additional protein sequences from SwissProt and RefSeq databases, as well as known contaminant sequences such as keratins and trypsin. To assess the false discovery rate (FDR) “decoy” proteins (reverse of the protein sequence) were added to the database. Protein identifications were accepted as follows: (i) For single spectrum to sequence assignments, we required the assignment to be the best

match and a minimum Mascot score of 37 and a 10× difference of the assignment over the next best assignment. On the basis of these criteria, the decoy search results indicated <1% false discovery rate (FDR). (ii) For multiple spectrum to sequence assignments and using the same parameters, the decoy search results indicate <0.1% FDR. To make our data accessible to the scientific community, we have uploaded the results of this study to TriTrypDB (<http://www.tritrypdb.org>).<sup>23</sup>

**Kinobead Profiling.** Procedures are essentially as described previously.<sup>11,17,24</sup> Kinobeads were prepared by immobilization of ATP-mimetics on sepharose beads, with the four versions differing in the identity of the immobilized kinase ligands, as described in Supplementary Table S8.

For kinobead profiling, compounds were dissolved in DMSO, added at various concentrations (0, 0.039, 0.156, 0.625, 2.5, and 10 μM) to 1-mL cell lysate samples, and incubated for 45 min at 4 °C. Subsequently, kinobeads were added to each sample and incubated for a further 60 min at 4 °C. The kinobeads were collected by centrifugation and washed with lysis buffer containing 0.2% Igepal-CA630, and bead-bound proteins were eluted with NuPAGE LDS buffer (Invitrogen) containing 50 mM DTT for 30 min at 50 °C followed by alkylation with 20 mg/mL iodoacetamide for 30 min. Samples were purified on 4–12% NuPAGE gels, stained with colloidal Coomassie blue, digested with trypsin, and subsequently labeled with TMT isobaric tagging reagents (ThermoFisher Scientific).<sup>15</sup> Tryptic peptides were separated over 4 h using nanoflow reversed-phase chromatography online coupled to an Orbitrap mass spectrometer. Peptide fragmentation was performed using PQD, and peptides were identified with Mascot and quantified as described.<sup>25</sup>

**Identification of the *T. brucei* 427 Strain Kinome.** The annotated proteins from *T. brucei brucei* 427 strain was obtained from TriTrypDB 3.0<sup>23</sup> and scanned through a highly sensitive and specific multilevel HMM library of the protein kinase superfamily,<sup>26</sup> followed by expert curation. Assignment of putative protein kinases to the main ePK and aPK groups was done by using the *E*-value cutoffs specific for each group as described previously.<sup>26,27</sup> This procedure identified and assigned 187 protein kinases (Supplementary Table S1).

**Phylogenetic Analysis.** The phylogeny of the kinase groups identified in *T. brucei* 427 strain was determined using the Phylogeny.fr platform<sup>28</sup> and comprised the following steps: Sequences were aligned with T-Coffee (v6.85) using pairwise alignment methods,<sup>29</sup> and ambiguous regions (*i.e.*, containing gaps and/or poorly aligned) were removed with Gblocks (v0.91b)<sup>30</sup> with low stringency (Min. seq. for flank pos.: 55%, Max. contig. nonconserved pos.: 8, Min. block length: 5, Gaps in final blocks: half). The phylogenetic tree was reconstructed using the maximum likelihood method implemented in the PhyML program (v3.0 aLRT)<sup>31</sup> with reliability for internal branch assessed using the aLRT test (minimum of SH-like and Chi2-based parametric).<sup>32</sup> Graphical representation and editing of the phylogenetic tree was performed with TreeDyn (v198.3).<sup>33</sup>

## ■ ASSOCIATED CONTENT

### ● Supporting Information

This material is free via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [m.d.urbaniak@dundee.ac.uk](mailto:m.d.urbaniak@dundee.ac.uk); [gerard.c.drewes@gsk.com](mailto:gerard.c.drewes@gsk.com).

### Notes

The authors declare the following competing financial interest(s): The authors T.M., D.E., M.B., and G.D. are employees of Cellzome AG, which contributed to the funding of this work by payment-in-kind.

## ■ ACKNOWLEDGMENTS

This work was funded by the Wellcome Trust (Grants 085622 and 077705, and Strategic award 083481). We thank F. Simeons, L. Stojanovski, and K. Read of the University of Dundee Drug Discovery Unit for assistance in the culture of *T. brucei* in rodents and M. Boesche (Cellzome) for performing mass spectrometry.

## ■ REFERENCES

- (1) Simarro, P., Diarra, A., Ruiz Postigo, J., Franco, J., and Jannin, J. (2011) The Human African Trypanosomiasis control and Surveillance Programme of the WHO 2000–2009: The Way Forward. *PLoS Negl. Trop. Dis.* 5, e1007.
- (2) Frearson, J. A., Wyatt, P. G., Gilbert, I. H., and Fairlamb, A. H. (2007) Target assessment for antiparasitic drug discovery. *Trends Parasitol.* 23, 589–595.
- (3) Naula, C., Parsons, M., and Mottram, J. C. (2005) Protein kinases as drug targets in *Trypanosomes* and *Leishmania*. *Biochem. Biophys. Acta* 1754, 151–159.
- (4) 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, 127–146.
- (5) 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. *Mol. Cell. Proteomics* 8, 1527–1538.
- (6) 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 Res.* 21, 915–924.
- (7) Davies, S. P., Reddy, H., Caivano, M., and Cohen, P. (2000) Specificity and mechanism of action of some commonly used protein kinase inhibitors. *Biochem. J.* 351, 95–105.
- (8) 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*. *J. Biol. Chem.* 285, 15356–15368.
- (9) Urbaniak, M. D. (2009) Casein kinase 1 isoform 2 is essential for bloodstream form *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 166, 183–185.
- (10) Ojo, K. K., Gillespie, J. R., Reichers, A. J., Napuli, A. J., Verlinde, C. L., Buckner, F. S., Gelb, M. H., Domostoj, M. M., Wells, S. J., Scheer, A., Wells, T. N., and Van Voorhis, W. C. (2008) Glycogen synthase kinase 3 is a potential drug target for African trypanosomiasis therapy. *Antimicrob. Agents Chemother.* 52, 3710–3717.
- (11) 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. *Nat. Biotechnol.* 25, 1035–1044.
- (12) Bantscheff, M., and Drewes, G. (2012) Chemoproteomic approaches to drug target identification and drug profiling. *Bioorg. Med. Chem.* 20, 1973–1978.
- (13) Urbaniak, M. D., Guther, M. L. S., and Ferguson, M. A. J. (2012) Comparative SILAC proteomic analysis of *Trypanosoma brucei* bloodstream and procyclic lifecycle stages. *PLoS One* 7, e36619.
- (14) 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, 482.
- (15) Thomson, A., Schafer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Naeumann, T., Johnstone, R., Mohammed, A. K., and Hamon, C. (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* 75, 1895–1904.



- (16) Ross, P., Huang, Y., Marchese, J., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Junhasz, P., Martin, S., Bartlett-Jones, M., He, F., Jacobson, A., and Pappin, D. (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* 3, 1154–1169.
- (17) Kruse, U., Pallasch, C. P., Bantscheff, M., Eberhard, D., Frenzel, L., Ghidelli, S., Maier, S. K., Werner, T., and Wendtner, C. M. (2011) Chemoproteomics-based kinome profiling and target deconvolution of clinical multi-kinase inhibitors in primary lymphocytic leukemia cells. *Leukemia* 25, 89–100.
- (18) Pfister, D. D., Burkhard, 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, 1126–1135.
- (19) Rotureau, B., Morales, M. A., Bastin, P., and Spath, G. F. (2009) The flagellum-mitogen-activated protein kinase connection in Trypanosomatids: a key sensory role in parasite signalling and development? *Cell. Microbiol.* 11, 710–718.
- (20) Cleghorn, L. A., Woodland, A., Collie, I. T., Torrie, L. S., Norcross, N., Luksch, T., Mpamhanga, C., Walker, R. G., Mottram, J. C., Brenk, R., Frearson, J. A., Gilbert, I. H., and Wyatt, P. G. (2011) Identification of inhibitors of the *Leishmania* cdc2-related protein kinase CRK3. *ChemMedChem* 6, 2214–2224.
- (21) 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 screening yield small molecule inhibitors of *Leishmania* CRK3:CYC6 cyclin-dependant kinase. *PLoS Negl. Trop. Dis.* 5, e1033.
- (22) Cross, G. A. M. (1984) Release and purification of *Trypanosoma brucei* variant surface glycoprotein. *J. Cell. Biochem.* 24, 79–90.
- (23) Aslett, M., Aurrecochea, 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., Jr., Subramanian, S., Thibodeau, R., Tivey, A., Treatman, C., Velarde, G., and Wang, H. (2010) TriTrypDB: a functional genomic resource for the Trypanosomatidae. *Nucleic Acid Res.* 38, D457–D462.
- (24) Bergamini, G., Bell, K., Shimamura, S., Werner, T., Cansfield, A., Mueller, K., Perrin, J., Rau, C., Ellard, K., Hopf, C., Doce, C., Leggate, D., Mangano, R., Mathieson, T., O'Mahony, A., Plavec, I., Rharbaoui, F., Reinhard, F., Savitski, M. M., Ramsden, N., Hirsch, E., Drewes, G., Rausch, O., Bantscheff, M., and Neubauer, G. (2012) A selective inhibitor reveals PI3K $\gamma$  dependence of T<sub>H</sub>17 cell differentiation. *Nat. Chem. Biol.* 8, 576–582.
- (25) Bantscheff, M., Boesche, M., Eberhard, D., Mathieson, T., Sweetman, G., and B., K. (2008) Robust and sensitive iTRAQ quantification on an LTQ Orbitrap mass spectrometer. *Mol. Cell. Proteomics* 7, 1702–1713.
- (26) Miranda-Saavedra, D., and Barton, G. J. (2007) Classification and functional annotation of eukaryotic protein kinases. *Proteins* 68, 893–914.
- (27) Martin, D. M., Miranda-Saavedra, D., and Barton, G. J. (2009) Kinomer v.1.0: a database of systematically classified eukaryotic protein kinases. *Nucleic Acid Res.* 37, D244–250.
- (28) Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon, S., Lefort, V., Lescot, M., Claverie, J. M., and Gascuel, O. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acid Res.* 36, 36.
- (29) Notredame, C., Higgins, D., and Heringa, J. (2000) T-Coffee: A novel method for multiple sequence alignment. *J. Mol. Biol.* 302, 205–217.
- (30) Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552.
- (31) Guindon, S., and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704.
- (32) Anisimova, M., and Gascuel, O. (2006) Approximate likelihood ratio test for branches: a fast, accurate and powerful alternative. *Syst. Biol.* 55, 539–552.
- (33) Chevenet, F., Brun, C., Banuls, A. L., Jacq, B., and Chisten, R. (2006) TreeDyn: towards dynamic graphic and annotation analyses of trees. *BMC Bioinformatics* 7, 439.
- (34) Durrant, J. D., Urbaniak, M. D., Ferguson, M. A. J., and McCammon, J. A. (2010) Computer-aided identification of *Trypanosoma brucei* uridine diphosphatase galactose 4'-epimerase inhibitors: towards the development of novel therapies for African sleeping sickness. *J. Med. Chem.* 53, 5025–5032.

# The Design and Synthesis of Potent and Selective Inhibitors of *Trypanosoma brucei* Glycogen Synthase Kinase 3 for the Treatment of Human African Trypanosomiasis

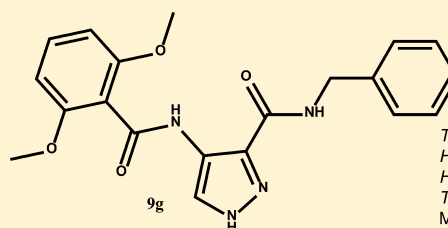
Robert Urich,<sup>†</sup> Raffaella Grimaldi,<sup>†</sup> Torsten Luksch,<sup>†</sup> Julie A. Frearson,<sup>†</sup> Ruth Brenk,<sup>\*,‡,§</sup> and Paul G. Wyatt<sup>\*,†</sup>

<sup>†</sup>Drug Discovery Unit, College of Life Sciences, University of Dundee, Sir James Black Centre, Dow Street, Dundee DD1 5EH, U.K.

<sup>‡</sup>Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Sir James Black Centre, Dundee DD1 5EH, U.K.

## **S** Supporting Information

**ABSTRACT:** Glycogen synthase kinase 3 (GSK3) is a genetically validated drug target for human African trypanosomiasis (HAT), also called African sleeping sickness. We report the synthesis and biological evaluation of aminopyrazole derivatives as *Trypanosoma brucei* GSK3 short inhibitors. Low nanomolar inhibitors, which had high selectivity over the off-target human CDK2 and good selectivity over human GSK3 $\beta$  enzyme, have been prepared. These potent kinase inhibitors demonstrated low micromolar levels of inhibition of the *Trypanosoma brucei* parasite grown in culture.



TbGSK3 IC<sub>50</sub> = 1 nM  
HsGSK3 IC<sub>50</sub> = 330 nM  
HsCDK2 IC<sub>50</sub> > 10  $\mu$ M  
*T. b. brucei* EC<sub>50</sub> = 6  $\mu$ M  
MRC5 EC<sub>50</sub> > 50  $\mu$ M

## ■ INTRODUCTION

Human African trypanosomiasis (HAT) or African sleeping sickness is a serious life threatening disease.<sup>1</sup> Around 60 million people in 36 African countries are currently in constant threat of infection. Although the reported number of cases has dropped over recent years, the actual number of unreported cases is estimated to be around 70000–80000.<sup>2</sup> HAT is caused by infection with *Trypanosoma brucei*, a vector-borne parasite, which is transmitted by the bite of tsetse flies. The symptoms of the disease occur in two main stages. In the first stage, known as the hemolymphatic phase, the parasites multiply in blood, subcutaneous tissues, and lymph, causing headaches, fever, itching, joint pains, and swelling of lymph nodes. In the second stage, or neurological phase, the trypanosomes cross the blood–brain barrier and invade the central nervous system. This phase entails confusion, change of behavior, reduced coordination, sensory disturbances, disturbance of sleep cycle, and finally death. Most available drugs for HAT display severe toxic side effects, require long periods of administration, and/or are expensive due to the logistics to reach rural African areas.<sup>3</sup> Further, resistance to all in use drugs has been observed in the laboratory and/or in the field,<sup>4</sup> resulting in an urgent requirement for better, safer, and inexpensive therapeutic alternatives to the current treatments.

Genetic knockdown studies have identified several proteins that are essential for the survival of the parasite, including members of the protein kinase (PK) family.<sup>5–8</sup> In *Trypanosoma brucei* PKs are essential in many fundamental cellular processes, e.g., proliferation, differentiation, and cell cycle control, and can

therefore be considered as potential drug targets for the treatment of HAT.<sup>7,9–12</sup>

In the *T. brucei* genome there are two kinases that are highly homologous to human glycogen synthase kinase 3 (*HsGSK3*): *TbGSK3* short and *TbGSK3* long.<sup>13</sup> RNA interference (RNAi) knockdown of *TbGSK3* has shown that *TbGSK3* short is critical for cell growth, with a role in the control of mitosis and/or cytokinesis.<sup>7,13</sup>

The ability to selectively inhibit *TbGSK3* over the off-target *HsGSK3* is highly desirable because mouse knockout studies revealed that the disruption of the murine GSK3 $\beta$  gene causes embryonic lethality; consequently, nonselective inhibitors are not applicable for use in infants and women of child bearing age.<sup>14,15</sup>

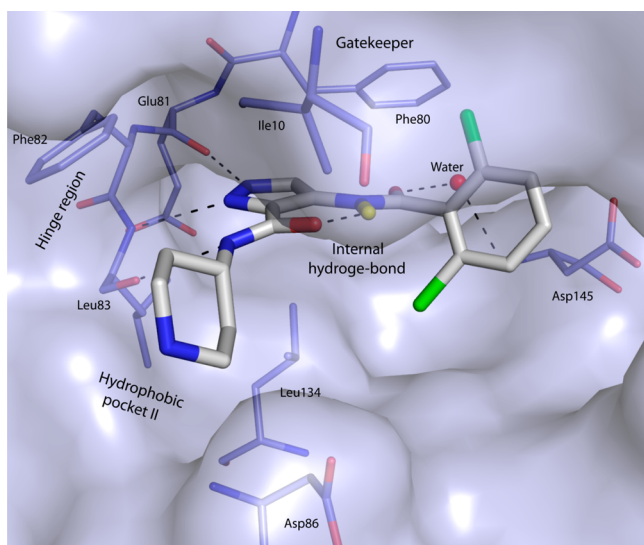
From a homology perspective, *TbGSK3* is not only very closely related to *HsGSK3 $\beta$*  but also to other human PKs such as cyclin dependent kinase 1 (*HsCDK1*) and cyclin dependent kinase 2 (*HsCDK2*).<sup>16</sup> *HsCDK2* and *HsCDK1* are essential for G1/G2 progression and S/M-phase entry of the cell cycle. Off-target inhibition of these human kinases will therefore result in cell cycle arrest and reduction of cellular proliferation and as such potentially lead to severe side effects.

Over the past decade, various groups and pharmaceutical companies have identified multiple series of *HsGSK3 $\beta$*  inhibitors.<sup>16,17</sup> Recently, Astex Therapeutics and researchers at the University of Osaka have developed a series of aminopyrazoles that are potent inhibitors of *HsGSK3 $\beta$* .<sup>18–20</sup>

Received: February 13, 2014

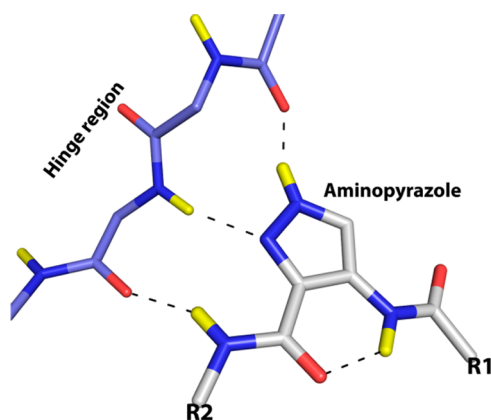
Published: September 8, 2014

Co-crystal structures of this series with *HsGSK3 $\beta$*  are not available to date; however, complex structures with the closely related *HsCDK2* have been determined.<sup>19</sup> In all structures, the pyrazole scaffold forms two hydrogen bonds to the hinge region of *HsCDK2* (Figure 1). Further, the NH group of the 3-



**Figure 1.** Co-crystal structure (2vu3) of AT7519<sup>19</sup> (carbon atoms in gray) bound to CDK2. The binding pocket of CDK2 is shown in light-blue surface representation. Key: red sphere, water molecule; black dashed lines, protein–ligand and water–ligand hydrogen bonds; yellow stick, hydrogen atom.

position amide forms an additional hydrogen-bond interaction to the backbone of Leu83. A water-mediated hydrogen bond from the amide carbonyl oxygen atom to the backbone NH of Asp145 is also observed. The R<sup>1</sup> residues (Figure 2) access the



**Figure 2.** Generic binding mode of the R<sup>1</sup> and R<sup>2</sup> substituted aminopyrazole scaffold (carbon atoms in gray).

gatekeeper region between the gatekeeper residue Phe80 and the catalytic Asp145 (Figure 1). The R<sup>2</sup> substituents occupy the hydrophobic pocket II, formed by the backbone of the linker region, Leu83, Phe82, and side chains of Ile10, Asp86, and Leu134. Finally, an intramolecular hydrogen bond between the R<sup>1</sup>-NH and R<sup>2</sup>-carbonyl group is present. The similarity of *HsCDK2*, *HsGSK3 $\beta$* , and *TbGSK3* indicates that aminopyrazoles will also bind into the ATP-binding site of the latter enzyme.<sup>13,16</sup>

Herein, we describe the design, synthesis, and biological evaluation of aminopyrazole inhibitors which bind to *TbGSK3* short. The inhibitors were also tested against the closely related off-targets *HsGSK3 $\beta$*  and *HsCDK2* and evaluated against a panel of mammalian protein kinases. The most potent compound has nanomolar affinity for *TbGSK3* short, is selective over *HsGSK3 $\beta$*  and *HsCDK2*, and clean in the kinase panel. By using computer-aided molecular modeling, we were able to rationalize the observed selectivity profile. Enzyme affinity correlated with inhibition of *T. b. brucei* proliferation, albeit a 100-fold offset in potency, was found. In light of these results, we discuss the value of *TbGSK3* short as a drug target for HAT.

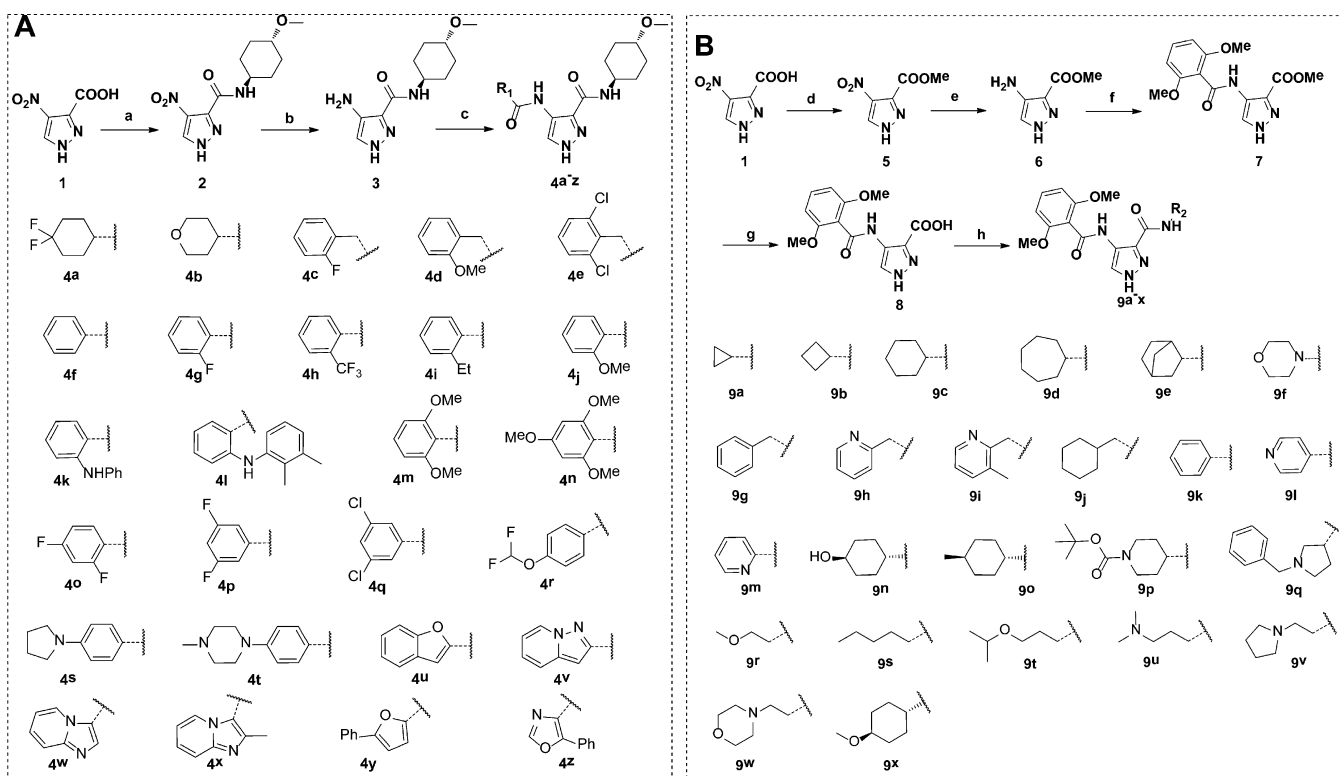
## RESULTS

**Starting Point.** The aminopyrazole derivatives developed by Astex Therapeutics and Yumiko Uno et al. for inhibition of *HsCDK2* and *HsGSK3 $\beta$*  enzymes were chosen as a starting point for the investigation of *TbGSK3* short inhibitors.<sup>18–20</sup> Aminopyrazoles analogues were generated by substituting at either R<sup>1</sup> or R<sup>2</sup> position (Figure 2) using two synthetic routes (Scheme 1A,B).<sup>19</sup>

**Differences in the ATP Binding Pockets of *TbGSK3*, *HsGSK3 $\beta$* , and *HsCDK2*.** A structural model of *TbGSK3* short, *HsGSK3 $\beta$* , and *HsCDK2* and to guide ligand design. An overlay of 42 *HsGSK3* crystal structures showed that there is low flexibility in the ATP binding site. Only the regions including Phe67 and Arg141 showed some mobility. Phe67 either points toward or away from the hinge region. Arg141 also spans a number of distinct conformations, including examples where it occupies space in the binding site (1J1B, 1J1C, 1O9U, 2OSK) and therefore could influence docking results. However, to allow for ligands of a significant size, we have used examples with Arg141 pointing out of the binding site. Therefore, two homology models for *TbGSK3* were generated representing both states of Phe67. As we were mainly interested in aminopyrazoles with less extended R<sup>1</sup> groups, the crystal structure with Phe67 pointing toward the hinge (with structure 1r0e as a representative) was more suited as model system. The selection of 1r0e instead of other members of this group (with Phe67 pointing toward the hinge) was arbitrary. For this analysis, all residues that are located within 6 Å of the ligand bound to the template structure (1r0e) were considered.

The binding pockets of *TbGSK3* and *HsGSK3 $\beta$*  differ by nine amino acid residues (Table 1, Figure 3A). Of the amino acid side chains that point toward the ligand, the most significant differences are the replacement of Tyr134 in *HsGSK3 $\beta$*  with Phe103, Leu132 with Met101, Gln72 with Leu36, and Tyr140 with His109. The binding pockets of *TbGSK3* and *HsCDK2* are more diverse. Here, in total 16 out of 26 amino acids were found to be different (Table 1, Figure 3B), the most important of these differences being the replacement of Lys20 in CDK2 with Leu36, Phe80 with Met101, His84 with Pro105, Lys89 with Arg110, and Ala144 with Cys170. Interestingly, most of the amino acid differences occurred in the hydrophobic pocket II and the gatekeeper region. Therefore, we decided to direct the optimization of the lead scaffold toward suitable interactions with amino acids which are located in these subpockets of the ATP binding site.

**Chemistry.** The synthesis of R<sup>1</sup> and R<sup>2</sup> substituted aminopyrazole derivatives started from 4-nitro-pyrazole-3-carboxylic acid **1** and is described by two different routes

Scheme 1. Synthetic Routes for 50 R<sup>1</sup> and R<sup>2</sup> Substituted Aminopyrazoles (4a–z and 9a–x)<sup>a</sup>

<sup>a</sup>Reagents and conditions used in routes A and B: (a) *trans*-4-methoxycyclohexylamine, EDC, HOBt, DMF, rt; (b) 10% Pd/C, H<sub>2</sub>, DMF, rt; (c) R<sup>1</sup>COOH, EDC, HOBt, DIPEA, rt; (d) SOCl<sub>2</sub>, MeOH, 0 °C, rt; (e) 10% Pd/C, H<sub>2</sub>, EtOH, rt; (f) 2,6-dimethoxybenzoyl chloride, Et<sub>3</sub>N, dioxane, rt; (g) NaOH, dioxane, H<sub>2</sub>O, rt; (h) R<sub>2</sub>NH<sub>2</sub>, polystyrene-bound carbodiimide, HOBt, acetonitrile, MW, 100 °C.

