



Steketee, Pieter Christiaan (2016) *Investigating the mode of action of AN5568, a novel therapeutic against African trypanosomiasis*. PhD thesis.

<http://theses.gla.ac.uk/7478/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten:Theses  
<http://theses.gla.ac.uk/>  
theses@ gla.ac.uk

# Investigating the mode of action of AN5568, a novel therapeutic against African trypanosomiasis

**Pieter Christiaan Steketeer, BSc, MRes**

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



University  
of Glasgow

Institute of Infection, Immunity & Inflammation

College of Medical, Veterinary & Life Sciences

University of Glasgow

March 2016

## Abstract

---

The protozoan parasite *Trypanosoma brucei* is the causative agent of Human African Trypanosomiasis (HAT) and Nagana disease in mammals. These diseases present a major socio-economic burden to large areas of sub-Saharan Africa. Current therapeutics involve complex and toxic regimens which can lead to fatal side-effects. In addition, there is evidence for drug resistance emerging in the field. Hence, there is a desperate need for novel therapies. Benzoxaboroles are a novel class of boron-containing compounds under development for use against a wide spectrum of diseases. AN5568 is a lead compound for the treatment of HAT, which has demonstrated effective clearance of both early- and late-stage trypanosomiasis in a murine model, and is currently undergoing clinical trials. However, the mechanism by which AN5568 kills *T. brucei* is elusive. In this study we sought to use 'omics'-based techniques to investigate the mode of action of AN5568 in a laboratory strain of *Trypanosoma brucei brucei*.

Cells treated with the benzoxaborole showed significant perturbations in methionine metabolism. In particular, there were increases in S-adenosyl-L-methionine, an essential methyl group donor involved in methyltransferase reactions. These changes were similar to those elicited by the non-specific methyltransferase inhibitor sinefungin. Changes were also observed in lipid metabolism, sugar nucleotide metabolism and glycerophosphatidylinositol biosynthesis. Further analyses were carried out to investigate the effect of AN5568 on cellular stress responses and cell morphology.

To further probe the mechanics of AN5568-treatment, a drug-resistant cell line was generated. This cell line showed cross-resistance with sinefungin, further supporting similar modes of action for these two drugs. Interestingly, the AN5568-resistant cell line exhibited upregulation of procyclic form-specific genes, as well as downregulation of blood-stream form-specific genes, which led to the hypothesis that the cell line had undergone a differentiation event. However, microscopy analysis showed that overall morphology of the cells still resembled those of bloodstream forms, despite them having acquired a procyclic-like metabolic physiology.

A secondary aim of this project was to elucidate the metabolic changes that lead to increased growth rates in *T. brucei* cells undergoing loss-of-heterozygosity on chromosome 10. This phenomenon, whereby a significant portion of the chromosome is lost, has been observed independently on multiple occasions in lab adapted *T. brucei* strains, yet how this alteration affects intracellular metabolism was hitherto unknown. Using two procyclic *T. brucei* cell lines, this study was able to show that the increased growth rates are glucose-dependent with a potential intracellular alteration in succinate and acetate production. These data have important implications for the field, where LOH has been observed in the clonally expanding *T.b. gambiense* type I.

# Contents

---

Abstract.....	i
Contents.....	ii
List of Figures .....	viii
List of Tables .....	xi
Acknowledgements.....	xii
Author’s declaration .....	xv
Abbreviations.....	xvi
Chapter 1. Introduction .....	1
1.1 General introduction.....	1
1.1.1 Taxonomy.....	1
1.1.2 Life Cycle .....	1
1.2 Human African Trypanosomiasis .....	3
1.2.1 Historical overview.....	3
1.2.2 Epidemiology.....	4
1.2.3 Clinical manifestations & diagnosis.....	5
1.2.4 Human mechanisms of resistance .....	7
1.3 Therapeutics/chemotherapy against HAT/AAT .....	7
1.3.1 Human African Trypanosomiasis (HAT) .....	7
1.3.2 Drugs currently in clinical trials.....	9
1.3.2.1 Fexinidazole .....	9
1.3.2.2 Benzoxaboroles – a unique scaffold with great potential .....	10
1.3.2.3 Benzoxaboroles for neglected tropical diseases.....	11
1.3.3 Animal African Trypanosomiasis (AAT) .....	12
1.4 Trypanocides and their mechanisms of action .....	12
1.4.1 Eflornithine ( $\alpha$ -difluoromethylornithine, DFMO) .....	12
1.4.2 Pentamidine .....	13
1.4.3 Melarsoprol.....	13
1.4.4 Suramin .....	13
1.5 Cell biology.....	14
1.5.1 Antigenic variation .....	14
1.5.2 Overview of <i>T. brucei</i> cell biology and organelles.....	14

1.5.3	Comparing bloodstream form and procyclic form cells.....	16
1.5.4	The glycosome .....	16
1.5.5	Kinetoplasts and the trypanosome mitochondrion.....	16
1.6	Genome and gene regulation in <i>T. brucei</i> .....	17
1.6.1	Genome structure and arrangement.....	17
1.6.2	Post-transcriptional regulation of gene expression.....	18
1.6.3	The trypanosome spliced leader.....	18
1.7	Genetic diversity in the trypanosome.....	20
1.7.1	Loss-of-heterozygosity: a potential key to genetic diversity .....	20
1.8	Metabolism in <i>T. brucei</i> .....	22
1.8.1	General metabolism.....	22
1.8.2	Energy metabolism: glycolysis and the incomplete citric acid cycle .....	23
1.8.3	L- methionine and the Yang cycle .....	26
1.8.4	The role of S-adenosyl-L-methionine in global metabolism .....	27
1.8.5	Methyltransferases .....	28
1.8.6	Polyamine pathway and oxidative stress.....	29
1.8.7	Metabolism during differentiation .....	30
1.9	Omics technologies .....	31
1.9.1	Genomics, transcriptomics and proteomics .....	31
1.9.2	Metabolomics .....	32
1.9.3	Other ‘omics.....	33
1.10	Techniques in drug target identification and validation.....	34
1.10.1	Affinity chromatography.....	35
1.10.2	Generation of resistant cell lines .....	36
1.10.3	RNAi-target sequencing (RIT-seq).....	37
1.10.4	Metabolomics .....	37
1.10.5	Systems biology.....	39
1.10.6	Drug affinity responsive target stability (DARTS).....	40
1.10.7	Stable isotope labelling by amino acids in cell culture (SILAC) .....	41
1.10.8	The cellular thermal shift assay and thermal profiling .....	41
1.10.9	CRISPR – the new kid on the block.....	42
1.10.10	Summary of drug target identification .....	43
1.11	Aims & Objectives for this project .....	44
Chapter 2.	Methods.....	45
2.1	Trypanosome tissue culture.....	45
2.1.1	Bloodstream form culture.....	45

2.1.2	Procyclic form culture .....	45
2.1.3	Differentiation of bloodstream forms to procyclic forms.....	46
2.1.4	Stabilate preparation and retrieval.....	46
2.1.5	Alamar blue assay .....	46
2.1.6	Isobologram assay.....	47
2.1.7	Transfection of trypanosomes .....	47
2.1.8	Generation and selection of AN5568-resistant <i>T. brucei</i> .....	48
2.2	Isolation of biological components from trypanosomes .....	48
2.2.1	DNA Extraction .....	48
2.2.2	RNA extraction .....	49
2.2.3	Protein extraction .....	49
2.3	Electrophoresis .....	50
2.3.1	DNA gel electrophoresis.....	50
2.3.2	Protein gel electrophoresis .....	50
2.3.3	Two-dimensional difference gel electrophoresis (2D DIGE) .....	50
2.4	Molecular biology .....	51
2.4.1	PCR .....	51
2.4.2	Genotyping.....	51
2.4.3	Plasmids .....	52
2.4.4	RNA interference constructs.....	52
2.4.5	Overexpression & re-expression constructs.....	53
2.5	Real-time qPCR.....	53
2.5.1	Reverse transcription .....	53
2.5.2	Primers and primer efficiency.....	54
2.5.3	Real-time PCR.....	54
2.6	Metabolomics .....	55
2.6.1	Metabolite extraction from <i>T. brucei</i> cultures for LC-MS.....	55
2.6.2	Liquid chromatography-mass spectrometry.....	55
2.6.3	LC-MS Data analysis .....	55
2.7	Proteomics .....	56
2.7.1	Drug affinity responsive target stability (DARTS).....	56
2.7.2	Protein mass spectrometry and analysis .....	57
2.8	Transcriptomics.....	57
2.8.1	Sample preparation .....	57
2.8.2	RNA-seq.....	57
2.8.3	Transcriptomics data analysis .....	57

2.9	Blotting.....	58
2.9.1	Sample preparation .....	58
2.9.2	Transfer & blotting.....	58
2.10	Microscopy.....	58
2.10.1	Mitotracker staining.....	58
2.10.2	Preparation and fixation of cells .....	59
2.10.3	Cell cycle and morphology analyses .....	59
2.10.4	Analysis of kinetoplast localisation .....	60
2.11	Chemicals .....	60
2.12	Computational methods .....	60
2.12.1	Data analysis and presentation.....	60
Chapter 3.	The mode of action of AN5568 .....	61
3.1	Introduction .....	61
3.1.1	AN5568 brings hope for HAT .....	61
3.2	<i>In vitro</i> activity of AN5568 on <i>T. brucei</i> .....	62
3.3	Morphological changes in AN5568-treated cells.....	64
3.4	Untargeted metabolomics analysis of AN5568-treated bloodstream form trypanosomes. .....	69
3.4.1	Methionine metabolism.....	72
3.4.2	Effect of AN5568-treatment on VSG biosynthesis.....	74
3.4.3	Lipid metabolism.....	76
3.4.4	Further changes of interest.....	79
3.4.5	Comparison of AN5568 & sinefungin metabolotypes .....	81
3.5	Metabolism in AN5568-treated procyclic cells.....	86
3.6	AN5568 elicits a similar metabolic response in <i>L. mexicana</i> .....	89
3.7	Targeting methionine metabolism .....	93
3.7.1	Changes in <sup>13</sup> C labeling patterns across the metabolome .....	95
3.7.2	Lysine is methylated in an AdoMet-dependent fashion, but is not involved in carnitine biosynthesis .....	97
3.8	Knock-down of TbCgm1, a splice-leader methyltransferase .....	98
3.9	ER stress leads to metabolic perturbations also seen in AN5568 treatment.....	103
3.9.1	ER stress results in upregulation of the trypanothione biosynthesis pathway .....	103
3.9.2	AdoMet increase is specific to AN5568 treatment.....	104
3.10	Using DARTS to probe the putative mode of action of AN5568.....	106
3.11	Lysates of treated parasites show enrichment of a protein subset .....	109
3.12	Analysis of the <i>T. brucei</i> complement of methyltransferases – the “methyltransferome”.. .....	115

3.13	Overexpression of 12 essential methyltransferases <i>in vitro</i> does not identify an AN5568 target .....	121
3.13.1	Domain analysis of 12 essential methyltransferases.....	121
3.13.2	Characterisation of 12 overexpression cell lines .....	124
3.14	Discussion.....	129
Chapter 4.	Generation and characterisation of an AN5568-resistant cell line.....	134
4.1	Introduction .....	134
4.2	Generation of AN5568 in <i>T. brucei</i> cells <i>in vitro</i> .....	135
4.3	RNA-seq analysis of the AN5568 <sup>R</sup> cell line .....	138
4.3.1	PCF-associated genes are upregulated .....	143
4.3.2	Downregulation of BSF-specific genes.....	146
4.3.3	Differential expression of <i>T. brucei</i> transporters .....	147
4.3.4	Pathway analysis of RNA-seq data.....	149
4.4	SNP and Indel analysis of the AN5568 <sup>R</sup> line.....	155
4.5	The AN5568 <sup>R</sup> line exhibits BSF morphology.....	159
4.6	Metabolomics analysis of the AN5568 <sup>R</sup> line .....	161
4.6.1	The AdoMet/5'-MTA phenotype is abolished in AN5568 <sup>R</sup> cells .....	162
4.6.2	Lipid and carbohydrate metabolism in the AN5568 <sup>R</sup> line.....	164
4.6.3	Lysine metabolism in the absence and presence of AN5568 .....	166
4.6.4	Differences in metabolism between untreated wild-type and untreated AN5568 <sup>R</sup> cells .....	167
4.6.5	Lipid metabolism is almost completely restored in AN5568 <sup>R</sup> cells.....	169
4.6.6	Comparison of metabolomics and transcriptomics datasets .....	171
4.7	Discussion.....	172
Chapter 5.	Loss-of-heterozygosity in African trypanosomes.....	176
5.1	Introduction .....	176
5.1.1	Aims & Objectives .....	178
5.2	Occurrence of LOH on chromosome 10.....	179
5.3	Growth phenotype in LOH lines is glucose-dependent, but not related to chromosome 10 glucose transporter arrays.....	179
5.4	Comparative genomics analyses of independent LOH lines.....	185
5.5	Metabolomics analyses of the TREU 927 LOH lines.....	188
5.6	Discussion.....	192
Chapter 6.	Discussion & final thoughts.....	196
6.1	The mode of action of AN5568 – Where do we stand?.....	196
6.2	The relationship between metabolism and its genetic determinants – LOH as a case study .....	200

6.3 Omics in the study of biological organisms – final thoughts .....201

References .....202

Appendix A. Differentiating trypanosome medium (DTM) .....235

Appendix B. SDM-80.....236

Appendix C. Creek’s minimal medium (CMM) .....237

Appendix D. Buffers and reagents .....238

Appendix E. Primers & Plasmids .....241

Appendix F. PCR and IEF cycling conditions .....246

## List of Figures

---

Figure 1-1: Life cycle of <i>Trypanosoma brucei</i> . .....	2
Figure 1-2: Epidemiology of HAT in the sub-Saharan African continent. ....	4
Figure 1-3: Deaths resulting from <i>T.b. gambiense</i> (Tbg) and <i>T.b. rhodesiense</i> (Tbr) infections between 1990 and 2014. ....	5
Figure 1-4: A timeline of drugs used against HAT, and when they were first brought to market. ....	8
Figure 1-5: Two-dimensional chemical structures of drugs commonly used against HAT. ....	10
Figure 1-6: Cell biology of <i>T. brucei</i> . ....	15
Figure 1-7: Map of PCR microsatellite analysis of chromosome 10. ....	21
Figure 1-8: Energy metabolism in <i>T. brucei</i> . ....	24
Figure 1-9: The Yang cycle. Components indicated in red are those that are known to be present in <i>T. brucei</i> . ....	27
Figure 1-10: Comparison of structures of AdoMet and sinefungin. ....	28
Figure 3-1: Sigmoidal dose-response curves carried out on AN5568-treated <i>T. brucei</i> . ....	62
Figure 3-2: Growth of BSF and PCF <i>T. brucei</i> during AN5568 treatment <i>in vitro</i> . ....	64
Figure 3-3: The effect of 2× EC <sub>50</sub> AN5568 on cellular morphology. ....	66
Figure 3-4: Effect of 10× EC <sub>50</sub> AN5568 treatment on <i>T. brucei</i> morphology. ....	68
Figure 3-5: Cell cycle analysis of AN5568-treated cells. ....	69
Figure 3-6: Volcano plot of metabolic changes in <i>T. brucei</i> after 6-hour treatment with 10× EC <sub>50</sub> AN5568. ....	71
Figure 3-7: Heat map showing the changes in metabolism occurring in <i>T. brucei</i> as a result of AN5568-treatment. ....	72
Figure 3-8: Changes in methylation metabolism after AN5568 treatment. ....	73
Figure 3-9: Metabolites involved in methyltransferase reaction pathways. ....	74
Figure 3-10: Metabolites involved in glycoprotein synthesis are upregulated after AN5568 treatment, but VSG expression is not impaired. ....	76
Figure 3-11: Fatty acid metabolism in cells treated with AN5568 for 6 hours at 10× EC <sub>50</sub> . ....	77
Figure 3-12: AN5568 causes decreased abundances in lipids derived from phosphatidylcholine (PC) and phosphatidylethanolamine (PE). ....	79
Figure 3-13: Comparative analysis of metabolic profiles of AN5568-treated and sinefungin-treated <i>T. brucei</i> . ....	82
Figure 3-14: Further analysis of metabolic changes in AN5568- and sinefungin-treated <i>T. brucei</i> . ....	85
Figure 3-15: Metabolic changes occurring after AN5568 treatment of PCF <i>T. brucei</i> . ....	87

Figure 3-16: Further metabolic changes in AN5568-treated PCF <i>T. brucei</i> parasites.....	88
Figure 3-17: Effect of AN5568 on <i>in vitro</i> growth of <i>L. mexicana</i> promastigotes. ....	90
Figure 3-18: AN5568 causes metabolic perturbations in <i>L. mexicana</i> that are similar to those observed in <i>T. brucei</i> .....	91
Figure 3-19: Changes in lipid metabolism after AN5568-treatment of <i>L. mexicana</i> promastigotes. .....	92
Figure 3-20: Heat plot showing metabolic changes upon AN5568 treatment in both intracellular metabolism and spent medium. ....	94
Figure 3-21: Tracing <sup>13</sup> C distribution in cells incubated with <sup>13</sup> C-(U)-L-methionine and treated with AN5568.....	96
Figure 3-22: AN5568 causes significant changes in methylated lysine metabolism. ....	98
Figure 3-23: Generation of a TbCgm <sup>RNAi</sup> cell line. ....	100
Figure 3-24: Metabolomics analysis of the TbCgmRNAi <i>T. brucei</i> cell line. ....	101
Figure 3-25: Comparative analysis of changes in nucleotide metabolism in the TbCgm <sup>RNAi</sup> cell line and AN5568-treated cells. ....	102
Figure 3-26: Metabolic changes in the trypanothione biosynthesis pathway.....	104
Figure 3-27: Changes seen in methionine metabolism are specific to AN5568-treatment, but modified lysine residues are not.....	105
Figure 3-28: Protein samples subjected to DARTS assay. ....	107
Figure 3-29: Protein lysates from AN5568-treated parasites visualised on protein gels show differences with wild-type lysates. ....	110
Figure 3-30: Comparative analysis of the soluble proteome from WT and AN5568-treated <i>T. brucei</i> cells. ....	112
Figure 3-31: Overview of MTases in the <i>T. brucei</i> genome. ....	119
Figure 3-32: Characterisation of an ODC overexpressor as well as 12 putative MTase overexpressors.....	127
Figure 4-1: Generation of an AN5568-resistant cell line. ....	136
Figure 4-2: AN5568 <sup>R</sup> undergoes several changes in drug sensitivity.....	137
Figure 4-3: RNAseq data quality analysis.....	139
Figure 4-4: RNA-seq sequencing data analysis of the AN5568-resistant cell line. ....	141
Figure 4-5: Volcano plot of changes in transcript abundance in the AN5568 <sup>R</sup> cell line. ....	142
Figure 4-6: RNAseq coverage of chromosome 6. ....	145
Figure 4-7: Heat map of abundances of transcripts involved in glycolytic metabolism.....	151
Figure 4-8: Heat map of abundances of transcripts involved in carboxylic acid metabolism. ....	152
Figure 4-9: Comparative analyses of the AN5568R line and “stumpy” expression.....	154
Figure 4-10: Number of variants (SNPs and indels) detected on each chromosome.....	156

Figure 4-11: Microscopy analysis of the AN5568<sup>R</sup> cell line.....161

Figure 4-12: Significant metabolic changes in methionine metabolism seen during AN5568 treatment are abolished in AN5568<sup>R</sup> cells. ....163

Figure 4-13: Further wild-type metabolic signatures that are altered in the AN5568<sup>R</sup> cell line. ...165

Figure 4-14: Modified lysine metabolism in benzoxaborole-resistant cells. ....167

Figure 4-15: Novel metabolic changes in the benzoxaborole-resistant cell line. ....169

Figure 4-16: Lipid metabolism in AN5568-resistant cells. ....170

Figure 5-1: Genotyping PCRs to identify the LOH 1 and LOH 2 cell lines.....179

Figure 5-2: Growth of WT and LOH *T. brucei* TREU 927 PCF cells in procyclic culture media. ....180

Figure 5-3: Comparative Analysis of the THT-1 and THT-2 glucose transporters from total genomic DNA and cDNA in WT and LOH *T. brucei* parasites. ....184

Figure 5-4: Predicted sequence and structure of mitochondrial malate dehydrogenase (mMDH, Tb927.10.2560) in the *T.b. brucei* TREU 927 LOH lines.....187

Figure 5-5: Hydrogen bonding around residue E293 of mMDH. ....188

Figure 5-6: Glycolytic metabolism in WT and LOH cells. ....189

Figure 5-7: Levels of metabolites involved in PCF central carbon metabolism. ....190

Figure 5-8: Phosphorylated pantothenate is absent from the LOH lines. ....191

Figure 5-9: Putative changes in central carbon metabolism occurring in LOH cell line, based on metabolomics data. ....193

## List of Tables

---

Table 1-1: Example of drug MoA studies using metabolomics.....	38
Table 3-1: EC <sub>50</sub> of AN5568 in various kinetoplastids.....	63
Table 3-2: Further changes in <i>T. brucei</i> metabolism after AN5568 treatment, that were not associated with one particular pathway.....	80
Table 3-3: Proteomics analysis of the enriched band observed in the DARTS assay. ....	108
Table 3-4: Peptides identified in a 2D-DiGE undergoing significant changes after AN5568-treatment.....	114
Table 3-5: MTases that have been identified and characterised in <i>T. brucei</i> . ....	116
Table 3-6: A comprehensive overview of the <i>T. brucei</i> methyltransferase repertoire. ....	118
Table 3-7: Putative essential MTases.....	122
Table 3-8: qPCR analysis of 12 MTase overexpression lines.....	125
Table 4-1: Top 15 genes significantly upregulated in the AN5568 <sup>R</sup> line.....	143
Table 4-2: Top 15 genes significantly downregulated in the AN5568 <sup>R</sup> cell line.....	147
Table 4-3: Transporter proteins that exhibited significant differential expression in AN5568-resistant <i>T. brucei</i> cells,.....	148
Table 4-4: SNP and indel counts for the complete RNA-seq data set. ....	155
Table 4-5: Homozygous SNPs unique to the AN5568 <sup>R</sup> cell lines.....	158
Table 4-6: Corresponding putative metabolite identifications for figure 4-16. ....	170
Table 5-1: Genes of interest located in the LOH region of <i>T. brucei</i> TREU 927 chromosome 10...	182

## Acknowledgements

---

First and foremost, I would like to thank Annette MacLeod and Mike Barrett for their constant guidance throughout the 3 years I have spent with their groups. It is thanks to Annette that I ended up in Glasgow in the first place, after a brief summer foray into trypanosome biology. In my time here I have been offered so many opportunities and met some incredible people, and that is thanks to you both. In addition you helped me realise just how great the world of the African trypanosome is. I hope to keep in touch with you for years to come, thank you.

To the present and past members of the MacLeod group: Paul, Anneli, Caroline, Willie, Heli and Taylor, and the Barrett group: Fede (grazie per tutto!), Fiona, Julie, Kevin, Andrew, Isabel, Al, Clement, Dave, Ed, Darren, Brunda, Jana, Kat, Hyun, and Gordon (keep on rockin'!). You have helped so much with friendly chats and discussions, immeasurable support, great advice and timeless evenings of pub fun. It has been an honour and a pleasure to work with you all.

I would like to thank everyone at the WTCMP. For 3 years it was my home, and I have come to consider you all my extended family. Special thanks to Fernando for introducing me to the Spanish way of life, and for being a great flatmate for at least a part of the PhD journey. Robyn, the best gig-buddy around. The Thursday and Friday football gangs, for the stress relief. Special thanks also to Jane for the constant advice and help in the lab, and to the members of the Mottram, McCulloch, de Koning and Waters groups for making my time here so memorable. Alex Mackay, thank you for all the work you do, it makes WTCMP a great place to be (Jazzademics was a real highlight). Thanks also to the members of Glasgow Polyomics, especially Richard Burchmore and Karl Burgess, for all their contributions.

I have learnt a lot about the beautiful world of bioinformatics, and for this I would like to thank the WTCMP bioinformatics team of Nick, Kathryn and John, as well as Willie. In particular I would like to thank Nick for his immeasurable patience and the countless hours of replying to all my emails. Who knew bioinformatics could be so much fun (once you begin to understand...)!

Thanks to Darren Monckton and Olwyn Byron, the directors of the Wellcome Trust PhD programme. You helped me settle and ease into the life of a PhD student, and have always been there to offer invaluable knowledge and wisdom. In addition, thanks to the entire WT 4-yr PhD group both past and present. The forum breakfasts were great, the retreats even greater.

A special mention to Kesha, my go-to pub buddy for 2 years, and the students, post-docs and PIs on level 2 as well. A big thanks to Nicky, Fraser and Gordon from the band. Sometimes loud music and non-science discussions were exactly what was needed.

Several other PIs and their labs made a significant contribution towards this project and I would like to take this opportunity to thank them: Keith Matthews and Paul Michels from the University of Edinburgh, Christine Clayton at the University of Heidelberg (and a big thanks to Daniela for her insightful discussions), Liam Morrison at the Roslin Institute, Paul Englund at John Hopkins University (and to Soo Hee for her help with the ER tracker troubles), and Fred Bringaud at the Université de Bordeaux for his help on the LOH project. In addition, thanks to Srini, Pamela, Hazel and Vanessa for a very successful and particularly enjoyable 5 weeks at the Novartis Institute for Tropical Disease.

I would never have made it this far without the support of my family. My sisters, Marina and Eva, thank you for being the best sisters anyone could ask for! I hope I will have more time to come and visit you both, and to do so often. Mum, you have been so supportive not just for the last 3 years, but all my life. You have always been there for me, and I thank you for listening to my constant chatter - *"Trouwens, trouwens!"* Dad, I am so, so sorry that you cannot be here to share this moment with me. Your encouragement, enthusiasm and everlasting smile, warmth and memories live on with me, and with us. I know you would be so proud, *"bedankt voor alles, makker."*

Finally, but most importantly by far, a loving thanks to Nina. You have survived my endless troubles of writing as well as late nights and weekends in the lab, and you have kept me grounded and sane. You make life so much easier to live, and I honestly don't know where I'd be without you. It's been one heck of a journey so far, I look forward to your company for the next step!

In loving memory,  
*Johan (Hans) Helenus Steketee*  
1945 - 2015

## Author's declaration

---

"I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution."

Signature: \_\_\_\_\_

Printed name: \_\_\_\_\_

## Abbreviations

---

<b>5'-MTA</b>	5'-methylthioadenosine/S-methyl-5'-thioadenosine
<b>5'-MTR</b>	S-methyl-5'-thio-D-ribose
<b>AAT</b>	Animal African trypanosomiasis
<b>ABC</b>	ATP-binding cassette
<b>AdoHcy</b>	S-adenosyl-L-homocysteine
<b>AdoHcyH</b>	S-adenosyl-L-homocysteine hydrolase
<b>AdoMet</b>	S-adenosyl-L-methionine
<b>AdoMetDC</b>	AdoMet decarboxylase
<b>ATP</b>	Adenosine triphosphate
<b>BSA</b>	Bovine serum albumin
<b>BSF</b>	Bloodstream form
<b>CATT</b>	Card agglutination test for trypanosomiasis
<b>CETSA</b>	Cellular thermal shift assay
<b>CMM</b>	Creek's minimal media
<b>CNS</b>	Central nervous system
<b>CoA</b>	Coenzyme A
<b>CRD</b>	Cross reactive determinant
<b>CSF</b>	Cerebro-spinal fluid
<b>DARTS</b>	Drug affinity responsive target stability
<b>dcAdoMet</b>	Decarboxylated AdoMet
<b>DIGE</b>	Difference gel electrophoresis
<b>DLP</b>	Dynammin-like protein
<b>DMSO</b>	Dimethyl sulfoxide
<b>DNA</b>	Deoxyribonucleic acid
<b>DNDi</b>	Drugs for neglected disease initiative
<b>DTT</b>	Dithiothreitol
<b>EC<sub>50</sub></b>	Half maximal effective concentration
<b>EP</b>	glutamic acid and proline
<b>ER</b>	Endoplasmic reticulum
<b>FBS</b>	Fetal bovine serum
<b>FCS</b>	Fetal calf serum

<b>GC-MS</b>	Gas chromatography-mass spectrometry
<b>GFP</b>	Green fluorescent protein
<b>GOI</b>	Gene of interest
<b>GPI</b>	Glycophosphatidylinositol
<b>GSS</b>	Glutathionylspermidine synthetase
<b>HAPT1</b>	High-affinity pentamidine transporter 1
<b>HAT</b>	Human African Trypanosomiasis
<b>Hcy</b>	L-homocysteine
<b>HRP</b>	Horseradish peroxidase
<b>Hsp</b>	Heat shock protein
<b>ICAT</b>	Isotope coded affinity tags
<b>ICMT</b>	Isoprenylcysteine carboxyl methyltransferase
<b>iTRAQ</b>	Isobaric tag for relative and absolute quantification
<b>IV</b>	Intravenous
<b>kDNA</b>	Kinetoplast DNA
<b>KO</b>	Knock-out
<b>LAPT1</b>	Low-affinity pentamidine transporter 1
<b>LC-MS</b>	Liquid chromatography-mass spectrometry
<b>LOH</b>	Loss-of-heterozygosity
<b>m/z</b>	Mass-to-charge ratio
<b>mMDH</b>	Mitochondrial malate dehydrogenase
<b>mRNA</b>	messenger RNA
<b>MS</b>	Mass spectrometry
<b>MTase</b>	Methyltransferase
<b>MTome</b>	Methyltransferome
<b>nm</b>	nanomolar
<b>NMR</b>	Nuclear magnetic spectroscopy
<b>NTD</b>	Neglected tropical disease
<b>ODC</b>	Ornithine decarboxylase
<b>PAD</b>	Protein associated with differentiation
<b>PC</b>	Phosphatidylcholine
<b>PCF</b>	Procyclic form
<b>PCR</b>	Polymerase chain reaction
<b>PE</b>	Phosphatidylethanolamine
<b>PEP</b>	Phospho <i>eno</i> pyruvate
<b>ppm</b>	Parts per million

<b>PYR</b>	Pyruvate kinase
<b>QTL</b>	Quantitative trait loci
<b>RBP</b>	RNA-binding protein
<b>RNA</b>	Ribonucleic acid
<b>RIT-seq</b>	
<b>RNAi</b>	RNA interference
<b>RNA-seq</b>	RNA sequencing
<b>rRNA</b>	ribosomal RNA
<b>RT</b>	Retention time
<b>SAGE</b>	Serial analysis of gene expression
<b>SAR</b>	Structure-activity relationship
<b>SILAC</b>	Stable isotope labeling by amino acids in cell culture
<b>SL</b>	Spliced leader
<b>SLS</b>	Spliced leader silencing
<b>SNP</b>	Single nucleotide polymorphism
<b>SSF</b>	Short stumpy form
<b>STIB</b>	Swiss Tropical Institute, Basel
<b>TB</b>	Tuberculosis
<b>Tb</b>	Trypanosoma brucei
<b>TCA</b>	Tricarboxylic acid/citric acid
<b>THT</b>	Trypanosome hexose transporter
<b>TLF</b>	Trypanosome lytic factor
<b>TREU</b>	Trypanosoma Research Edinburgh University
<b>tRNA</b>	transfer RNA
<b>TrypSyn</b>	Trypanothione synthetase
<b>TYW3</b>	tRNA-yW synthesising protein
<b>UTR</b>	Untranslated region
<b>VSG</b>	Variable surface glycoprotein
<b>WHO</b>	World Health Organisation
<b>WT</b>	Wild-type
<b>yW</b>	Wybutosine

# Chapter 1. Introduction

---

## 1.1 General introduction

---

### 1.1.1 Taxonomy

*Trypanosoma brucei* is a single-celled protozoan parasite and the causative agent of Human African Trypanosomiasis (HAT, also referred to as Sleeping Sickness) as well as Nagana disease (also referred to as Animal African Trypanosomiasis, AAT) in mammals (Lejon *et al.*, 2013). This parasite is a member of Order Kinetoplastida, which also includes other human pathogens including *Trypanosoma cruzi*, which causes Chagas' disease, and *Leishmania spp*, the parasitic cause of a range of diseases collectively referred to as the Leishmaniasis, of which the visceral form, or kala azar, is fatal (Rassi *et al.*, 2010, Barrett & Croft, 2012). Together, these parasites are responsible for a significant socio-economic burden across the globe. Taking into account veterinary parasites including *T. congolense*, *T. vivax*, *T. evansi* and *T. equiperdum*, all subspecies of the *Trypanosoma* genus, widens this burden considerably (Field & Carrington, 2009). The taxonomy of the human-infective subspecies of African trypanosomes can be broken down into two subspecies of *T. brucei*: *T.b. gambiense* and *T.b. rhodesiense* are prevalent in west and central Africa or east and South Africa respectively (Koffi *et al.*, 2009). A third *T. brucei* subspecies, *T.b. brucei*, is infective to non-human vertebrates and is distributed across the sub-Saharan African continent (Balmer *et al.*, 2011). The West African *T.b. gambiense* can be further divided into two subspecies, type 1 and type 2 (Tait *et al.*, 1984, Hide *et al.*, 1990).

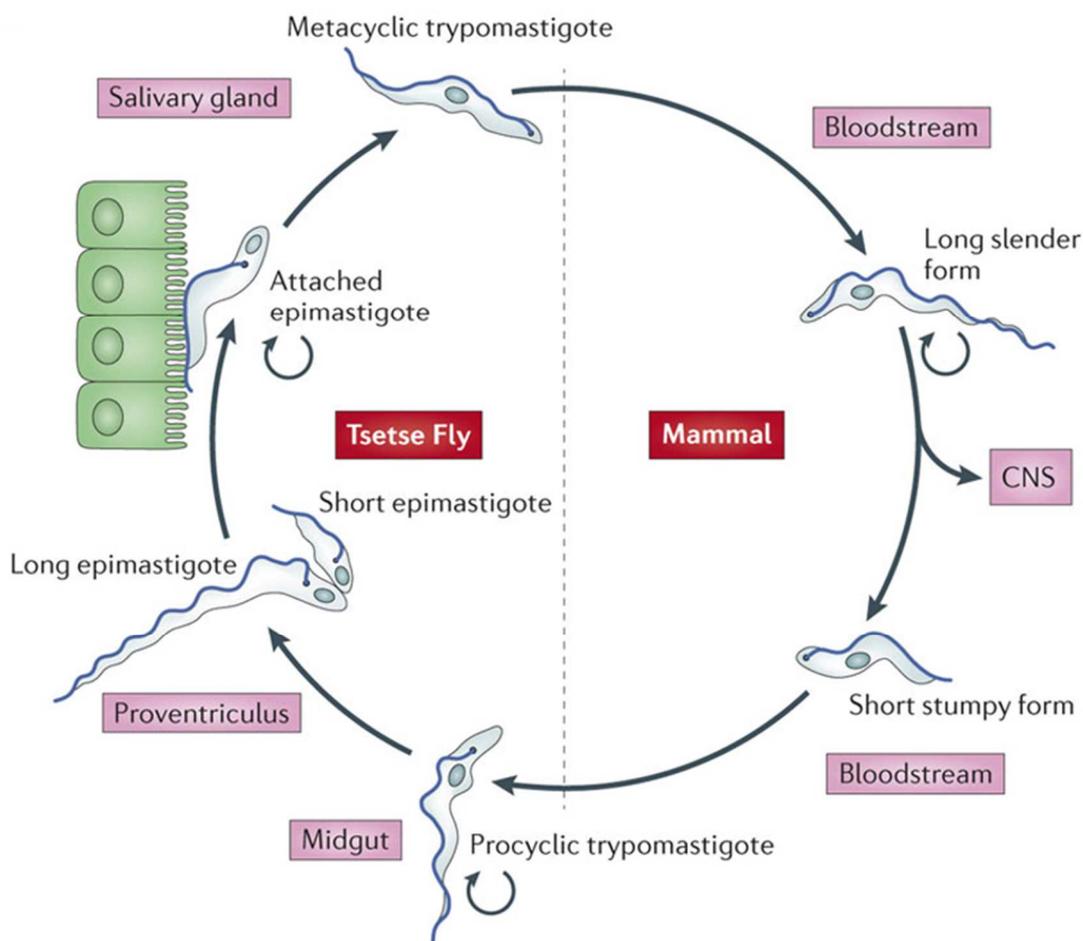
Previous analyses of the three aforementioned subspecies of this parasite have found that *T.b. gambiense* is genetically distinct from both *T.b. rhodesiense* and *T.b. brucei* (Baker, 1995). *T.b. rhodesiense* and *T.b. brucei* are morphologically indistinguishable, and are only separated due to the human-infective nature of *T.b. rhodesiense*, which arises due to a single gene, SRA, which encodes a protein that defends the parasite against trypanosome lytic factors in human serum, as explained in more detail below (Balmer *et al.*, 2011).

### 1.1.2 Life Cycle

The African trypanosomes undergo a complex life cycle involving multiple hosts and extreme morphological changes in order to successfully colonise them (fig. 1-1). At least six developmental forms have been described with significantly contrasting cell architecture, intracellular transport, primary metabolism and gene expression (fig. 1-1). The long slender bloodstream form (BSF) infects mammalian hosts, primarily thriving on glucose present in the blood, before penetrating the

mammalian blood-brain barrier (Mulenga *et al.*, 2001). These rapidly dividing cells possess a protein coat of variable surface glycoproteins (VSGs), which is utilised to evade the host immune system. At high densities, BSF cells differentiate into a non-dividing form referred to as short stumpy form (SSF) (fig. 1-1) (Matthews *et al.*, 2004). Whilst the mechanisms that initiate differentiation cascades are not fully known, they are thought to involve bacterial-like quorum sensing and a secreted molecule known as stumpy induction factor (SIF), the molecular identity of which, despite intense efforts, remains elusive (Vassella *et al.*, 1997, Fenn & Matthews, 2007).

SSF parasites are taken up by the definitive host, the tsetse fly, and rapidly differentiate into procyclic form (PCF) trypomastigotes. Several significant changes occur in protein expression and metabolism at this stage. VSG expression is abolished and the parasites express EP procyclin, and at a later stage, GPEET procyclin (Vassella *et al.*, 2001). These are non-variable forms of surface proteins, which aid in evasion against the rudimentary tsetse immune system (Richardson *et al.*, 1988). In addition, central carbon metabolism shifts from glycolysis to a semi-complete TCA cycle that utilises proline as a carbon source (Bringaud *et al.*, 2006).



**Figure 1-1: Life cycle of *Trypanosoma brucei*.** BSF parasites in the mammalian host reside in the blood, lymph, and in later stages of infection, the CNS. Some BSF cells differentiate into SSF cells which are pre-adapted to life in the tsetse fly vector. Once taken up by feeding tsetse flies, the PCF trypanosomes colonise the fly midgut, and continue to divide. PCFs differentiate into the epimastigote stage and travel to the tsetse salivary glands where they differentiate once more into a trypomastigote stage that is pre-adapted for mammalian infection. Figure is taken from Langousis & Hill (2014).

At this stage, cells are again proliferative. PCFs subsequently migrate to the salivary glands, where they differentiate once more to an epimastigote, and importantly, sexual recombination has been shown to occur in several trypanosome species, with the exception of one class of *T.b. gambiense*, which is thought to be clonal (Koffi *et al.*, 2009, Weir *et al.*, 2016). Cells then differentiate to a metacyclic trypomastigote stage, which are non-dividing and infectious to mammalian hosts (Roditi & Lehane, 2008). Parasites infect new hosts during tsetse fly feeding, to re-initiate the life cycle.

All African trypanosome species exhibit a life cycle similar to that described, with the exception of the animal-infective *T. vivax* and *T. evansi*. These species do not undergo a PCF stage, instead colonising the tsetse fly salivary glands as epimastigotes only. In contrast, the other animal infective African trypanosome, *T. congolense* exhibits a life cycle similar to that of *T. brucei* (Jackson *et al.*, 2015).

## 1.2 Human African Trypanosomiasis

---

### 1.2.1 Historical overview

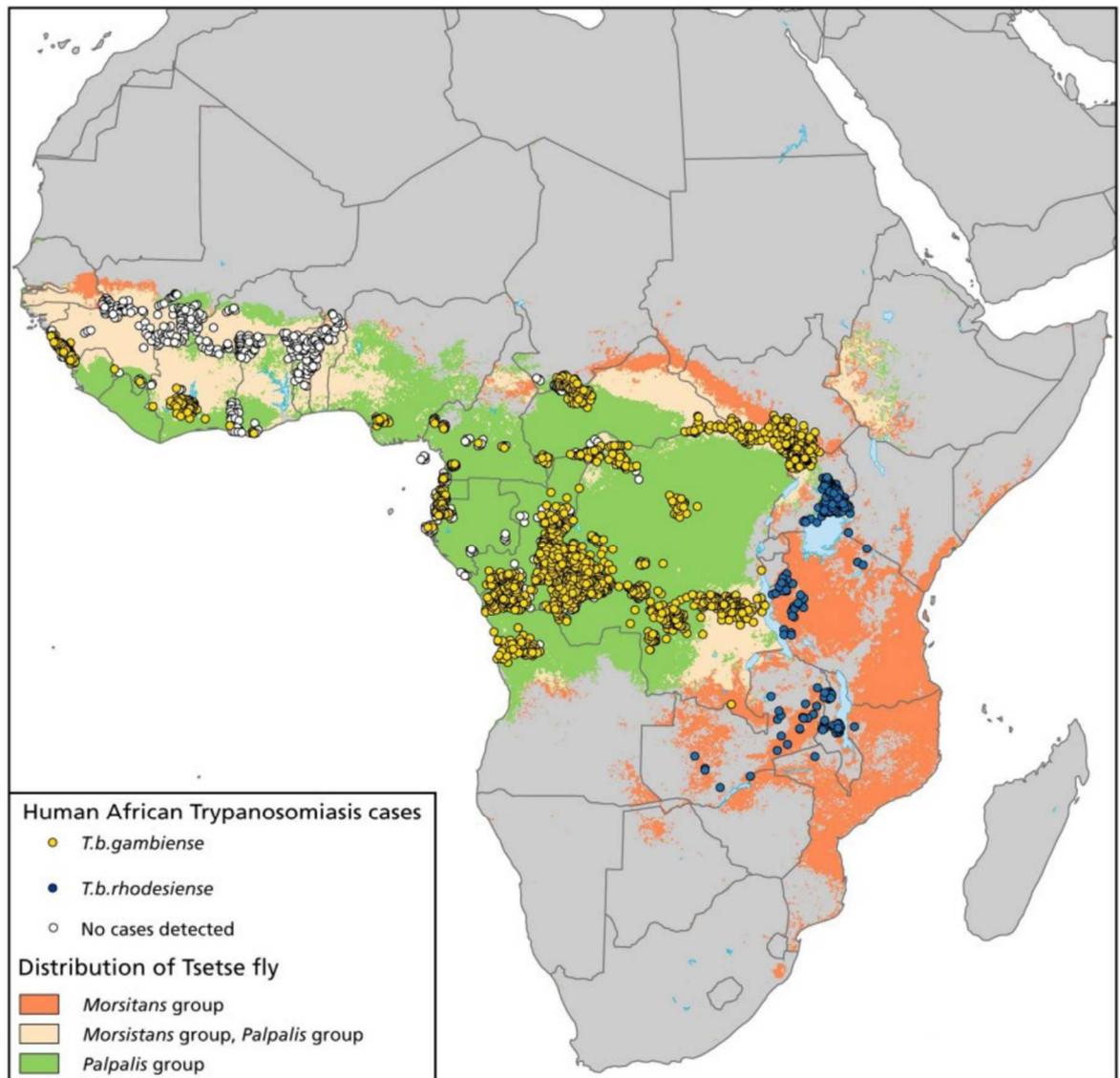
Phylogenetic analyses of the trypanosome have concluded that the salivarian trypanosomes separated from other trypanosomes around 300 million years ago (Haag *et al.*, 1998). Whilst these parasites inhabited the guts of early insects, they were not spread to mammals until the appearance of the tsetse fly, some 35 million years ago (Steverding, 2008).

The first detailed documented report on sleeping sickness was made by naval surgeon Jon Atkins in 1721 (Cox, 2002). More than 100 years later, in 1841, the trypanosome was discovered by Gabriel Valentin, who observed them during routine observations of salmon blood (Steverding, 2008). However, it was not until research carried out by David Bruce that trypanosomes were proven to be the cause of Nagana (Cox, 2002). The first trypanosome infection that was isolated from a human patient was in 1902, by Robert Michael Ford, who, with the physician Joseph Everett Dutton, named the trypanosome *Trypanosoma brucei gambiense* (Dutton, 1902). Several years later Stephens and Fantham described *T.b. rhodesiense* and further work established the role of the tsetse fly in transmission, building on previous studies by David Bruce (Kleine, 1909, Stephens & Fantham, 1910).

During the evolution of the trypanosome, most African mammalian wildlife has evolved tolerance to the parasite, and these animals show no symptoms. In stark contrast, domesticated mammals do not possess mechanisms of trypanotolerance and the majority succumb to Nagana disease (Murray *et al.*, 1982). Humans, have evolved mechanisms of resistance to all subspecies of African trypanosome, with the exception of *T.b. gambiense* and *T.b. rhodesiense* (de Greef & Hamers, 1994, Capewell *et al.*, 2013a).

## 1.2.2 Epidemiology

HAT, as well as the majority of AAT, is restricted to areas where the tsetse fly is prevalent, in sub-Saharan Africa (fig. 1-2). This area, known as the tsetse belt, comprises some 8 million km<sup>2</sup> (Wertheim *et al.*, 2012). Only *T. vivax* and *T. evansi*, which can be transmitted mechanically by other biting flies, and *T. equiperdum*, transmitted venereally in horses, are transmitted outside of this area of Africa (Jones & Davila, 2001, Desquesnes *et al.*, 2013).



**Figure 1-2: Epidemiology of HAT in the sub-Saharan African continent.** This figure, from 2012, shows the areas where *T. b. gambiense* (yellow spots) and *T. b. rhodesiense* (blue spots), the human-infective trypanosomes, are prevalent. In addition, the figure shows how the epidemiology of the disease is restricted by the area inhabited by the tsetse fly vector. Taken from Wertheim *et al.* (2012).

*T. b. gambiense* infection is responsible for approximately 98% of HAT cases in 24 countries, with many of these (~75%) occurring in the Democratic Republic of the Congo. In contrast, *T. b. rhodesiense* is only responsible for 2% of all HAT cases and affects populations in 13 countries (Simarro *et al.*, 2011). In the case of *T. b. rhodesiense*, Uganda and Malawi carry the highest burden

of the disease with the former presenting the only country to report both human variants of the parasite (WHO, 2015a).

Historically, there have been several epidemics of HAT including the most recent occurrence between 1970 and 1990 (WHO, 2015b). This was blamed on social insecurity as well as human conflicts. However, concentrated efforts by WHO as well as other non-governmental organisations, meant that after 38,000 cases were reported in 1998, these numbers dropped to <10,000 in 2009, 6,314 in 2013 and 3,796 in 2014 (fig. 1-3) (WHO, 2015a). *T. gambiense* HAT was targeted for elimination by 2020 under the WHO Roadmap on NTDs (WHO, 2012), although this date has since been revised to 2030 (WHO, 2013). Evidence suggests significant underreporting of the disease across sub-Saharan Africa and there are approximately 65 million individuals at risk across the continent (WHO, 2015b).

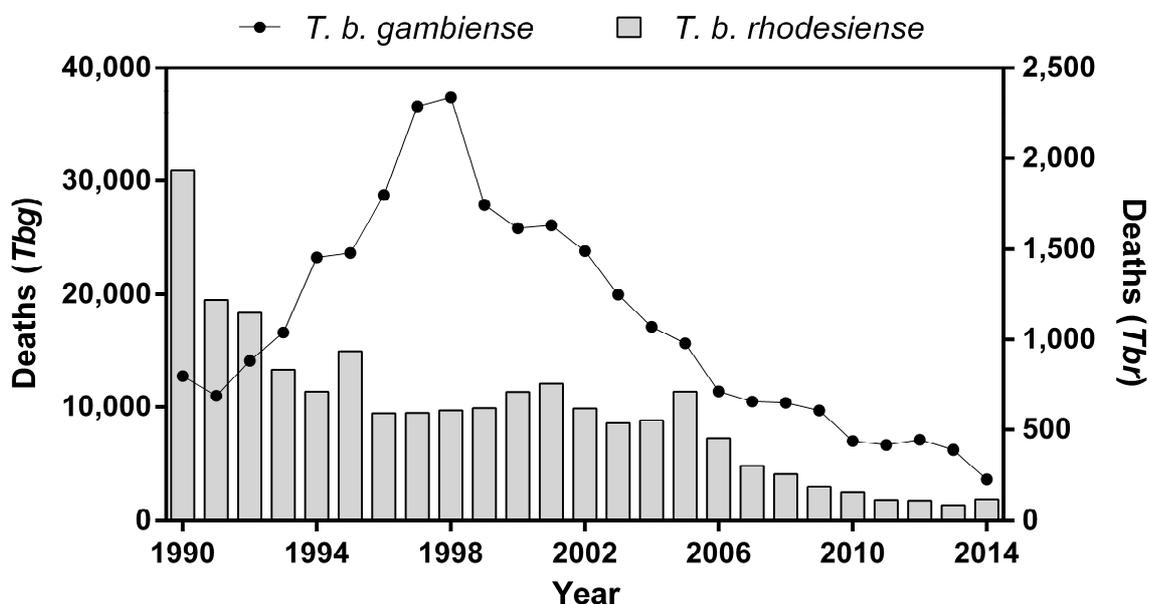


Figure 1-3: Deaths resulting from *T. b. gambiense* (Tbg) and *T. b. rhodesiense* (Tbr) infections between 1990 and 2014. Data taken from a WHO report (WHO, 2015a). Whilst the 1990s saw a resurgence of deaths as a result of *T. b. gambiense* infection (likely due to increased reporting), in general, there has been a decreasing trend in human deaths caused by African trypanosomes.

Restricting the distribution of the fly is one method of disease control (Tirados *et al.*, 2015). The main genus of the vector is *Glossina spp.*, which survive at specific temperatures (16°C–38°C) and humidity (50%–80%) (Franco *et al.*, 2014). It is due to these factors that HAT has not spread further north past the Sahara. Furthermore, wild mammals provide a significant reservoir, in particular for *T. b. rhodesiense*, and to a lesser extent, *T. b. gambiense* (Welburn *et al.*, 2001, Njiokou *et al.*, 2010).

### 1.2.3 Clinical manifestations & diagnosis

Whilst HAT is commonly referred to as Sleeping Sickness, this only partially describes the clinical symptoms of *T. brucei* infections (Kennedy, 2013). Clinically, HAT infection can be broken into two

stages. During the early stage (stage 1), also known as the haemolymphatic stage, the parasite infects the bloodstream and lymphatic vessels, where it thrives on glucose present in mammalian blood (Brun *et al.*, 2010). This is followed by the late stage (stage 2), the CNS stage, where the parasite penetrates the blood-brain barrier (Masocha *et al.*, 2007).

The clinical manifestations arising from HAT exposure can vary significantly between the two human-infective variants (Kennedy, 2004). Whilst *T.b. gambiense* is responsible for a chronic infection, often lasting months or years, *T.b. rhodesiense* infection is acute and severe, and can lead to fatalities in a matter of weeks to months (Kennedy, 2004). However, in both cases, early stage symptoms can be non-specific such as headaches, weight loss, fatigue and intermittent fevers (Kennedy, 2013). As the infection progresses, symptoms gradually become more severe and can include, but are not limited to, lymphadenopathy, cardiac complications, conjunctivitis, endocrine dysfunction, alopecia, hepatomegaly and splenomegaly (Lejon *et al.*, 2013).

During the late stage of the disease, symptoms become increasingly severe. This can mainly be attributed to neurological complications as a result of the parasite spreading to the CNS (Reviewed by (Kennedy, 2004)). These manifestations include mental or motor system disturbances as well as the typical sleep disturbances for which the disease is named, although the latter does not occur in every case (Blum *et al.*, 2006).

Rapid diagnosis is crucial to ensure patient recovery. However, because the early stage symptoms are not specific to trypanosome infection, early diagnosis is challenging and other tropical diseases such as malaria, leishmaniasis, TB and others, must also be excluded (Kennedy, 2004). The most commonly used assay for diagnosis is a card agglutination test for trypanosomiasis (CATT) (Magnus *et al.*, 1978). However, this assay has several disadvantages. In particular, the test does not apply to *T.b. rhodesiense* as it uses *T.b. gambiense*-specific antigens, and there is limited sensitivity (Dukes *et al.*, 1992). In addition, false positive results are often reported (Chappuis *et al.*, 2005, Brun *et al.*, 2010). For these reasons, microscopic identification of parasites in patient blood or lymph is often required for diagnosis (Mitashi *et al.*, 2012, Mukadi *et al.*, 2013), but low parasitaemias, especially in the case of *T.b. gambiense*, can confound this (Kennedy, 2004). Other diagnostic tools such as PCR have also been utilised, although their practicality in the field remains debatable (Mugasa *et al.*, 2012). Rapid tests, using lateral flow devices, are showing increasing promise, although parasitological identification post serology still seems necessary (Sternberg *et al.*, 2014).

Different therapeutic regimens exist depending on the subspecies that is responsible for a HAT case, as described below. This makes accurate diagnosis crucial, especially in areas such as Uganda, where both variants of the human-infective trypanosomes are prevalent (Kennedy, 2013). In

addition, differential diagnosis of early and late stage HAT is equally important, as different therapies exist for both, as described in [section 1.3](#) (Lejon *et al.*, 2013).

#### 1.2.4 Human mechanisms of resistance

The mechanisms of resistance to human serum-mediated lysis for both human-infective strains have been elucidated (de Greef & Hamers, 1994, Capewell *et al.*, 2013a). African trypanosome species such as *T.b. brucei* succumb to the effects of a human serum component called trypanosome lytic factor (TLF) (Raper *et al.*, 1999). To date, two of these factors have been described, termed TLF-1 and TLF-2. The mechanism of TLF-1 mediated trypanosome killing is understood to a better extent than TLF-2. TLF-1 is taken up via endosomes through the trypanosome haptoglobin haemoglobin receptor (TbHpHbR), and trafficked to the lysosome (Drain *et al.*, 2001, Vanhollebeke *et al.*, 2008). Here, TLF is activated by acidic pH, which leads to lysosomal membrane destabilisation and degradation. It is thought that apolipoprotein L-1 (apoL-1), one of the components of TLF-1, forms pores in the lysosomal membrane, allowing the contents to leak into the cytoplasm (Vanhamme *et al.*, 2003, Perez-Morga *et al.*, 2005). Ultimately, this leads to trypanosome cell death.

Interestingly, resistance to this killing mechanism has arisen on three independent occasions, in *T.b. rhodesiense* and *T.b. gambiense* (Stephens *et al.*, 2012). In the case of the former, it was shown that this species possesses a VSG-like molecule, termed the serum resistance (SRA) gene (de Greef & Hamers, 1994, Xong *et al.*, 1998). Current hypotheses suggest that whilst TLF-1 uptake still occurs, the molecule does not localise to the lysosome, a process that is likely inhibited by the binding of SRA to apoL-1 (Vanhamme *et al.*, 2003). The mechanism for *T.b. gambiense* resistance to human serum is not as clear, although it has been shown that another VSG-like molecule called TgsGP confers serum resistance (Capewell *et al.*, 2013a, Uzureau *et al.*, 2013). In addition, a study also reported downregulation of the TbHpHbR gene in human serum resistant *T.b. gambiense*, which could lead to a reduction in TLF-1 uptake (Kieft *et al.*, 2010). *T.b. brucei* does not possess either of the human serum resistance genes, which explains why this subspecies does not infect humans (Xong *et al.*, 1998).

### 1.3 Therapeutics/chemotherapy against HAT/AAT

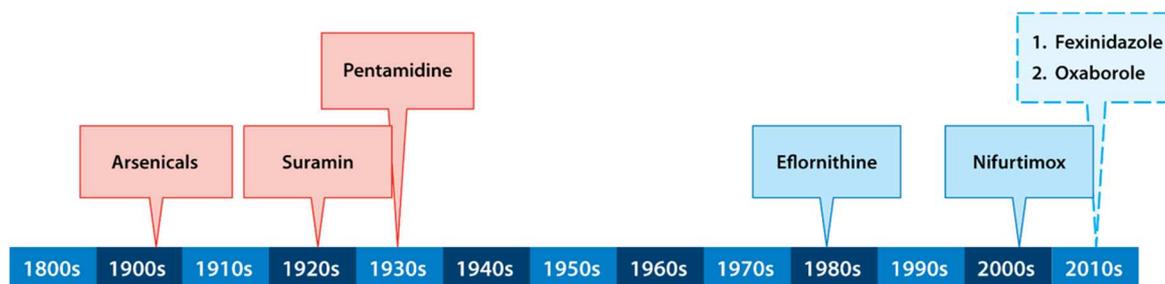
---

#### 1.3.1 Human African Trypanosomiasis (HAT)

HAT was until recently considered to be fatal in close to 100% of untreated cases, although recent evidence in the case of *T.b. gambiense* has challenged this theory (Kabore *et al.*, 2011, Jamonneau *et al.*, 2012, Ilboudo *et al.*, 2014). In any case, this fact underlies the desperate need for therapeutics to combat the disease. The majority of current therapies target the mammalian stage of the

parasite, although some work has also focused on blocking the transmission cycle of *T. brucei* as a form of preventative medicine (Tirados *et al.*, 2015, Vale *et al.*, 2015).

Current regimens used to treat HAT consist of species- and stage-dependent chemotherapeutics which are outlined in figure 1-4. Treatments for early stage HAT usually involve pentamidine and suramin for *T.b. gambiense* and *T.b. rhodesiense* respectively (Kennedy, 2013). Pentamidine is a diamidine administered intramuscularly, with good efficacy (cure rate ~93-98%) (Balasegaram *et al.*, 2006). However, treatment with pentamidine can lead to potential cardiac malfunctions as well as hypotension and hypoglycaemia amongst other side-effects (Babokhov *et al.*, 2013). Similarly, side effects to suramin treatment, which is administered by intravenous injection, have been described in the form of renal failure, anaphylactic shock and neuronal complications (Babokhov *et al.*, 2013). Both these therapies involve strict regimens of at least seven days, thereby requiring patients to remain in a clinical setting for the duration of the therapy. This is challenging to achieve in a rural environment (Mwanakasale *et al.*, 2013).



**Figure 1-4: A timeline of drugs used against HAT, and when they were first brought to market.** Eflornithine and nifurtimox, both for the treatment of *T.b. gambiense* late-stage infection, are newest drugs currently available. Red boxes indicate reports of resistance in the field. Dotted border indicates drugs currently in clinical trials. Figure taken from Horn & Duraisingh (2014).

Late stage HAT presents an even greater challenge to combat, given that drugs developed against this particular stage must cross the blood-brain barrier. The most recent drug brought to market for late-stage *T.b. gambiense* infection is eflornithine, an analogue of ornithine, which targets the ornithine decarboxylase enzyme (Phillips & Wang, 1987, Vincent *et al.*, 2012). Developed initially as an anti-tumor drug as well as a therapy against facial hirsutism, eflornithine is thought to have an efficacy of up to 95% , although it has only limited potency against late-stage *T.b. rhodesiense* infection (Babokhov *et al.*, 2013).

The only front-line therapeutic specifically against late-stage *T.b. rhodesiense* is melarsoprol, a trivalent arsenical (Friedheim, 1949). This drug can elicit significant toxicities, leading to fatal side effects such as encephalopathy. It is estimated that up to 10% of patients succumb not to the parasite, but rather the treatment (Pepin & Milord, 1994). Melarsoprol is a relatively old drug compared with eflornithine, having been introduced in 1949, and resistance to this compound has

been reported in the field, with some *T.b. gambiense* endemic areas showing almost 30% treatment failure (Lejon *et al.*, 2013).