(Scheme 1) based on previous work from Wyatt et al.<sup>19</sup> In route A, **1** was coupled with *trans*-4-methoxycyclohexylamine using EDC as the activating agent. Reduction of the subsequent intermediate **2** by hydrogenation in the presence of palladium on carbon generated amino pyrazole **3**. The conversion to compounds **4a–z** was accomplished by coupling of **3** with a suitable selection of carboxylic acids. In route B, after esterification of the carboxyl group of **1**, the nitro group of intermediate **5** was reduced to afford amine **6**. Treatment of **6** with 2,6-dimethoxybenzoyl chloride under standard conditions, followed by base hydrolysis of the ester, provided acid **8**. In the final step, **8** was coupled with appropriate amines in a microwave reaction using polystyrene-bound carbodiimide to yield final compounds **9a–x**.

**Activity and Selectivity of R<sup>1</sup> Substituted Compounds.** Twenty-six R<sup>1</sup> substituted aminopyrazole analogues (Table 2) were made according to the synthetic route shown in Scheme 1A. A range of R<sup>1</sup> groups varying in size and polarity was chosen to probe whether the differences in the gate keeper region between *Tb*GSK3, *Hs*GSK3, and *Hs*CDK2 could be exploited to derive selective and potent *Tb*GSK3 inhibitors (Figure 3).

**Enzyme Activity.** All compounds showed good potency against *Tb*GSK3 (<1 μM). An unsubstituted phenyl ring (**4f**) provided on average a 20-fold improvement of inhibition potency relative to saturated six-membered ring systems (**4a** and **4b**) and benzyl groups (**4c**, **4d**, and **4e**). In general, a variety of different aryl and heteroaryl rings (**4g–4y** compared to **4a** and **4b**) in the R<sup>1</sup> position led to significantly improved potency against *Tb*GSK3. Additionally, a wide variety of

substituents were tolerated on the phenyl ring. In general, *ortho*-substituted phenyl rings gave the best improvement in activity compared to the unsubstituted phenyl group (**4j**, **4k**, **4m**, and **4n**). The methoxyphenyl moieties in **4j**, **4m**, and **4n**, which had *Tb*GSK3 IC<sub>50</sub> values of 4, 2, and 3 nM, respectively, were the most favorable substituents. These derivatives were approximately 10-fold more potent than the unsubstituted phenyl compound **4f**. Only the 2,4,6-trimethoxy derivative **4n** showed >10-fold selectivity over *Hs*GSK3. Interestingly, this was also the most selective compound for *Hs*CDK2 (>1000-fold).

To rationalize the observed selectivity, all analogues were docked into the binding sites of *Tb*GSK3, *Hs*GSK3β, and *Hs*CDK2 and their poses were visually analyzed. For most compounds, a binding mode similar to that observed for AT7519<sup>19</sup> in *Hs*CDK2 (Figure 1) was predicted in *Tb*GSK3, *Hs*GSK3β, and *Hs*CDK2. One important difference between *Hs*CDK2, *Tb*GSK3, and *Hs*GSK3β is the gatekeeper residue (Table 1, Figure 3A,B). While *Hs*GSK3 and *Tb*GSK3 enzymes have Leu or Met, respectively, in this position, in *Hs*CDK2 Phe is present. As a consequence, the gatekeeper region of *Hs*CDK2 (located between Phe80 and Asp145) is more restricted compared to the other two enzymes. This resulted in a higher energy, out of plane conformation of the amide group of **4f** when binding into this pocket (Figure 4), while a low energy conformation was found when binding into *T. brucei* and human GSK3 (not shown). Further, without induced fit adaptations, the bulky R<sup>1</sup>-substituents such as the 2,6-dimethoxybenzamide group of **4m**, the 2,4,6-trimethoxybenzamide groups **4n**, and the phenylaminobenzamide groups of **4k**

**Table 1.** Differences in the Binding Pockets of *Tb*GSK3, *Hs*GSK3 $\beta$ , and *Hs*CDK2<sup>a</sup>

<i>Tb</i> GSK3	<i>Hs</i> GSK3 $\beta$	<i>Hs</i> CDK2
V25	V61	<b>K9</b>
A26	<b>I62</b>	<b>I10</b>
G27	G63	G11
Q28	<b>N64</b>	<b>E12</b>
G29	G65	G13
T30	<b>S66</b>	T14
F31	F67	<b>Y15</b>
V34	V70	V18
L36	<b>Q72</b>	<b>K20</b>
A47	A83	A31
K49	K85	K33
E61	E97	<b>I52</b>
M65	M101	<b>L55</b>
V77	V110	V64
M101	<b>L132</b>	<b>F80</b>
E102	<b>D133</b>	E81
F103	<b>Y134</b>	F82
I04	V135	<b>L83</b>
P105	P136	<b>H84</b>
E106	E137	<b>Q85</b>
T107	T138	<b>D86</b>
H109	<b>Y140</b>	<b>K88</b>
R110	R141	<b>K89</b>
K154	K183	K129
H156	<b>Q185</b>	<b>Q131</b>
N157	N186	N132
L159	L188	L134
C170	C199	<b>A144</b>
D171	D200	D145

<sup>a</sup>Amino acids of *h*GSK3 $\beta$  or *Hs*CDK2 which differ in *Tb*GSK3 short are shown in boldface.

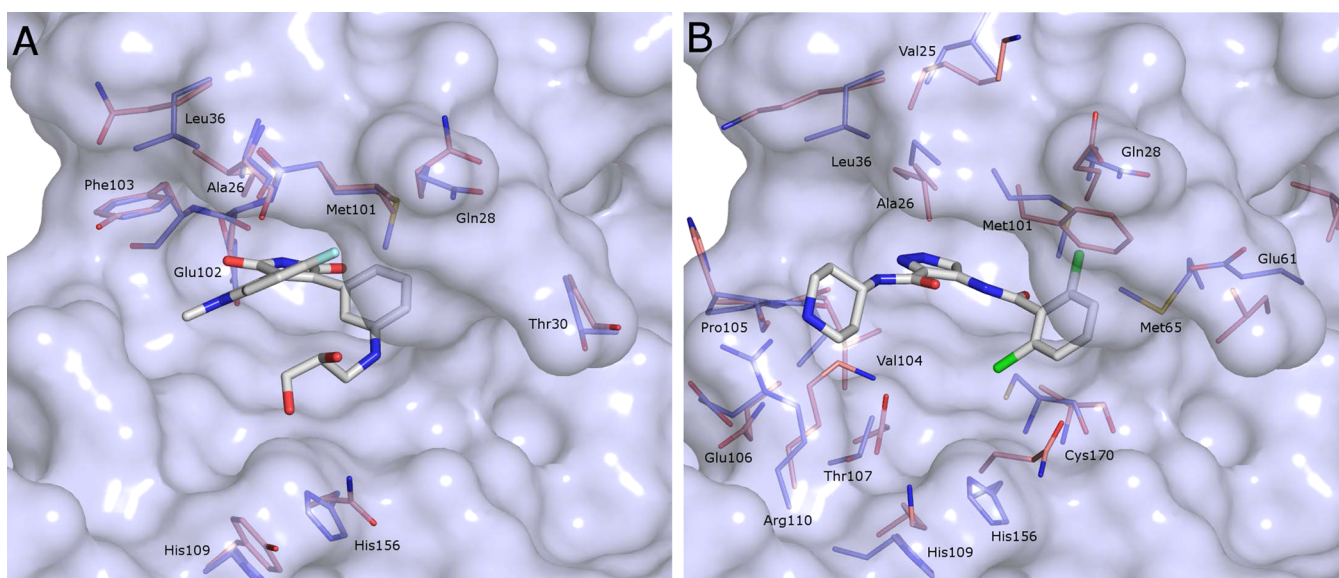
and **4l** can only be accommodated by the gatekeeper region of *Tb*GSK3 and *Hs*GSK3 $\beta$  but not the narrower *Hs*CDK2 gatekeeper region. These observations might explain the reduced binding affinity of **4f** for *Hs*CDK2 compared to *Hs*GSK3. Of note, this explanation is further supported by the report by Wyatt et al.,<sup>19</sup> which found that the aryl groups which are located in the same position need to twist in order to provide potent CDK2 activity.

**Antiparasitic Activity.** All R<sup>1</sup> substituted compounds were tested for their ability to inhibit the proliferation of bloodstream form (BSF) *T. b. brucei* in culture. As an initial indication of potential toxicity, compounds **4a–4z** were additionally tested against proliferating human fetal lung fibroblast cells (MRC5 cell line). Four compounds (**4g**, **4j**, **4m**, and **4y**) had EC<sub>50</sub> values <1  $\mu$ M and a further 11 compounds had EC<sub>50</sub> values <3  $\mu$ M against BSF *T. b. brucei* (Table 2). The EC<sub>50</sub> values correlated well with enzyme activity ( $R^2 = 0.73$ , Figure 5). However, a 100-fold drop from enzyme to cellular activity was observed. Selectivity over the MRC5 cells was achieved with compounds **4m** (60-fold), **4s** (>19-fold), **4x** (12-fold), and **4n** (>9-fold), however, the majority of compounds showed a poor selectivity over MRC5 cells.

#### Activity and Selectivity of R<sup>2</sup> Substituted Compounds.

As **4m** was the most potent inhibitor of *Tb*GSK3 and proliferation of *T. b. brucei* cells, we retained the 2,6-dimethoxybenzamide group at position R<sup>1</sup> for optimization of the R<sup>2</sup> substituent. R<sup>2</sup> substituted aminopyrazole analogues (**9a–9x**) were made according to the synthetic route shown in Scheme 1B to explore the structural requirements for improvement of antiparasitic activity and selectivity over the closely related human kinases.

**Enzyme Activity.** The majority of variations led to potent *Tb*GSK3 inhibitors, indicating that chemical diversity at this position was well tolerated (Table 3). One of the SAR trends observed was that six-membered saturated rings (**9c**) and seven-membered saturated rings (**9d**) were favored over their three- and four-membered equivalents (**9a** and **9b**). Further, it

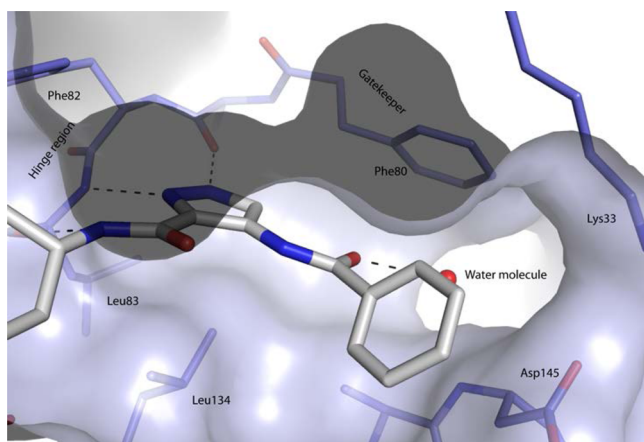


**Figure 3.** Superposition of the binding sites of the homology model of *Tb*GSK3 short (blue carbon atoms) with (A) the *Hs*GSK3 $\beta$  crystal structure (PDB code 1r0e) and (B) the *Hs*CDK2 crystal structure (PDB code 2vu3). The solvent accessible surface of *Tb*GSK3 short is shown in light blue. Only residues that differ between the binding pockets are shown. For orientation, the ligands bound to crystal structures are also displayed. Amino acid residue labels are for *Tb*GSK3.

Table 2. Kinase Inhibitory Activity and Antiproliferative Efficacy of R<sup>1</sup> Substituted Aminopyrazoles

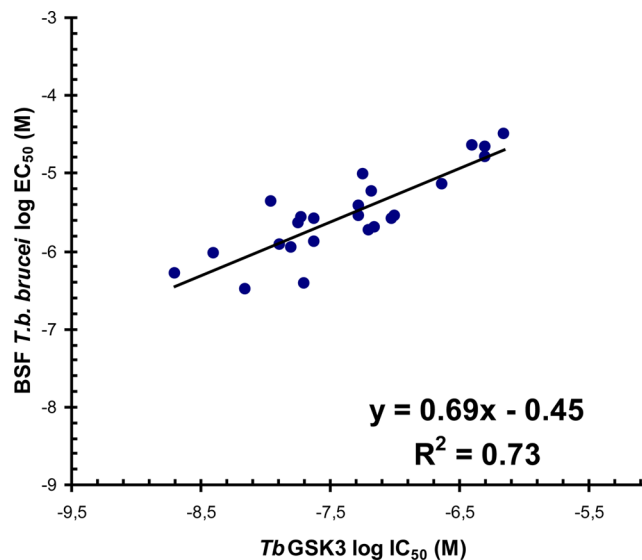
compd	IC <sub>50</sub> (μM)		IC <sub>50</sub> (μM)			EC <sub>50</sub> (μM)	
	TbGSK3 <sup>a</sup>	HsGSK3β <sup>a</sup>	ratio HsGSK3β/TbGSK3	HsCDK2 <sup>a</sup>	ratio HsCDK2/TbGSK3	<i>T. b. brucei</i> <sup>b</sup>	MRC5 <sup>c</sup>
4a	0.50	0.008	0.16	1.0	2	16	>50
4b	0.69	0.008	0.0012	0.38	0.55	32	>50
4c	0.39	0.14	0.35	0.82	2.1	23	>50
4d	0.50	0.21	0.42	0.85	1.7	22	40
4e	0.23	0.2	0.87	0.29	1.3	7.3	3.1
4f	0.024	<0.005	>0.21	0.038	<1.6	1.3	0.8
4g	0.020	<0.005	0.25	0.014	<0.7	0.4	1.0
4h	0.018	<0.013	<0.72	0.1	5.6	2.3	3.3
4i	0.053	0.02	0.38	0.07	1.3	2.8	3.4
4j	0.004	<0.013	<3.3	0.1	25	0.9	1.5
4k	0.011	0.02	1.8	>10	>910	4.4	35
4l	0.066	0.03	0.45	>10	>150	5.8	16
4m	0.002	<0.005	<2.5	0.19	95	0.5	31
4n	0.003	0.09	30	3.1	1000	5.9	>50
4o	0.053	<0.005	<0.09	0.22	4.2	3.8	11
4p	0.024	<0.005	<0.21	0.083	3.5	2.6	0.6
4q	0.057	<0.005	<0.08	0.63	11	9.6	13
4r	0.016	<0.005	<0.31	0.27	17	1.1	2.5
4s	0.019	<0.005	<0.26	0.13	6.8	2.7	>50
4t	0.070	<0.005	<0.071	1.0	14	2.0	20
4u	0.013	<0.005	<0.38	nd	nd	1.2	0.8
4v	0.063	<0.005	<0.080	0.042	0.67	1.9	1.0
4w	0.094	<0.005	<0.053	0.15	1.6	2.6	5.9
4x	0.10	0.042	0.42	2.2	22	2.9	34
4y	0.007	<0.005	<0.72	0.01	1.4	0.3	0.1
4z	0.92	0.005	0.0058	22	24	>50	28

<sup>a</sup>Data represents the average of two or more experiments. <sup>b</sup>Concentration required to inhibit the growth of *T. b. brucei* in culture by 50% over 72 h. <sup>c</sup>Concentration required to inhibit the growth of MRC5 cells in culture by 50% over 72 h.



**Figure 4.** Predicted binding mode of **4f** in HsCDK2. Putative hydrogen bonds are shown as black dotted lines. Docking results suggested that the phenyl ring of compound **4f** needs to be significantly twisted out of plane by approximately 60° compared to the amide in order to fit into the gatekeeper region..

was noted that the replacement of the cyclohexane of **9c** with a phenyl ring or 4-pyridine, to give **9k** or **9l**, gave a 6-fold decrease in potency against *TbGSK3*. The 2-pyridine analogue (**9m**), on the other hand, was much less active (50-fold) against *TbGSK3*. Homologation of aromatic (**9g**) and saturated six-membered (**9j**) rings by one carbon atom produced inhibitors with 1 nM activity for *TbGSK3*. For aliphatic side chain derivatives **9r–9w**, the pentanyl and 1-isopropoxypropyl analogues had IC<sub>50</sub> values of 1 nM. The impact of replacing the



**Figure 5.** Correlation between the inhibition of recombinant *TbGSK3* and bloodstream form *T. b. brucei* proliferation by R<sup>1</sup> substituted aminopyrazole derivatives (**4a–z**).

amide group (3-position) with carboxylic acid and ester groups was investigated with compounds **8** and **7**. Compound **8** containing a carboxylate group in the 4-position showed a dramatic loss in activity (IC<sub>50</sub> >50 μM). The ester group of compound **7** on the other hand was better tolerated (IC<sub>50</sub> 0.5 μM). Interestingly, compared with **4a–4z**, a majority of R<sup>2</sup> substituted analogues (**9a–9x**) showed selectivity over

Table 3. Kinase Inhibitory Activity and Antiproliferative Efficacy of R<sup>2</sup> Substituted Aminopyrazoles

compd	IC <sub>50</sub> (μM)		IC <sub>50</sub> (μM)			EC <sub>50</sub> (μM)	
	TbGSK3 <sup>a</sup>	HsGSK3β <sup>a</sup>	ratio HsGSK3β/TbGSK3	HsCDK2 <sup>a</sup>	ratio HsCDK2/TbGSK3	<i>T. b. brucei</i> <sup>b</sup>	MRC5 <sup>c</sup>
9a	0.012	0.22	18	>10	>830	19	>50
9b	0.008	0.08	10	2.4	300	12	>50
9c	0.001	0.05	50	1.2	1200	4.1	35
9d	0.001	0.02	20	2.0	2000	4.5	42
9e	0.018	nd	nd	>10	560	16	>50
9f	0.081	0.45	5.6	>10	120	50	>50
9g	0.001	0.33	330	>10	10000	5.9	50
9h	0.015	0.32	21	>10	670	20	>50
9i	0.14	0.87	6.2	>10	71	>50	>50
9j	0.001	nd	nd	nd	nd	7.7	>50
9k	0.006	0.07	12	4.3	720	11.5	>50
9l	0.004	0.12	30	1.3	330	8.2	>50
9m	0.32	0.94	2.9	>10	31	>50	>50
9n	0.002	0.07	35	1.6	800	6.4	34
9o	0.001	nd	nd	>4.8	4800	6.7	45
9p	0.006	0.14	23	4.7	780	12	>50
9q	0.008	0.08	10	>10	1300	8.9	38
9r	0.034	0.3	8.8	>10	290	43	>50
9s	0.001	0.1	100	4.8	4800	7.3	>50
9t	0.001	nd	nd	>10	10000	6.6	>50
9u	0.33	0.47	1.4	>10	30	>50	>50
9v	0.32	0.66	2.1	>10	31	>50	>50
9w	0.054	0.63	12	>10	190	>50	>50
9x	0.002	<0.005	<2.5	0.19	95	0.5	31
8	>50	>10		>10		>50	>50
7	0.52	4.5	8.7	>10	19	>50	>50

<sup>a</sup>Data represents the average of two or more experiments. <sup>b</sup>Concentration required to inhibit the growth of *T. b. brucei* in culture by 50% over 72 h. <sup>c</sup>Concentration required to inhibit the growth of MRC5 cells in culture by 50% over 72 h.

*HsCDK2* and *HsGSK3β*. These are the most selective *TbGSK3* inhibitors described to date.

The predicted binding mode of **9g** in *TbGSK3* offers an explanation for the observed selectivity (Figure 6). In the highest scoring docking pose, the core scaffold adopts a similar binding mode as observed for AT7519 in *HsCDK2* (Figure 1).<sup>19</sup> In addition, the docking results suggested that the hydrophobic pocket II of *TbGSK3* was occupied by the *N*-benzylamide group of **9g** in such a way that its phenyl moiety formed T-shaped edge-to-face interactions with the side chain of Phe103 and hydrophobic interactions with Leu36 and Ala26. In *hGSK3β*, Phe103 is replaced with a Tyr (Table 1, Figure 3A), resulting in steric clash and electrostatic repulsion toward the benzyl moiety of **9g**. Further, Leu36 is substituted with Gln72 in *hGSK3β* and Lys20 in *HsCDK2*, diminishing hydrophobic interactions between the benzyl moiety of **9g** and these residues. Overall, these changes together with differences in the gatekeeper region of *HsCDK2* (see above) are likely to be responsible for the high selectivity of **9g** for *TbGSK3* over *hGSK3* and *HsCDK2*. Similar observations regarding the R<sup>2</sup> group placement and selectivity were also made for compound **9h** supporting this model.

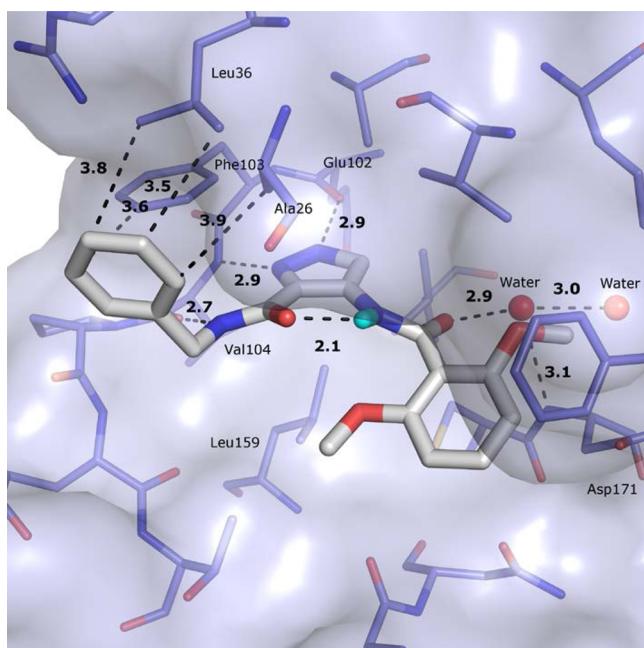
**Antiparasitic Activity.** The R<sup>2</sup> substituted compounds were tested against BSF *T. b. brucei* and MRC5 cells. As for the R<sup>1</sup>-substituted analogues, a good correlation between the EC<sub>50</sub> and IC<sub>50</sub> values and a 100-fold drop in activity between the biochemical and cell assay was observed (Figure 7). Compound **9c** had an EC<sub>50</sub> for *T. b. brucei* of 4 μM (Table 3). Limited selectivity (>7-fold) over MRC5 cells was achieved with compounds **9c**, **9d**, **9g**, **9s**, and **9t**. It was found that the

compounds showed selective inhibition of *TbGSK3* over *HsGSK3* (>20-fold) and *HsCDK2* (>1200-fold).

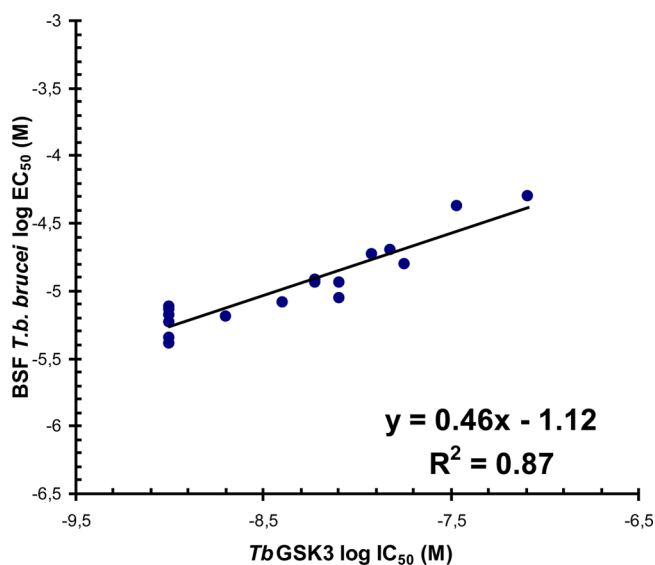
**Human Kinase Selectivity Profile.** PK inhibitors frequently inhibit multiple kinases, often leading to off-target toxic effects. To assess the selectivity of the aminopyrazole inhibitors, remaining activity at 10 μM concentration was measured for compounds **4f**, **4m**, and **4y** against a panel of 80 human PKs and for compound **9g** against 124 human PKs. Compounds **4m** and **9g** were found to be highly specific (Table 4). Compound **4m** inhibited only two PKs, namely *GSK3β* and *CDK2*, at more than 80%. **9g** showed activity against three PKs: *GSK3β*, *MAPKAP-K2*, and *MINK1* at more than 80%. Compound **4f** was found to inhibit seven PKs and compound **4y** 15 PKs by greater than 80% at 10 μM.

## DISCUSSION

In this work, we exploited the knowledge of the previously described aminopyrazoles inhibitors of *HsCDK2* and *HsGSK3β*<sup>20</sup> to identify selective inhibitors of the *TbGSK3* short isoform. This kinase has been shown using genetic manipulation studies to be essential for the survival of the *T. b. brucei* parasite.<sup>13</sup> However, we wanted to confirm if antiparasitic activity could be gained using selective, small molecule inhibitors of *TbGSK3*. The ability to selectively inhibit *TbGSK3* over *HsGSK3* and *HsCDK2* is essential to avoid potential side effects. Therefore, more than 50 aminopyrazole derivatives were synthesized and screened against *TbGSK3*, *HsGSK3β*, *HsCDK2*, and proliferating *T. b. brucei* and human cells in culture.



**Figure 6.** Proposed binding mode of **9g** in the homology model of *TbGSK3* (blue carbon atoms) overlaid on the *HsGSK3β* crystal structure (pink carbon atoms). Both ligand and protein are represented as sticks and color coded by atom types. Ligand carbon atoms are shown in gray, protein carbon atoms of *TbGSK3* are shown in blue, and *HsGSK3β* carbon atoms in salmon. Amino acid residue labels are for *TbGSK3*. Hydrogen bonds and hydrophobic interactions are shown as black dotted lines, with interaction distances in angstroms. *TbGSK3* amino acids which are involved in hydrophobic interactions with the benzyl group are marked in bold. The gold sphere represents the center of the phenyl ring.



**Figure 7.** Correlation between the inhibition of recombinant *TbGSK3* and bloodstream from *T. b. brucei* proliferation using *R*<sup>2</sup> substituted aminopyrazole derivatives (**9a–x**).

The results (Table 2 and Table 3) showed that almost all compounds were highly potent *TbGSK3* inhibitors. The activity could be rationalized using the homology models and subsequent molecular docking studies (Figure 4 and Figure 6). The aminopyrazole derivatives make three H-bond interactions with the kinase hinge region, driving much of the

**Table 4.** Kinase Profiling against a Panel of Mammalian Kinases<sup>a</sup>

PKs	<b>4f</b>	<b>4m</b>	<b>4y</b>	<b>9g</b>
MKK1	54	83	17	76
ERK2	<b>18</b>	51	2	92
JNK1	31	77	8	71
JNK2	51	82	15	63
ERK8	14	27	7	47
MAPKAP-K2	92	33	90	<b>13</b>
GSK3b	33	<b>0</b>	<b>0</b>	<b>10</b>
CDK2	<b>1</b>	<b>9</b>	<b>1</b>	38
MELK	32	100	13	96
DYRK1A	<b>6</b>	84	<b>1</b>	79
DYRK2	<b>3</b>	63	<b>1</b>	28
DYRK3	24	100	2	80
PIM1	46	100	13	92
PIM3	<b>19</b>	94	<b>1</b>	96
HIPK2	14	100	4	93
IGF-1R	96	100	15	79
MINK1	nd	nd	nd	<b>18</b>

<sup>a</sup>Numbers represent average percentage of activity compared to the control at 10  $\mu$ M. In this table, only kinases with activity values <20% are shown (for full table see Supporting Information). PK activity values <20% are marked in bold.

potency of the compounds against the three kinases studied. Selectivity could be derived from substitution at both *R*<sup>1</sup> and *R*<sup>2</sup> positions.