More recently, eflornithine has been used in combination with nifurtimox, a 5-nitrofuranyl drug that can be administered orally (Legros *et al.*, 2002). Whilst nifurtimox is not effective as a monotherapy, patients can be treated with a lower dose of eflornithine, whilst maintaining a similar efficacy and cure rate, if nifurtimox is used simultaneously (Priotto *et al.*, 2009).

As shown in fig 1-4, the majority of drugs used to treat HAT were introduced almost half a century ago, with no significant improvements since. As several reviewers have argued, this can be attributed to underinvestment by governments and pharmaceuticals, given the weak prospect of financial gain (Kennedy, 2013). However, the turn of the 21<sup>st</sup> century has brought changes, due in large parts to the efforts of many non-governmental organisations such as the Drugs for Neglected Tropical Disease initiative (DNDi) and the Bill and Melinda Gates foundation, as well as academic partnerships with significant input from several pharmaceutical companies (Barrett, 2010). Indeed, as mentioned previously, these efforts are helping to lower the number of annual cases, and with two new drugs in clinical trials for HAT, the future is looking brighter.

### **1.3.2 Drugs currently in clinical trials**

#### **1.3.2.1 Fexinidazole**

Fexinidazole (1-methyl-2-((p-(methylthio)phenoxy)methyl)-5-nitroimidazole) (fig. 1-5) was previously under preclinical development in the early 1980s as a broad-spectrum antimicrobial agent (Torreele *et al.*, 2010). At this time, this drug was reported to exhibit anti-trypanosomal activity, but its development as a therapeutic against HAT was not pursued (Jennings & Urquhart, 1983). Currently, the drug is in phase III trials in the Democratic Republic of the Congo, as well as Central African Republic, and close to registration (Tarral *et al.*, 2014). Whilst the target of fexinidazole is not currently known, it is thought that the compound acts in a similar fashion to other nitro prodrugs like nifurtimox, which requires activation in the cell via a trypanosomal mitochondrial nitroreductase (Sokolova *et al.*, 2010, Wyllie *et al.*, 2015). Several studies have looked in detail at the potency of fexinidazole, as well as its sulfoxide and sulfone derivatives which are rapidly produced upon administration of the parent compound *in vivo* (Torreele *et al.*, 2010, Kaiser *et al.*, 2011, Tarral *et al.*, 2014). Doses of the compound up to 3,600 mg are safe and well tolerated, and also shows activity against visceral leishmaniasis and Chagas disease (Tarral *et al.*, 2014). One potential disadvantage of the drug is that generation of cross-resistance to nifurtimox in the field could have a knock-on effect leading to decreased efficacy of fexinidazole (Sokolova *et al.*, 2010).

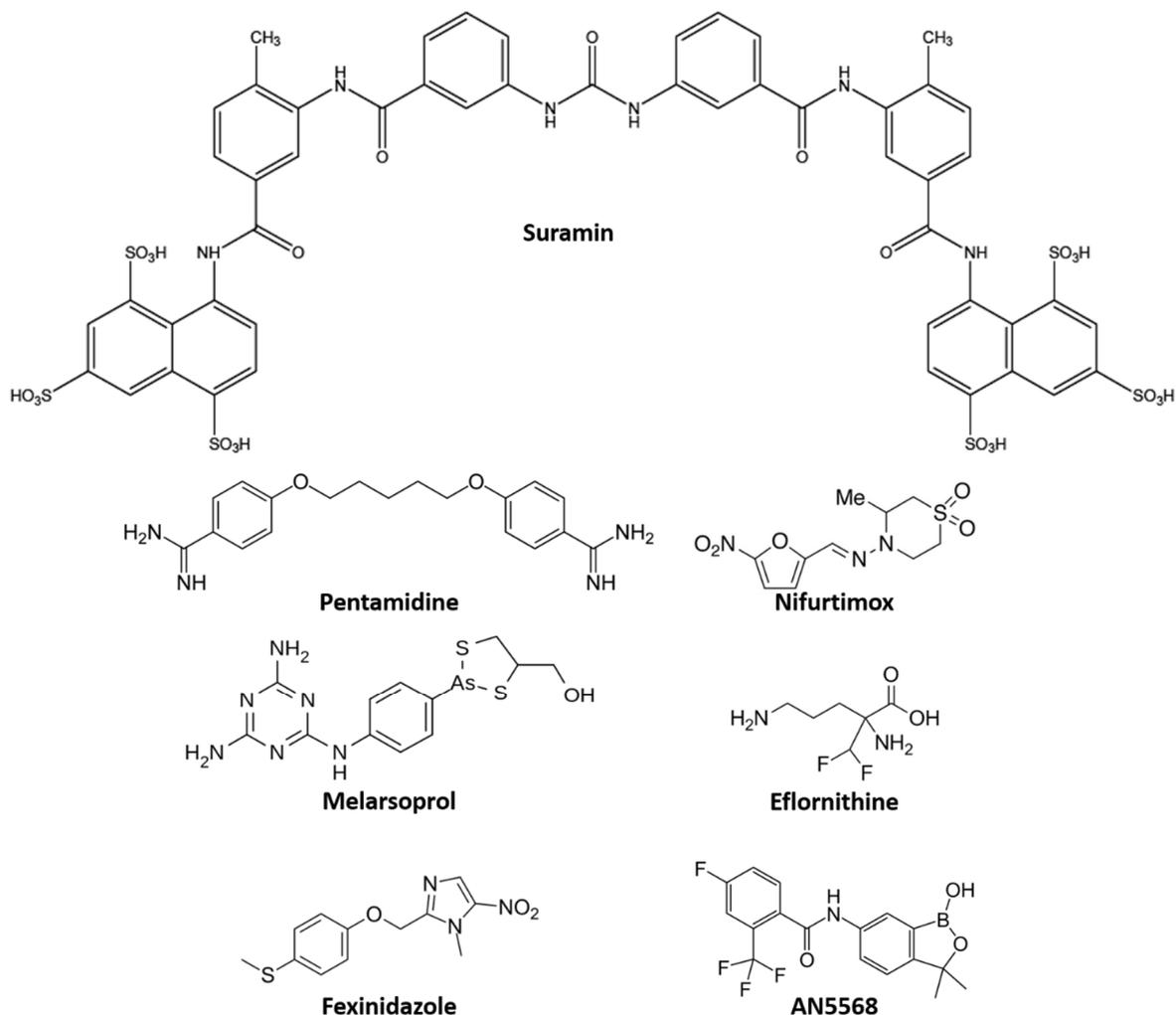


Figure 1-5: Two-dimensional chemical structures of drugs commonly used against HAT.

### 1.3.2.2 Benzoxaboroles – a unique scaffold with great potential

Benzoxaboroles have only recently come to wider attention, although this class of compounds was first synthesised more than half a century ago (Baker *et al.*, 2011, Liu *et al.*, 2014). These compounds possess a unique chemistry offered by the presence of boron in their structure (Nare *et al.*, 2010). The structural motif has been used in a wide range of applications including, but not limited to, antivirals (Mahalingam *et al.*, 2011), antifungals (Markham, 2014), antibacterials (Goldstein *et al.*, 2013, Hernandez *et al.*, 2013), kinase inhibitors (Akama *et al.*, 2013b, George *et al.*, 2015), anti-inflammatories (Akama *et al.*, 2013a, Dong *et al.*, 2013), and antiparasitics (Ding *et al.*, 2010, Zhang *et al.*, 2015). In many cases, there are few toxicity or instability problems related to the in the compounds *in vivo*. The use of benzoxaboroles is by no means limited to the development of novel therapeutics, as is reviewed by Liu and colleagues (Liu *et al.*, 2014). Indeed, the high affinity for sugars and diols that benzoxaboroles possess has led to their application in materials science for chromatographic separation-based work (Li *et al.*, 2012).

In many cases there is only a basic knowledge regarding the SAR of these compounds in a therapeutic setting, and there is much less known about the exact mechanisms of drug action on

pathogens. Several studies have looked at boronic acids, from which benzoxaboroles are derived, but to date, only one benzoxaborole target has been validated (Rock *et al.*, 2007). Tavaborole, (AN2690), is used as a treatment against onychomycosis, a fungal infection affecting the toenail, and has recently been approved by the FDA (FDA, 2014). Through the generation and characterisation of AN2690-resistant yeast mutants as well as the resolution of crystallography structures, Rock and colleagues were able to show that AN2690 inhibits the cytoplasmic leucyl-tRNA synthetase by formation of a stable tRNA<sup>LEU</sup>-AN2690 adduct localised in the editing site of the enzyme (Rock *et al.*, 2007). This adduct was mediated by the boron from the oxaborole ring which bound to the diol group of the 3'-terminal adenosine (Rock *et al.*, 2007), thus highlighting the importance of the boron group within the structure, as well as the high affinity benzoxaboroles exhibit for diol groups. However, the group did not discuss what led to the attachment of AN2690 to tRNA<sup>LEU</sup> rather than other molecular structures exhibiting exposed diol groups.

### 1.3.2.3 Benzoxaboroles for neglected tropical diseases

Anacor Pharmaceuticals has been developing drugs against neglected tropical diseases (NTDs) based on the benzoxaborole scaffold, in collaboration with the DNDi as well as the Gates foundation. Recent publications highlighted potential new scaffolds with great potency against *Plasmodium falciparum*, the causative agent of the most virulent strain of malaria (Zhang *et al.*, 2011b). Working with the Medicines for Malaria Venture (MMV), two benzoxaboroles were selected for pre-clinical screening. Further efforts to develop therapeutics against leishmaniasis and Chagas disease, as well as tuberculosis, river blindness and shigellosis, are also underway (Anacor, 2015). In addition, drugs against NTDs in animals, such as *T. congolense* and *T. vivax*, are under development, and leads have been selected for pre-clinical screening (Barrett, personal communication).

Of the compounds in the Anacor NTD portfolio, the most promising lead for NTDs so far is a benzoxaborole currently awaiting phase II clinical trials as a therapeutic against HAT. This compound, AN5568 (formerly known as SCYX-7158) (fig. 1-5), exhibits low levels of host toxicity, and readily clears both *T.b. gambiense* and *T.b. rhodesiense* infections in mouse models (Jacobs *et al.*, 2011). More importantly, it is able to clear parasites in both early and late stage HAT. Further *in vivo* studies concentrating on reproductive toxicology have also determined there are no adverse effects on male and female fertility, mating behaviour or peri/postnatal development (Giannotti *et al.*, 2014). Whilst the drug appears to be secreted in maternal milk, Giannotti and colleagues found no adverse symptoms in F1 progeny (Giannotti *et al.*, 2014). These data, in addition to those published previously, demonstrate that AN5568 possesses a good safety profile (Jacobs *et al.*, 2011, Giannotti *et al.*, 2014). In addition, unlike current therapies, AN5568 is orally bioavailable (Jacobs

*et al.*, 2011). These factors make the benzoxaborole a very exciting candidate as a therapeutic against HAT.

### 1.3.3 Animal African Trypanosomiasis (AAT)

The outlook for therapeutics against AAT, or Nagana, is much poorer compared to HAT. The two drugs currently used for the majority of cattle cases are diminazene aceturate and isometamidium chloride. The latter was introduced in 1961, but remains the most recent novel therapy against AAT (Berg *et al.*, 1961). In some cases, ethidium bromide, marketed as homidium, is also used, although widespread resistance has been reported (Scott & Pegram, 1974). Treatment failure, presumably due to emerging resistance, has been increasingly reported in recent years for all drugs (Geerts *et al.*, 2001). Due to differences between trypanosome species, the drugs commonly used against HAT have little effect against Nagana (Auty *et al.*, 2015). Clearly, there is also a critical need for novel therapeutics against the animal trypanosomes.

## 1.4 Trypanocides and their mechanisms of action

---

Drug target deconvolution (also referred to as drug target identification, or mechanism of action [MoA] studies) has become increasingly important with regards to the development of novel therapies. Understanding the mechanisms by which drugs acts to kill pathogens leads to a greater understanding of the biology of the organism, which in turn can lead to the development of more potent therapies with increased efficacy and reduced side- and off-target effects (Hughes *et al.*, 2011). In the study of trypanosomatids, there is a severe lack of knowledge regarding drug MoAs. This is in part due to the lack of optimised techniques for investigating drug MoAs in the kinetoplastids, and also due to lack of research specifically aimed at identifying detailed mechanisms.

### 1.4.1 Eflornithine ( $\alpha$ -difluoromethylornithine, DFMO)

As mentioned previously, eflornithine is the current front-line therapy against *T.b. gambiense* and targets ornithine decarboxylase (ODC) (Li *et al.*, 1998). This compound is an analogue of ornithine, a precursor for polyamine and subsequently, trypanothione biosynthesis in the parasite. Recent metabolomics analyses by the Barrett lab have confirmed the effect of eflornithine on this pathway, leading to a build-up of putrescine and a decrease in spermidine (Vincent *et al.*, 2012). Earlier work by Wang and colleagues demonstrated that an ODC-null mutant, whilst unable to establish an *in vivo* infection in mice, was desensitised to eflornithine (Li *et al.*, 1998). In a separate study, Wang showed that eflornithine-resistant cells retained ODC activity, but exhibited significantly reduced uptake of the drug (Phillips & Wang, 1987). These data were supported by more recent work carried out by Vincent and colleagues (Vincent *et al.*, 2010, Vincent *et al.*, 2012).

### 1.4.2 Pentamidine

Pentamidine is a diamidine that has been used in both HAT and AAT. However, the mode of action of this drug is still not fully known (Berger *et al.*, 1993, Barrett *et al.*, 2007). Uptake is carrier-mediated through the P2 amino purine transporter, the high-affinity pentamidine transporter (TbHAPT1) (now known to be encoded by the TbAQP2 gene), and the low-affinity pentamidine transporter (TbLAPT1) (De Koning, 2001, de Koning & Jarvis, 2001, Munday *et al.*, 2014). Published data indicates that diamidines are nucleic-acid binding drugs that localise to the mitochondrion, where they destroy the kinetoplast, and the mitochondrial DNA (Ludewig *et al.*, 1994). Whilst this drug remains effective in the field, loss of the kinetoplast in trypanosomes, which has been observed *in vitro* and in other trypanosome species like *T. evansi* (Lai *et al.*, 2008), could lead to widespread resistance (Horn & Duraisingh, 2014). In reality, loss of transporters have played a more significant role in generating resistance (Graf *et al.*, 2013).

### 1.4.3 Melarsoprol

Melarsoprol has multiple modes of action in both host and parasite (Alsford *et al.*, 2012). In trypanosomes, melarsoprol is metabolised to melarsen oxide, which binds to trypanothione, forming a stable adduct which competitively inhibits the enzyme trypanothione reductase, an essential protein involved in the regulation of the thiol/disulphide redox balance in trypanosomatids (Fairlamb *et al.*, 1989). In addition, the oxide derivative of melarsoprol also binds other thiol groups, thereby inhibiting many enzymes in both the host and the parasite, leading to the widespread side effects seen in HAT patients treated with the drug (Fairlamb *et al.*, 1992).

### 1.4.4 Suramin

Whilst several potential MoAs have been proposed for suramin, none of these have been proven experimentally (Barrett *et al.*, 2007). The drug possesses significant points of negative charge, leading to electrostatic interactions with many enzymes (Barrett *et al.*, 2007). In addition, the drug is not able to enter the cell by passive diffusion, and endocytosis was proposed as a route of entry (Vansterkenburg *et al.*, 1993). A recent study investigating drug MoA using the RIT-seq approach implicated the invariant surface glycoprotein, ISG75, as a major receptor for suramin uptake. In addition, drug resistance was observed in cells lacking components of the endocytosis pathway (Alsford *et al.*, 2012).

Previous work indicates that suramin is far more potent against BSF trypanosomes, compared with PCF parasites (Scott *et al.*, 1996). This led to the hypothesis that glycolysis could be targeted, as glycolytic metabolism is essential to BSF trypanosomes only (Fairlamb & Bowman, 1977, Besteiro *et al.*, 2005). However, suramin has been shown to inhibit a wide range of enzymes, and it is

unknown whether inhibition of glycolysis is the primary mechanism resulting in trypanosome death (Alsford *et al.*, 2012).

## 1.5 Cell biology

---

### 1.5.1 Antigenic variation

Trypanosomes live in an extracellular environment and are continuously exposed to the host immune system. To counter this, the parasite has evolved a unique and effective method to evade host immunity, using a switchable array of cell surface proteins called variable surface glycoproteins (VSGs), which were first discovered in 1969 (Vickerman & Luckins, 1969). Approximately 10 million copies of these 60 kDa homodimeric proteins cover the cell surface of the parasite at any one time, and the protein accounts for up to 20% of the entire cell proteome (Horn, 2014). VSGs are not transmembrane proteins but instead, are anchored to the cell surface membrane by glycosylphosphatidylinositol (GPI) moieties (Ferguson *et al.*, 1988). The host immune system recognises VSGs as a foreign antigen, which it combats with high levels of B-cell derived antibodies (Campbell *et al.*, 1977, de Gee *et al.*, 1983, Reinitz & Mansfield, 1990). However, trypanosomes possess many hundreds of genes coding for VSGs. These, in addition, can undergo recombination, leading to further variation (Marcello & Barry, 2007, Hall *et al.*, 2013). By predominantly expressing only one VSG at a time, the host immune response remains concentrated against one VSG structure. This results in the killing of high numbers of trypanosomes, whilst those expressing different VSGs survive to re-initiate infection (Horn, 2014). The VSG coat is highly immunogenic, but it also serves to protect non-variant surface proteins from immune effectors by physically obstructing access to these proteins (Horn, 2014). It is the continuous switching of dominantly expressed VSGs and the subsequent host immune responses generated against each VSG structure that are responsible for the intermittent parasitaemia often observed in the clinic (Black *et al.*, 1985, Vickerman, 1985, Emmer *et al.*, 2010).

Interestingly, whilst the human immune system does not possess antibodies capable of recognising epitopes of trypanosome molecules other than VSG, cattle have been shown to possess antibodies with an exceptionally long CDR H3 loop, the most diverse portion of a classical antibody (Wang *et al.*, 2013). It is thought these could potentially bypass the VSG coat and recognise other molecules on the trypanosome cell surface, making them important in the context of novel therapeutics and, potentially, vaccines (L. Morrison, personal communication).

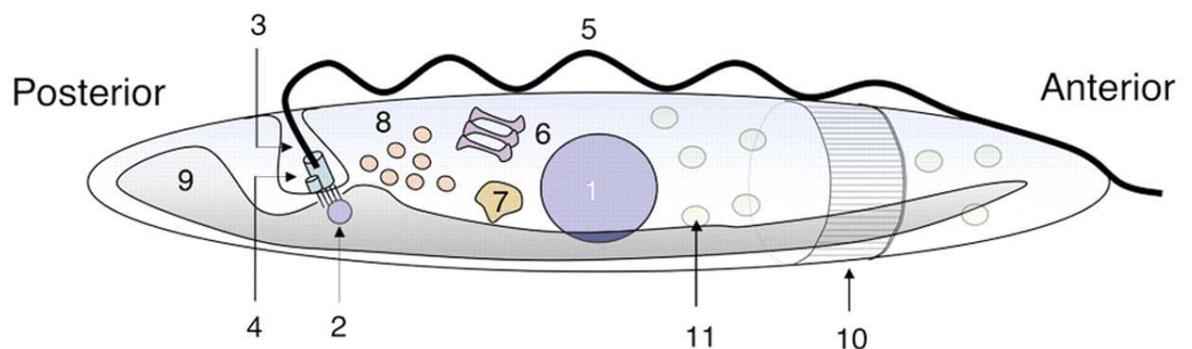
### 1.5.2 Overview of *T. brucei* cell biology and organelles

The majority of the *T. brucei* forms are spindle-shaped, with a polarized microtubule cytoskeleton (fig. 1-6), and possess a single elongated flagellum, which emerges through a flagellar pocket (Matthews, 2005, Field & Carrington, 2009). This pocket is a small plasma membrane invagination

close to the basal body and is the only site for endo- and exo-cytosis (Overath & Engstler, 2004) (fig. 1-6). The motility of the parasite is driven by a flagellar wave that initiates at the tip of the flagellum and moves towards the base (Weisse *et al.*, 2012). It has recently been shown that BSF trypanosome locomotion is adapted to the density of blood (Heddergott *et al.*, 2012). Interestingly, Heddergott and colleagues used high-speed video microscopy to show that the presence of blood cells is required for the parasite to achieve maximum velocity in the bloodstream (Heddergott *et al.*, 2012). Thus, motility is not solely due to flagellar motion. The flagellum itself is made up of a conventional axonemal structure along with a paraflagellar rod (Vaughan & Gull, 2003).

Whilst *T. brucei* possesses a conventional nucleus, endoplasmic reticulum and Golgi body, the cell has an elongated single mitochondrion, rather than multiple smaller mitochondria as observed in yeast and mammalian cells (Acestor *et al.*, 2009) (fig. 1-6). In BSF trypanosomes, the mitochondrion does not contain cristae, although the organelle is much more active in PCF cells, and here it contains an extensive electron transport chain (Matthews, 2005).

*T. brucei*, along with the other kinetoplastids *T. cruzi* and *Leishmania spp.*, possess several unique organelles. A major part of glycolysis is contained in the glycosome, described in more detail below. In addition, the complex network of mitochondrial DNA is encompassed by the kinetoplast, from which this group of organisms takes its name (Shlomai, 2004). Finally, trypanosomatids also possess acidocalcisomes, which are organelles rich in orthophosphate (Pi), PPi and PolyP, complexed with cations and basic amino acids (Docampo *et al.*, 2005).



- |                                  |                              |
|----------------------------------|------------------------------|
| 1. Nucleus                       | 7. Lysosome                  |
| 2. Kinetoplast                   | 8. Endosomes                 |
| 3. Flagellar pocket              | 9. Mitochondrion             |
| 4. Basal body and probasal body  | 10. Microtubule cytoskeleton |
| 5. Axoneme and paraflagellar rod | 11. Glycosomes               |
| 6. Golgi                         |                              |

**Figure 1-6: Cell biology of *T. brucei*.** A simplified diagram showing the locations of the major organelles and structural features of *T. brucei*. Taken from Matthews (2005).

### 1.5.3 Comparing bloodstream form and procyclic form cells

Compared to BSF cells, PCFs are larger and longer, and the mechanics of cell division differ between the two life cycle stages (Wheeler *et al.*, 2013). There are also extreme metabolic differences between BSF and PCF forms. In particular, whilst BSF parasites rely primarily on glycolysis for the generation of ATP, PCFs utilise proline, generating ATP via oxidative phosphorylation (Besteiro *et al.*, 2005).

The localisation of organelles also differs between life cycle stages. In BSF cells, the kinetoplast is localised at the posterior end of the parasite. In contrast, the kinetoplast moves inwards in PCF cells and is localised midway between the nucleus and the posterior end (Matthews, 2005). Another well characterised event during differentiation, is the downregulation of the VSG surface coat, as this is replaced by EP procyclin, and later on, by GPEET procyclin (Butikofer *et al.*, 1997, Vassella *et al.*, 2001).

### 1.5.4 The glycosome

BSF trypanosomes rely heavily on glucose as a primary carbon source (Mazet *et al.*, 2013). To enable high flux through the glycolytic pathway, the parasite has evolved a unique and essential organelle which compartmentalises the first seven steps of glycolysis (Visser *et al.*, 1981). This peroxisome-like organelle, known as the glycosome, contains primarily (~90%) glycolytic enzymes including hexokinases, phosphofructokinase, fructose-bisphosphate aldolase and glyceraldehyde-3-phosphate dehydrogenase (Misset *et al.*, 1986). Previous studies have shown that glucose is toxic to glycosome-deficient trypanosomes (Furuya *et al.*, 2002, Guerra-Giraldez *et al.*, 2002). A *T. brucei* cell typically contains approximately 65 glycosomes in its BSF stage (Tetley & Vickerman, 1991). This number is drastically reduced in PCF stage parasites, reflecting the BSF stage parasite's dependency on this organelle for central metabolism (Haanstra *et al.*, 2015). In addition to glycolysis, the glycosome also contains a number of other pathways including lipid synthesis, the pentose phosphate pathway (PPP), purine salvage and pyrimidine synthesis (Michels *et al.*, 2006).

### 1.5.5 Kinetoplasts and the trypanosome mitochondrion

*T. brucei* possess a large, complex mitochondrial DNA compartment called the kinetoplast. This DNA (referred to as kDNA) consists of two types of circular DNAs termed minicircles and maxicircles (Hines & Ray, 2010). Whilst maxicircles encode ribosomal RNA and mitochondrial-specific genes, similar to other eukaryotes, minicircles code for a myriad of small RNAs referred to as guide RNAs, and number to approximately 5,000 (Carpenter & Englund, 1995, Lukes *et al.*, 2005). These guide RNAs act as templates in the RNA editing of maxicircle transcripts, primarily in the form of uridine additions/deletions (Feagin *et al.*, 1988, Sturm & Simpson, 1990).

In addition to the kDNA, the kinetoplast also contains a variety of house-keeping proteins unique to trypanosomatids. For example, there are four histone-like molecules called kinetoplast associated proteins 1-4 (Xu *et al.*, 1996). More recent work has also identified the key players involved in the unique mechanisms of RNA editing (Ammerman *et al.*, 2013). These kinetoplastid-specific phenomena could arguably become an important therapeutic target in the near future.

## 1.6 Genome and gene regulation in *T. brucei*

---

### 1.6.1 Genome structure and arrangement

*T. brucei* possesses a diploid genome divided into 11 pairs of classical chromosomes of varying size (0.9 – 5.7 Mb) (Berriman *et al.*, 2005). The genome is thought to code for approximately 10,000 genes, many of them multiple copies (Berriman *et al.*, 2005). In addition, *T. brucei* contains varying numbers of intermediate and mini chromosomes made up of approximately 0.3-0.9 Mb and 0.05-0.1 Mb respectively (Berriman *et al.*, 2005). These chromosomes contain the VSG gene repositories, and genes here are usually transcribed by RNA polymerase (Pol) I (Shea *et al.*, 1987, Zomerdijk *et al.*, 1991). In contrast, all other genes are transcribed by RNA polymerase (Pol) II. Pol I initiation can be regulated epigenetically, enabling stage-specific expression, and more importantly in the case of VSGs, exclusive expression of one gene at a time (Clayton, 2014).

Gene arrangement in African trypanosomes differs from mammalian cells, as well as other protozoans. This is due to two phenomena. Firstly, almost all genes are arranged in poly-cistronic arrays that can contain tens to hundreds of genes (Imboden *et al.*, 1987, Berriman *et al.*, 2005). These genes are constitutively transcribed and their expression is regulated at the post-transcriptional level through complex interplay of RNA transcription and degradation (De Gaudenzi *et al.*, 2011). Secondly, trypanosomes do not appear to possess introns, and trans-splicing only occurs during addition of a short 39-nucleotide spliced leader (described below) to the 5' ends of mature mRNAs (Perry *et al.*, 1987). This process occurs simultaneous to polyadenylation of the 3' ends (Huang & van der Ploeg, 1991). To date, only two genes have been discovered that harbour introns: poly(A) polymerase (PAP) and RNA helicase, which are both *cis*-spliced (Mair *et al.*, 2000, Berriman *et al.*, 2005).

In many ways, the arrangement of the *T. brucei* genome is linked to the significant differential expression that occurs in the various extracellular environments the parasite encounters in its different life cycle stages. Several studies have investigated the transcriptomes of BSF and PCF trypanosomes by microarray (Brems *et al.*, 2005, Koumandou *et al.*, 2008, Jensen *et al.*, 2009), as well as RNA-seq (Kolev *et al.*, 2010, Siegel *et al.*, 2010, Hamidou Soumana *et al.*, 2015). These data have shown that most transcripts are present at 1-10 copies per cell, although some are present at much higher quantities (>200) (Kolev *et al.*, 2010). One of the advantages of genes arranged in poly-

cistronic arrays is that expression of stage specific genes can be altered rapidly (Roditi *et al.*, 1989, Archer *et al.*, 2011).

The genomes of several *T. brucei* strains have been sequenced, including *T.b. brucei* isolates, with TREU 927 being the gold standard reference (Berriman *et al.*, 2005). In addition, sequencing has been carried out on *T.b. gambiense* (DAL972) (Jackson *et al.*, 2010). However, the majority of protein coding regions are known, there is a lack of annotation, with almost 50% of genes annotated as “putative” or “unknown” (Salavati & Najafabadi, 2010, Shateri Najafabadi & Salavati, 2010). It is also thought there are non-coding RNAs in the *T. brucei* genome, but these have not been characterised experimentally (Zheng *et al.*, 2013).

### 1.6.2 Post-transcriptional regulation of gene expression

Due to the constitutive and poly-cistronic nature of *T. brucei* transcription, there is a lack of regulation of gene expression at the transcription level (Clayton, 2014). Instead, regulation occurs on the post-transcriptional level and can include trans-splicing, polyadenylation and mRNA degradation, amongst others (Clayton, 2014). In addition, it is thought that sequence elements in the 5' and 3'-untranslated regions (UTRs) play an important role in mRNA stability and degradation (Irmer & Clayton, 2001, Robles & Clayton, 2008). This hypothesis is supported by the fact that the trypanosome genome encodes a large number of RNA-binding proteins (RBPs) (Wurst *et al.*, 2009).

### 1.6.3 The trypanosome spliced leader

In 1982, researchers discovered that all trypanosome VSG mRNAs possessed an identical short nucleotide sequence on the 5' end (Boothroyd & Cross, 1982). This 39-nucleotide sequence was termed the spliced leader (SL) and further work showed it was present on all mature mRNAs (Parsons *et al.*, 1984). Whilst SL splicing has since been shown to be present in other eukaryotes such as nematodes, it has not been observed in higher eukaryotes such as vertebrates, including humans (Douris *et al.*, 2010, Gunzl, 2010). This makes the SL splicing pathway an attractive target for HAT therapy and much work has revolved around attempting to identify the enzymes and spliceosome factors that trypanosomes possess, and how they differ from their mammalian counterparts (Gunzl, 2010).

Published data indicates that trypanosomes possess all 5 conserved U spliceosomal snRNAs (U1, U2, U4, U5 and U6), which are involved in the generation of mature mRNAs (Gunzl, 2010). Humans possess up to 170 protein factors often associated with spliceosomes, but it is not known how many of these are present in *T. brucei*, mainly due to lack of annotation and significant sequence divergence (Gunzl, 2010).

The trypanosome SL sequence itself has been shown to possess a highly conserved and complex pattern of methylation on the first 4 bases (Perry *et al.*, 1987, Bangs *et al.*, 1992). This capping

process is crucial as it protects the mRNA from degradation, and it is also implicated in translation initiation (Zeiner *et al.*, 2003). The first base, an inverted guanosine bound to the rest of the SL via a 5'-5' triphosphate linkage, is termed cap0 (m<sup>7</sup>G), and is modified by TbCet1, an RNA triphosphatase, as well as a bifunctional guanylyltransferase-methyltransferase, TbCgm1 (Hall & Ho, 2006a, Ruan *et al.*, 2007a, Takagi *et al.*, 2007). Cap0 formation and methylation is essential (Takagi *et al.*, 2007). However, RNAi knock-down of TbCgm1 in PCF trypanosomes does not lead to an observed growth defect (Monnerat *et al.*, 2009). A further methyltransferase (MTase) called TbCmt1 has also been described, and is closely related to TbCgm1, although it does not possess guanylyltransferase capabilities (Hall & Ho, 2006a).

The cap1, cap2, cap3 and cap4 positions immediately after the cap0 m<sup>7</sup>G are also methylated and to date, three 2'-O-ribose MTases have been identified (Zamudio *et al.*, 2009). TbMtr1 methylates the cap1 position and is distantly related to MTase families also found in humans, although it acts specifically on *T. brucei* SL (Zamudio *et al.*, 2007, Mittra *et al.*, 2008). TbMtr2 methylates the cap2 nucleotide (Arhin *et al.*, 2006b). TbMtr3 methylates the third nucleotide and in addition, is thought to methylate the cap4 position, as TbMtr3 KOs lack 2'-O-ribose methylation on cap4 (Arhin *et al.*, 2006a, Zamudio *et al.*, 2009). However, this could be because the cap4 position is only methylated after methylation of cap3. Whilst enzymes acting on cap0 are essential, data indicates that the 2'-O-ribose MTases are not (Zamudio *et al.*, 2006, Zamudio *et al.*, 2009). For example, TbMtr2 dKOs are viable, albeit with a lack of cap2 modifications (Arhin *et al.*, 2006b).

One recent study discovered an mRNA decapping and recapping pathway that has important implications for the functional reactivation of mRNA, as well as mRNA stability and turnover (Ignatochkina *et al.*, 2015). Specifically, this study characterised a cytoplasmic capping enzyme called TbCe1. Ignatochkina and colleagues showed that TbCe1 possesses 5'-monophosphate RNA (pRNA) kinase activity, enabling it to recap uncapped SL RNA (Ignatochkina *et al.*, 2015). In the same study, the authors discovered an SL decapping enzyme called TbDcp2, which also localised to the cytoplasm and is involved in the 5'-3' exonuclease-mediated degradation of mature mRNA (Ignatochkina *et al.*, 2015).

African trypanosomes have been shown to possess an SL silencing (SLS) pathway that is related to the highly conserved unfolding protein response (UPR) pathway found in other eukaryotes (Goldshmidt *et al.*, 2010, Michaeli, 2012, Michaeli, 2015). In response to persistent ER stress, which can occur due to several acute intracellular changes leading to protein un- or mis-folding, an apoptotic cascade is activated that leads to programmed cell death through the inhibition of SL *trans*-splicing. This process leads to significant degradation of, and reduction in all RNAs (Goldshmidt *et al.*, 2010). Several other intracellular changes also occur during this process including increase in cytoplasmic Ca<sup>2+</sup>, and DNA fragmentation (Goldshmidt *et al.*, 2010).

## 1.7 Genetic diversity in the trypanosome

---

The generation and maintenance of genetic diversity is key to survival. In parasites like *T. brucei*, the genome is shaped by the requirement to constantly evolve in order to survive in the various hosts in which the parasite lives (Stephens *et al.*, 2012). Leigh Van Valen's Red Queen hypothesis (Van Valen, 1973), covered in detail elsewhere (Liow *et al.*, 2011), states that this constant evolution is necessary to overcome the host immune responses, which themselves evolve to destroy the parasite (Stephens *et al.*, 2012). The evolutionary mechanisms of human serum resistance described in section 1.2.4 are a good example of this theory. Our understanding of genetic diversity of *T. brucei* in the field is crucial in order to target the parasite (Tait *et al.*, 2011). For example, varying genetics could in theory explain the differing efficacies of drugs, eflornithine in particular, against *T.b. gambiense* and *T.b. rhodesiense* (Iten *et al.*, 1997). Furthermore, genetic diversity has been implicated in differing pathologies such as splenomegaly and hepatomegaly (Morrison *et al.*, 2009, Capewell *et al.*, 2015b).

Several years ago it was demonstrated that *T. brucei* undergoes sexual recombination during the epimastigote stage in the salivary glands of the tsetse fly, prior to the development of metacyclic parasites (Gibson & Stevens, 1999, Tait *et al.*, 2002, Tait *et al.*, 2007). In addition, genetic diversity in the field has been measured using genotyping markers (Tait *et al.*, 2011). Whilst *T. brucei* have been shown to exhibit high levels of genetic diversity, *T.b. rhodesiense* is mixed, depending on geographical location, and *T.b. gambiense* shows low levels of diversity. The latter, however, has been shown to possess distinct genetic traits in different geographical foci (Tait *et al.*, 2011).

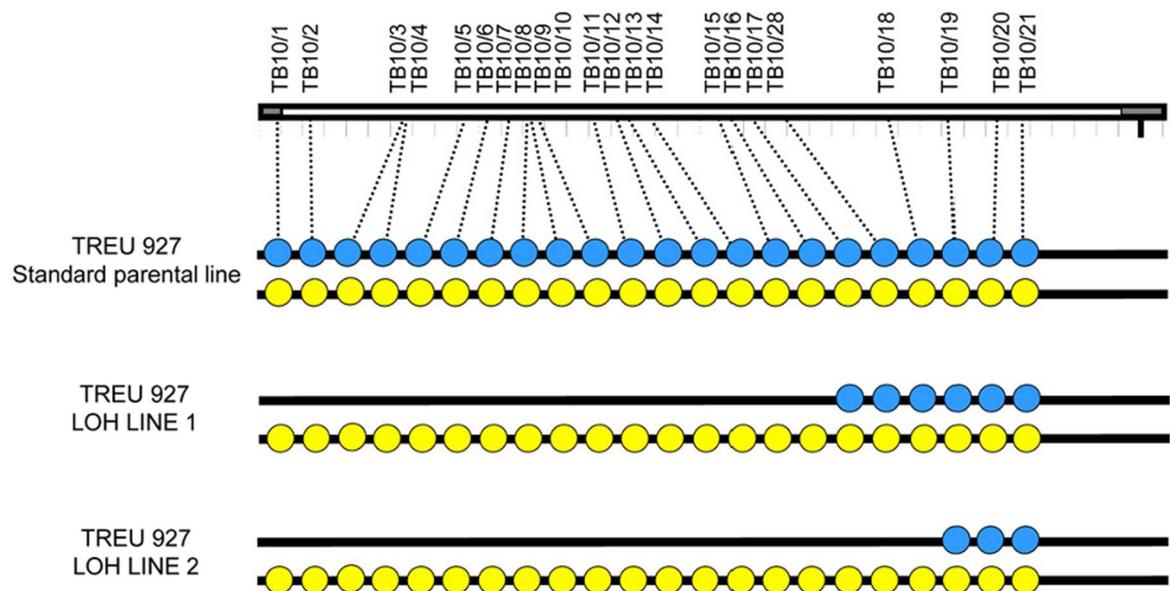
Whilst meiotic recombination does aid in generating diversity, there are mechanisms involving mitotic recombination that may also contribute to this phenomenon. Examples include, but are not limited to, multiple ploidy states, which has been reported in *C. albicans* and *C. neoformans* (Bennett *et al.*, 2014), and loss-of-heterozygosity (LOH).

### 1.7.1 Loss-of-heterozygosity: a potential key to genetic diversity

Loss of heterozygosity is a previously reported genetic phenomenon where heterozygous loci are rendered homozygous either due to loss of one allelic variant, or by mitotic recombination (Ryland *et al.*, 2015). Generation of genetic diversity can occur in heterozygous diploids as a result of homozygosis of short genomic regions, or large portions of a chromosome (Bennett *et al.*, 2014). This can be due to recombination, double chromosome cross-overs or break-induced replication (BIR) involving strand invasion (Llorente *et al.*, 2008). In the case of the latter, break events that occur in a chromosome can result in LOH from the break-point all the way to the telomeric regions of the chromosome (Bennett *et al.*, 2014). Double-strand breaks and their repair in *T. brucei* have been studied in depth, especially in the context of driving antigenic variation (Alsford *et al.*, 2009).

This phenomenon has been reported in several organisms, including *C. albicans* (Diogo *et al.*, 2009) and *S. cerevisiae* (Andersen *et al.*, 2008). One recent publication provided evidence to show that environmental perturbations that induce cellular stress can also lead to LOH (Forche *et al.*, 2011).

Interestingly, LOH has been observed in an *in vitro* setting in several independent cell lines including two TREU 927 *T.b. brucei* isolates, a melarsoprol-resistant STIB 247 *T.b. brucei* line and a STIB 386 *T.b. gambiense* line (A. Cooper, personal communication). In the three *T.b. brucei* isolates, a significant increase in growth rate was also observed (A. Cooper, personal communication). Genotyping experiments as well as genomics analyses of two TREU 927 LOH lines showed the occurrence of LOH on chromosome 10 (fig. 1-7). In the laboratory strains, up to 50% of the chromosome had mutated to a homozygous genotype, and the current hypothesis that an underlying genetic alteration is responsible for increased growth rates is currently under investigation by the MacLeod group.



**Figure 1-7: Map of PCR microsatellite analysis of chromosome 10 in a wild-type TREU 927 PCF line, and two LOH lines.** All sites selected were those that usually display heterozygosity in TREU 927, as indicated by blue and yellow circles. In both LOH lines, a large portion of the chromosome showed the presence of only one allele, thus indicating homozygosity in these loci. Figure taken from Cooper (2009).

LOH was recently shown to be one of the major factors influencing genetic diversity in the asexual *T.b. gambiense* type 1 (Weir *et al.*, 2016). Genomics analyses by Weir and colleagues showed that type 1 strains exhibit low levels of intra group diversity and geographically restricted sub-populations. Inspection of SNP distributions showed large areas of chromosomes to be devoid of SNPs, confirming that LOH played a role in the genomic evolution of these parasites (Weir *et al.*, 2016). Furthermore, the authors hypothesise that LOH is intimately linked to the Meselson Effect, whereby chromosome homologues in asexual organisms evolve and diverge independently (Weir *et al.*, 2016). Thus, whilst LOH removes heterozygosity, independent evolution re-introduces SNPs,

thereby driving genetic diversity. It is currently not known whether these events also lead to phenotypic advantages such as the increased growth rates seen in the laboratory strains.

## 1.8 Metabolism in *T. brucei*

---

### 1.8.1 General metabolism

As an ancient eukaryote, metabolism in *T. brucei* is biologically relevant not only due to its ability to cause disease, but also to those attempting to understand the evolution of metabolism. The parasite's metabolism has evolved to exploit the host for many nutrients, including amino acids, fatty acids, sterols, sugars and purines (Bringaud *et al.*, 2006, Smith & Butikofer, 2010). The key differences between the mammalian infective and insect forms have been studied for many years, and it is known that whilst BSF parasites carry out metabolism streamlined towards acquiring energy through glycolysis, PCF cells readily use various sugars and amino acids from their extracellular environment for energy production, with an emphasis on mitochondrial metabolism (Besteiro *et al.*, 2005, Bringaud *et al.*, 2006, Bringaud *et al.*, 2012).

Lipids constitute 8-11% of the *T. brucei* dry weight (Richmond *et al.*, 2010), and lipid metabolism in the trypanosomatids bears some key differences compared to more conventional eukaryotes (Reviewed by (Smith & Butikofer, 2010)). Historically, research on *T. brucei* lipids focused on glycosylphosphatidylinositol (GPI) biosynthesis, as these form the structures that anchor VSGs to the trypanosome surface (Smith & Butikofer, 2010). In particular, myristate formation and turnover was of high interest as a potential therapeutic target (Buxbaum *et al.*, 1996, Werbovets & Englund, 1997). However, focus has recently shifted to other areas of lipid metabolism that are unique to the parasite (Richmond *et al.*, 2010). *T. brucei* possesses all major phospholipid classes present in other eukaryotes, and whilst eukaryotic cells typically utilise type I or type II synthases to generate fatty acids, *T. brucei* uses a system of 4 elongases to generate fatty acids, including myristate in the BSF stage (Patnaik *et al.*, 1993, Smith & Butikofer, 2010).

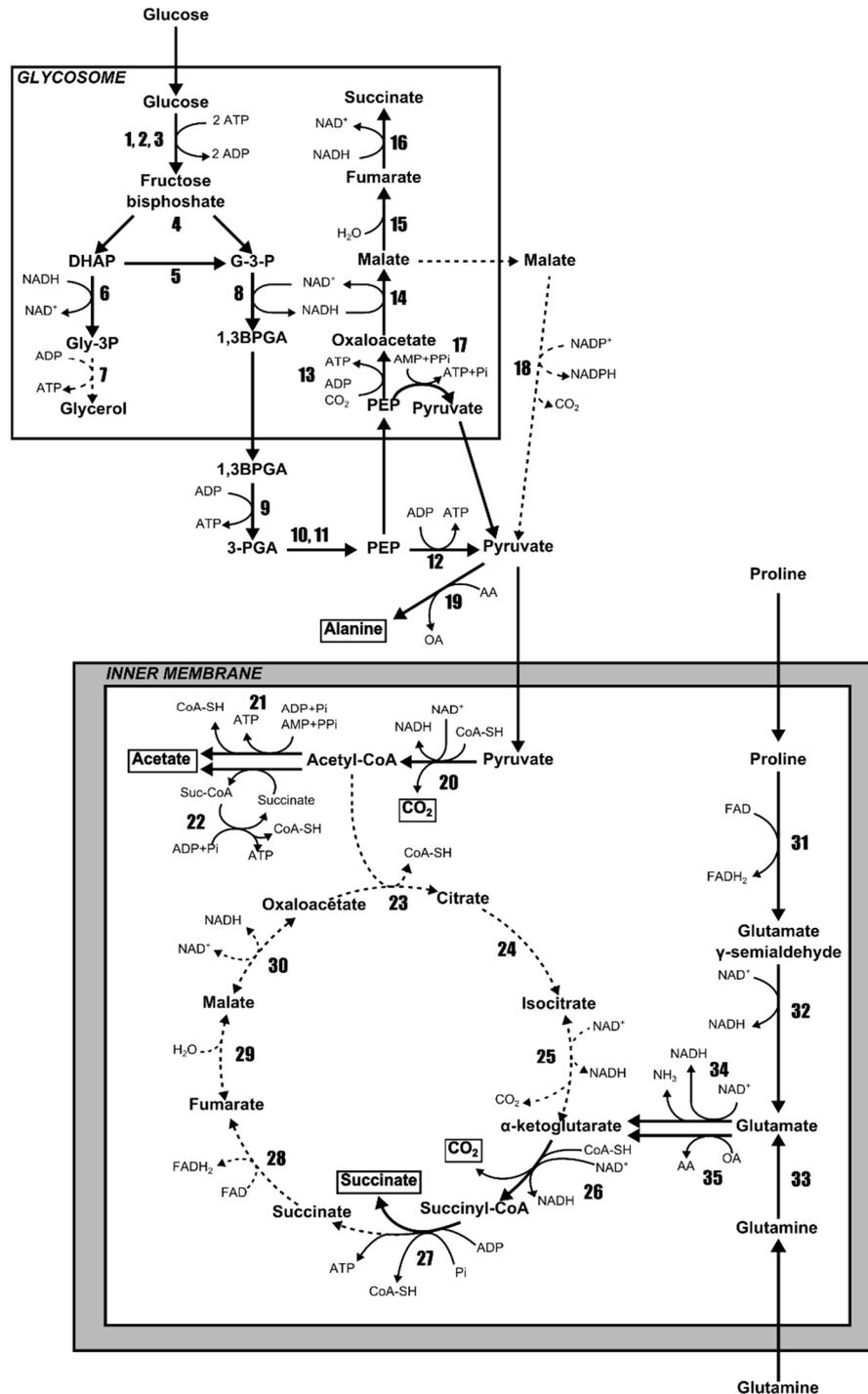
The glycerophospholipids phosphatidylcholine (PC) and phosphatidylethanolamine (PE) comprise 45-60% and 10-20% of the total *T. brucei* phospholipid content respectively (Richmond *et al.*, 2010). Choline and ethanolamine, which are taken up from the extracellular environment, form the head groups of these phospholipids and these are incorporated into the Kennedy pathway to generate PC and PE (Smith & Butikofer, 2010).

*T. brucei* was thought to possess a large complement of nutrient transporters, which was confirmed upon the completion of the sequencing of its genome (Berriman *et al.*, 2005), as well as that of the human-infective *T.b. gambiense* (Jackson *et al.*, 2010) and *T. congolense* and *T. vivax* (Jackson *et al.*, 2013). This reflects the parasite's dependency on its extracellular environment for many metabolic

factors such as amino acids and vitamins. Indeed, *T. brucei* has the largest number of transporters of the three TriTryps, probably due to the fact that unlike both *Leishmania spp.* and *T. cruzi*, it remains extracellular and thus, has greater access to these nutrients (Berriman *et al.*, 2005).

### **1.8.2 Energy metabolism: glycolysis and the incomplete citric acid cycle**

BSF trypanosomes rely primarily on glycolysis for ATP production (Michels *et al.*, 2006). As mentioned previously, several of the glycolytic steps proceed in a peroxisome-like organelle called the glycosome (Visser *et al.*, 1981). However, in this organelle, net production of ATP is zero, and ATP production through the action of pyruvate kinase using phosphoenolpyruvate (PEP) is the main generator of ATP-derived energy (Michels *et al.*, 2006). The majority of pyruvate is shuttled out of the cell as a waste product (fig. 1-8).



**Figure 1-8: Energy metabolism in *T. brucei*.** BSF trypanosomes rely on glucose for ATP production, carrying out enzymatic reactions 1-12. In contrast, PCF trypanosomes preferentially utilise proline to generate succinate, although they are known to utilise glucose through glycolysis in the presence of the sugar. Abbreviations: AA, amino acid; 1,3BPGA, 1,3-bisphosphoglycerate; CoA-SH, coenzyme A; DHAP, dihydroxyacetone phosphate; G-3-P, glyceraldehyde-3-phosphate; Gly-3P, glycerol 3-phosphate; OA, 2-oxoacid; PEP, phosphoenolpyruvate; 3-PGA, 3-phosphoglycerate; Pi, inorganic phosphate; PPi, inorganic pyrophosphate; Suc-CoA, succinyl-CoA. Enzymes: 1, hexokinase; 2, glucose 6-phosphate isomerase; 3, phosphofructokinase; 4, aldolase; 5, triose-phosphate isomerase; 6, glycerol 3-phosphate dehydrogenase; 7, glycerol kinase; 8, glyceraldehyde 3-phosphate dehydrogenase; 9, phosphoglycerate kinase; 10, phosphoglycerate mutase; 11, enolase; 12, pyruvate kinase; 13, phosphoenolpyruvate carboxykinase; 14, glycosomal malate dehydrogenase; 15, fumarase; 16, NADH-dependent fumarate reductase; 17, pyruvate phosphate dikinase; 18, malic enzyme; 19, alanine aminotransferase; 20, pyruvate dehydrogenase; 21, unknown; 22, acetate:succinate CoA transferase; 23, citrate synthase; 24, aconitase; 25, isocitrate dehydrogenase; 26, ketoglutarate dehydrogenase; 27, succinyl-CoA synthase; 28, NADH-dependent fumarate reductase; 29, fumarase; 30, malate dehydrogenase; 31, L-proline dehydrogenase; 32, pyrroline-5 carboxylate dehydrogenase; 33, L-glutamine deaminase; 34, glutamate dehydrogenase; 35, glutamate aminotransferase.

Whilst the paradigm of BSF *T. brucei* central carbon metabolism has historically been based around the uptake of glucose and excretion of the main glycolytic waste product, pyruvate, recent data has emerged that questions the accuracy of this theory (Mazet *et al.*, 2013). Mazet and colleagues demonstrated that inhibition of acetate production from pyruvate was lethal to the parasite, suggesting that pyruvate is not only excreted from the cell, but rather like PCF metabolism, is also used for mitochondrial metabolism in the form of succinate and acetate production (Mazet *et al.*, 2013). In addition, the mitochondrion is the location of the trypanosome alternative oxidase (TbAOX) that aids in the maintenance of redox balance in the glycosome through the action of a glycerol 3-phosphate, dihydroxyacetone shunt (Oppendoes *et al.*, 1977, Creek *et al.*, 2015).

Work was recently published by the Barrett group showing the distribution of heavy carbon isotopes ( $^{13}\text{C}$ ) derived from glucose, in the BSF parasite (Creek *et al.*, 2015). By culturing *T. brucei* in the presence of universally (U)- $^{13}\text{C}$ -labeled glucose as well as (U)- $^{12}\text{C}$ -glucose in a 50:50 ratio, and subsequently analysing the cell pellets by liquid chromatography mass spectrometry (LC-MS), Creek and colleagues observed glucose-derived carbon labelling in over 150 metabolites ranging from pentose phosphate pathway components to sugar phosphates and nucleotides (Creek *et al.*, 2015). In addition to 50% labelling seen in pyruvate, carbon isotopes were also incorporated into succinate, alanine and acetyl-CoA, thereby supporting the previous findings by Mazet and colleagues (Creek *et al.*, 2015).

The glycosome also harbours the pentose phosphate pathway (PPP) (Kerkhoven *et al.*, 2013). This pathway, although less active compared to glycolysis, plays an important role in producing NADPH, protecting the parasite from oxidative stress. The link between glycolysis and PPP metabolism in trypanosomatids was highlighted by a study which showed that *Leishmania mexicana* cells lacking glucose transporters were more sensitive to oxidative stress (Rodriguez-Contreras *et al.*, 2007).

In PCF trypanosomes, energy metabolism occurs primarily in the mitochondrion, which houses a semi-complete citric acid cycle (tricarboxylic acid [TCA] cycle) (Besteiro *et al.*, 2005, van Hellemond *et al.*, 2005) (fig. 1-8). In addition, pyruvate is further metabolised by the pyruvate dehydrogenase complex into acetyl-CoA, before being converted to acetate. PCFs also utilise threonine, which is taken up from the extracellular environment, to produce acetate (Millerioux *et al.*, 2013).

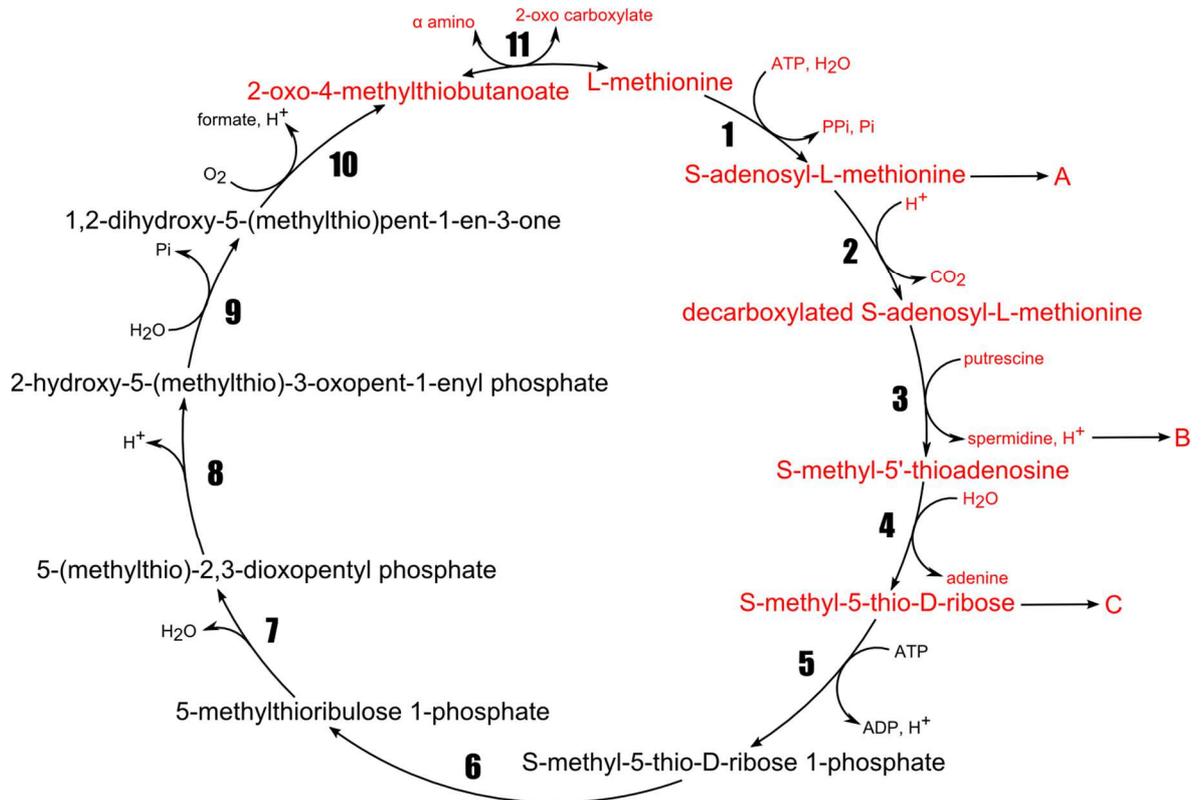
In addition to TCA metabolism, PCFs are also able to switch to glucose catabolism for energy generation, akin to their mammalian infective counterparts (Bringaud *et al.*, 2006). In laboratory culture, PCFs are routinely cultured in high levels (>5 mM) of glucose, and proline metabolism is down-regulated under these conditions (Lamour *et al.*, 2005). Therefore, by regulating the proline and glucose levels, different areas of central carbon metabolism are activated in this life cycle stage (Lamour *et al.*, 2005). Importantly, tsetse fly haemolymph contains negligible amounts of glucose,

although this depends on the fly taking bloodmeals, and parasites using proline as the main carbon source are thought to more accurately reflect PCF metabolism in the field (Besteiro *et al.*, 2005)

### 1.8.3 L- methionine and the Yang cycle

L-methionine is a highly-important sulfur-containing amino acid involved in many key processes in *T. brucei*, akin to all eukaryotes (Hasne & Barrett, 2000). It is the first residue in all proteins and in plants and prokaryotes and also forms an integral part of the Yang cycle (fig. 1-9). This cycle involves the degradation of L-methionine through S-adenosyl-L-methionine (AdoMet), S-methyl-5'-thioadenosine (5'-methylthioadenosine, 5'-MTA) and S-methyl-5-thio-D-ribose (5'-MTR) (Adams & Yang, 1977, Yung *et al.*, 1982). In plants and bacteria, this cycle proceeds with further degradation of 5'-MTR to 2-oxo-4-methylthiobutanoate through 6 steps (fig. 1-9), before L-methionine is regenerated (Yung *et al.*, 1982). Thus, in these organisms, the cycle is maintained through the continuous biosynthesis and degradation of L-methionine. L-methionine degradation is essential in trypanosomatids, as has been shown through the use of 5'-alkyl-substituted analogues of 5'-MTA (Bacchi *et al.*, 1991).

In contrast to other eukaryotes, evidence suggests that *T. brucei* lacks the ability to synthesise L-methionine *de novo*, and instead, this amino acid is taken up from its extracellular surroundings (F. Achcar, personal communication). Ongoing work using (U)-<sup>13</sup>C-L-methionine has shown that the Yang cycle is incomplete in *T. brucei*, and 5'-MTR can be observed in spent medium samples, suggesting that L-methionine degradation only proceeds up to this point (F. Achcar & K. Johnston, personal communication) (fig. 1-9). Whilst there were previous suggestions that the full methionine recycling pathway exists in *T. brucei*, the stable-isotope labelling observations do not support this hypothesis (F. Achcar, personal communication).



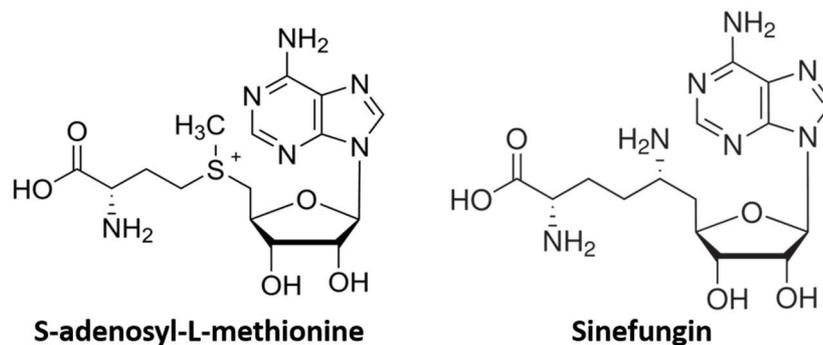
**Figure 1-9: The Yang cycle.** Components indicated in red are those that are known to be present in *T. brucei*. Enzymes: 1, S-adenosyl-L-methionine synthase; 2, S-adenosyl-L-methionine decarboxylase; 3, spermidine synthase; 4, S-methyl-5'-thioadenosine nucleosidase; 5, S-methyl-5-thio-D-ribose kinase; 6, S-methyl-5-thio-D-ribose 1-phosphate isomerase; 7, 5-methylthioribulose 1-phosphate dehydratase; 8, 5-(methylthio)-2,3-dioxopentyl phosphate enolase; 9, 2-hydroxy-5-(methylthio)-3-oxopent-1-enyl phosphate phosphatase; 10, acireductone dioxygenase; 11, 2-oxo-4-methylthiobutanoate aminotransferase. Letters correspond to the following pathways: A. Methyltransferase reaction, B. Polyamine synthesis pathway and trypanothione biosynthesis, C. Excretion of S-methyl-5-thio-D-ribose (F. Achcar, personal communication).

There are a number of important metabolic pathways that branch from the L-methionine degradation pathway where its components play an integral role. Firstly, as described in [section 1.8.6](#), decarboxylated AdoMet (dcAdoMet) is fed into the polyamine pathway, where it is a substrate for spermidine synthase (Bitonti *et al.*, 1984, Willert & Phillips, 2012). Secondly, AdoMet itself is a crucial substrate involved in methyltransferase (MTase) reactions, where it acts as a methyl group donor. MTase reactions are found in all living organisms, yet whilst the process of methyl transfer is conserved, enzymes regulating these reactions are remarkably divergent (Schubert *et al.*, 2003).

#### 1.8.4 The role of S-adenosyl-L-methionine in global metabolism

As mentioned, AdoMet is a crucial substrate in MTase reactions, where it provides a methyl group. Methylation is an essential process involved in many areas of biology such as gene expression, signal transduction, protein sorting, protein activation and nucleic acid processing (Fontecave *et al.*, 2004). In more complex eukaryotic systems, methylation also plays an important role in regulation of hormones and neurotransmitters (Hunter *et al.*, 2009, Nugent *et al.*, 2015). Characterised examples of MTases in *T. brucei* include protein arginine MTases as well as histone MTases in the

context of gene regulation (Janzen *et al.*, 2006, Fisk *et al.*, 2010, Frederiks *et al.*, 2010, Lott *et al.*, 2014). Substrates of methylation are varied and include nucleic acids and proteins. Many of these processes are essential, and this has been shown through the use of sinefungin, an analogue of AdoMet containing an amino methylene group in place of the methylated sulfonium group that usually acts as the methyl donor (Devkota *et al.*, 2014). Most eukaryotes are highly sensitive to this compound, reflecting the deep conservation of methyl transfer reactions (Borchardt *et al.*, 1979, Ferrante *et al.*, 1984, Ferrante *et al.*, 1988, Chrebet *et al.*, 2005, Yadav *et al.*, 2014).



**Figure 1-10: Comparison of structures of AdoMet and sinefungin.** In AdoMet, the methyl group that is donated to the substrate in MTase reactions, forms parts of the methylated sulfonium group in the AdoMet structure. In contrast, sinefungin possesses an amino methylene group in this position and thereby competitively inhibits AdoMet-dependent MTases.

Whilst the key role of AdoMet appears to be its use in MTase reactions, it is important in many other metabolic processes. AdoMet is utilised as a source of methylene groups (Grogan & Cronan, 1997), ribosyl groups (Van Lanen *et al.*, 2003), and 5'-deoxyadenosyl radicals (Jarrett, 2003). As mentioned previously, AdoMet is also an important source of aminopropyl groups in the polyamine synthesis pathway (Willert & Phillips, 2012). Indeed, AdoMet is one of the most commonly utilised substrates in eukaryotic cells, second only to ATP (Fontecave *et al.*, 2004).

### 1.8.5 Methyltransferases

Methyltransferases (MTases) catalyse methyl transfer reactions. Whilst they all catalyse similar reactions, MTases are structurally divergent and, amino acid sequence identity can be as low as 10% (Schubert *et al.*, 2003). The majority of MTases are AdoMet-dependent, although other methyl donors including methyl tetrahydrofolate, methylamines, methanethiol and chloromethane have been described (Coulter *et al.*, 1993, Ragsdale, 2008).

Several years ago, researchers classified MTases based on structure (Schubert *et al.*, 2003). AdoMet-dependent-MTases were subsequently classified further based on substrate specificity and the atom targeted for methylation (Martin & McMillan, 2002). Protein methylation is the second most common posttranslational modification in eukaryotic biology, after phosphorylation (Fontecave *et al.*, 2004). In addition, methylation of nucleic acids, tRNA in particular, are essential

modification processes in structure maintenance and stabilisation (Swinehart & Jackman, 2015). Given their structural divergence and importance, MTases are important therapeutic targets, as exemplified in the field of oncology (reviewed by (Subramaniam *et al.*, 2014, Hamamoto & Nakamura, 2016)).

To date, two genome-wide MTase analyses have been carried out, one in humans (Petrossian & Clarke, 2011), and the other in *Saccharomyces cerevisiae* (Wlodarski *et al.*, 2011). Both studies highlighted the diversity of MTases. In *T. brucei*, MTases described in the literature are concentrated around several specific processes. Firstly, histone modifications, where both DOT homologues have been characterised (Janzen *et al.*, 2006, Figueiredo *et al.*, 2008, Frederiks *et al.*, 2010). MTases involved in SL methylation have also been described in detail, although the *T. brucei* genome contains further uncharacterised capping enzymes, the annotations of which can be visualised on the TriTrypDB project website (Aslett *et al.*, 2010). Another area of MTase biology that has been described are the protein arginine MTases, of which there are several that exhibit homology to their mammalian counterparts (Pelletier *et al.*, 2005, Goulah *et al.*, 2006, Pasternack *et al.*, 2007, Fisk *et al.*, 2009, Fisk *et al.*, 2010, Lott *et al.*, 2014, Wang *et al.*, 2014a). However, there are likely to be many more MTases encoded in the *T. brucei* genome, which can be found through bioinformatical means, in order to collate a “methyltransferome”, similar to the work done by the Mottram group on the *T. brucei* “kinome” (Parsons *et al.*, 2005, Jones *et al.*, 2014a).

### 1.8.6 Polyamine pathway and oxidative stress

Polyamines are positively charged amino hydrocarbons essential to most eukaryotes. Whilst most eukaryotic species synthesize putrescine, spermidine and spermine, all components of the polyamine synthesis pathway, trypanosomatids do not synthesize the latter (Willert & Phillips, 2012). Putrescine is synthesised by the decarboxylation of ornithine, catalysed by ODC (Vincent *et al.*, 2012). Putrescine is then converted to spermidine by the addition of an aminopropyl group transferred from decarboxylated AdoMet (dcAdoMet) (Bitonti *et al.*, 1984). This reaction is catalysed by spermidine synthase (fig. 1-9) (Willert & Phillips, 2012).

Trypanosomatids have evolved an unusual end product of the polyamine pathway, where spermidine fuses with two glutathione molecules to generate trypanothione ( $N^1, N^8$ -bis(glutathionyl)spermidine), a redox factor that is essential in maintaining redox balance and is used in oxidative stress responses (Fairlamb *et al.*, 1985). The first step is carried out by glutathionylspermidine synthetase (GSS) and the second reaction that conjugates a second glutathione molecule to the structure is catalysed by trypanothione synthetase (TrypSyn) (Henderson *et al.*, 1990).

Almost every enzyme involved in the synthesis of trypanothione has been characterised in the context of gene deletion studies (Huynh *et al.*, 2003, Xiao *et al.*, 2009, Pratt *et al.*, 2014). In addition, inhibitors of AdoMet decarboxylase (AdoMetDC) have proven potent against the parasite in both *in vitro* and *in vivo* experiments (Barker *et al.*, 2009, Hirth *et al.*, 2009). AdoMetDC in particular has attracted interest as a drug target, because it possesses key structural differences compared to mammalian AdoMetDC (Willert *et al.*, 2007). Upon the publication of the *T. brucei* genome, an AdoMetDC-like gene, now called prozyme, was discovered (El-Sayed *et al.*, 2005b). This protein was shown to form a heterodimer with AdoMetDC (Willert *et al.*, 2007). This type of AdoMetDC protein has so far only been identified in *T. brucei* and *T. cruzi*, making it an intriguing therapeutic target, as it catalyses the production of a precursor involved in the essential polyamine synthesis pathway (Willert & Phillips, 2012).

Regulation at the mRNA and translation level is crucial for the function of the polyamine synthesis pathway (Clayton, 2014). AdoMetDC and ODC in particular have been shown to exhibit long half-lives, when compared to the eukaryotic counterparts (Phillips *et al.*, 1987, Willert & Phillips, 2008). It is thought that this plays a role in the selective nature of eflornithine action *in vivo*. The half-life of ODC in *T.b. rhodesiense* is much shorter compared to that in *T.b. gambiense*, and this could explain why eflornithine is less potent against the former (Iten *et al.*, 1997).

### 1.8.7 Metabolism during differentiation

Differentiation from BSF to SSF trypanosomes is initiated by an external signal termed SIF (Vassella *et al.*, 1997). Response to this factor generates several intracellular signalling cascades that bring around a significant remodelling in morphology as well as intracellular metabolism, involving several kinases and the cAMP pathway (Jones *et al.*, 2014a, Tagoe *et al.*, 2015).

Glycolysis is a key process in BSF parasites and interestingly, silencing of the glycosomal enzymes of glycolysis leads to changes in surface protein expression, which suggests that the downregulation of glycolysis and the differential expression of stage-dependent surface proteins are intrinsically linked (Morris *et al.*, 2002). Glycosomal turnover is also altered in differentiating trypanosomes, ultimately leading to a reduction in number of these organelles (Herman *et al.*, 2008, Bauer *et al.*, 2013). GPEET procyclin, the major procyclic surface molecule, exhibits differential expression depending on mitochondrial activity and is directly related to expression of key mitochondrial enzymes such as the acetate, succinate CoA transferase/succinyl-CoA synthetase (ASCT) (Vassella *et al.*, 2004).

The Barrett group has carried out metabolomics analyses of a *pleomorphic T. brucei* line, to investigate metabolic changes in differentiating cells (J. Anderson *et al.*, unpublished). As expected, a general downward trend was seen in metabolites involved in BSF-specific metabolism, including

in phenylpyruvate, a pyruvate derivative resulting from a phenylalanine transaminase reaction, which is commonly found in the serum of trypanosome-infected hosts (Li *et al.*, 2011). Sphingolipid metabolism also exhibited a decreasing trend. In contrast, several metabolic components of PCF-specific metabolism, such as mitochondrial metabolites, were found to increase during differentiation (J. Anderson *et al.*, unpublished). There were also unexpected, and so-far unexplained, changes observed in differentiating cells. For example, L-cystathionine was found to increase, as well as 5'-MTA, several modified lysines such as acetyl-L-lysine and trimethyl-L-lysine, sn-glycero-3-phosphocholine and other phospholipids (J. Anderson *et al.*, unpublished).

## 1.9 Omics technologies

---

Investigations into *T. brucei* metabolism have benefited from the use of 'omics' technologies. Modernisation of computer science and experimental techniques have been exploited by many areas of science, especially with regards to the deconvolution of large, memory-intensive datasets. This has enabled large-scale studies of entire cells, organisms and systems (Berger *et al.*, 2013). Through the analysis of biological processes in their entirety, a much more accurate and detailed picture of the process or system in question can be obtained, at a higher resolution than previously possible. The term 'omics', is often used to refer to these types of datasets, and the four classical omics technologies are summarised below.

### 1.9.1 Genomics, transcriptomics and proteomics

Analysis of full genomes, is now an integral part of biological research. Since the completion of the Human Genome Project, thousands of genomes from a wide range of species have been sequenced, improving our understanding of organisms including prokaryotes, eukaryotes and archaea has been gained. One of the key parts of genomics analyses is the accurate and detailed annotations of genomes, including protein coding genes, regulatory regions, regulator binding sites, UTRs, introns and exons and non-coding RNA regions.

The abundance of mRNA transcripts can provide important insights into differential expression, and is often, but not always, correlated with the levels of protein expression (Zhang *et al.*, 2010). Given that the transcriptome can vary dynamically in response to external stimuli, it is often a better indicator of intracellular dynamics, compared to the relatively static genome (Zhang *et al.*, 2010).

Historically, transcriptomics analyses were based on understanding differential expression of mRNA transcripts, using techniques such as microarrays (Kagnoff & Eckmann, 2001). However, transcriptomics data is frequently carried out through sequencing technologies, allowing the analysis of both transcript abundance and sequence (Ozsolak *et al.*, 2009). Importantly, RNA-seq thereby provides a very detailed overview of splicing variants in organisms that possess introns as

well as exons (Trapnell *et al.*, 2013). In the case of *T. brucei*, this is less important as they only possess two known introns (Mair *et al.*, 2000).

The entire complement of proteins in a cell or system is referred to as the proteome. As mentioned above, the transcriptome does not always reflect the proteome. For example, *T. brucei* RNA undergoes significant posttranscriptional regulation. Therefore, not all mRNA is translated to protein (Clayton, 2014). The main strategy for proteomics investigations is the use of 2-dimensional PAGE (2D-PAGE) or 2-dimensional difference gel electrophoresis (2D-DIGE) (Timms & Cramer, 2008). These experiments separate proteins according to the isoelectric point as well as mass (Timms & Cramer, 2008). Protein identification is generally done by mass spectrometry (Ummanni *et al.*, 2011). Whilst the principle of studying proteomics shows significant potential, there are still some disadvantages, the main one being that only a small portion of the proteome is usually identified (Mirza *et al.*, 2007). In addition, insoluble proteins are often unwittingly removed during the generation of a protein lysate (Mirza *et al.*, 2007). Recently, proteomics analyses have been aided by several technological developments that allow for quantification of relative protein abundances in multiple samples: Examples are stable isotope labelling-based isotope coded affinity tags (ICAT) and isobaric tag for relative and absolute quantification (iTRAQ) (Haqqani *et al.*, 2008, Yan *et al.*, 2008).

### 1.9.2 Metabolomics

The field of metabolomics aims to investigate the complex interplay of metabolic pathways occurring in a biological system, and to understand how external environments and pressures such as drug treatment can impact this system (Creek & Barrett, 2014). The metabolome is defined as all small molecules (<1,500 Da) in a cell or system (Dunn *et al.*, 2013). This can include a large variety of molecules including, but not limited to, peptides, sugars, oligonucleotides, organic acids, ketones, amines, amino acids, drugs, xenobiotics, lipids and steroids.

Historically, the concept of metabolomic studies dates back as far as 2,000 B.C., where evidence suggests that Chinese doctors used patient urine to diagnose diabetes (Eknoyan & Nagy, 2005). In the 17<sup>th</sup> century, Santorio Santorio, regarded by many as the founding father of metabolism research, published a study on “*insensible perspiration of the body*” (Eknoyan, 1999). However, many early studies were based on investigating individual components of metabolism. It was not until the early 20<sup>th</sup> century, after the assembly of the first mass spectrometer at the University of Cambridge, that multiple metabolites could be studied. The first mention of the word “metabolome” was a publication by Oliver and colleagues in 1998 (Oliver *et al.*, 1998). Since then, rapid modernisation of the technologies required to study the metabolome have allowed researchers to study them in their entirety, culminating in the publication of the first draft of the human metabolome in 2007 (Wishart *et al.*, 2007).

Currently, several methods exist to study the metabolome, primarily involving either mass spectrometry (MS) or nuclear magnetic resonance (NMR) spectroscopy (Bringaud *et al.*, 2015). The most common NMR platform analyses  $^1\text{H}$  nuclei within sampled metabolites (termed  $^1\text{H}$ -NMR) by detecting the spin properties of the nuclei (Creek *et al.*, 2012a). Whilst reproducibility of experimental data, as well as easy of sample preparation without sample loss, are advantages to this technique, it offers poorer resolution compared to mass spectrometry.

Mass spectrometry involves mass measurements through the ionisation of metabolites and subsequent detection of the molecule mass-charge ratios ( $m/z$ ) (Creek *et al.*, 2012a). Generally, mass spectrometers consist of an ion source, a mass analyser and a detector. More modern MS systems incorporate an orbitrap analyser to increase resolution, with mass accuracies less than 1 ppm (Kamleh *et al.*, 2008). To further resolve complex mixtures of cellular metabolites, MS systems are generally coupled to liquid- or gas-chromatography (LC & GC respectively) columns which further separate metabolites and aid in the reduction of ion suppression. Peaks of similar  $m/z$  ratios can then be separated by retention time (Misra & van der Hooft, 2016).

The methodology used for metabolomics studies often depends on the biological questions and system under investigation. Neither platform offers 100% coverage (Creek *et al.*, 2012a). There are two commonly used approaches to metabolomics experiments. An untargeted approach is hypothesis generating, often comparing a biological system under two different conditions, such as drug- and vehicle-treated (Creek & Barrett, 2014). In contrast, a targeted approach is aimed at investigating a particular pathway, often with the use of stable isotope (e.g.  $^{13}\text{C}$ ,  $^{15}\text{N}$ ) labelling of particular metabolites in order to investigate pathway interactions and isotope distribution throughout a system (Chokkathukalam *et al.*, 2014). Important considerations must also be made with regards to data analysis, for which many different pipelines exist (Misra & van der Hooft, 2016).

### 1.9.3 Other 'omics

Whilst the four main 'omics' described above are considered to be the classical omics fields, the last few years have seen a rapid increase in the use of '-omics'-related terms in the literature. Some estimates suggest there are in excess of 100 individual omics fields described (Baker, 2013).

Whether these new -omes will be taken seriously over time remains to be seen. There lies some sense in coining '-omics' terms for subsets of proteins, or as in another example, describing all known genome modifications (i.e. epigenomics). However, several novel 'omics' describe very specific fields of biology, and whether these are valid, stand alone omics areas is debateable. The rapid expansion of 'omics' research has prompted one evolutionary biologist to say: "it [omics] is a language parasite" (Hotz, 2012). To that extent, Eisen also outlined a theory of "badomics", even going so far as to develop approaches to distinguish "badomes" from "goodomes" (Eisen, 2012). In

this publication, Eisen argues that, in order for newer omics fields to be recognised and taken seriously, they must be a valid effort to further an area of science hitherto scarcely studied, rather than revisiting old hypotheses (Eisen, 2012).

The term ‘Omics’ itself could in theory be applied to any sort of “totality”. As Prof. Barrett mentions in an opinion piece for the *New Statesman*, the total scientific literature could be called a “bibliome”, and continuing in a similar vein, a collection of portraits, such as those at the National Gallery, is a “National Portraitome” (Barrett, 2015). Expanding further still, if all opinions on a particular topic were to be taken into account, would that be an “opiniome” (Watson, 2015)?

Whilst these interpretations highlight the absurdity of some recently emerging ‘omics’ topics, it is also important to realise that perhaps this term reflects the universal aim of modern science, to generate large amounts of finely detailed information on a whole cell or organism, and more importantly, to integrate these datasets in order to further our understanding of the system in question. Whether this ultimate aim is achieved could define the success or failure of the ‘omics’ era of biology. In the context of NTDs, systems biology and omics analyses have concentrated on genome, transcriptome, proteome and metabolome, with the odd mention of ‘kinome’ (Parsons *et al.*, 2005, Jones *et al.*, 2014a). Importantly, all have been utilised with great success, especially in the area of drug discovery.

## 1.10 Techniques in drug target identification and validation

---

As mentioned in [section 1.4](#), drug target deconvolution is crucial for the development of novel therapeutics. Discovery of novel therapeutic compounds is typically carried out in several different, but complementary, approaches. Structure-based drug design involves the validation of an essential protein drug target in a pathogen, and the subsequent development of small molecules aimed at that target (Khalaf *et al.*, 2014). This process is cyclical, involving many rounds of optimisation of the small molecule structure and observations of its interactions with the protein target, often through the use of 3-dimensional X-ray structures (Deschamps, 2005). This approach is greatly aided by bioinformatics analyses that can predict drug:protein interactions, as well as genome sequencing to identify novel therapeutic targets (Reguera *et al.*, 2014). The latter has been achieved for several trypanosomatids (Berriman *et al.*, 2005, El-Sayed *et al.*, 2005a, Ivens *et al.*, 2005, Jackson *et al.*, 2010), meaning that target-based approaches are now feasible for many NTDs (Reguera *et al.*, 2014). The second approach is termed phenotypic screening. This involves making observations of phenotypic effects that are induced by compounds, and further optimising compounds showing significant selectivity for further structure optimisation. Gene modification in the pathogen can be used to validate essential targets. Phenotypic screening is more commonly applied in an *in vitro* setting, as *in vivo* screens do not offer high throughput (Reguera *et al.*, 2014).

Both the aforementioned methods carry advantages and disadvantages. For example, whilst target-based approaches allow the refinement of the interaction between drug and target, many of the compounds often present low selectivity against the pathogen (Reguera *et al.*, 2014). In addition, a target-based approach does not take into account the cell membrane the compound must cross to get to the pathogen target. Similarly, neither approach can predict further biological barriers, such as the blood-brain barrier in late-stage HAT. One of the major disadvantages of phenotypic screens is that they do not directly identify the MoA of lead compounds (Lee & Bogyo, 2013).

To this extent, many techniques are now available to aid in drug target identification and many more are currently under development, providing an ever-improving field with many complementary techniques to aid in this challenging, but exciting field of biology. The main techniques currently available are summarised here.

### 1.10.1 Affinity chromatography

Affinity chromatography remains one of the most widely used techniques for the purposes of drug target identification (Lomenick *et al.*, 2011b). The technique involves the immobilisation of the active compound to an affinity tag, often through the action of biotin attachment to streptavidin, and is carried out in a similar fashion to immunoprecipitation assays (Annis *et al.*, 2007). Subsequent to drug immobilisation, a protein isolate pull down is carried out using a lysate from the pathogen required. This is followed by several wash steps to remove unbound protein and an elution step using excess drug or denaturing agents. The development of mass-spectrometry-based chromatography assays has greatly aided in the identification process, which is normally coupled to a protein gel assay (Slon-Usakiewicz *et al.*, 2005, Annis *et al.*, 2007).

There are many examples where affinity chromatography has been utilised successfully for drug target deconvolution. One *T. brucei* study used this technique to identify an interaction between the trypanocidal cymelarsan and glycerol 3-phosphate dehydrogenase (Denise *et al.*, 1999). A more recent Anacor study on a benzoxaborole, SCYX-5070, described an affinity chromatography approach using drug immobilised on sepharose beads to reveal interactions between the compound and mitogen-activated protein kinases (MAPKs) as well as cdc2-related kinases (CRKs) (Mercer *et al.*, 2011).

The main, and perhaps obvious, disadvantage of affinity chromatography, given that the drug is immobilised prior to the ligand binding step, is that a certain degree of SAR knowledge is required. In cases where the wrong functional group is used to bind an affinity tag, the chances of false negatives as well as false positives are significantly increased due to either non-specific protein binding, or lack of ligand-protein binding due to steric hindrance caused by the immobilisation.

Therefore, this method is less effective in drug discovery approaches in cases where no SAR knowledge exists (Lomenick *et al.*, 2011a).

### 1.10.2 Generation of resistant cell lines

Another commonly used method for drug target identification is the generation of resistance in a lab-adapted pathogen, followed by comparative analyses of the wild-type and resistant lines. Resistance can be generated in one of two ways: the first (and more common) method involves incubation of the pathogen *in vitro* with very low dose of the drug or compound. Small increments in drug concentrations over time, mimic selection pressure, and resistance is generated in the cell line through alterations in the genome or protein expression (Nzila & Mwai, 2010). The second method involves induced mutagenesis through the addition of powerful chemical mutagens to *in vitro* cultures. The live cells recovered from this treatment are then subject to the drug or compound, and resistant mutants are recovered for downstream analysis (Vincent *et al.*, 2010). Several other techniques including RIT-seq and CRISPR (both explained below) have refined this approach.

Mechanisms of resistance in the trypanosomatids (*Leishmania spp*, *T. cruzi* and the African trypanosomes) can involve reduced drug uptake, increased drug efflux, reduced metabolism in the drug activation process, biochemical inactivation of drug, target mutation, target overexpression, metabolic compensatory pathways and increased uptake of downstream metabolites, amongst others (Horn & Duraisingh, 2014). The advent of omics techniques discussed previously, have greatly enhanced the study of resistant cell lines by allowing a detailed overview of both wild-type and resistant cell lines on multiple levels (Canuto *et al.*, 2014, Wyllie *et al.*, 2015). To exhibit resistant phenotypes, pathogens often undergo point mutations, gene deletions, or significant changes in expression of the target protein (Graf *et al.*, 2013, Ranade *et al.*, 2013). In cases where these mutations affect protein:drug binding, subsequent analyses often incorporate structural biology approaches to determine the underlying mechanisms or resistance, as was done by Rock and colleagues to elucidate the MoA of Tavaborole, an anti-fungal benzoxaborole (Rock *et al.*, 2007). Whilst the main findings of these studies are often more relevant to resistance mechanisms, they can lead to the identification of target proteins that are otherwise unknown (Rock *et al.*, 2007).

One disadvantage of this method of drug target identification is that the mechanism of resistance can mask the genuine mode of action of the drug. For example, loss of expression of a specific transporter or subsets thereof is often sufficient to generate drug resistance in trypanosomes *in vitro*. Indeed, there are several examples of this phenomenon involving an amino acid transporter (TbAAT6) in eflornithine resistance (Vincent *et al.*, 2010), loss of the P2 (TbAT1) transporter in diminazene resistance (de Koning *et al.*, 2004, Teka *et al.*, 2011), and both the TbAT1 and the

TbAQP2 transporter in pentamidine resistance (Baker *et al.*, 2012, Graf *et al.*, 2013, Munday *et al.*, 2014).

### 1.10.3 RNAi-target sequencing (RIT-seq)

RIT-seq was developed specifically as a gene knock-down tool for *T. brucei* (Alsford *et al.*, 2011, Alsford *et al.*, 2012). It is based on genome wide knock-down of functional genes using RNA interference (RNAi). A genome-wide knock-down is easier to achieve as RNAi targeting does not require full coding sequences. Instead, an RNAi plasmid library was generated, containing randomly sheared genomic fragments (Alsford *et al.*, 2011).

This assay has been used successfully to unravel pathways involved in drug action for several drugs with previously unknown mechanisms (Alsford *et al.*, 2012). Importantly, with a genome-wide approach, multiple genes involved in the drug mode of action can be elucidated. However, one caveat of this approach is that it exclusively identifies non-essential genes (Alsford *et al.*, 2011). This is due to the fact that knock-down of essential genes will kill the parasite regardless of drug treatment, whereas non-essential gene knock-downs that confer drug resistance do survive (Alsford *et al.*, 2012). A more desirable approach based on a genome-wide study would be to create an overexpression library, as overexpression of target proteins can confer resistance in the pathogen (Begolo *et al.*, 2014). Surviving parasites could then be analysed regardless of the essentiality of the overexpressed gene. This approach is challenging, however, as plasmids containing full open reading frames for each of the >8,000 genes must be generated, compared to the RIT-seq approach requiring only gene fragments.

Another disadvantage associated with this approach is that more than 50% of genes found to be essential by knock-down experiment, carry an unknown function (Shateri Najafabadi & Salavati, 2010). Therefore, in many cases, candidates have to be selected and characterised before their role can be elucidated.

### 1.10.4 Metabolomics

As discussed above, metabolomics provides perhaps the most dynamic and accurate observation of a cellular system compared to other omics technologies. Metabolites change in abundance in response to many external as well as internal stimuli (Creek & Barrett, 2014). Metabolomics is ideally suited to investigating the mode of action of novel drugs by allowing the examination of these changes. Indeed, in the study of tropical diseases, the MOAs of several drugs have been elucidated through the application of this tool (table 1-1). It should be noted, however, that the application of metabolomics is well suited to a great variety of organisms for the study of drug modes of action (Reviewed in (Keun & Athersuch, 2007, Rabinowitz *et al.*, 2011, Russell *et al.*, 2013)

Organism	Compound/drug	Target identified	Reference
<i>T. brucei</i>	Eflornithine	Ornithine decarboxylase	(Vincent <i>et al.</i> , 2012)
<i>T. brucei</i>	Nifurtimox	Involves nucleotide and oxidative stress metabolism	(Vincent <i>et al.</i> , 2012)
<i>T. brucei</i>	5-fluoro-2'-deoxyuridine	Thymidine synthase inhibition	(Ali <i>et al.</i> , 2013)
<i>T. brucei</i>	5-fluoroorotate & 5-fluorouracil	RNA modification	(Ali <i>et al.</i> , 2013)
<i>T. cruzi</i>	Benznidazole	Thiol metabolism	(Trochine <i>et al.</i> , 2014)
<i>T. cruzi</i>	Bestatin	Leucine aminopeptidase	(Trochine <i>et al.</i> , 2015)
<i>L. infantum</i>	Antimony (III)	Oxidative stress	(Canuto <i>et al.</i> , 2012, Rojo <i>et al.</i> , 2015)
<i>L. donovani</i>	Milefosine	Polyamine pathway	(Canuto <i>et al.</i> , 2014)
<i>P. falciparum</i>	Atovaquone & CK-2-68	Dihydroorotate dehydrogenase	(Biagini <i>et al.</i> , 2012)
<i>P. falciparum</i>	Fosmidomycin	Methylerythritol phosphate cytidyltransferase (IspD) and deoxyxylulose phosphate	(Zhang <i>et al.</i> , 2011a)
<i>P. falciparum</i>	Eflornithine & MDL73811	Polyamine synthesis	(van Brummelen <i>et al.</i> , 2009)
<i>P. falciparum</i>	Pyrimethamine	Dihydrofolate reductase	A. Srivastava <i>et al.</i> , unpublished
<i>P. falciparum</i>	Dihydroartemisinin	Haemoglobin catabolism	(Cobbold <i>et al.</i> , 2015)

**Table 1-1: Example of drug MoA studies using metabolomics.** Adapted from Creek & Barrett (2014).

There are various methodologies employed for metabolomics-based approaches, which are outlined in section 1.9.2. Individual experiments can therefore be altered to suit the requirements of the investigation. In trypanosome biology in particular, metabolomics was successfully employed by Vincent and colleagues to confirm the target of eflornithine, a drug used in combination with nifurtimox in the treatment of late-stage *T. b. gambiense* infection (Vincent *et al.*, 2012) (table 1-1). In addition to confirming the preconceived hypotheses of drug treatment, metabolomics analysis also led to the discovery of further perturbations occurring in treated cells, that were hitherto unknown (Vincent *et al.*, 2012). This highlights one of the main advantages of metabolomics as a tool for drug target deconvolution.

Furthermore, metabolomics-based approaches carry the advantage that mechanisms involving both protein targets as well as non-protein targets, such as nucleic acids, can be identified. One study investigating treatment of *T. brucei* *in vitro* with 5-fluoroorotate and 5-fluorouracil showed the presence of fluorinated uracil residues in the treatment groups, confirming the mechanism of action of these two compounds (Ali *et al.*, 2013). These phenomena cannot be detected using techniques such as the recently developed DARTS and thermal profiling, as discussed below.

Whilst mass spectrometry data often contains a significant amount of background noise, recent work has gone to great lengths to counter this issue with the development of minimal media for the culture of *T. brucei* (Creek *et al.*, 2013). Indeed, similar work is currently underway for *in vitro* culture of *L. mexicana* as well as *T. congolense* (M. Barrett, personal communication). Interestingly,

by mimicking the extracellular environment more accurately, drug sensitivity was shown to be more akin to that reported *in vivo* (Creek *et al.*, 2013). In addition, it is theorised that a minimal essential medium restricts the ability of the parasite to activate many compensatory mechanisms, potentially masking drug efficacy.

### 1.10.5 Systems biology

As mentioned previously, systems biology aims to integrate omics data on multiple levels in order to produce an *in silico* overview of a cell or system. These models can be used to predict metabolic networks, and in the context of drug discovery, predict pathways that are essential to the system in question. Systems biology is mainly used to identify novel drug targets in a target-based drug discovery approach. Many methods in drug discovery are limited exclusively to unique proteins and pathways in the pathogen, which leads to several disadvantages including limited number of targets if the pathogen is similar to the host (e.g. cancer), as well as the disqualification of potentially valid targets (Haanstra & Bakker, 2015).

Genome-wide metabolic models now exist for several organisms, including *T. brucei* (Achcar *et al.*, 2014a, Achcar *et al.*, 2014b), and some pathways in this organism that the host also possesses have been shown have been predicted to be more essential in the pathogen, when compared to the host (Haanstra *et al.*, 2011). In the aforementioned paper, Haanstra and colleagues showed that a 50% inhibition of glycolysis was sufficient to kill *T. brucei*, as this is the sole source of ATP. Whilst the host system also possesses the glycolytic pathway, 50% inhibition is not enough to affect host cells to the same degree. This validates glycolysis as a potential drug target that would not be identified by other approaches, highlighting the role systems biology could play in drug discovery (Haanstra *et al.*, 2011). One complication of this approach is the complexity of the mammalian host, where differing cell types react in varied ways to signals such as hormones and chemokines. These factors must be taken into account to further enhance the model accuracy, but this has proven challenging (Haanstra & Bakker, 2015).

Another method by which computer predictions aid drug discovery, is predicting the effect of multiple targets being inhibited by combining drugs (Kandoi *et al.*, 2015). The phenotype-based approaches for target identification often utilise gene knock-outs or knock-downs, as previously mentioned. Crucially, gene knock-downs are primarily carried out for one gene/protein at a time, due to the fact that targeting more genes increases the complexity of the cloning approaches used in the lab. In contrast, metabolic models can easily investigate the effect of multiple target inhibition *in silico* (Haanstra *et al.*, 2011). This is important, because whilst one gene might not appear to be essential in a pathogen, inhibiting multiple proteins in the same pathway could affect the cell to a much greater degree (Haanstra & Bakker, 2015).

An important consideration for the use of systems biology in drug discovery is that there is a requirement for in depth knowledge concerning the pathogen under investigation. Significantly improved predictions are calculated when enzyme kinetics and metabolic networks are understood to a very high degree (Haanstra & Bakker, 2015). Metabolism, and glycolysis in particular, in *T. brucei* has received much attention for a long time, and in silico modelling taking into account uncertain enzyme kinetics and parameters has allowed a much more accurate depiction of glycolytic metabolism in the parasite (Achcar *et al.*, 2012), as well as other important pathways (Kerkhoven *et al.*, 2013). These models can then be validated experimentally, which was done in both of the aforementioned cases (Achcar *et al.*, 2012, Kerkhoven *et al.*, 2013).

### 1.10.6 Drug affinity responsive target stability (DARTS)

One of the major disadvantages of structure-based approaches in drug discovery is that techniques such as affinity chromatography often require immobilisation of the compound under investigation (Lomenick *et al.*, 2009). For this reason, there is increasing interest in the development of assays that are effective in the absence of drug immobilisation.

One of these assays was developed by Lomenick and colleagues and is termed Drug Affinity Responsive Target Stability (DARTS) (Lomenick *et al.*, 2009). This concept is based on the theory that binding of ligands to targets results in a thermodynamic shift, whereby the conformation of protein and ligand exhibits higher stability and resistance to proteolytic attack compared to unbound conformations (Lomenick *et al.*, 2011b). The DARTS technique has already shown promise in multiple publications, either elucidating novel drug targets, or confirming known targets as a proof-of-principle (Gao *et al.*, 2015, Kost *et al.*, 2015). In addition, the same assay was used to show that  $\alpha$ -ketoglutarate inhibits ATP synthase in *C. elegans*, thereby extending the organism's lifespan (Chin *et al.*, 2014).

Whilst the assay shows good potential, it still carries numerous disadvantages, the majority of which can be linked to the mechanisms of ligand-protein binding. First and foremost, the amount of proteinases used must be carefully optimised. Excessive amounts can cause widespread proteolysis of proteins including the drug target. In contrast, insufficient amounts can lead to a high number of undigested proteins (Lomenick *et al.*, 2011a). In *T. brucei*, VSGs are by far the most abundant proteins in the proteome, and this could lead to bias in the resulting protein gel. Furthermore, the *T. brucei* proteome, under normal conditions, possesses a great variation of stabilities, one example of which is fructose bisphosphate aldolase, which exhibits a high level of proteolytic resistance (Clayton, 1988).

### 1.10.7 Stable isotope labelling by amino acids in cell culture (SILAC)

The use of stable isotopes in complex biochemical or protein mixtures, coupled to mass spectrometry, have generally allowed for high resolution identification of both proteins and metabolites. One method that takes advantage of this technology is termed stable isotope labelling by amino acids in cell culture (SILAC), and was first described more than a decade ago (Ong *et al.*, 2002).

SILAC was not originally developed for drug discovery specifically. In the case of trypanosomatid biology, this method was used to great success to investigate changes in the proteome between different life cycle stages of *T. brucei* (Gunasekera *et al.*, 2012, Urbaniak *et al.*, 2012a). However, more recent adaptations of SILAC include the comparative analyses of drug-treated and negative control cell cultures in order to dissect the mechanisms of drug action. One recent example of the use of SILAC involved studies on a compound called Withaferin A, which is isolated from a medicinal plant (Narayan *et al.*, 2015). Another SILAC-based method was pioneered by researchers at the Board Institute (Ong *et al.*, 2009). In this case, SILAC was combined with affinity enrichment, where lysates were pulled down over beads containing small molecule drugs. Quantitative proteomics was then used to probe for drug target enrichment (Ong *et al.*, 2009). As a proof of concept, drug targets of numerous kinase inhibitors and immunophilin binders were investigated, with a high success rate. Indeed, in addition to previously known targets, several new drug targets were also identified (Ong *et al.*, 2009).

### 1.10.8 The cellular thermal shift assay and thermal profiling

The cellular thermal shift assay (CETSA) was developed by Jafari and colleagues as a method to investigate drug-ligand interaction without requiring immobilised drugs (Martinez Molina *et al.*, 2013, Jafari *et al.*, 2014). In a similar approach to the DARTS method outlined above, CETSA is based on the principle of increased thermodynamic stability during drug-ligand interactions (Jafari *et al.*, 2014). However, whilst the DARTS assay uses a proteolysis approach, the CETSA method relies on protein aggregation at high temperatures. Protein aggregation causes the molecules to become insoluble, and they are removed from the proteome during the experiment. By generating a melting curves of proteins of interest either in the presence or absence of drug, the group observed increased melting temperatures for proteins interacting with a drug (Jafari *et al.*, 2014). As proof-of-principle, the group investigated melting curves for the targets of methotrexate and raltitrexed, which are dihydrofolate reductase and thymidylate synthase inhibitors, respectively (Martinez Molina *et al.*, 2013). Protein lysates were heated at varying temperatures and the proteins of interest were subsequently probed by western blot. The band intensities were quantified relative to the lowest temperature of the experiment, in order to generate a melting curve for each protein of interest.

Building on this approach to drug target identification, a collaborative group of researchers based at GSK in Germany and the Karolinska Institutet in Sweden, developed a high-throughput method based on CETSA, in order to visualise the entire proteome (Savitski *et al.*, 2014). For this experiment, entire protein lysates, rather than one protein of interest, were assayed by mass spectrometry. Samples in the presence or absence of a drug were incubated at 10 different temperatures and labelled using TMT10 isotope tags. Relative protein quantification could then be carried out with just 2 mass spectrometry runs (Savitski *et al.*, 2014). This process was, as shown by several studies, carried out with great success. In addition to altered thermodynamic stability of target proteins, the experiments were also able to identify off-targets in the lysates (Savitski *et al.*, 2014). However, one significant criticism of this study was the lack of replicates and it is likely that this expensive assay would require lengthy optimisation for new systems such as protozoan pathogens.

The authors discussed several other areas where improvement is required. Most importantly, the assay has only been used on soluble protein isolates. This disregards proteins such as transmembrane proteins that might be targeted. Transporters form an important subset of these proteins, and the assay requires significant improvement to enable incorporation of this subset of proteins for investigation. In addition, and much like the DARTS assay, only compounds that target one or multiple proteins can be studied, whereas the assay is rendered useless for those that target, for example, nucleic acids (Savitski *et al.*, 2014).

### **1.10.9 CRISPR – the new kid on the block**

The RNA guided nuclease CRISPR-Cas9 was recently shown to be a very useful tool for gene editing and its use in genetic manipulation of a wide variety of host organisms has been highly publicised (Cong *et al.*, 2013, Mali *et al.*, 2013), with 661 publications on this topic in 2015 alone. The system involves the bacterial clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 nuclease, which is part of the prokaryote adaptive immune response (Moore, 2015). This nuclease can be targeted using a guide RNA with very high specificity to edit genes in other cell types such as mammalian cells, or protozoans (Sollelis *et al.*, 2015).

The CRISPR-Cas9 system shows great promise as a phenotypic screening method to validate new therapeutic targets, especially on a whole genome, high-throughput scale (Koike-Yusa *et al.*, 2014). Indeed, the advantages of this system compared to, for example, the RIT-seq approach are evident as it allows the genome-wide targeting of both essential and non-essential genes (Moore, 2015). It remains to be seen whether CRISPR-Cas9 causes a similar revolution in the study of trypanosomatids, although the early indications are very positive (Lander *et al.*, 2015, Peng *et al.*, 2015, Sollelis *et al.*, 2015, Zhang & Matlashewski, 2015).

### 1.10.10 Summary of drug target identification

Taken together, it is clear new techniques have recently emerged that can be applied to drug discovery, both in target- and structure-based approaches. Whilst the techniques discussed above are perhaps most applicable to the study of the trypanosomatid drugs, the list is by no means finite and other published methods used in drug target deconvolution studies include the yeast three-hybrid system (Rezwan & Auerbach, 2012), mammalian three-hybrid system (Caligiuri *et al.*, 2006), phage display (Omidfar & Daneshpour, 2015), mRNA display (Josephson *et al.*, 2014), protein microarrays (Schax *et al.*, 2014), biochemical fractionation/suppression (Peterson *et al.*, 2006), genome-wide association studies (GWAS) (Grover *et al.*, 2014), drug-induced haploinsufficient profiling (HIP) (Giaever *et al.*, 1999), kinobeads (Kim & Sim, 2010) and further chemoproteomics systems aside from affinity chromatography (Urbaniak *et al.*, 2012b). With this newfound wealth of complementary techniques, the future of drug target and mode of action elucidation looks positive. However, for NTD research, it is of utmost importance that some of these techniques are adapted to relevant organisms. This will have the knock-on effect of generating new therapeutic targets and mechanistic understandings, which will significantly aid in the fight against these debilitating diseases.

## 1.11 Aims & Objectives for this project

---

With the advent of omics technologies coupled to more classical molecular biology approaches, there is a great opportunity to gain detailed insights into the biology of the African trypanosome, in a hypothesis generating manner. This opportunity coincides with the emergence of several new therapeutic leads in fexinidazole and the benzoxaborole AN5568. Both of these have so far shown little to no adverse effects in mammalian cells, show good clearance in murine models for both human-infective *T. brucei* subspecies, and importantly, are orally bioavailable (Torreele *et al.*, 2010, Jacobs *et al.*, 2011). Whilst the mode of action of fexinidazole has been reported in terms of its metabolism and effect on the trypanosome (Jennings & Urquhart, 1983, Kaiser *et al.*, 2011), no MoA has been described for AN5568. This knowledge is crucial, as it could aid in understanding the SAR of the drug, which in turn could lead to improved small molecule inhibitors that bind the target with higher potency and minimise resistance.

Similarly, given the occurrence of LOH in the field, its discovery in a laboratory environment presents a great opportunity to develop an understanding of the mechanisms involved in the metabolic alterations leading to increased growth rates in cell lines exhibiting LOH. This understanding could lead to new hypothesis about the generation and evolution of genetic diversity in the African trypanosomes. Metabolomics analysis tools provide enhanced methods that could aid in the understanding of these mechanisms.

With the above in mind, the main aims of this study were as follows:

1. To elucidate the MoA of AN5568 using a combination of omics and molecular biology approaches.
2. To generate and characterise an AN5568-resistant cell line with a view to complementing the first aim, as well as develop an understanding of the mechanisms that could lead to benzoxaborole resistance.
3. To investigate the changes in metabolism underlying differential growth in trypanosomes possessing LOH on chromosome 10.
- 4.

## Chapter 2. Methods

---

### 2.1 Trypanosome tissue culture

---

#### 2.1.1 Bloodstream form culture

Bloodstream form (BSF) culture was carried out using the 427 Lister strain, a monomorphic strain of *T.b. brucei* well adapted to *in vitro* conditions. Lister 427 express VSG221 with a relatively low switch rate (between  $10^{-6}$ – $10^{-7}$  switches per generation) (Lamont *et al.*, 1986). For RNAi experiments, the 2T1 cell line, derived from Lister 427, was used (Alsford *et al.*, 2011).

Cells were cultured at a density between  $2 \times 10^4$  cells/mL and  $2 \times 10^6$  cells/mL, and incubated at 37°C in a humidified incubator with 5% CO<sub>2</sub>. Densities were routinely assessed using a 2-cell counting chamber (Hawksley). To maintain stable cultures, cell lines were typically discarded after a maximum of 30 passages, and substituted for a new isolate from the same stabilate batch. Cells were passaged, on average, every two days and maximum three. This approach limits a build-up of genetic abnormalities that can occur as a result of *in vitro* culture.

BSF parasites were cultured in HMI-9, a medium supplied in powder form (Gibco). The media was supplemented with 10% foetal bovine serum (FBS). For targeted metabolomics experiments, Creek's Minimal Media (CMM) was used. This medium is based on HMI-9, but contains fewer components, without reducing cellular growth. CMM thereby reduces background noise in experiments such as mass spectrometry (Creek *et al.*, 2013). In addition, this medium allows for the addition of isotope-labeled components in place of their standard counterparts. CMM was supplemented with gold standard FBS (10%) (PAA).

#### 2.1.2 Procyclic form culture

The procyclic form (PCF) trypanosomes used in this study were either Lister 427 (also identified as 29-13), or Trypanosomiasis Research Edinburgh University (TREU) 927, the strain used for the reference genome of *T. brucei* (Berriman *et al.*, 2005). PCFs were typically cultured in SDM-79, supplemented with 10% FBS and 7.5 µg/mL hemin. In some cases, a more defined medium, SDM-80 was used (Appendix B). SDM-80 allows more control over the concentrations of proline and glucose, the main carbon sources used by the parasite, added to media. PCF cell density was kept between  $5 \times 10^5$  cells/mL and  $1 \times 10^7$  cells/mL and grown at 27°C. PCFs do not require artificial levels of CO<sub>2</sub> or humidified incubators.

### 2.1.3 Differentiation of bloodstream forms to procyclic forms

In some experiments, attempts were made to differentiate BSF cells into PCF cells *in vitro*. In these experiments, cells were cultured in differentiating trypanosome medium (DTM) containing 3 mM cis-aconitate (Overath *et al.*, 1986) (Appendix A). These cells were incubated at 27°C, under conditions similar to PCF trypanosomes. Cell lines that survived subsequent passages were kept in DTM lacking cis-aconitate.

### 2.1.4 Stabilate preparation and retrieval

Where novel cell lines were made, these were subsequently stabilized in order to maintain a cell line library. Cell cultures in mid-log phase of growth were diluted in a 1:1 ratio with medium containing 20% glycerol, resulting in a cell culture aliquot of 10% glycerol. Up to ten 1 mL aliquots were then made in 2 mL cryovials (Alpha Laboratories, UK). These were kept in wool at -80°C for 24 hours to slowly freeze the cells, and subsequently they were transferred to liquid nitrogen for long-term storage. In order to revive stabilized cell lines, 1 mL cell aliquots were quickly thawed and transferred to 9 mL culture medium warmed to the required temperatures. Drugs and selective antibiotics were added after the first passage.

### 2.1.5 Alamar blue assay

To obtain *in vitro* EC<sub>50</sub> values for specific compounds targeting *T. brucei*, the alamar blue assay was applied (adapted from (Raz *et al.*, 1997)). This colorimetric assay was carried out in solid white flat-bottomed 96-well plates. Compounds were added starting with the highest concentration (typically 100 µM) and serially diluted over either 23, or 11 wells, leaving one negative control. Subsequently, 100 µL cells at a density of  $4 \times 10^4$  cells/mL was added to each well, resulting in a final density of  $2 \times 10^4$  cells/mL. Bloodstream form parasites were incubated for 48 hours at 37°C, 5% CO<sub>2</sub>. Subsequently, 20 µL of alamar blue reagent (resazurin sodium salt, 0.49 mM in 1× PBS, pH 7.4) was added to each well, and the plate incubated for another 24 hours. When carried out on slower growing PCF parasites, plates were incubated for 72 hours prior to addition of alamar blue reagent, and then incubated for another 48 hours.

Reduction of the alamar blue reagent was measured as a function of cell viability on a BMG FLUOstar OPTIMA microplate reader (BMG Labtech GmbH, Germany) with  $\lambda_{\text{excitation}} = 544$  nm and  $\lambda_{\text{emission}} = 590$  nm. The raw values were plotted against the log value of each concentration of drug or compound, and EC<sub>50</sub> values were calculated using a non-linear sigmoidal dose-response curve. Each assay was performed in duplicate, and the final EC<sub>50</sub> values presented represent a mean of three or four independent experiments.

### 2.1.6 Isobologram assay

To investigate drug-drug interactions, isobolograms were carried out using a fixed-ratio protocol previously described (Fivelman *et al.*, 2004). This assay uses the same principles as an alamar blue assay to test for cell viability over a range of drug concentrations but test two drugs simultaneously.

For both drugs, the top concentrations used were chosen for the EC<sub>50</sub> to fall near the midpoint of a 12-part two-fold dilution series. We typically started with 16× EC<sub>50</sub> and concentrations are indicated where applicable. Once drug stocks were prepared in parasite medium, fixed ratio solutions of drugs were made as follows: 10:0, 9:1, 8:2, 7:3, 6:4, 5:5, 4:6, 3:7, 2:8, 1:9, 0:10. These stocks were then added to the first column of 3 solid white flat-bottomed 96-well plates in duplicate. The compounds were then serially diluted with the final well of each row left blank as a negative control. Finally, cells were added at a final concentration of  $2 \times 10^4$  cells/mL. Plates were then incubated for 48 hours and alamar blue reagent added as described above. The plates were read after another 24 hour incubation on a BMG FLUOstar OPTIMA microplate reader (BMG Labtech GmbH, Germany) with  $\lambda_{\text{excitation}} = 544$  nm and  $\lambda_{\text{emission}} = 590$  nm.

For each fixed ratio, one EC<sub>50</sub> value was generated for each drug. These values were then used to obtain fractional inhibitory concentration (FICs) indices, which are defined as the EC<sub>50</sub> values of drug in combination, divided by the EC<sub>50</sub> values of those drugs acting alone (Hall *et al.*, 1983). The isobologram was then made by plotting the FICs of one drug against the other. An overall mean  $\Sigma$ FIC was calculated for each combination by adding the values of each individual FIC of either drug and taking an average. Interactions were then defined as described by Seifert and colleagues (Seifert *et al.*, 2011):

- $\Sigma\text{FIC} \leq 0.5$  = synergy
- $\Sigma\text{FIC}$  between  $> 0.5$  and  $\leq 4$  = indifference
- $\Sigma\text{FIC} > 4$  = antagonism

### 2.1.7 Transfection of trypanosomes

To transfect trypanosomes with genetic constructs *in vitro*, a previously described method was used (Schumann Burkard *et al.*, 2011). Up to  $4 \times 10^7$  cells in mid-log growth were centrifuged at  $1,500 \times g$  for 10 minutes, and resuspended in 100  $\mu\text{L}$  transfection buffer (Appendix D) (Schumann Burkard *et al.*, 2011). Cells were then transferred to a transfection cuvette along with 10  $\mu\text{L}$  H<sub>2</sub>O, containing 10  $\mu\text{g}$  linearised plasmid DNA. Cells were electroporated using an Amaxa Nucleofector II (Lonza), set to the programme X-001. After electroporation, cells were transferred by Pasteur pipette to 30 mL warm HMI-9. After a 10 minute incubation at 37°C, 5% CO<sub>2</sub>, cells were cloned by dilution in 24-well plates at 1:20 and 1:40 dilutions, and left at 37°C, 5% CO<sub>2</sub>. Selective antibiotics were added at the required concentration after 24 hours. Successfully transfected cells were harvested after 7

days. When <30% of wells contained cells, these were assumed to be clonal. If not, another round of cloning by dilution was carried out. For each transfection, two different clones were used in downstream experiments, unless mentioned otherwise. All resulting cell lines were given nomenclature in accordance with classifications suggested by Clayton and colleagues (Clayton *et al.*, 1998).

### 2.1.8 Generation and selection of AN5568-resistant *T. brucei*

To generate *in vitro* resistance to the benzoxaborole AN5568, the 427 Lister cell line was used in HMI-9. First, RNA was extracted from the parental line, to store for downstream analyses. Resistance generation was initiated by growing cells in 170 nM of the compound. Once a cell line was established in this concentration, cells were then transferred to 6 wells of a 24-well plate at a density of  $1 \times 10^5$  cells/mL. Six incrementing concentrations of AN5568 were then added, to obtain six 2 mL cultures. Typically, increments of 0.2  $\mu$ M were used, and cells were incubated at 37°C and 5% CO<sub>2</sub>. Cells were routinely observed under a microscope and after 2 weeks, live cells in the highest concentration of AN5568 were taken and cultured for 2 passages in the same concentration of the benzoxaborole, before transfer to 6 wells in a 24-well plate, from which point the process was repeated using the new AN5568 concentration as the lowest. A wild-type cell line was grown in the absence of drug as a “highly-passaged” control, in order to detect transcriptome changes related to *in vitro* culture adaptation. Once sufficient resistance was deemed to have been generated, cells were cloned by dilution and 4 isolates were recovered. Drug resistance was measured by alamar blue.

## 2.2 Isolation of biological components from trypanosomes

---

### 2.2.1 DNA Extraction

Genomic DNA extractions from trypanosomes were carried out to use in PCR reactions, to genotype and to verify construct integration. Between  $1 \times 10^7$  and  $2 \times 10^7$  cells were centrifuged at  $1,500 \times g$  for 10 minutes and DNA was extracted from pellets using a commercially available kit (Nucleospin® Tissue, Macherey-Nagel GmbH), according to the manufacturer’s instructions.

For experiments that required highly purified and concentrated DNA, an ethanol precipitation was carried out on DNA samples purified using the aforementioned approach. Firstly, 1/10 $\times$  volume of 3 M sodium acetate (pH 5.2) was added, followed by 2.5 $\times$  volumes 100% ethanol. After a brief vortex, this mixture was left at room temperature for at least one hour. Subsequently, the sample was centrifuged for 10 minutes at  $16,060 \times g$ , 4°C. Supernatant was removed and the DNA pellet was washed with 70% ethanol before centrifugation for another 2 minutes under the aforementioned conditions. After removal of the supernatant, the DNA pellet was left to dry in a

laminar flow hood. The desired volume of filter-sterilised ddH<sub>2</sub>O was then added to resuspend the DNA. Concentration was confirmed using a NanoDrop (Thermo Scientific).

### 2.2.2 RNA extraction

For RNA extractions, between  $4 \times 10^7$  and  $5 \times 10^7$  cells were centrifuged for 10 minutes at  $1,500 \times g$ . After removal of the supernatant, cells were resuspended in remaining medium, transferred to a 1.5 mL eppendorf tube and spun again for 5 minutes. Cell pellets were then resuspended in 1 mL TRIzol® (Life Technologies, Thermo Fischer Scientific). Samples were stored at  $-80^\circ\text{C}$  until extractions were carried out according to the kit instructions. Briefly, 200  $\mu\text{L}$  chloroform was added, before tubes were vigorously shaken and centrifuged to separate the aqueous phase containing RNA. This RNA was further purified, treated with DNase, and extracted into RNase-free H<sub>2</sub>O using a commercially available kit (Nucleospin® RNA, Macherey-Nagel GmbH). RNA concentration was assessed with a NanoDrop (Thermo Scientific).

### 2.2.3 Protein extraction

Proteome isolation from trypanosomes was carried out in differing buffers, dependent on the downstream applications. Typically,  $1 \times 10^8$  cells in mid-log phase were centrifuged for 10 minutes at  $1,500 \times g$ ,  $4^\circ\text{C}$ . Supernatant was removed and cells transferred to an eppendorf tubes before they were washed twice with ice-cold  $1\times$  phosphate buffered saline (PBS) (Appendix D). After removing the supernatant, cell lysates were resuspended in protein extraction lysis buffer (Appendix D), and left on ice for 30 minutes with intermittent vortexing. If required, samples were broken up by sonication. Otherwise, the protein concentration was measured using a Bradford assay (Bio-Rad Laboratories Inc), according to the kit instructions.

Fractionation of the proteome was achieved using a hypotonic lysis buffer (Appendix D). Cells were resuspended in this buffer after the PBS wash, and incubated on ice for 5 minutes, prior to centrifugation at  $20,000 \times g$  for 10 minutes. The soluble fraction was removed and stored at  $-20^\circ\text{C}$ . The pellet was then resuspended in more hypotonic lysis buffer and spun down again to obtain the wash fraction. Finally, the remaining pellet, containing the membrane fraction, was resuspended in ice-cold protein extraction lysis buffer. All samples were stored at  $-20^\circ\text{C}$ .

For 2D DIGE experiments, cell lysates were washed and resuspended in DIGE lysis buffer (Appendix D) supplemented with a cOmplete™ ULTRA protease inhibitor cocktail (Roche, Switzerland). These samples were then sonicated in three 1-second cycles followed by incubations on ice to avoid overheating of the sample. After a 10-minute incubation at room temperature, samples were centrifuged for 10 minutes at  $13,000 \times g$  and supernatant was transferred to a fresh eppendorf tube.

To hyperconcentrate and clean the protein lysates for 2D DIGE, an acetone precipitation was performed. Four volumes of ice-cold 100% acetone were added to the protein samples, which were then incubated at  $-20^{\circ}\text{C}$  overnight to allow protein to precipitate. Samples were centrifuged at  $13,000 \times g$  for 10 minutes at  $4^{\circ}\text{C}$ . The supernatant was removed and pellets were washed with 80% ice-cold acetone before being centrifuged again under the same conditions. This washing step was repeated twice to remove contaminants. After the supernatant was removed for the final time, pellets were left to air-dry in a laminar flow hood. Finally, the pellets were resuspended in the desired volume of DIGE lysis buffer and protein content was evaluated using a Bradford assay (Bio-Rad Laboratories Inc.), according to the manufacturer's instructions.

## 2.3 Electrophoresis

---

### 2.3.1 DNA gel electrophoresis

DNA gel electrophoresis was typically performed on 1% UltraPure™ agarose gels (Life Technologies) that were made with  $1\times$  TAE (Appendix D) containing a 1:10,000 dilution of SYBRsafe (Invitrogen, Life Technologies). Gels were run in tanks filled with  $1\times$  TAE at 110 V. In the first and last lanes, a commercial 1 kb or 100 bp DNA ladder was added (Promega). Gels were visualised by fluorescence, using a BioDoc-It® imaging system (UVP).

### 2.3.2 Protein gel electrophoresis

Protein samples were mixed in a 1:4 ratio with SDS loading buffer (Appendix D) and then boiled at  $100^{\circ}\text{C}$  for 10 minutes. 20  $\mu\text{L}$  of each sample was then loaded onto a NuPAGE® Novex® 4-12% Bis-Tris protein gel (Life Technologies), set up in an Xcell SureLock® Mini tank system with  $1\times$  SDS running buffer (Appendix D). The outer lanes of the gel contained dual colour protein standard ladders (Bio-Rad Laboratories Inc.). Electrophoresis was run at 200 V for approximately 40 minutes. This allowed an expected current of 110–125 mA. Subsequent to electrophoresis, gels were either used for Western blots (section 2.9), or proteins were visualised by Coomassie.

For Coomassie staining, gels were incubated at room temperature on a shaker in Coomassie staining solution (Appendix D) for at least 4 hours, or overnight. The gel was then destained in destaining solution (Appendix D), which was replaced every 30 minutes for 2-3 hours. Gels were visualised using a myECL™ Imager system (Thermo Fischer Scientific).