From a homology perspective, *HsCDK2* is the most closely related kinase to *HsGSK3β*.<sup>16</sup> Although, the enzymes only share approximately 33% amino acid identity, their ATP binding pockets are highly conserved,<sup>16</sup> resulting in the majority of known *HsCDK2* inhibitors also potently inhibiting *HsGSK3β*. The results demonstrated that the required profile could be achieved, with several compounds with high affinity (<18 nM) for *TbGSK3* showing high selectivity (>500-fold) over *HsCDK2*. Docking studies provided a number of important insights into the binding modes and the selectivity profile of aminopyrazole derivatives. First, the docking results suggested that if the phenyl ring of compound **4f** is planar with the amide group at *R*<sup>1</sup> position it cannot bind into the truncated gate keeper region of *HsCDK2*, defined by Phe80 in *HsCDK2*, compared to Met101 in *TbGSK3*. To fit into this region of *HsCDK2*, the phenyl ring needs to significantly twist out of plane of the amide, with a torsion angle of approximately 60°, resulting in reduced binding affinity (Figure 4). To stabilize this twist, di-*ortho*-substituents on the *R*<sup>1</sup> phenyl group are required to cause a steric/electronic clash with the carbonyl of the amide bond. However, this region of the pocket in *HsCDK2* is narrow (in a plane perpendicular to the hinge backbone and the pyrazole core), only allowing small *ortho*-substituents (such as in compound **9g**) on the phenyl group. In contrast, the wider gatekeeper regions of *hGSK3* and *TbGSK3* can tolerate large substituents such as the *ortho*-dimethoxy groups of compound **4m** or **4n** and the *ortho*-phenylaminobenzamide groups of **4k** or **4l**. Second, the highest increase in selectivity (>10000-fold) over *HsCDK2* was achieved by accessing the hydrophobic pocket II. Exploitation of hydrophobic interactions in these two pockets not only reliably increased ligand-binding affinity but also impacted on the selectivity profile of these compounds. On the basis of the biological results of compound **9g** and structural modeling studies, we have shown that selectivity over



HsGSK3 can be achieved by exploiting the Phe103Tyr, Leu36Gln, and Ala26Ile active site differences in the hydrophobic pocket II of *TbGSK3* enzyme (Figure 6). Taken together, the region between the gate keeper residue and the catalytic aspartate of the DFG loop, together with the hydrophobic pocket II, are the key areas to exploit to achieve high selectivity over *HsCDK2*.

The *TbGSK3* short enzyme  $IC_{50}$  values correlated well with the *T. b. brucei* antiproliferative  $EC_{50}$  activities of the described substituted aminopyrazole inhibitors (Figures 5 and 7), indicating that the compounds act on target. However, a 100-fold drop in cell activity was observed, compared to that in the *TbGSK3* assay (1  $\mu$ M). The calculated physical properties (MW < 473; log *P* -0.4–3.6; PSA < 130 Å) of the series of compounds suggests this loss of activity was not driven by lack of cellular penetration. In addition, the compound series was observed to be highly chemically stable under the range of synthetic conditions used during the chemistry campaign, suggesting that chemical degradation was not responsible for the loss of activity in the proliferation assay. Although metabolism by the parasite cannot be ruled out, the high degree of correlation between enzyme inhibition and antiparasitic activity suggests this is not the case, as it would be not expected that all compounds be metabolized to a constant extent. Therefore, the drop of activity was probably due to the high ATP concentration (millimolar range) in the cell compared to the kinase assay conditions.<sup>21</sup> Furthermore, chemical proteomic profiling conducted in parasite cell extracts confirmed that compound **4m** binds the endogenous *TbGSK3* short with nanomolar affinity and very few other kinase targets with much lower affinity in the micromolar range.<sup>22</sup>

## CONCLUSION

In this study, we have developed a series of substituted aminopyrazole amides as *TbGSK3* short inhibitors starting from a compound series initially designed by Astex Therapeutics to inhibit *HsCDK2* and *HsGSK3 $\beta$* . SAR investigation and optimization successfully provided 18 low nanomolar ( $IC_{50}$  < 10 nM) inhibitors of *TbGSK3* with high selectivity (>10000-fold) over *HsCDK2*. With compound **9g**, we have shown that good (330-fold) selectivity over *HsGSK3* can be achieved by targeting the hydrophobic pocket II. Compound **9g** is the most selective *TbGSK3* inhibitor described to date.<sup>13,23–25</sup> In addition, **9g** proved to be highly selective against a panel of 124 human PKs, showing >90% inhibition at 10  $\mu$ M against only one PK, *HsGSK3 $\beta$* . Molecular modeling has also shown that despite overall conservation in sequence and conformation between the three PKs (*HsGSK3 $\beta$* , *HsCDK2*, and *TbGSK3*), the binding pockets have distinct features that determine their specificity for particular compounds. Further, we have shown that enzymatic inhibition correlates well with cell efficacy over a wide range of concentrations and a representative member of this series binds the endogenous *TbGSK3* with nanomolar potency, indicating that compounds definitely act on target.<sup>22</sup> However, a general 100-fold drop in activity between target and cellular activities resulted at best in compounds with low micromolar antiparasitic activity. Taken together, this data suggests that specific ATP competitive hinge binders of *TbGSK3* short require low picomolar potency to obtain nanomolar antiproliferative activity against *T. brucei*. This leads us to the conclusion that alternative strategies are required. First, non-ATP competitive approaches to inhibition of *TbGSK3*, through

irreversible hinge binders or allosteric inhibitors, could be pursued. However, these approaches have potential downsides, through the introduction of a reactive functionality or an increased chance of resistance causing mutations, respectively. Second, a polypharmacology approach through the inhibition of a number of essential *T. brucei* kinases in addition to *TbGSK3* could be investigated, although obtaining selectivity over human kinases would be more problematical. However, the aminopyrazole compounds (**4a–4z** and **9a–9x**) reported here represent an excellent start for chemistry optimization of selective *TbGSK3* short inhibitors and an outstanding probe for studying the physiological functions of *TbGSK3* short in *T. brucei* parasites.

## EXPERIMENTAL SECTION

**Molecular Modeling. Homology Modeling.** Sequence alignments between *T. brucei* and *HsGSK3 $\beta$*  were generated using ClustalW.<sup>26</sup> Subsequently, Modeler 9.2<sup>27</sup> was used to build homology models of *TbGSK3* short, whereas the *HsGSK3 $\beta$*  crystal structure (PDB code 1r0e) served as template. Modeler was run with default settings, and only the highest-scoring structure was used for further analysis and modeling.

**Ligand Docking.** FlexX 2.0.1 (BioSolveIT GmbH) was used to dock ligands flexible into protein binding sites.<sup>28</sup> The active sites were defined as the areas within 7 Å of the co-crystallized ligands of *HsCDK2* (PDB code 2vu3)<sup>19</sup> and *HsGSK3 $\beta$*  (PDB code 1r0e)<sup>29</sup> or the equivalent residues in the homology model of *TbGSK3*. In all three structures, protonation states of amino acids and the orientations of the protons of hydroxyl and amine groups of active-site residues were manually assigned using the FlexX GUI. A highly conserved water molecule (H<sub>2</sub>O 82 in 1r0e or H<sub>2</sub>O 2134 in 2vu3) was kept in all three protein structures used for docking. Docking was carried out using default settings, and only the highest scoring binding modes were visually analyzed.

All figures of protein binding sites were prepared using PyMol.<sup>30</sup>

**Potency Screen Assays.** For compound potency determinations, a radiometric 96-well Flashplate assay (PerkinElmer) was adopted. Compounds were solubilized in DMSO at a top concentration of 3 mM and serially diluted to achieve 10-point titration of final assay concentrations from 30  $\mu$ M to 0.3 nM with a final DMSO concentration of 1% (v/v). The reaction mixtures contained 1  $\mu$ M biotinylated GSP2 substrate, 1  $\mu$ M ATP, 3.7 KBq/well [ $\gamma$ -<sup>33</sup>P]-ATP and 2.5 nM *TbGSK3* in the *TbGSK3* kinase assay buffer. GSK3 inhibitors were screened for selectivity assessment also against *HsGSK3 $\beta$* . For *HsGSK3* assay, the reaction mixes contained 1  $\mu$ M biotinylated GSP2 substrate, 2  $\mu$ M ATP, 7.4 KBq/well [ $\gamma$ -<sup>33</sup>P]-ATP and 15 nM *HsGSK3 $\beta$*  in the *TbGSK3* kinase assay buffer (25 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 5 mM DTT, 0.02% CHAPS, 2 U/mL heparin). For *HsCDK2*/cyclin A assay, the reaction mixtures contained 1 mM CDK5 biotinylated peptide substrate (Biotin-C<sub>6</sub>-PKTPKAKKLL), 1  $\mu$ M ATP, 7.4 KBq/well [ $\gamma$ -<sup>33</sup>P]-ATP and 2 nM *HsCDK2*/cyclin A in the kinase assay buffer (50 mM Tris-HCl, pH 7.5, 10 mM MgCl<sub>2</sub>, 2 mM DTT, 100 mM NaCl, 0.2 mM EGTA, 0.02% (v/v) Brij35).

**Statistical Evaluation of Assay Reproducibility.** The statistical significance of the compound potency ( $IC_{50}$ ) was based on the performance of standard molecules which have been tested to a high replication. In the case of *TbGSK3* short assay, the standard compound GW8510 was tested 93 times across 9 independent runs. The average  $pIC_{50}$  value was 8.26 with a SD (standard deviation) of 0.23. The minimum significant ratio (MSR) of 0.4 was evaluated considering the following formula:

$$MSR = 2.3 \times SD / \sqrt{N}$$

where SD is the standard deviation and *N* the number of replicate values routinely used for the assay (2 in our case).<sup>31</sup> This implies that a difference of >0.4 in  $pIC_{50}$  can be considered statistically significant for this assay.

In the case of *Hs*GSK3 assay, the standard compound GW8510 was tested 47 times across 5 independent runs, with an average pIC<sub>50</sub> value of 8.10 and a SD of 0.21. This implies that a difference of >0.3 in pIC<sub>50</sub> can be considered statistically significant for this assay.

For the *Hs*CDK2 assay, the analysis was performed using two different standards (GW8510 and staurosporine) tested respectively five times in a single run and 19 times in 2 independent runs. This implies that a difference of >0.3 in pIC<sub>50</sub> can be considered statistically significant for this assay.

**Mammalian Kinase Profiling.** Selected compounds were screened against a panel of mammalian kinases routinely run by the Division of Signal Transduction Therapy (DSTT) at the University of Dundee in duplicate at 10 μM.<sup>32</sup> Enzymes included in the panel and assay conditions are reported in the literature. All biochemical assays are run below the K<sub>m</sub><sup>app</sup> for the ATP for each enzyme, allowing comparison of inhibition across the panel.

**Trypanosome and MRC5 Proliferation Assay.** Measurement of inhibition of the proliferation of MRC5 (human lung fibroblast) cells and *T. b. brucei* bloodstream stage cells was performed using a modification of the cell viability assay previously described.<sup>33</sup> Compounds (50 μM to 0.5 nM) were incubated with 2 × 10<sup>3</sup> cells/well in 0.2 mL of the appropriate culture medium (MEM with 10% fetal bovine serum for MRC5 cells) in clear 96-well plates. Plates were incubated at 37 °C in the presence of 5% CO<sub>2</sub> for 69 h. Resazurin was then added to a final concentration of 50 μM, and plates were incubated as above for a further 4 h before being read on a BioTek flx800 fluorescent plate reader.

**Chemistry. General Experimental Details.** <sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded on either a Bruker Avance DPX 300 or 500 MHz spectrometer. Chemical shifts (δ) are expressed in parts per million (ppm) and coupling constants (J) are in hertz (Hz). Signal splitting patterns are described as singlet (s), broad singlet (br s), doublet (d), triplet (t), quartet (q), quintuplet (quin), sextuplet (sex), septet (sept), multiplet (m), or combinations thereof. LCMS (liquid chromatography mass spectrometry) analyses were performed with either an Agilent HPLC 1100 series connected to a Bruker Daltonics MicrOTOF or an Agilent Technologies 1200 series HPLC connected to an Agilent Technologies 6130 quadrupole LCMS, and both instruments were connected to an Agilent diode array detector. LCMS chromatographic separations were conducted with a Phenomenex Gemini C18 column, 50 mm × 3.0 mm, 5 μm particle size; mobile phase/acetonitrile +0.1% HCOOH 80:20 to 5:95 over 3.5 min, and then held for 1.5 min; flow rate 0.5 mL min<sup>-1</sup>. High resolution electrospray measurements (HRMS) were performed with a Bruker Daltonics MicrOTOF mass spectrometer. Thin layer chromatography (TLC) was carried out on Merck silica gel 60 F254 plates using UV light and/or KMnO<sub>4</sub> for visualization. Column chromatography was performed using RediSep 4 or 12 g silica prepacked columns. When applicable, all glassware was oven-dried overnight and all reactions were carried out under dry and inert conditions (Argon atmosphere).

All in this work synthesized compounds had a measured purity of greater than 95% (measured on analytical HPLC-MS system). M<sup>+</sup> data are given below to substantiate the purity and integrity of the compounds. <sup>1</sup>H NMR, <sup>13</sup>C NMR, and HRMS experiments were also used to confirm compound identity and purity.

**N-((1*r*,4*r*)-4-Methoxycyclohexyl)-4-nitro-1*H*-pyrazole-3-carboxamide (2).** A mixture of 4-nitro-3-pyrazolecarboxylic acid (1) (2.33 g, 14.8 mmol), *trans*-4-methoxy-cyclohexylamine (2.39 g, 18.5 mmol), EDC (3.55 g, 18.5 mmol), and HOBt (2.50 g, 18.5 mmol) in DMF (75 mL) was stirred at ambient temperature for 16 h. The mixture was reduced in vacuo and partitioned between saturated aqueous sodium bicarbonate and EtOAc. The organic layer was washed (water, brine), dried (MgSO<sub>4</sub>), and reduced in vacuo to give a yellow oil, which was purified by column chromatography, eluting 0–100% EtOAc in petroleum ether to give 2. Yield: 3.12 g (solid), 62%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 14.02 (s, 1H), 8.73 (s, 1H), 8.57 (d, J = 7.81 Hz, 1H), 3.74 (m, 1H), 3.24 (s, 3H), 3.11 (m, 1H), 1.95 (dd, J = 55.8, 10.9 Hz, 4H), 1.27 (m, 4H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 159.04, 141.49, 132.17, 131.44, 77.40, 55.01, 47.58, 29.67, 29.30. LRMS (ES<sup>+</sup>): m/z 269 [M + H]<sup>+</sup>.

**4-Amino-N-(4-methoxycyclohexyl)-1*H*-pyrazole-3-carboxamide (3).** A solution of 2 (1.13 g, 4.2 mmol) in DMF (100 mL) was treated with 10% palladium on carbon then shaken under hydrogen at room temperature and atmospheric pressure for 5 h. The reaction mixture was diluted with EtOAc, filtered through Celite, washing with further EtOAc, and the filtrate reduced in vacuo to give crude 3 as brown oil. Yield: 982 mg, 98%. <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ (ppm) 7.23 (s, 1H), 3.84 (m, 1H), 3.37 (s, 3H), 3.23 (m, 1H), 2.07 (dd, J = 45.8, 11.2 Hz, 4H), 1.38 (m, 4H). <sup>13</sup>C NMR (MeOD-*d*<sub>4</sub>) δ (ppm) 165.58, 134.15, 133.13, 118.22, 79.72, 56.15, 48.58, 31.37. LRMS (ES<sup>+</sup>): m/z 239 [M + H]<sup>+</sup>.

**General Method for Variation of Substituent R<sup>1</sup>: Example 4-Benzamido-N-(4-methoxycyclohexyl)-1*H*-pyrazole-3-carboxamide (4f).** A mixture of benzoic acid (0.051 g, 0.42 mmol), 3 (0.1 g, 0.42 mmol), EDC (0.096 g, 0.5 mmol), and HOBt (0.068 g, 0.5 mmol) in DMF (10 mL) was stirred at ambient temperature for 16 h. The mixture was reduced in vacuo and partitioned between saturated aqueous sodium bicarbonate and EtOAc. The organic layer was washed (water, brine), dried (MgSO<sub>4</sub>), and reduced in vacuo to give a creamy solid 4f, which was purified by column chromatography. Evaporation of the appropriate fraction yielded the desired compound as an amorphous solid. Yield: 39 mg, 27%. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ (ppm) 10.65 (s, 1H), 8.52 (s, 1H), 8.01 (d, J = 7.2 Hz, 2H), 7.57 (t, J = 7.2 Hz, 1H), 7.51 (m, 2H), 6.86 (d, J = 8.3 Hz, 1H), 4.01 (m, 1H), 3.39 (s, 3H), 3.22 (m, 1H), 2.16 (m, 4H), 1.42 (m, 4H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ (ppm) 164.36, 163.21, 133.54, 133.28, 132.03, 128.83, 127.24, 123.74, 120.82, 78.13, 55.91, 47.48, 30.69, 30.11. LRMS (ES<sup>+</sup>): m/z 343 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>18</sub>H<sub>23</sub>N<sub>4</sub>O<sub>3</sub> [M + H]<sup>+</sup> 343.1765, found 343.1751.

**4-(4,4-Difluorocyclohexanecarboxamido)-N-(4-methoxycyclohexyl)-1*H*-pyrazole-3-carboxamide (4a).** Yield: 90 mg (solid) 56%. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ (ppm) 9.86 (s, 1H), 8.35 (s, 1H), 6.83 (d, J = 8.3 Hz, 1H), 3.95 (m, 1H), 3.39 (s, 3H), 3.21 (m, 1H), 2.44 (m, 1H), 2.22 (m, 2H), 2.11 (m, 6H), 1.94 (m, 2H), 1.81 (m, 2H), 1.40 (m, 4H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ (ppm) 171.61, 163.21, 133.28, 123.22, 122.55 (t, J = 239.4 Hz), 120.71, 78.07, 55.91, 47.50, 42.74, 32.85 (t, J = 23.5 Hz), 30.35, 25.73. LRMS (ES<sup>+</sup>): m/z 385 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>18</sub>H<sub>27</sub>F<sub>2</sub>N<sub>4</sub>O<sub>3</sub> [M + H]<sup>+</sup> 385.2046, found 385.2036.

**N-(4-Methoxycyclohexyl)-4-(tetrahydro-2*H*-pyran-4-carboxamido)-1*H*-pyrazole-3-carboxamide (4b).** Yield: 94 mg (solid), 64%. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ (ppm) 9.86 (s, 1H), 8.34 (s, 1H), 6.81 (d, J = 8.1 Hz, 1H), 4.07 (m, 2H), 3.95 (m, 1H), 3.47 (m, 2H), 3.39 (s, 3H), 3.20 (m, 1H), 2.59 (m, 1H), 2.13 (d, J = 11.0 Hz, 4H), 1.91 (m, 4H), 1.40 (m, 4H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ (ppm) 171.97, 163.30, 133.22, 123.23, 120.71, 78.12, 67.26, 55.92, 47.50, 42.20, 30.64, 30.12, 29.10. LRMS (ES<sup>+</sup>): m/z 351 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>17</sub>H<sub>27</sub>N<sub>4</sub>O<sub>4</sub> [M + H]<sup>+</sup> 351.2027, found 351.2011.

**4-(2-(2-Fluorophenyl)acetamido)-N-(4-methoxycyclohexyl)-1*H*-pyrazole-3-carboxamide (4c).** Yield: 56 mg (solid), 71%. <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ (ppm) 8.20 (s, 1H), 7.38 (t, 7.7 Hz, 1H), 7.34–7.26 (m, 1H), 7.18–7.05 (m, 2H), 3.83 (m, 1H), 3.80 (s, 2H), 3.33 (s, 3H), 3.16 (m, 1H), 2.06 (br d, J = 11.5 Hz, 2H), 1.97 (br d, J = 12.0 Hz, 2H), 1.34 (m, 4H). <sup>13</sup>C NMR (CD<sub>3</sub>OD) δ (ppm) 169.57, 164.14 (d, J = 159.5 Hz), 134.23, 132.98, 130.64 (d, J = 10.9 Hz), 125.75, 123.77, 123.20 (d, J = 18.2 Hz), 121.96, 116.48 (d, J = 21.7 Hz), 79.69, 56.14, 48.68, 37.61, 31.41, 31.27. LRMS (ES<sup>+</sup>): m/z 375 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>19</sub>H<sub>24</sub>FN<sub>4</sub>O<sub>3</sub> [M + H]<sup>+</sup> 375.1827, found 375.1817.

**N-(4-Methoxycyclohexyl)-4-(2-(2-methoxyphenyl)acetamido)-1*H*-pyrazole-3-carboxamide (4d).** Yield: 43 mg (solid), 53%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 13.13 (s, 1H), 9.76 (s, 1H), 8.15 (s, 1H), 8.00 (d, J = 8.3 Hz, 1H), 7.31–7.23 (m, 2H), 7.02 (d, J = 8.0 Hz, 1H), 6.93 (t, J = 7.4 Hz, 1H), 3.82 (s, 3H), 3.79–3.70 (m, 1H), 3.24 (s, 3H), 3.07 (m, 1H), 2.01 (d, J = 11.4 Hz, 2H), 1.79 (d, J = 11.4 Hz, 2H), 1.45 (m, 2H), 1.21 (m, 2H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 167.30, 162.43, 157.02, 132.38, 130.90, 128.59, 123.21, 122.30, 120.44, 119.70, 110.88, 77.70, 55.41, 55.04, 46.70, 38.17, 30.26, 29.73. LRMS (ES<sup>+</sup>): m/z 387 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>20</sub>H<sub>27</sub>N<sub>4</sub>O<sub>4</sub> [M + H]<sup>+</sup> 387.2027, found 387.2028.

**4-(2-(2,6-Dichlorophenyl)acetamido)-N-(4-methoxycyclohexyl)-1*H*-pyrazole-3-carboxamide (4e).** Yield: 63 mg (solid), 81%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 13.18 (br s, 1H), 9.86 (s, 1H), 8.14 (s,

1H), 8.05 (d, *J* = 7.9 Hz, 1H), 7.51 (d, *J* = 8.2 Hz, 2H), 7.36 (t, *J* = 8.2 Hz, 1H), 4.10 (s, 2H), 3.76 (m, 1H), 3.24 (s, 3H), 3.08 (s, 1H), 2.01 (m, 2H), 1.81 (m, 2H), 1.45 (m, 2H), 1.19 (m, 2H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 170.16, 162.31, 135.49, 131.62, 129.68, 128.30, 128.10, 122.23, 122.09, 77.68, 55.04, 46.77, 38.23, 30.25, 29.74. LRMS (ES<sup>+</sup>): *m/z* 425 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>19</sub>H<sub>23</sub>Cl<sub>2</sub>N<sub>4</sub>O<sub>3</sub> [M + H]<sup>+</sup> 425.1142, found 425.1147.

**4-(2-Fluorobenzamido)-N-(4-methoxycyclohexyl)-1H-pyrazole-3-carboxamide (4g).** Yield: 97 mg (solid), 64%. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ (ppm) 10.89 (d, *J* = 12.2 Hz, 1H), 8.45 (s, 1H), 8.07 (m, 1H), 7.44 (m, 1H), 7.22 (m, 1H), 7.13 (m, 1H), 6.73 (d, *J* = 8.3 Hz, 1H), 3.96 (m, 1H), 3.29 (s, 3H), 3.11 (m, 1H), 2.05 (m, 4H), 1.31 (m, 4H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ (ppm) 162.80, 160.84 (d, *J* = 246.9 Hz), 160.53, 133.81 (d, *J* = 10.9 Hz), 131.77, 124.79, 123.19, 121.42, 120.65 (d, *J* = 10.7 Hz), 116.46 (d, *J* = 21.8 Hz), 78.19, 55.88, 47.27, 30.75, 30.14. LRMS (ES<sup>+</sup>): *m/z* 361 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>18</sub>H<sub>22</sub>FN<sub>4</sub>O<sub>3</sub> [M + H]<sup>+</sup> 361.1670, found 361.1645.

**N-(4-Methoxycyclohexyl)-4-(2-(trifluoromethyl)benzamido)-1H-pyrazole-3-carboxamide (4h).** Yield: 67 mg (solid), 78%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 13.40 (s, 1H), 10.22 (s, 1H), 8.32 (s, 1H), 8.29 (d, *J* = 8.6 Hz, 1H), 7.90–7.88 (m, 1H), 7.84–7.81 (m, 1H), 7.78–7.75 (m, 2H), 3.72 (m, 1H), 3.23 (s, 3H), 3.06 (m, 1H), 2.00 (br d, *J* = 12.5 Hz, 2H), 1.76 (br d, *J* = 12.5 Hz, 2H), 1.45 (m, 2H), 1.14 (m, 2H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 163.34, 162.73, 135.06, 133.01, 132.78, 130.73, 128.39, 126.63 (d, *J* = 5.2 Hz), 126.04 (d, *J* = 29.1 Hz), 123.57 (d, *J* = 276.3 Hz), 122.06, 120.30, 77.66, 55.06, 46.97, 30.28, 29.67. LRMS (ES<sup>+</sup>): *m/z* 411 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>19</sub>H<sub>22</sub>F<sub>3</sub>N<sub>4</sub>O<sub>3</sub> [M + H]<sup>+</sup> 411.1639, found 411.1621.

**4-(2-Ethylbenzamido)-N-(4-methoxycyclohexyl)-1H-pyrazole-3-carboxamide (4i).** Yield: 63 mg (solid), 81%. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ (ppm) 10.04 (s, 1H), 8.39 (s, 1H), 7.53 (d, *J* = 7.6 Hz, 1H), 7.39 (t, *J* = 7.6 Hz, 1H), 7.30 (d, 7.7 Hz, 1H), 7.25 (t, *J* = 7.5 Hz, 1H), 6.87 (d, *J* = 8.3 Hz, 1H), 3.92 (m, 1H), 3.36 (s, 3H), 3.16 (m, 1H), 2.90 (q, *J* = 7.6 Hz, 2H), 2.08 (br d, *J* = 10.0 Hz, 4H), 1.35 (m, 4H), 1.26 (t, *J* = 7.6 Hz, 3H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ (ppm) 167.60, 163.29, 142.86, 134.89, 133.19, 130.65, 129.69, 127.15, 126.11, 123.25, 120.94, 78.18, 55.86, 74.45, 30.58, 30.12, 26.44, 15.84. LRMS (ES<sup>+</sup>): *m/z* 371 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>20</sub>H<sub>27</sub>N<sub>4</sub>O<sub>3</sub> [M + H]<sup>+</sup> 371.2078, found 371.2078.

**4-(2-Methoxybenzamido)-N-(4-methoxycyclohexyl)-1H-pyrazole-3-carboxamide (4j).** Yield: 35 mg (solid), 45%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 13.25 (s, 1H), 11.77 (s, 1H), 8.44 (s, 1H), 8.15 (s, 1H), 8.04 (s, 1H), 7.63 (s, 1H), 7.29 (s, 1H), 7.18 (s, 1H), 4.15 (s, 3H), 3.88 (m, 1H), 3.30 (s, 3H), 3.16 (m, 1H), 2.08 (m, 2H), 1.91 (m, 2H), 1.55 (m, 2H), 1.29 (m, 2H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 162.43, 160.94, 157.47, 133.49, 131.33, 122.29, 120.83, 120.59, 120.28, 112.32, 77.73, 56.05, 55.05, 46.72, 30.28, 29.83. LRMS (ES<sup>+</sup>): *m/z* 373 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>19</sub>H<sub>25</sub>N<sub>4</sub>O<sub>4</sub> [M + H]<sup>+</sup> 373.1870, found 373.1873.