### 2.3.3 Two-dimensional difference gel electrophoresis (2D DIGE)

For the 2D DIGE experiment, protein lysates were prepared using the aforementioned protocols (section 2.2.3). This experiment was carried out with a control sample and a treatment sample. The protein concentration of each sample was measured using a Bradford assay (Bio-Rad Laboratories Inc), and adjusted to 5 mg/mL. For each sample, a 10  $\mu\text{L}$  aliquot was prepared (50  $\mu\text{g}$ ). The pH was

checked using pH strips (Macherey-Nagel GmbH), and if pH was judged to be between 8 and 9, either CyDye Fluor Cy5 or CyDye Fluor Cy3 (both GE Healthcare) were added to each aliquot at a final concentration of 400 pmol. Sample and dye were mixed by pipetting and incubated in the dark, on ice, for 30 minutes, before the reaction was stopped by addition of 1  $\mu$ L 10 mM lysine. The aliquots were then combined together and diluted with rehydration buffer (Appendix D). This sample, containing both control and treatment lysates bound to different fluorescent markers, was pipetted onto an immobilised pH gradient (IPG) gel (Bio-Rad Laboratories Inc.). Once rehydration of the strip was complete, isoelectric focusing was set up with an IPGphor isoelectric focusing system (GE Healthcare) and run overnight (Cycling conditions: Appendix F). Upon IEF completion, the strip was equilibrated with strip equilibration buffer, and washed twice, before laying on top of an SDS-page gel (Appendix D) which was run at 1 W overnight. Spots were analysed by fluorescence using a Typhoon 9410 Variable Mode Imager (GE Healthcare). If gel extraction was required, another gel was run using pooled samples of 250  $\mu$ g of each sample. This gel was run alongside the gel containing CyDyes, but stained with colloidal Coomassie (Appendix D) upon completion of gel electrophoresis, spots of interest were first found using DeCyder 2-D Differential Analysis Software (GE Healthcare), and then excised under sterile conditions. Excised parts were placed in wells of a sterile 96 well plate. Isolation of protein samples from gel as well as mass spectrometry, was carried out by Richard Burchmore at Glasgow Polyomics.

## 2.4 Molecular biology

---

### 2.4.1 PCR

Polymerase chain reaction set up was dependent on the application. For the amplification of genes for the generation of expression or knock-down constructs, a high-fidelity proofreading polymerase (New England Biolabs Ltd) was used in 50  $\mu$ L reactions. The cycling conditions for HF polymerase PCRs are shown in Appendix F. For routine PCR assays throughout molecular cloning experiments, such as colony screens and construct integration confirmations, GoTaq polymerase (Promega) was used in 20  $\mu$ L reactions using cycling conditions in Appendix F. Primers are listed in Appendix E.

### 2.4.2 Genotyping

Genotyping was routinely carried out on chromosome 10 of the TREU 927 PCF lines and its LOH derivatives. This was done to confirm cultures were labelled correctly and no further genetic changes had occurred. Two genotyping primer-sets, 3778 and CRAM, were used for this. The PCR conditions are summarised in Appendix F. Each PCR reaction contained the following: 45 mM Tris-HCl (pH 8.0), 11 mM  $(\text{NH}_4)_2\text{SO}_4$ , 4.5 mM  $\text{MgCl}_2$ , 6.7 mM 2-mercaptoethanol, 4.4  $\mu$ M EDTA, 113  $\mu$ g/mL BSA, 1 mM dATP, 1 mM dGTP, 1 mM dTTP, 1 mM dCTP (Custom PCR MasterMix), 1  $\mu$ M of each primer (Appendix E), 0.5 units Taq DNA polymerase and 1  $\mu$ L DNA template at a

concentration of 2 ng/ $\mu$ L. The 3778 and CRAM PCR products were run on 3% and 1% agarose gels respectively and gels were visualised by fluorescence, using a BioDoc-It<sup>®</sup> imaging system (UVP).

### 2.4.3 Plasmids

The various plasmids used in this study are outlined in Appendix E. In all cases the gene of interest was amplified by PCR, before adding A-overhangs and ligation into the intermediate vector pGEM-T easy (Promega). Insertion of fragments into this plasmid disrupts a LacZ gene, allowing for blue-white screening of bacterial colonies. Plasmids were transformed under standard conditions and plated onto LB agar plates containing 150  $\mu$ g/mL ampicillin. Colony screens were carried out by PCR and single positive colonies were grown overnight at 37°C in LB broth supplemented with 100  $\mu$ g/mL ampicillin. Extraction of plasmid DNA from overnight cultures was carried out using a commercial kit, without protocol modification (Nucleospin<sup>®</sup> Plasmid, Macherey-Nagel, GmbH). Purified plasmid was eluted with 50  $\mu$ L elution buffer.

To ligate the PCR products into their final plasmid, they were extracted from pGEM-T easy through a restriction digest followed by an agarose gel extraction. Final vectors were treated with Antarctic phosphatase (New England BioLabs) for 1 hour at 37°C, to avoid self-ligation. PCR products were then ligated into the corresponding vector using T4 DNA ligase (Promega).

### 2.4.4 RNA interference constructs

RNA interference was carried out using published methods developed by Alsford *et al.* (Alsford *et al.*, 2011). The system is based on the pHellsgate4 plasmid, and is termed pGL2084 (Helliwell *et al.*, 2002). It contains two sets of opposing AttP1 and AttP2 recombination sites flanking counter-selectable markers. A single PCR product with flanking AttB1/2 sites was recombined in opposing directions into pGL2084, to generate a stem loop RNAi construct. Successful integration of PCR-products into the vector remove counter-selectable markers, allowing the vector to be amplified in DH5 $\alpha$  *E. coli* cultures. Confirmation of successful integration was determined through restriction digests, and the vector was linearised using AscI. Finally, linearised DNA was concentrated with an ethanol precipitation and resuspension in sterile distilled H<sub>2</sub>O (500 ng/ $\mu$ L).

RNAi constructs in pGL2084 were transfected into the 2T1 cell line (Alsford *et al.*, 2011). Under normal conditions, this cell line contains puromycin and phleomycin resistant cassettes, and is grown in 0.2  $\mu$ g/mL and 0.5  $\mu$ g/mL of these antibiotics respectively. Upon successful transfection and integration of the RNAi construct, puromycin resistance is removed, and cells become hygromycin resistant (2.5  $\mu$ g/mL), and puromycin sensitive, whilst maintaining phleomycin resistance. RNAi-mediated knockdown was induced by the addition of 1  $\mu$ g/mL tetracycline from a 1 mg/mL stock. During induction experiments, tetracycline was added to cultures every 24 hours.

To confirm successful gene knock-downs, RNA was extracted from cells in the presence or absence of tetracycline over a period of 5 days. RNA was reverse-transcribed and qPCRs were carried out using the cDNA as described in section 2.5.

### 2.4.5 Overexpression & re-expression constructs

Two plasmids were used to overexpress *T. brucei* proteins. The primary plasmid utilised in this study was pURAN, a ribosomal array-targeting plasmid. This vector was a gift from the Stephen Hajduk group in Georgia, USA. The vector contained one EcoRI restriction site into which genes of interest were cloned with T4 DNA ligase. Colonies containing the pURAN::GOI construct were screened by PCR using one gene-specific and one construct-specific primer. The plasmid was linearised with BstXI or Aval before transfection into *T. brucei*. Successful transfectants were screened using a G418 selectable marker at a concentration of 2.5 µg/mL. Integration of the construct was then assessed by PCR using one gene-specific primer and one rRNA-targeting primer. Increases in GOI transcript abundance was assessed by qPCR.

In some cases, pURAN-mediated overexpression was detrimental to *T. brucei* cells, as judged by inability to grow. In these cases, the pRM482 plasmid was used (Trenaman *et al.*, 2013). This construct is targeted to the tubulin array of the *T. brucei* genome. Promoters here lead to constitutive expression of genes encoded in the construct, but to a much lesser degree than that ribosomal locus integration.

## 2.5 Real-time qPCR

---

### 2.5.1 Reverse transcription

To confirm successful knock-down, over-expression or re-expression of genetic constructs transfected into *T. brucei* cells, real-time quantitative PCR (qPCR) was used. This process involves reverse transcription of total cellular RNA to cDNA, followed by qPCR analysis of genes of interest alongside an endogenous control.

All experiments involving RNA were carried out in RNase-free conditions, using certified RNase-free or DEPC-treated H<sub>2</sub>O as well as RNase-free components (Appendix D). RNA was extracted from cells as described in section 2.2.2. Reverse transcription was initiated by adding 250 ng random primers (Invitrogen), 1 µL of 10 mM dNTPs (Roche) and H<sub>2</sub>O to 1 µg RNA in a 14 µL reaction in a PCR tube. This mixture was heated to 65°C for 5 minutes to allow denaturing of RNA and binding of random primers, and was then incubated on ice for 1 minute. For the reverse transcription, a commercial kit, Superscript III<sup>®</sup>, was used (Invitrogen). The RT reaction was supplemented with 4 µL first strand buffer, 1 µL 0.1 M DTT and 1 µL Superscript III<sup>®</sup>. A negative control lacking the reverse transcriptase was also set up for each sample. All reactions were then incubated at room temperature for 5

minutes before a 60 minute incubation at 50°C. The reaction was subsequently inactivated by heating it to 70°C for 15 minutes. Finally, template RNA was removed by the addition of 2U *E. coli* RNase H (Life Technologies) and incubation at 37°C for 20 minutes. Samples were stored at -20°C and cDNA prepared in this way was used directly in qPCR reactions.

### 2.5.2 Primers and primer efficiency

To guarantee accurate data interpretation, each newly acquired primer-set for qPCRs was assessed for their efficiency. Primers were designed using the Primer Express software, version 3.0 (Life Technologies, USA), and were supplied by Eurofins Genomics (Germany). All qPCR primers used in this study are described in table E-2 (Appendix E). Wild-type *T. brucei* Lister 427 DNA was used as template for primer efficiency analysis.

Primers were diluted to 3 µM stock, and 25 µL reactions were set up on 96-well MicroAmp® optical 96-well qPCR reaction plates (Applied Biosystems) with 12.5 µL SYBR® Green master-mix (Life Technologies), 2.5 µL of each primer, 6.5 µL H<sub>2</sub>O and 1 µL template DNA. For each primer set, 8 reactions were set up in triplicates with the following amounts template DNA (from 2 ng/µL stock): 2 ng, 1 ng, 500 pg, 250 pg, 125 pg, 62.5 pg, 31.25 pg and negative control. The qPCR was run using an Applied Biosystems 7500 RT PCR system with cycling conditions described in (Appendix F).

Results were downloaded in .csv files containing the Ct values for each reaction well. These Ct values were then plotted against Log<sub>10</sub> of the corresponding DNA concentration in order to obtain an efficiency slope. The following equations were then used to calculate the primer efficiency:

$$E = 10^{-1/slope}$$

$$Efficiency = (E - 1) \times 100$$

The theoretical maximum of 100%, using the aforementioned calculations, would indicate doubling of the amount of product with each cycle, during the logarithmic phase of the reaction (Bustin *et al.*, 2009). However, it is not uncommon for primer efficiency to be higher than 100% (Bustin *et al.*, 2009). Using these calculations, primer sets were judged to be accurate if efficiency was >90%. Otherwise they were discarded and new sets designed.

### 2.5.3 Real-time PCR

Real-time PCRs were set up in triplicate with the following reaction in each well: 12.5 µL SYBR® Green Master Mix (Life Technologies), 2.5 µL forward primer, 2.5 µL reverse primer, 5 µL RNase-free H<sub>2</sub>O and 2.5 µL cDNA. RT-PCR was carried out using an Applied Biosystems 7500 RT PCR system with cycling conditions described in (Appendix F).

## 2.6 Metabolomics

---

### 2.6.1 Metabolite extraction from *T. brucei* cultures for LC-MS

Samples for metabolomics analysis were acquired by rapidly quenching  $8 \times 10^7$  cells in log phase in a dry-ice/ethanol bath, to  $\sim 4^\circ\text{C}$ . For each sample group, four replicates were grown independently. In experiments involving stable isotope labelling, three replicates were used and isotope-labeled compounds ( $^{13}\text{C}$ -(U)-L-methionine, enrichment 99%, Cambridge Isotope Laboratory Inc., cat: CLM-893-H-0.1) were added at the moment the time course was initiated. After quenching, samples were centrifuged for 10 minutes at  $1,500 \times g$ ,  $4^\circ\text{C}$ , and all experimental steps hereafter were done on ice to keep the samples cold. Subsequently, 5  $\mu\text{L}$  supernatant was transferred to an eppendorf containing 200  $\mu\text{L}$  extraction solvent (Appendix D), and the rest removed. Cells were resuspended in the remaining supernatant and transferred to an eppendorf so that they could be centrifuged at  $1,500 \times g$  for another 5 minutes. Remaining supernatant was carefully removed, and the cells resuspended in 200  $\mu\text{L}$  extraction solvent. All samples, including a blank and fresh medium control, were then left in a shaker at  $4^\circ\text{C}$  for one hour. Subsequently, samples were spun down at  $16,060 \times g$  for 10 minutes, and the supernatant was transferred to a 2 mL screw-top tube. From each sample, 10  $\mu\text{L}$  was transferred to an empty tube to produce a quality control sample. Finally, air in the tubes was displaced with argon gas to prevent oxidation of metabolites, and samples were stored at  $-80^\circ\text{C}$  until they were analysed by liquid chromatograph-mass spectrometry.

### 2.6.2 Liquid chromatography-mass spectrometry

All mass spectrometry of metabolite samples was carried out by Glasgow Polyomics. Metabolomics samples were separated by HPLC using a ZIC-pHILIC (polymer-based hydrophilic interaction liquid chromatography) column (Merck). Two solvents were used in the column. Solvent A was 20 mM ammonium carbonate in  $\text{H}_2\text{O}$  and solvent B was 100% acetonitrile. Mass detection was carried out using an Exactive Orbitrap mass spectrometer (Thermo). The mass spectrometer was run in positive and negative mode with an injection volume of 10  $\mu\text{L}$  and a flow rate of 100  $\mu\text{L}/\text{minute}$ .

### 2.6.3 LC-MS Data analysis

Raw data produced by the mass spectrometer was converted to the mzXML format using msconvert (Chambers *et al.*, 2012). This step also split the polarity of the data. Files were then converted to peakML files with XCMS, which uses the Centwave function to pick peaks, converting every individual file to the peakML output. Data was combined across replicate samples and related peaks were annotated using mzMatch, an R-based analysis package. mzMatch output was processed using Ideom (Scheltema *et al.*, 2011, Creek *et al.*, 2012b), an excel-based annotation tool, which uses known standards to assign putative annotations to the filtered peaks.

In experiments where  $^{13}\text{C}$ -isotope labelling was used, mzXML files were analysed using mzMatch-ISO (Chokkathukalam *et al.*, 2013). This R-based tool requires a tab-delimited file containing metabolites of interest and their corresponding formulae, which was obtained from the KEGG database (Okuda *et al.*, 2008). This list is then used by mzMatch-ISO to look for labelled metabolites which possess identical retention times, but differ in their mass. Peaks were subsequently manually filtered based on noise and intensity.

## 2.7 Proteomics

---

### 2.7.1 Drug affinity responsive target stability (DARTS)

The DARTS assay was recently described by Lomenick and colleagues (Lomenick *et al.*, 2009, Lomenick *et al.*, 2011a, Pai *et al.*, 2015), who used it to successfully validate several small molecule drug targets in both mammalian and yeast cells. The protocol outlined by the group was adapted for use with *T. brucei* cells.

In this study,  $1 \times 10^8$  Lister 427 *T. brucei* cells were used for one experiment. Cells were centrifuged at  $1,500 \times g$  for 10 minutes. Supernatant was removed and cells were washed twice in ice-cold  $1 \times$  PBS, before resuspension in DARTS lysis buffer (Appendix D). After an hour-long incubation at  $4^\circ\text{C}$  with intermittent vortexing, the sample was centrifuged at  $16,060 \times g$  for 30 minutes at  $4^\circ\text{C}$ , to remove cellular debris. The protein content was then measured with a Bradford assay (Bio-Rad Laboratories Inc), and adjusted to 1 mg/mL by adding additional buffer.

The sample was then split into three aliquots of 297  $\mu\text{L}$  in 1.5 mL eppendorf tubes (the remaining sample was discarded). To one of each aliquot was added either the AN5568 benzoxaborole at varying concentration, or the drug vehicle (i.e. DMSO). Samples were mixed, briefly centrifuged, and incubated at  $37^\circ\text{C}$  for one hour.

Each aliquot was subsequently split into 6 aliquots of 50  $\mu\text{L}$ . Pronase (Sigma-Aldrich), which was kept as a 10 mg/mL stock, was added at the following final concentrations to 5 of the 6 samples for each sample group: 50  $\mu\text{g}/\text{mL}$ , 10  $\mu\text{g}/\text{mL}$ , 5  $\mu\text{g}/\text{mL}$ , 1  $\mu\text{g}/\text{mL}$  and 0.5  $\mu\text{g}/\text{mL}$ . One sample from each sample group was left as a negative control. Samples were incubated with pronase for exactly 30 minutes at room temperature. In order for this to be carried out successfully, a timer was started after pronase was added to the first sample. The next pronase digest was then set up 30 seconds later and this was repeated until all digests had been set up. After 30 minutes, protease inhibitor (Roche) was added to the first sample. Protease inhibitor was then added to each subsequent sample every 30 seconds in the same order as the digests had been initiated. Samples were then left on ice until the experiment was complete.

To visualise the DARTS experiment results, 9  $\mu\text{L}$  from each sample was added to 3  $\mu\text{L}$  SDS loading buffer and boiled at 100°C for 10 minutes. All samples were then run on a NuPAGE® Novex® 4-12% Bis-Tris protein gel (Life Technologies, USA). Sample to which the same concentration of pronase has been added, were run next to each other (for example, DMSO and benzoxaborole samples digested with 5  $\mu\text{g}/\text{mL}$  were separated in parallel, so they could be directly compared). Protein gels were run and stained with Coomassie as described in section 2.3.2.

### 2.7.2 Protein mass spectrometry and analysis

Protein gel extractions, trypsin digests and mass spectrometry was carried out by Glasgow Polyomics. Submitted gel pieces from either protein gels or 2D-DiGE gels were washed twice with 100 mM ammonium bicarbonate (ABC), followed by destaining using 50% acetonitrile/100 mM ABC. Gels were incubated in acetonitrile for 10 minutes and subsequently dried in a speed vac. Samples were then re-hydrated using 25 mM ABC and incubated overnight.

Digested proteins were dried and analysed by LC-MS/MS. An UltiMate 3000 Rapid Separation LC system (Thermo) was used to separate samples and this was coupled to an electrospray ionization tandem mass spectrometer (Amazon ETD, Bruker Daltonics). Raw MS/MS data was analysed using the MASCOT server (Hirose *et al.*, 1993) and protein sequences were identified using the TREU 927 (v5.0) *T. brucei* reference proteome.

## 2.8 Transcriptomics

---

### 2.8.1 Sample preparation

RNA was extracted from samples as described in section 2.2.2, using approximately  $10^8$  cells to ensure the isolation of surplus RNA. The total RNA was purified using a commercial kit (Nucleospin® RNA, Macherey-Nagel). A DNase treatment step was included in the kit protocol. RNA was stored at -80°C and transported to Glasgow Polyomics on dry ice.

### 2.8.2 RNA-seq

Purified RNA was first subjected to a quality control by NanoDrop (Thermo Fischer), and then sequenced by Glasgow Polyomics. The RNA library was prepared using PolyA selection using the TruSeq stranded mRNA sample prep kit (Illumina®). Sequencing was carried out using Illumina NextSeq500 sequencing apparatus.

### 2.8.3 Transcriptomics data analysis

Raw sequencing files, in .file format, were processed using a pipeline as follows: Raw reads were trimmed using cutadapt to trim adapter sequences (Martin, 2011). The read quality and coverage was then assessed using FastQC (Andrews, 2010). Once all sequencing data was judged to be of

good quality, the reads were then aligned to the TREU927 genome (TriTrypDB, v9.0) using the TopHat RNAseq data processing pipeline (Trapnell *et al.*, 2009, Trapnell *et al.*, 2012, Ghosh & Chan, 2016), which incorporates Bowtie2 to align reads and Samtools to remove duplicates (Li *et al.*, 2009). For sequencing analysis, the data was further processed using the genome analysis tool-kit (GATK) (McKenna *et al.*, 2010), and SNPs as well as indels were filtered using the SnpEff/SnpSift package (Cingolani *et al.*, 2012). To analyse transcript abundance the cufflinks package was used (Trapnell *et al.*, 2012, Ghosh & Chan, 2016).

## 2.9 Blotting

---

### 2.9.1 Sample preparation

Protein samples for Western blots were prepared as described in [section 2.3.2](#). Protein was extracted either in protein extraction lysis buffer, or in hypotonic lysis buffer for fractionation. Samples were diluted in 4x sample loading buffer (Appendix D) and boiled at 100°C for 10 minutes, with one exception; For Western blots to study the cross-reactive determinant (CRD), samples were not boiled to avoid membrane protein aggregation in the membrane fractions. Protein gel electrophoresis was carried out as described above.

### 2.9.2 Transfer & blotting

Protein was transferred to a nitrocellulose membrane (Amersham Hybond ECL, GE Healthcare Life Sciences) that was equilibrated in transfer buffer (Appendix D). Transfer was achieved using a XCell II™ blot module (Invitrogen) in an Xcell SureLock® Mini tank (Invitrogen). The blot was kept at 30 V for ~2 hours. Once protein was transferred, the gel was discarded and blot was placed in a 50 mL falcon tube and blocked overnight at 4°C on a roller mixer containing 5 mL western blot blocking solution (Appendix D). The blot was then incubated with a primary antibody ( $\alpha$ -VSG221, a gift from the Mottram group) for 1 hour at room temperature. This was followed by three 10-minute washes in PBS-T and another hour incubation containing the secondary antibody conjugated to HRP. The blot was then washed three times in PBS-T and treated with the developing reagent (Thermo Fischer Scientific). Chemiluminescence was detected using a myECL™ Imager system (Thermo Fischer Scientific), with an exposure time between 1 and 15 minutes depending on the target. Typically,  $\alpha$ -enolase was used for the loading control as previously described (Hannaert *et al.*, 2003).

## 2.10 Microscopy

---

### 2.10.1 Mitotracker staining

For some experiments, mitochondria were stained using Mitotracker Red (Invitrogen) prior to fixation and mounting. To achieve this, cells (typically 1 mL at  $\sim 5 \times 10^5$  cells/mL) were incubated for 5 minutes at 37°C, 5% CO<sub>2</sub>, with a final concentration of 100 nM Mitotracker. Upon completion,

cells were centrifuged for 5 minutes at  $1,500 \times g$  and washed twice in fresh medium before cells were fixed as described below.

### 2.10.2 Preparation and fixation of cells

Trypanosomes were grown to mid-log phase before control and treatment cultures were started at a density of  $2 \times 10^5$  cells/mL. For each time-point, at least 3 mL culture was transferred to a falcon tube, centrifuged at  $1,500 \times g$  and resuspended in 1 mL fresh, warm, HMI-9. Subsequently, cells were spun down again, washed twice with sterile  $1 \times$  PBS, and finally resuspended in 500  $\mu$ L PBS.

Samples were fixed by adding a final concentration of 2% formaldehyde, and incubated for 15 minutes at room temperature. Next, another 500  $\mu$ L PBS was added to dilute fixative, and samples were washed once with sterile PBS. Cells were then resuspended in 50  $\mu$ L PBS and transferred onto a poly-L-lysine-coated slide, which was left to air-dry in a safety cabinet.

Dried slides were rehydrated and washed in PBS, and a counterstain consisting of  $1 \times$  PBS with 3  $\mu$ M 4,6-diamidino-2-phenylindole (DAPI) was applied to the slide, before mounting with a coverslip that was sealed with clear nail varnish.

Slides were kept at  $4^\circ\text{C}$  for up to a week until analysis was carried out using a Zeiss axioscope (Scope.A1, Zeiss).

### 2.10.3 Cell cycle and morphology analyses

To ascertain whether compounds affected the *T. brucei* cell cycle, cells were prepared for microscopy analysis as described above. Cells were then counterstained with DAPI and classified according to the numbers of nuclei and kinetoplasts they had, as a direct correlation to phases of the cell cycle, as described in several publications (Woodward & Gull, 1990, McKean, 2003, Siegel *et al.*, 2008). Cells in G1 phase have one nucleus and one kinetoplast (1N1K). Kinetoplast replication (S-phase) then initiates (1N2K) prior to nuclear division (2N2K). Finally, completion of the cell cycle leads to two cells in G1 phase.

To analyse any potential changes in cell cycle, >300 cells were counted in multiple samples and the number of cells in different cell cycle stages, as well as those in 2N1K and MNMK ('M' defined as 'multiple' organelles) phase, were calculated as percentages of the total number of counted cells.

To measure the distance between nucleus and kinetoplast, images were obtained from DAPI stained samples and imported into the Fiji software (Schindelin *et al.*, 2012). Distances were measured after the scale was set using the "measure" tool. For each sample group, 30 measurements were taken from three independent microscopy experiments.

#### 2.10.4 Analysis of kinetoplast localisation

To measure the mean distance between nucleus and kinetoplast, DAPI-stained slides were prepared for the relevant cell lines, as described above. For each sample, images were under a DAPI filter and DC, at 10× magnification. These were imported into the Fiji software (Schindelin *et al.*, 2012). Distances between nucleus and kinetoplast were measured in ~300 cells per sample, using the Analyse → Measure function. Mean and error was then calculated in Microsoft Excel.

### 2.11 Chemicals

---

This study used several experimental compounds. The benzoxaborole AN5568, previous identifier SCYX-7158, was supplied by Anacor Pharmaceuticals Inc., and was stored at -20°C in DMSO at a stock concentration of 100 mM. Ornidyl was supplied by Sanofi Aventis. Unless described specifically, all powder reagents were obtained from Sigma-Aldrich.

### 2.12 Computational methods

---

#### 2.12.1 Data analysis and presentation

Graphical representation of data was created using the Graphpad Prism software (v6.0, GraphPad Software, [www.graphpad.com](http://www.graphpad.com)). Statistical analyses were carried out using Graphpad Prism, SPSS and Microsoft Excel. Further computational analysis was carried out using R (R Core Team, 2013).

Analysis of microscopy images was carried out using the Fiji software (Schindelin *et al.*, 2012), an ImageJ package containing all available extensions. This software was also used for further processing of protein gels and blots.

Metabolite identification was done using Ideom (Scheltema *et al.*, 2011, Creek *et al.*, 2012b). Where targeted metabolomics using stable isotopes was carried out, mzMatch-ISO (Chokkathukalam *et al.*, 2013) was used for metabolite identification, and all hits were then confirmed using Ideom. Putatively identified metabolites were also analysed using the Metacyc (Caspi *et al.*, 2006, Caspi *et al.*, 2015), Kegg (Okuda *et al.*, 2008), and HMDB (Wishart *et al.*, 2007) databases. Further metabolomics figures, including heat maps and PCA plots, were either created using Metaboanalyst (Xia *et al.*, 2015) Microsoft Excel, R or GraphPad Prism.

Genomic and transcriptomic analysis was visualised with IGV (Robinson *et al.*, 2011). For sequence alignments, the CLC Genomics Workbench software v7.0 (QIAGEN) was used.

## Chapter 3. The mode of action of AN5568

---

### 3.1 Introduction

---

#### 3.1.1 AN5568 brings hope for HAT

AN5568 was identified as an effective inhibitor of *T. brucei* infection from a library screen involving benzoxaborole 6-carboxamides. The compound is based on another benzoxaborole called SCYX-6759, which exhibited only partial inhibition of stage 2 trypanosomiasis (Nare *et al.*, 2010). Both *in vitro* as well as *in vivo* assays have shown AN5568 to be highly potent against *T.b. gambiense* and *T.b. rhodesiense* (Wring *et al.*, 2010, Jacobs *et al.*, 2011, Giannotti *et al.*, 2014). Importantly, pharmacokinetic characterisation of the drug showed that brain exposure was high, with a  $C_{max}$  of  $>10 \mu\text{g/mL}$  and  $\text{AUC}_{0-24 \text{ hr}} >100 \mu\text{g}\cdot\text{h/mL}$  (Jacobs *et al.*, 2011). Further *in vitro* analysis showed that AN5568 exhibited high CNS permeability in the MDCK-MDR1 CNS transport assay ( $414.8 \text{ nm/s}$ ) (Wring *et al.*, 2014), an important consideration with regards to development of drugs for stage 2 HAT. In addition, CNS concentrations of the compound were maintained above the MIC for 20 hours, meaning the benzoxaborole could be administered orally once or twice daily (Wring *et al.*, 2014). This is consistent with *in vivo* studies, which have shown 100% cure rates only after 3 or more days of dosing (Jacobs *et al.*, 2011, Wring *et al.*, 2014).

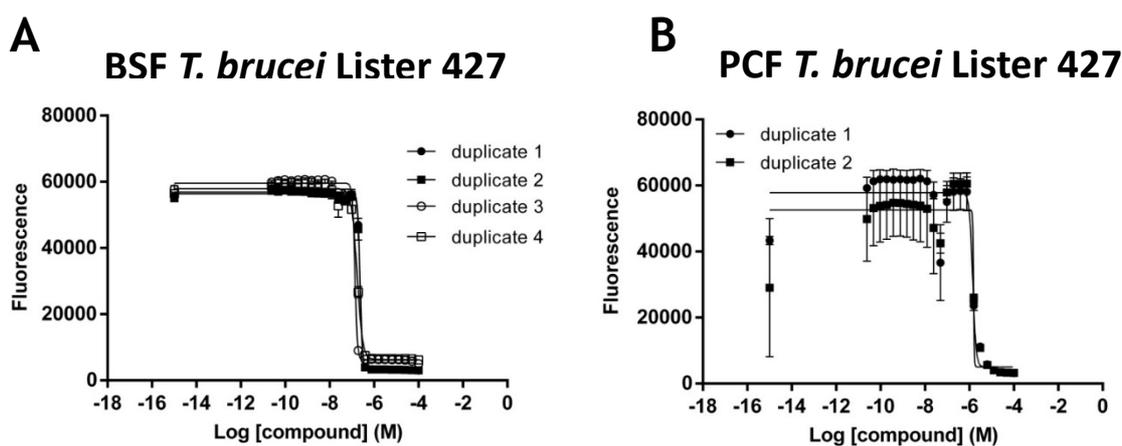
Whilst the pharmacokinetics of AN5568 are well understood, little work has been carried out to understand its mode of action, which is currently not known. This understanding is crucial, as it will give novel insights into trypanosome biology, as well as indicating any potential structure alterations that could lead to higher drug efficacy as well as reduce the possibility of drug resistance arising. Previous work in the Barrett group has suggested that AN5568 causes several major and highly specific perturbations in the metabolome of *T. brucei* cells *in vitro* (D. Creek, unpublished). In particular, significant increases were seen in S-adenosyl-L-methionine, an important methyl-group donor in methyltransferase reactions. Its breakdown product, 5'-methylthioadenosine was also increased in drug-treated cells. These results led to the hypothesis that AN5568 might interfere with methyltransferase reactions.

In this chapter we investigated the putative mode of action of AN5568. Firstly, we repeated the preliminary metabolomics analyses resulting in similar findings to those previously seen. We further probed this data set, with an emphasis on understanding the underlying determinants leading to changes in methionine metabolism. Initial characterisation of drug-treated cells *in vitro* showed signs of cell cycle inhibition, and through utilising a combination of omics-based approaches we

were able to show how the drug impacts various metabolic pathways. In addition, we performed an *in silico* study of the trypanosome “methyltransferome” to gain a better understanding of the methyltransferase complement of the parasite.

### 3.2 *In vitro* activity of AN5568 on *T. brucei*

To investigate AN5568 efficacy *in vitro*, an alamar blue assay was carried out to determine the  $EC_{50}$  (the concentration necessary to kill 50% of cells) for AN5568 in Lister 427 *T. brucei*. The  $EC_{50}$  for BSF cells was  $193 \pm 48$  nM. In contrast, the  $EC_{50}$  of AN5568 in PCF cells was  $1,500 \pm 90$  nM (fig. 3-1). This suggests the compound exhibits a far higher efficacy against BSF trypanosomes. Whilst this could mean the benzoxaborole targets a protein or pathway that is more active in BSF cells, it may also reflect slower uptake and metabolism, both of which are decreased in PCF cells compared to BSF. Previous work showed that a daily dose of 5 mg/kg/day for 4 days was sufficient for a 100% cure-rate *in vivo* (Jacobs *et al.*, 2011).



**Figure 3-1:** Sigmoidal dose-response curves carried out on AN5568-treated *T. brucei*. Alamar blue plates were set up as described in the methods section, and the  $EC_{50}$  of AN5568 was calculated in both BSF and PCF parasites *in vitro*. Whilst procyclic cells have an  $EC_{50}$  of 1,500 nM, BSFs were found to be almost ten-fold more sensitive, with an  $EC_{50}$  of 190 nM.

The drug was also tested on the other trypanosomatids (table 3-1). Alamar blue assays were carried out on *Leishmania mexicana* promastigotes, a kinetoplastid parasite closely related to *T. brucei*, but kept in culture in its vector life cycle stage. Here, we found the  $EC_{50}$  to be far higher, at  $37.8 \pm 3.7$   $\mu$ M (table 3-1). Further work was done on this parasite in section 3.6.

The  $EC_{50}$  concentrations are important to know because further *in vitro* studies would be based on the use of consistent concentrations of the benzoxaborole. The next step was to analyse the impact of drug treatment on cellular growth. Cultures were set up in triplicate for both BSF and PCF Lister 427 cell lines. For both cell lines, incubations with AN5568 were set up at concentrations of  $\frac{1}{2} \times EC_{50}$ ,  $1 \times EC_{50}$ ,  $2 \times EC_{50}$ ,  $5 \times EC_{50}$  and  $10 \times EC_{50}$ , along with a control containing DMSO only. Cells were cultured

at a density of  $2 \times 10^5$  cells/mL, so that a 10-fold decrease in density could be seen using a cell counting chamber. Cell growth was assessed by haemocytometer for up to 30 hours.

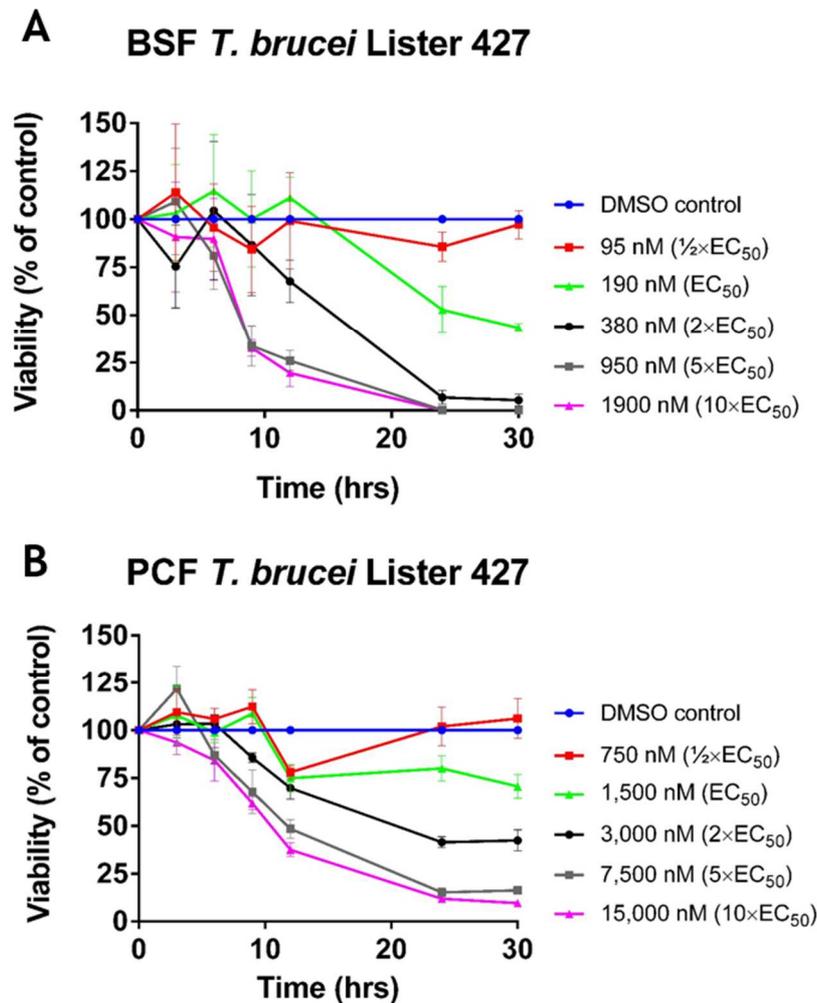
Organism	Approximate EC <sub>50</sub>	Origin of data
<i>T. brucei</i> – BSF	190 nM	This study
<i>T. brucei</i> – PCF	1,500 nM	This study
<i>T. congolense</i> – BSF	500 nM	F. Giordani
<i>T. cruzi</i>	7.5 $\mu$ M	R. Jacobs (Anacor)
<i>L. mexicana</i> - promastigote	38 $\mu$ M	This study

**Table 3-1: EC<sub>50</sub> of AN5568 in various kinetoplastids.** Where the value was not obtained by this study, it was instead given by the named individual in a personal communication.

BSF trypanosomes were particularly susceptible to higher concentrations of AN5568, with cell growth halted by 12 hours in concentrations of  $2 \times EC_{50}$  and above (fig. 3-2A). Indeed, cells incubated with 950 nM AN5568, or higher, show growth phenotypes by the 6-hr time-point. In addition, there were no viable cells observed at 24 hours when cells were treated with  $\geq 5 \times EC_{50}$ . This suggests that BSF cells in these concentrations of the benzoxaborole are killed before 24 hours.

In contrast, PCF cells were less affected by benzoxaborole exposure. Despite a noticeable decrease in growth in high concentrations, live cells were observed after 30 hours exposure to  $10 \times EC_{50}$  (fig. 3-2B). It appeared that cell populations treated with  $\leq 5 \times EC_{50}$  recovered from AN5568 exposure after 24 hours.

These data indicate that AN5568 is trypanocidal in BSF cells, but its effect on PCF cells is largely trypanostatic, even at high doses. PCFs are therefore less sensitive to the compound, which supports the alamar blue results.

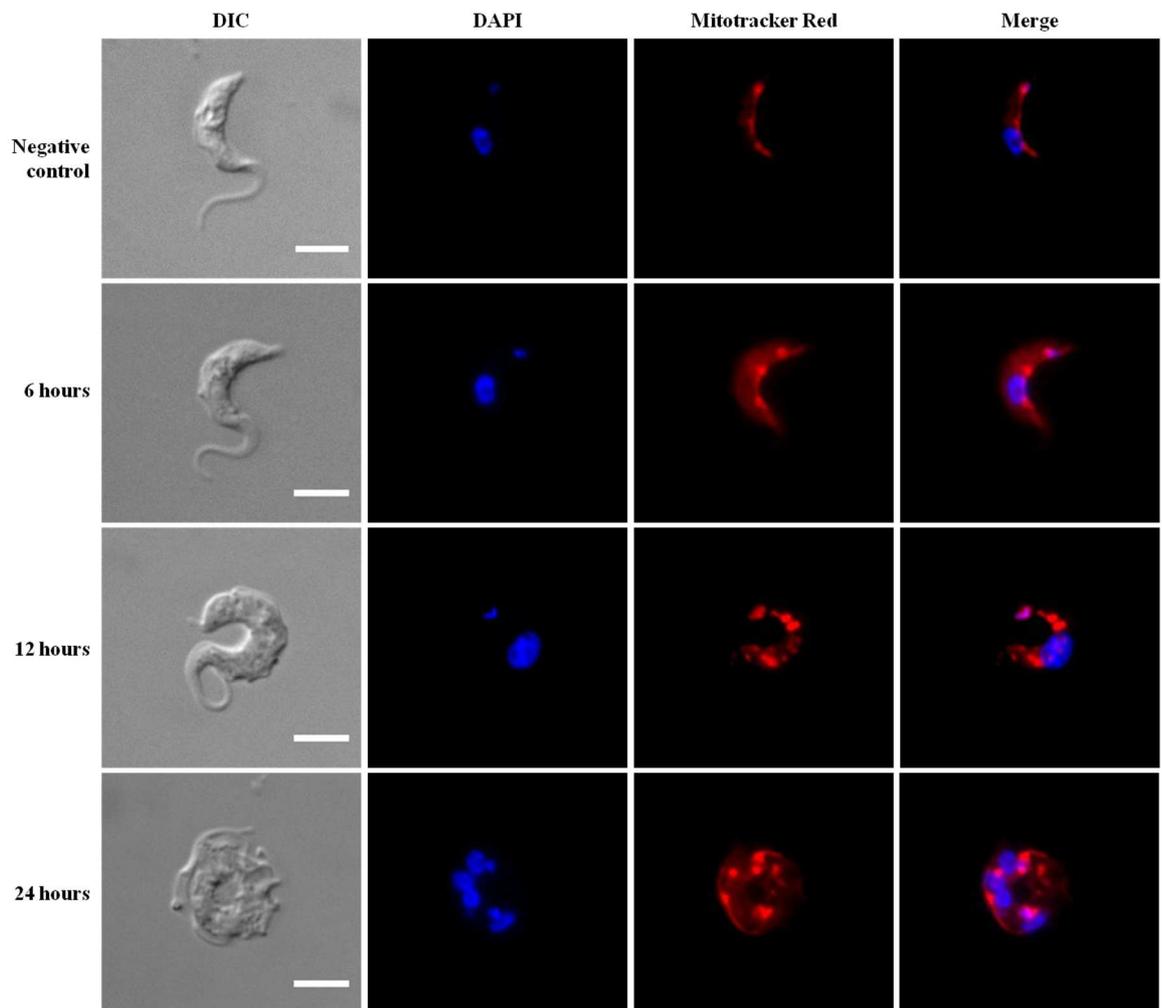


**Figure 3-2: Growth of BSF and PCF *T. brucei* during AN5568 treatment *in vitro*.** Growth curves were set up, and cultures observed over a period of 30 hours, to assess parasite growth in the presence of various concentrations of the benzoxaborole drug. Cell viability was calculated as a percentage of a DMSO control. In both cases, colours correspond to the following relative concentrations: red:  $\frac{1}{2} \times EC_{50}$ , green:  $1 \times EC_{50}$ , black:  $2 \times EC_{50}$ , grey:  $5 \times EC_{50}$ , purple:  $10 \times EC_{50}$ . A) Concentrations of  $5 \times EC_{50}$  and above abolished all growth within 24 hours in BSF *T. brucei*. Growth cultures were rendered static in concentrations above  $EC_{50}$ , whilst there was no difference in parasite growth when cells were treated with  $\frac{1}{2} \times EC_{50}$  concentration of AN5568. B) Whilst parasite growth was limited upon treatment with high concentrations of AN5568, live cells were still observed after 30 hours, even in cultures given the highest dose of the drug.

### 3.3 Morphological changes in AN5568-treated cells

To analyse morphological changes in BSF parasites after AN5568 treatment, fluorescent microscopy was employed. Cell cycle analysis was carried out as described in the methods, using DAPI staining as well as bright field microscopy. In addition, mitochondrial staining was carried out using a commercial antibody (Mitotracker Red, Invitrogen), to assess further changes to this organelle. The use of ER tracker (Life technologies) was also attempted, however, this was unsuccessful in fixed cells. Analysis was done with AN5568 concentrations of  $2 \times EC_{50}$  and  $10 \times EC_{50}$ , (380 nM & 1,900 nM respectively) and a negative control was supplemented with an equal volume of DMSO. Microscopy analyses were carried out at time points of 6 hours, 12 hours and 24 hours post-treatment.

At  $2\times EC_{50}$  concentration of AN5568, no significant changes in morphology were observed at the 6 hour time-point (fig 3-3). Cells were comparable to the control under visualisation of both DAPI and mito-tracker. However, by 12 hours, many cells under this dosage exhibited a rounded shape. Under DIC, these cells appeared to have a compromised plasma membrane. Whilst DAPI showed the location and morphology of both nuclei and kinetoplasts to be normal, the mitochondrial shape differed to the wild-type mitochondria. These phenotypes were more extreme at 24 hours post-treatment (fig. 3-3). In addition, many cells exhibited multiple nuclei, but not multiple kinetoplasts, at this time-point. A similar phenomenon has been reported after depletion of mitochondrial acyl carrier protein (Clayton *et al.*, 2011). However, in this case, there were a variety of defects in the kinetoplast, including organelle enlargement, which was not observed here. In addition, RNAi of a mitochondrial DNA helicase, TbPIF1, resulted in the complete loss of kinetoplast DNA (Liu *et al.*, 2010). Again, this was not a phenomenon that occurred after AN5568-treatment. Cells with multiple flagella were also seen, albeit low in number. DIC microscopy showed disintegration of the plasma membrane, although we could not confirm whether this was as a direct result of the benzoxaborole action, or rather, a general cell death phenotype. There were many cells showing abnormal mitochondria after 24 hours (fig. 3-3). Rather than the linear mitochondrion seen in WT cells, the cells here exhibited patches of high intensity stain, which could suggest breaking up, or leakage, of the mitochondrial matter (fig. 3-3).

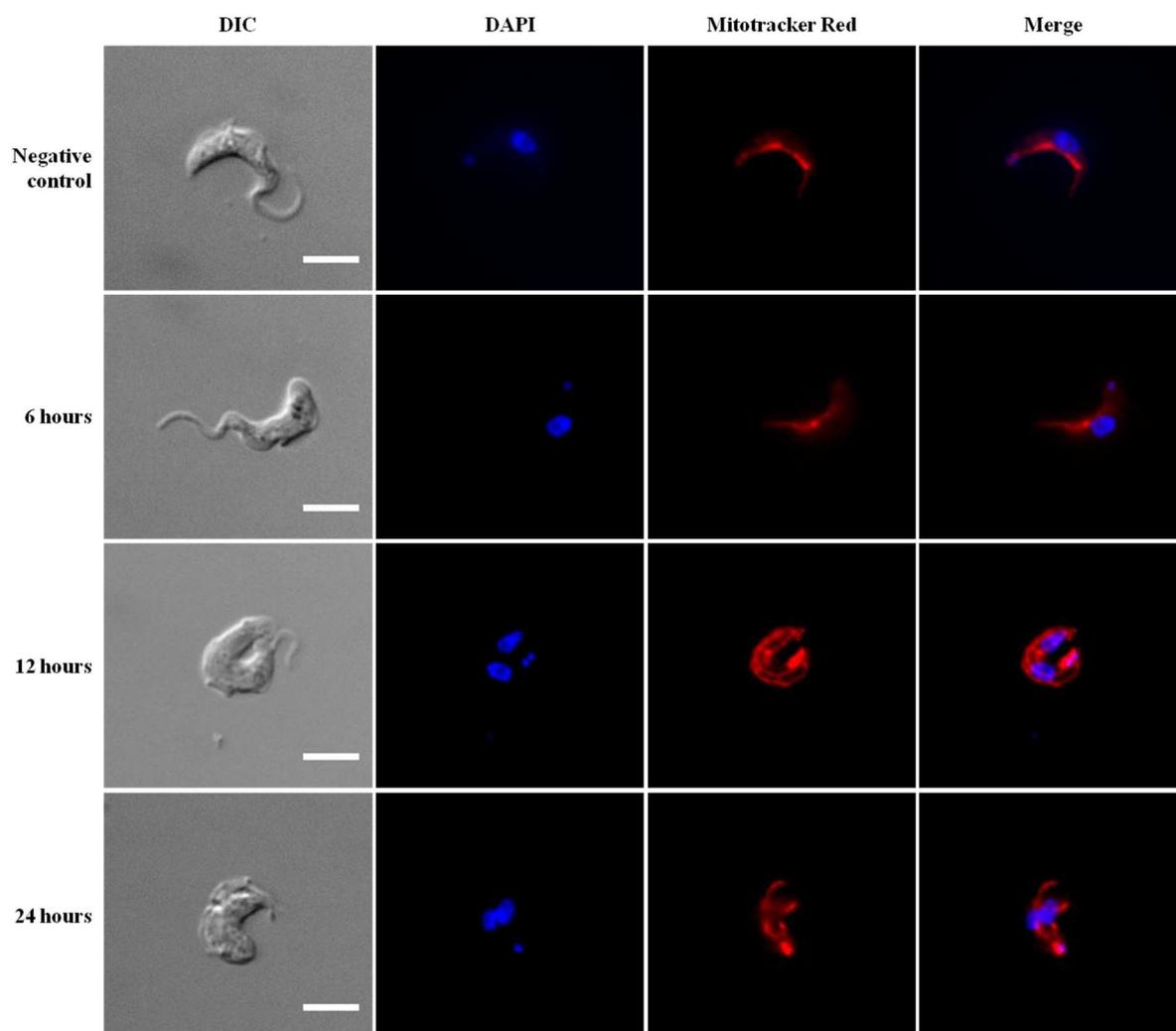


**Figure 3-3: The effect of  $2\times EC_{50}$  AN5568 on cellular morphology.** *T. brucei* cells were incubated with  $2\times EC_{50}$  concentration of AN5568 and cells were isolated for analysis by microscopy at specific time-points. Under normal conditions, cells show 1 nucleus and 1 kinetoplast, or other conformations depending on the stage of cellular division. In addition, the mitochondrion appears as a thin line along the long axis of the cell when stained by Mitotracker Red. After 6 hours of drug exposure, both DIC and DAPI staining showed the cells to appear normal. However, the mitochondrion appeared to be swelling at this time-point. At the 12-hour time-point, cells were observed to be swelling under direct light. This is a common morphological phenotype in response to trypanocidal compounds. In addition, whilst the nuclei and kinetoplast still appeared to be mostly normal in both number and morphology, the mitochondria were showing abnormal phenotypes. In particular, the organelles seemed to be breaking up and become ‘patchy’. This phenotype was also apparent 24 hours after AN5568 exposure. Here, cells were also showing significant rounding and compromised membranes. Moreover, DAPI staining indicated a build-up of nuclei, but not kinetoplasts, in a significant number of cells, suggesting a cellular division defect. Scale bar represents 5  $\mu\text{m}$ .

Work was also undertaken to analyse the morphological changes to the parasite when treated with  $10\times EC_{50}$  ( $\sim 2\ \mu\text{M}$ ). Cells were analysed at the same time-points as mentioned above. DIC analysis showed that the cells exhibited the same rounding of the cell body after 12 hours, as seen after treatment with  $2\times EC_{50}$  for the same amount of time (fig. 3-4). By 24 hours, it was challenging to find any live cells resembling BSF *T. brucei*. Much of the culture was reduced to numerous flagella and cellular debris, making this time-point challenging to analyse at this concentration of benzoxaborole. DAPI staining did not reveal the same phenotypes as those seen at  $2\times EC_{50}$ . Live cells contained either one or two nuclei, as well as one or two kinetoplasts, suggesting that at this concentration, cells were killed before a significant cell division phenotype was visible. Mitochondrial morphology appears to undergo similar changes to those seen at lower drug

concentrations (fig. 3-4). By 12 hours there appears to more mitochondrial leakage, although this could also occur as a result of organelle replication before cytokinesis. After 24 hours of drug treatment, the mitochondria appeared 'patchy' as seen in figure 3-3 as well.

In addition to visual analysis of drug-treated cells, the nuclei and kinetoplasts in untreated and treated cells were quantified. This type of analysis, termed 'NK analysis', allows the assessment of changes in the cell division cycle (Woodward & Gull, 1990). For each sample, up to 300 cells were counted in triplicate. Cells with 1 nucleus and 1 kinetoplast (1N1K) were deemed not to be undergoing cell division. Cells with one nucleus but 2 kinetoplasts (1N2K) had replicated kinetoplasts, but not nuclei. Cells containing two nuclei as well as two kinetoplasts (2N2K) were towards the end of the DNA replication cycle and were undergoing cytokinesis. In some cases two flagella could be seen at the 2N2K stage. We also noted cells that contained a conformation of either 2N1K, or multiple DNA-containing organelles (MNMK), which suggests a defect in cytokinesis or DNA replication (Woodward & Gull, 1990).

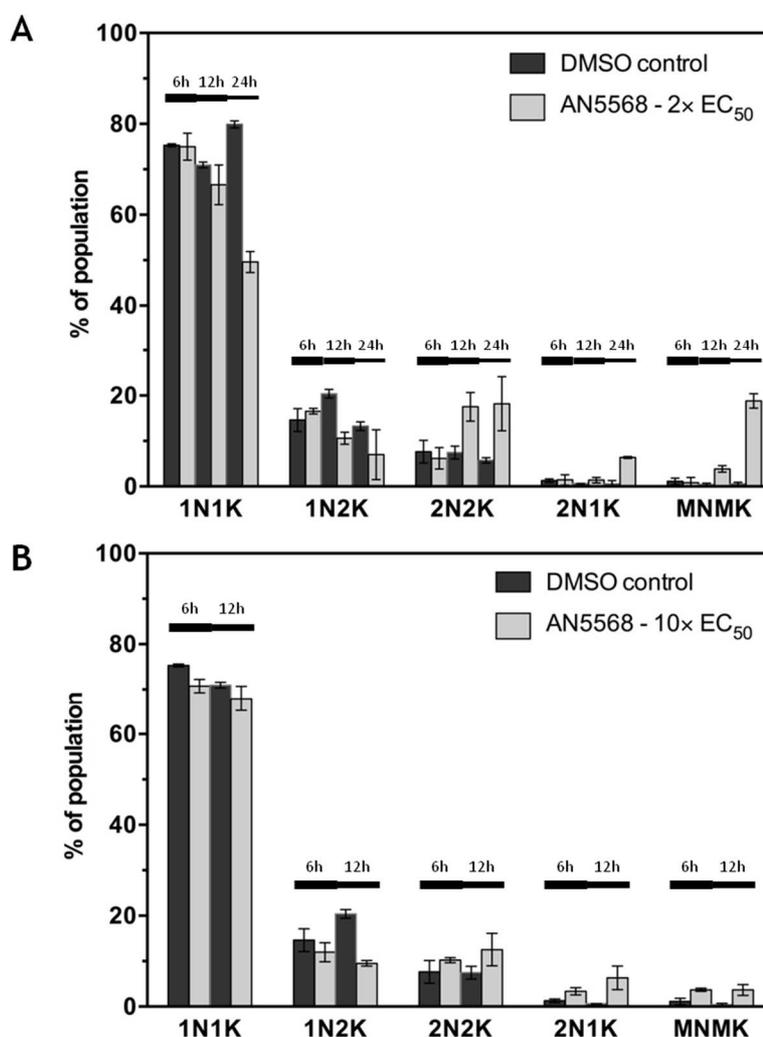


**Figure 3-4: Effect of 10x EC<sub>50</sub> AN5568 treatment on *T. brucei* morphology.** Microscopy analysis was also carried out on cells treated with 10x EC<sub>50</sub> AN5568, at similar time-points to those observed with 2x EC<sub>50</sub>. Similarly, cells were observed to show wild-type morphology after 6 hours of drug exposure. DAPI and Mitotracker staining did not raise any phenotypes of interest. By 12 hours, cells were very rounded and showing staining phenotypes similar to those observed after 24 hours of incubation with 2x EC<sub>50</sub>. In the representative image shown, cells did show normal replication of nuclei, kinetoplasts and mitochondria, suggesting the cell division defects were not as severe. Analysis of the 24 hour time-point proved challenging by the lack of cells left in the sample. When they were seen, there was significant swelling and rounding of the cell body. Whilst DAPI staining still did not show severe phenotypes, the mitochondrial stain exhibited leakage throughout the cell, indicating that the mitochondrial membrane had been compromised, most likely as a result of cell death. Scale bar represents 5  $\mu$ m.

Treatment with benzoxaborole does not cause any visible changes until at least 12-hours post-treatment (fig. 3-3 & 3-4). Similarly, changes in NK distribution were also observed after 12 hours, when treated with 2x EC<sub>50</sub> AN5568 (fig. 3-5A). This phenotype was more apparent after 24-hours, where an increase was also observed in MNMK cells. Indeed, up to 20% of cells exhibited multiple nuclei and/or kinetoplasts at this time-point.

When treated with 10x EC<sub>50</sub>, a similar increase in 2N1K cells was observed by 12 hours. There was also a 10% reduction in 1N2K cells at this time-point. Unfortunately, there were so few cells left by 24 hours that NK analyses were not possible.

The microscopy analysis suggests that AN5568 has a significant impact on the parasite's ability to divide. In particular, there are indications that cytokinesis is impaired, leading to cells containing multiple nuclei as seen in figures 3-3 and 3-4. Interestingly, increases in 2N1K, and MNMK cells, suggests that whilst replication of the nucleus occurs, there could be further complications in kinetoplast division as a direct result of drug-treatment.



**Figure 3-5: Cell cycle analysis of AN5568-treated cells.** For both microscopy experiments, DAPI staining was analysed in more detail to identify potential defects in cell division as a result of AN5568 treatment. A) Exposure to 2× EC<sub>50</sub> AN5568 caused a build-up of 2N2K cells by 12 hours. In addition, there were increases in 2N1K and MNMK (M represent multiple) cells by 24 hours, along with a corresponding decrease of cells in the 1N1K state, compared to WT cells. B) These phenotypes were less obvious in cells treated with 10× EC<sub>50</sub> concentration of the benzoxaborole. Whilst there were increases in 2N2K and 2N1K cells, the changes were not as high as those seen in cells treated with a lower dose of the drug. Unfortunately, not enough cells were seen after 24 hours of 10× EC<sub>50</sub> exposure, making cell cycle analysis at this time-point all but impossible.

### 3.4 Untargeted metabolomics analysis of AN5568-treated bloodstream form trypanosomes

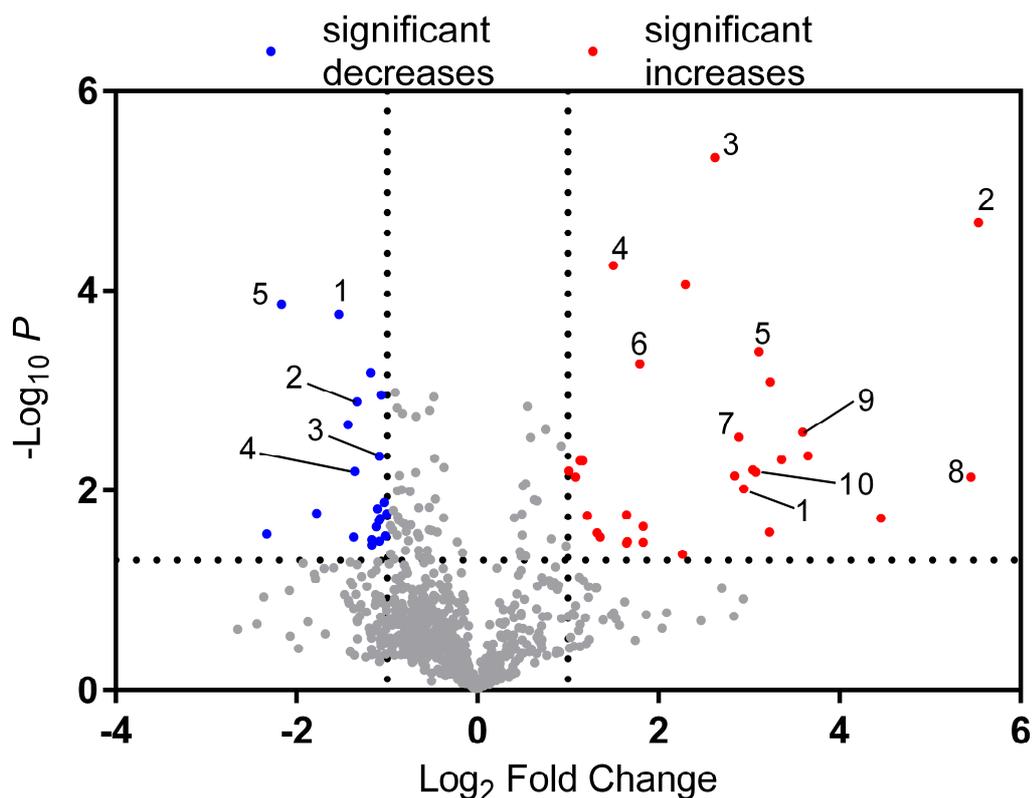
As described previously, there are several methods that can be utilised to investigate and validate drug targets. Recent advances in metabolomics allow us to observe all small molecule metabolites in a cell/system. This gives a distinct advantage over the use of genomics or transcriptomics because

the metabolome is much more dynamic, allowing the observation of significant perturbations very quickly (Veenstra, 2012). Given the advantages and potential of this method, we chose to apply it to the investigation of the mode of action of AN5568.