**N-(4-Methoxycyclohexyl)-4-(2-(phenylamino)benzamido)-1H-pyrazole-3-carboxamide (4k).** Yield: 41 mg (solid), 45%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 13.32 (br s, 1H), 10.93 (s, 1H), 9.53 (s, 1H), 8.31 (s, 1H), 8.28 (d, *J* = 8.5 Hz, 1H), 7.68 (dd, *J* = 8.0, 1.4 Hz, 1H), 7.43 (m, 1H), 7.33–7.28 (m, 3H), 7.16–7.14 (m, 2H), 7.01–6.96 (m, 2H), 3.82 (m, 1H), 3.35 (s, 1H), 3.25 (s, 3H), 3.09 (m, 1H), 2.03 (br d, *J* = 12.0 Hz, 2H), 1.81 (br d, *J* = 12.0 Hz, 2H), 1.48 (m, 2H), 1.20 (m, 2H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 164.84, 162.97, 144.57, 141.62, 132.81, 129.33, 128.04, 122.32, 122.00, 120.01, 119.63, 119.26, 118.30, 116.49, 77.74, 55.06, 46.95, 30.33, 29.73. LRMS (ES<sup>+</sup>): *m/z* 434 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>24</sub>H<sub>28</sub>N<sub>5</sub>O<sub>3</sub> [M + H]<sup>+</sup> 434.2187, found 434.2165.

**4-(2-((2,3-Dimethylphenyl)amino)benzamido)-N-(4-methoxycyclohexyl)-1H-pyrazole-3-carboxamide (4l).** Yield: 53 mg (solid), 55%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 13.35 (br s, 1H), 10.89 (s, 1H), 9.53 (s, 1H), 8.33 (s, 1H), 7.63 (dd, *J* = 8.1, 1.4 Hz, 1H), 7.34 (m, 1H), 7.11–7.09 (m, 2H), 6.99–6.97 (m, 1H), 6.87 (m, 1H), 6.82 (dd, *J* = 8.4, 1.0 Hz, 1H), 3.84 (m, 1H), 3.35 (s, 1H), 3.25 (s, 3H), 3.09 (m, 1H), 2.29 (s, 3H), 2.14 (s, 3H), 2.03 (br d, *J* = 12.0 Hz, 2H), 1.82 (br d, *J* = 12.0 Hz, 2H), 1.49 (m, 2H), 1.21 (m, 2H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 165.36, 163.08, 147.05, 138.87, 137.74, 132.95, 132.72,

130.25, 127.52, 125.96, 125.74, 122.41, 120.79, 119.98, 117.44, 115.36, 114.63, 77.74, 55.06, 46.95, 30.33, 29.75, 20.25, 13.64. LRMS (ES<sup>+</sup>): *m/z* 462 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>26</sub>H<sub>32</sub>N<sub>5</sub>O<sub>3</sub> [M + H]<sup>+</sup> 462.2500, found 462.2503.

**4-(2,6-Dimethoxybenzamido)-N-(4-methoxycyclohexyl)-1H-pyrazole-3-carboxamide (4m).** Yield: 73 mg, 43% (solid). <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 13.28 (s, 1H), 9.75 (s, 1H), 8.29 (s, 1H), 8.20 (d, *J* = 8.4 Hz, 1H), 7.39 (t, *J* = 8.5 Hz, 1H), 6.75 (d, *J* = 8.5 Hz, 2H), 3.76 (s, 6H), 3.70 (m, 1H), 3.22 (s, 3H), 3.06 (m, 1H), 2.00 (m, 2H), 1.76 (m, 2H), 1.45 (m, 2H), 1.13 (m, 2H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 162.79, 161.21, 156.88, 132.21, 131.08, 122.52, 119.81, 115.03, 104.28, 77.66, 55.81, 55.08, 46.93, 30.29, 29.66. LRMS (ES<sup>+</sup>): *m/z* 403 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>20</sub>H<sub>27</sub>N<sub>4</sub>O<sub>5</sub> [M + H]<sup>+</sup> 403.1976, found 403.1960.

**N-(4-Methoxycyclohexyl)-4-(2,4,6-trimethoxybenzamido)-1H-pyrazole-3-carboxamide (4n).** Yield: 76 mg (solid), 84%. <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ (ppm) 8.36 (s, 1H), 6.17 (s, 2H), 3.84 (s, 3H), 3.82 (s, 6H), 3.35 (s, 3H), 3.22 (m, 1H), 2.08 (m, 4H), 1.36 (m, 4H). <sup>13</sup>C NMR (CD<sub>3</sub>OD) δ (ppm) 164.90, 164.72, 164.35, 160.74, 134.07, 124.36, 122.58, 108.39, 92.09, 79.74, 57.05, 56.89, 56.62, 48.71, 31.61, 31.40. LRMS (ES<sup>+</sup>): *m/z* 433 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>21</sub>H<sub>29</sub>N<sub>4</sub>O<sub>6</sub> [M + H]<sup>+</sup> 433.2082, found 433.2065.

**4-(2,4-Difluorobenzamido)-N-(4-methoxycyclohexyl)-1H-pyrazole-3-carboxamide (4o).** Yield: 114 mg (solid), 72%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 13.24 (s, 1H), 10.89 (d, *J* = 10.2 Hz, 1H), 8.33 (s, 1H), 8.09 (q, *J* = 8.7 Hz, 1H), 8.05 (d, *J* = 8.4 Hz, 1H), 7.39 (m, 1H), 7.24 (m, 1H), 3.82 (m, 1H), 3.25 (s, 3H), 3.10 (m, 1H), 2.02 (d, *J* = 10.3 Hz, 2H), 1.84 (d, *J* = 11.6 Hz, 2H), 1.47 (q, *J* = 12.0 Hz, 2H), 1.22 (q, *J* = 12.5 Hz, 2H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 164.3 (q, *J* = 250.6, 10.9 Hz), 162.71, 160.3 (q, *J* = 250.8, 14.5 Hz), 158.03, 133.26, 133.18, 132.82, 122.12, 120.50, 117.39, 117.31, 112.59, 112.42, 104.96, 104.75, 104.52, 79.12, 78.84, 78.58, 77.71, 55.07, 46.89, 30.09, 29.76. LRMS (ES<sup>+</sup>): *m/z* 379 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>18</sub>H<sub>21</sub>F<sub>2</sub>N<sub>4</sub>O<sub>3</sub> [M + H]<sup>+</sup> 379.1576, found 379.1567.

**4-(3,5-Difluorobenzamido)-N-(4-methoxycyclohexyl)-1H-pyrazole-3-carboxamide (4p).** Yield: 157 mg (solid), 99%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 13.39 (s, 1H), 10.77 (s, 1H), 8.34 (d, *J* = 8.5 Hz, 1H), 8.30 (s, 1H), 7.59 (m, 1H), 7.52 (m, 2H), 3.84 (m, 1H), 3.24 (s, 3H), 3.09 (m, 1H), 2.02 (d, *J* = 10.7 Hz, 2H), 1.81 (d, *J* = 10.7 Hz, 2H), 1.49 (q, *J* = 11.8 Hz, 2H), 1.20 (q, *J* = 11.8 Hz, 2H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 162.8, 162.5 (q, *J* = 250.7, 14.4 Hz), 160.3, 137.16 127.75, 126.59, 124.12, 122.13, 118.94, 110.2 (q, *J* = 22.0, 14.5 Hz), 107.5 (t, *J* = 27.0 Hz), 77.71, 55.06, 46.99, 30.29, 29.76. LRMS (ES<sup>+</sup>): *m/z* 379 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>18</sub>H<sub>21</sub>F<sub>2</sub>N<sub>4</sub>O<sub>3</sub> [M + H]<sup>+</sup> 379.1576, found 379.1568.

**4-(3,5-Dichlorobenzamido)-N-(4-methoxycyclohexyl)-1H-pyrazole-3-carboxamide (4q).** Yield: 147 mg (solid), 85%. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ (ppm) 10.34 (s, 1H), 8.23 (s, 1H), 7.60 (d, *J* = 1.9 Hz, 2H), 7.31 (t, *J* = 1.8 Hz, 1H), 7.05 (s, 1H), 6.58 (d, 1H), 3.77 (m, 1H), 3.15 (s, 3H), 2.98 (m, 1H), 1.91 (m, 4H), 1.18 (m, 4H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 163.12, 161.74, 136.46, 135.73, 131.89, 128.34, 126.03, 125.79, 121.09, 78.10, 55.91, 47.48, 30.69, 30.08. LRMS (ES<sup>+</sup>): *m/z* 411 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>18</sub>H<sub>21</sub>Cl<sub>2</sub>N<sub>4</sub>O<sub>3</sub> [M + H]<sup>+</sup> 411.0985, found 411.0966.

**4-(4-(Difluoromethoxy)benzamido)-N-(4-methoxycyclohexyl)-1H-pyrazole-3-carboxamide (4r).** Yield: 96 mg (solid), 56%. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ (ppm) 10.65 (s, 1H), 8.48 (s, 1H), 8.02 (d, *J* = 8.6 Hz, 2H), 7.24 (d, *J* = 8.6 Hz, 2H), 6.84 (d, *J* = 8.3 Hz, 1H), 6.62 (t, *J* = 7.3 Hz, 1H), 4.00 (m, 1H), 3.39 (s, 3H), 3.22 (m, 1H), 2.15 (m, 4H), 1.42 (m, 4H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ (ppm) 163.28, 163.20, 154.0 (t, *J* = 2.9 Hz), 133.56, 130.28, 129.24, 123.64, 120.74, 119.22, 118.94, 115.47, 112.01, 78.11, 55.94, 47.53, 30.67, 30.12. LRMS (ES<sup>+</sup>): *m/z* 409 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>19</sub>H<sub>23</sub>F<sub>2</sub>N<sub>4</sub>O<sub>4</sub> [M + H]<sup>+</sup> 409.1682, found 409.1649.

**N-(4-Methoxycyclohexyl)-4-(4-(pyrrolidin-1-yl)benzamido)-1H-pyrazole-3-carboxamide (4s).** Yield: 73 mg (solid), 42%. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ (ppm) 10.44 (s, 1H), 8.49 (s, 1H), 7.89 (d, *J* = 8.8 Hz, 2H), 6.85 (d, *J* = 8.3 Hz, 1H), 6.62 (d, *J* = 8.8 Hz, 2H), 4.01 (m, 1H), 3.39 (m, 7H), 3.20 (m, 1H), 2.15 (m, 4H), 2.06 (m, 4H), 1.41 (m, 4H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ (ppm) 164.74, 163.48, 150.26, 133.26,

128.99, 124.26, 120.52, 119.42, 111.20, 78.20, 55.92, 47.66, 47.39, 30.73, 30.18, 25.46. LRMS (ES<sup>+</sup>): *m/z* 412 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>25</sub>H<sub>30</sub>N<sub>5</sub>O<sub>3</sub> [M + H]<sup>+</sup> 412.2343, found 412.2338.

*N*-(4-Methoxycyclohexyl)-4-(4-(4-methylpiperazin-1-yl)-benzamido)-1*H*-pyrazole-3-carboxamide (**4t**). Yield: 37 mg (solid), 20%. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ (ppm) 10.48 (s, 1H), 8.47 (s, 1H), 7.90 (d, *J* = 8.8 Hz, 2H), 6.94 (d, *J* = 8.8 Hz, 2H), 6.83 (m, 1H), 4.00 (m, 1H), 3.39 (s, 3H), 3.37 (t, *J* = 5 Hz, 4H), 3.20 (m, 1H), 2.60 (t, *J* = 5 Hz, 4H), 2.38 (s, 3H), 2.14 (m, 4H), 1.41 (m, 4H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ (ppm) 164.17, 163.44, 153.54, 133.33, 128.81, 124.05, 122.92, 120.61, 114.27, 78.17, 55.92, 54.79, 47.62, 47.41, 46.11, 30.72, 30.17. LRMS (ES<sup>+</sup>): *m/z* 441 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>23</sub>H<sub>33</sub>N<sub>6</sub>O<sub>3</sub> [M + H]<sup>+</sup> 441.2609, found 441.2597.

4-(Benzofuran-2-carboxamido)-*N*-(4-methoxycyclohexyl)-1*H*-pyrazole-3-carboxamide (**4u**). Yield: 134 mg (solid), 83%. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ (ppm) 10.76 (s, 1H), 8.53 (s, 1H), 7.70 (d, *J* = 7.5 Hz, 1H), 7.66 (d, *J* = 8.4 Hz, 1H), 7.59 (d, *J* = 0.9 Hz, 1H), 7.46 (t, *J* = 7.8 Hz, 1H), 7.33 (t, *J* = 7.5 Hz, 1H), 6.85 (d, *J* = 8.4 Hz, 1H), 4.07 (m, 1H), 3.40 (s, 3H), 3.22 (m, 1H), 2.16 (m, 4H), 1.43 (m, 4H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ (ppm) 163.03, 156.21, 155.22, 148.18, 133.76, 129.23, 127.56, 127.24, 123.77, 122.59, 121.20, 112.41, 111.23, 78.15, 55.91, 47.41, 30.74, 30.14. LRMS (ES<sup>+</sup>): *m/z* 383 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>20</sub>H<sub>23</sub>N<sub>4</sub>O<sub>4</sub> [M + H]<sup>+</sup> 383.1714, found 383.1703.

*N*-(3-((4-Methoxycyclohexyl)carbamoyl)-1*H*-pyrazol-4-yl)-pyrazol[1,5-*a*]pyridine-2-carboxamide (**4v**). Yield: 115 mg (solid), 72%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 13.30 (s, 1H), 11.01 (s, 1H), 8.83 (d, *J* = 7.0 Hz, 1H), 8.37 (s, 1H), 8.15 (d, *J* = 8.4 Hz, 1H), 7.82 (d, *J* = 8.9 Hz, 1H), 7.34 (t, *J* = 7.7 Hz, 1H), 7.09 (t, *J* = 6.9 Hz, 1H), 3.85 (m, 1H), 3.26 (s, 3H), 3.11 (m, 1H), 2.04 (d, *J* = 10.3 Hz, 2H), 1.84 (d, *J* = 10.7 Hz, 2H), 1.49 (m, 2H), 1.23 (m, 2H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 162.70, 158.15, 146.85, 141.15, 132.80, 128.99, 124.60, 122.00, 120.08, 119.14, 114.53, 97.79, 77.74, 55.06, 46.95, 30.33, 29.79. LRMS (ES<sup>+</sup>): *m/z* 383 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>19</sub>H<sub>23</sub>N<sub>6</sub>O<sub>3</sub> [M + H]<sup>+</sup> 383.1826, found 383.1812.

*N*-(3-((4-Methoxycyclohexyl)carbamoyl)-1*H*-pyrazol-4-yl)-imidazo[1,2-*a*]pyridine-3-carboxamide (**4w**). Yield: 11 mg (solid), 7%. <sup>1</sup>H NMR (MeOD-*d*<sub>4</sub>) δ (ppm) 9.57 (d, *J* = 6.9 Hz, 1H), 8.31 (d, *J* = 8.3 Hz, 2H), 7.74 (d, *J* = 9.0 Hz, 1H), 7.57 (td, *J* = 6.9, 1.2 Hz, 1H), 7.19 (t, *J* = 6.9, 1.2 Hz, 1H), 3.94 (m, 1H), 3.39 (s, 3H), 3.27 (m, 1H), 2.12 (m, 4H), 1.50 (m, 2H), 1.39 (m, 2H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 162.89, 156.52, 147.37, 136.35, 127.61, 127.41, 122.12, 119.89, 117.57, 117.46, 114.43, 77.71, 55.07, 46.90, 30.30, 29.76. LRMS (ES<sup>+</sup>): *m/z* 383 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>19</sub>H<sub>23</sub>N<sub>6</sub>O<sub>3</sub> [M + H]<sup>+</sup> 383.1826, found 383.1811.

*N*-(3-((4-Methoxycyclohexyl)carbamoyl)-1*H*-pyrazol-4-yl)-2-methylimidazo[1,2-*a*]pyridine-3-carboxamide (**4x**). Yield: 47 mg (solid), 28%. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ (ppm) 10.88 (s, 1H), 10.24 (s, 1H), 9.58 (d, *J* = 6.9 Hz, 1H), 8.51 (s, 1H), 7.67 (d, *J* = 9.0 Hz, 1H), 7.41 (t, *J* = 7.9 Hz, 1H), 6.99 (t, *J* = 6.9 Hz, 1H), 6.80 (d, *J* = 7.9 Hz, 1H), 4.01 (m, 1H), 3.39 (s, 3H), 3.21 (m, 1H), 3.02 (s, 3H), 2.14 (m, 4H), 1.41 (m, 4H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 162.73, 157.27, 146.18, 145.49, 127.61, 127.42, 122.35, 120.04, 116.31, 114.46, 113.62, 77.74, 55.06, 46.77, 30.24, 29.81, 16.52. LRMS (ES<sup>+</sup>): *m/z* 397 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>20</sub>H<sub>25</sub>N<sub>6</sub>O<sub>3</sub> [M + H]<sup>+</sup> 397.1983, found 397.1972.

*N*-(4-Methoxycyclohexyl)-4-(5-phenylfuran-2-carboxamido)-1*H*-pyrazole-3-carboxamide (**4y**). Yield: 107 mg (solid), 62%. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ (ppm) 10.61 (s, 1H), 8.49 (s, 1H), 7.84 (m, 2H), 7.48 (t, *J* = 7.5 Hz, 2H), 7.38 (t, *J* = 7.4 Hz, 1H), 7.31 (t, *J* = 3.6 Hz, 1H), 6.85 (d, *J* = 8.5 Hz, 1H), 6.80 (d, *J* = 3.6 Hz, 1H), 4.07 (m, 1H), 3.40 (s, 3H), 3.22 (m, 1H), 2.16 (m, 4H), 1.42 (m, 4H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ (ppm) 163.03, 156.39, 155.72, 146.33, 133.74, 129.49, 128.94, 128.80, 124.72, 122.85, 120.91, 117.13, 107.30, 78.18, 55.88, 47.25, 30.78, 30.14. LRMS (ES<sup>+</sup>): *m/z* 409 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>22</sub>H<sub>25</sub>N<sub>4</sub>O<sub>4</sub> [M + H]<sup>+</sup> 409.1870, found 409.1832.

*N*-(3-((4-Methoxycyclohexyl)carbamoyl)-1*H*-pyrazol-4-yl)-5-phenyloxazole-4-carboxamide (**4z**). Yield: 71 mg (solid), 41%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 13.30 (s, 1H), 11.20 (s, 1H), 8.68 (s, 1H), 8.38 (s, 1H), 8.30 (m, 2H), 8.14 (d, *J* = 8.4 Hz, 1H), 7.58–7.51 (m, 3H), 3.83 (m, 1H), 3.26 (s, 3H), 3.11 (m, 1H), 2.04 (d, *J* = 10.5 Hz, 2H),

1.84 (d, *J* = 11.0 Hz, 2H), 1.49 (q, *J* = 12.5 Hz, 2H), 1.23 (q, *J* = 11.5 Hz, 2H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 162.64, 157.34, 151.99, 150.43, 132.90, 130.23, 128.50, 128.07, 127.92, 126.62, 121.89, 120.02, 77.74, 55.06, 47.00, 30.33, 29.76. LRMS (ES<sup>+</sup>): *m/z* 410 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>21</sub>H<sub>24</sub>N<sub>5</sub>O<sub>4</sub> [M + H]<sup>+</sup> 410.1823, found 410.1807.

*Methyl 4-Nitro-1*H*-pyrazole-3-carboxylate* (**5**). A 100 mL three-necked round-bottomed flask equipped with a magnetic stirring bar and fitted with a dropping funnel was charged with 4-nitro-1*H*-pyrazole-3-carboxylic acid (4.0 g, 25.5 mmol) and methanol (40 mL). The flask was cooled to 0 °C, and thionyl chloride (2.1 mL, 28.9 mmol) was added to the vigorously stirred solution over a period of 10 min. The mixture was stirred for an additional 12 h at room temperature, after which time TLC indicated complete consumption of the starting acid. The reaction mixture was concentrated under reduced pressure at 40 °C and the residue treated with toluene and reconcentrated (3 × 20 mL) under reduced pressure at 40 °C to give methyl ester **5** as an off-white solid. Yield: 4.42 g, 99%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 14.39 (br s, 1H), 9.98 (s, 1H), 3.90 (s, 3H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 161.15, 138.13, 133.20, 130.90, 52.84. LRMS (ES<sup>+</sup>): *m/z* 172 [M + H]<sup>+</sup>.

*Methyl-4-amino-1*H*-pyrazole-3-carboxylate* (**6**). A 100 mL round-bottomed flask equipped with digital thermometer and stirrer was charged with 10% palladium on carbon (0.621 g) under argon. In a separate vessel, a slurry of methyl ester **5** (4.42 g, 25.8 mmol) in ethanol (45 mL) was warmed to 35 °C to effect dissolution and the solution added to the catalyst under argon. Following a nitrogen–hydrogen purge sequence, an atmosphere of hydrogen was introduced and the reaction mixture maintained at 30 °C until the reaction completion (6 h) was noted by <sup>1</sup>H NMR analysis. Following a purge cycle, the reaction mixture under argon was filtered and the liquors concentrated under reduced pressure to give amine **6** as a solid. Yield: 3.57 g, 98%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 12.83 (br s, 1H), 7.10 (s, 1H), 4.83 (br s, 2H), 3.78 (s, 3H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 160.39, 136.94, 128.43, 115.59, 50.88. LRMS (ES<sup>+</sup>): *m/z* 142 [M + H]<sup>+</sup>.

*Methyl-4-(2,6-dimethoxybenzamido)-1*H*-pyrazole-3-carboxylate* (**7**). A solution of amine **6** (3.57 g, 25.3 mmol) in 1,4-dioxane (50 mL) under argon was treated with triethylamine (4.3 mL, 31 mmol) followed by 2,6-dimethoxybenzoyl chloride (6.13 g, 30.6 mmol) such that the internal temperature was maintained in the range 20–25 °C. The reaction mixture was stirred at 25 °C until the reaction was complete (12 h) by TLC analysis. The reaction mixture was filtered, the filter-cake washed with 1,4-dioxane, and the combined filtrates progressed to next stage without further isolation.

To obtain analytical data for compound **7** and also to determine the yield of this reaction, a 2 g sample was taken out of the homogeneous filtrate solution (total weight of this solution is 91g). The 2 g sample was then concentrated under reduced pressure until dryness. The crude product (~192 mg) was purified by column chromatography (DCM/MeOH). Evaporation of the appropriate fractions yielded finally the desired compound **7** as an amorphous solid (161 mg). Therefore, in the whole filtrate contained 7.33 g of compound **7**. A 5 mg sample was used for to obtain analytical data; the rest was redissolved for use in the next reaction. Yield: 7.33 g, 95%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) 13.68 (br s, 1H), 9.16 (s, 1H), 8.31 (s, 1H), 7.41 (t, *J* = 8.4 Hz, 1H), 6.76 (d, *J* = 8.4 Hz, 2H), 3.83 (s, 3H), 3.77 (s, 6H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 163.86, 161.55, 157.07, 131.27, 129.97, 123.61, 120.41, 114.66, 104.35, 55.84, 51.63. LRMS (ES<sup>+</sup>): *m/z* 306 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>14</sub>H<sub>16</sub>N<sub>3</sub>O<sub>5</sub> [M + H]<sup>+</sup> 306.1084, found 306.1081.

4-(2,6-Dimethoxybenzamido)-1*H*-pyrazole-3-carboxylic Acid (**8**). To a solution of sodium hydroxide (3.32 g, 83 mmol) in water (20 mL) was charged a solution of ester **7** in one portion (7.33 g, 24.0 mmol); the solution of crude **7** from the previous reaction, plus 156 mg of redissolved pure **7**). The reaction mixture was stirred at 25 °C until completion as determined by TLC analysis. The reaction mixture was concentrated under reduced pressure at 45 °C, the oily residue diluted with water and acidified to pH 1 with concentrated hydrochloric acid, such that the temperature was maintained below 30 °C. The resulting

precipitate was collected by filtration, washed with water, pulled dry on the filter, and subsequently washed with heptanes. The filter cake was charged to a 200 mL rotary evaporator flask and drying completed azeotropically with toluene. Yield: 6.22 g, 89%.  $^1\text{H}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 13.44 (br s, 2H), 9.17 (br s, 1H), 8.29 (s, 1H), 7.40 (t,  $J$  = 8.4 Hz, 1H), 6.76 (d,  $J$  = 8.4 Hz, 2H), 3.77 (s, 6H). LRMS (ES $^+$ ):  $m/z$  292 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{13}\text{H}_{14}\text{N}_3\text{O}_5$  [M + H] $^+$  292.0928, found 292.0920.

**General Method for Variation of Substituent R $^2$ : Example N-Cyclohexyl-4-(2,6-dimethoxybenzamido)-1H-pyrazole-3-carboxamide (9c).** A mixture of carboxylic acid (50 mg, 0.17 mmol, 1.2 equiv), amine (14 mg, 0.14 mmol, 1.0 equiv), hydroxybenzotriazole (19 mg, 0.14 mmol, 1.0 equiv), polymer supported-carbodiimide (105 mg, 0.14 mmol, 1.0 equiv), and acetonitrile was heated by microwave irradiation for 10 min at 100 °C. The final product (9c) was isolated from the reaction mixture by filtering through a short column of Si-carbonate under gravity, which scavenged the excess carboxylic acid and hydroxybenzotriazole. No further purification was required. Removal of the solvent under reduced pressure yielded the required compounds as amorphous solids. Yield: 49 mg (solid), 67%.  $^1\text{H}$  NMR (CD $_3$ OD)  $\delta$  (ppm) 8.33 (s, 1H), 7.42 (t,  $J$  = 8.5 Hz, 1H), 6.75 (d,  $J$  = 8.4 Hz, 2H), 3.86 (s, 6H), 3.82 (m, 1H), 1.88 (m, 4H), 1.68 (d,  $J$  = 12.8 Hz, 1H), 1.40 (m, 4H), 1.27 (m, 1H).  $^{13}\text{C}$  NMR (CD $_3$ OD)  $\delta$  (ppm) 165.02, 164.71, 159.24, 134.17, 132.97, 123.89, 121.82, 115.64, 105.24, 56.50, 49.36, 33.77, 26.58, 26.19. LRMS (ES $^+$ ):  $m/z$  373 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{19}\text{H}_{25}\text{N}_4\text{O}_4$  [M + H] $^+$  373.1870, found 373.1850.

**N-Cyclopropyl-4-(2,6-dimethoxybenzamido)-1H-pyrazole-3-carboxamide (9a).** Yield: 42 mg (solid), 64%.  $^1\text{H}$  NMR (CD $_3$ OD)  $\delta$  (ppm) 8.32 (s, 1H), 7.42 (t,  $J$  = 8.4 Hz, 1H), 6.76 (d,  $J$  = 8.4 Hz, 2H), 3.86 (s, 6H), 2.79 (m, 1H), 0.80 (m, 2H), 0.65 (m, 2H).  $^{13}\text{C}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 161.32, 161.20, 131.10, 122.44, 122.30, 115.01, 104.28, 55.82, 22.13, and 22.02 (d, rotamers), 5.55. LRMS (ES $^+$ ):  $m/z$  331 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{16}\text{H}_{19}\text{N}_4\text{O}_4$  [M + H] $^+$  331.1401, found 331.1385.

**N-Cyclobutyl-4-(2,6-dimethoxybenzamido)-1H-pyrazole-3-carboxamide (9b).** Yield: 56 mg (solid), 94%.  $^1\text{H}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 13.28 (s, 1H), 9.71 (s, 1H), 8.60 (d,  $J$  = 8.1 Hz, 1H), 8.30 (s, 1H), 7.40 (t,  $J$  = 8.4 Hz, 1H), 6.74 (d,  $J$  = 8.4 Hz, 2H), 4.37 (sex,  $J$  = 8.3, 1H), 3.76 (s, 6H), 2.13 (m, 4H), 1.63 (m, 2H).  $^{13}\text{C}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 162.53, 161.23, 156.93, 132.14, 131.07, 122.61, 119.84, 115.07, 104.32, 55.82, 54.86, 43.39, 29.81, 14.62. LRMS (ES $^+$ ):  $m/z$  345 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{17}\text{H}_{21}\text{N}_4\text{O}_4$  [M + H] $^+$  345.1557, found 345.1548.

**N-Cycloheptyl-4-(2,6-dimethoxybenzamido)-1H-pyrazole-3-carboxamide (9d).** Yield: 60 mg (solid), 90%.  $^1\text{H}$  NMR (CDCl $_3$ )  $\delta$  (ppm) 12.19 (br s, 1H), 9.94 (s, 1H), 8.41 (s, 1H), 7.29 (t,  $J$  = 8.5 Hz, 1H), 6.98 (d,  $J$  = 8.1 Hz, 1H), 6.57 (d,  $J$  = 8.5 Hz, 2H), 4.05 (m, 1H), 3.80 (s, 6H), 2.00–1.94 (m, 2H), 1.66–1.45 (m, 10H).  $^{13}\text{C}$  NMR (CDCl $_3$ )  $\delta$  (ppm) 161.75, 160.98, 155.95, 131.18, 129.57, 121.16, 119.75, 112.84, 102.24, 54.19, 48.30, 33.28, 26.12, 22.39. LRMS (ES $^+$ ):  $m/z$  387 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{20}\text{H}_{27}\text{N}_4\text{O}_4$  [M + H] $^+$  387.2027, found 387.2043.

**N-(Bicyclo[2.2.1]heptan-2-yl)-4-(2,6-dimethoxybenzamido)-1H-pyrazole-3-carboxamide (9e).** Yield: 49 mg (solid), 52%.  $^1\text{H}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 11.80 (br s, 1H), 9.72 (s, 1H), 8.28 (s, 1H), 8.00 (br d,  $J$  = 6.9 Hz, 1H), 7.40 (t,  $J$  = 8.4 Hz, 1H), 6.75 (d,  $J$  = 8.4 Hz, 2H), 3.76 (s, 6H), 3.66 (m, 1H), 2.21 (br s, 1H), 2.15 (br s, 1H), 1.63–1.37 (m, 5H), 1.16–0.99 (m, 3H).  $^{13}\text{C}$  NMR (CD $_3$ OD)  $\delta$  (ppm) 165.11, 164.82, 159.28, 133.97, 133.00, 124.01, 122.33, 115.79, 105.34, 56.59, 53.91, 43.80, 40.38, 37.06, 36.28, 29.28, 27.54. LRMS (ES $^+$ ):  $m/z$  385 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{20}\text{H}_{25}\text{N}_4\text{O}_4$  [M + H] $^+$  385.1870, found 385.1857.

**4-(2,6-Dimethoxybenzamido)-N-morpholino-1H-pyrazole-3-carboxamide (9f).** Yield: 34 mg (solid), 53%.  $^1\text{H}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 13.30 (s, 1H), 9.66 (s, 1H), 9.55 (s, 1H), 8.31 (s, 1H), 7.39 (t,  $J$  = 8.4 Hz, 1H), 6.75 (d,  $J$  = 8.6 Hz, 2H), 3.76 (s, 6H), 3.62 (m, 4H), 2.84 (m, 4H).  $^{13}\text{C}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 161.26, 161.18, 156.93, 131.41, 131.10, 122.90, 119.80, 115.01, 104.32, 65.98, 55.82, 54.31. LRMS (ES $^+$ ):  $m/z$  376 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{17}\text{H}_{22}\text{N}_5\text{O}_5$  [M + H] $^+$  376.1615, found 376.1620.

**4-(2,6-Dimethoxybenzamido)-N-(3-(dimethylamino)propyl)-1H-pyrazole-3-carboxamide (9g).** Yield: 41 mg (solid), 63%.  $^1\text{H}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 13.15 (brs, 1H), 9.70 (s, 1H), 8.99 (t,  $J$  = 6.3 Hz, 1H), 8.32 (s, 1H), 7.38 (t,  $J$  = 8.4 Hz, 1H), 7.31–7.29 (m, 4H), 7.22 (m, 1H), 6.74 (d,  $J$  = 8.4 Hz, 2H), 4.41 (d,  $J$  = 6.4 Hz, 2H), 3.75 (s, 6H).  $^{13}\text{C}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 163.46, 161.29, 156.93, 139.45, 131.95, 131.10, 128.21, 127.26, 126.71, 122.58, 120.18, 115.01, 104.32, 55.82, 41.66. LRMS (ES $^+$ ):  $m/z$  381 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{20}\text{H}_{21}\text{N}_4\text{O}_4$  [M + H] $^+$  381.1557, found 381.1543.

**4-(2,6-Dimethoxybenzamido)-N-(pyridin-2-ylmethyl)-1H-pyrazole-3-carboxamide (9h).** Yield: 28 mg (solid), 42%.  $^1\text{H}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 13.33 (br s, 1H), 9.66 (s, 1H), 8.96 (t,  $J$  = 5.9, 1H), 8.50 (m, 1H), 8.34 (s, 1H), 7.74 (td,  $J$  = 7.7, 1.9 Hz, 1H), 7.39 (t,  $J$  = 8.4 Hz, 1H), 7.30 (d,  $J$  = 7.8 Hz, 1H), 7.25 (m, 1H), 6.73 (d,  $J$  = 8.4 Hz, 2H), 4.53 (d,  $J$  = 5.6 Hz, 2H), 3.75 (s, 6H).  $^{13}\text{C}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 163.65, 161.32, 158.14, 156.93, 148.75, 136.68, 131.89, 131.10, 122.61, 122.06, 120.84, 120.18, 114.98, 104.32, 55.82, 43.54. LRMS (ES $^+$ ):  $m/z$  382 [M + H] $^+$ .

**4-(2,6-Dimethoxybenzamido)-N-((3-methylpyridin-2-yl)methyl)-1H-pyrazole-3-carboxamide (9i).** Yield: 30 mg (solid), 44%.  $^1\text{H}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 13.30 (br s, 1H), 9.68 (s, 1H), 8.72 (t,  $J$  = 5.0 Hz, 1H), 8.39 (d,  $J$  = 5.0 Hz, 1H), 8.34 (s, 1H), 7.60 (d,  $J$  = 7.6 Hz, 1H), 7.39 (t,  $J$  = 8.4 Hz, 1H), 7.23 (dd,  $J$  = 7.6, 5.0 Hz, 1H), 6.75 (d,  $J$  = 8.4 Hz, 2H), 4.54 (d,  $J$  = 5.0 Hz, 2H), 3.76 (s, 6H), 2.29 (s, 3H).  $^{13}\text{C}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 163.28, 161.29, 156.96, 153.98, 145.81, 137.64, 132.10, 131.13, 130.64, 122.47, 122.24, 120.04, 114.98, 104.32, 55.84, 40.96, 17, 13. LRMS (ES $^+$ ):  $m/z$  396 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{20}\text{H}_{22}\text{N}_5\text{O}_4$  [M + H] $^+$  396.1666, found 396.1656.

**N-(Cyclohexylmethyl)-4-(2,6-dimethoxybenzamido)-1H-pyrazole-3-carboxamide (9j).** Yield: 56 mg (solid), 84%.  $^1\text{H}$  NMR (CD $_3$ OD)  $\delta$  (ppm) 8.35 (s, 1H), 7.37 (t,  $J$  = 8.4 Hz, 1H), 6.71 (d,  $J$  = 8.4 Hz, 2H), 3.82 (s, 6H), 3.17 (d,  $J$  = 7.0 Hz, 2H), 1.77–1.64 (m, 5H), 1.55 (m, 1H), 1.28–1.13 (m, 3H), 1.00–0.89 (m, 2H).  $^{13}\text{C}$  NMR (CD $_3$ OD)  $\delta$  (ppm) 165.64, 165.00, 159.22, 134.15, 132.99, 123.89, 122.10, 115.69, 105.31, 56.58, 46.09, 39.42, 32.00, 27.57, 27.03. LRMS (ES $^+$ ):  $m/z$  386 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{20}\text{H}_{27}\text{N}_4\text{O}_4$  [M + H] $^+$  387.2027, found 387.2008.

**4-(2,6-Dimethoxybenzamido)-N-phenyl-1H-pyrazole-3-carboxamide (9k).** Yield: 48 mg (solid), 76%.  $^1\text{H}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 13.51 (s, 1H), 10.31 (s, 1H), 9.65 (s, 1H), 8.40 (s, 1H), 7.78 (d,  $J$  = 7.7 Hz, 2H), 7.41 (t,  $J$  = 8.5 Hz, 1H), 7.31 (t,  $J$  = 7.8 Hz, 2H), 7.09 (t,  $J$  = 7.3 Hz, 1H), 6.76 (d,  $J$  = 8.5 Hz, 2H), 3.77 (s, 6H).  $^{13}\text{C}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 162.29, 161.40, 157.02, 138.10, 132.32, 131.19, 128.49, 123.81, 123.02, 120.69, 120.37, 114.98, 104.40, 55.88. LRMS (ES $^+$ ):  $m/z$  367 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{19}\text{H}_{19}\text{N}_4\text{O}_4$  [M + H] $^+$  367.1401, found 367.1402.

**4-(2,6-Dimethoxybenzamido)-N-(pyridin-4-yl)-1H-pyrazole-3-carboxamide (9l).** Yield: 10 mg (solid), 16%.  $^1\text{H}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 13.63 (s, 1H), 10.73 (s, 1H), 9.55 (s, 1H), 8.43 (d,  $J$  = 6.4 Hz, 3H), 7.83 (d,  $J$  = 5.5 Hz, 2H), 7.42 (t,  $J$  = 8.3 Hz, 1H), 6.77 (d,  $J$  = 8.4 Hz, 2H), 3.78 (s, 6H). LRMS (ES $^+$ ):  $m/z$  368 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{18}\text{H}_{18}\text{N}_5\text{O}_4$  [M + H] $^+$  368.1353, found 368.1347.

**4-(2,6-Dimethoxybenzamido)-N-(pyridin-2-yl)-1H-pyrazole-3-carboxamide (9m).** Yield: 10 mg (solid), 16%.  $^1\text{H}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 13.61 (s, 1H), 9.63 (s, 1H), 9.49 (s, 1H), 8.42 (s, 1H), 8.38 (br d,  $J$  = 4.6 Hz, 1H), 8.09 (d,  $J$  = 8.4 Hz, 1H), 7.83 (m, 1H), 7.42 (t,  $J$  = 8.4 Hz, 1H), 7.18 (m, 1H), 6.77 (d,  $J$  = 8.5 Hz, 2H), 3.78 (s, 6H).  $^{13}\text{C}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 161.87, 161.55, 157.01, 150.45, 148.29, 138.41, 131.62, 131.30, 123.04, 120.90, 120.15, 114.66, 113.86, 104.35, 55.84. LRMS (ES $^+$ ):  $m/z$  368 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{18}\text{H}_{18}\text{N}_5\text{O}_4$  [M + H] $^+$  368.1353, found 368.1335.

**4-(2,6-Dimethoxybenzamido)-N-((1r,4r)-4-hydroxycyclohexyl)-1H-pyrazole-3-carboxamide (9n).** Yield: 51 mg (solid), 77%.  $^1\text{H}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 13.21 (br s, 1H), 9.75 (s, 1H), 8.28 (s, 1H), 8.09 (d,  $J$  = 8.4 Hz, 1H), 7.39 (t,  $J$  = 8.4 Hz, 1H), 6.75 (d,  $J$  = 8.4 Hz, 2H), 4.53 (br d,  $J$  = 3.8 Hz, 1H), 3.76 (s, 6H), 3.66 (m, 1H), 3.40–3.33 (m, 1H), 1.82 (m, 2H), 1.73 (m, 2H), 1.44 (m, 2H), 1.19 (m, 2H).  $^{13}\text{C}$  NMR (DMSO- $d_6$ )  $\delta$  (ppm) 162.67, 161.25, 156.91, 131.07, 122.56, 115.10, 104.32, 68.11, 55.82, 47.00, 34.19, 30.00. LRMS (ES $^+$ ):  $m/z$  389 [M + H] $^+$ . HRMS (ES $^+$ ): calcd for  $\text{C}_{19}\text{H}_{25}\text{N}_4\text{O}_5$  [M + H] $^+$  389.1819, found 389.1800.

*cis/trans*-(4-(2,6-Dimethoxybenzamido)-*N*-(4-methylcyclohexyl)-1*H*-pyrazole-3-carboxamide) (**9o**). Yield: 49 mg (solid), 74%. <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ (ppm) 8.36 (s, 0.5H), 8.33 (s, 0.5H), 7.39 (t, *J* = 8.5 Hz, 1H), 6.73 (d, *J* = 8.5 Hz, 2H), 4.05 (m, 0.5H), 3.84 (s, 6H), 3.76 (m, 0.5H), 1.96–1.32 (m, 8H), 1.06 (m, 1H), 0.98 (d, *J* = 6.5 Hz, 1.5 H), 0.92 (d, *J* = 6.6 Hz, 1.5H). <sup>13</sup>C NMR (CD<sub>3</sub>OD) δ (ppm) 165.06, 165.02, 164.76, 164.64, 159.25, 134.07, 134.01, 132.97, 123.99, 123.91, 122.30, 122.12, 115.70, 115.66, 105.39, 105.29, 56.54, 49.52, 46.86, 35.15, 33.67, 33.12, 31.53, 31.01, 30.09, 22.63, 21.25. LRMS (ES<sup>+</sup>): *m/z* 387 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>20</sub>H<sub>27</sub>N<sub>4</sub>O<sub>4</sub> [M + H]<sup>+</sup> 387.2027, found 387.2007.

*tert*-Butyl 4-(4-(2,6-Dimethoxybenzamido)-1*H*-pyrazole-3-carboxamido)piperidine-1-carboxylate (**9p**). Yield: 50 mg (solid), 41%. <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ (ppm) 8.35 (s, 1H), 7.37 (t, *J* = 8.4 Hz, 1H), 6.71 (d, *J* = 8.4 Hz, 2H), 4.08–3.99 (m, 3H), 3.82 (s, 6H), 2.90 (m, 2H), 1.89 (m, 2H), 1.51 (m, 2H), 1.46 (s, 9H). <sup>13</sup>C NMR (CD<sub>3</sub>OD) δ (ppm) 165.00, 164.89, 159.22, 156.39, 134.00, 132.99, 124.03, 122.16, 115.72, 105.35, 81.17, 56.61, 47.67, 44.39, and 43.61 (br d, rotamers), 32.69, 28.79. LRMS (ES<sup>+</sup>): *m/z* 474 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>23</sub>H<sub>32</sub>N<sub>5</sub>O<sub>6</sub> [M + H]<sup>+</sup> 474.2347, found 474.2324.

*N*-(1-Benzylpyrrolidin-3-yl)-4-(2,6-dimethoxybenzamido)-1*H*-pyrazole-3-carboxamide (**9q**). Yield: 69 mg (solid), 89%. <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ (ppm) 8.32 (s, 1H), 7.36 (t, *J* = 8.4 Hz, 1H), 7.32–7.22 (m, 5H), 6.68 (d, *J* = 8.4 Hz, 2H), 4.52 (m, 1H), 3.78 (s, 6H), 3.58 (d, *J* = 2.5 Hz, 2H), 2.76 (m, 2H), 2.58 (dd, *J* = 10.0, 4.3, 1H), 2.43 (q, *J* = 8.1 Hz, 1H), 2.27 (m, 1H), 1.75 (m, 1H). <sup>13</sup>C NMR (CD<sub>3</sub>OD) δ (ppm) 165.06, 165.00, 159.22, 139.26, 133.92, 132.99, 130.33, 129.50, 128.49, 124.01, 122.16, 115.67, 105.32, 61.26, 61.17, 56.58, 53.83, 49.19, 32.54. LRMS (ES<sup>+</sup>): *m/z* 450 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>24</sub>H<sub>28</sub>N<sub>5</sub>O<sub>4</sub> [M + H]<sup>+</sup> 450.2136, found 450.2113.

4-(2,6-Dimethoxybenzamido)-*N*-(2-methoxyethyl)-1*H*-pyrazole-3-carboxamide (**9r**). Yield: 39 mg (solid), 65%. <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ (ppm) 8.35 (s, 1H), 7.38 (t, *J* = 8.4 Hz, 1H), 6.72 (d, 8.4 *J* = 8.4 Hz, 2H), 3.83 (s, 6H), 3.53 (s, 4H), 3.37 (s, 3H). <sup>13</sup>C NMR (CD<sub>3</sub>OD) δ (ppm) 165.76, 165.05, 159.26, 134.10, 132.99, 115.71, 105.32, 72.07, 59.03, 56.56, 39.52. LRMS (ES<sup>+</sup>): *m/z* 349 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>16</sub>H<sub>21</sub>N<sub>4</sub>O<sub>5</sub> [M + H]<sup>+</sup> 349.1506, found 349.1504.

4-(2,6-Dimethoxybenzamido)-*N*-pentyl-1*H*-pyrazole-3-carboxamide (**9s**). Yield: 52 mg (solid), 84%. <sup>1</sup>H NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 13.23 (br s, 1H), 9.74 (s, 1H), 8.38 (br s, 1H), 8.29 (s, 1H), 7.39 (t, *J* = 8.5 Hz, 1H), 6.75 (d, *J* = 8.5 Hz, 2H), 3.76 (s, 6H), 3.19 (m, 2H), 1.49 (quint, *J* = 7.2 Hz, 2H), 1.31–1.22 (m, 4H), 0.85 (t, *J* = 7.0 Hz, 3H). <sup>13</sup>C NMR (DMSO-*d*<sub>6</sub>) δ (ppm) 163.39, 161.20, 156.93, 132.25, 131.07, 122.35, 119.79, 115.05, 104.32, 55.82, 37.95, 28.77, 28.57, 21.78, 13.87. LRMS (ES<sup>+</sup>): *m/z* 361 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>18</sub>H<sub>25</sub>N<sub>4</sub>O<sub>4</sub> [M + H]<sup>+</sup> 361.1870, found 361.1868.

4-(2,6-Dimethoxybenzamido)-*N*-(3-isopropoxypropyl)-1*H*-pyrazole-3-carboxamide (**9t**). Yield: 64 mg (solid), 95%. <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ (ppm) 8.35 (s, 1H), 7.38 (t, *J* = 8.5 Hz, 1H), 6.71 (d, *J* = 8.5 Hz, 2H), 3.83 (s, 6H), 3.58 (sept, *J* = 6.1 Hz, 1H), 3.53 (t, *J* = 6.1 Hz, 2H), 3.45 (t, *J* = 6.6 Hz, 2H), 1.83 (quin, *J* = 6.4 Hz, 2H), 1.15 (d, *J* = 6.1 Hz, 6H). <sup>13</sup>C NMR (CD<sub>3</sub>OD) δ (ppm) 165.58, 165.00, 159.22, 134.17, 132.99, 123.82, 121.99, 115.69, 105.32, 73.08, 67.50, 56.61, 38.06, 30.72, 22.49. LRMS (ES<sup>+</sup>): *m/z* 391 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>19</sub>H<sub>27</sub>N<sub>4</sub>O<sub>5</sub> [M + H]<sup>+</sup> 391.1976, found 391.1961.

4-(2,6-Dimethoxybenzamido)-*N*-(3-(dimethylamino)propyl)-1*H*-pyrazole-3-carboxamide (**9u**). Yield: 49 mg (solid), 76%. <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ (ppm) 8.23 (s, 1H), 7.26 (t, *J* = 8.4 Hz, 1H), 6.59 (d, *J* = 8.4 Hz, 2H), 3.71 (s, 6H), 3.26 (t, *J* = 7.5 Hz, 2H), 2.26 (t, *J* = 7.5 Hz, 2H), 2.11 (s, 6H), 1.65 (q, *J* = 7.5 Hz, 2H). <sup>13</sup>C NMR (CD<sub>3</sub>OD) δ (ppm) 165.70, 164.94, 159.22, 134.15, 132.99, 123.92, 122.05, 115.75, 105.35, 58.25, 56.61, 45.52, 38.23, 28.32. LRMS (ES<sup>+</sup>): *m/z* 376 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>18</sub>H<sub>26</sub>N<sub>5</sub>O<sub>4</sub> [M + H]<sup>+</sup> 376.1979, found 376.1965.

4-(2,6-Dimethoxybenzamido)-*N*-(2-(pyrrolidin-1-yl)ethyl)-1*H*-pyrazole-3-carboxamide (**9v**). Yield: 57 mg (solid), 86%. <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ (ppm) 8.34 (s, 1H), 7.38 (t, *J* = 8.4 Hz, 1H), 6.71 (d, *J* = 8.4 Hz, 2H), 3.83 (s, 6H), 3.51 (t, *J* = 6.9 Hz, 2H), 2.69 (t, *J* = 6.9 Hz, 2H), 2.59 (m, 4H), 1.80 (m, 4H). <sup>13</sup>C NMR (CD<sub>3</sub>OD) δ (ppm) 165.69, 164.97, 159.22, 134.09, 132.99, 123.89, 122.01, 115.72, 105.32,

56.58, 56.32, 55.08, 38.64, 24.31. LRMS (ES<sup>+</sup>): *m/z* 388 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>19</sub>H<sub>26</sub>N<sub>5</sub>O<sub>4</sub> [M + H]<sup>+</sup> 388.1979, found 388.1965.

4-(2,6-Dimethoxybenzamido)-*N*-(2-morpholinoethyl)-1*H*-pyrazole-3-carboxamide (**9w**). Yield: 37 mg (solid), 53%. <sup>1</sup>H NMR (CD<sub>3</sub>OD) δ (ppm) 8.34 (s, 1H), 7.40 (t, *J* = 8.4 Hz, 1H), 6.73 (d, *J* = 8.4 Hz, 2H), 3.84 (s, 6H), 3.70 (t, *J* = 4.5 Hz, 4H), 3.5 (t, *J* = 6.6, 2H), 2.57 (t, *J* = 6.6 Hz, 2H), 2.52 (br t, *J* = 4.5 Hz, 4 Hz). <sup>13</sup>C NMR (CD<sub>3</sub>OD) δ (ppm) 165.67, 165.00, 159.23, 134.06, 132.96, 123.86, 122.02, 115.72, 105.32, 67.80, 58.60, 56.55, 54.67, 36.50. LRMS (ES<sup>+</sup>): *m/z* 404 [M + H]<sup>+</sup>. HRMS (ES<sup>+</sup>): calcd for C<sub>19</sub>H<sub>26</sub>N<sub>5</sub>O<sub>5</sub> [M + H]<sup>+</sup> 404.1928, found 404.1929.

## ■ ASSOCIATED CONTENT

### Supporting Information

Kinase profiling of **4f**, **4m**, **4y**, and **9g** against a panel of mammalian kinases; calculated physicochemical properties of all synthesized aminopyrazole derivatives (PDF and PDB). This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Authors

\*For P.G.W.: phone, +44 (0)1382 386231; e-mail, [p.g.wyatt@dundee.ac.uk](mailto:p.g.wyatt@dundee.ac.uk).

\*For R.B.: e-mail, [brenk@uni-mainz.de](mailto:brenk@uni-mainz.de).

### Present Address

§Johannes Gutenberg-Universität Mainz Institut für Pharmazie und Biochemie Staudinger Weg 5 D-55128 Mainz, Germany

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank the Wellcome Trust (grant 077705 and strategic award WT083481) for financial support for these studies. We also thank Iain Collie, Irene Hallyburton, and Bhavya Rao for carrying out the *T. b. brucei* and MRC5 proliferation studies, Daniel James for data management, and Gina McKay for performing HRMS analyses and for assistance with performing other NMR and MS analyses.

## ■ ABBREVIATIONS USED

GSK, glycogen synthase kinase; HAT, human African trypanosomiasis; CDK, cyclin dependent kinase; *T. brucei*, *Trypanosoma brucei*; *T. b. brucei*, *Trypanosoma brucei brucei*; SAR, structure–activity relationships; PK, protein kinase

## ■ REFERENCES

- (1) Pink, R.; Hudson, A.; Mouries, M. A.; Bendig, M. Opportunities and challenges in antiparasitic drug discovery. *Nature Rev. Drug Discovery* **2005**, *4*, 727–740.
- (2) Walton, J. G.; Jones, D. C.; Kiuru, P.; Durie, A. J.; Westwood, N. J.; Fairlamb, A. H. Synthesis and Evaluation of Indatraline-Based Inhibitors for Trypanothione Reductase. *ChemMedChem* **2011**, *6*, 321–328.
- (3) Fairlamb, A. H. Chemotherapy of human African trypanosomiasis: current and future prospects. *Trends Parasitol.* **2003**, *19*, 488–494.
- (4) Alsford, S.; Eckert, S.; Baker, N.; Glover, L.; Sanchez-Flores, A.; Leung, K. F.; Turner, D. J.; Field, M. C.; Berriman, M.; Horn, D. High-throughput decoding of antitrypanosomal drug efficacy and resistance. *Nature* **2012**, *482*, 232–236.
- (5) Allocco, J. J.; Donald, R.; Zhong, T.; Lee, A.; Tang, Y. S.; Hendrickson, R. C.; Liberator, P.; Nare, B. Inhibitors of casein kinase I

block the growth of *Leishmania major* promastigotes in vitro. *Int. J. Parasitol.* **2006**, *36*, 1249–1259.

(6) Alsford, S.; Turner, D. J.; Obado, S. O.; Sanchez-Flores, A.; Glover, L.; Berriman, M.; Hertz-Fowler, C.; Horn, D. High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Res.* **2011**, *21*, 915–924.

(7) Jones, N. G.; Thomas, E. B.; Brown, E.; Dickens, N. J.; Hammarton, T. C.; Mottram, J. C. Regulators of *Trypanosoma brucei* cell cycle progression and differentiation identified using a kinome-wide RNAi screen. *PLoS Pathog.* **2014**, *10*, e1003886.

(8) Naula, C.; Parsons, M.; Mottram, J. C. Protein kinases as drug targets in trypanosomes and Leishmania. *Biochim. Biophys. Acta* **2005**, *1754*, 151–159.

(9) Barquilla, A.; Saldivia, M.; Diaz, R.; Bart, J.-M.; Vidal, I.; Calvo, E.; Hall, M. N.; Navarro, M. Third target of rapamycin complex negatively regulates development of quiescence in *Trypanosoma brucei*. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14399–14404.