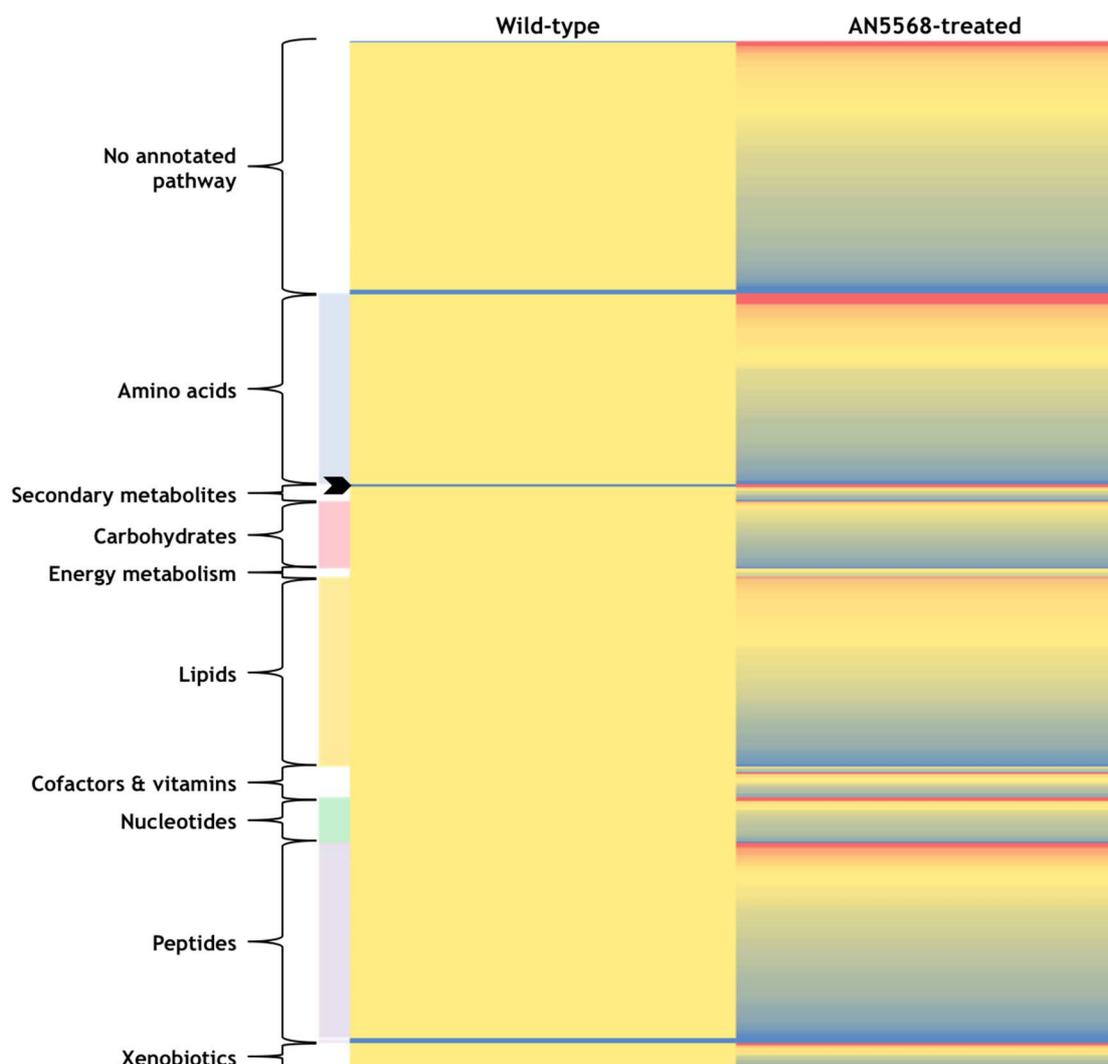
Growth analysis of drug-treated cells indicated that effects might be visible by 6 hours post-treatment, as cell growth is static by this point at a drug concentration of  $10\times EC_{50}$ , even if morphology remains rather unchanged. This is because the proteome is less dynamic than the metabolome, and changes in the metabolome will occur much faster upon drug treatment. Preliminary metabolomics analysis by the Barrett group indicated perturbations in S-adenosyl-L-methionine (AdoMet) metabolism (Darren Creek, personal communication). Therefore, mass spectrometry analysis was carried out on cells treated with AN5568 as well as a sample group treated with sinefungin, an AdoMet. Sinefungin is a non-specific inhibitor of MTase reactions, and would therefore show similarities with AN5568 should our hypotheses be correct. Both drug-treated samples were incubated for 6 hours with a drug concentration of  $10\times EC_{50}$ , alongside a DMSO control. Samples were prepared as described in the methods and analysed by liquid chromatography mass spectrometry (LC-MS). The resulting data was processed using the mzMatch/Ideom pipeline.

Processing of data using Ideom led to the identification of 840 putative metabolites. Of these, 50 were deemed to show significant changes ( $\text{Log}_2$  fold-change =  $<-1, >1, P<0.05$  [*t*-test]), as shown in figure 3-6. Amongst the significantly altered metabolites, we noticed an enrichment of the methionine degradation pathway (fig. 3-6-3-8). There were also several significant increases found in amino acid, as well as nucleotide metabolism, which were analysed in depth below.

In addition, several lipids were significantly decreased after drug-treatment (fig. 3-6). This could suggest either degradation of cell surface membrane, or an organelle membrane. This result was especially interesting in the context of possible mitochondrial damage, and is covered in detail in section 3.4.3. The majority of metabolites were, however, unchanged. This suggests that the benzoxaborole possesses a highly specific target, or mode of action, bypassing many metabolic pathways present in the parasite. In addition, it indicates a high level of robustness in the *T. brucei* metabolome. Components of glycolysis and carbohydrate metabolism in general, were no different than in wild-type samples.



**Figure 3-6: Volcano plot of metabolic changes in *T. brucei* after 6-hour treatment with 10× EC<sub>50</sub> AN5568.** Metabolite samples were prepared from drug-treated as well as DMSO control samples, as described in the methods. Analysis was carried out using the mzMatch/Ideom pipeline. A volcano plot was generated where all red dots indicate metabolites undergoing significant increases as assessed by a combination of a Student's t-test ( $P < 0.05$ ), and a fold change higher than 2× ( $\text{Log}_2(\text{fold-change}) = > 1$ ). Blue dots indicate significant decreases as calculated by the same t-test and a fold change lower than -2× ( $\text{Log}_2(\text{fold-change}) = < -1$ ). Numbers at red dots correspond to the following metabolites that were significantly increased: 1) S-adenosyl-L-methionine (m/z: 398.1374, RT: 18.02 mins, 7.67-fold), 2) 1,2-Dihydroxy-5-(methylthio)pent-1-en-3-one (m/z: 162.0350, RT: 17.24 mins, 46.33-fold), 3) 5'-methylthioadenosine (m/z: 297.0896, RT: 7.75 mins, 6.17-fold), 4) N6-acetyl-L-lysine (m/z: 188.1162, RT: 14.70 mins, 2.83-fold), 5) adenine (m/z: 135.0546, RT: 7.71 mins, 12.55-fold), 6) 4-hydroxy-4-methylglutamate (m/z: 177.0637, RT: 15.3 mins, 3.46-fold), 7) N6,N6,N6-trimethyl-L-lysine (m/z: 188.1524, RT: 23.54 mins, 7.4-fold), 8) cyclic ADP-ribose (m/z: 541.0608, RT: 16.90 mins, 43.78-fold), 9) aminoacetone (m/z: 73.0528, RT: 7.67 mins, 12.06-fold), 10) 8-amino-7-oxononanoate (m/z: 187.1208, RT: 13.75 mins, 8.43-fold). Numbers at blue dots correspond to the following metabolites that were significantly decreased: 1) [PC(14:0)] 1-tetradecanoyl-sn-glycero-3-phosphocholine (m/z: 467.3014, RT: 4.75 mins, 0.35-fold), 2) sn-glycerol 3-phosphate (m/z: 172.0136, RT: 16.44 mins, 0.4-fold), 3) D-glucosamine 6-phosphate (m/z: 259.0457, RT: 17.68 mins, 0.47-fold), 4) 2-deoxy-D-ribose 5-phosphate (m/z: 214.0242, RT: 16.45 mins, 0.39-fold), 5) Asp-Asp-Cys-Pro (peptide) (m/z: 448.1256, RT: 17.64 mins, 0.22-fold).

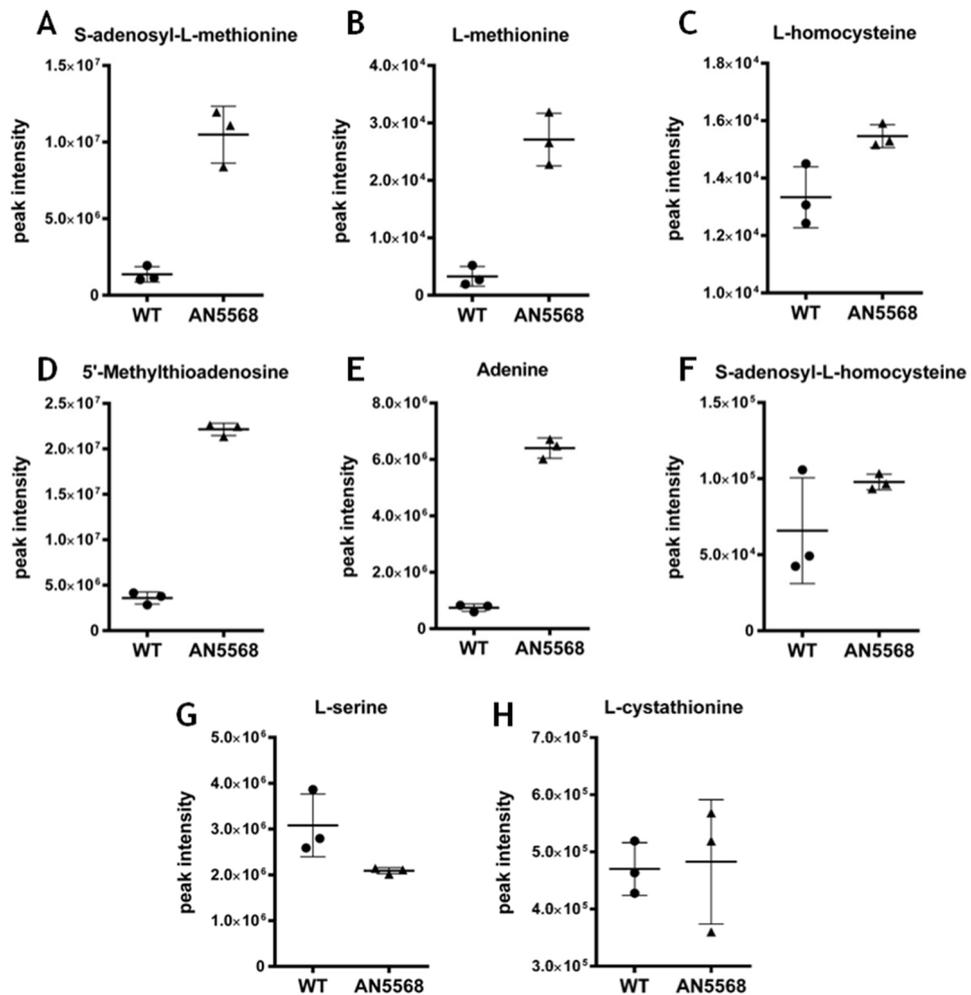


**Figure 3-7:** Heat map showing the changes in metabolism occurring in *T. brucei* as a result of AN5568-treatment. Black arrow: AN5568. Whilst both increases and decreases were shown to occur in the majority of areas of metabolism, there were few increases in lipid metabolism. In addition, 206 metabolites were purported not to be definitively involved in a particular metabolic pathway (shown as 'no annotated pathway'). Red colours indicate increased metabolite peak intensities in drug-treated cells. Blue indicates peak intensity decreases whilst yellow indicates no difference.

### 3.4.1 Methionine metabolism

After 6 hours, there were significant changes in methionine metabolism in response to AN5568 treatment (fig 3-7 & 3-8). The largest changes were seen in L-methionine ( $m/w$ : 149.0510, RT: 18.05 mins), S-adenosyl-L-methionine (AdoMet,  $m/z$ : 398.1374, RT: 18.01 mins) and 5'-methylthioadenosine (5'-MTA,  $m/z$ : 297.0896, RT: 7.75 mins). In addition, there was also a significant increase in adenine, a molecule produced during the hydrolysis of 5'-MTA. These molecules are all part of the methionine degradation pathway in *T. brucei*, which is thought to progress from L-methionine to 5'-methylthio-D-ribose (fig. 3-9). 5'-methylthio-D-ribose is thought to be excreted by the parasite, albeit at very small quantities (F. Achcar, personal communication), and was not detected in this experiment. Importantly, methionine is required for the biosynthesis of AdoMet, which plays a significant role in MTase reactions by supplying methyl groups in these reactions. In addition, decarboxylation of AdoMet results in the formation of decarboxylated

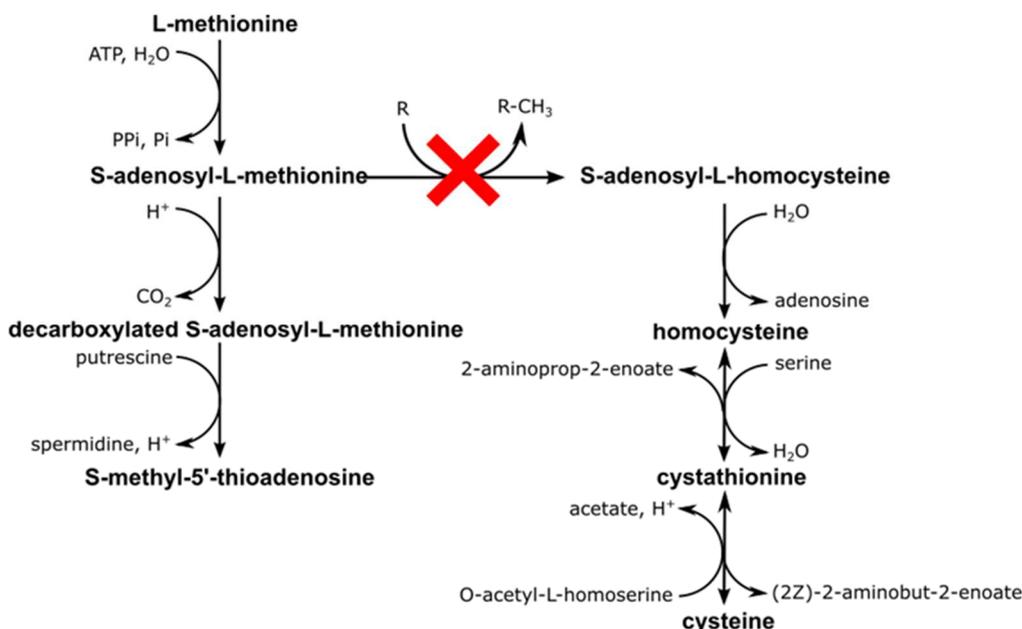
AdoMet (dcAdoMet), which feeds the polyamine pathway, a crucial and essential stress response pathway for the parasite (Olin-Sandoval *et al.*, 2012). Again, dcAdoMet was not detected by mass spectrometry. However, the product of the aminopropyl transfer reaction that consumes dcAdoMet is 5'-MTA, which was significantly increased in AN5568-treated cells, most likely due to the breakdown of accumulating AdoMet.



**Figure 3-8: Changes in methylation metabolism after AN5568 treatment.** A major pattern was observed with regards to the methionine degradation pathway. Several metabolites with significant changes were all found to be associated with methionine and methylation metabolism. AdoMet (A) had been seen to increase previously and this result was repeated here. L-methionine (B) 5'-MTA (D) and adenine (E) levels showed a similar increase, suggesting a build-up of these metabolites all belonging to the same pathway. Both AdoHcy (F) and Hcy (C) were found to increase, which at first, did not appear to support our MTase inhibition hypothesis. However, both the peaks for these metabolites were noisy and the wild-type samples showed significant variation. Finally, L-serine (G), utilised in the conversion of L-homocysteine to L-cystathionine, showed a decrease after AN5568-treatment, suggesting it was being used at a higher rate compared to wild-type cells. However, levels of L-cystathionine (H) were not found to be altered in AN5568-treated cells, although there was a high amount of variation in this peak, for the drug-treated samples.

Previous work investigating the mode of action of eflornithine also looked at the methionine degradation pathway (Vincent *et al.*, 2012). However, the metabolomics results of that study did not find similar metabolic alterations in AdoMet, even though 5'-MTA did show increases. Data from eflornithine studies led to the hypothesis that the benzoxaborole does not target the polyamine pathway directly. Indeed, most of the components in the trypanothione synthesis

pathway were actually increased after AN5568 treatment (fig. 3-14C). In contrast, many of these metabolites are decreased in abundance after eflornithine treatment (Vincent *et al.*, 2012).



**Figure 3-9: Metabolites involved in methyltransferase reaction pathways.** The uptake of L-methionine is crucial for *T. brucei*, as its breakdown product, 5'-MTA, is involved in the polyamine synthesis pathway. In addition, L-methionine is converted to AdoMet, which is an important methyl group donor in MTase reactions. The results generated by the metabolite extraction of AN5568-treated cells indicated that a specific MTase reaction could be inhibited (red cross). However, the particular reaction might not be observed in metabolomics datasets, should the methylation occur on proteins or nucleic acids. In addition, it is unknown what effect this could have on the downstream metabolites such as AdoHcy, Hcy, L-cystathionine and L-cysteine.

Instead, given the high levels of AdoMet, one possible target could be an MTase reaction. Methylation and demethylation of DNA, RNA and proteins is an important regulatory modification across all organisms (Fontecave *et al.*, 2004). AdoMet is crucial in these reactions, as it provides a methyl group that is transferred to the target substrate. In fact, there are very few instances where AdoMet is not the donor for an MTase reaction (Coulter *et al.*, 1993, Ragsdale, 2008). MTases are also highly specific, meaning the inhibition can occur of one MTase only.

Interestingly, we did not see many abundance changes in the metabolites found after the methyltransferase reaction (fig. 3-8). AdoMet is converted to AdoHcy, which did not show a corresponding drop in peak intensity. Similarly, both L-homocysteine ( $m/z$ : 135.0354, RT: 19.12 mins) and L-cystathionine ( $m/z$ : 222.0674, RT: 19.12 mins) did not decrease after AN5568 treatment, with the former slightly increased compared to the WT control.

### 3.4.2 Effect of AN5568-treatment on VSG biosynthesis

Further analysis of the metabolomics dataset also showed increased peak intensities for several components of the glycoprotein synthesis pathway upon benzoxaborole treatment (fig. 3-10). Glycoprotein synthesis in *T. brucei* is crucial to the parasite, which requires the GPI-anchored

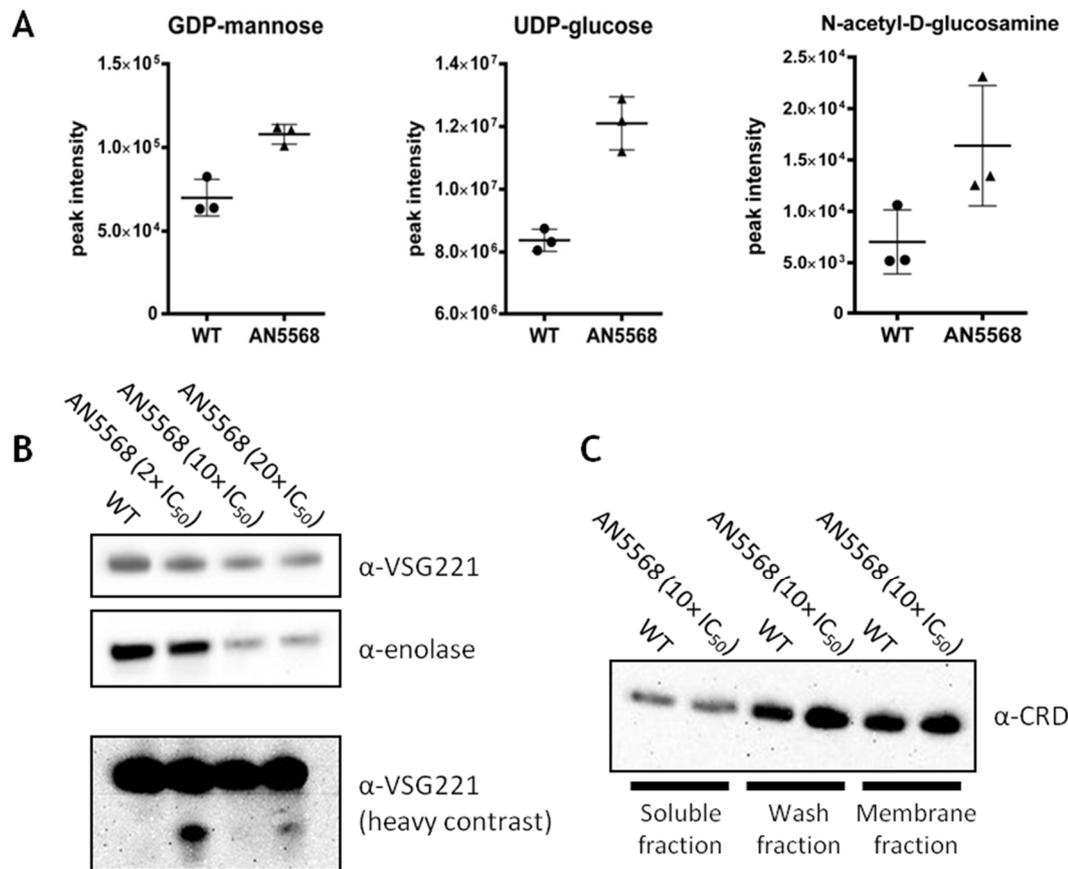
glycoproteins for the expression of VSGs on its cell surface, in addition to a wide variety of other cell surface proteins (Crossman *et al.*, 2002). Indeed, several recent studies have highlighted that inhibition of glycoprotein synthesis can lead to a general arrest in translation initiation as well as a spliced leader silencing response pathway, similar to the mammalian unfolding protein response (Smith *et al.*, 2009, Goldshmidt *et al.*, 2010)

Upon exposure to AN5568, we found increased abundance of 3 important metabolites involved in glycoprotein biosynthesis (fig. 3-10A). GDP-mannose ( $m/w$ : 605.0774, RT: 20.59 mins) is formed by a side-branch of glycolysis from mannose 6-phosphate (Smith *et al.*, 1996). The mannose molecules are then bound to N-acetyl-glucosamine (GlcNAc,  $m/w$ : 221.0899, RT: 15.29 mins), with the loss of GDP in the reaction (Kuettel *et al.*, 2012). Both of these metabolites were increased after AN5568 treatment, which led us to believe there might be changes in glycoprotein synthesis. To further confirm changes in glycoprotein synthesis, we also found a putative UDP-glucose ( $m/z$ : 566.0551, RT: 18.71 mins) to be increased in drug-treated cells (fig. 3-10A). Whilst this metabolite was predicted to be UDP-glucose, it carries the same mass as UDP-galactose, and these two metabolites are impossible to differentiate using LC-MS. UDP glucose is utilised in *T. brucei* by a UDP-glucose:glycoprotein glucosyltransferase, in a reaction that helps to protect the parasite against stress (Izquierdo *et al.*, 2009). UDP-glucose is also converted to UDP-galactose, which is used for the galactosylation of glycoproteins (Urbaniak *et al.*, 2006). In any case, increased intensity of this peak suggests the parasite is undergoing significant perturbations in glycoprotein biosynthesis.

To determine whether VSG biosynthesis was disrupted by AN5568 treatment, several western blots were carried out to assess their expression over a 6-hour period in various concentrations of the drug (fig. 3-10B & 3-10C). The Lister 427 laboratory strain of *T. b. brucei* is known to express VSG221, with a very low VSG switching rate (Lamont *et al.*, 1986). Interestingly, VSG expression was not found to change over a period of 6 hours post-treatment, even at  $20\times$  EC<sub>50</sub> concentrations of the benzoxaborole (fig. 3-10B). By altering contrast significantly, an extra band was seen in samples treated with  $2\times$  EC<sub>50</sub> AN5568, but this was dismissed as background noise, or non-specific binding (fig. 3-10B). One interesting finding was that the loading control used for VSG blots,  $\alpha$ -enolase, seemed to be less abundant in cells treated with high concentrations of AN5568 (fig. 3-10B). Indeed, when the metabolome data was analysed to specifically look for 2-phospho-D-glycerate and phosphoenolpyruvate, the substrate and product of enolase respectively, both were found to be decreased in AN5568-treated cells (Supplementary [S]-1).

In addition, further blots were analysed for the cross-reactive determinant (CRD) epitope. This marker corresponds to the glycosyl-phosphatidylinositol (GPI) membrane protein that anchors VSGs to the cell membrane (Shak *et al.*, 1988). In this experiment, *T. brucei* cells were fractionated to look for differences between the soluble fraction and the membrane fraction, which normally

contains the highest levels of CRD. We hypothesised that, should GPI biosynthesis be inhibited, we would see a less intense band for CRD in the membrane fraction. However, we found this not to be the case (fig. 3-10C). In all 3 fractions that were observed, there were no unexpected differences in CRD expression between the two sample groups. In both groups there was weaker expression in the membrane fraction. These results suggest that neither VSG expression, nor glycoprotein biosynthesis are affected by AN5568. However, it would be worthwhile to analyse protein lysates from cells incubated for longer time-points as expression might change more significantly over time.



**Figure 3-10: Metabolites involved in glycoprotein synthesis are upregulated after AN5568 treatment, but VSG expression is not impaired.** A) Three metabolites, all known to be involved in glycoprotein biosynthesis, were found to be upregulated in AN5568-treated cells. The majority of glycoproteins expressed on the *T. brucei* cell membrane are glycosyl-phosphatidylinositol moieties which anchor the VSG, leading us to hypothesise that VSG expression could be impaired post-treatment. B) Western blots carried out on protein lysates taken from AN5568-treated cells at a 6-hour time-point showed no significant changes in VSG expression at this time-point. When the contrast of the resulting blot was increased, an extra band was seen in the 2× EC<sub>50</sub>-treated sample. However, this was dismissed as background noise, as this band was not seen in the other two drug-treated conditions. C) In addition to VSG expression, blots were probed for the cross-reacting determinant of the GPI anchor. For this experiment, protein lysates were fractionated to isolate the membrane fraction and separate it from the soluble fraction. Most of the cellular GPI is found in the membrane, and no changes in expression could be seen in cells treated with 10× EC<sub>50</sub> for 6 hours.

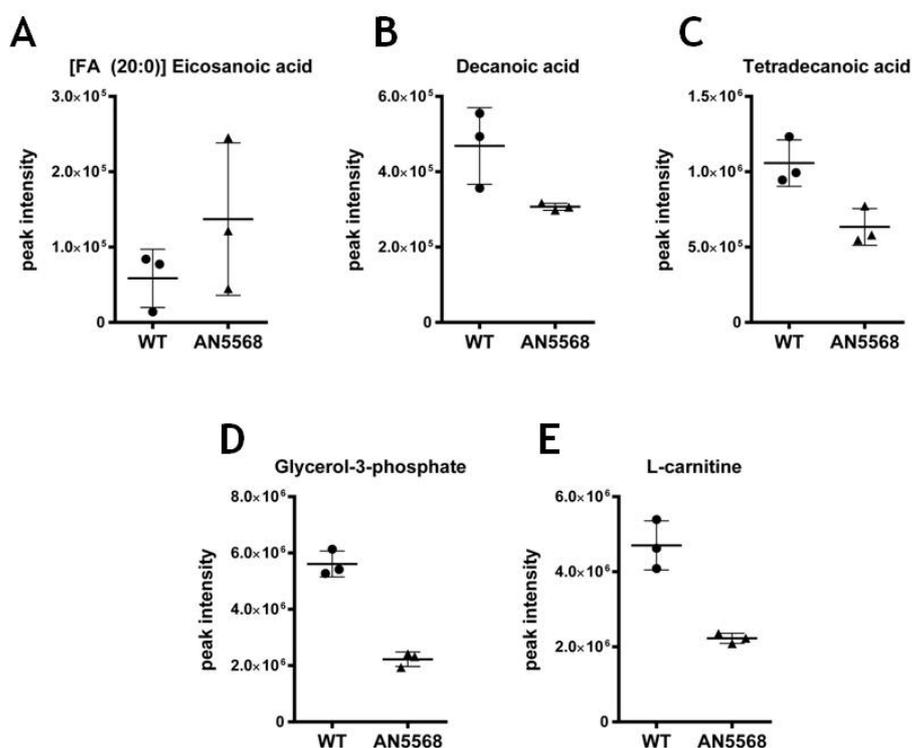
### 3.4.3 Lipid metabolism

The analysis of lipids and long-chain fatty acids using pHILIC columns are not optimal due to high amounts of background noise and poor mass peaks (Richmond *et al.*, 2010). In addition, accurate

identification of these molecules is difficult in comparison to the identification of nucleotides and amino acids, as they can undergo a wide variety of modifications. Nevertheless, attempts were made to find any patterns of change in these metabolites.

Firstly, we aimed to analyse possible changes in the mitochondria, given the microscopy analysis that had been carried out before. This organelle is an important compartment for the  $\beta$ -oxidation of fatty acids in all eukaryotes, in order to generate acetyl-CoA. (Millerioux *et al.*, 2012). Whilst in *T. brucei* this catabolic process is more important in the procyclic form, it is still required for the generation of GPI synthesis in BSF cells. Therefore, to further probe the hypothesis that mitochondrial metabolism is targeted by AN5568, we set out to identify any changes in lipids often associated with the mitochondria.

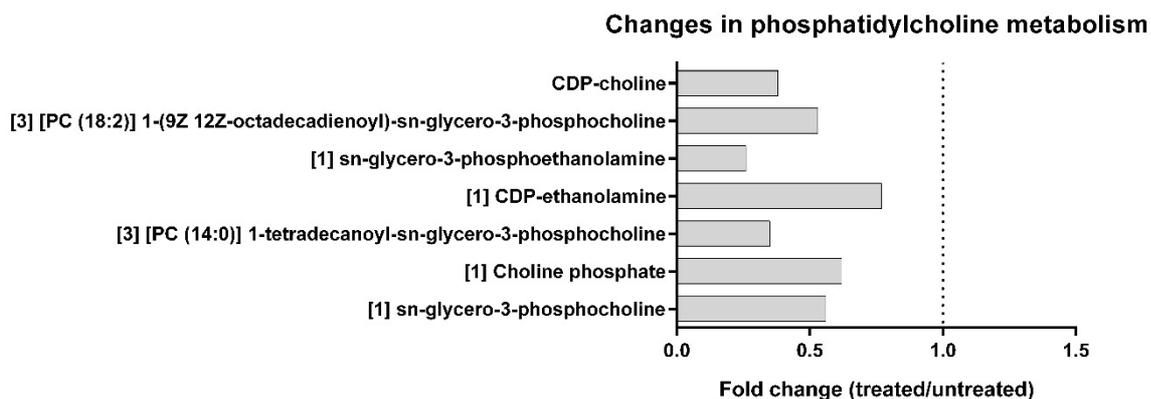
Fatty acids were seen on the mass spectrometer, although several of the putative metabolites had many isomers (fig. 3-11 & 3-12). Interestingly, these fatty acids all appeared to have a decreased abundance in AN5568-treated cells. However, it is impossible to confirm with these data whether this is due to mitochondrial disintegration or cell death in general.



**Figure 3-11: Fatty acid metabolism in cells treated with AN5568 for 6 hours at  $10\times EC_{50}$ .** In general, long chain fatty acids showed decreased abundance, although there were putative fatty acids showing increased abundance, such as eicosanoic acid (A). However, identification of these metabolites was complex, and many showed high numbers of isomers. Therefore, the identification of [FA (20:0)] eicosanoic acid (A, 11 isomers), decanoic acid (B, 9 isomers) and tetradecanoic acid (C, 15 isomers) is putative at best. Interestingly, glycerol-3-phosphate, a precursor of cardiolipin synthesis, was significantly decreased in benzoxaborole-treated cells (D). In addition, L-carnitine, an important molecule for the transport of fatty acyl chains into mitochondria, was also significantly decreased in abundance (E), thereby supporting the microscopy observations regarding damage in the mitochondria upon AN5568 treatment.

Another result of possible intrigue was a sharp decrease seen in L-carnitine abundance (fig. 3-11E). This molecule is crucial for the transport of acyl groups from fatty acids into the mitochondrial matrix for  $\beta$ -oxidation (Gilbert & Klein, 1982, Klein *et al.*, 1982). Therefore, its decrease suggests this process has been downregulated either intentionally, or as a result of drug treatment. We were unsuccessful in finding carnitine derivatives such as acylcarnitines in this experiment. However, several acylcarnitines were identified putatively in further experiments involving *Leishmania mexicana* promastigotes (section 3-6). Perhaps the most interesting type of lipid to further investigate in the context of mitochondrial biology, would be cardiolipins (IUPAC name: 1,3-bis(sn-3'-phosphatidyl)-sn-glycerol), as these usually constitute approximately 20% of the inner mitochondrial membrane, a higher concentration than in any other organelle membrane (Serricchio & Butikofer, 2012). However, its size (monoisotopic mass: >1,000) makes it difficult to detect by pHILIC mass spectrometry due to the possibility of complex fragmentations as well as lipids possessing a decreased ionising capability, and it was not found in this study. However, glycerol-3-phosphate, a precursor of the cardiolipin synthesis pathway, was found at decreased levels in AN5568-treated cells (fig. 3-11D). Whilst this could be indicative of decreased mitochondrial lipids, it is clear that glycerol-3-phosphate has diverse roles in the cell including as a component of glycolysis, and caution must be taken before reaching any conclusions with regards to this particular metabolite.

A trend of general decrease was seen in phosphatidylcholine (PC) and phosphatidylethanolamine (PE) metabolism (fig. 3-12). These molecules are all important components of membranes of both the whole cell, as well as individual organelles, and are the largest class of phospholipids in *T. brucei* (Smith & Butikofer, 2010). Whilst there was a clear pattern, we were unsure whether this was a direct result of the benzoxaborole, or perhaps a general result of cell death. Ongoing work in the Barrett group is currently observing the metabolic signatures of trypanosome cell death (D. Kim, personal communication), and the outcome of these analyses should be analysed in the context of drug treatment as well. Furthermore, the rounding of cells seen after 12 hours of AN5568 treatment at 10 $\times$  EC50 (fig. 3-4) could suggest that the outer membrane of the cell becomes severely compromised after drug treatment, which might explain the general loss of lipids in the drug-treated samples as they diffuse into the extracellular environment.



**Figure 3-12: AN5568 causes decreased abundances in lipids derived from phosphatidylcholine (PC) and phosphatidylethanolamine (PE).** These lipids are amongst the most abundant phospholipids found in *T. brucei* and are synthesised in the ER and on the ER membrane. These lowered abundances suggest reduced ER activity, but could also indicate loss of lipids due to cell death. The number of isomers predicted by the Ideom software are given in square brackets.

### 3.4.4 Further changes of interest

An untargeted metabolomics approach allowed us to look at all results in a hypothesis generating manner, rather than with preconceived ideas. Using this approach, we isolated several metabolic perturbations of interest, other than those mentioned previously. These are shown in table 3-2.

Importantly, we were able to identify the benzoxaborole in metabolite extracts from AN5568-treated parasites ( $m/z$ : 367.1005, RT: 4.63 mins). A further isotope of the drug, corresponding to an  $^{11}\text{B}$  isotope, was also found. Using the known isotope distribution of the common boron isotopes  $^{10}\text{B}$  and  $^{11}\text{B}$ , we attempted to mine the metabolomics data to find other metabolites with a similar isotope distribution, as this could give an insight into potential metabolic activation of the drug. Whilst the mzMatch output data showed several possible molecules, the peak intensities were low and they were dismissed as background noise after analysis of the raw data. We were therefore unable to confirm whether the drug requires metabolic activation.

There were several further perturbations in amino acid metabolism. Several acetylated or hydroxylated amino acids, such as N6-acetyl-L-lysine, were increased, perhaps indicating increased post-translational modifications occurring as a response to benzoxaborole treatment. Modification on lysine residues are often found in proteins such as histones, although it is unclear what could cause increases in free acetyl-L-lysine. There was also increased (R)-2-hydroxyglutarate ( $m/z$ : 148.0371, RT: 17.51 mins) present in AN5568-treated cells.

Measured Mass	Retention time	Predicted formula	No. isomers	Putative metabolite ID	Fold change in AN5568-treated cells	FDR
73.05280	7.67	C <sub>3</sub> H <sub>7</sub> NO	6	Aminoacetone	12.06	0.01213
188.11619	14.70	C <sub>8</sub> H <sub>16</sub> N <sub>2</sub> O <sub>3</sub>	7	N6-Acetyl-L-lysine	2.83	0.00056
148.03710	17.51	C <sub>5</sub> H <sub>8</sub> O <sub>5</sub>	18	(R)-2-Hydroxyglutarate	2.18	0.02035
173.08010	16.17	C <sub>6</sub> H <sub>11</sub> N <sub>3</sub> O <sub>3</sub>	1	5-Guanidino-2-oxopentanoate	2.01	0.00020
129.07899	7.76	C <sub>6</sub> H <sub>11</sub> NO <sub>2</sub>	9	N4-Acetylaminobutanal	1.71	0.00156
187.12085	13.75	C <sub>9</sub> H <sub>17</sub> NO <sub>3</sub>	3	8-Amino-7-oxononanoate	8.43	0.03460
244.08837	7.71	C <sub>10</sub> H <sub>16</sub> N <sub>2</sub> O <sub>3</sub> S	1	Biotin	0.66	0.10260
177.06370	15.30	C <sub>6</sub> H <sub>11</sub> NO <sub>5</sub>	1	4-Hydroxy-4-methylglutamate	3.46	0.00207
194.07902	15.07	C <sub>7</sub> H <sub>14</sub> O <sub>6</sub>	18	1-O-Methyl-myo-inositol	0.73	0.01799
230.01903	18.15	C <sub>5</sub> H <sub>11</sub> O <sub>8</sub> P	16	D-Ribose 5-phosphate	0.60	0.00100
161.06868	18.75	C <sub>6</sub> H <sub>11</sub> NO <sub>4</sub>	10	N-Methyl-L-glutamate	0.75	0.01297
367.10048	4.63	C <sub>17</sub> H <sub>14</sub> BF <sub>4</sub> NO <sub>3</sub>	1	AN5568	246.02	0.03408
252.13595	3.91	C <sub>14</sub> H <sub>20</sub> O <sub>4</sub>	1	ubiquinol-1	1.63	0.00513
165.06522	10.61	C <sub>6</sub> H <sub>7</sub> N <sub>5</sub> O	5	7-Methylguanine	0.78	0.01211
111.04324	13.34	C <sub>4</sub> H <sub>5</sub> N <sub>3</sub> O	1	Cytosine	0.62	0.00002
138.03168	18.75	C <sub>7</sub> H <sub>6</sub> O <sub>3</sub>	7	4-Hydroxybenzoate	1.76	0.00653

**Table 3-2: Further changes in *T. brucei* metabolism after AN5568 treatment, that were not associated with one particular pathway.** Additional analysis of the Ideom results file showed several other metabolites that were significantly changed in AN5568-treated cells. Whilst these metabolites could not be linked to one particular pathway, they could offer further clues of specific changes occurring in the *T. brucei* metabolome as a result of drug treatment.

The analyses also showed a large increase in putatively identified aminoacetone. This compound, also known as 1-amino-2-propanone, is a secondary amino ketone and is a product of aminoacetone synthase (glycine C-acetyltransferase) (Linstead *et al.*, 1977). A older study showed that the activity of this enzyme is tightly coupled to L-threonine dehydrogenase in vertebrates, and is involved in L-threonine catabolism (Aoyama & Motokawa, 1981), an important mitochondrial metabolic network in PCF trypanosomes, that has also recently been found to be essential in BSF parasites (Mazet *et al.*, 2013). Linstead and colleagues discussed in their study that L-threonine breakdown proceeds via 2-amino-3-oxobutyric acid, which spontaneously decarboxylates to aminoacetone in a pathway that ultimately generates glycine and acetate (Linstead *et al.*, 1977). Several studies have more recently identified aminoacetone as a source of methylglyoxal in diabetes patients, where it increases in cases of nutritional deprivation (Sartori *et al.*, 2008). In the context of AN5568-treatment, increase in aminoacetone could suggest perturbations in mitochondrial metabolism, or in acetate production, which could be consistent with the general downward trend in lipid abundance.

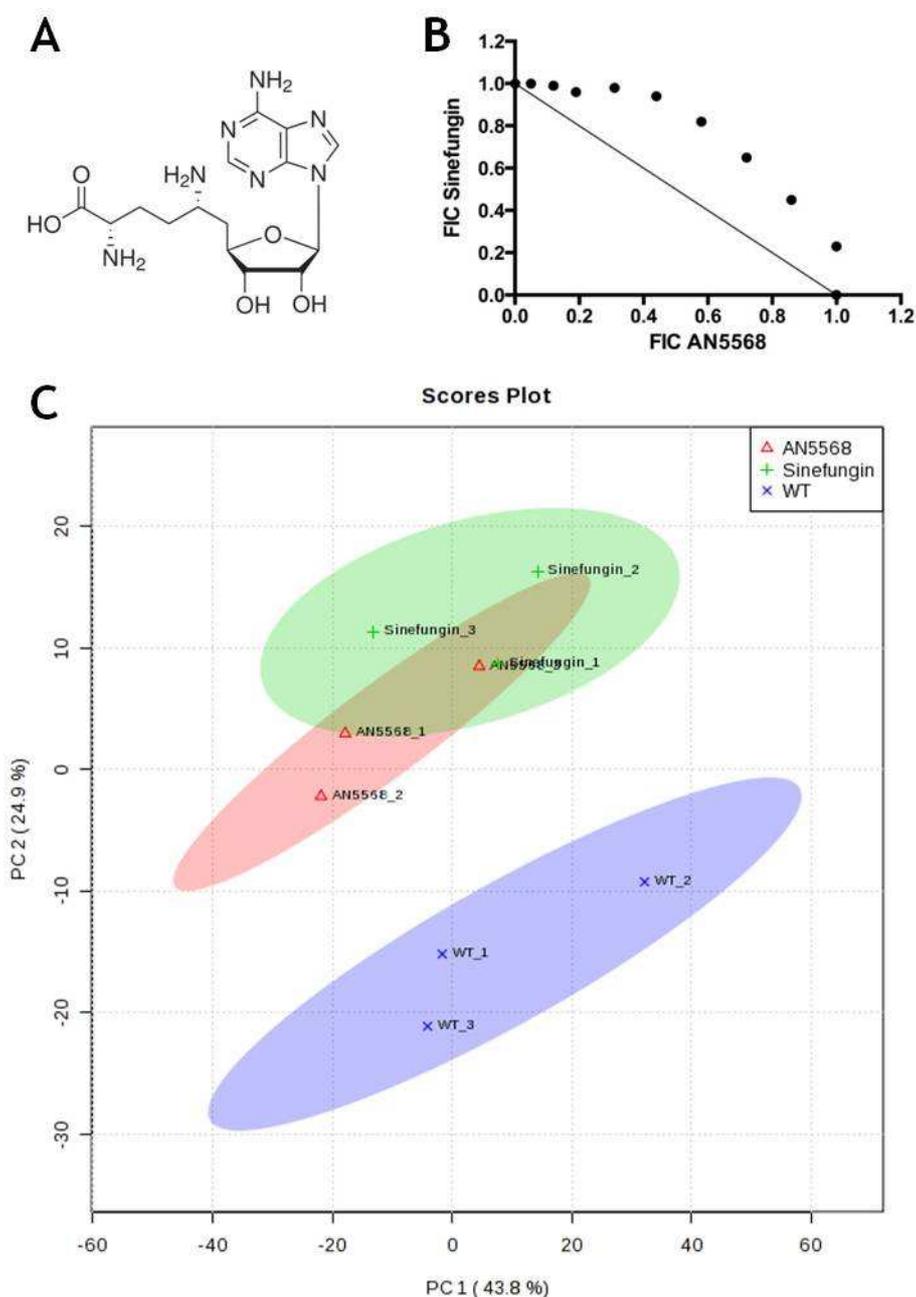
Carbohydrate metabolism was largely unchanged. Glycolysis components were not significantly different in drug-treated samples. However, we found metabolites in the pentose phosphate pathway to be decreased, with a putatively identified ribose 5-phosphate (*m/z*: 230.0190, RT: 18.15

mins) a good example of this. However, it must be noted that ribose 5-phosphate could be a fragment of a larger structure. In addition, it is impossible to distinguish ribose from other pentose sugars which are not involved in the PPP.

### 3.4.5 Comparison of AN5568 & sinefungin metabolotypes

Previous data generated by the Barrett group using AN5568 support the data presented here. In particular, high levels of S-adenosyl-L-methionine and 5'-MTA were seen during preliminary experiments (Creek & Barrett, unpublished). The prevailing theory that a methyltransferase reaction might be the drug target led to the comparative analysis of the metabolic profile of AN5568-treated and sinefungin-treated parasites. Sinefungin is a naturally occurring AdoMet analogue (fig. 3-13A). Whilst structurally related to AdoMet, it contains an amino methylene group in place of the methylated sulfonium group that usually acts as the methyl donor (Devkota *et al.*, 2014). Thus, the molecule inhibits AdoMet-dependent methyltransferase reactions in a non-specific/non-selective manner (Vedel *et al.*, 1978).

After calculating the  $EC_{50}$  of sinefungin to be approximately 1.0 nM in Lister 427 trypanosomes, we looked at the possibility of AN5568-sinefungin interactions by carrying out an isobologram (fig. 3-13B). Isobolograms show whether there are synergistic, antagonistic, or additive effects between two drugs. This experiment suggests an antagonistic drug-drug interaction because the addition of one drug causes the overall efficacy of the other to diminish. Further analyses of these data showed that the  $\Sigma FIC$  for AN5568 and sinefungin were 5.28 and 8.02 respectively. The mean  $\Sigma FIC$  was therefore 6.65, which indicates significant antagonism between the 2 drugs (Seifert *et al.*, 2011). This result indicates that the two drugs might act on similar protein targets or metabolic pathways. Another possibility is that the two drugs are transported through the same routes. Sinefungin is non-specific, affecting the AdoMet binding pocket on MTases, and affects most organisms to a similar degree. Thus, the mode of action of this drug probably differs from that of AN5568. Furthermore, given the contrasting structures, should these two drugs target similar classes of proteins, one would assume they do so with different mechanisms, especially since the benzoxaborole maintains a strong affinity for diols, which has not been shown for sinefungin.



**Figure 3-13: Comparative analysis of metabolic profiles of AN5568-treated and sinefungin-treated *T. brucei*.** A) The structure of sinefungin. This compound is an AdoMet analogue, and therefore binds strongly to the AdoMet-dependent MTase domain of MTases and inhibits them. B) A fixed-ratio isobologram was carried out using sinefungin and AN5568. Data showed that there is significant antagonism between the two drugs, suggesting they could act on the same pathway, or interfere with each other's mechanisms of drug action. C) Comparative metabolomics analysis was carried out to assess any similarities between the metabolomes of sinefungin-treated and AN5568-treated cells. Whilst the wild-type samples separated well, both drug-treated samples showed overlap, thereby supporting the results obtained from isobologram experiments.

Having established a potential link between the two drugs, we next looked at the metabolic profile of sinefungin-treated cells and compared them to both a wild-type, DMSO-treated control and an AN5568-treated sample group. Interestingly, a PCA plot of the 9 samples suggested an overlap between the two drug-treated samples, again highlighting potential similarities between the modes of action of AN5568 and sinefungin (fig. 3-13C). We further analysed this by looking at single metabolites in detail (fig. 3-14).

Given the metabolic changes occurring in AN5568-treated cells, and the hypothesis that methyltransferase inhibition should give rise to changes in methionine metabolism, the changes in these pathways were investigated first (fig. 3-14A). As expected, both AdoMet and 5'-MTA were significantly increased in comparison to a DMSO control, in both drug-treated sample groups. However, the fold change in sinefungin-treated cells was overall lower than that in AN5568-treated cells. For AdoMet, there was a 4.10-fold increase in sinefungin-treated cells, whilst the increase was 7.67-fold in AN5568-treated cells. Similarly, 5'-MTA was increased 3.66-fold and 6.17-fold in sinefungin-treated and AN5568-treated samples respectively. Whilst both sample groups were treated at  $10\times EC_{50}$  for 6 hours, the absolute concentrations of the drugs contrasted significantly, which could lead to more extreme perturbations for some metabolites.

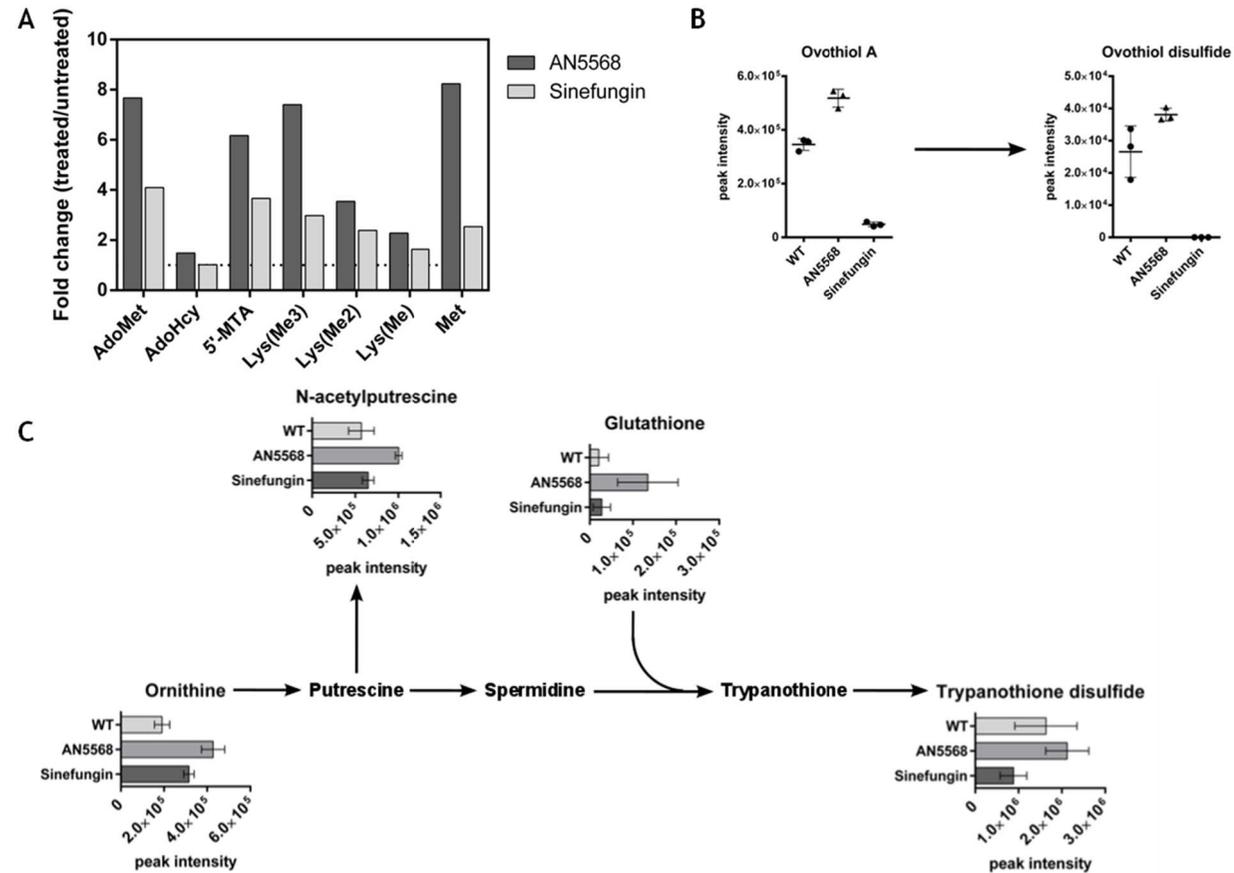
Somewhat interestingly, there was no apparent change in S-adenosyl-L-homocysteine (AdoHcy,  $m/z$ : 384.1216, RT: 15.32 mins) levels after sinefungin treatment (fig. 3-14A). Given the high levels of AdoMet and 5'-MTA resulting from methyltransferase inhibition after sinefungin treatment, we hypothesised there would be a corresponding decrease in AdoHcy and downstream metabolites such as L-homocysteine (Hcy) and L-cystathionine. However, sinefungin treatment does not lead to these expected changes. It is possible that the balance of the hydrolysis reaction between AdoHcy and Hcy is highly regulated, potentially by lysine acetylation on S-adenosyl-L-homocysteine hydrolase (AdoHcyH), the enzyme that catalyses this reaction (Wang *et al.*, 2014b). Therefore, MTase inhibition could leave the AdoHcy pool only moderately changed, whilst causing the high levels of AdoMet and its degradation products, as seen in these data. Interestingly, acetylated free L-lysine was found to increase in AN5568-treated samples, but we could not deduce whether this was linked to AdoHcyH. Nonetheless, given the metabolic profile of sinefungin-treated cells, it is likely that methyltransferase inhibition, albeit with a different mechanism, could be a target of AN5568. Methyltransferases are a highly divergent class of proteins with very small sequences conserved across the group, such as the binding sites of the AdoMet substrate (Wlodarski *et al.*, 2011). However, members of this class can possess similar domains depending on the methylation target. For example, most lysine MTases possess a SET domain (Schubert *et al.*, 2003). Should this domain be targeted, it is likely that many lysine MTases would be inhibited.

Importantly, there were also marked differences between the metabolic profiles of AN5568- and sinefungin-treated cells. Sinefungin, unlike AN5568, is a non-specific inhibitor of AdoMet-dependent MTases, thereby affecting a high number of metabolic pathways. One example is the biosynthesis of ovothiol (fig. 3-14B). This metabolite is thought to be synthesized in situ through AdoMet-dependent methylation of mercaptohistidine (Ariyanayagam & Fairlamb, 2001). Whilst we could only see a putatively labelled mercaptohistidine with low confidence, putative peaks predicted to correspond to Ovothiol A ( $m/z$ : 201.0572, RT: 17.55 mins) as well as ovothiol disulfide

( $m/z$ : 400.0987, RT: 17.38 mins), the reduced form of the metabolite, were found in the datasets (fig. 3-14B). Our data showed a large reduction in this metabolite after sinefungin-treatment, suggesting inhibition of the corresponding methyltransferase by the compound.

In addition, sinefungin left the polyamine pathway largely unchanged (fig. 3-14C). This was in stark contrast to AN5568, which appeared to cause upregulation of this pathway, possibly due to the stress-inducing effects of the benzoxaborole. This pathway is upregulated in response to oxidative stress with the end product, trypanothione, being a highly potent reducing agent (Fairlamb *et al.*, 1985). These data would indicate that the cells are undergoing significant amount of oxidative stress, supported by the production of both trypanothione disulfide and ovothiol disulfide.

These data also showed increased levels of ornithine ( $m/z$ : 132.0898, RT: 18.04 mins) in the AN5568-treated sample group. It is currently thought that ornithine is not synthesised in situ, but is transported from the extracellular environment (M. Barrett, personal communication), suggesting increased ornithine transport upon drug treatment (fig. 3-14C). We hypothesise the reason for this is increased polyamine synthesis in response to drug treatment. This pathway is involved in the oxidative stress response, and is unique in *T. brucei*, in particular due to the generation of trypanothione (Fairlamb *et al.*, 1985, Henderson *et al.*, 1990). One theory could be that mitochondrial defects, such as those observed by microscopy (fig. 3-3 & 3-4), cause leakage of mitochondrial material including reactive oxygen species generated by oxidative phosphorylation. This phenomenon could indeed lead to a strong oxidative stress response, and this might be worth investigating further.



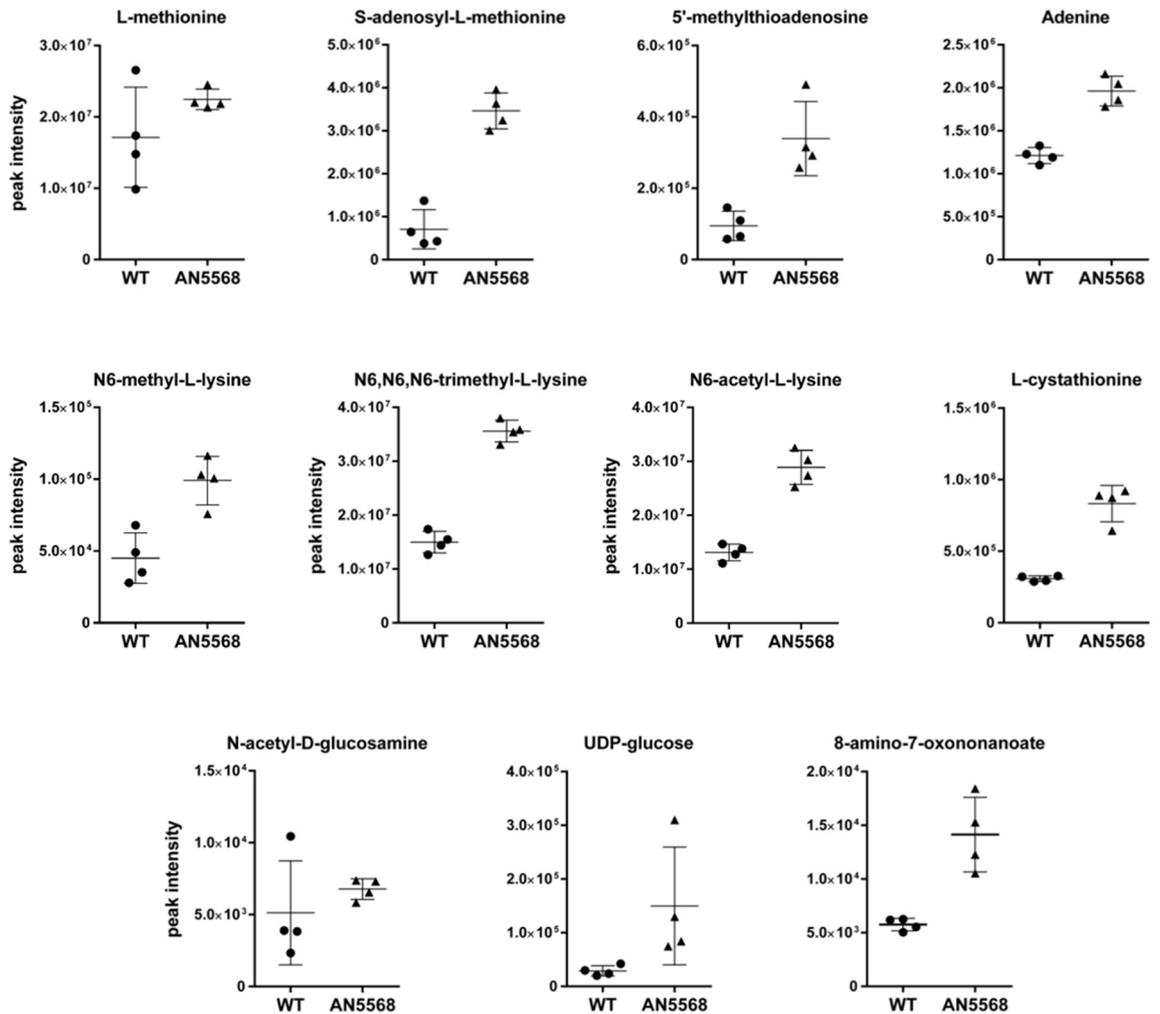
**Figure 3-14: Further analysis of metabolic changes in AN5568- and sinefungin-treated *T. brucei*.** A) The metabolic phenotypes observed in methionine metabolism were consistent between both sample groups when compared to wild-type cells. However, AN5568 treatment led to much higher fold changes. Interestingly, AdoHcy did not decrease as expected in sinefungin-treated cells, suggesting AdoHcy is a poor marker for MTase inhibition. Abbreviations: AdoMet - S-adenosyl-L-methionine, AdoHcy - S-adenosyl-L-homocysteine, 5'-MTA - 5'-methylthioadenosine, Lys(Me3) - N6,N6,N6-trimethyl-L-lysine, Lys(Me2) - N6,N6-dimethyl-L-lysine, Lys(Me) - N6-methyl-L-lysine, Met - L-methionine. B) Sinefungin is a non-specific MTase inhibitor which led to defects in a wide range of MTase reactions. Ovothiol is a metabolite utilised in oxidative stress responses, and is generated by the methylation of mercapto-histidine. Whilst AN5568 caused increases in ovothiol, and its reduced form ovothiol disulphide, sinefungin treatment caused significant decreases in both metabolites, indicating that the MTase reaction had been severely affected by the drug. C) AN5568 also seemed to cause an upregulation of the trypanothione biosynthesis pathway, indicating a response to oxidative stress. This was not seen in sinefungin-treated cells.