(10) Domenicali Pfister, D.; Burkard, G.; Morand, S.; Renggli, C. K.; Roditi, L.; Vassella, E. A Mitogen-activated protein kinase controls differentiation of bloodstream forms of *Trypanosoma brucei*. *Eukaryotic Cell* **2006**, *5*, 1126–1135.

(11) Mony, B. M.; MacGregor, P.; Ivens, A.; Rojas, F.; Cowton, A.; Young, J.; Horn, D.; Matthews, K. Genome-wide dissection of the quorum sensing signalling pathway in *Trypanosoma brucei*. *Nature* **2014**, *505*, 681–685.

(12) Vassella, E.; Kramer, R.; Turner, C. M.; Wankell, M.; Modes, C.; van den Bogaard, M.; Boshart, M. Deletion of a novel protein kinase with PX and FYVE-related domains increases the rate of differentiation of *Trypanosoma brucei*. *Mol. Microbiol.* **2001**, *41*, 33–46.

(13) Ojo, K. K.; Gillespie, J. R.; Riechers, A. J.; Napuli, A. J.; Verlinde, C. L.; Buckner, F. S.; Gelb, M. H.; Domostoj, M. M.; Wells, S. J.; Scheer, A.; Wells, T. N.; Van Voorhis, W. C. Glycogen synthase kinase 3 is a potential drug target for African trypanosomiasis therapy. *Antimicrob. Agents Chemother.* **2008**, *52*, 3710–3717.

(14) Hoefflich, K. P.; Luo, J.; Rubie, E. A.; Tsao, M. S.; Jin, O.; Woodgett, J. R. Requirement for glycogen synthase kinase-3beta in cell survival and NF-kappaB activation. *Nature* **2000**, *406*, 86–90.

(15) Kugimiya, F.; Kawaguchi, H.; Ohba, S.; Kawamura, N.; Hirata, M.; Chikuda, H.; Azuma, Y.; Woodgett, J. R.; Nakamura, K.; Chung, U. I. GSK-3beta controls osteogenesis through regulating Runx2 activity. *PLoS One* **2007**, *2*, e837.

(16) Khanfar, M. A.; Hill, R. A.; Kaddoumi, A.; El Sayed, K. A. Discovery of novel GSK-3beta inhibitors with potent in vitro and in vivo activities and excellent brain permeability using combined ligand- and structure-based virtual screening. *J. Med. Chem.* **2010**, *53*, 8534–8545.

(17) Meijer, L.; Flajolet, M.; Greengard, P. Pharmacological inhibitors of glycogen synthase kinase 3. *Trends Pharmacol. Sci.* **2004**, *25*, 471–480.

(18) Uno, Y.; Iwashita, H.; Tsukamoto, T.; Uchiyama, N.; Kawamoto, T.; Kori, M.; Nakanishi, A. Efficacy of a novel, orally active GSK-3 inhibitor 6-Methyl-N-[3-[[3-(1-methylethoxy)propyl]-carbamoyl]-1H-pyrazol-4-yl]pyridine-3-carboxamide in tau transgenic mice. *Brain Res.* **2009**, *1296*, 148–163.

(19) Wyatt, P. G.; Woodhead, A. J.; Berdini, V.; Boulstridge, J. A.; Carr, M. G.; Cross, D. M.; Davis, D. J.; Devine, L. A.; Early, T. R.; Feltell, R. E.; Lewis, E. J.; McMenamin, R. L.; Navarro, E. F.; O'Brien, M. A.; O'Reilly, M.; Reule, M.; Saxty, G.; Seavers, L. C.; Smith, D. M.; Squires, M. S.; Trewartha, G.; Walker, M. T.; Woolford, A. J. Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (AT7519), a novel cyclin dependent kinase inhibitor using fragment-based X-ray crystallography and structure based drug design. *J. Med. Chem.* **2008**, *51*, 4986–4999.

(20) Wyatt, P. G.; Berdini, V.; Gill, A. L.; Trewartha, G.; Woodhead, A. J.; Navarro, E. F.; O'Brien, M. A.; Phillips, T. R. Preparation of pyrazole derivatives for use in pharmaceutical compositions for the inhibition of cyclin dependent kinases and glycogen synthase kinases. Patent Application WO2006077419 A1, 2006.

(21) Hofer, A.; Ekanem, J. T.; Thelander, L. Allosteric regulation of *Trypanosoma brucei* ribonucleotide reductase studied in vitro and in vivo. *J. Biol. Chem.* **1998**, *273*, 34098–34104.

(22) Urbaniak, M. D.; Mathieson, T.; Bantscheff, M.; Eberhard, D.; Grimaldi, R.; Miranda-Saavedra, D.; Wyatt, P.; Ferguson, M. A.; Frearson, J.; Drewes, G. Chemical proteomic analysis reveals the drugability of the kinome of *Trypanosoma brucei*. *ACS Chem. Biol.* **2012**, *7*, 1858–1865.

(23) Oduor, R. O.; Ojo, K. K.; Williams, G. P.; Bertelli, F.; Mills, J.; Maes, L.; Pryde, D. C.; Parkinson, T.; Van Voorhis, W. C.; Holler, T. P. *Trypanosoma brucei* glycogen synthase kinase-3, a target for anti-trypanosomal drug development: a public–private partnership to identify novel leads. *PLoS Negl. Trop. Dis.* **2011**, *5*, e1017.

(24) Ojo, K. K.; Arakaki, T. L.; Napuli, A. J.; Inampudi, K. K.; Keyloun, K. R.; Zhang, L.; Hol, W. G.; Verlinde, C. L.; Merritt, E. A.; Van Voorhis, W. C. Structure determination of glycogen synthase kinase-3 from *Leishmania major* and comparative inhibitor structure–activity relationships with *Trypanosoma brucei* GSK-3. *Mol. Biochem. Parasitol.* **2011**, *176*, 98–108.

(25) Woodland, A.; Grimaldi, R.; Luksch, T.; Cleghorn, L. A.; Ojo, K. K.; Van Voorhis, W. C.; Brenk, R.; Frearson, J. A.; Gilbert, I. H.; Wyatt, P. G. From on-target to off-target activity: identification and optimization of *Trypanosoma brucei* GSK3 inhibitors and their characterisation as anti-*Trypanosoma brucei* drug discovery lead molecules. *ChemMedChem* **2013**, *8*, 1127–1137.

(26) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680.

(27) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.

(28) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.

(29) Gong, L.; Hirschfeld, D.; Tan, Y. C.; Heather Hogg, J.; Peltz, G.; Avnur, Z.; Dunten, P. Discovery of potent and bioavailable GSK-3beta inhibitors. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 1693–1696.

(30) *The PyMOL Molecular Graphics System*, Version 0.99; Schrödinger, LLC.

(31) Eastwood, B. J.; Farmen, M. W.; Iversen, P. W.; Craft, T. J.; Smallwood, J. K.; Garbison, K. E.; Delapp, N. W.; Smith, G. F. The minimum significant ratio: a statistical parameter to characterize the reproducibility of potency estimates from concentration-response assays and estimation by replicate-experiment studies. *J. Biomol. Screening* **2006**, *11*, 253–261.

(32) Bain, J.; Plater, L.; Elliott, M.; Shpiro, N.; Hastie, C. J.; McLauchlan, H.; Klevernic, I.; Arthur, J. S.; Alessi, D. R.; Cohen, P. The selectivity of protein kinase inhibitors: a further update. *Biochem. J.* **2007**, *408*, 297–315.

(33) Raz, B.; Iten, M.; Grether-Buhler, Y.; Kaminski, R.; Brun, R. The Alamar Blue assay to determine drug sensitivity of African trypanosomes (*T. b. rhodesiense* and *T. b. gambiense*) in vitro. *Acta Trop.* **1997**, *68*, 139–147.

DOI: 10.1002/cmdc.201300072

# 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

Andrew Woodland,<sup>[a]</sup> Raffaella Grimaldi,<sup>[a]</sup> Torsten Luksch,<sup>[a]</sup> Laura A. T. Cleghorn,<sup>[a]</sup> Kayode K. Ojo,<sup>[b]</sup> Wesley C. Van Voorhis,<sup>[b]</sup> Ruth Brenk,<sup>[a]</sup> Julie A. Frearson,<sup>[a]</sup> Ian H. Gilbert,<sup>\*[a]</sup> and Paul G. Wyatt<sup>\*[a]</sup>

Human African trypanosomiasis (HAT) is a life-threatening disease with approximately 30 000–40 000 new cases each year. *Trypanosoma brucei* protein kinase GSK3 short (*Tb*GSK3) is required for parasite growth and survival. Herein we report a screen of a focused kinase library against *T. brucei* GSK3. From this we identified a series of several highly ligand-effi-

cient *Tb*GSK3 inhibitors. Following the hit validation process, we optimised a series of diaminothiazoles, identifying low-nanomolar inhibitors of *Tb*GSK3 that are potent in vitro inhibitors of *T. brucei* proliferation. We show that the *Tb*GSK3 pharmacophore overlaps with that of one or more additional molecular targets.

## Introduction

Human African trypanosomiasis (HAT), also known as African sleeping sickness, is a parasitic disease caused by protozoan parasites of the species *Trypanosoma brucei* and is fatal if untreated. HAT is endemic in certain regions of sub-Saharan Africa, with around 50 million people at risk of infection across 25 countries. The number of reported cases of HAT has fallen recently and is now at about 10 000 reported new cases per year; however, the actual number of cases is estimated to be much higher (30 000–40 000 new cases per year).<sup>[1–3]</sup>

Following infection by the bite of a tsetse fly, patients initially suffer from phase 1 disease, in which they experience episodes of fever, headache, sweating, and swelling of the lymph nodes. Phase 2 disease results from the spread of infection into the central nervous system (CNS). Patients begin to experience a disturbance in their circadian rhythm, resulting in bouts of fatigue alternating with manic periods, which progress to

daytime slumber and nighttime insomnia, with progressive mental deterioration leading to coma and death. Generally the disease is diagnosed only when it has already progressed to the phase 2 CNS stage.


HAT is a neglected disease, because despite millions of people being under the threat of infection, there is no commercial market to justify funding drug development. There are only two stand-alone drugs available for the treatment of late-stage sleeping sickness: melarsoprol and eflornithine. However, both drugs have serious limitations such as toxicity, complex parenteral administration, which is poorly suited to a rural African setting, low and variable brain penetration, the development of resistant parasites,<sup>[4]</sup> and patient compliance.<sup>[5]</sup> A combination therapy of nifurtimox and eflornithine was recently approved for the treatment of stage 2 HAT primarily due to a cost benefit and improved convenience of the new treatment over eflornithine alone. Unfortunately, resistance to nifurtimox develops rapidly in the laboratory.<sup>[6–8]</sup>


In recent years a number of drug development initiatives funded by foundations and/or governments have begun to address the need for improved drugs to treat stage 2 HAT.<sup>[9]</sup> Two new oral clinical candidates were recently developed: fexinidazole,<sup>[10]</sup> a nitroimidazole derivative that is currently in clinical development, and SCYX-7158,<sup>[11]</sup> a benzoxaborole derivative that has been selected for entry into clinical development. However, owing to the high rates of attrition in drug discovery and the requirement for multiple drugs to combat the development of resistant parasites, the pipeline must be further enhanced.

There is a lack of validated drug discovery targets and lead compounds for HAT and other neglected diseases.<sup>[12]</sup> Protein kinases have been explored as possible targets for HAT, as they

[a] Dr. A. Woodland, R. Grimaldi, Dr. T. Luksch, Dr. L. A. T. Cleghorn, Dr. R. Brenk, Prof. J. A. Frearson, Prof. I. H. Gilbert, Prof. P. G. Wyatt  
Drug Discovery Unit (DDU)  
Division of Biological Chemistry and Drug Discovery  
College of Life Sciences, University of Dundee  
Sir James Black Centre, DD1 5EH (UK)  
E-mail: i.h.gilbert@dundee.ac.uk  
p.g.wyatt@dundee.ac.uk

[b] Dr. K. K. Ojo, Prof. W. C. Van Voorhis  
Division of Allergy and Infectious Diseases, Department of Medicine  
University of Washington, Seattle, WA 98195 (USA)

 Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cmdc.201300072>.

 © 2013 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.



play important roles in virtually every cellular event from cell division to stress response.<sup>[13]</sup> Kinases are druggable targets, and crystal structures have been published for many of them.<sup>[14]</sup> Bioinformatics searches of the *T. brucei* genome identified 176 parasite protein kinases,<sup>[15,16]</sup> making this family an attractive source of novel drug discovery targets for the treatment of HAT and other parasitic diseases.<sup>[17–19]</sup>

Human GSK3 $\beta$  (*HsGSK3 $\beta$* ) is involved in the regulation of a vast array of cellular processes in eukaryotes: insulin signaling, growth factors, nutrient levels, cell fates during embryonic development, cell division, apoptosis, and microtubule function.<sup>[20]</sup> *HsGSK3 $\beta$*  has been investigated as a drug target for many diseases, from diabetes to neurodegenerative diseases. To aid development, the crystal structure of *HsGSK3 $\beta$*  has been solved, and high-affinity small-molecule inhibitors of *HsGSK3 $\beta$*  have been developed.<sup>[14,21–24]</sup> Whilst the precise role of the trypanosome orthologue GSK3 short kinase (*TbGSK3*) in the bloodstream form of *T. brucei* has yet to be determined in terms of parasite biology, the importance of this enzyme has been demonstrated by RNA interference experiments that showed decreased growth rates for parasites in vitro culture.<sup>[25,26]</sup>

Herein we report our studies on the identification and optimisation of *TbGSK3* inhibitors with potent antiparasitic activity and highlight their potential for the development of new therapies for the treatment of HAT.

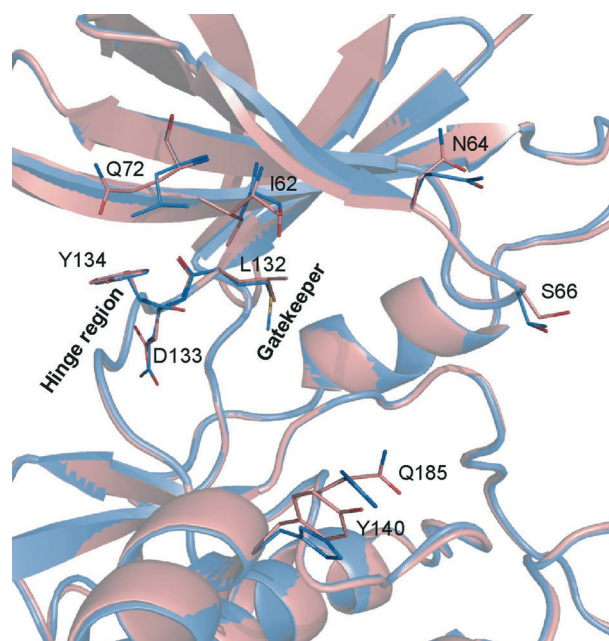
## Results and Discussion

### Homology modelling

Crystal structures of *Leishmania major*<sup>[27]</sup> and human GSK3 have been published. The tertiary structure of *LmGSK3* (PDB code 3E3P) is very similar to that of *HsGSK3 $\beta$*  (PDB code 1R0E), but several binding pocket residues are missing in the *Leishmania* crystal structure, as no clearly defined electron density was present. In addition, no ligand is bound in the *LmGSK3* crystal structure. Therefore, we selected the crystal structure of *HsGSK3 $\beta$*  as the template to build a homology model for *TbGSK3*. The *TbGSK3* sequence is 52% identical and 71% similar to the sequence in the *HsGSK3 $\beta$*  structure (PDB code 1R0E). Hence *HsGSK3 $\beta$*  provides a template for 91% of the *TbGSK3* sequence (amino acids 20–348) which allowed a reliable model to be built (Figure 1). Analysis of the ATP binding pockets revealed amino acid differences that could be exploited to design selective inhibitors (Table 1).

### Hit discovery

Recombinant *TbGSK3* was produced as previously described.<sup>[25]</sup> The kinetic parameters were determined by measuring initial reaction velocities in a matrix experiment of varied ATP and peptide substrate concentrations. The  $K_M$  value of *TbGSK3* for the substrate with sequence YRRAAVPPSPSLSAHSSPHQ[pS]E-DEEE (GSP2) and ATP were  $8.4 \pm 1.3$  and  $11.0 \pm 1.8$   $\mu\text{M}$ , respectively, with no evidence of cooperativity (Figure 2). The deter-



**Figure 1.** Superposition of the *HsGSK3 $\beta$*  crystal structure (PDB code 1R0E) with the homology model of *TbGSK3*. The human structure is shown in pink, the *TbGSK3* homology model in blue; both structures are shown in ribbon representation. Binding pocket residues of *TbGSK3* that differ from those of *HsGSK3 $\beta$*  are represented as sticks. Regions of the kinase site are labelled according to previously published conventions.<sup>[28]</sup>

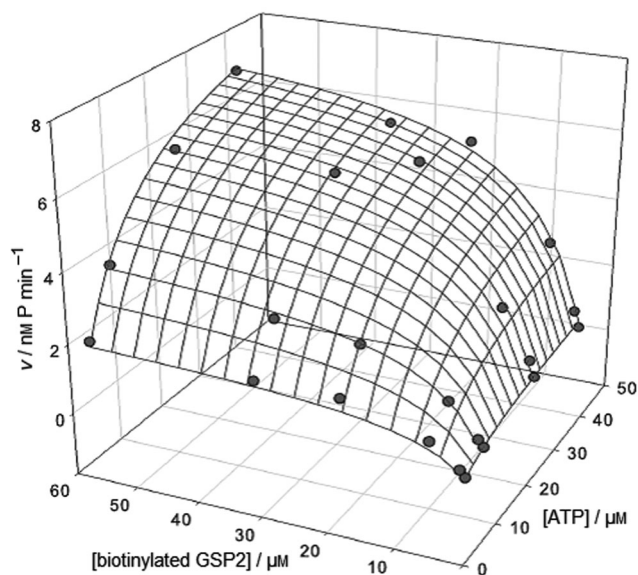
Table 1. Residue differences in the ATP binding sites of <i>HsGSK3<math>\beta</math></i> and <i>TbGSK3</i> .		
Kinase pocket region	<i>HsGSK3<math>\beta</math></i> <sup>[a]</sup>	<i>TbGSK3</i> <sup>[b]</sup>
Hinge region	D133	E
	Y134	F
Gatekeeper	L132	M
	Y140	H
Adenine pocket	Q185	H
	I62	A
G-loop	Q72	L
	N64	Q
	S66	T

[a] PDB code 1R0E. [b] Homology model.

mined  $K_M$  values for GSP2 and ATP are similar to those previously reported of 2.4 and 4.5  $\mu\text{M}$ , respectively.<sup>[25]</sup>

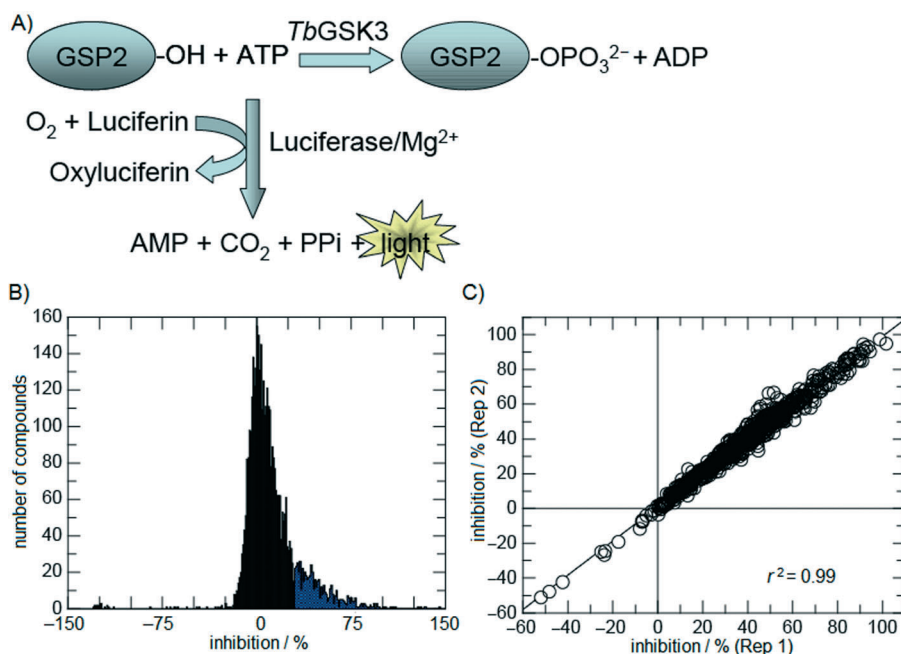
For the primary screen, a 384-well KinaseGlo (Promega) luminescence-based assay was used as previously described.<sup>[25]</sup> In this assay, luminescence is inversely related to kinase activity and directly related to ATP depletion (Figure 3 A). GW8510 was used as a standard inhibitor<sup>[6]</sup> (figure S1, Supporting Information). The DDU kinase set of 4110 compounds<sup>[29]</sup> was screened in single point at 25  $\mu\text{M}$  providing robust data ( $Z' = 0.61 \pm 0.06$ ; Table 2). From this, 567 compounds with a percentage of inhibition  $> 30\%$  were retested in a duplicate-point screen, to give 517 reconfirmed compounds with inhibition values  $\geq 30\%$  (12.8% of the library; Figure 3 B,C).

For hit validation and all subsequent compound potency determinations, a radiometric 96-well Flashplate assay (PerkinElm-



**Figure 2.** Determination of kinetic parameters for *TbGSK3*.  $K_M$  values for the ATP and GSP2 substrates were determined in a time-course matrix experiment in which the initial velocities ( $v$ ) were determined as product formed ( $P$ , nM) per minute. The grid represents the best fit obtained by globally fitting the data to the equation for the random-order rapid equilibrium model with cooperation parameter ( $\alpha$ ) equal to 1.

er) was adopted (Figure 4A). Although the KinaseGlo format has many advantages for screening of chemical libraries, its reliance upon ATP consumption means it requires a level of substrate consumption  $>10\%$  to achieve an acceptable signal window; it is therefore unsuitable for accurate  $IC_{50}$  determina-



**Figure 3.** Format and results of the primary screen. A) A KinaseGlo assay format was adopted for the primary screen based on luminescence detection. B) Distribution of the percent inhibition of the focused kinase set. Hits were selected by setting 3 standard deviation units from the average of high controls as threshold ( $\geq 30\%$  inhibition). C) 567 selected hits were retested in duplicate, and the two replicate values showed high correlation.

Table 2. Assay conditions and screening statistics.		
	KinaseGlo	Flashplate
<i>TbGSK3</i>	7.5 nM	2.5 nM
GSP2	3.2 $\mu$ M	1 $\mu$ M
ATP	1 $\mu$ M ( $< K_M$ )	1 $\mu$ M ( $< K_M$ )
Z'	$0.61 \pm 0.06$	$0.80 \pm 0.08$
GW8510 $IC_{50}$	$10 \pm 0.2$ nM	$6 \pm 4$ nM

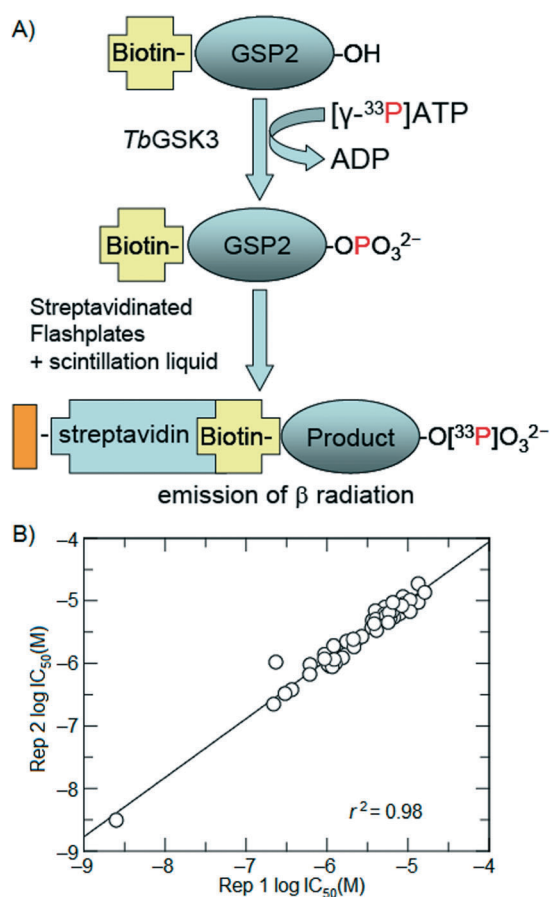
tions. The Flashplate assay was performed at an ATP concentration below the enzyme's  $K_M$  value for ATP, so that the  $K_i^{app}$  approximated the measured  $IC_{50}$  value, aiding the assessment of selectivity. Potency evaluation (10-point curves) was carried out in duplicate for the 100 most potent compounds. As in the primary assay, the potency assay format yielded highly robust data ( $Z' = 0.80 \pm 0.08$ ; Table 2). The compounds exhibited a range of potencies for *TbGSK3* which were highly reproducible, with an  $r^2$  value of 0.99 for two replicates (Figure 4B), with 15 compounds having  $IC_{50}$  values  $<1 \mu$ M. The identity and purity of hits taken into potency determination were confirmed by LC-MS.

#### Prioritisation of hits and series definition

To prioritise the hits, hit compounds were rank ordered by their ligand efficiencies against *TbGSK3*, in which ligand efficiency =  $[-RT \ln(IC_{50})] / N_{non-H atoms}^{[30]}$ . Subsequently, the highest-priority compounds were grouped into series based on structural similarity (Figure 5), resulting in a number of different compound series. Several of the *TbGSK3* inhibitors identified have ligand efficiencies  $>0.4 \text{ kcal mol}^{-1}$  per heavy atom, beyond the commonly used guideline of  $0.3 \text{ kcal mol}^{-1}$  per heavy atom, which relates to an optimised lead with an  $IC_{50}$  value of 10 nM and 38 non-hydrogen atoms ( $M_r \sim 500 \text{ Da}$ ).<sup>[30]</sup> In particular, diaminothiazole **1** is a highly ligand-efficient starting point, with a *TbGSK3* ligand efficiency of  $0.52 \text{ kcal mol}^{-1}$  per heavy atom.

#### Hit validation

Five promising hit series containing ligands with ligand efficiencies  $>0.3 \text{ kcal mol}^{-1}$  per heavy atom were progressed into hit series validation. Series 2 was based around a diaryl urea and exemplified by compound **2**. Compounds of this series were also identified as broad-spectrum kinase inhibitors with toxic-



**Figure 4.** Format and results of the potency test. A) A flashplate radiometric assay format was adopted for the potency determination based on selective capture of the biotinylated phosphorylated product by the streptavidin-coated plates. B) The 100 most potent hits were tested in 10-data point curves in two independent determinations; the correlation plot between the  $\log IC_{50}$  values of the replicates is reported.

ity toward mammalian cell lines during our previously published investigations of the CRK3 kinase, so no further work was conducted with this compound series.<sup>[31]</sup>

Series 3 was based around 2-amino-1,3,5-triazines; 36 examples of this series were contained in the screening library with compound 3 being the only active example (Figure 5). Due to the limited commercial availability for representatives of this series and the apparent requirement for specific substituents, this series was not pursued further.

The 8-aminoimidazo[1,2-*a*]pyrazines (series 4) were also identified in a previous DDU project and were found to be broad-spectrum kinase inhibitors that are toxic to human cell lines (MRC5 cells). In addition, the lead compound 4 has an unfavourable calculated  $\log P$  value of 4.8 (Figure 5). Therefore, no further work was carried out on this series.

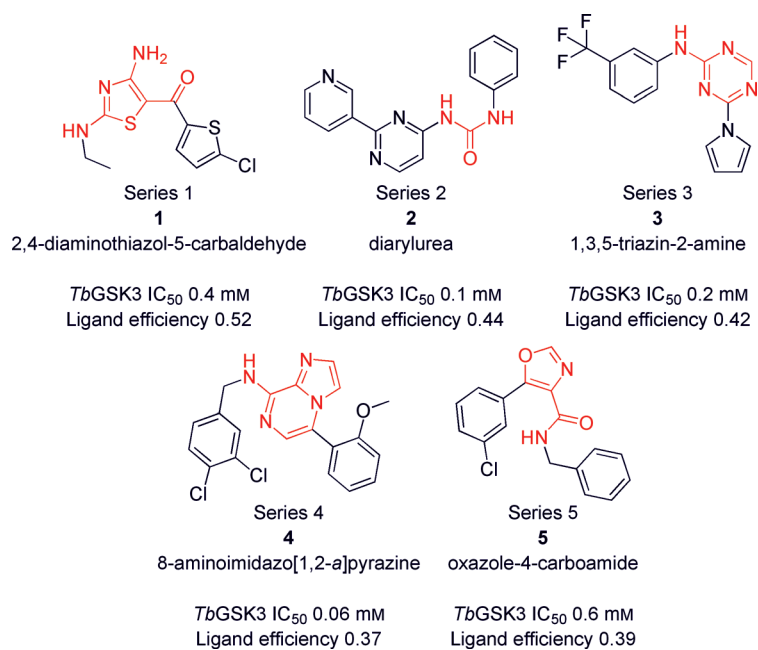
Eleven oxazole-4-carboxamides (series 5) were identified in the high-throughput screen (HTS), with compound 5 inhibiting *TbGSK3* at a sub-micromolar  $IC_{50}$  value (Figure 5). Over 900 examples of this com-

ound series were commercially available at the time of our study, and more than 100 compounds were selected for purchasing and testing. However, none of these compounds demonstrated activity in the *T. brucei* cell assay. This, combined with the relatively poor *TbGSK3*  $IC_{50}$  value of 0.1  $\mu M$  observed within the series after testing more than 140 examples, as well as the flat SAR, it was decided that this series would not be pursued any further.

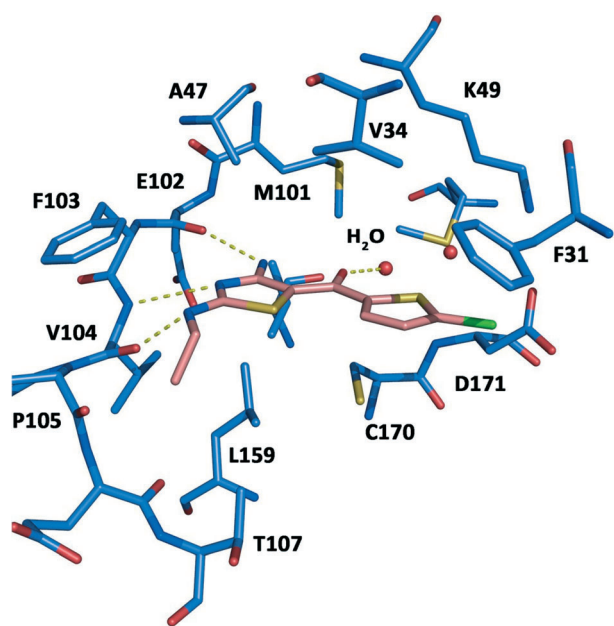
Compound 1 (series 1) was the only 2,4-diaminothiazol-5-carbaldehyde present in the DDU kinase screening set at the time of screening (Figure 5). Although 1 was a singleton, it has good predicted physical properties ( $M_r=288$  Da,  $\log P=2.1$ ,  $TPSA=68$  Å<sup>2</sup>), a ligand efficiency of 0.52 kcal mol<sup>-1</sup> per heavy atom for *TbGSK3*, and demonstrated promising activity in a *T. brucei* proliferation assay ( $EC_{50}$  2  $\mu M$ ). Of slight concern is the presence of a ketone functionality, which has the potential to interact with nucleophiles within the cell; this would have to be monitored during compound development. Based on these considerations, it was decided to progress this compound to hit validation. As a side note, compound 1 is also a very effective *HsGSK3* inhibitor, with an  $IC_{50}$  value of 5 nM ( $n=2$ ) and an outstanding ligand efficiency of 0.67 kcal mol<sup>-1</sup> per heavy atom. Thus, 1 may also be an excellent starting point for a human GSK3 drug discovery programme.

### Structure-guided design

A potential binding mode for compound 1 in the homology model of *TbGSK3* was generated using Moloc (Gerber Molecular Design, Switzerland; Figure 6). In the proposed binding mode the 2,4-diaminothiazole moiety forms three hydrogen bonds with the protein backbone of the *TbGSK3* hinge (Glu102, Phe103, and Val104). Furthermore, the thiazole group



**Figure 5.** Series classification. Common features representative of each series are highlighted in red. Ligand efficiencies are given as kcal mol<sup>-1</sup> per heavy atom.



**Figure 6.** Proposed binding mode for **1** in the homology model of *TbGSK3*. Both, ligand and protein are represented in sticks and colour-coded by atom type. Ligand carbon atoms are shown in salmon, and protein carbon atoms in light blue. Putative hydrogen bonds to the ligand are shown as yellow dotted lines (other hydrogen bonds within the active are not shown for clarity).

is sandwiched between Leu159 and Ala47 and undergoes hydrophobic interactions with both amino acids. In addition, the ligand carbonyl group is proposed to be in plane with the core thiazole and forms an intramolecular hydrogen bond with the primary amino group. A water molecule, observed in many *HsGSK3* $\beta$  structures, was kept for the binding mode generation and is predicted to mediate a hydrogen bond between the inhibitor carbonyl lone pair and the backbone amide NH group of Asp171. The presence of an analogous water molecule in the recently solved X-ray crystal structure of *LmGSK3* (PDB code 3E3P) gives further evidence that the water molecule is conserved and plays an important role in *GSK3*–ligand binding. The 2-chlorothiophene moiety is predicted to form lipophilic interactions with Phe31 and Cys170, while the ethyl group points toward Leu159, and the hydrophobic part of Thr107 potentially contributes van der Waals interactions to the binding affinity.

Docking was carried out in order to guide the hit expansion, with the aim of confirming the aminothiazoles as a series of *TbGSK3* inhibitors, improving in vitro potency, and to build SAR.

A series of 112 commercially available 2,4-diaminothiazoles with various substituents were docked into the homology model of *TbGSK3* using FlexX.<sup>[32]</sup> The docking solutions were visually inspected, and 21 compounds were selected for purchase (Table 3).

### Biological results and structure–activity relationships

The biological data for early hit optimisation are summarised in Table 3. The hit compound **1** has an ethyl group at  $R^1$  which may form hydrophobic interactions with Leu159 and the hydrophobic part of Thr107 (Figure 6, Table 1). Although no direct analogues were available, the closely related methyl analogue **6** is around 50-fold less active than **1**, suggesting that lipophilic bulk in the  $R^1$  region is beneficial. Introducing an aromatic group into  $R^1$  and truncating  $R^2$  to methyl was tolerated, with an  $IC_{50}$  value of 1  $\mu\text{M}$  for **7**. Increasing the size of  $R^1$  while maintaining a lipophilic group at  $R^2$  gave a 10-fold improvement in activity, with several examples demonstrating  $IC_{50}$  values of <100 nM for *TbGSK3*, such as compounds **8**, **9**, and **10**.

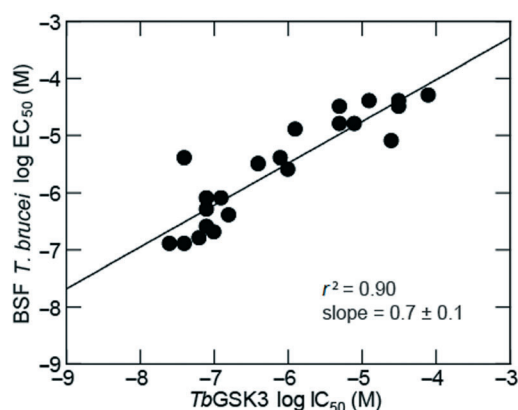
In general the 2,4-diaminothiazol-5-carbaldehydes tested were found to be more potent inhibitors of *HsGSK3* $\beta$  than they are of *TbGSK3*, confirming that selectivity between *Hs* and *TbGSK3* can be achieved, albeit initially in the undesired direction. In contrast, in cellular assays, the lead molecules are selective antiparasitic agents as exemplified by **9**, which has an  $EC_{50}$  value of 0.13  $\mu\text{M}$  against *T. brucei*, but no activity against MRC5 cells ( $EC_{50}$  > 50  $\mu\text{M}$ ).

A correlation plot of cell efficacy (bloodstream form (BSF) *T. brucei* log $EC_{50}$ ) against enzyme potency (*TbGSK3* log $IC_{50}$ ) gave a strong correlation for the early members of this series (correlation coefficient: 0.90) with a 5-fold drop off between the *TbGSK3* and *T. brucei* activities (Figure 7 and Supporting Information table S1). Considering that the physiological level of

**Table 3.** Activity of key compounds from series 1.

Compd			$IC_{50}$ [ $\mu\text{M}$ ]		$EC_{50}$ [ $\mu\text{M}$ ]	
	$R^1$	$R^2$	<i>TbGSK3</i>	<i>HsGSK3</i> $\beta$	<i>T. brucei</i>	MRC5
<b>1</b>			0.4	0.005	2	> 15
<b>6</b>	Me		25	ND <sup>[a]</sup>	7	> 50
<b>7</b>		Me	1	ND <sup>[a]</sup>	13	25
<b>8</b>			0.03	0.007	0.2	13
<b>9</b>			0.05	0.003	0.1	> 50
<b>10</b>			0.1	0.007	0.7	4

[a] Not determined.



**Figure 7.** Correlation between inhibition of *TbGSK3* and inhibition of *T. brucei* cell growth for the initial set of compounds. Supporting Information table S1 lists the compounds used to derive the correlation plots along with the *TbGSK3* log  $IC_{50}$  and *T. brucei* log  $EC_{50}$  values.

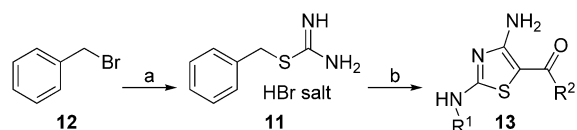
ATP in *T. brucei* is in the millimolar range, whilst in our *TbGSK3* potency assay, the concentration was 1  $\mu\text{M}$ , we expected the  $IC_{50}$  value for *TbGSK3* in cells to drop off by at least 100-fold according to the Cheng–Prusoff equation [Eq. (1)].<sup>[33]</sup> In addition, other factors such as protein binding or the requirement for a high level of enzyme inhibition to achieve a phenotypic effect, as observed previously for other *T. brucei* targets, could even result in a  $>100$ -fold drop off.<sup>[34]</sup> The much lower observed difference between  $IC_{50}$  and  $EC_{50}$  suggested that the mode of action of series 1 may not be just through inhibition of *TbGSK3*.

$$IC_{50} = K_i \left( 1 + \frac{[ATP]}{K_M^{ATP}} \right) \quad (1)$$

The small difference between potency against the enzyme and the cell activity for this series led us to consider that there may be more than one mechanism of action driving the cell activity. Substituted 2,4-diaminothiazoles have been described, and examples are known to be potent inhibitors of *HsGSK3 $\beta$* ,<sup>[35]</sup> several CDKs,<sup>[36,37]</sup> p25,<sup>[37]</sup> and PDE4B.<sup>[38]</sup> Interestingly, there are also reports of a 2,4-diaminothiazole (DAT1) which binds to and disrupts microtubules.<sup>[39]</sup> Homologues of these targets are present in *T. brucei* and could therefore be modulated by compounds of this series. Additionally, prolific kinase inhibitors often show toxicity toward cells in culture. Compound **8** was profiled at 10  $\mu\text{M}$  against the mammalian protein kinase panel at the University of Dundee, which at the time of testing consisted of 76 mammalian kinases. In agreement with our biochemical assays, compound **8** potently inhibited *HsGSK3 $\beta$* ; it also inhibited CDK2, MKK1, ERK8, and HIPK2 by  $>90\%$  at this concentration (Supporting Information table S2). The human protein kinase profile was sufficiently clean for an early-stage kinase inhibitor project, and we decided to monitor the kinase profile as the series was developed.

## Hit-to-lead development

Based on the promising data obtained for the series we decided to progress the project into hit-to-lead development. A stepwise solid-supported two-step synthetic route to diaminothiazoles has been published, starting from the non-commercial reagent, benzyl carbamimidothioate hydrobromide **11**.<sup>[40]</sup> We modified the synthesis so that both steps are carried out in a one-pot, solution-phase reaction (Scheme 1). Benzyl bromide



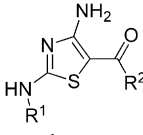



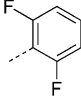

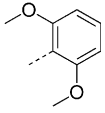

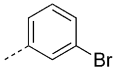






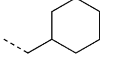

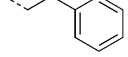



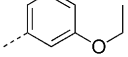
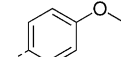
**Scheme 1.** Reagents and conditions: a) thiourea (1 equiv), EtOAc, 120 °C, 5 min, b) isothiocyanate (1.1 equiv), DIPEA (1.2 equiv), DMF, RT, 24 h, then 2-bromoketone (1.2 equiv), DIPEA (2.1 equiv), RT, 1 h.

**12** was treated with thiourea to give benzyl carbamimidothioate hydrobromide **11** in 89% yield. Benzyl carbamimidothioate **11** was allowed to react with an isothiocyanate in the presence of Hünig's base, 2-bromoketones were then added along with an additional equivalent of Hünig's base to give the desired 2,4-diaminothiazoles **13** in a three-component one-pot synthesis with yields ranging from 3 to 69%.

The most potent compounds identified in the early work were 2,4-diamino-5-ketothiazoles **13** bearing aromatic groups in both the  $R^1$  and  $R^2$  position, such as compound **8** ( $IC_{50} = 0.03 \mu\text{M}$ , Table 3). Although we had identified several *TbGSK3* inhibitors with  $IC_{50}$  values less than 1  $\mu\text{M}$ , we wished to improve the selectivity over *HsGSK3 $\beta$* . Our design explored  $R^1$  and  $R^2$  groups of varying size, shape, and polarity to probe the limits of the ATP binding pockets of the *TbGSK3* and *HsGSK3 $\beta$*  enzymes to identify regions where selectivity could be achieved (Table 4).

Initial work focused on the  $R^2$  group. Introduction of *ortho* substituents, which would be expected to twist the  $R^2$  group out of plane with the ketothiazole core, was well tolerated for small groups such the fluorine-substituted analogue **14** ( $IC_{50} = 0.02 \mu\text{M}$ ). The larger and more electron-rich 2,6-dimethoxyphenyl **15** ( $IC_{50} = 0.4 \mu\text{M}$ ) was over 10-fold less active than **14** against *TbGSK3*. Lipophilic bulk in the *meta* position of the  $R^2$  substituent was tolerated, but provided no significant potency gains, with 3-bromo analogue **16** ( $IC_{50} = 0.1 \mu\text{M}$ ) showing similar activity to the phenyl analogue **8** ( $IC_{50} = 0.03 \mu\text{M}$ ).

The  $R^1$  substituent was then investigated. The 2,6-dimethylphenyl derivative **17** ( $IC_{50} = 0.2 \mu\text{M}$ ) is approximately 10-fold less active against *TbGSK3* than **8** ( $IC_{50} = 0.03 \mu\text{M}$ ). Increasing the *meta* and *para* lipophilic bulk at  $R^1$  had no benefit, with 3,4-dimethylphenyl **18** similar to **8**. Replacing the  $R^1$  phenyl ring with cyclohexyl **19** ( $IC_{50} = 0.19 \mu\text{M}$ ) caused a small (sixfold) decrease in potency against the enzyme. Insertion of one or two methylene units between the amine and the cyclohexyl and aromatic groups gave small decreases in activity against *TbGSK3* (**20**,  $IC_{50} = 0.7 \mu\text{M}$ ; **21**,  $IC_{50} = 0.5 \mu\text{M}$ ). As part of our strat-

Table 4. Activity of key compounds from series 1.						
Compd			IC <sub>50</sub> [μM]		EC <sub>50</sub> [μM]	
	R <sup>1</sup>	R <sup>2</sup>	TbGSK3	HsGSK3β	<i>T. brucei</i>	MRC5
8			0.03	0.007	0.2	13
14			0.02	0.005	0.2	1.6
15			0.4	2	0.4	20
16			0.1	0.005	0.2	> 50
17			0.2	0.2	9	35
18			0.06	0.002	0.2	14
19			0.2	0.002	0.1	11
20			0.7	0.02	0.4	41
21			0.5	0.003	6	38
22			12	ND <sup>[a]</sup>	0.3	> 50
23			0.04	0.007	4.1	0.3

[a] Not determined.

egy to test the limits of the *TbGSK3* and *HsGSK3β* pockets, the tetrahedral lipophilic *tert*-butyl ketone **22** was synthesised. This compound is ~300-fold less active than **8** toward *TbGSK3*; however, it is a potent inhibitor of *T. brucei* growth in vitro, with an EC<sub>50</sub> value of 0.3 μM, and is also highly selective (> 150-fold) over the human MRC5 cell line.

### Biological characterisation of 2,4-diaminothiazol-5-carbaldehydes

Studies of the mechanism of inhibition by compounds from the diaminothiazole series (**19** and **14**) confirmed that they are ATP-competitive inhibitors of *TbGSK3*, as evident from the Lineweaver–Burk plots (Figure 8). The calculated  $K_i^{\text{app}}$  values ( $0.25 \pm 0.03$  and  $0.05 \pm 0.01$  μM for **19** and **14**, respectively) correlated very well with the determined IC<sub>50</sub> values in the biochemical assay (Table 4) as expected, considering that the level of ATP in the assay was below the  $K_M$  value for ATP.

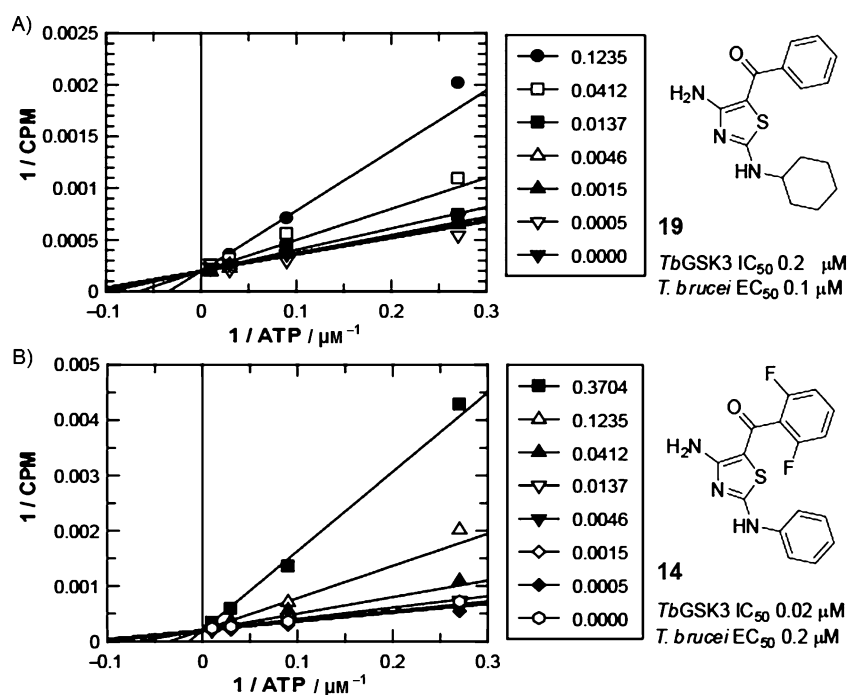
During analysis of the SAR it became apparent that the inhibition of cellular (*T. brucei*) growth was no longer tracking with inhibition of the enzyme, *TbGSK3*. This was particularly apparent when a correlation plot of *T. brucei* logEC<sub>50</sub> values against *TbGSK3* logIC<sub>50</sub> values was produced for all of the compounds generated by this stage in the programme (Figure 9 and Supporting Information table S1). At the extremes of this plot, the ratio of *T. brucei* EC<sub>50</sub> values/*TbGSK3* IC<sub>50</sub> values for **23** is 115, whilst for **22** it is 0.017, a difference of ~6700-fold (Table 4). We believe that the reason for this variation is that an essential molecular target or targets is present in the *T. brucei* parasite which can be modulated by compounds with a similar pharmacophore to that required for *TbGSK3* inhibition.

Earlier in the project we had demonstrated that **8** was not a prolific kinase inhibitor (Table 3 and Supporting Information table S2). To determine whether **22** is a broad-spectrum inhibitor of protein kinases we tested **22** along with **14** and **19** in the University of Dundee kinase panel at a concentration of 10 μM (Supporting Information table S2). Most members of

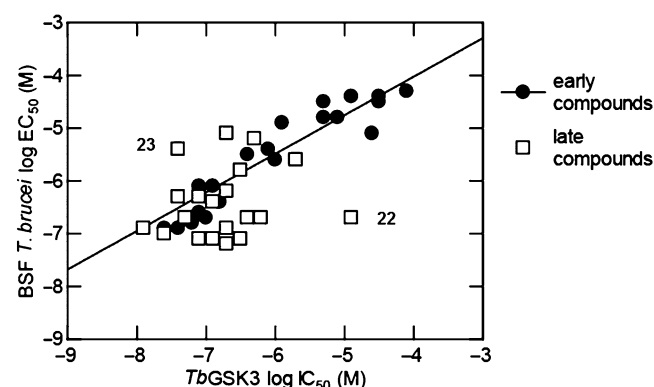
series 1 showed > 90% inhibition against several of the kinases tested. Interestingly, **22** was the least broad-spectrum kinase inhibitor tested and it did not inhibit any kinase at the > 90% level; it only inhibited CDK2 and GSK3 at > 70%. These encouraging data demonstrate that **22** selectively modulates the activity of the unknown parasitic target(s) over a broad range of mammalian kinases, and that these targets can be modulated by drug-like molecules. However, owing to the complex pharmacology of these compounds, and as it is not clear what the off-targets are, the only option will be to optimise these compounds phenotypically, rather than through protein screening.

### Conclusions

In summary, we have shown that a screen of a focused kinase library led to the identification of ligand-efficient inhibitors of *TbGSK3* with sub-micromolar potency. The chemotypes we identified have physicochemical properties consistent with the



**Figure 8.** Mechanism of inhibition was determined for A) 19 and B) 14. Rates (CPM) were determined at the reported inhibitor concentrations ( $\mu\text{M}$ ) with four varied concentrations of ATP at saturating concentration of the other substrate. The resulting Lineweaver–Burk plots were examined for diagnostic patterns for competitive inhibition and globally fitted to the equation for competitive inhibition.



**Figure 9.** Correlation between inhibition of *TbGSK3* and inhibition of *T. brucei* cell growth for the initial set of compounds. The later compounds show weaker correlation. Supporting Information table S1 lists the compounds used to derive the correlation plots along with the *TbGSK3*  $\log \text{IC}_{50}$  and *T. brucei*  $\log \text{EC}_{50}$  values. Figure 9 was generated using both the early and later examples of series 1.

development of orally bioavailable compounds. For the most active series based on 2,4-diaminothiazoles, the compounds were also active against the parasite *T. brucei*. Some of the examples had potencies in the order of 100 nM against the parasite, which represents a very good starting point for a drug discovery programme against HAT. However, it quickly became apparent that this series has one or more molecular targets in addition to *TbGSK3*, which contributes strongly to the trypanocidal activity of the compounds. Interestingly, although based on a kinase scaffold, this series had a reasonably clean profile against the 79 mammalian kinases investigated. In the absence

of a clear understanding of which molecular target(s) are responsible for anti-trypanosomal activity, the only way to optimise this compound is phenotypically. We shall report results for this in due course.

## Experimental Section

**Homology modelling:** Sequence alignments between *TbGSK3* $\beta$  (Tb927.10.13780 in GeneDB) and *HsGSK3* $\beta$  were generated with ClustalW.<sup>[41]</sup> Subsequently, Modeller 9.2<sup>[42]</sup> was used to build homology models of *TbGSK3* short, whereas the human *GSK3* $\beta$  crystal structure (PDB code 1R0E) served as template.<sup>[43]</sup> Modeller was run with default settings, and only the highest-scoring structure was used for further analysis and modelling. The quality of the model was assessed with QMEAN.<sup>[44,45]</sup> A total QMEAN score of 0.726 and Z-score of  $-0.50$  were calculated, indicating that the model is of good quality.

**Generation of putative binding modes:** An initial binding mode for compound 1 was generated by manually docking the ligand into the ATP binding pocket of the *TbGSK3* homology model with the requirement to establish hydrogen bonds to the hinge residues using Moloc. Its position was subsequently optimised with the MAB force field as implemented in Moloc.<sup>[46]</sup> To allow for uncertainties in the homology model, the amino acid side chains facing the binding site were kept flexible during minimisation. For the hit expansion, ligands were docked into the active site of the homology model using FlexX.<sup>[32]</sup> A highly conserved water molecule (H<sub>2</sub>O 82 in 1R0E) was kept in the structure used for docking. For the active site determination, a radius of 13 Å around the ligand bound in 1R0E was selected, always using complete amino acids. Ligand conformations were calculated using CORINA.<sup>[47]</sup> Compounds were docked using default settings.

***TbGSK3* biochemical characterisation:** *TbGSK3* short protein with an N-terminal maltose binding protein fusion was cloned, expressed, and purified as previously reported.<sup>[25]</sup> For biochemical characterisation of *TbGSK3*, the kinase assay buffer (25 mM Tris-HCl pH 7.5, 10 mM MgCl<sub>2</sub>, 5 mM DTT, 0.02% CHAPS, 2 U mL<sup>-1</sup> heparin) and mix of ATP/[ $\gamma$ -<sup>33</sup>P]ATP and GSP2 substrate (YRRRAVPPSPSL-SAHSSPHQ[pS]EDEEE; Pepceuticals) were used in a radiometric format (filterplate assay). The  $K_M$  values for the two substrates (ATP and GSP2) were determined by varying the concentrations of both substrates in a time-course matrix experiment.

**Compound library selection** was described previously.<sup>[29]</sup>

**Primary screen assays:** For the primary screening of the focused kinase inhibitor library, a 384-well KinaseGlo (Promega) luminescence-based assay was used as previously described.<sup>[25]</sup> The reactions contained 7.5 nM *TbGSK3*, 3.2  $\mu\text{M}$  substrate peptide (Pepceuticals), 1  $\mu\text{M}$  ATP, and 25  $\mu\text{M}$  test inhibitor compound in optimised

kinase assay buffer. DMSO and an eight-point titration of GW8510 (Sigma) from 1  $\mu\text{M}$  to 50  $\mu\text{M}$  were included as negative and positive controls. Reactions were incubated at room temperature for 1 h and stopped by the addition of KinaseGlo reagent. Plates were then sealed, and the signal was left to stabilise for 1 h in the dark before luminescence was determined using a TopCount NXT HTS counter (PerkinElmer).