### 3.5 Metabolism in AN5568-treated procyclic cells

---

Metabolomics analyses of AN5568-treated PCF *T. brucei* were also carried out, in order to ascertain whether the benzoxaborole caused similar effects in parasites at this point in the *T. brucei* life cycle. For this study, PCF 427 Lister cells were used, in order to maintain consistency with the trypanosome strain, allowing comparisons between BSF and PCF cells. As shown by the initial *in vitro* study, this experiment required a significantly higher concentration of the benzoxaborole due to a high EC<sub>50</sub>. Therefore, 15 μM, equal to 10× EC<sub>50</sub>, was used for drug treatment, and cultures were incubated for 8 hours before metabolite extractions were carried out, as this was deemed to be an optimal time point given the preliminary *in vitro* data. Two sample groups were studied: A drug treated sample group of 4 independent replicates and a sample group consisting of 4 replicates treated with an equal volume of DMSO.

Interestingly, the metabolic profiles of AN5568-treated PCF cells showed similar trends to those seen in BSF parasites. Again, AdoMet (*m/z*: 398.1386, RT: 13.88 mins) and 5'-MTA (*m/z*: 297.0898, RT: 7.69 mins) were highly increased after drug treatment. In addition, methylated lysines were found in higher abundance than in control samples (fig. 3-15). Also comparable with BSF drug treatment were the lack of changes in methylated arginine, in contrast to lysine. Indeed, acetylated lysine was also increased in AN5568-treated PCF cells, as they were in BSF cells (fig. 3-15). GDP-mannose, which was found to increase in BSF parasites treated with AN5568, was not detected in this experiment. However, UDP-N-acetyl-D-glucosamine, which is also involved in GPI biosynthesis, was increased in a similar manner as that seen in AN5568-treated BSF cells (fig. 3-15). Finally, an increase in L-cystathionine (*m/z*: 222.0677, RT: 14.25 mins) was again seen (fig. 3-15) in this dataset.

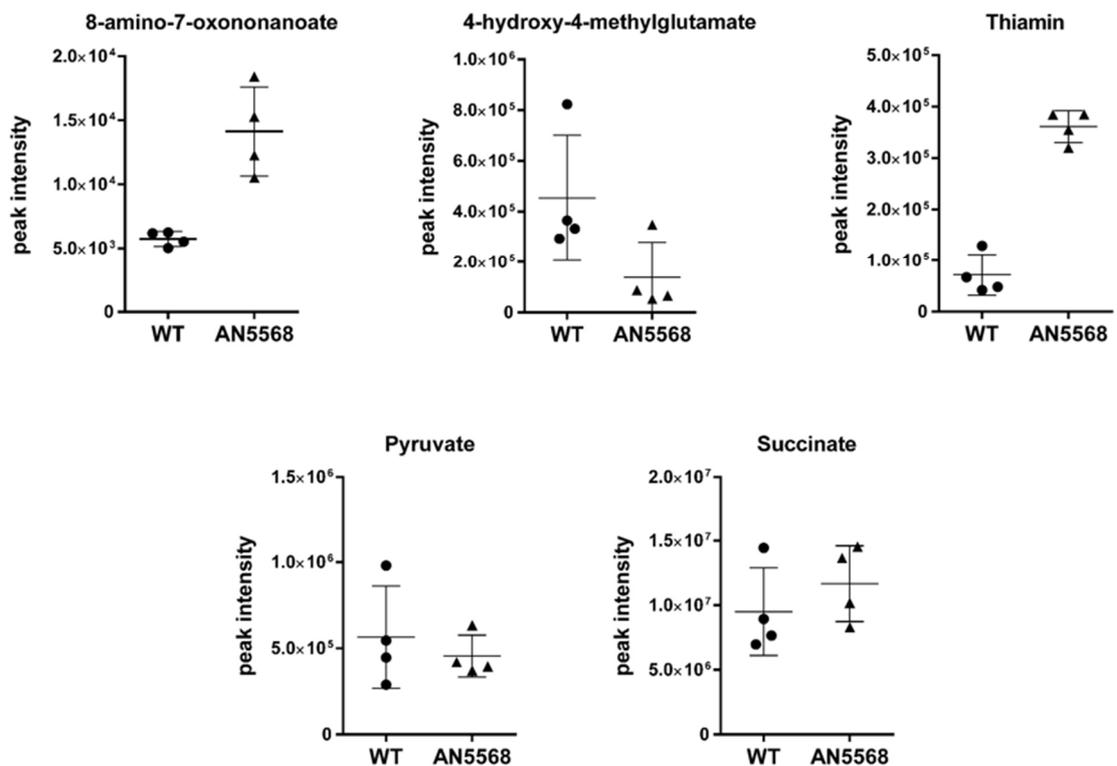


**Figure 3-15: Metabolic changes occurring after AN5568 treatment of PCF *T. brucei*.** Many of the changes seen in BSF cells subsequent to treatment with AN5568 were also observed in PCF cells. Increases were seen in AdoMet (m/z: 398.1386, RT: 13.88 mins), 5'-MTA (m/z: 297.0898, RT: 7.69 mins) and adenine (m/z: 135.0544, RT: 10.01 mins), as well as a slight increase in L-methionine (m/z: 149.0511, RT: 11.34 mins). Another similarity was widespread increases in abundance of modified lysine residues such as N6-methyl-L-lysine (m/z: 160.1211, RT: 18.59 mins), N6,N6,N6-trimethyl-L-lysine (m/z: 188.1227, RT: 17.69 mins) and N6-acetyl-L-lysine (m/z: 188.1160, RT: 12.05 mins). There was an increase in L-cystathionine (m/z: 222.0677, RT: 14.25 mins), which is generated downstream from the MTase reaction. There were similar changes in carbohydrate metabolism, with increases seen in N-acetyl-D-glucosamine (m/z: 221.0901, RT: 11.54 mins) and UDP-glucose (m/z: 566.0551, RT: 13.61 mins), and finally, an increase in 8-amino-7-oxononanoate (m/z: 187.1210, RT: 11.65 mins), the role of which is unknown in *T. brucei*.

The PCF dataset was also analysed for key differences between BSF and PCF AN5568-treated cells (fig.3-16). An increase was observed in a metabolite predicted to be 3-demethylubiquinol, which is found in the ubiquinol biosynthesis pathway (Chaudhuri *et al.*, 2006). Interestingly, this metabolite undergoes an MTase reaction with AdoMet as donor, leading to the formation of ubiquinol-9, which is a substrate for the trypanosome alternative oxidase (TbAOX), an important component of the mitochondrial electron chain (Chaudhuri *et al.*, 1995, Fang & Beattie, 2003). An increase in 3-demethylubiquinol could suggest inhibition of the 3-demethylubiquinol 3-O-methyltransferase.

Other differences included high levels of thiamin in PCF cells treated with AN5568, which was not observed in BSF cells. Thiamin is an important component in energy metabolism, and it is

thought that the metabolite is synthesized by the parasite, as uptake has not been observed (Stoffel *et al.*, 2006). Another interesting change was 4-hydroxy-4-methylglutamate (fig. 3-16). During metabolomics analysis with BSF parasites, this metabolite, derived from glutamate, was found to increase subsequent to AN5568-treatment. In contrast, it was decreased following AN5568-treatment in PCF cells. There is no literature on this metabolite in protozoan parasites, although it has been found to accumulate in plants (Kakule *et al.*, 2015). Here, it is proposed that the metabolite is generated using pyruvate as a precursor, in a reaction catalysed by an aldolase (Kakule *et al.*, 2015). Pyruvate is produced at high levels as a glycolytic waste product in BSF *T. brucei*. In PCF cells pyruvate is used in pathways synthesising either succinate or acetate. This has also recently been shown to occur in BSF cells (Mazet *et al.*, 2013). Whilst acetate was not detected in this experiment, both succinate and pyruvate were found not to undergo any significant changes following treatment with AN5568 (fig.3-16).



**Figure 3-16: Further metabolic changes in AN5568-treated PCF *T. brucei* parasites.** Much like BSF *T. brucei*, 8-amino-7-oxononanoate was increased in AN5568-treated cells. In contrast, 4-hydroxy-4-methylglutamate was decreased in AN5568-treated PCFs. Another change not observed in BSF cells was increased levels of thiamin. Finally, components of energy metabolism such as pyruvate and succinate appeared unchanged in drug-treated cells, as seen in BSFs.

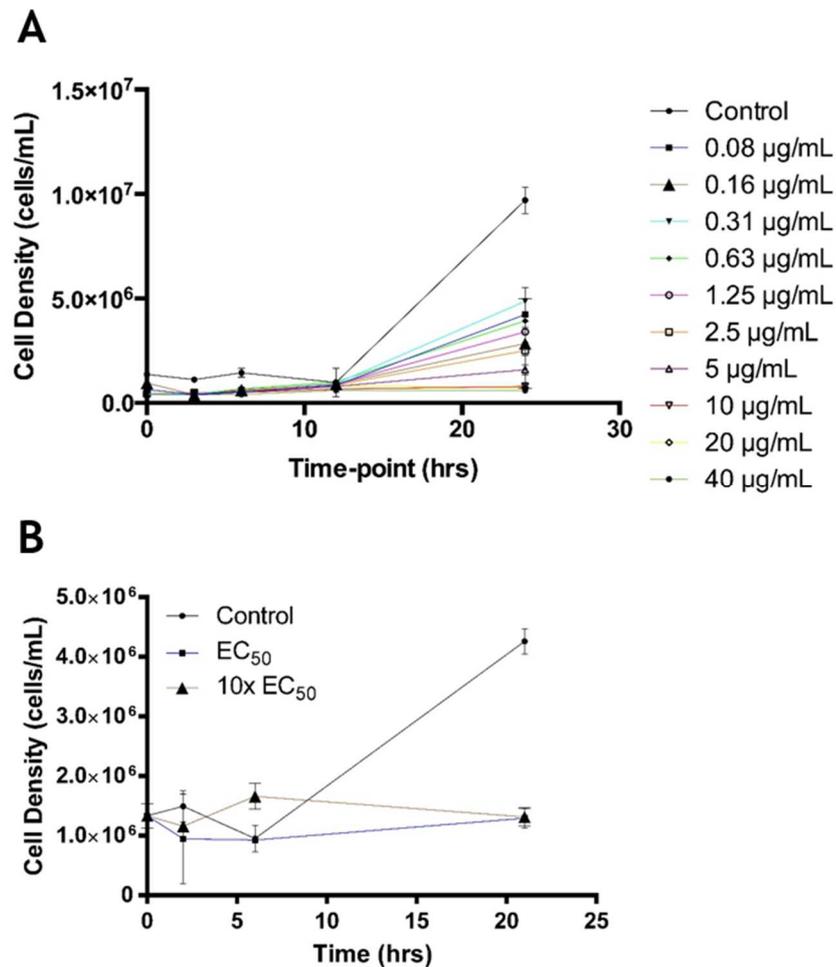
The results presented here indicate that similar changes occur in PCF cells, in addition to a couple of key differences that should be analysed further. Unfortunately, these differences occur with metabolites whose roles are not understood well in trypanosomatids. As mentioned above, the concentration of drug required to elicit the same metabolic responses are significantly higher for PCF cells, suggesting that the target is either less essential in procyclics, or it is present in higher abundance. There is also a chance that high concentrations of the

benzoxaborole lead to lethal non-specific effects on the cell, although the metabolomics data does not support this hypothesis. Finally, another difference between BSF and PCF cells that could influence the efficacy of the drug is the differential expression of transporters between the two life cycle stages. However, as mentioned previously, SAR knowledge on this benzoxaborole is scarce, and the uptake mechanisms are currently unknown.

### 3.6 AN5568 elicits a similar metabolic response in *L. mexicana*

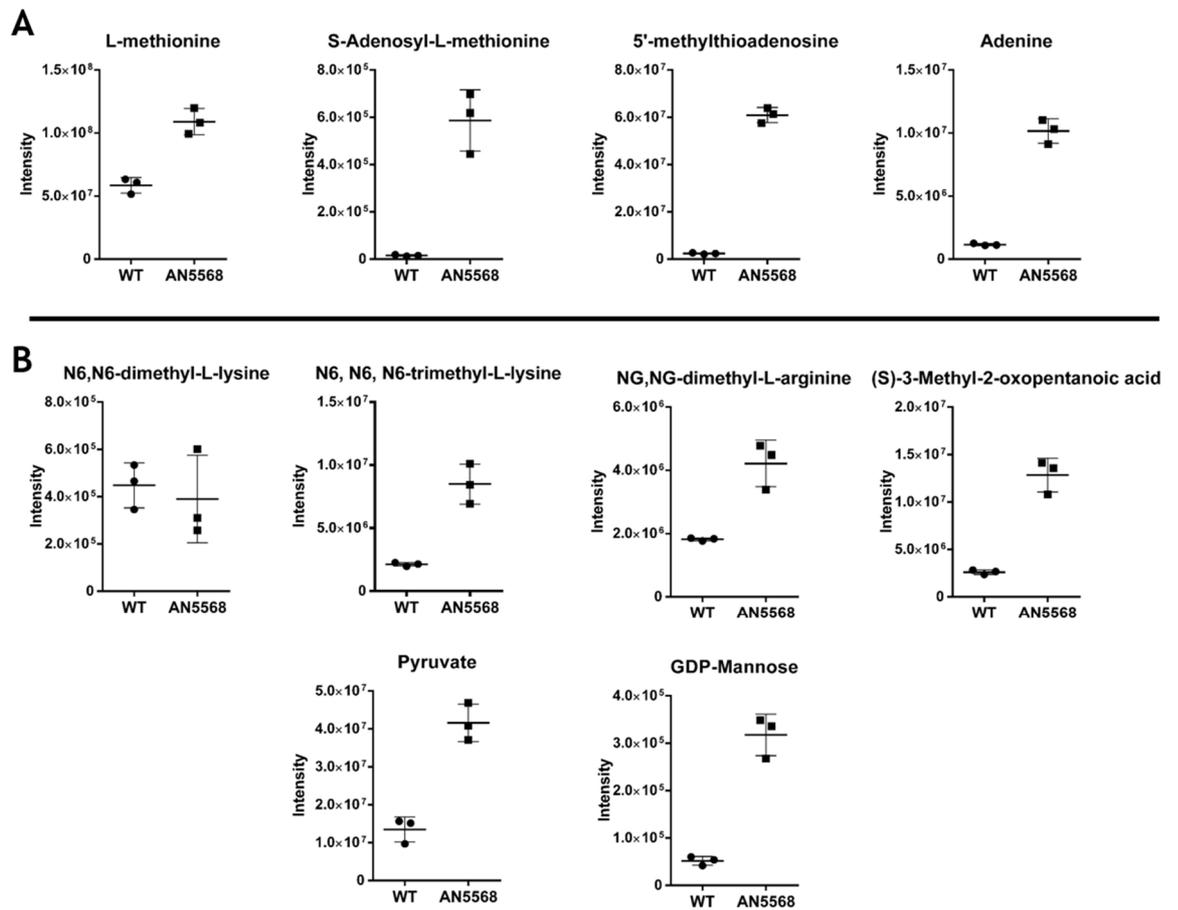
---

*Trypanosoma brucei* is a member of the kinetoplastidea, which also includes *Trypanosoma cruzi*, the causative agent of Chagas disease in South America, and *Leishmania spp.*, parasites responsible for the Leishmaniases. To determine whether AN5568 is active against these other protozoans, collectively called the TriTryps, we sought to analyse the metabolome of AN5568-treated leishmania cells. *L. mexicana* promastigotes are easily cultivated in laboratory culture, and a drug-treatment incubation was set-up after the EC<sub>50</sub> was calculated to be  $37.8 \pm 3.7 \mu\text{M}$ . Firstly, growth of cells was measured in various concentrations of the benzoxaborole (fig. 3-17). As was observed with PCF *T. brucei* cells, the benzoxaborole did not completely kill *L. mexicana* parasites, even after 24 hours (fig. 3-17). Whilst cell division seemed to have been abolished, live cells were seen even at the highest concentration of benzoxaborole used in both growth curves. Furthermore, growth was static even at  $10\times$  EC<sub>50</sub> AN5568 (fig. 3-17B).



**Figure 3-17: Effect of AN5568 on *in vitro* growth of *L. mexicana* promastigotes.** Cells were grown in the presence of varying concentrations of the benzoxaborole. A) In the first experiments, concentrations were replicated from a previous publication by Jacobs and colleagues (Jacobs *et al.*, 2011). This showed that concentrations as low as 0.08  $\mu\text{g/mL}$  were sufficient to cause a decrease in growth rate after 24 hours. The extent of cellular growth decreased with every doubling of the AN5568 concentration. However, at 40  $\mu\text{g/mL}$ , there were still live cells visible in the cultures. B) The  $\text{EC}_{50}$  of AN5568 in *L. mexicana* was calculated to be  $\sim 38 \mu\text{M}$ , which corresponds to 14  $\mu\text{g/mL}$ . Interestingly, cells treated with this concentration were equally resilient as those treated with  $10\times \text{EC}_{50}$  (140  $\mu\text{g/mL}$ ). In both cases, cell density was stationary, although live cells were observed 24 hours after drug treatment.

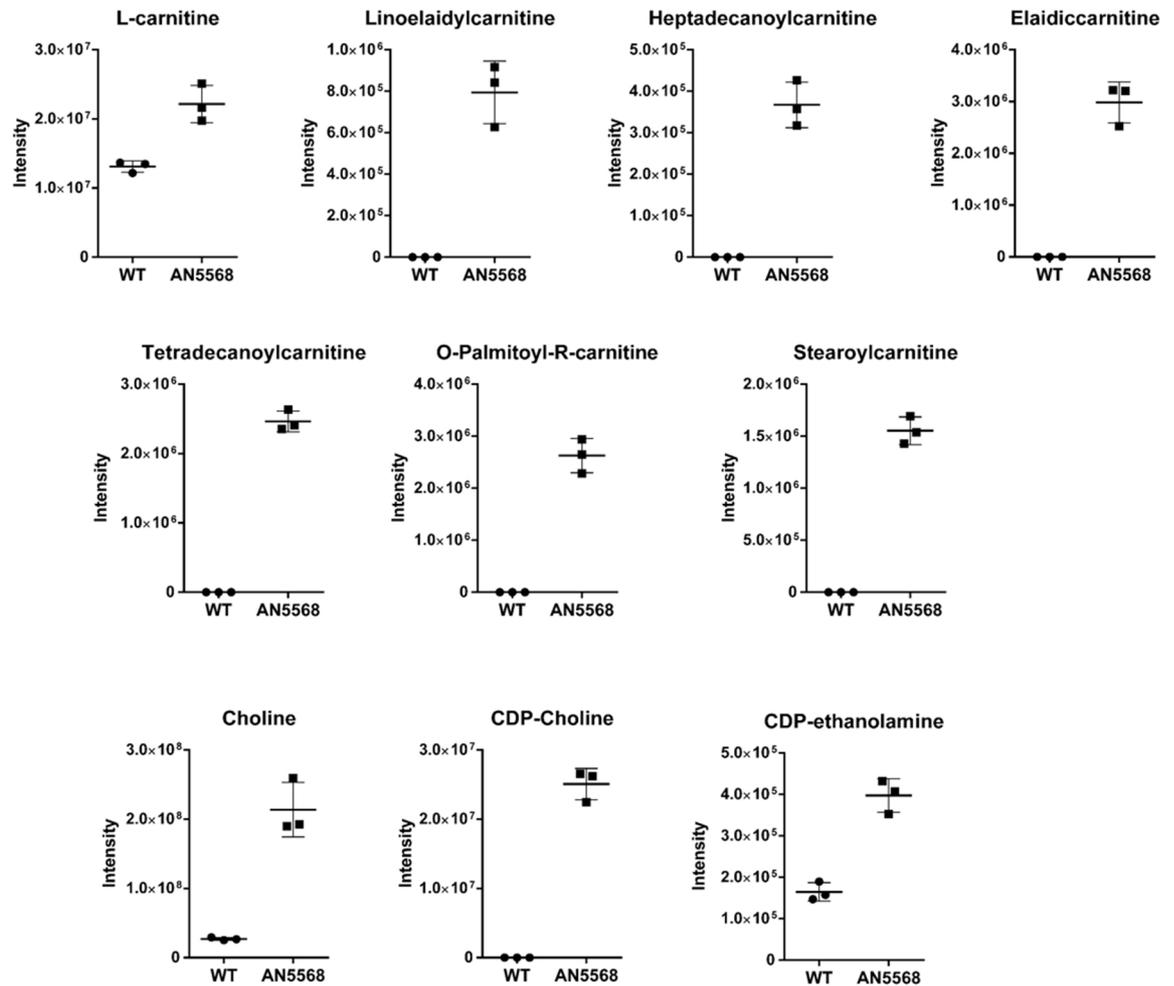
Metabolite extracts were prepared after a 6-hour incubation with  $10\times \text{EC}_{50}$  concentration of the benzoxaborole. Mass spectrometry analysis was carried out by Glasgow Polyomics, and data was processed using the same mzMatch/Ideom pipeline as described in the methods. Interestingly, the metabolic changes post-treatment were very similar to those observed in *T. brucei* (fig.3-18). AdoMet ( $m/z$ : 398.1372, RT: 111.14 mins), 5'-MTA ( $m/z$ : 297.0896, RT: 6.97 mins) and adenine ( $m/z$ : 135.0545, RT: 8.30 mins), were again found at high levels in drug-treated samples (fig. 3-18). In addition, there was a similar increase in modified lysines in drug-treated cells, although dimethyl-L-lysine ( $m/z$ : 174.1368, RT: 15.11 mins) was not found to increase in this experiment (fig. 3-18).



**Figure 3-18: AN5568 causes metabolic perturbations in *L. mexicana* that are similar to those observed in *T. brucei*.** The most significant metabolic phenotypes in BSF and PCF *T. brucei* treated with AN5568 were increased levels of AdoMet, 5'-MTA and adenine. These metabolites, along with L-methionine, were again increased in *L. mexicana* promastigote samples after  $10 \times EC_{50}$  ( $380 \mu\text{M}$ ) drug treatment (A). Whilst N6,N6-dimethyl-L-lysine did not show variation between the two sample groups, N6,N6,N6-trimethyl-L-lysine was increased after drug treatment, similar to *T. brucei* cells. Interestingly, NG,NG-dimethyl-L-arginine was found to increase subsequent to AN5568-treatment, in contrast to *T. brucei*, where no increase was detected. Further changes were seen in (S)-3-methyl-2-oxopentanoic acid, pyruvate and GDP-mannose.

Again, key differences were observed between AN5568-treated *T. brucei* and *L. mexicana* (fig. 3-19). One of the most intriguing results in this dataset was high levels of putative carnitine derivatives, which were not seen in the control group (fig. 3-19). These derivatives were also not seen in other metabolomics experiment and could appear due to defects in lipid transport across the mitochondrial membranes. In addition, there were increases in putative lipid metabolites (fig. 3-19). This was in stark contrast to the *T. brucei* results, which indicated decreased lipids (fig. 3-11-3-12). There are several reasons this phenomenon could have occurred. *L. mexicana* are a different species of kinetoplastid, and thus exhibit differences in the overall "lipidome", as reviewed by Smith and Butikofer (Smith & Butikofer, 2010). In addition, significant metabolic changes occur throughout the differential life cycles of trypanosomatids. In the case of *Leishmania*, promastigotes colonise the insect vector, whilst amastigotes are mammalian infective. Therefore, the parasites used in this study are more

similar to PCF *T. brucei*, and comparisons would be predicted to be weaker between *L. mexicana* promastigotes and *T. brucei* BSF parasites.



**Figure 3-19: Changes in lipid metabolism after AN5568-treatment of *L. mexicana* promastigotes.** There were significant perturbations in carnitine metabolism, with several derivatives observed only in AN5568-treated cells. L-carnitine itself was increased in abundance in drug-treated cells, as well as linoelaidylcarnitine, heptadecanoylcarnitine, elaidiccarnitine, tetradecanoylcarnitine, O-palmitoyl-R-carnitine and stearoylcarnitine. This might reflect increased mitochondrial metabolism in AN5568-treated cells.

With regards to key differences in lipid biosynthesis, one interesting feature of *Leishmania* is their ability to synthesize phosphatidylcholine (PC) from phosphatidylethanolamine (PE). This is a conserved process in eukaryotic cells that forms part of the Kennedy pathway (Gibellini *et al.*, 2009). In *Leishmania*, PE can be methylated to PC, but the enzyme for this reaction, PE N-MTase, has not been found in the trypanosome genome (Smith & Butikofer, 2010). PE itself is a product of the Kennedy pathway that involves conversion of ethanolamine to CDP-ethanolamine. In addition, a similar pathway converts choline to CDP-choline, before PC is generated (Farine *et al.*, 2015). Of the metabolites that were found in AN5568-treated *T. brucei*, the majority were found in lower abundance (fig. 3-12). However, they were all increased in *L. mexicana* cells, which could indicate that the benzoxaborole targets different pathways in each organism. On

the other hand, this result in particular could also suggest that the benzoxaborole possesses several targets including, for example, the PE N-MTase in *Leishmania*.

### 3.7 Targeting methionine metabolism

---

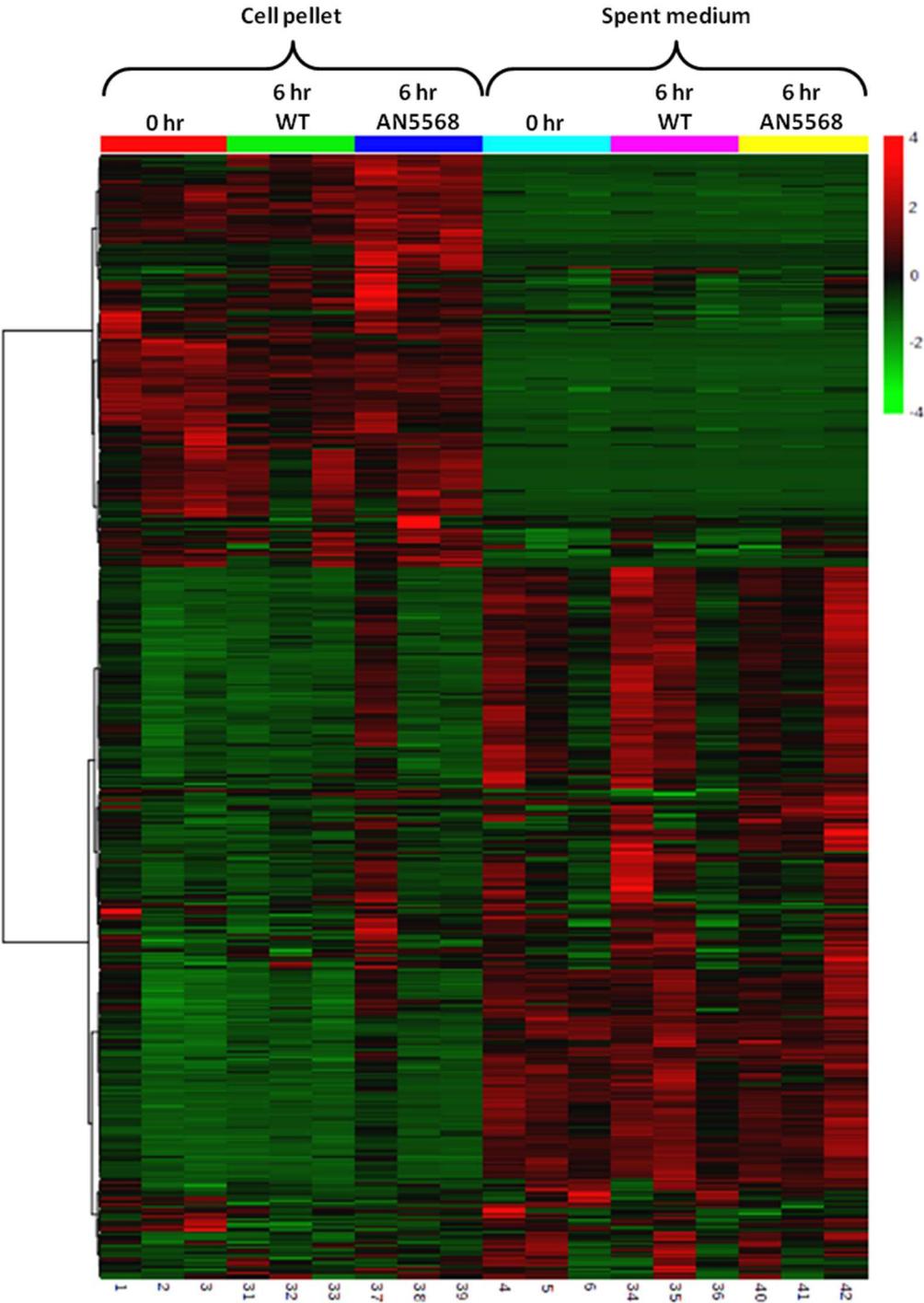
Currently, it is not thought that *T. brucei* recycles, or synthesises, L-methionine (Hasne & Barrett, 2000). It must therefore transport L-methionine into the cytoplasm from its extracellular environment. This methionine is then incorporated into AdoMet, as discussed in chapter 1. As we were able to observe several significant changes in methionine metabolism, we hypothesized that the benzoxaborole might affect methylation. To further investigate this hypothesis, we pursued a targeted metabolomics experiment utilizing methionine universally labelled with stable carbon isotopes ( $^{13}\text{C}$ ). This way, AdoMet becomes labelled with 5 carbon isotopes which can be traced throughout the metabolome to search for changes between control and AN5568-treated sample groups.

For this experiment, cells were grown in Creek's minimal medium (CMM), a semi-defined medium containing far less components when compared to standard culturing media such as HMI-9 (Hirumi & Hirumi, 1989, Creek *et al.*, 2013). In addition, this medium allowed us to add 100%  $^{13}\text{C}$ -(U)-methionine at a final concentration of 200  $\mu\text{M}$ . This methionine was added and cell pellets, as well as spent medium, were analysed after 6 hours, in addition to a non-labelled control at 0 hours. Labelling patterns were analysed using mzMatch-ISO, an R-based package that calculates isotope labelling distributions (Chokkathukalam *et al.*, 2013). The results were analysed in the context of methionine metabolism, using metabolic pathways from Kegg and Metacyc (Caspi *et al.*, 2006, Okuda *et al.*, 2008, Caspi *et al.*, 2015).

Using this pipeline, several hundred metabolites were identified by mzMatch-ISO after scanning the data for each metabolite in the KEGG database (Okuda *et al.*, 2008). However, a surprisingly small number of these showed any isotope labelling from L-methionine, suggesting that carbons from this source are only used by the parasite for very specific pathways, or that further labelling is a slow process and thus, requires longer than 6 hours before isotope labels are incorporated. There were 17 labeled metabolites detected in positive mode, whilst a further 8 were identified in negative mode. These data are presented in the supplementary data (S2). The mzMatch-ISO output data was subsequently manually filtered to remove peaks with high amounts of background noise.

Whilst AdoMet is a widely used methyl donor, it is mainly used as a substrate for nucleic acid or protein methylation, which could explain the low number of  $^{13}\text{C}$ -labelled metabolites. In addition, the experimental data presented here was generated from a 6 hour time-course,

which might not have been sufficient for some methylation pathways to incorporate <sup>13</sup>C isotopes. For example, we could not detect arginine methylation (S2).



**Figure 3-20:** Heat plot showing metabolic changes upon AN5568 treatment in both intracellular metabolism and spent medium. Both intracellular metabolites and spent medium metabolites were analysed in this experiment. There was a significant difference in metabolic profile between cell pellets and spent medium, which was expected, given the difference between the two sample groups. The heat map shows large amounts of variation within sample groups (for example, the first replicate taken from the AN5568-treated 6-hour time-point), which was important to consider for downstream analyses. Heat map was generated by Euclidian distance measures using the Ward clustering algorithm, in the Metaboanalyst tool (Xia *et al.*, 2015).

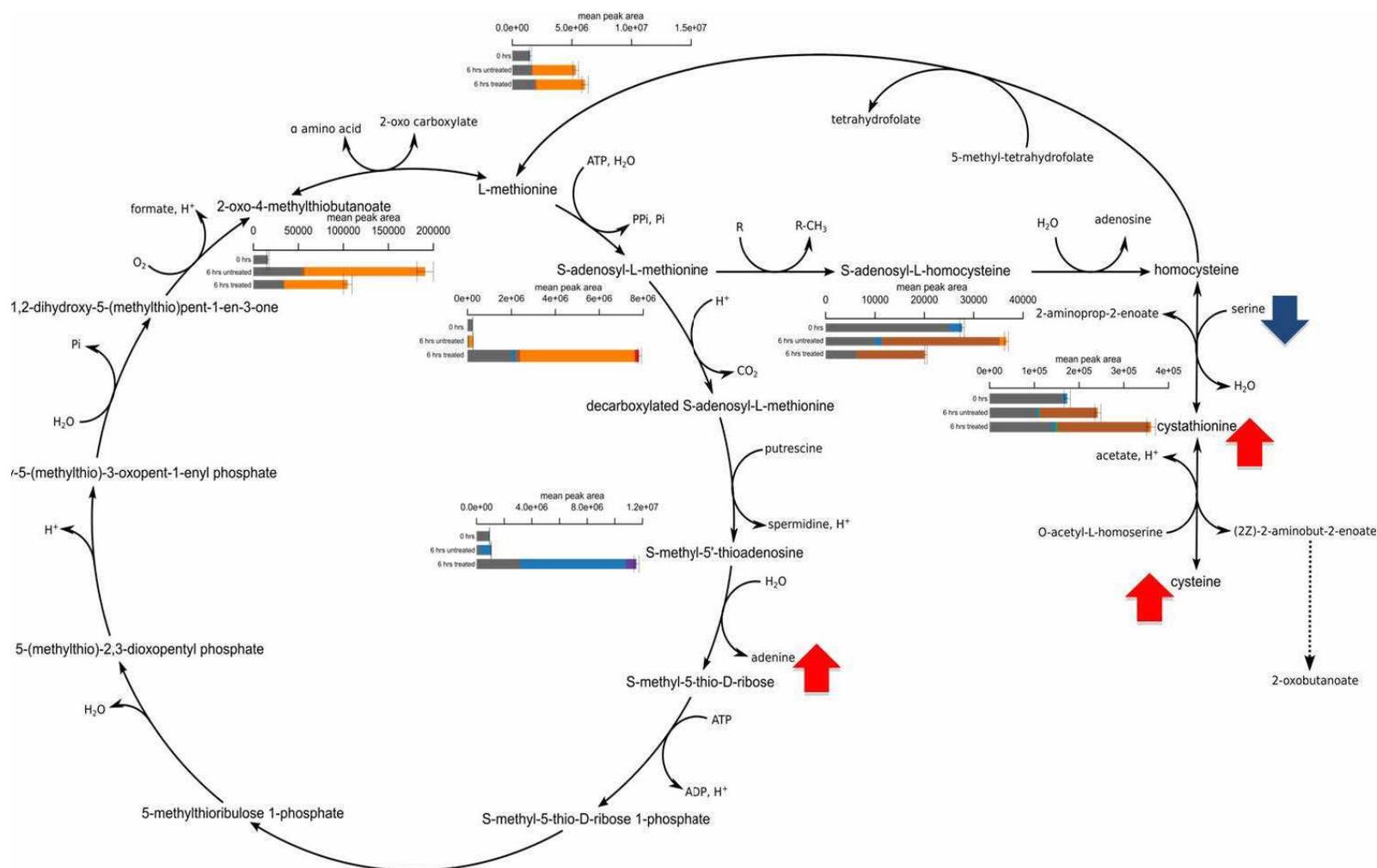
### 3.7.1 Changes in $^{13}\text{C}$ labeling patterns across the metabolome

We first looked at methionine metabolism in detail (fig. 3-21). After 6 hours post-treatment, we saw similar changes to those in previous metabolomics experiments. In particular, the high increases seen in AdoMet ( $m/z$ : 298.1370, RT: 13.49 mins) and 5'-MTA ( $m/z$ : 297.0892, RT: 13.50 mins) were apparent. This highlighted the reproducibility of the results, and helped to confirm these changes are real, and significant. However, L-methionine itself ( $m/z$ : 149.0512, RT: 10.92 mins) was not increased in the same manner as recorded previously (fig. 3-21). This could be due to the change in culture mediums, although at 200  $\mu\text{M}$ , HMI-9 contains similar amounts of L-methionine as CMM, excluding the serum contents. The other result that was highly reproducible was the decrease observed in AdoHcy ( $m/z$ : 384.1222, RT: 12.08 mins).

Using the mzMatch-ISO package (Chokkathukalam *et al.*, 2013), we next analysed the patterns of  $^{13}\text{C}$  distribution, and how these isotopes disseminated through the metabolome after addition of  $^{13}\text{C}$ -(U)-L-methionine (fig. 3-21). After 6 hours, there were no changes in labelling distribution in metabolites involved in methionine degradation. In addition, whilst the culture medium contained 100%  $^{13}\text{C}$ -(U)-L-methionine, intracellular L-methionine was not 100% labelled after 6 hours (fig. 3-21), suggesting uptake is slow compared to, for example, glucose, which shows 100% labelling within minutes post-incubation (Creek *et al.*, 2015).

Both AdoMet and 5'-MTA showed approximately 75% labelling in the presence and absence of AN5568 (fig. 3-21), whilst AdoHcy and L-cystathionine were approximately 50% labelled.  $^{13}\text{C}$  isotopes were also used in the methylation reactions of L-lysine, an important post-translational modification, discussed in detail below. In contrast, we could not find any labelling of arginine after 6 hours in the presence, or absence of AN5568, suggesting the methylation of this amino acid is a much slower process compared to lysine methylation. Arginine methylation has been shown to originate from AdoMet and thus, L-methionine (F. Achcar, personal communication).

The labelling data presented here indicate that there is no specific point in the metabolome where carbon distribution originating from L-methionine is perturbed. In particular, AdoHcy is still labelled after AN5568-treatment, which suggests that methyltransferase reactions are still occurring, albeit at a potentially lower overall rate. Indeed, this could highlight the specificity of the AN5568 protein target. This is supported by the observation of methylated lysines. An in depth study of methylation patterns in the proteome could potentially indicate more specific changes, should a particular enzyme be inhibited. This type of study has recently been completed for the *S. cerevisiae* proteome, where a significant number of previously unannotated methylation events were uncovered (Wang *et al.*, 2015).



**Figure 3-21: Tracing  $^{13}\text{C}$  distribution in cells incubated with  $^{13}\text{C}$ -(U)-L-methionine and treated with AN5568.** Cells were incubated with  $200\ \mu\text{M}$   $^{13}\text{C}$ -(U)-L-methionine for 6 hours. In addition, one sample group was incubated with  $10\times\ \text{EC}_{50}$  AN5568. Firstly, methionine metabolism was analysed. After 6 hours, there were still significant increases in the metabolites previously seen to increase. In addition, intracellular L-methionine was not 100% labelled, suggesting that uptake is relatively slow compared to, for example, glucose uptake. Interestingly, whilst only small amounts of AdoMet and 5'-MTA were detected in WT samples, they showed almost 100% labelling, suggesting these metabolites all originate from L-methionine. In contrast, these metabolites were only ~75% labelled in drug-treated samples, suggesting they might be recycled, or the build-up initiates very quickly after drug treatment. AdoHcy was also detected, with 4  $^{13}\text{C}$  labels, as expected. In addition, L-cystathionine had 4  $^{13}\text{C}$  labels, and was found to increase, as seen previously. This means that MTase reactions are still occurring, but perhaps only very specific reactions are being affected. This would be masked in these data. Arrows indicate metabolites that were found to change in abundance, but were not labelled. Where no arrow or graph is shown, the metabolite was not detected.

### 3.7.2 Lysine is methylated in an AdoMet-dependent fashion, but is not involved in carnitine biosynthesis

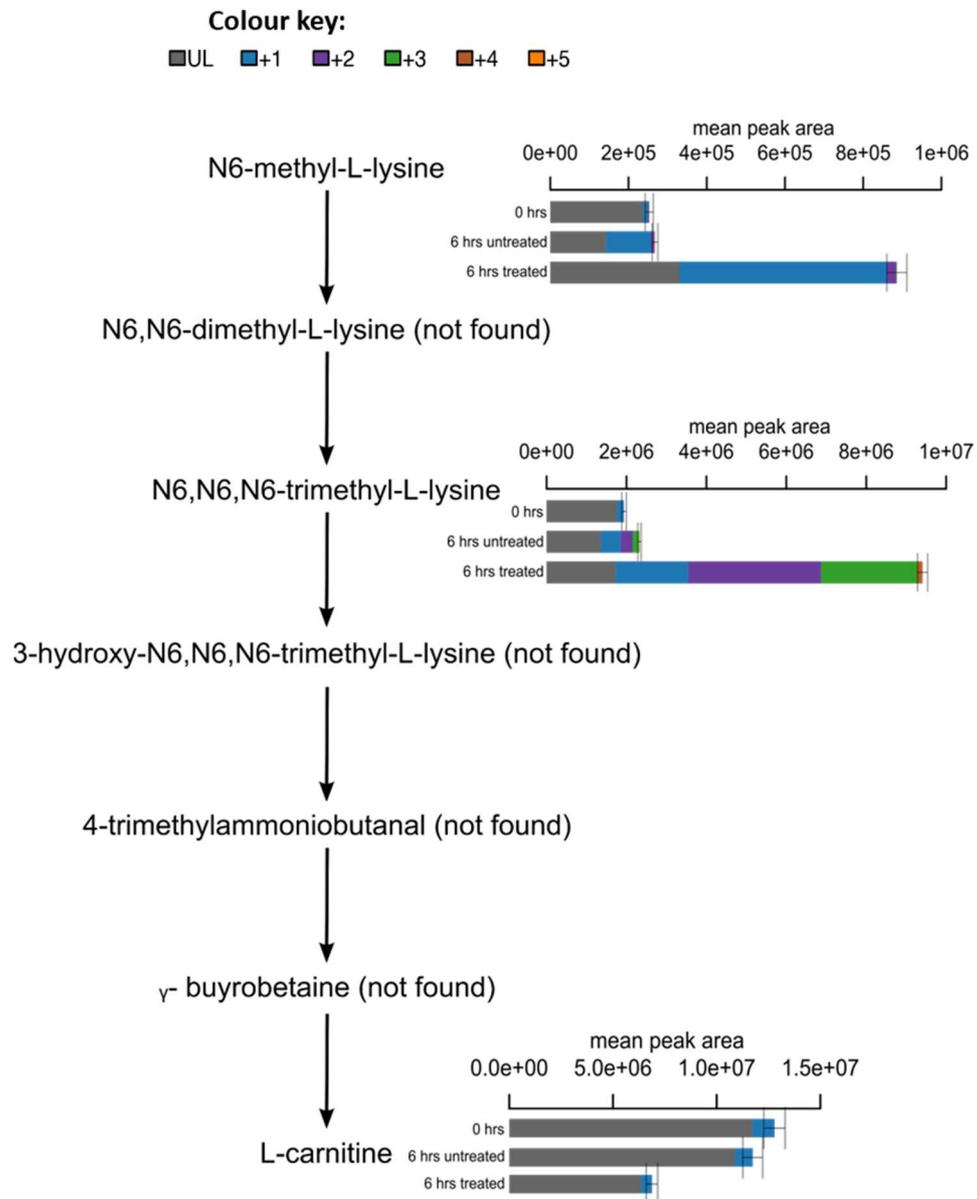
An interesting result regarding lysine methylation was uncovered from the dataset (fig. 3-22). Lysine methylation plays an important biological role in all living organisms, including *T. brucei* as a post-translational modification (Janzen *et al.*, 2006). In the metabolomics dataset, both mono-methylated lysine and tri-methylated lysine were found to be highly increased after AN5568 treatment. Furthermore, a large percentage of these metabolites were labelled with heavy carbon isotopes, suggesting that the methylation is AdoMet dependent (fig. 3-22).

In mammalian cells, methylation of lysine is part of the carnitine biosynthesis pathway, which is also referred to as the L-lysine degradation pathway (Hoppel *et al.*, 1980). Carnitine is an important molecule used to transport fatty acids across the mitochondrial membrane (Klein *et al.*, 1982). Therefore, we looked at L-carnitine and its derivatives in AN5568-treated parasites. Interestingly, we did not find the same methylation pattern in L-carnitine (fig. 3-22). This suggests that the L-carnitine biosynthesis pathway does not exist in *T. brucei*. Indeed, previous studies have suggested that the parasite obtains this metabolite exclusively from its external environment (Gilbert & Klein, 1982, Klein *et al.*, 1982). The results shown here support these previous conclusions.

Given these results, it is important to understand the reasons behind the increased levels of free methylated lysine residues. These modifications are crucial in the activation of proteins, for example, on histones (Janzen *et al.*, 2006, Gassen *et al.*, 2012). It is also possible that these lysines are involved in the synthesis of tRNA, which is known to involve many complex PTMs (Swinehart & Jackman, 2015). To understand the origin of these methylated lysines, it would be necessary to look at PTMs on a proteomic level, allowing the analysis of methylation changes across the proteome. This is possible using mass spectrometry technologies as well as SILAC based approaches (Olsen *et al.*, 2016). Another method that could be used to look at lysine methylation patterns are to use commercial antibodies and carry out a Western blot. Whilst this was attempted during this study, we were not able to optimise these in time for the completion of the project. This would be an interesting starting-point to analyse changes in lysine methylation upon AN5568-treatment.

Interestingly, whilst L-carnitine does not seem to originate from lysine methylation in *T. brucei*, there was a decrease in the abundance of this metabolite after AN5568 treatment (0.66-fold change relative to untreated). This is an important point of discussion because L-carnitine is crucial in aiding the transport of long chain acyl groups from fatty acids into the mitochondrial matrix (Gilbert & Klein, 1982). Given the data mentioned above, the decreased levels of L-carnitine support the theory that there are significant changes in the mitochondrion upon

AN5568 treatment. In addition, the observed decreases in lipid metabolism in general also correlate with L-carnitine downregulation.



**Figure 3-22: AN5568 causes significant changes in methylated lysine metabolism.** AN5568-treated cells were observed to have highly increased levels of both mono-methylated and tri-methylated L-lysine. This is a common, and highly conserved, posttranslational modification that has been shown to play roles in a myriad of cellular processes. Interestingly, it was also observed that much of the methylation originated from L-methionine in MTase reactions. In mammalian cells, methyl-L-lysine degradation leads to the formation of L-carnitine, which was previously found to decrease after AN5568 treatment. However, we could not find any evidence of  $^{13}\text{C}$ -labeled L-carnitine, nor its derivatives. Indeed, previous studies have suggested that L-carnitine is solely obtained through uptake in *T. brucei*, and the results presented here support this theory.

### 3.8 Knock-down of TbCgm1, a splice-leader methyltransferase

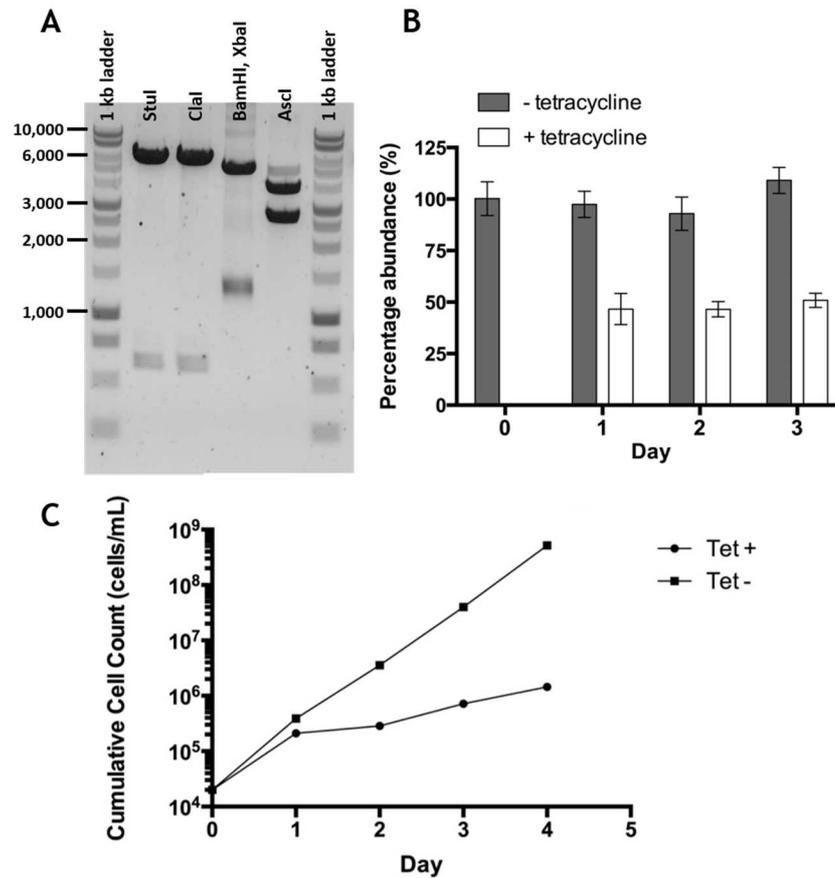
An intense literature search for important and unique mechanisms of methylation in the trypanosomatids, led to the hypothesis that spliced leader methylation might be targeted by the benzoxaborole. As described in the introduction, spliced leader methylation is uniquely

complex in the tri-tryps, in comparison to other eukaryotes (Gunzl, 2010). A 39-nucleotide sequence is spliced to every mRNA transcribed, and we hypothesised that this process requires a significant quantity of AdoMet as the methyl-group supplying substrate. Indeed, this has led to previous discussions that the spliced leader complex machinery could be targeted by therapeutics against the parasite (Takagi *et al.*, 2007).

Previous studies have identified four MTases involved in methylation of the spliced leader cap. TbCgm1 (Tb927.7.2080) methylates the inverted guanosine cap, and is thought to be essential in BSF trypanosomes (Hall & Ho, 2006a, Ruan *et al.*, 2007a, Takagi *et al.*, 2007). TbMtr1, TbMtr2 and TbMtr3 (Tb927.10.7940, Tb927.11.4890 and Tb927.9.12.040 respectively) are 2'-O-ribose MTases that methylate the first 4 bases of the spliced leader cap (Zamudio *et al.*, 2007, Mittra *et al.*, 2008, Zamudio *et al.*, 2009). Importantly, several studies have found these to be non-essential. Whilst methylation on the corresponding bases is abolished, there are no growth phenotypes associated with the knock-down of these 2'-O-ribose MTases (Zamudio *et al.*, 2006).

Given the aforementioned findings, we chose to attempt an RNAi-mediated knock-down of the TbCgm1 MTase, and comparatively analyse the metabolic profile of this knock-down to that of AN5568-treated cells. For this experiment, the 2T1 cell line was used (Alsford *et al.*, 2011), and a construct targeting this enzyme was made using the pGL2084 plasmids as described in [section 2.4.4](#). Integration of the construct into the vector was confirmed by a series of restriction digests (fig. 3-23A).

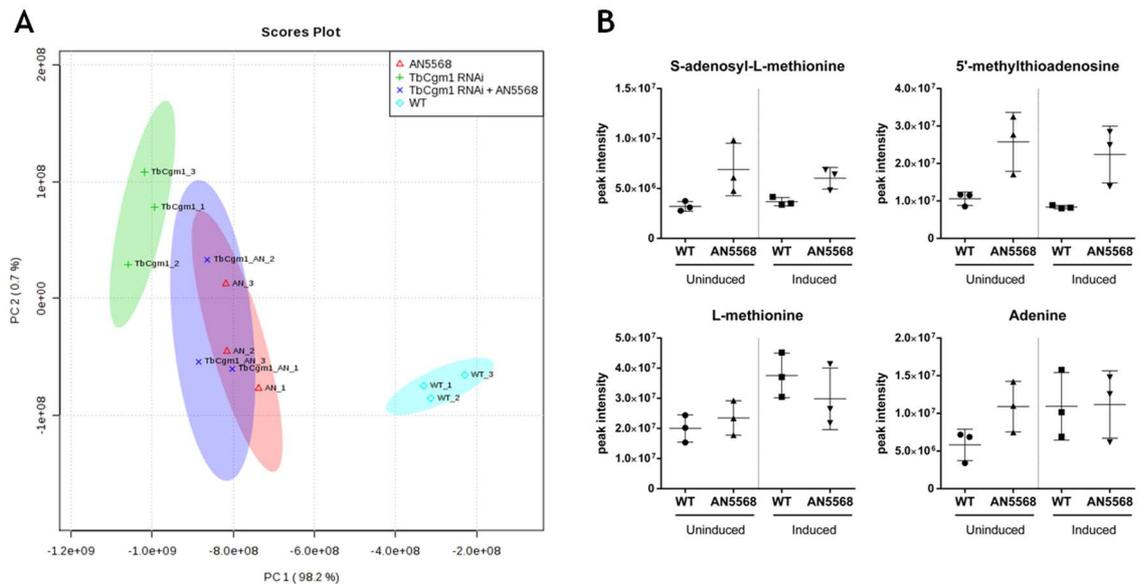
A successful transfectant was isolated and characterised. qPCR analysis showed a 50% reduction in TbCgm1 transcript abundance by 24 hours (fig 3-23B). Consecutive days of tetracycline-mediated RNAi induction showed that 50% knock down was the highest amount possible (3-23B). In addition, growth analysis of the RNAi-induced cell line, compared to a culture grown in the absence of tetracycline, showed the appearance of a growth phenotype by 48 hours, with a significant inhibition of cellular growth (fig. 3-23C). Thus, we concluded that TbCgm1 protein to be essential in BSF *T. brucei* parasites.



**Figure 3-23: Generation of a  $TbCgm^{RNAi}$  cell line.** An RNAi knock-down was generated using the 2T1 cell line (Alford *et al.*, 2011). A) A combination of restriction digests confirmed the insertion of two identical gene constructs into the pGL2084 plasmid. B) RNA was isolated every 24 hours over a period of 3 days post-induction. qPCRs were carried out on both uninduced and induced  $TbCgm1^{RNAi}$  cells, which confirmed a 50% knock-down of the  $TbCgm1$  gene. C) Comparative growth analysis of induced and uninduced cells showed a significant growth defect in the presence of tetracycline, thereby proving that the  $TbCgm1$  gene was indeed essential to BSF trypanosomes.

The next step was to carry out untargeted metabolomics of the  $TbCgm1^{RNAi}$  line and compare the RNAi induction metabolic phenotype to that of AN5568 cells. Tetracycline-mediated RNAi-induction was done for 36 hours, before an uninduced culture was treated with 10× EC<sub>50</sub> benzoxaborole for 6 hours. In addition, an uninduced sample containing DMSO, as well as an RNAi-induced sample treated with AN5568 were also prepared. Samples were then extracted and run on a mass spectrometer as described in the methods. Data was processed using the mzMatch/Ideom pipeline.

The triplicate samples in the dataset were compared first by analysis of variance using a PCA plot (fig. 3-24A). As expected, both drug-treated samples correlated closely. This also suggests that the effect of the drug is much stronger than the phenotype produced by RNAi-mediated knock-down of  $TbCgm1$ . However, the RNAi-induced cell line by itself did not correlate very well with either drug-treated sample, suggesting that the overall metabolotypes were different.

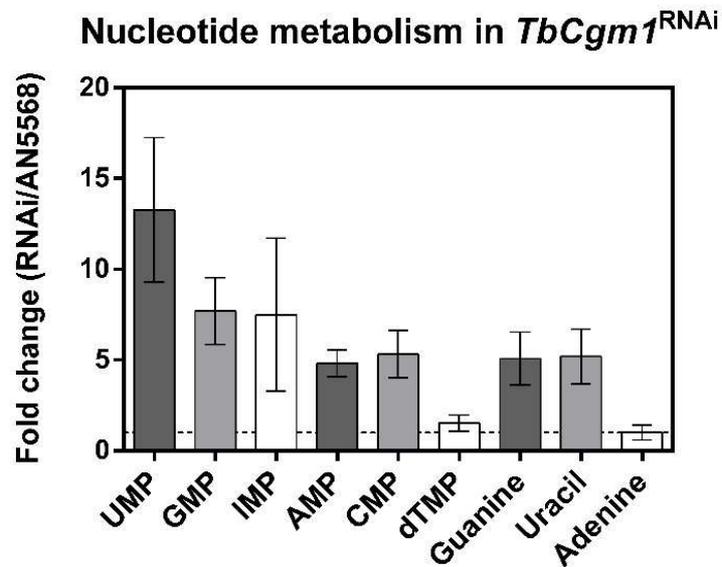


**Figure 3-24: Metabolomics analysis of the *TbCgmRNAi T. brucei* cell line.** Metabolite extractions were carried out on uninduced and induced cells, both in the presence and absence of 10× EC<sub>50</sub> concentrations of AN5568 (6 hour incubation). Interference was induced for a total of 36 hours before extractions were carried out. A) A PCA plot of the four sample groups showed that both the drug-treated sample groups localised closely together, whilst good separation was achieved from the uninduced cell line. There was no significant overlap of the induced RNAi line in the absence of drug with any drug treated samples, suggesting that the metabolic phenotypes of these cells were somewhat different from those treated with AN5568. Moreover, this indicated that the drug treatment caused a much stronger phenotype than RNAi-mediated knock-down of *TbCgm1*. Legend: WT: *TbCgm1*<sup>RNAi</sup> in the absence of tetracycline; AN: AN5568-treated *TbCgm1*<sup>RNAi</sup> cells in the absence of tetracycline; *TbCgm1*: *TbCgm1*<sup>RNAi</sup> in the presence of tetracycline; *TbCgm1*\_AN: AN5568-treated *TbCgm1*<sup>RNAi</sup> cells in the presence of tetracycline. B) Some of the candidate metabolites were analysed using the mzMatch/Ideom pipeline. Both AdoMet and 5'-MTA were not found to be altered in RNAi-induced cells, indicating that the *TbCgm1* protein was most likely not the target. However, *TbCgm1* knock-down did lead to increased levels of L-methionine and adenine, most likely due to perturbations in nucleotide metabolism, as well as ablation of protein synthesis, both of which are known to occur as a result of spliced leader methylation inhibition.