**Potency screen assays:** For hit validation and all subsequent compound potency determinations, a radiometric 96-well Flashplate assay (PerkinElmer) was adopted. Compounds were solubilised in DMSO at a top concentration of 3 mM and serially diluted to achieve 10-point titration of final assay concentrations from 30  $\mu\text{M}$  to 0.3 nM with a final DMSO concentration of 1% (v/v). The reaction mixtures contained 1  $\mu\text{M}$  biotinylated GSP2 substrate, 1  $\mu\text{M}$  ATP, 3.7 KBq [ $\gamma$ - $^{33}\text{P}$ ]ATP per well, and 2.5 nM *TbGSK3* in the *TbGSK3* kinase assay buffer. GSK3 inhibitors were screened for selectivity assessment against *HsGSK3 $\beta$*  as well. For the *HsGSK3 $\beta$*  assay the reaction mixtures contained 1  $\mu\text{M}$  biotinylated GSP2 substrate, 2  $\mu\text{M}$  ATP, 7.4 KBq [ $\gamma$ - $^{33}\text{P}$ ]ATP per well, and 15 nM *HsGSK3 $\beta$*  in the *TbGSK3* kinase assay buffer (25 mM Tris-HCl pH 7.5, 10 mM  $\text{MgCl}_2$ , 5 mM DTT, 0.02% CHAPS, 2  $\text{U mL}^{-1}$  heparin).

**Mammalian kinase profiling:** Selected compounds were screened against a panel of mammalian kinases routinely run by the Division of Signal Transduction Therapy (DSTT) at the University of Dundee in duplicate at 10  $\mu\text{M}$ . Enzymes included in the panel and assay conditions are reported.<sup>[48]</sup> All biochemical assays were carried out below the  $K_M^{\text{ATP}}$  value for ATP for each enzyme, allowing comparison of inhibition across the panel.

**Trypanosome and MRC5 proliferation assay:** Measurement of inhibition of the proliferation of MRC5 (human lung fibroblast) cells and *T. brucei* bloodstream-stage cells was performed using a modification of a cell viability assay previously described.<sup>[49]</sup> Compounds (50  $\mu\text{M}$  to 0.5 nM) were incubated with  $2 \times 10^3$  cells per well in 0.2 mL of the appropriate culture medium (MEM with 10% foetal bovine serum for MRC-5 cells) in clear 96-well plates. Plates were incubated at 37 °C in the presence of 5%  $\text{CO}_2$  for 69 h. Resazurin was then added to a final concentration of 50  $\mu\text{M}$ , and plates were incubated as above for a further 4 h before being read on a BioTek flx800 fluorescent plate reader.

**Data analysis:** Determination of the *TbGSK3* kinetic parameters was carried out as described previously.<sup>[50]</sup>  $\text{IC}_{50}$  values were determined using a four-parameter equation in XLFit 4.2. To establish mode of inhibition, rates were determined at 10 inhibitor concentrations with four varied concentrations of ATP in saturating concentration of GSP2. The resulting Lineweaver–Burk plots were examined for diagnostic patterns for competitive, mixed, or uncompetitive inhibition. Graphs and analyses were carried out using Graft 6.0. The correlation between in vitro  $\text{IC}_{50}$  against *TbGSK3* in *T. brucei* cells and inhibitor potency ( $K_i$ ) for ATP-competitive inhibitors was determined according to the Cheng–Prusoff equation [Eq. (1)].

## Chemistry

$^1\text{H}$  NMR spectra were recorded on a Bruker Avance DPX 500 instrument unless otherwise stated. Chemical shifts ( $\delta$ ) are expressed in ppm. Signal splitting patterns are described as singlet (s), broad singlet (brs), doublet (d), triplet (t), quartet (q), multiplet (m), or combination thereof. Low-resolution electrospray (ES) mass spectra were recorded on a Bruker MicroTof mass spectrometer, run in positive ion mode, using either MeOH, MeOH/ $\text{H}_2\text{O}$  (95:5), or  $\text{H}_2\text{O}/$

$\text{CH}_3\text{CN}$  (1:1)+0.1% formic acid as the mobile phase. High-resolution electrospray MS measurements were performed on a Bruker MicroTof mass spectrometer. LC–MS analyses were performed with an Agilent HPLC 1100 (Phenomenex Gemini Column 5 m  $\text{C}_{18}$  110A 50 $\times$ 3.0 mm, eluting with 20% MeOH/ $\text{H}_2\text{O}$ , 0–3 min) and a diode array detector in series with a Bruker MicroTof mass spectrometer. Column chromatography was performed using RediSep 4 or 12 g silica pre-packed columns.

The following compounds were purchased from commercial suppliers and their purity and identity confirmed by LC–MS analysis (all compounds were >85% pure based on a diode array detector). Compounds **1** and **2** were purchased from ChemBridge Corporation; **3** was purchased from Maybridge; **4** was purchased from BioFocus; **5** was purchased from ChemDiv Inc.; **6**, **9**, and **23** were purchased from TimTec; **7** and **10** were purchased from Specs.

**Benzyl carbamimidothioate hydrobromide (11):** Benzyl bromide (16.0 mL, 135 mmol), thiourea (10.0 g, 131 mmol), and EtOAc (75 mL) were combined, and then heated at 120 °C in a microwave reactor for 5 min. The reaction was allowed to cool to room temperature, and the resulting solid was collected by filtration to give benzyl carbamimidothioate hydrobromide **11** as a white solid (29.3 g, 118.6 mmol, 89% yield);  $^1\text{H}$  NMR (500 MHz,  $[\text{D}_6]\text{DMSO}$ ):  $\delta$  = 9.07 (brs, 4H), 7.44–7.32 (m, 5H), 4.49 ppm (s, 2H).

**4-Amino-2-(phenylamino)thiazol-5-yl(phenyl)methanone (8):** **General procedure A** (used for the synthesis of compounds **8**, **14**–**22**): Benzyl carbamimidothioate hydrobromide (100 mg, 0.4 mmol), phenyl isothiocyanate (52  $\mu\text{L}$ , 0.43 mmol), and *N,N*-diisopropylethylamine (DIPEA; 76  $\mu\text{L}$ , 0.43 mmol) were added to DMF (5 mL), and the resulting mixture was stirred at room temperature for 24 h. 2-Bromoacetophenone (96 mg, 0.48 mmol), DIPEA (139  $\mu\text{L}$ , 0.80 mmol), and DMF (2 mL) were then added, and the mixture was stirred at room temperature for 1 h, after which the reaction was quenched by the addition of aqueous HCl (1 M, 4 mL). The product was extracted with EtOAc (3 $\times$ 3 mL), and the combined extracts were back washed with LiCl (2 $\times$ 3 mL of a 5% w/w solution in  $\text{H}_2\text{O}$ ), brine (3 mL), and then dried with  $\text{MgSO}_4$ , and the solvent was removed under reduced pressure. Purification by column chromatography on silica eluting with petroleum ether (PE) 40–60 °C and EtOAc gave **8** as a yellow solid (47 mg, 0.15 mmol, 39% yield);  $^1\text{H}$  NMR (500 MHz,  $[\text{D}_6]\text{DMSO}$ ):  $\delta$  = 10.80 (brs, 1H), 8.22 (brs, 2H), 7.69–7.67 (m, 2H), 7.62 (d,  $J$  = 7.7 Hz, 2H), 7.51–7.46 (m, 3H), 7.39–7.36 (m, 2H), 7.09 ppm (tt,  $J$  = 7.4 and 1 Hz, 1H);  $^{13}\text{C}$  NMR (125 MHz, DMSO):  $\delta$  = 182.7, 167.2, 165.6, 141.9, 139.5, 130.4, 129.1, 128.4, 126.7, 123.4, 119.0, and 92.2 ppm; LC–MS  $m/z$  = 296  $[\text{M} + \text{H}]^+$ ,  $t_R$  = 4.28 min, purity 88% (Agilent, 20–90%  $\text{CH}_3\text{CN}$ , ES+ on an acidic method).

**(4-Amino-2-(phenylamino)thiazol-5-yl)(2,6-difluorophenyl)methanone (14):** General procedure A gave **14** as a yellow solid (32 mg, 0.10 mmol, 10% yield);  $^1\text{H}$  NMR (500 MHz,  $[\text{D}_6]\text{DMSO}$ ):  $\delta$  = 10.89 (brs, 1H), 8.31 (brs, 2H), 8.08 (brs, 1H), 7.57–7.50 (m, 3H), 7.39–7.35 (m, 2H), 7.23–7.19 (m, 2H), 7.13–7.10 ppm (m, 1H); LC–MS  $m/z$  = 332  $[\text{M} + \text{H}]^+$ ,  $t_R$  = 4.25 min, purity 90% (Agilent, 20–90%  $\text{CH}_3\text{CN}$ , ES+ on an acidic method).

**(4-Amino-2-(phenylamino)thiazol-5-yl)(2,6-dimethoxyphenyl)methanone (15):** General procedure A gave **15** as a yellow solid (9 mg, 0.03 mmol, 3% yield);  $^1\text{H}$  NMR (500 MHz,  $[\text{D}_6]\text{DMSO}$ ):  $\delta$  = 10.58 (brs, 1H), 7.79 (s, 2H), 7.54 (d,  $J$  = 8.0 Hz, 2H), 7.54–7.52 (m, 3H), 7.06 (t,  $J$  = 7.3 Hz, 1H), 6.69 (d,  $J$  = 8.4, 2H), 3.71 ppm (6H, s); LC–MS  $m/z$  = 356  $[\text{M} + \text{H}]^+$ ,  $t_R$  = 4.10 min, purity 100% (Agilent, 20–90%  $\text{CH}_3\text{CN}$ , ES+ on an acidic method).



**(4-Amino-2-(phenylamino)thiazol-5-yl)(3-bromophenyl)methanone (16):** General procedure A gave **16** as a yellow solid (211 mg, 0.56 mmol, 69% yield);  $^1\text{H NMR}$  (500 MHz,  $[\text{D}_6]\text{DMSO}$ ):  $\delta$  = 10.85 (brs, 1H), 8.29 (brs, 2H), 7.81 (t,  $J$  = 1.7 Hz, 1H), 7.71 (ddd,  $J$  = 1.0, 2.0, 8.0 Hz, 1H), 7.69–7.67 (m, 1H), 7.63 (d,  $J$  = 7.8 Hz, 2H), 7.46 (t,  $J$  = 7.8 Hz, 1H), 7.39–7.36 (m, 2H), 7.12–7.08 ppm (m, 1H); LC–MS  $m/z$  = 376  $[M+H]^+$ ,  $t_R$  = 4.58 min, purity 100% (Agilent, 20–90%  $\text{CH}_3\text{CN}$ , ES+ on an acidic method).

**(4-Amino-2-((2,6-dimethylphenyl)amino)thiazol-5-yl)-(phenyl)methanone (17):** General procedure A gave **17** as a yellow solid (35 mg, 0.11 mmol, 11% yield);  $^1\text{H NMR}$  (500 MHz,  $[\text{D}_6]\text{DMSO}$ ):  $\delta$  = 10.11 (brs, 1H), 8.40 (brs, 1H), 7.97 (brs, 1H), 7.54–7.16 (m, 8H), 2.20 ppm (s, 6H); LC–MS  $m/z$  = 324  $[M+H]^+$ ,  $t_R$  = 4.39 min, purity 95.6% (Agilent, 20–90%  $\text{CH}_3\text{CN}$ , ES+ on an acidic method).

**(4-Amino-2-((3,4-dimethylphenyl)amino)thiazol-5-yl)-(phenyl)methanone (18):** General procedure A gave **18** as a yellow solid (100 mg, 0.30 mmol, 30% yield);  $^1\text{H NMR}$  (500 MHz,  $[\text{D}_6]\text{DMSO}$ ):  $\delta$  = 10.60 (brs, 1H), 8.31 (brs, 1H), 8.11 (brs, 1H), 7.67–7.65 (m, 2H), 7.51–7.45 (m, 3H), 7.34 (d,  $J$  = 8.3 Hz, 1H), 7.31 (s, 1H), 7.12 (d,  $J$  = 8.2 Hz, 1H), 2.21 (s, 3H), 2.19 ppm (s, 3H); LC–MS  $m/z$  = 324  $[M+H]^+$ ,  $t_R$  = 4.55 min, purity 88% (Agilent, 20–90%  $\text{CH}_3\text{CN}$ , ES+ on an acidic method); HRMS  $m/z$   $[M+H]^+$  calcd for  $\text{C}_{18}\text{H}_{18}\text{N}_3\text{OS}$ : 324.1165, found: 324.1158.

**(4-Amino-2-(cyclohexylamino)thiazol-5-yl)(phenyl)methanone (19):** General procedure A gave **19** as a yellow solid (59 mg, 0.19 mmol, 19% yield);  $^1\text{H NMR}$  (500 MHz,  $[\text{D}_6]\text{DMSO}$ ):  $\delta$  = 8.57 (brs, 1H), 8.49 (brs, 1H), 7.89 (brs, 1H), 7.63–7.61 (m, 2H), 7.47–7.42 (m, 3H), 3.70 (brs, 1H), 1.91 (d,  $J$  = 10.7 Hz, 2H), 1.76–1.70 (m, 2H), 1.57 (d,  $J$  = 12.9 Hz, 1H), 1.32–1.13 ppm (m, 5H); LC–MS  $m/z$  = 302  $[M+H]^+$ ,  $t_R$  = 4.42 min, purity 88% (Agilent, 20–90%  $\text{CH}_3\text{CN}$ , ES+ on an acidic method); HRMS  $m/z$   $[M+H]^+$  calcd for  $\text{C}_{16}\text{H}_{20}\text{N}_3\text{OS}$ : 302.1322, found: 302.1321.

**(4-Amino-2-((cyclohexylmethyl)amino)thiazol-5-yl)-(phenyl)methanone (20):** General procedure A gave **20** as a yellow solid (136 mg, 0.43 mmol, 43% yield);  $^1\text{H NMR}$  (500 MHz,  $[\text{D}_6]\text{DMSO}$ ):  $\delta$  = 8.63 (brs, 1H), 8.43 (brs, 1H), 7.84 (brs, 1H), 7.63–7.61 (m, 2H), 7.48–7.43 (m, 3H), 3.17–3.03 (m, 2H), 1.17–1.52 (m, 6H), 1.23–1.11 (m, 3H), 0.94–0.87 ppm (m, 2H); LC–MS  $m/z$  = 316  $[M+H]^+$ ,  $t_R$  = 4.63 min, purity 97% (Agilent, 20–90%  $\text{CH}_3\text{CN}$ , ES+ on an acidic method); HRMS  $m/z$   $[M+H]^+$  calcd for  $\text{C}_{17}\text{H}_{22}\text{N}_3\text{OS}$ : 316.1478, found: 316.1472.

**(4-Amino-2-(phenethylamino)thiazol-5-yl)(phenyl)methanone (21):** General procedure A gave **21** as a yellow solid (77 mg, 0.24 mmol, 24% yield);  $^1\text{H NMR}$  (500 MHz,  $[\text{D}_6]\text{DMSO}$ ):  $\delta$  = 8.73 (brs, 1H), 8.47 (brs, 1H), 7.87 (brs, 1H), 7.63–7.61 (m, 2H), 7.48–7.43 (m, 3H), 7.32–7.29 (m, 2H), 7.26–7.21 (m, 3H), 2.90 (s, 1H), 2.87 (t,  $J$  = 7.2 Hz, 2H), 2.74 ppm (d,  $J$  = 0.5 Hz, 1H); HRMS  $m/z$   $[M+H]^+$  calcd for  $\text{C}_{14}\text{H}_{18}\text{N}_3\text{OS}$ : 324.1165, found: 324.1153.

**1-(4-Amino-2-(cyclohexylamino)thiazol-5-yl)-2,2-dimethylpropan-1-one (22):** General procedure A gave **22** as a yellow powder (171 mg, 0.61 mmol, 61% yield);  $^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$  = 5.34 (d,  $J$  = 7.1 Hz, 1H), 3.30–3.28 (m, 1H), 2.00 (dd,  $J$  = 13.6, 3.3 Hz, 1H), 1.70 (dt,  $J$  = 13.6, 4.0 Hz, 2H), 1.57 (dt,  $J$  = 13.2, 4.0 Hz, 1H), 1.38–1.33 (m, 2H), 1.20 (s, 9H, tBu-H), 1.19–1.17 (m, 1H), 0.81 (t,  $J$  = 7.1 Hz, 1H), 0.79–0.76 ppm (m, 1H);  $^{13}\text{C NMR}$  (DMSO, 125 MHz)  $\delta$  = 128.6, 128.4, 119.5, 53.4, 53.3, 32.0, 27.0, 24.9, 24.3, 17.9, 16.6, 12.2 ppm; LC–MS  $m/z$  = 282  $[M+H]^+$ ,  $t_R$  = 4.60 min, purity 96% (Agilent, 20–90%  $\text{CH}_3\text{CN}$ , ES+ on an acidic method); HRMS  $m/z$   $[M+H]^+$  calcd for  $\text{C}_{14}\text{H}_{24}\text{N}_3\text{OS}$ : 282.1635, found: 282.1641.

## Acknowledgements

We thank the Wellcome Trust for financial support for these studies (grant ref. 077705 and strategic award WT083481). We also thank Iain Collie, Irene Hallyburton, and Bhavya Rao (DDU, University of Dundee) for carrying out the *T. brucei* and MRC5 proliferation studies, and Daniel James for data management. W.C.V.V. and K.K.O. were supported by US NIH grants R01AI089441 and R01AI080625.

**Keywords:** antiprotozoal agents • GSK3 • medicinal chemistry • protein kinases • *Trypanosoma brucei*

- [1] K. Stuart, R. Brun, S. Croft, A. Fairlamb, R. E. Gurtler, J. McKerrow, S. Reed, R. Tarleton, *J. Clin. Invest.* **2008**, *118*, 1301–1310.
- [2] R. Brun, R. Don, R. T. Jacobs, M. Z. Wang, M. P. Barrett, *Future Microbiol.* **2011**, *6*, 677–691.
- [3] P. P. Simarro, G. Cecchi, M. Paone, J. R. Franco, A. Diarra, J. A. Ruiz, E. M. Fevre, F. Courtin, R. C. Mattioli, J. G. Jannin, *Int. J. Health Geographics* **2010**, *9*, 57.
- [4] E. Matovu, T. Seebeck, J. C. Enyaru, R. Kaminsky, *Microbes Infect.* **2001**, *3*, 763–770.
- [5] M. P. Barrett, D. W. Boykin, R. Brun, R. R. Tidwell, *Br. J. Pharmacol.* **2007**, *152*, 1155–1171.
- [6] G. Priotto, S. Kasparian, W. Mutombo, D. Ngouama, S. Ghorashian, U. Arnold, S. Ghabri, E. Baudin, V. Buard, S. Kazadi-Kyanza, M. Ilunga, W. Mutangala, G. Pohlig, C. Schmid, U. Karunakara, E. Torreele, V. Kande, *Lancet* **2009**, *374*, 56–64.
- [7] O. Yun, G. Priotto, J. Tong, L. Flevaud, F. Chappuis, *PLoS Neglected Trop. Dis.* **2010**, *4*, e720.
- [8] A. Y. Sokolova, S. Wyllie, S. Patterson, S. L. Oza, K. D. Read, A. H. Fairlamb, *Antimicrob. Agents Chemother.* **2010**, *54*, 2893–2900.
- [9] E. Chatelain, J. R. Ioset, *Drug Des. Development Ther.* **2011**, *5*, 175–181.
- [10] E. Torreele, B. B. Trunz, D. Tweats, M. Kaiser, R. Brun, G. Mazue, M. A. Bray, B. Pecoul, *PLoS Neglected Trop. Dis.* **2010**, *4*, e923.
- [11] R. T. Jacobs, B. Nare, S. A. Wring, M. D. Orr, D. Chen, J. M. Sliagar, M. X. Jenks, R. A. Noe, T. S. Bowling, L. T. Mercer, C. Rewerts, E. Gaukel, J. Owens, R. Parham, R. Randolph, B. Beaudet, C. J. Bacchi, N. Yarllett, J. J. Plattner, Y. Freund, C. Ding, T. Akama, Y. K. Zhang, R. Brun, M. Kaiser, I. Scandale, R. Don, *PLoS Neglected Trop. Dis.* **2011**, *5*, e1151.
- [12] A. R. Renslo, J. H. McKerrow, *Nat. Chem. Biol.* **2006**, *2*, 701–710.
- [13] P. Cohen, *Nat. Rev. Drug Discovery* **2002**, *1*, 309–315.
- [14] B. D. Marsden, S. Knapp, *Curr. Opin. Chem. Biol.* **2008**, *12*, 40–45.
- [15] I. R. Nett, D. M. Martin, D. Miranda-Saavedra, D. Lamont, J. D. Barber, A. Mehler, M. A. Ferguson, *Mol. Cell. Proteomics* **2009**, *8*, 1527–1538.
- [16] M. Parsons, E. A. Worthey, P. N. Ward, J. C. Mottram, *BMC Genomics* **2005**, *6*, 127.
- [17] C. Naula, M. Parsons, J. C. Mottram, *Biochim. Biophys. Acta Proteins Proteomics* **2005**, *1754*, 151–159.
- [18] P. Rotella, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 6788–6793.
- [19] C. Doerig, *Biochim. Biophys. Acta Proteins Proteomics* **2004**, *1697*, 155–168.
- [20] P. Cohen, S. Frame, *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 769–776.
- [21] L. Meijer, M. Flajolet, P. Greengard, *Trends Pharmacol. Sci.* **2004**, *25*, 471–480.
- [22] P. Cohen, M. Goedert, *Nat. Rev. Drug Discovery* **2004**, *3*, 479–487.
- [23] R. Dajani, E. Fraser, S. M. Roe, N. Young, V. Good, T. C. Dale, L. H. Pearl, *Cell* **2001**, *105*, 721–732.
- [24] E. ter Haar, J. T. Coll, D. A. Austen, H. M. Hsiao, L. Swenson, J. Jain, *Nat. Struct. Biol.* **2001**, *8*, 593–596.
- [25] K. K. Ojo, J. R. Gillespie, A. J. Riechers, A. J. Napuli, C. L. Verlinde, F. S. Buckner, M. H. Gelb, M. M. Domostoj, S. J. Wells, A. Scheer, T. N. Wells, W. C. Van Voorhis, *Antimicrob. Agents Chemother.* **2008**, *52*, 3710–3717.
- [26] S. Alsford, D. J. Turner, S. O. Obado, A. Sanchez-Flores, L. Glover, M. Beriman, C. Hertz-Fowler, D. Horn, *Genome Res.* **2011**, *21*, 915–924.

- [27] K. K. Ojo, T. L. Arakaki, A. J. Napuli, K. K. Inampudi, K. R. Keyloun, L. Zhang, W. G. J. Hol, C. L. M. J. Verlinde, E. A. Merritt, W. C. Van Voorhis, *Mol. Biochem. Parasitol.* **2011**, *176*, 98–108.
- [28] J. J. L. Liao, *J. Med. Chem.* **2007**, *50*, 409–424.
- [29] R. Brenk, A. Schipani, D. James, A. Krasowski, I. H. Gilbert, J. Frearson, P. G. Wyatt, *ChemMedChem* **2008**, *3*, 435–444.
- [30] A. L. Hopkins, C. R. Groom, A. Alex, *Drug Discovery Today* **2004**, *9*, 430–431.
- [31] L. A. Cleghorn, A. Woodland, I. T. Collie, L. S. Torrie, N. Norcross, T. Luksch, C. Mpamhanga, R. G. Walker, J. C. Mottram, R. Brenk, J. A. Frearson, I. H. Gilbert, P. G. Wyatt, *ChemMedChem* **2011**, *6*, 2214–2224.
- [32] M. Hendlich, A. Bergner, J. Gunther, G. Klebe, *J. Mol. Biol.* **2003**, *326*, 607–620.
- [33] Y. Cheng, W. H. Prusoff, *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.
- [34] D. Spinks, H. B. Ong, C. P. Mpamhanga, E. J. Shanks, D. A. Robinson, I. T. Collie, K. D. Read, J. A. Frearson, P. G. Wyatt, R. Brenk, A. H. Fairlamb, I. H. Gilbert, *ChemMedChem* **2011**, *6*, 302–308.
- [35] A. N. Bowler, F. B. Hansen (Novo Nordisk A/S) WO2003011843A1, **2003**.
- [36] R. Lin, P. J. Connolly, S. Huang, S. K. Wetter, Y. Lu, W. V. Murray, S. L. Emanuel, R. H. Gruninger, A. R. Fuentes-Pesquera, C. A. Rugg, S. A. Middleton, L. K. Jolliffe, *J. Med. Chem.* **2005**, *48*, 4208–4211.
- [37] J. K. Laha, X. Zhang, L. Qiao, M. Liu, S. Chatterjee, S. Robinson, K. S. Kosik, G. D. Cuny, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 2098–2101.
- [38] P. N. Ibrahim, H. Cho, B. England, S. Gilletter, D. R. Artis, R. Zuckerman, Z. Chao (Plexxikon, Inc.) US20060041006, **2006**.
- [39] S. Sengupta, S. L. Smitha, N. E. Thomas, T. R. Santhoshkumar, S. K. Devi, K. G. Sreejalekshmi, K. N. Rajasekharan, *Br. J. Pharmacol.* **2005**, *145*, 1076–1083.
- [40] K. G. Sreejalekshmi, S. K. C. Devi, K. N. Rajasekharan, *Tetrahedron Lett.* **2006**, *47*, 6179–6182.
- [41] J. D. Thompson, D. G. Higgins, T. J. Gibson, *Nucleic Acids Res.* **1994**, *22*, 4673–4680.
- [42] A. Sali, T. L. Blundell, *J. Mol. Biol.* **1993**, *234*, 779–815.
- [43] J. Allard, T. Nikolcheva, L. Gong, J. Wang, P. Dunten, Z. Avnur, R. Waters, Q. Sun, B. Skinner, PDB ID: 1R0E, **2004**, <http://www.rcsb.org/pdb/home/home.do>.
- [44] P. Benkert, S. C. E. Tosatto, D. Schomburg, *Proteins* **2008**, *71*, 261–277.
- [45] P. Benkert, M. Kuenzli, T. Schwede, *Nucleic Acids Res.* **2009**, *37*, W510–W514.
- [46] P. R. Gerber, K. Muller, *J. Comput. Aided Mol. Des.* **1995**, *9*, 251–268.
- [47] J. Sadowski, J. Gasteiger, *Chem. Rev.* **1993**, *93*, 2567–2581.
- [48] J. Bain, L. Plater, M. Elliott, N. Shpiro, C. J. Hastie, H. McLauchlan, I. Klevernic, J. S. Arthur, D. R. Alessi, P. Cohen, *Biochem. J.* **2007**, *408*, 297–315.
- [49] B. Rätz, M. Iten, Y. Grether-Bühler, R. Kaminski, R. Brun, *Acta Trop.* **1997**, *68*, 139–147.
- [50] J. Ma, C. Benz, R. Grimaldi, C. Stockdale, P. Wyatt, J. Frearson, T. C. Hammaron, *J. Biol. Chem.* **2010**, *285*, 15356–15368.

---

Received: February 17, 2013

Revised: May 15, 2013

Published online on June 14, 2013