We next analysed metabolites independently, to assess the metabolic changes occurring during *TbCgm1* knock-down, and compare these changes to those occurring in drug-treated cells. Specifically, we targeted the changes that were known to occur during AN5568 treatment (fig. 3-24B). There were no significant changes in either AdoMet levels, or 5'-MTA levels, in the *TbCgm1* knock-down (fig. 3-24B). However, interestingly, there was a large increase in L-methionine, as well as adenine, both of which were seen to increase in AN5568-treated cells as well.

The reason for these changes might well vary from why they occur in AN5568-treated cells. Inhibition of the *TbCgm1* MTase is known to abolish the formation of mature mRNAs, and therefore, protein translation (Hall & Ho, 2006a, Takagi *et al.*, 2007). This has several knock-on effects in the metabolome. Firstly, inhibition of protein synthesis could explain high levels of L-methionine, which is no longer incorporation into peptide chains. Hypothetically, there would be significant changes in nucleotide metabolism in general, should both transcription and translation cease to function.

Indeed, when we looked closer at nucleotide metabolism in the *TbCgm1<sup>RNAi</sup>* line and compared it to AN5568-treated cells, we found large increases in almost every purine and pyrimidine that could be detected (fig. 3-25). The majority of these do not change in AN5568-treated cells, and therefore, this led to the conclusion that the MTase *TbCgm1* cannot be the protein target of the benzoxaborole. However, the metabolic perturbations that result from spliced leader methylation inhibition are intriguing in their own right, and should be studied further.



**Figure 3-25:** Comparative analysis of changes in nucleotide metabolism in the *TbCgm1<sup>RNAi</sup>* cell line and AN5568-treated cells. The *TbCgm1<sup>RNAi</sup>* cells showed high increases in several metabolites associated with nucleotide metabolism, especially in comparison with cell treated with AN5568 alone.

The data from the *TbCgm1* knock-down were important to analyse, especially because it allowed us to understand whether the changes seen in AN5568-treated cells are indicative of spliced leader methylation inhibition. Given the data, we concluded that spliced leader methylation is not the target of the benzoxaborole. In particular, the lack of AdoMet and 5'-MTA increase in *TbCgm1<sup>RNAi</sup>* is intriguing, because we had hypothesised these pathways would have been severely affected.

There were other candidates under consideration for RNAi experiments. Recent publications have given further insight into the activation pathways associated with spliced leader function (Zamudio *et al.*, 2009). In addition, there were other MTases that could have been worth knocking-down to observe metabolic changes. However, given the technical demands of developing these lines, and the costs associated with repeated mass spectrometry experiments, it was decided to search for other methods that could probe the possible mode of action of AN5568 further.

### 3.9 ER stress leads to metabolic perturbations also seen in AN5568 treatment

---

Further literature searches with the metabolomics results in mind led to another hypothesis centred on ER stress induction. As discussed previously, ER stress and protein unfolding predominantly leads to a conserved unfolding protein response (Smith *et al.*, 2009, Goldshmidt *et al.*, 2010). We hypothesised that ablation of transcription could lead to increased levels of AdoMet, due to the absence of spliced leader formation, and therefore, methylation. To observe the metabolic changes occurring as a result of ER stress, and more importantly, to compare these changes to those occurring in AN5568-treated cells, another metabolomics experiment was set-up, using the ER stress inducer dithiothreitol (DTT). This compound contains two S-H groups. It is a very strong reducing agent that acts by attacking disulphide bridges in two sequential thiol-disulphide exchange reactions which leave a 6-membered ring (Tiengwe *et al.*, 2015). By attacking disulphide bridges, DTT disrupts protein folding, leading to the conserved response described above.

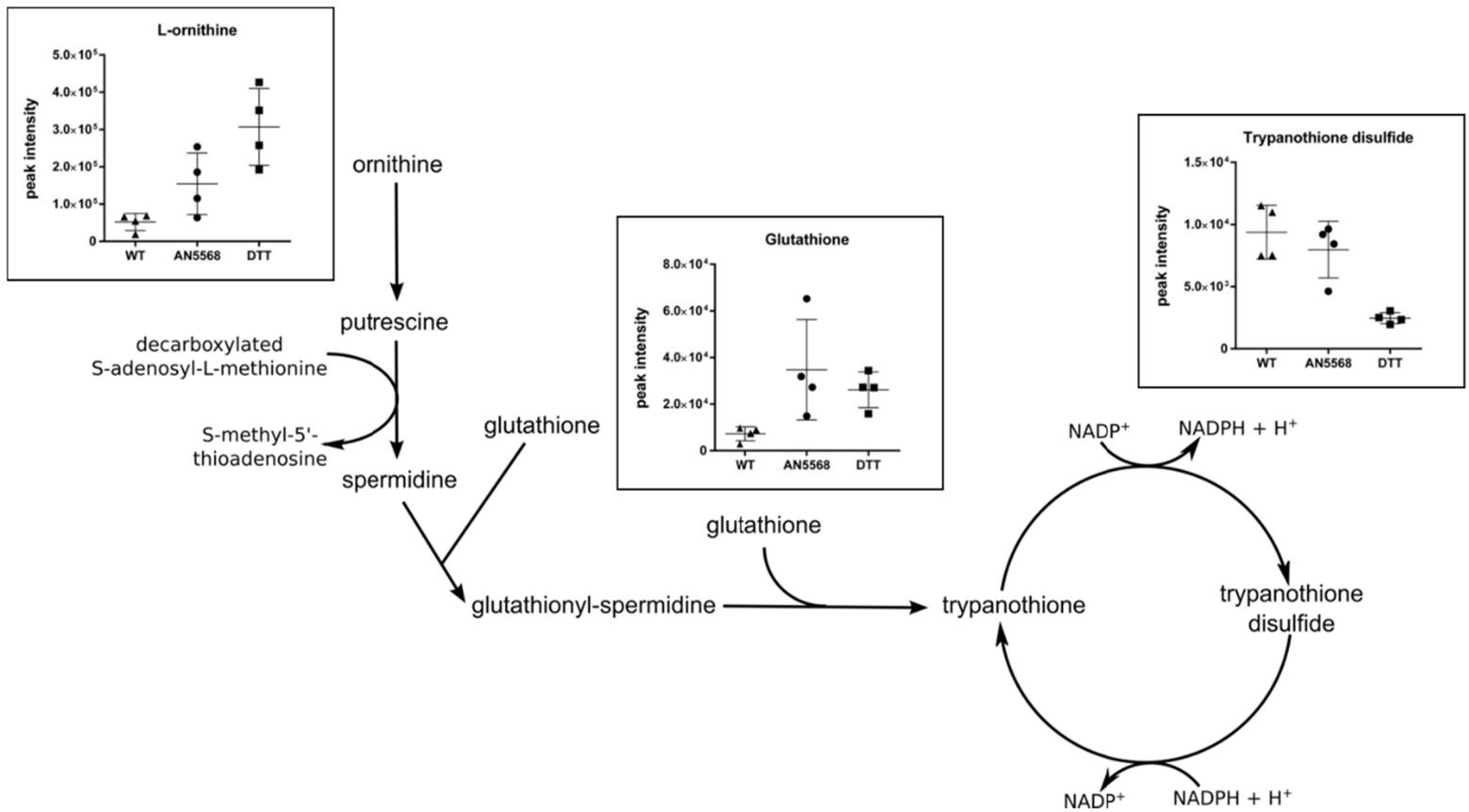
#### 3.9.1 ER stress results in upregulation of the trypanothione biosynthesis pathway

Analyses of the dataset revealed that several components of the trypanothione biosynthesis pathway, which is involved in the oxidative stress response, were upregulated in DTT-treated cells. Both ornithine and glutathione were increased in similar fashion to AN5568-treated cells, suggesting this response was not specific to benzoxaborole treatment (fig. 3-26). In addition, trypanothione disulphide, which consists of two trypanothione molecules bound by a disulphide bridge, was decreased in DTT-treated cells, to a much greater extent than in those treated with AN5568 (fig. 3-26). This was most likely due to the effect of DTT in disulphide bridge disruption.

Whilst lower levels of trypanothione disulphide were unsurprising, upregulation of the oxidative stress response pathway in the presence of a reducing agent was unexpected. It is possible that this upregulation was due to the need for the parasite to synthesise trypanothione disulphide, to counter the effect of the stress inducer.

**3.9.2 AdoMet increase is specific to AN5568 treatment**

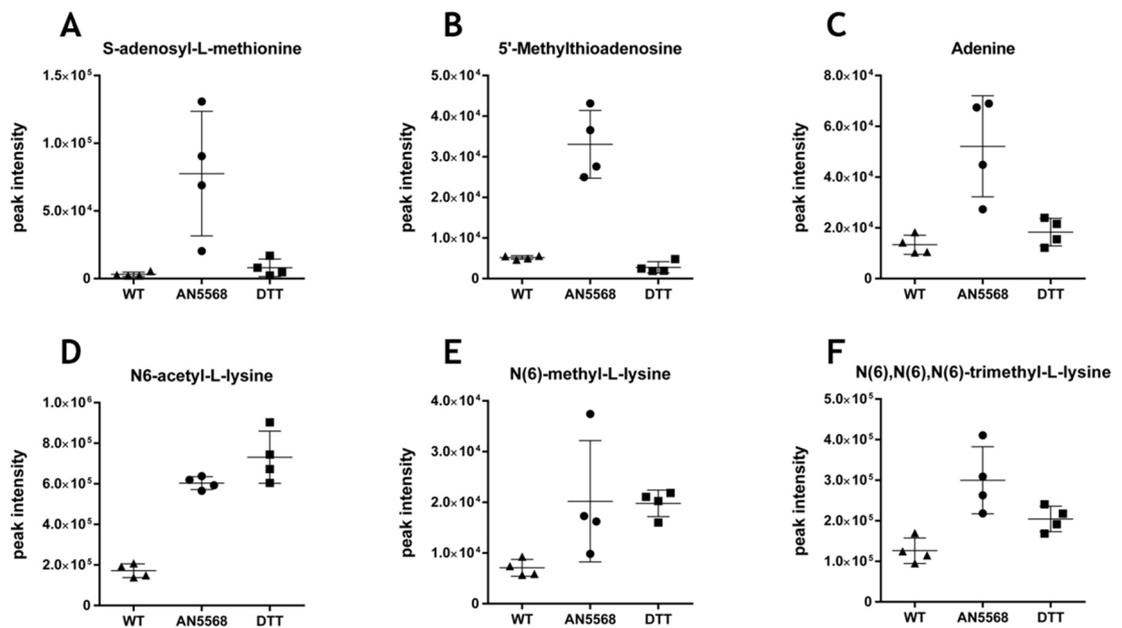
Whilst several similarities were found between the metabolic profiles of DTT- and AN5568-treated cells, the most significant changes in benzoxaborole treated cells were not observed in the DTT sample group. In particular, AdoMet, 5'-MTA and adenine, which were again highly



**Figure 3-26: Metabolic changes in the trypanothione biosynthesis pathway.** There were similar perturbations observed between AN5568- and DTT-treated cells. L-ornithine, an important precursor for the generation of spermidine, was increased in both sample groups, compared to untreated cells, as was glutathione. Trypanothione disulfide did not show much change in AN5568-treated cells. However, this metabolite was significantly reduced in DTT-treated cells, most likely due to the reducing nature of the compound, which breaks disulfide bonds in a non-specific manner.

increased in AN5568-treated cells, did not show changes in DTT-treated parasites (fig. 3-26). This suggests that the changes occurring in methionine metabolism are not due to the conserved SLS pathway response.

However, other changes first observed in cells treated with the benzoxaborole were shown to occur in DTT-treated cells as well (fig. 3-27). In addition to the aforementioned upregulation of the trypanothione biosynthesis pathway, modified lysine residues were also sharply increased in DTT-treated cells (fig. 3-27)



**Figure 3-27: Changes seen in methionine metabolism are specific to AN5568-treatment, but modified lysine residues are not.** The increases in AdoMet (A), 5'-MTA (B) and adenine (C) which are the most significant metabolic phenotypes in AN5568-treated cells, do not occur after DTT treatment, suggesting these changes are specific to the MoA of the benzoxaborole. Conversely, modified lysine residues such as N6-acetyl-L-lysine (D), N6-methyl-L-lysine (E) and N6,N6,N6-tri-methyl-L-lysine (F) were also found to increase in abundance after DTT treatment. This implies that these changes could be related to cellular stress responses.

The results certainly help to understand how conserved the metabolic changes are in response to the benzoxaborole. It would appear as though lysine modifications are not related to the AN5568 mode of action in particular, but could form part of a more generalised cellular stress response. Indeed, this posttranslational modification has been implicated in many cellular functions including protein degradation, gene expression, protein activation and regulation of protein-protein interactions. This is reflected by the significant number of lysine MTases that have been characterised in both humans and yeast.

Whilst these data provided a useful insight into the non-specific changes occurring in AN5568-treatment, the DTT dataset is also crucial to our understanding of trypanosome biology in its own right. Further analyses will be carried out to unravel the metabolic changes associated with ER stress.

### 3.10 Using DARTS to probe the putative mode of action of AN5568

---

A vital area of NTD research is drug target identification and validation. As discussed in chapter 1, investigations in this field often take one of two approaches: target-based approaches and phenotype screening approaches (Reguera *et al.*, 2014). For reasons described in details in chapter 1, the approach taken by Anacor in the context of HAT drug discovery was phenotypic screen based on the benzoxaborole 6-carboxamides class of boron-containing compounds. Whilst this led to the identification of AN5568 as a lead compound, the screen did not identify the protein target of the parasite.

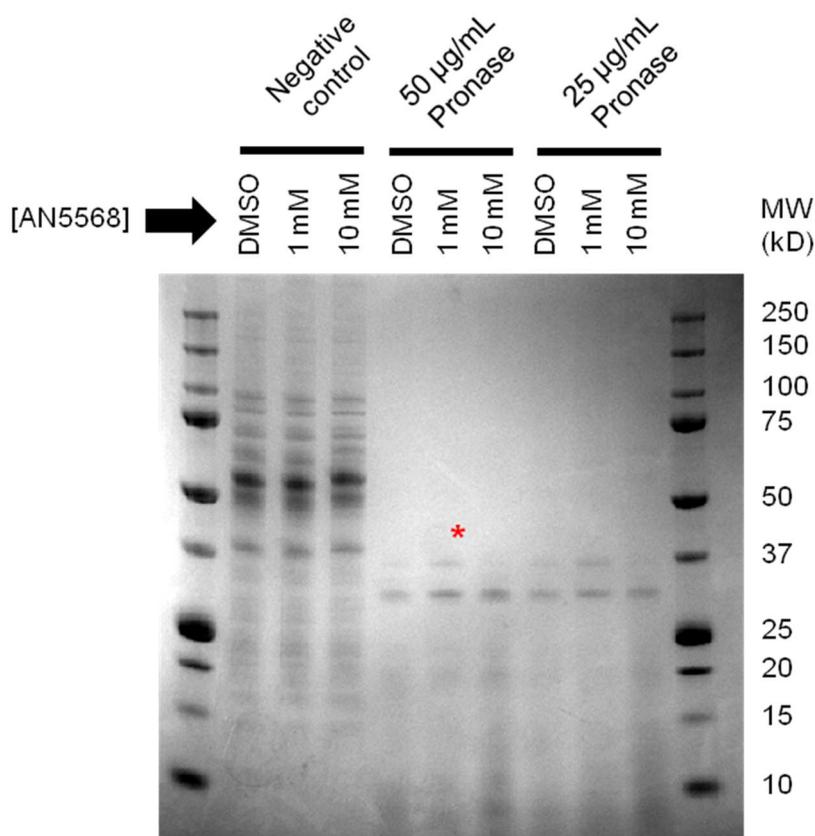
There are a variety of methods currently used to investigate drug mode of action, and more specifically, for drug target identification, as described in [section 1.10](#). The most commonly used method is affinity chromatography. This method involves immobilization of the drug to beads. These beads are added to a column through which a protein lysate pull-down can be carried out, thereby allowing the protein target to bind to the compound of interest (Annis *et al.*, 2007). However, immobilisation of drugs becomes problematic if there is a lack of knowledge concerning the compound structure-activity relationship (SAR) (Lomenick *et al.*, 2009). Novel experimental approaches, published recently, have attempted to utilise compounds without the necessity of immobilising them. One example published by Lomenick and colleagues, is based on the theory that proteins have an increased thermodynamic stability when bound to ligands (Lomenick *et al.*, 2009, Pai *et al.*, 2015). By subjecting a control protein lysate as well as a drug treated lysate to the same amount of proteolysis, Lomenick and colleagues argue that target proteins would show a lesser degree of disintegration in the drug-treated sample (Lomenick *et al.*, 2009).

This process has been shown to work by the group themselves, and several other publications have shown a moderate success rate (Lomenick *et al.*, 2011b, Chin *et al.*, 2014). We therefore attempted to use this assay in *T. brucei* to see if it would be possible to find any proteins of interest with AN5568. Optimisation was required for the final concentration of the benzoxaborole, as well as the protease concentrations added to the samples and their incubation time. The final conditions that were chosen are outlined in the methods section ([section 2.7.1](#))

Subsequent to optimisation, the DARTS experiment was run using final concentrations of both 1 mM and 10 mM AN5568. The protein lysates were extracted using trypanosome lysis buffer, rather than M-PER lysis buffer as suggested by the original publications (Pai *et al.*, 2015). Samples were incubated for 1 hour and pronase concentrations were added as described in [section 2.7.1](#). The pronase-treated lysates were then run on a protein gel, and a Coomassie stain showed a band of increased intensity in the 1 mM AN5568 sample (fig. 3-28). This band,

as well as the corresponding band in the DMSO control, were extracted from the gel and gel extraction, as well as mass spectrometry, was carried out by Glasgow Polyomics.

Interestingly, we did not see the same enrichment when cells were treated with 10 mM AN5568, though this was probably because the drug precipitated at this concentration, removing a large number of proteins from the sample. When concentrations of pronase lower than 25  $\mu\text{g}/\text{mL}$  were used, there were many more bands detected (data not shown), and therefore, we determined both 50  $\mu\text{g}/\text{mL}$  and 25  $\mu\text{g}/\text{mL}$  to be optimal for the DARTS assay.



**Figure 3-28: Protein samples subjected to DARTS assay.** The soluble protein content was isolated from *T. brucei* cells and divided into three samples. These were incubated with either AN5568 or DMSO alone and subsequently divided into further sample groups for pronase degradation. Samples were subsequently analysed on a protein gel, and an enriched band was seen at approximately 37 kDa, in the sample treated with 1 mM AN5568 (red star). This band was excised from the gel, along with the corresponding control band, and sent for proteomics analysis at Glasgow Polyomics.

Sample	GeneID	Description	Sequence coverage	Peptide	Delta	E-value
DMSO	Tb.927.10.5620	Fructose-bisphosphate aldolase (ALD)	13.70%	TGETFPQYLR	-0.1888	0.0054
				ATAEQVAEYTVK	-0.1247	0.6000
				GEQMTAGLDGYIK	-0.2622	0.0260
	Tb10.v4.0241	Variant surface glycoprotein (VSG), putative	2.70%	DGRQHRLR	-0.2273	0.7300
				DGRQHRLR	-0.2113	0.1100
				DGRQHRLR	-0.1653	0.0470
	Tb927.2.4590	Branched-chain amino acid aminotransferase, putative	21.50%	DSILSLVR	-0.1578	0.0500
				LHVEEER	-0.1754	0.1000
Tb927.6.3150	Hydin	0.50%	MPMTIDVPAKQK	-0.2372	0.2800	
Tb11.16.0004	Variant surface glycoprotein	3.50%	RIAEALAPK	-0.3600	0.3700	
Tb927.6.1000	Cysteine peptidase, clan A,	1.80%	APAAVDWR	-0.1090	0.4500	
Tb927.4.3470	Hypothetical protein,	1.00%	GSLHCR	0.0047	0.5700	
Tb927.1.3910	TPR repeat	0.90%	AMKQCMQR	-0.0620	0.4700	
AN5568	Tb.927.10.5620	Fructose-bisphosphate aldolase (ALD)	16.70%	TGETFPQYLR	-0.0868	0.2000
				GLLADESTGSCSK	-0.0393	0.0005
				GEQMTAGLDGYIK	-0.1182	0.1900
	Tb927.2.4590	Branched-chain amino acid aminotransferase, putative	16.10%	DSILSLVR	-0.3018	0.0049
				LHVEEER	-0.1734	0.1900
	Tb927.11.5570	DNA replication licensing factor MCM5, putative	0.90%	TYINGGR	-0.1751	0.4500
				TYINGGR	-0.1411	0.2600
TYINGGR				-0.1291	0.0150	
Tb927.1.4010	Primase 2	0.60%	CPQKYSR	-0.3205	0.3400	
Tb927.6.550	Hypothetical protein	7.00%	MLTTSFMVVEPK	0.0740	0.7000	
Tb927.10.3080	Methionine biosynthetic protein, putative	3.00%	STLDEATSLMR	0.3465	0.9500	

**Table 3-3: Proteomics analysis of the enriched band observed in the DARTS assay.** Several peptides were identified in both the control sample and the drug-treated sample. Several of these are known to be highly abundant in the *T. brucei* proteome.

Proteomics analysis of both gel bands showed several protein hits, with several being unique to either the drug-treated, or control band (table 3-3). Both bands contained several sequences that were mapped to a fructose-bisphosphate aldolase (ALD) (ID: Tb927.10.5620). The appearance of ALD in both protein samples is interesting, because one previous study showed this enzyme to be partially protease resistant (Clayton, 1987). This could explain why it does not disappear from the protein lysate after pronase treatment. In addition, studies in *A. thaliana*, a plant model organism, have shown that a chloroplastic homologue of this enzyme is methylated by lysine MTases (Mininno *et al.*, 2012). Whether the same occurs with ALD in *T. brucei* is currently unknown, but the appearance of this protein in both sample groups led us to discard it.

Peptides belonging to branched-chain amino acid aminotransferase (ID: Tb927.2.4590) were also seen in both gel extracts, and therefore not deemed to be of interest. In addition, the control band contained peptides mapped to two different VSGs, as well as a peptide corresponding to a cysteine peptidase (table 3-3). Again, these results were disregarded as the peptides identified in the drug-treated gel extract were the main concern, and result of interest.

Here, there were several uniquely mapped peptides that were not seen in the control sample. Firstly, a DNA replication factor, MCM5 (Tb927.11.5570). The product of this gene has previously been shown to be part of the MCM complex (Kim *et al.*, 2013). This complex is involved in many DNA-specific roles including replication initiation, damage response, transcription and chromatin structure, and is highly conserved across all eukaryotes (reviewed by (Forsburg, 2004). Due to this conservation, it is unlikely that MCM5 is the benzoxaborole target, and this peptide was disregarded. Another protein identified in the drug-treated gel extract was primase 2 (ID: Tb927.1.4010). One previous study was found regarding this protein and in it, the authors showed this protein is targeted to the mitochondrion (Hines & Ray, 2010). In addition, Hines and Ray found that knock-down by RNAi resulted in a loss of kinetoplast (k)DNA (Hines & Ray, 2011). This is certainly interesting, because microscopy analysis presented here indicates ablation of kinetoplast replication in response to AN5568 treatment. In addition, the mitochondria appear to lose their normal linear shape. However, there was no evidence of complete loss of kinetoplast DNA during, or after, AN5568 treatment. Furthermore, the low percentage coverage seen in the DARTS assay could signify that this identification was an artefact and this protein was not investigated further. Another protein identified in the drug-treated sample was a hypothetical protein (Tb927.6.550). No further information could be found regarding this protein, even after running the full sequence through databases such as pfam, interpro, SMART and SCOP. For this reason, this protein was not followed up.

Finally, and by far most interesting was the appearance of a protein thought to be involved in methionine metabolism in the drug-treated band. This gene (ID: Tb927.10.3080) showed 3% sequence coverage in the dataset, but importantly, was not detected in the corresponding control band. We therefore chose to investigate this further through *in vitro* overexpression experiments, which were carried out as shown in [section 3.13](#).

### **3.11 Lysates of treated parasites show enrichment of a protein subset**

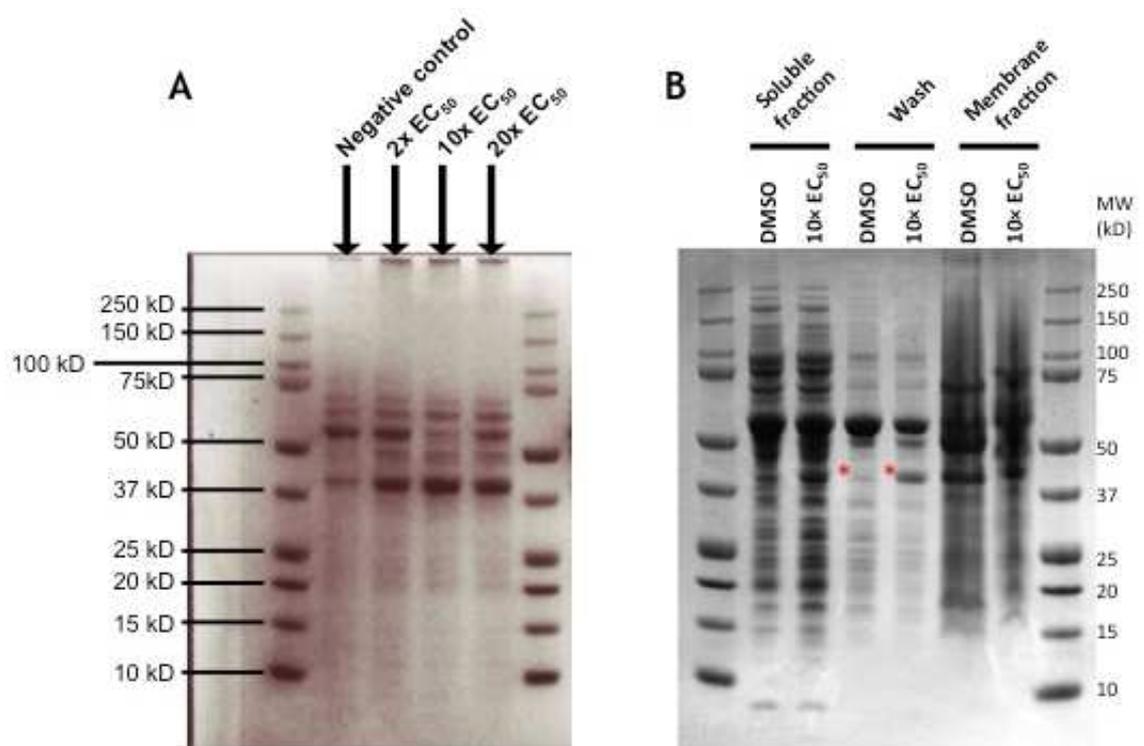
---

Whilst carrying out the DARTS experiment, it was subsequently discovered that incubation of parasites with the benzoxaborole caused a significant change in the protein lysate, when run directly on a protein gel (fig. 3-29A). This phenotype was not seen when extracted lysates were incubated with AN5568 and then run on a protein gel, suggesting that the drug could require activation before causing the metabolic phenotypes seen in earlier experiments.

The gel band that was enriched in the drug-treated protein lysate was found to localise at a molecular weight of just over 37 kDa (fig. 3-29A). Interestingly, this matched the mass of VSG221 in the absence of any glycosylation (Aitchison *et al.*, 2005). However, as western blots found in previous experiments, there was no detectable change in VSG expression after drug

treatment (fig. 3-10). In order to better understand this change in the proteome, new protein lysates were prepared by isolating the soluble and membrane fractions in separate samples. Included with these samples was a further wash fraction containing soluble protein. These samples were then run on a protein gel in similar fashion to previous experiments (fig. 3-29B).

In this experiment, results showed the presence of an enrichment of proteins at the same molecular weight as found previously (fig. 3-29). Interestingly, this enrichment was only seen in the soluble fraction, and to a greater extreme in the wash fraction. As well as concluding that proteomic changes were occurring upon drug-treatment, this experiment also highlighted the reproducibility of this phenomenon in drug-treated protein lysates. To attempt to identify the proteins present in this band, both the drug-treated gel band (red star in fig. 3-29) and the corresponding control band were cut out of the gel. These bands were then sent to Glasgow Polyomics for protein extraction and subsequent proteomics analysis.

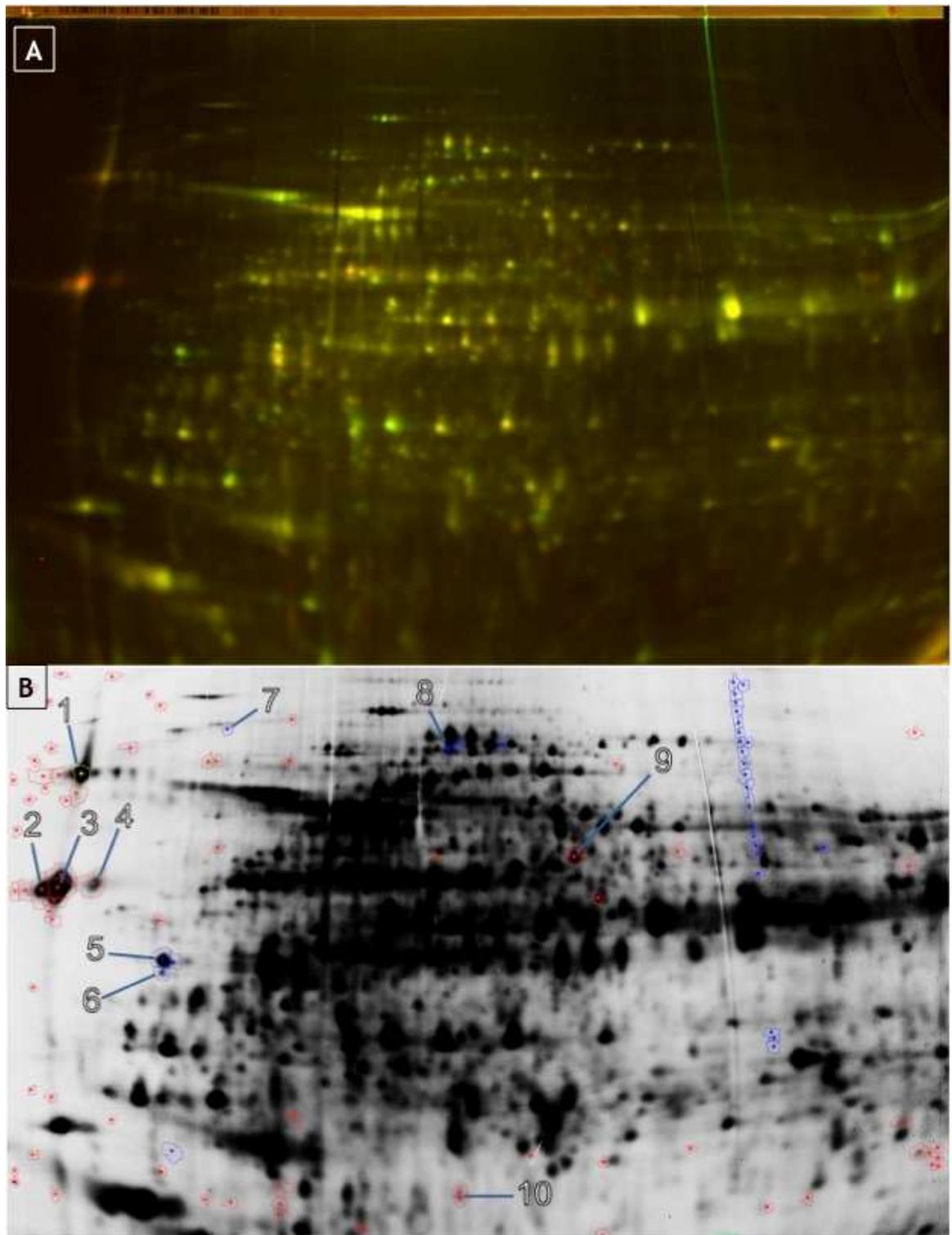


**Figure 3-29: Protein lysates from AN5568-treated parasites visualised on protein gels show differences with wild-type lysates.** A) Cultures of *T. brucei* Lister 427 were treated with varying concentrations of the benzoxaborole for 6 hours. The soluble proteome was then isolated and prepared for protein gel analysis as described in methods (section 2.7.2). Gels were visualised with Coomassie to show an enrichment of proteins in the treated samples at a molecular weight of just above 37 kDa. B) Further analysis of fractionated lysates confirmed the presence of this enriched subset of proteins in both the soluble and wash fractions. It did not appear to be present in the membrane fraction of the lysates.

For both bands, over 60 different proteins were identified from the wash fraction. These results are shown in the supplementary data (S3). This was a much larger dataset than anticipated. Given the nature of proteomics analysis, which allows identification of peptides, but not their amounts, it was also very difficult to assess whether any proteins had changes in their

abundance compared to the control sample. Therefore, rather than filter through this large dataset, analyse individual proteins and select candidates, it was decided to attempt a 2-dimensional difference gel electrophoresis (DiGE) experiment, in order to produce greater separation with the aim of pinpointing the proteins causing visible changes to a higher resolution. This experiment was set up as described in the methods. A control and drug-treated lysate were labelled with Cy5 and Cy3 respectively, and subsequently separated by isoelectric point (IEP). One gel was used for initial analysis of protein shifts, and was visualised using a Typhoon laser scanner (fig. 3-30).

The gel images were next analysed using commercial software, to pick out significant changes between the two conditions. Ten point of interest were chosen, and manually extracted from a second gel containing a combination of lysates from both the control and drug-treated samples, and these were sent once more to Glasgow Polyomics for extraction and mass spectrometry analysis.



**Figure 3-30: Comparative analysis of the soluble proteome from WT and AN5568-treated *T. brucei* cells.** Two protein lysates from a DMSO and AN5568-treated cell culture were tagged with Cy5 and Cy3 respectively and combined on to an IEP separation column. The lysates were then separate by size on an SDS protein gel. A) A merged composite of the DMSO-control (green) and AN5568-treated (red) lysates. Yellow points are those where no change is judged to have occurred post-treatment. In contrast, red and green spots represent proteins that are more abundant in the AN5568-treated sample and the DMSO-control respectively. B) A commercially available software, DeCyder, was used to pick out spots of interest that showed a minimum of two-fold change between control and AN5568-treated lysates. Red contours indicate spots with a higher volume in AN5568-treated cells, whilst blue contours are of higher volume in the control lysate. Ten spots (indicated by number) were eventually extracted from the gel and analysed by mass spectrometry at Glasgow Polyomics. The line of blue contours in the top-right of the gel was judged to be a smear and thus, disregarded in the analysis.

Several peptides were identified in each spot, and the top hits, along with their fold changes, are summarised in table 3-4. Several of the spots were found to contain peptides originating from multiple proteins. In particular, spot numbers 7, 8 and 9 matched several peptides. This is most likely due to the fact they were extracted from a portion of the gel that was heavily stained, indicating that the majority of the soluble proteome migrated to this area of the gel based on pI and size. However, for the majority of spots analysed by mass spectrometry, only one or two significant hits were identified.

Most strikingly was the enrichment of both serine and cysteine peptidase precursors in AN5568-treated cells (table 3-4). Spot 1 was identified as a serine peptidase, an enzyme that cleaves peptide bonds, which contains a serine as the nucleophilic amino acid in its active site (Moss *et al.*, 2015). In *T. brucei*, this serine peptidase is part of a tandem array of at least 3 peptidases that have been linked to suramin efficacy (Alsford *et al.*, 2012). Recent analyses of the entire complement of serine peptidases has identified these genes as serine carboxypeptidase III precursors (Moss *et al.*, 2015).

Spots 2, 3 and 4 were all identified as cysteine peptidases, of which there is an array totalling some 11 near-identical genes (Moss *et al.*, 2015). The differences in the spots indicate that they were separated based on pI, which could be due to posttranslational modifications such as phosphorylation or acetylation (Zhu *et al.*, 2005). Current literature on *T. brucei*, as well as other fields of biology, have not linked any of the metabolic phenotypes seen after AN5568 treatment to inhibition of cysteine/serine peptidase activation, and therefore, it is unknown whether these phenomena could be linked.

Amongst the proteins that decreased in abundance after AN5568 was a nascent polypeptide associated complex subunit, which was found at two different molecular weights (spots 5 & 6, table 3-4). This chaperone complex is one of the first cytosolic proteins to contact new polypeptide chains as they emerge from the ribosome and it aids in prevention of inappropriate interactions with signal recognition peptides (Funfschilling & Rospert, 1999). Whilst the complex is not thought to be essential in yeast, there is evidence to suggest it may aid in the targeting of proteins to the mitochondria (Funfschilling & Rospert, 1999). There is scarce literature to explain the potential reasons for downregulation of this complex, although it could be due to reduced protein synthesis in the nucleus requiring targeting to the mitochondria. One study, again carried out in yeast, showed that one of the sub-units is downregulated in response to ubiquitin ligase (Rsp5) under stress conditions (Hiraishi *et al.*, 2009). With this in mind, it is probable that similar mechanisms occur in the kinetoplastids, and the large amounts of cellular stress caused by AN5568 leads to a downregulation of machinery involved with protein

translation. Indeed, given the conclusions by Hiraishi and colleagues, one could suggest that the difference between spots 5 and 6 could be related to ubiquitination.

There was also downregulation of a heat shock protein (Hsp)-70 in response to AN5568-treatment. Interestingly, this particular Hsp has been identified as a glycosomal Hsp (Guther *et al.*, 2014). Whilst Hsps are generally upregulated in response to environmental stress (Bringaud *et al.*, 1995), downregulation of this Hsp70 could suggest decreased metabolism, or deterioration of the glycosome as a result of benzoxaborole treatment. This could be confirmed with western blotting.

Spot number	fold change (AN5568/ctrl)	Significant ID's	Description	Score
1	1.91	Tb927.10.1040	Serine peptidase, Clan SC, family S10 (CBP1)	138
2	3.15	Tb927.6.1000	Cysteine peptidase, Clan CA, family C1, Cathepsin L-like (CP)	132
3	3.47	Tb927.6.1000	Cysteine peptidase, Clan CA, family C1, Cathepsin L-like (CP)	460
4	3.85	Tb927.6.1000	Cysteine peptidase, Clan CA, family C1, Cathepsin L-like (CP)	55
		Tb927.3.5050	60S ribosomal protein L4	46
5	-2.13	Tb927.9.8100	Nascent polypeptide associated complex subunit	315
6	-2.11	Tb927.9.8100	Nascent polypeptide associated complex subunit	49
		Tb927.11.550	Hypothetical protein SCD6.10 (P-body formation)	40
7	-2.19	Tb927.8.4140	Hypothetical protein	38
		Tb927.4.3470	Hypothetical protein	36
8	-2.12	Tb927.11.11330	Heat shock protein (Hsp) 70	1397
		Tb927.11.11290	Heat shock protein (Hsp) 70	423
		Tb927.11.13500	paraflagellar rod component 1	213
9	2.21	Tb927.10.6510	Chaperonin HSP60, mitochondrial precursor	496
		Tb927.8.4970	Paraflagellar rod protein (PFR2)	196
10	3.57	Tb927.10.5390	Hypothetical protein	147
		Tb927.4.1080	V-type ATPase A subunit	112

**Table 3-4: Peptides identified in a 2D-DiGE undergoing significant changes after AN5568-treatment.** Spot number corresponds to those identified in figure 3-30B. Out of the 10 most significant results, 4 were found to be decreased in AN5568-treated cells, with the rest showing increased abundance. In some cases, more than the first significant hit were included in the analysis, if the overall score was relatively low compared to the other hits in the peptide sample. Domain analyses were carried out for the hypothetical proteins, using the Interpro, Pfam and SMART databases.

Finally, two more spots with increased volume in AN5568-treated cells were identified as mitochondrial chaperonin HSP60 precursor, and a hypothetical protein (spots 9 and 10 respectively, table 3-4).

Results from the DiGE experiment, whilst not as conclusive as was initially hoped, do provide an interesting insight into the proteomic changes occurring in response to AN5568. Interestingly, a knock-down of ATG4.2, a cysteine peptidase in *Leishmania major*, showed a remarkably similar mitochondrial phenotype to *T. brucei* cells incubated with AN5568 (Williams

*et al.*, 2013). In this paper, Williams and colleagues found that the mitochondria exhibited a “patchy” phenotype after an RNAi-mediated ATG4.2 knock-down. However, the orthologue of this protein in *T. brucei* is Tb927.6.1690, not the CA cysteine peptidase identified in the DiGE experiment.

Unfortunately, the experiments shown here were carried out late on during this project, and hence, more data and planning is necessary in order to follow up on the leads identified with the 2D DiGE experiment. This work is currently ongoing in the Barrett group.

The main hypotheses on AN5568 mode of action, up to this point, were as follows: the compound causes significant, yet specific changes in cellular metabolism, characterised by a sharp increase in AdoMet, 5'-MTA and adenine. Further significant changes in 8-amino-7-oxononanoate, AdoHcy, L-cystathionine and lysine methylation support the hypothesis that methyl metabolism is targeted. Additional data from sinefungin-treatment showed overlap with AN5568. With this knowledge, it was decided to explore the *T. brucei* complement of MTase, in order to obtain a better picture of the repertoire of these enzymes and with it, possible ideas on AN5568 targets.

### **3.12 Analysis of the *T. brucei* complement of methyltransferases – the “methyltransferome”**

---

Methyltransferases are a large class of proteins catalysing the methylation of substrates including protein targets as well as nucleic acids (Schubert *et al.*, 2003). It is estimated that the overwhelming majority of the methyl groups used in these reactions come from AdoMet (Fontecave *et al.*, 2004). The substrates of MTases vary widely, and their biological roles have generated great interest in a diverse range of biological systems and models, ranging from oncology to tropical disease (Schubert *et al.*, 2003, Petrossian & Clarke, 2011, Wlodarski *et al.*, 2011, Subramaniam *et al.*, 2014, Hamamoto & Nakamura, 2016). To gain a better understanding of the MTase repertoire present in *T. brucei*, we performed a literature search. To our surprise, only 22 MTases have been identified and characterised in *T. brucei* (Table 3-5). This is in stark contrast to, for example, kinases, many more of which have been studied in detail in recent years (Parsons *et al.*, 2005, Jones *et al.*, 2014b). Furthermore, a search in the TriTrypDB database using a key term of “methyltransferase” produced an output of 129 putative genes.

GeneID	Name	Putative function	Reference
Tb927.1.4690	PRMT 1	type I arginine MTase	(Pelletier <i>et al.</i> , 2005)
Tb927.10.640	PRMT 5	type II arginine MTase	(Pasternack <i>et al.</i> , 2007)
Tb927.7.5490	PRMT 7	type III arginine MTase	(Pasternack <i>et al.</i> , 2007) (Fisk <i>et al.</i> , 2009) (Wang <i>et al.</i> , 2014a)
Tb927.5.3960	PRMT 6	type I arginine MTase	(Pasternack <i>et al.</i> , 2007) (Fisk <i>et al.</i> , 2010) (Wang <i>et al.</i> , 2014a)
Tb927.10.3560	PRMT	putative arginine MTase	(Pasternack <i>et al.</i> , 2007)
Tb927.10.3560	PRMTs	looked at all arginine MTases	(Lott <i>et al.</i> , 2014)
Tb927.1.4690	PRMTs	looked at all arginine MTases	(Goulah <i>et al.</i> , 2006)
Tb927.7.2080	TbCgm1	cap guanylyltransferase-methyltransferase	(Ruan <i>et al.</i> , 2007a)
Tb927.10.4500	TbCmt1	cap methyltransferase	(Hall & Ho, 2006a)
Tb927.10.7940	TbMtr1	2'-O-ribose MTase	(Zamudio <i>et al.</i> , 2007, Mittra <i>et al.</i> , 2008)
Tb927.11.4890	TbMtr2	2'-O-ribose MTase	(Hall & Ho, 2006b) (Arhin <i>et al.</i> , 2006b)
Tb927.9.12040	TbMtr3	2'-O-ribose MTase	(Arhin <i>et al.</i> , 2006a) (Zamudio <i>et al.</i> , 2006)
Tb927.8.1920	DOT1A	Histone MTase	(Janzen <i>et al.</i> , 2006)
Tb927.1.570	DOT1B	Histone MTase	(Figueiredo <i>et al.</i> , 2008)
Tb927.11.3200	DOT1A/B	Histone MTase	(Janzen <i>et al.</i> , 2006)
Tb927.9.2800	MGMT	O6-methylguanine-MTase	Genois
Tb927.10.12600	TbHEN1	HUA enhancer	(Shi <i>et al.</i> , 2014)
Tb927.8.5720	TbTRM5	tRNA MTase	(Paris <i>et al.</i> , 2013)
Tb11.02.5090	TbTgs	PRIP interacting protein, Trimethylguanosine (TMG) RNA cap synthase	(Benarroch <i>et al.</i> , 2010) (Ruan <i>et al.</i> , 2007b)
Tb927.10.6910			
Tb927.10.6950	N/A	sterol 24-c-MTases	(Perez-Moreno <i>et al.</i> , 2012)
Tb927.8.7120			
Tb927.11.3270			
Tb927.9.12690	N/A	Prenyl protein specific carboxyl MTase	(Buckner <i>et al.</i> , 2002)

**Table 3-5: MTases that have been identified and characterised in *T. brucei*.** A literature search of MTases characterised in *T. brucei* showed publications covering 32 genes, with varying levels of characterisation. Unsurprisingly, MTases involved in spliced leader methylation have received by far the most attention, followed closely by arginine MTases. Whilst histone MTases such as DOT1A & DOT1B have been investigated, it is interesting to note that most lysine MTases in *T. brucei* have yet to receive similar amounts of attention.

Based on the hypothesis that AN5568 might affect an essential MTase in *T. brucei*, we sought to carry out an *in silico* study of the entire MTase complement in *T. brucei*. This type of study has been carried out in both *Saccharomyces cerevisiae* and in the human genome (Petrossian & Clarke, 2011, Wlodarski *et al.*, 2011). However, both studies disagreed on the term used to describe the MTase complement, with Petrossian & Clarke coining the name “methyltransferasome”, whilst Wlodarski and colleagues called it the “methyltransferome”. It was decided to use the latter name in this study.

The first list, generated through TriTrypDB (v9.0) (Aslett *et al.*, 2010), was analysed and used for the subsequent overexpression experiments detailed below. However, more in depth analysis of MTases in *T. brucei* was carried out using domain searches. For this, the pfam, interpro and SMART databases (Finn *et al.*, 2014, Jones *et al.*, 2014b, Letunic *et al.*, 2015) were used to generate a list of all genes containing MTase domains (table 3-6). The list was refined further by including EC number analysis from TriTrypDB with the search term “2.1.1” (methyltransferase).

The final list for the *T. brucei* MTase complement totalled 145 genes, all predicted to contain MTase domains (table 3-6). Most genes contained just one MTase domain, whilst several were predicted to contain multiple MTase domains. In addition, many of the predicted MTases also possessed other domains related to substrate binding, as well as transmembrane domains.

We divided the “methyltransferome” into groups depending on the major folds present in the MTase domains, as was carried out by Wlodarski and colleagues (Wlodarski *et al.*, 2011). These classes have previously been made based on the catalytic domain structure (Schubert *et al.*, 2003), as explained in [section 1.8.5](#). Rossmann-like folds are the largest class of MTases and consist of a seven-stranded beta sheet joined by several alpha helices (Schubert *et al.*, 2003).

<b>Fold</b>	<b>Substrate</b>	<b>GeneID (Name)</b>	
<b>Rossmann-like</b>	<b>DNA</b>	Tb927.3.1360, Tb927.7.6620	
		Tb11.02.5090 ( <b>TbTgs1</b> ), Tb927.7.2080 ( <b>TbCgm1</b> ), Tb927.10.4500 ( <b>TbCMT1</b> ), Tb927.10.7940 ( <b>TbMtr1</b> ), Tb927.11.4890 ( <b>TbMtr2</b> ), Tb927.9.12040 ( <b>TbMtr3</b> ),	
	<b>RNA</b>	<b>mRNA</b>	Tb927.11.5060, Tb927.11.6720, Tb927.2.2450, Tb927.2.4550, Tb927.4.4170, Tb927.5.490, Tb927.8.2500, Tb11.02.5090b, Tb11.v5.0219, Tb11.v5.0415
		<b>tRNA</b>	Tb927.8.5720 ( <b>TbTRM5</b> ), Tb927.1.3050, Tb927.10.3080, Tb927.11.3890, Tb927.11.5240, Tb927.4.4030, Tb927.9.9120
		<b>rRNA</b>	Tb927.8.3130, Tb927.6.1610, Tb927.10.3690
	<b>sno/ snRNA</b>	Tb927.10.14630, Tb927.10.14750, Tb927.10.7500	
	<b>Protein</b>		Tb927.8.1920 ( <b>TbDOT1A</b> ), Tb927.1.570 ( <b>TbDOT1B</b> ), Tb927.11.3200 ( <b>TbDOT1A/B</b> ), Tb927.1.4690 ( <b>TbPRMT1</b> ), Tb927.10.640 ( <b>TbPRMT5</b> ), Tb927.7.5490 ( <b>TbPRMT7</b> ), Tb927.10.3560 ( <b>TbPRMTx</b> ), Tb927.10.12980, Tb927.10.4460, Tb927.10.6680, Tb927.10.7850, Tb927.11.15630, Tb927.11.9090, Tb927.11.9210, Tb927.3.3230, Tb927.3.4540, Tb927.4.3840, Tb927.5.2050, Tb927.5.2420, Tb927.6.2270, Tb927.6.2330, Tb927.9.12780, Tb927.9.6430, Tb927.9.6720
		<b>Small molecule</b>	Tb11.v5.0692, Tb927.10.9390, Tb927.11.10510, Tb927.5.3000
		<b>Lipid</b>	Tb10.v4.0247, Tb11.v5.0496, Tb927.10.6910, Tb927.10.6950

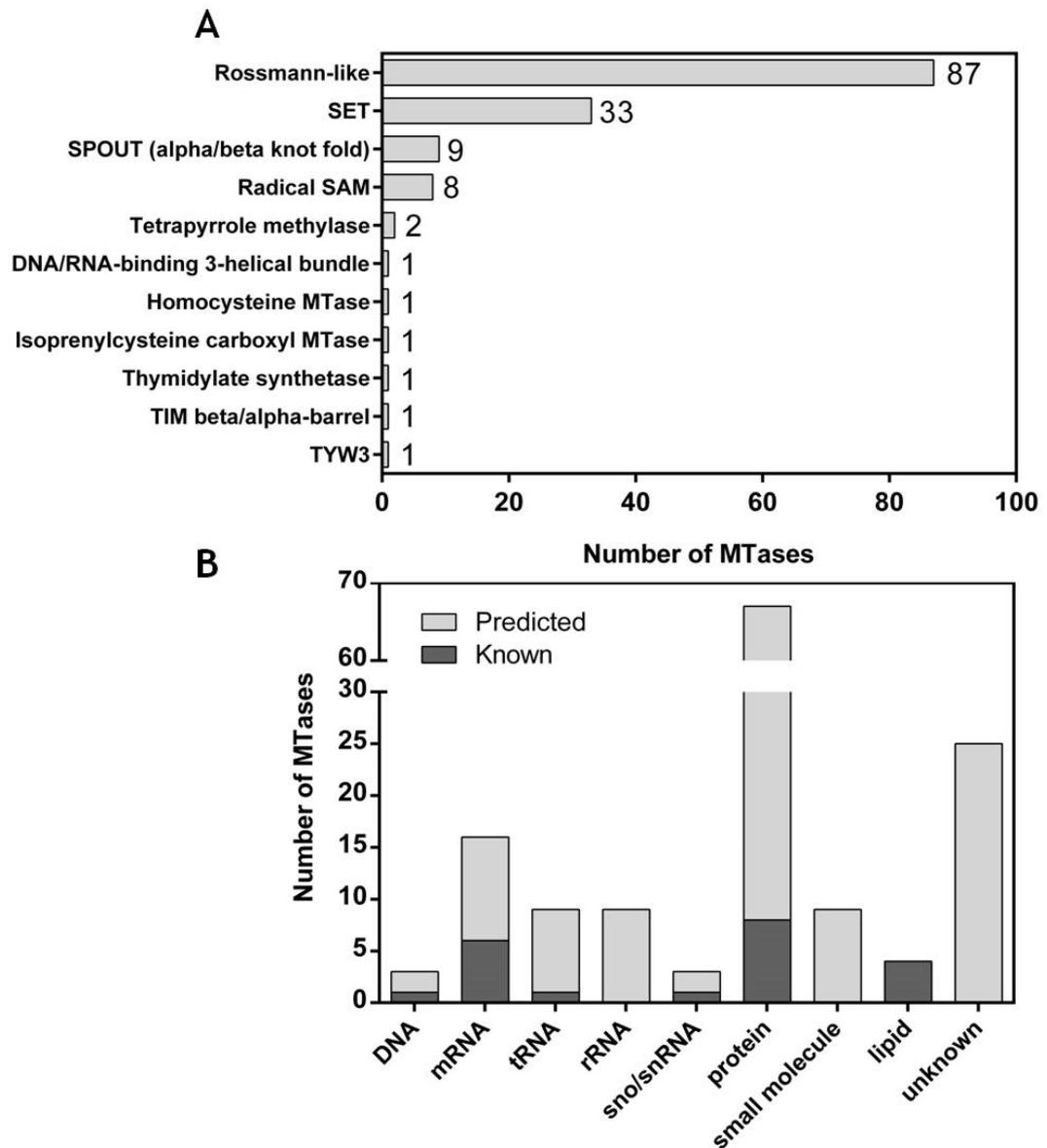
Fold	Substrate	GeneID (Name)
	Unknown	Tb927.10.12600 ( <b>TbHEN1</b> ), Tb11.v5.0775, Tb11.v5.0832, Tb927.1.1120, Tb927.10.12270, Tb927.10.15420, Tb927.10.1800, Tb927.10.2590, Tb927.10.7560, Tb927.10.9860, Tb927.11.15290, Tb927.11.9200, Tb927.3.4940, Tb927.4.1900, Tb927.5.2600, Tb927.6.2540, Tb927.7.3340, Tb927.7.6890, Tb927.8.2670, Tb927.8.5420, Tb927.9.10580, Tb927.9.11750, Tb927.9.14200, Tb927.9.2320, Tb927.9.7000
SET domain	Protein	Tb11.v5.0422, Tb11.v5.0590, Tb927.1.4720, Tb927.10.11130, Tb927.10.12880, Tb927.10.3730, Tb927.10.4600, Tb927.10.8060, Tb927.10.8100, Tb927.10.9680, Tb927.11.13560, Tb927.11.5120, Tb927.3.1860, Tb927.3.3370, Tb927.3.750, Tb927.4.2440, Tb927.4.3310, Tb927.4.4300, Tb927.5.2770, Tb927.5.3500, Tb927.6.3610, Tb927.6.910, Tb927.7.2040, Tb927.7.5620, Tb927.8.2490, Tb927.8.2690, Tb927.8.2710, Tb927.8.6470, Tb927.8.6530, Tb927.9.10100, Tb927.9.11350, Tb927.9.13470, Tb927.9.1510
SPOUT	RNA	tRNA Tb927.6.4420
		rRNA Tb11.v5.0683, Tb927.1.4810, Tb927.11.15830, Tb927.2.4980, Tb927.7.5120, Tb927.8.5040, Tb927.8.6560, Tb927.9.12850
Radical SAM	Protein	Tb11.v5.0520, Tb11.v5.0635, Tb927.10.15010, Tb927.6.3510, Tb927.7.7130, Tb927.8.3310, Tb927.8.5770, Tb927.9.4950
Tetrapyrrole methylase	Small molecule	Tb927.4.4650, Tb927.9.12100
DNA/RNA-binding 3-helical bundle	DNA	Tb927.9.2800 ( <b>MGMT</b> )
Hcy MTase	Small molecule	Tb927.1.1270
ICMT	Protein	Tb927.9.12690
Thymidylate synthetase	Small molecule	Tb927.7.5480
TIM beta/alpha-barrel	Small molecule	Tb927.8.2610
TYW3	RNA	tRNA Tb927.11.12670

**Table 3-6: A comprehensive overview of the *T. brucei* methyltransferase repertoire.** MTases were divided into groups based on A) their MTase domain structure, and B) their substrate, if known. Identification of MTases was based on their domains that were identified using pfam and interpro (Finn *et al.*, 2014, Jones *et al.*, 2014b). These are available on the TriTrypDB website.

Using the Pfam database, Rossmann-like folds were assigned to the pfam clan CL0063 (FAD/NAD(P)-binding Rossmann fold superfamily). This was by far the largest group of MTases in the *T. brucei* genome, with 88 representatives (fig. 3-31A). The second largest group was MTases belonging to the SET domain class, which is known to include mostly methylases of lysine residues in proteins (Wlodarski *et al.*, 2011).

The most studied group of MTases in *T. brucei* is the mRNA MTase group (fig. 3-31B). Indeed, many of these proteins have been characterised as methylases of the mRNA spliced leader cap, which, as mentioned previously, is unique and complex in the kinetoplastids (Gunzl, 2010).

However, interestingly, many of these proteins have not been found to be essential to the parasite (Zamudio *et al.*, 2006). Indeed, the absence of methylation does not appear to inhibit protein synthesis in all cases, although TbCgm1, for which an RNAi knock-down was generated, did appear to have reduced mRNA synthesis. This has been shown for three other mRNA methylators: TbMtr1, TbMtr2 and TbMtr3 act on the first, second and third bases of the spliced leader respectively, but all are non-essential (Zamudio *et al.*, 2006, Zamudio *et al.*, 2009).



**Figure 3-31: Overview of MTases in the *T. brucei* genome.** A) MTases were grouped according to the catalytic MTase domain they possessed. The most abundant fold is the Rossmann-like fold (Pfam: CL0063), most often associated with AdoMet-dependent MTases. Many lysine MTases, which contain a SET domain, were also predicted. Several putative MTases were predicted to contain a SPOUT domain, or a radical SAM domain. Other classes that have been covered in literature elsewhere were also found here. In most cases, only one MTase was predicted for each domain, similar to the MTase complement in *Saccharomyces cerevisiae*. B) Putative MTases encoded in the *T. brucei* genome were grouped based on their predicted substrates, dependent on the domains identified in each gene. The majority of MTases were predicted to act on amino acid residues in proteins, especially those containing the SET domain known to interact with lysines. In addition, at least 16 predicted MTases are predicted to catalyse mRNA methylation. There was also a large group of MTases for which the substrates remain unknown, and must be investigated further.

Four of the putative MTases (Tb10.v4.0247, Tb11.v5.0496, Tb927.10.6910 & Tb927.10.6950) were found to contain sterol MTase C-terminal domains (Pfam: PF08498), and were therefore predicted to be sterol MTases. These can be involved in the methylation of sterols, which is an essential process in lipid membranes (Perez-Moreno *et al.*, 2012).

In some cases, proteins previously found to be MTases did not contain pfam domains relating to MTase activity. One example is Tb927.8.7120, which was predicted to be a sterol MTase (Perez-Moreno *et al.*, 2012). Instead, this protein was found to possess a squalene synthesis domain. Similarly, Tb927.11.3270 contained a squalene epoxidase domain, but no MTase domain. In another case, we could not find any arginine MTase domains in the PRMT6 gene (Tb927.5.3960).

These cases highlight the need for a more refined study of methyltransferases in the *T. brucei* genome. This type of study of the MTase complement should incorporate two analyses: Firstly, domain prediction allows for the identification of signalling peptides that could give clues as to the localisation of all the MTases. Indeed, this information can also be gained through the study of homologues in other organisms. For example, two distant homologues of a mitochondrial MTase that dimethylates 12S rRNA in humans and yeast (TFB1M), Tb927.6.1610 & Tb927.10.3690, were found based on their similarity to the yeast and human proteins. Based on this finding, many other characteristics of these MTases can potentially be found.

Secondly, datasets exist that give very detailed information of protein expression as well as transcript abundance of thousands of genes throughout the varying life cycle stages of the parasite. These datasets could be used to find important MTases involved in the regulation of life-cycle-specific genes.

One interesting findings from this study is the very high number of protein MTases *T. brucei* possesses (fig. 3-31B). Our analyses predict 67 protein MTases, which is much higher compared to organisms such as yeast (32) (Wlodarski *et al.*, 2011), but similar to the number in humans (at least 57 SET domain proteins humans) (Petrossian & Clarke, 2011). One possible reason for this is due to the complex life cycle the parasite undergoes, requiring a significant number of changes in morphology, gene expression and metabolism. Protein methylation is well established as a posttranslational modification that can induce these necessary changes. Given this information, one would hypothesize that this gene set could contain unique proteins and therefore, potential drug targets against the parasite.

Further work to refine the *T. brucei* methyltransferome, as well as those of the other TriTryps would be highly beneficial to the scientific community. Firstly, this dataset will contain druggable proteins that can be used to develop novel therapeutics against these parasites. In

addition, the kinetoplastidea in particular, are an ancient eukaryote. Therefore, a great deal of biological understanding can be gained through the study of both *Trypanosoma spp.*, and *Leishmania spp.*

The human and yeast genomes were shown to contain 208 and 86 MTases respectively (Petrossian & Clarke, 2011, Wlodarski *et al.*, 2011), and a good starting point would be to find any homologues in the TriTryp genomes. Additional work would include analysis of all genes containing MTase domains, especially with regards to any conserved secondary and tertiary structure using prediction algorithms such as Hidden Markov Models (HMM) (Katz *et al.*, 2003, Petrossian & Clarke, 2009).

The development of systems such as RITseq have been very useful for the study of *T. brucei* (Alsford *et al.*, 2012). Recently, it was used with great success to investigate the *T. brucei* kinase repertoire, referred to as the “kinome” (Jones *et al.*, 2014a). A similar approach to the methyltransferase complement would provide great insight into the varied roles these proteins are thought to play, including, but not limited to, differentiation, DNA transcription, protein synthesis, and posttranslational modifications (Petrossian & Clarke, 2009, Subramaniam *et al.*, 2014, Swinehart & Jackman, 2015). Using the RITseq dataset that was previously published elsewhere (Alsford *et al.*, 2011), we analysed the first list of 129 MTases to look for enzymes that were deemed essential, i.e. their knock-down was lethal.

### **3.13 Overexpression of 12 essential methyltransferases *in vitro* does not identify an AN5568 target**

---

#### **3.13.1 Domain analysis of 12 essential methyltransferases**

Using RITseq analysis datasets available on TriTrypDB (Aslett *et al.*, 2010), we determined 12 MTases that were shown to be essential in the original dataset of 129 genes found in the initial “methyltransferome” scan (Table 3-7). As a proof-of-principle, to confirm that overexpression of a target protein could lead to drug resistance, a cell line overexpressing ornithine decarboxylase (ID: Tb427tmp.01.5300), the target of eflornithine (Vincent *et al.*, 2012), was generated. For each of the MTases, the sequence was run through the SMART database to identify any known protein domains (Letunic *et al.*, 2015). The results from this experiment are presented in table 3-7.

Gene – TREU 927	Gene – Lister 427	Given ID	Putative description
Tb927.10.3080	Tb427.10.3080	MT 1	AdoMet-dependent MTase
Tb927.10.7560	Tb427.10.7560	MT 2	AdoMet-dependent MTase
Tb927.10.9020	Tb427.10.9020	MT 3	Gcd10p MTase
Tb927.11.13250	Tb427tmp.01.4830	MT 4	Elongation factor MTase
Tb927.11.9690	Tb427tmp.01.1460	MT 5	MORN repeat protein
Tb927.3.2890	Tb427.03.2890	MT 6	MORN repeat protein
Tb927.4.1900	Tb427.04.1900	MT 7	AdoMet-dependent MTase
Tb927.5.2050	Tb427.05.2050	MT 8	Lysine MTase
Tb927.7.4320	Tb427.07.4320	MT 9	mRNA cap triphosphatase
Tb927.8.5040	Tb427.08.5040	MT 10	Alpha/beta-knot fold MTase
Tb927.6.2270	Tb427.06.2270	MT 11	AdoMet-dependent di-proline MTase
Tb927.10.7850	Tb427.10.7850	MT 12	Rab-like GTPase
Tb927.11.13730	Tb427tmp.01.5300	ODC	Ornithine decarboxylase

**Table 3-7: Putative essential MTases.** The genes chosen for further analysis were deemed to be essential using the RITseq datasets generated by Horn & Alford (Alford *et al.*, 2011). In total, 12 genes were chosen to be overexpressed to test whether they could be potential targets of AN5568. Whilst TriTrypDB generated these genes in the search for MTases, some were found to contain unrelated domains. These were still treated as putative MTases, and thus included in the subsequent overexpression experiments.

A predicted MTase previously isolated from the previously discussed DARTS assay (table 3-3) (ID: Tb927.10.3080) was shown to contain a methyltransferase domain (pfam: PF02475, IPR: 003402) thought to be involved in wybutosine biosynthesis. This is a heavily modified nucleoside present in <sup>Phe</sup>tRNA, and thus required for protein synthesis (Kalhor *et al.*, 2005, Noma *et al.*, 2006). By analysing the RITseq data published by Horn & Alford, we were able to confirm that this protein was essential, and it was selected for downstream analyses. There is no literature to suggest the presence of the wybutosine biosynthesis pathway in *T. brucei*, although it is conserved across most eukaryotes (Kalhor *et al.*, 2005, Noma *et al.*, 2006).

The second gene (ID: Tb927.10.5760) was shown to contain a domain matching 4 different pfam domain classes of AdoMet-dependent methyltransferases (pfam: PF13489, PF12847, PF13649 & PF08241). The third gene (Tb427.10.9020) possesses a predicted domain involved in a methyltransferase complex with Gcd14p (pfam: 04189), which is involved with modifying tRNA residues. A homologue of this domain has been characterised in *S. cerevisiae* where it has been shown to be an essential 1-methyladenosine MTase (Anderson *et al.*, 2000). The fourth methyltransferase, Tb427tmp.01.4830, was predicted to contain an elongation factor domain (pfam: PF00009). This domain is part of a family that is involved in GTP-dependent binding of aminoacyl tRNAs to ribosomes and aids in the binding of initiation factors such as eIF2 gamma (Moller *et al.*, 1987).

Two genes filtered from this subset (Tb427tmp.01.1460 & Tb427.03.2890) were predicted to contain several membrane occupation and recognition nexus (MORN) repeat domains (pfam: PF02493). Proteins containing this domain are thought to function as linkers between

membrane regions and the cytoskeleton, as has been shown in the apicomplexan parasite *Toxoplasma gondii*, where it is thought the proteins play important roles in cytokinesis (Gubbels *et al.*, 2006). In addition, this domain can also be found on a subset of histone-lysine MTases, specifically the SETD7 MTase that is found in many eukaryotes (Couture *et al.*, 2006).

The seventh protein (Tb427.04.1900) contained two predicted MTase domains. The first was an AdoMet-dependent MTase domain (pfam: PF012847), as was the second (pfam: PF12589). Interestingly, the latter domain has been shown to be deleted in a neurodevelopmental disorder in humans known as Williams-Beuren Syndrome (Doll & Grzeschik, 2001). However, it is unknown what role, if any, this domain plays in trypanosomatids. Another essential MTase, Tb427.5.2050, was shown to contain a lysine MTase domain (pfam: PF10294). Further study indicated that this domain is also found in nicotinamide N-MTases (Katz *et al.*, 2003).

mRNA capping enzymes are vital in trypanosome biology, and one of the proteins found to be essential by RITseq, Tb427.07.4320, was predicted to contain a beta domain that codes for an mRNA cap triphosphatase (pfam: PF02940). Normally, proteins with this domain also contain a catalytic alpha chain that possesses guanylyltransferase activity (Lima *et al.*, 1999). However, neither pfam, nor SMART, could predict any alpha chains in the gene examined here.

A further essential protein, Tb427.8.5040, was predicted to possess an EMG1/NEP1 MTase domain (pfam: PF03587). This MTase belongs in the alpha/beta-knot MTases and this domain in particular was previously shown to be involved in 40S ribosome biogenesis in *S. cerevisiae* and the domain in question is conserved across most eukaryotes and archaea (Leulliot *et al.*, 2008). Two other putative MTases were predicted to be essential by RITseq. The first, Tb427.6.2270, a predicted AdoMet-dependent di-proline MTase (pfam: PF05891). In other organisms such as yeast, this MTase domain is involved in methylation of ribosomal N-terminal domains (Webb *et al.*, 2010). This gene also contained a second predicted AdoMet-dependent MTase domain, but this domain has not been annotated or characterised.

Finally, Tb927.10.7850 was found not to contain any domains as predicted by the Pfam database, but did contain a Rab-like GTPase domain when analysed using the SMART database. The proteins are involved in GTPase activation, an important and conserved signalling mechanism.

In all cases, these *T. brucei* genes and the proteins they encode had not been characterised before, and there was no mention in the literature of any. Indeed, several of the processes or pathways they are predicted to be involved in have not been confirmed to exist in the trypanosomatids, perhaps highlighting the lack of knowledge regarding MTases in these parasites.

It must be stressed that further bioinformatics analyses of these 12 proteins carried out later in the study, showed that not all of these were purported to be MTases, but instead, were enzymes containing domains that might interact directly or indirectly with MTases. It was surprising to find that several of the putative MTases filtered out by TriTrypDB did not in fact contain recognised MTase domains, and it would be worthwhile to repeat this search for essential MTases with the new knowledge gained from the generation of a dataset containing every predicted MTase, as shown in table 3-6. However, further work was carried out in order to overexpress all genes summarised in table 3-7, initially with the overexpression of each gene in the lab-adapted Lister 427 *T.b. brucei* strain.

### 3.13.2 Characterisation of 12 overexpression cell lines

To determine whether any of the 12 chosen genes could be a potential target of AN5568, we chose to overexpress the genes *in vitro* in Lister 427 wild-type *T. brucei* cells. This was carried out primarily with the ribosomal (r)RNA targeting pURAN plasmid (a gift from S. Hadjuk). As the 427 Lister strain was used, it was chosen to overexpress the homologues of the corresponding 427 genes in these cell lines. This also allowed easy amplification of DNA from 427 Lister total DNA. All genes were cloned into pURAN after sequencing in the intermediate vector pGEM-T easy, as described in the methods ([section 2.4.5](#)).

For two of the chosen genes, (IDs: Tb427.10.7560 & Tb427.05.2050) attempts to transfect a pURAN construct into parasites were unsuccessful and, therefore, a different plasmid, pRM482, was chosen. This plasmid targets the tubulin array, which is constitutively active, but at a lower overall rate of transcription (Trenaman *et al.*, 2013) compared to the rRNA loci. Unsuccessful transfection could be due to the detrimental effect resulting from high levels of overexpression of a protein usually present in much lower quantities. In addition, high levels of expression could also compromise the structure of the endoplasmic reticulum, leading to swollen ERs as well as ER stress. This could explain the lack of stable transfected clones for the two aforementioned genes.

Important to note is the lack of tags on the proteins being overexpressed. It is common to express proteins with tags such as GFP or HA. This enables localisation of the protein intracellularly, and more importantly, to confirm overexpression on a protein level, rather than just transcript level, as a qPCR does. However, the tags run the risk of leading to improper folding of the protein due to interference with the protein chain's C- or N-terminal domains. For this reason, no tags were transfected along with the proteins of interest.

Upon transfection of each MTase overexpression construct, two clonal transfectants were isolated for each MTase, as well as the transfectant for *ODC*. To confirm overexpression of each MTase, qPCR analysis was carried out on each clone (table 3-8). Changes in mRNA abundance

were found to be highly variable in cell lines transfected with pURAN, as well as pRM482 constructs, with fold changes ranging from <2-fold all the way to >34-fold. This is most likely due to the nature of the construct, which is not specifically targeted to one locus in the rRNA or tubulin arrays. Instead, the construct can recombine in many places within the same locus, leading to the high variability seen in table 3-8. Satisfied that overexpression was achieved, we next sought to characterise both cellular growth, as well as changes in drug resistance/sensitivity.

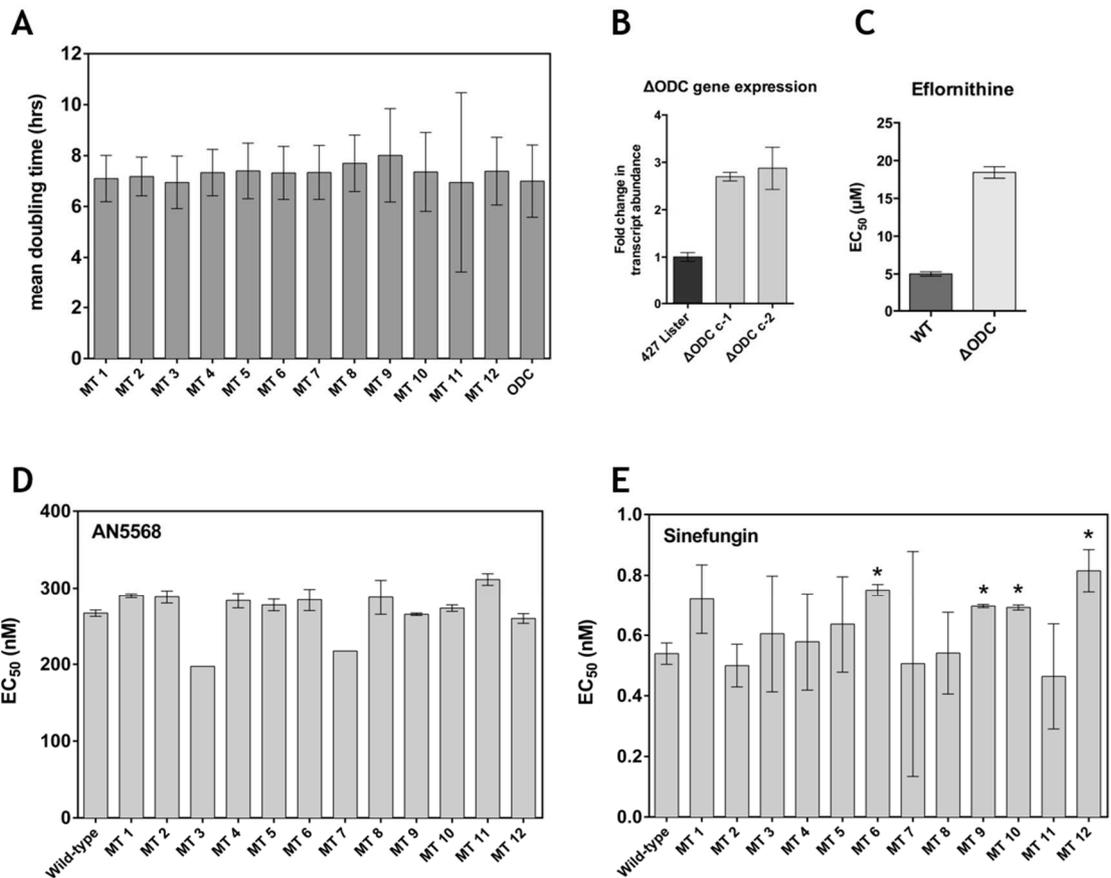
GeneID	Fold change (overexpressor/WT ctrl)
MT 1 - clone 1	2.01 ± 0.13
MT 1 - clone 2	2.29 ± 0.05
MT 2 - clone 1	21.61 ± 5.98
MT 2 - clone 2	1.49 ± 0.09
MT 3 - clone 1	10.90 ± 8.15
MT 3 - clone 2	8.90 ± 3.84
MT 4 - clone 1	2.89 ± 0.71
MT 4 - clone 2	2.32 ± 0.19
MT 5 - clone 1	4.64 ± 2.35
MT 5 - clone 2	6.90 ± 0.53
MT 6 - clone 1	1.86 ± 1.86
MT 6 - clone 2	4.16 ± 1.54
MT 7 - clone 1	6.58 ± 1.58
MT 7 - clone 2	2.47 ± 0.41
MT 8 - clone 1	1.71 ± 0.03
MT 8 - clone 2	3.14 ± 0.06
MT 9 - clone 1	4.82 ± 0.56
MT 9 - clone 2	2.81 ± 0.21
MT 10 - clone 1	3.41 ± 0.19
MT 10 - clone 2	3.22 ± 0.05
MT 11 - clone 1	1.55 ± 0.09
MT 11 - clone 2	1.71 ± 0.04
MT 12 - clone 1	16.42 ± 0.37
MT 12 - clone 2	34.52 ± 0.95
ODC - clone 1	2.70 ± 0.09
ODC - clone 2	2.88 ± 0.45

**Table 3-8: qPCR analysis of 12 MTase overexpression lines.** RNA was isolated and reverse-transcribed from each overexpression line made. Transcript abundance of the gene of interest was then analysed by qPCR. Results showed a significant amount of variation in the levels of overexpression, ranging from 2-fold to as high as 34-fold for one cell line. Two clones were characterised for each construct. Cell lines are given by their ID, corresponding full gene IDs are given in table 3-7.

Mean doubling time was next calculated as described by Sykes & Avery (Sykes & Avery, 2009b), and is shown in figure 3-32A. All cell lines were counted every 24 hours during exponential growth (between  $2 \times 10^4$  and  $2 \times 10^6$  cells/mL) over a period of 5 days, and mean doubling time was then calculated for each 24-hour period. Previous studies on the 427 Lister cell line has

shown mean doubling time to vary between 6 and 7 hours (Sykes & Avery, 2009a). Every MTase overexpression line that was analysed showed a mean doubling time within these limits, indicating that none of the protein overexpression was detrimental, nor advantageous, to the modified cell lines.

Previous studies in *T. brucei* have suggested that overexpression of proteins can cause resistance to drugs that target the protein in question (Ranade *et al.*, 2013). To confirm whether overexpression of a protein target with the pURAN construct can in fact cause resistance to a specific drug, we first analysed the ODC overexpressor (fig. 3-32B & 3-32C). Firstly, qPCR analysis of total cDNA showed that for two isolated clones of the overexpressor, there was just under 3-fold change in transcript abundance (table 3-8 & fig. 3-32B), suggesting that the construct was successfully integrated and overexpressing ODC at the transcript level. Next, an alamar blue assay was run for eflornithine (Ornidyl®), using a wild-type Lister 427 *T. brucei* line as a control. Whilst the WT control showed an EC<sub>50</sub> of ~5 µM, the combined average of the two ODC overexpressors exhibited an EC<sub>50</sub> of more than 15 µM, equivalent to a 3-fold increase in drug resistance (fig. 3-32C). This result showed that overexpression can change drug resistance, and therefore allowed use to proceed with the characterisation of the 12 MTase overexpressors.



**Figure 3-32: Characterisation of an ODC overexpressor as well as 12 putative MTase overexpressors.** Once successful integration of the constructs, as well as overexpression, had been confirmed, the cell lines were characterised with regards to growth and drug sensitivity/resistance. A) Growth curves carried out for all 12 lines showed that none exhibited any significant changes in their doubling times, when compared to wild-type Lister 427 cells which have been shown to possess a doubling time of approximately 6-7 hours (Sykes & Avery, 2009a). B) As a proof-of-concept, a cell line overexpressing ornithine decarboxylase (ODC) was generated. C) The ODC overexpressor showed an increased resistance to Eflornithine, an ODC inhibitor, by more than 3-fold. D) Subsequent analysis of the 12 putative MTase overexpressors showed that none had gained any significant resistance to AN5568 when compared to a wild-type cell line, although two cell lines did appear to show increased sensitivity. E) Interestingly, several of the putative MTase overexpressors showed increased resistance to sinefungin, a non-specific MTase inhibitor. Whilst not entirely unexpected, this result was interesting by itself, and warrants further study (\* =  $P < 0.05$ , Student's *t*-test).

To ascertain whether any of the MTase overexpressors had altered resistance or sensitivity to the benzoxaborole, alamar blue assays were carried out in 12-well format, as described in the methods. For this assay, the same AN5568 sample was used for each cell line, so that the results could be compared. The increased AN5568 (250 nM) EC<sub>50</sub> for Lister 427 WT cells was likely due to the compound degrading over time (fig. 3-32D). However, when compared to the EC<sub>50</sub> values of all other MTase overexpressors for AN5568, no statistically (Student's *t*-test) significant differences were found. In addition to investigating changes in AN5568 resistance, and given that the proteins being overexpressed were putative MTase, further drug resistance assays were carried out with sinefungin, a non-specific AdoMet-dependent MTase inhibitor discussed earlier in this chapter. In similar fashion to the AN5568 sensitivity assays, alamar blue experiments were carried out for each cell line (fig. 3-32E). Results showed a significant amount of variation within cell lines, for unknown reasons. Interestingly, four of the twelve MTase

overexpressors showed a significant increase in sinefungin resistance (fig. 3-32E). The MORN-repeat protein (MT 6,  $p = 0.0005$ ), putative mRNA cap-triphosphatase (MT 9,  $p = 0.0250$ ), putative alpha/beta-knot MTase (MT 10,  $p = 0.0008$ ) and rab-like GTPase (MT 12,  $p = 0.070$ ) all showed increased resistance to sinefungin (Gene IDs: Tb427.03.2890, Tb427.07.4320, Tb427.05.5040 & Tb427.10.8750 respectively). These differences were not completely unexpected, but it is interesting that these genes in particular seem to be involved in dealing with sinefungin exposure, whilst other MTases are not. Overexpression of an mRNA cap triphosphatase could increase the parasite's capability for protein expression, thereby providing a larger pool of essential proteins, and diminishing the potency/efficacy of the compound. In addition, whilst there is no literature on this mRNA cap triphosphatase, others have been identified and characterised to a greater extent in *T. brucei* (Takagi *et al.*, 2007), suggesting there might be some redundancy between related genes.

It was interesting to observe that not every MTase could confer cellular resistance to sinefungin when overexpressed. It is likely that high amounts of overexpression would give the parasites the ability to “mop” up the sinefungin, without compromising metabolism. This mechanism would not be specific to a particular methyltransferase because sinefungin is non-specific. In addition, some methyltransferases might be more essential

Unfortunately, overexpression of these 12 candidate genes did not lead to a potential AN5568 target. However, it is encouraging to know that overexpression of the correct target should provide an idea of the possible targets, as shown in the case of ODC overexpression and eflornithine-treatment. The RITseq approach pioneered by David Horn and Sam Alford has recently been applied to the study of drug mode of action, with interesting findings (Alford *et al.*, 2012). However, one of the main disadvantages of this approach is that RNAi-mediated knock-down of essential proteins is highly detrimental to the parasite. With this in mind, it is likely that RITseq will only identify non-essential proteins involved in drug mode of action. In contrast, essential protein targets that are overexpressed do not appear to cause severe growth phenotypes given the experiments presented here. Therefore, this approach could in theory, identify direct protein drug targets. The main issue associated with this method is that, unlike RITseq, the full ORF of the protein must be transfected into the parasite. This means that development of a complete overexpression library of the entire proteome, which is thought to contain upwards of 9,000 genes, would take a considerable amount of effort and time, and would require very elaborate and well-organised experimental set-up. However, several groups collaborating in Dundee are currently proceeding with an overexpression library generated with a random shotgun approach, which is easier and more efficient to carry out (M. Barrett, personal communication).

### 3.14 Discussion

---

The novel benzoxaborole, AN5568, has been shown to have a high efficacy both *in vivo* and *in vitro*, and importantly, is active against both *T. b. rhodesiense* and *T. b. gambiense* (Jacobs *et al.*, 2011). The drug is currently undergoing clinical trials, and along with another novel compound, fexinidazole, could become the first new anti-trypanosomal on the market in a generation (Brun *et al.*, 2011). However, the MoA of AN5568 is currently unknown.

Here, we attempted to utilise a range of omics-based approaches in combination with molecular biology techniques, in order to dissect the MoA of this novel benzoxaborole. Firstly, we show that the drug retains potency in an *in vitro* setting, against BSF *T. brucei* Lister 427 cells. In contrast, the compound efficacy was drastically reduced in PCF trypanosomes, indicating that this drug is better suited to the treatment of the mammalian infective stages. Indeed, even after a 30 hour treatment at  $10\times EC_{50}$ , live cells were observed in PCF cultures. In addition, *T. brucei* BSF are more sensitive to AN5568 compared to other trypanosomatids such as *L. mexicana* and *T. cruzi*.

Several interesting BSF phenotypes were observed during microscopy analyses of AN5568-treated cells. Firstly, drug-mediated action causes significant complications in cell division. This was highlighted by incubating cells for 24 hours with a sub-lethal dosage of the benzoxaborole (fig. 3-3), but also occurred in lethal concentrations (fig. 3-4). By 12 hours, cells become rounded and mitochondrial morphology was severely compromised, and quantification of nuclei and kinetoplasts hinted at defects in cytokinesis, as well as kDNA replication, as shown by the increase in 2N1K and MNMK cells subsequent to drug-treatment. These phenomena have previously been reported in KO's or knock-downs of specific targets involved in mitochondrial metabolism and nucleotide replication (Ferguson *et al.*, 1994, Liu *et al.*, 2010, Clayton *et al.*, 2011). For example, knock-down of the mitochondrial acyl carrier protein led to alterations in mitochondrial morphology (Clayton *et al.*, 2011), and kinetoplast replication is compromised when expression of important maintenance proteins are abolished (Hines & Ray, 2010, Liu *et al.*, 2010).

In order to obtain a better understanding of intracellular metabolism during AN5568 treatment, cell pellets were analysed by LC-MS. In BSF cells, upregulation of methionine metabolism was by far the most significant change. This result, in addition to the significant antagonism between simefungin and AN5568, suggests that the benzoxaborole potentially affects an MTase reaction pathway. Given that PCF cells require a higher dose of the compound, we hypothesize that the drug targets a BSF-specific MTase. Several other significant perturbations occurred during AN5568-treatment. For example, a general downward trend was observed in lipid metabolism. In particular, phospholipids such as PC and PE and other metabolites derived from these

phospholipids decreased significantly. Furthermore, modified lysines, such as acetyl-L-lysine and methylated lysines, exhibited significant increases in drug-treated cells. We hypothesise that cells treated with AN5568 respond through the activation of the oxidative stress response pathway. This hypothesis is supported by upregulation of the trypanothione biosynthesis pathway that was observed in metabolomics analysis.

Finally, there were changes in metabolites with unknown roles in the trypanosome, such as aminoacetone, 5-guanidino-2-oxopentanoate and 8-amino-7-oxononanoate. The changes in these metabolites were significant and they warrant further study to understand their role in *T. brucei* metabolism.

Further metabolomics analysis was carried out using stable-isotope labelled  $^{13}\text{C}$ -(U)-L-methionine, in the hope that changes in isotope labelling distributions would be observed. However, only 18 metabolites that incorporated carbon isotopes were observed. There are several reasons for this. Firstly, the experiment was carried out over a period of 6 hours. In this time, AdoMet and AdoHcy incorporated  $^{13}\text{C}$  isotopes, as did methylated L-lysine residues. This suggests that methylation was still occurring in the presence of AN5568. However, metabolites such as methyl-arginine, which is known to receive a methyl group from AdoMet (Lott *et al.*, 2014), was not labelled. Therefore, the experiment could be repeated over a longer duration, with a lower concentration of AN5568, to ensure that cell cultures survive longer. Secondly, the majority of methylation occurs on proteins and nucleic acids. Proteins, apart from short peptides (<3 residues), are not detected with the LC-MS platform, or the metabolite extraction protocol that was used. Therefore, MS such as MALDI-TOF could be employed to investigate patterns of protein distribution, potentially coupled to SILAC for comparative proteomics. This has been carried out successfully in trypanosomatids in the context of phosphorylation detection (Urbaniak *et al.*, 2013).

Several of the observed metabolic changes were followed up by comparative metabolomics analyses using genetically modified cell lines as well as other drug treatments with known targets. For the former, an RNAi construct was generated to knock-down TbCgm1, an essential SL MTase that modifies cap0 (Hall & Ho, 2006a, Takagi *et al.*, 2007). This pathway is unique and essential in *T. brucei*, although the cap1, cap2, cap3 and cap4 methylators are currently not thought to be essential (Zamudio *et al.*, 2006). Once successful knock-down had been confirmed, comparative metabolomics analyses were carried out. These showed that L-methionine and adenine increased drastically, probably due to a decrease in protein synthesis as a result of transcription inhibition. However, there were no changes in AdoMet or 5'-MTA, such as those observed in AN5568-treated cells. Interestingly, knock-down of TbCgm1 resulted in high levels of nucleotides. This supports the theory that SL methylation inhibition significantly

affects the cell's transcriptional machinery, as well as ability to generate mature mRNAs (Zamudio *et al.*, 2009).

To confirm activation of stress responses in AN5568-treated cells, comparative metabolomics analyses were carried out with DTT, a known ER stress inducer. Here, we show that DTT causes significant upregulation of the trypanothione biosynthesis pathway, with increases observed in glutathione and L-ornithine. There was also a significant decrease in trypanothione disulphide, due to the effect of DTT in breaking disulphide bridges (Tiengwe *et al.*, 2015). In addition, the increases in modified lysines seen in AN5568-treated cells were also observed in DTT-treated cells, suggesting these are involved in the oxidative stress response. This supports the hypothesis that AN5568 causes widespread cellular stress responses, and these metabolites might not be directly involved in the benzoxaborole target inhibition. In contrast, no changes were seen in AdoMet, 5'-MTA or adenine, and we hypothesise that these changes can be attributed to the effect of AN5568 target inhibition.

As described in [chapter 1](#), several methods exist for the identification of drug targets. Prior to the study presented here, we were informed that affinity chromatography experiments had been carried out unsuccessfully to identify the AN5568 MoA (R. Jacobs, personal communication). This was likely due to a lack of knowledge concerning SAR. We therefore attempted to carry out drug target identification techniques without the requirement of drug immobilisation, in the form of the DARTS assay. This assay was developed several years ago by Lomenick and colleagues and has been successfully used in mammalian system as well as prokaryotes (Lomenick *et al.*, 2009, Chin *et al.*, 2014). To our knowledge, the data presented here is the first description of the DARTS assay in a protozoan parasite.

Using the DARTS assay, several hits were identified, although none were hypothesised to be involved in the AN5568 MoA. Indeed, proteins such as ALD and the branched chain aminotransferase identified in both protein samples are reported to be highly abundant in the *T. brucei* proteome (M. Barrett, personal communication). One of the proteins identified was a putative methionine biosynthetic protein (table 3-3). Further analyses showed that this protein contained a wybutosine MTase domain. Wybutosine, a heavily modified tRNA, has not been described in *T. brucei* and this gene remains to be characterised, although several publications have described the yeast homologue (Kalhor *et al.*, 2005, Noma *et al.*, 2006). As part of a small overexpression library, this gene (Tb927.1.3080) was overexpressed to look for changes in AN5568 sensitivity. However, none was found. The primase 2 identified in the DARTS assay has been reported to be localised in the *T. brucei* mitochondrion (Hines & Ray, 2011). This study showed that primase 2 was an essential component in kDNA maintenance (Hines & Ray, 2011). It would therefore be of interest to confirm whether or not AN5568 targets this protein.

There are several limitations with the DARTS approach. Firstly, it requires a lengthy optimisation process for each novel organism and/or compound that is tested, because it is unknown whether the unidentified target protein undergoes significant alterations in thermodynamic stability when bound to a ligand or drug. Furthermore, the concentrations of pronase, or proteolysis agent, required differ for each compound. Indeed, we found that the pronase concentration added must be sufficient to degrade most of the proteome, without compromising the drug target protein to the same extent. In addition, there are differing results depending on whether the drug-lysate incubation is carried out in living cells, or with a protein extract. This is especially important if the drug requires activation.

During the DARTS analysis, we found that running AN5568-treated protein extracts directly on a protein generated a band enrichment at ~37 kDa (fig. 3-29). Subsequently, a 2D-DiGE experiment was carried out to enable higher resolution of protein differences between WT and drug-treated cell lysates (fig. 3-30). Through this experiment, it was found that expression of cysteine and serine peptidases, or their precursors, are upregulated in response to AN5568 treatment (table 3-4). Cysteine peptidases in particular have been studied as a potential drug target in *T. brucei* (Troberg *et al.*, 1999). Several studies have shown that these enzymes influence the ability of the cell to differentiate (Williams *et al.*, 2006, Santos *et al.*, 2007). In *T. brucei*, disruption of the natural inhibitor of cysteine peptidase gene leads to increased clan CA cysteine peptidase expression (Santos *et al.*, 2007). Cells possessing increased cysteine peptidase expression were found to undergo more efficient VSG surface coat exchange, although there was no difference in the cells' ability to commit to differentiation (Santos *et al.*, 2007). Whilst further work is required to validate the proteomics findings in AN5568-treated cells, the aforementioned studies suggest that increased cysteine peptidase expression could result from increased autophagy or activation of differentiation pathways in *T. brucei*.

The 2D-DiGE experiment also revealed downregulation of Hsp70 in AN5568-treated *T. brucei*. This gene has been reported to be localised in the glycosome (Guther *et al.*, 2014) and its downregulation could be associated with a decrease in energy metabolism due to the effects of the benzoxaborole. Furthermore, a putative chaperonin Hsp60 that is targeted to the mitochondrion, was upregulated in AN5568-treated cells. These findings again hint at the potential stress that *T. brucei* undergoes upon treatment with AN5568. One note of interest is that several of these proteins are regulated by a *T. brucei* zinc finger protein called ZC3H11 (Droll *et al.*, 2013, Singh *et al.*, 2014). It is therefore possible that AN5568 inhibits a transcription regulator that leads to widespread perturbations in the cell.

During the course of this study, the decision was made to carry out a more detailed study of the *T. brucei* MTase complement, which has not been done for the trypanosomatids. This

complement, known as the “methyltransferome”, identified 145 MTases encoded in the *T. brucei* genome (table 3-6). In similar fashion to both *S. cerevisiae* and *H. sapiens*, the MTases differed in both type/fold and substrate. There were 33 MTases containing SET domains, which are thought to be involved in lysine methylation and in particular, histone methylation (Herz *et al.*, 2013). This number was larger than both of the previous MTome studies (Petrossian & Clarke, 2011, Wlodarski *et al.*, 2011), and could reflect the requirement of *T. brucei* to change global transcription levels in response to differentiation to different life cycle stages. There were also predicted radical SAM MTases, tetrapyrrole methylases, a Hcy MTase, a thymidylate synthetase, a TIM beta/alpha barrel MTase and TYW3, none of which have been described in *T. brucei*. Given the importance and conserved roles of MTases in cellular function, an in-depth description and characterisation of the MTase complement in the tri-tryps would be beneficial to the field.

The final experiment in this chapter was to overexpress 12 essential putative methyltransferases and test the sensitivity of these overexpressors to both AN5568 and the MTase inhibitor sinefungin. Using an ornithine decarboxylase overexpressor, we were able to show that overexpression of a drug target generates resistance to the specific drug (fig. 3-32). This has previously been reported in the context of methionyl-tRNA synthetase inhibitors (Ranade *et al.*, 2013). Unfortunately, none of the 12 overexpression lines generated resistance against AN5568, although 4 of these were significantly desensitised to sinefungin. This is most likely due to higher abundance of a protein “absorbing” or “mopping” excess sinefungin. We therefore conclude that none of these 12 proteins are targeted by AN5568.

The volume of data produced during this study was significant, and several datasets were not fully discussed but instead, the results deemed important in the context of AN5568 target identification were presented in a way that highlights key findings of each mass spectrometry experiment. Each dataset was carefully analysed by looking at individual pathways that were hypothesised to have undergone perturbations, and comparative metabolomics analyses proved highly beneficial to dissecting the mode of action of AN5568. The next step of the project was to generate and characterise an AN5568-resistance cell line with the expectation that this could greatly aid in the identification of the benzoxaborole MoA in *T. brucei*.

## Chapter 4. Generation and characterisation of an AN5568-resistant cell line

---

### 4.1 Introduction

---

As described in the chapter 1, several methods exist that can be utilised to investigate drug modes of action. Some were employed in chapter 3, with reasonable success. However, when there is limited information regarding the SAR of a compound, as often occurs with candidates selected from large multi-compound screens, many of these methods are rendered somewhat useless. This is especially true with affinity chromatography, which requires docking of compounds to affinity beads, before a protein lysate pull-down is performed (Denise *et al.*, 1999).

One method that has been used with great success on many occasions is the generation and selection of a drug-resistant cell line, and subsequent investigation of the biological and molecular changes that have occurred to cause increased drug-resistance. This was the method used by Rock and colleagues to elucidate the mechanism by which the anti-fungal AN2690 inhibits tRNA<sup>LEU</sup> synthetase (Rock *et al.*, 2007). Generating resistant cell lines has also led to many significant discoveries in *T. brucei* and the related tri-tryps, *Trypanosoma cruzi* and *Leishmania spp.* (Campos *et al.*, 2013, Berg *et al.*, 2015).

For example, loss of the P2 transporter was associated with resistance against the mitochondrial targeting DB75 (Lanteri *et al.*, 2006). Another study by Ranade and colleagues found that resistance against methionyl-tRNA targeted drugs was generated by high levels of target protein expression (Ranade *et al.*, 2013). Furthermore, both the MacLeod and Barrett groups have studied drug resistance in detail in both African trypanosomes and *Leishmania*, especially in the cases of melarsoprol and amphotericin B respectively (A. MacLeod & M. Barrett, personal communication).

It is important to note that especially in the case of *T. brucei*, resistance is often generated by the downregulation, or loss of expression, of uptake transporters specific to the drug, as was demonstrated by Vincent and colleagues when they characterised an Eflornithine-resistant *T. brucei* line (Vincent *et al.*, 2010). In these cases, whilst important discoveries are made with regards to potential resistance mechanisms in the field, it does not aid in furthering our understanding of the drug mode of action, as the drug is no longer taken up by the cell. Whilst this method of resistance sometimes has a detrimental effect on growth-rates, it allows the

parasite to survive in high concentrations of the drug in question as shown by the aforementioned studies.

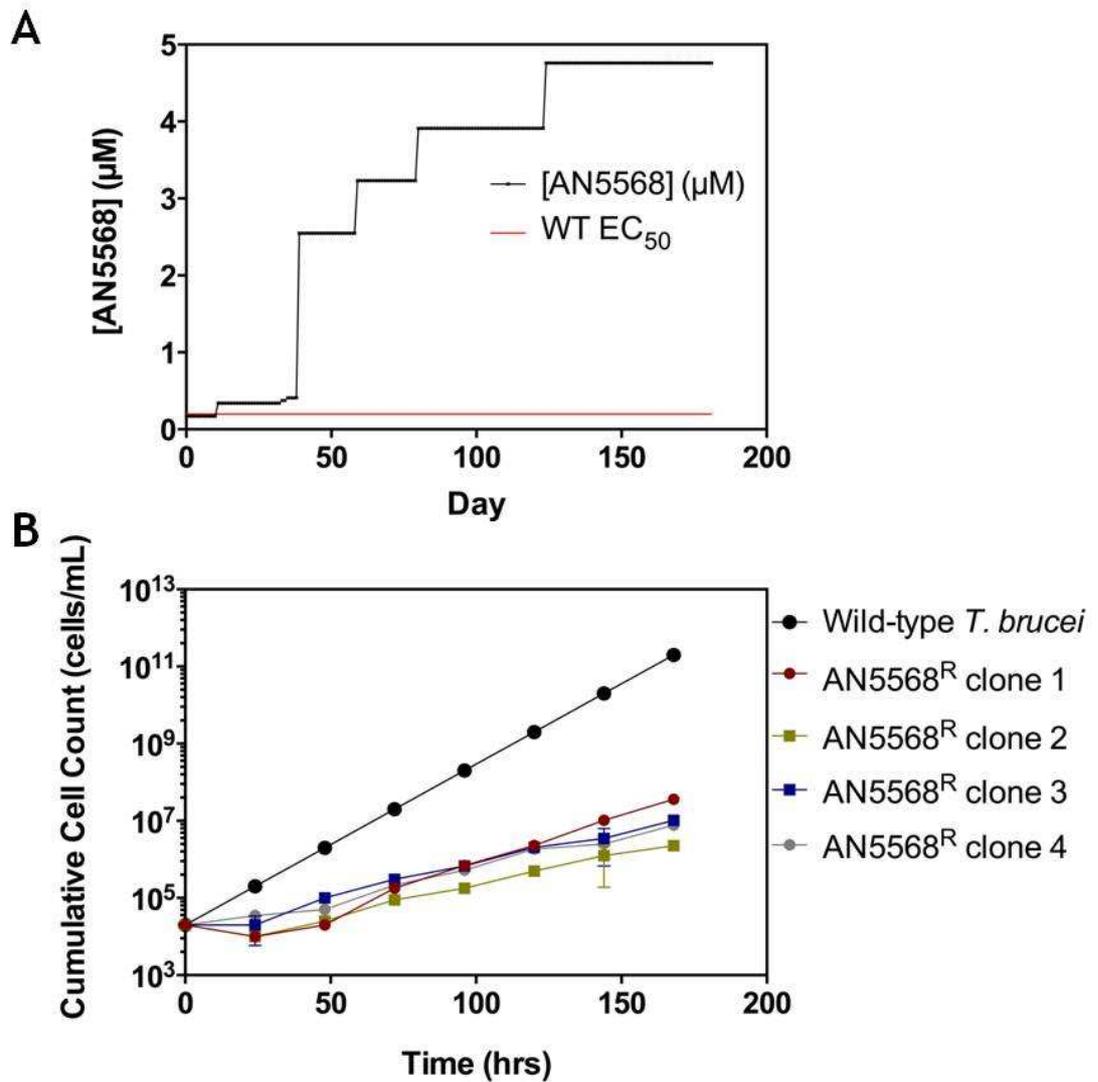
Combined with the modernisation of technologies allowing the characterisation of new cell lines, generating drug resistance is a simple (albeit painstakingly slow), but effective tool to understand drug mode of action. Therefore, we decided to utilise this experimental technique, in order to try and gain a better understanding of the mechanisms by which AN5568 kills *T. brucei*, and if and how the parasite could generate resistance against it. Building on the knowledge and hypotheses developed in the earlier chapter, we developed a targeted approach. Again, we utilised the expertise and specialities of the Barrett and MacLeod groups, namely metabolomics, genomics and transcriptomics, to characterise this resistant cell line.

## 4.2 Generation of AN5568 in *T. brucei* cells *in vitro*

---

The project, in keeping with previous work, utilised the Lister 427 strain of *T.b. brucei* as a parental strain. With the knowledge of the EC<sub>50</sub>, cells were grown in the presence of very low concentrations of AN5568 initially. As we did not know how well cells would grow, the experiment was initiated by keeping cells in several wells of a 24-well plate at varying concentrations of the benzoxaborole. The highest concentration that allowed growth of cells initially, was ~170 nM. This was just under the EC<sub>50</sub> value reported in Chapter 3 ([section 3.2](#)). From here on, higher concentrations of AN5568 were added if and when cells were deemed to be resistant to the lower concentrations (fig. 4-1A). After a period of ~200 days, cells were stably growing in 4.8 μM AN5568, approximately 24-fold the EC<sub>50</sub> of wild-type cells. At this point, cells were cloned by dilution, and 4 clones were retrieved for further investigation.

Growth analysis was carried out on the AN5568<sup>R</sup> cell line, to assess whether there were any differences compared to wild-type cells (fig. 4-1B). This experiment showed that all resistant clones had a very slow growth rate compared to wild-type Lister 427 *T.b. brucei*. This suggests the resistance comes at a fitness cost, resulting in slower growth and potentially, slower metabolism.

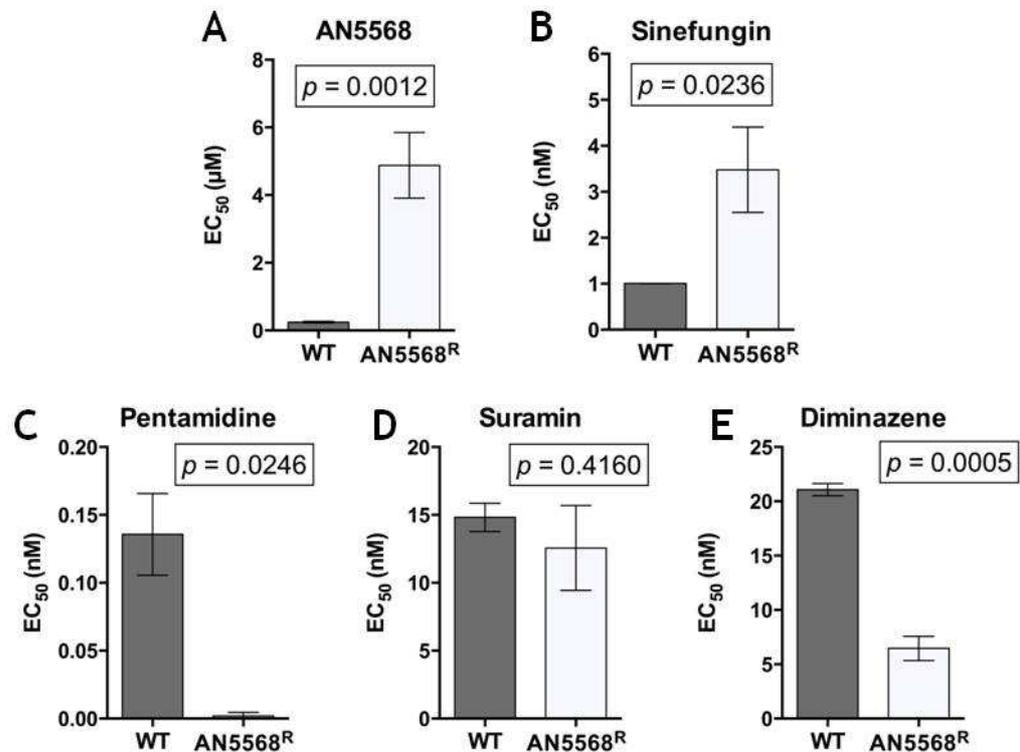


**Figure 4-1: Generation of an AN5568-resistant cell line.** A) Resistance was generated in a Lister 427-derived cell line, by adding increasing increments of the benzoxaborole to cell cultures *in vitro*. Whilst parasites could only grow at very low concentrations ( $<\text{EC}_{50}$ ) and a very slow rate initially, there was a significant increase in the concentration they could tolerate, after approximately 46 days, suggesting a resistance mechanism was generated at this point. Red line indicates mean wild-type  $\text{EC}_{50}$ . B) Once cells were thought to have a stable resistance phenotype, they were cloned by dilution, and 4 clones were assessed for their growth. Whilst separate AN5568<sup>R</sup> clones grew at similar rates, overall this was much slower than WT Lister 427 cells.

To test whether the observed resistance against AN5568 was indeed quantitative, we carried out Alamar blue assays with the AN5568<sup>R</sup> line and a WT Lister 427 control (fig. 4-2). Overall, the AN5568<sup>R</sup> clones showed a 15-20-fold increase in  $\text{EC}_{50}$  value against the benzoxaborole, thereby proving the cell lines were resistant, especially compared to wild-type cells (fig. 4-2A). Further analyses were also performed with other drugs, to test any cross-resistance that might have arisen.

In Chapter 3 (section 3.4.5), the contrasts and similarities between the modes of action of sinefungin and AN5568 were discussed in detail. We were therefore very interested to find that the AN5568<sup>R</sup> lines also exhibited some resistance against sinefungin, albeit at a lesser extreme than the benzoxaborole (fig. 4-2B). There was a 4-fold increase in the sinefungin  $\text{EC}_{50}$

concentrations between wild-type cells and the AN5568<sup>R</sup> lines, which was significant by t-test ( $P=0.0236$ ).



**Figure 4-2: AN5568<sup>R</sup> undergoes several changes in drug sensitivity.** A) As expected, the AN5568<sup>R</sup> line, once cloned and assayed for drug resistance using an alamar blue assay, showed a significantly increased EC<sub>50</sub> compared to wild-type cells. Indeed, the mean EC<sub>50</sub> was  $4.88 \pm 0.56 \mu\text{M}$ , approximately 19-fold higher than WT cells. B) Interestingly, AN5568<sup>R</sup> also showed an increased resistance against the AdoMet-dependent MTase inhibitor sinefungin, which was shown to act in similar fashion to AN5568 (chapter 3). Three other drugs were assayed with the AN5568<sup>R</sup> line. Whilst pentamidine (C) and diminazene (E) both showed increased sensitivity compared to WT cells, there was no significant change in suramin (D) resistance.

Further analyses were carried out using pentamidine, suramin and diminazene (fig. 4-2). Whilst resistant cells showed hypersensitivity to pentamidine (fig. 4-2C) and diminazene (fig. 4-2E), there was no significant change in suramin resistance. It was interesting to observe that the generation of resistance to AN5568 also resulted in altered sensitivity to several other drugs. In the cases of pentamidine and diminazene, resistance has previously been shown to result from altered expression of cell surface receptors (de Koning *et al.*, 2004, Bridges *et al.*, 2007). Therefore, we were interested to find out whether a similar phenomenon had occurred in the case of AN5568-resistance. In addition, it would be interesting to know whether resistance to sinefungin could also be due to a loss of transporter, as this has not been investigated before. Having established changes in drug sensitivity and *in vitro* growth of the benzoxaborole-resistant line, we next sought to characterise the cell line using omics-based approaches, to follow up on these objectives, and to generate new hypotheses about the mechanisms involved in AN5568 resistance.

### 4.3 RNA-seq analysis of the AN5568<sup>R</sup> cell line

---

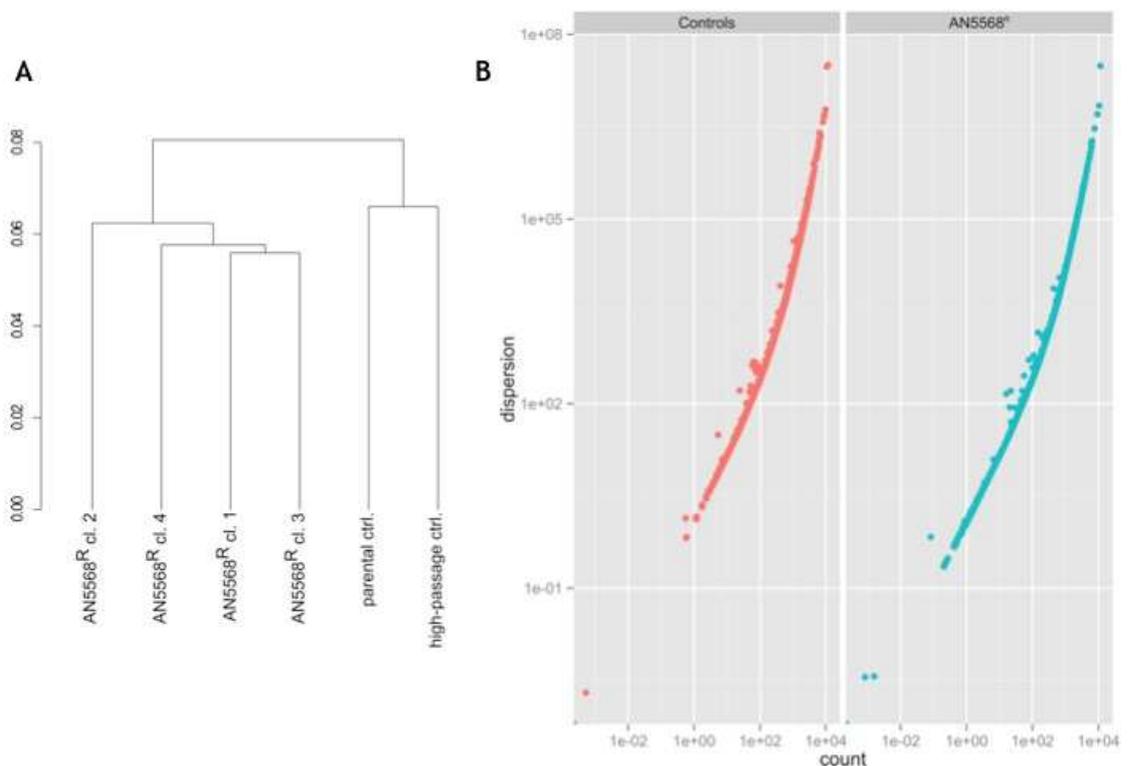
The use of RNA-seq, which allows the observation of the transcriptome, has both advantages and disadvantages compared to DNA-seq. The transcriptome is known to be significantly more dynamic than the genome with regards to environmental responses (Macaulay & Voet, 2014). Therefore, to get a more accurate picture of the cellular response to drugs, as well as changes that render cells drug-resistant, RNA-seq was judged to be preferable to DNA-seq. However, sequencing analysis would be more difficult as only coding regions and their untranslated regions could be analysed for polymorphisms and other forms of mutation. RNA-seq was also chosen instead of other transcriptome analysis techniques such as microarray or EST/cDNA sequencing due to the high resolution offered by this platform, in addition to the fact it enables high-throughput analysis with automated bioinformatics pipelines similar to those that are already in use for DNA-seq analysis (Mortazavi *et al.*, 2008).

For this experiment, 6 RNA samples were prepared: Four AN5568<sup>R</sup> clones, along with an RNA sample from the parental line from which the AN5568<sup>R</sup> line was derived. This sample was isolated before generation of resistance was initiated. In addition, another RNA sample was prepared from a cell line that underwent a high number of passages without the benzoxaborole. This was done in order to rule out false positive mutations that might arise from adaptations to long-term *in vitro* tissue culture. The 6 samples were submitted to Glasgow Polyomics for RNA sequencing. Libraries were prepared using the Illumina TruSeq mRNA library prep kit, which involves a poly-adenylation targeting step to ensure enrichment of mRNAs. Paired end sequencing (2×75 bp) was carried out using the Illumina NextSeq 500 platform (Illumina®).

The resulting raw data was processed and analysed with the aid of the Wellcome Trust Centre for Molecular Parasitology Bioinformatics team. This was done using the bioinformatics pipeline described in the methods. Briefly, all reads were trimmed to remove adapters, and a quality control was carried out on each sample using FastQC (Andrews, 2010). Once the sequencing quality was judged to be satisfactory for each sample, processing was done using the TopHat pipeline (Trapnell *et al.*, 2009), and transcriptome analysis was carried out using the cufflinks package (Trapnell *et al.*, 2010). Finally, the cummeRbund tool package was used to visualise the data (Trapnell *et al.*, 2012). The widely used reference genome as well as transcriptome for the TREU 927 strain (v9.0) was used for RNA-seq alignments. This reference genome is commonly used as it provides by far the best assembly and annotation, although it does contain some differences compared to the Lister 427 cell line (Brems *et al.*, 2005). Therefore, we anticipated a high number of polymorphisms (SNPs and indels) arising from this analysis. This is discussed in detail in [section 4.4](#).

A dendrogram generated using cummeRbund showed that both controls were more closely correlated to each other than any of the AN5568<sup>R</sup> samples, in terms of differential gene expression data (fig. 4-3A). Importantly, this meant that long-term *in vitro* culture had not significantly altered the high-passage control, as it still clustered with the parental control. This led to the assumption that the genetic or transcriptomic drift caused by drug resistance was greater than the effect of long-term culture.

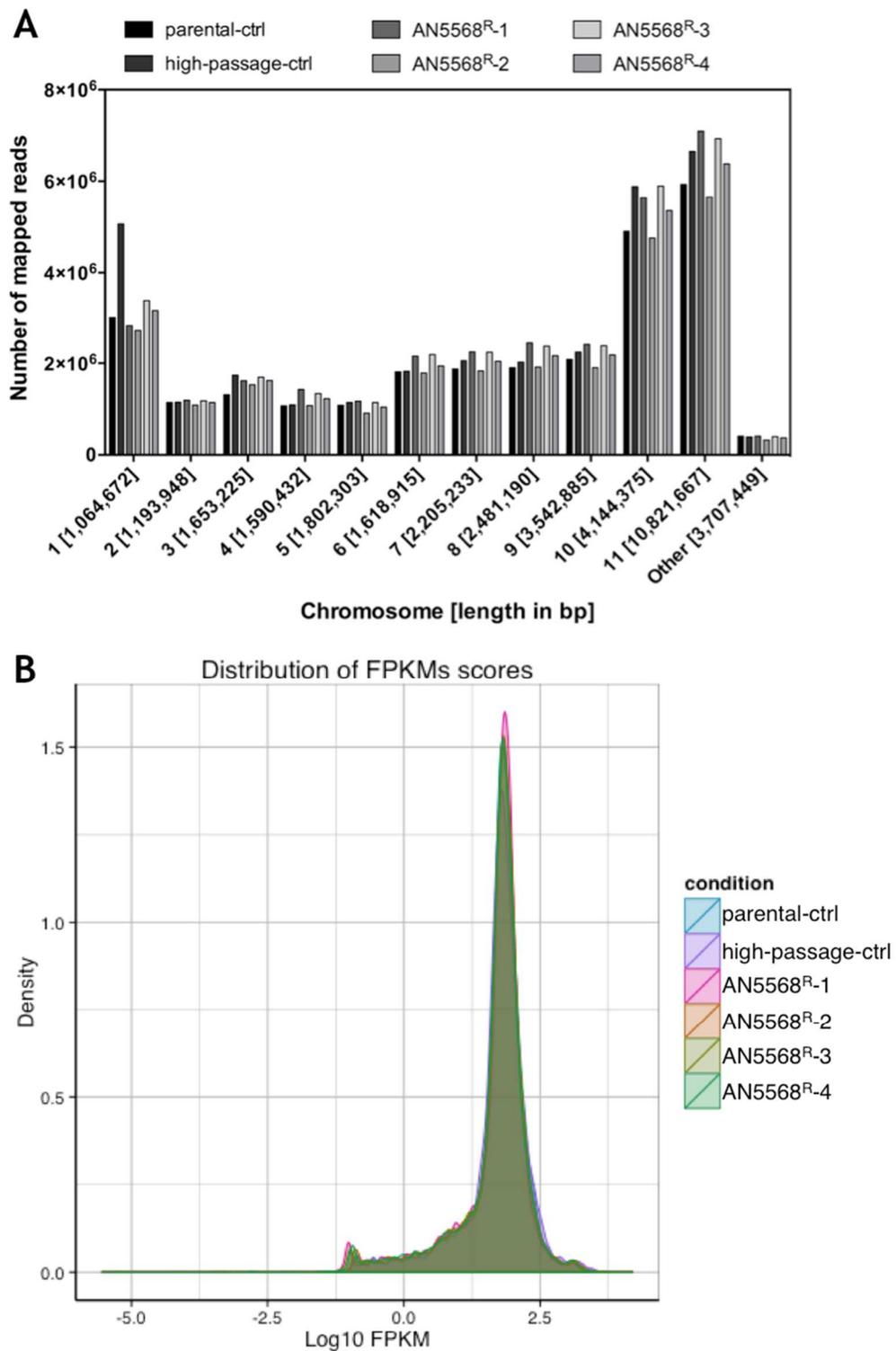
In addition, dispersion, a measure of the quality of the data, of the two sample groups was analysed (fig. 4-3B). Here, deviation from the threshold (dispersion) was plotted against the read count in FPKM (fragments per kilobase of transcript per million mapped reads). Thus, with high FPKM, a specific fragment, or gene, is hypothesised to show increased dispersion. This is what was found from the RNA-seq data generated from the 6 cell lines. Indeed, both sample groups showed similar dispersion patterns, and no over-dispersion was observed in either sample group.



**Figure 4-3: RNAseq data quality analysis.** A) A dendrogram showed clustering of samples was the same as hypothesised. Whilst all four AN5568<sup>R</sup> clones clustered together, they were found to be dissimilar to either control. Similarly, both controls clustered together, suggesting that high passage number had minimal effect on gene expression profiles. B) A plot of dispersion against FPKM counts showed the data to be of high quality with little over-dispersion observed in either sample group.

The next step was to analyse the mapping of the sequencing reads (fig. 4-4). In total, just under 174 million reads were mapped from the 6 samples. Between 25 and 32 million reads were mapped to the genome for each sample. Coverage of the reads was calculated using the bedtools package which found a median of 100% coverage (measured as percentage of bases

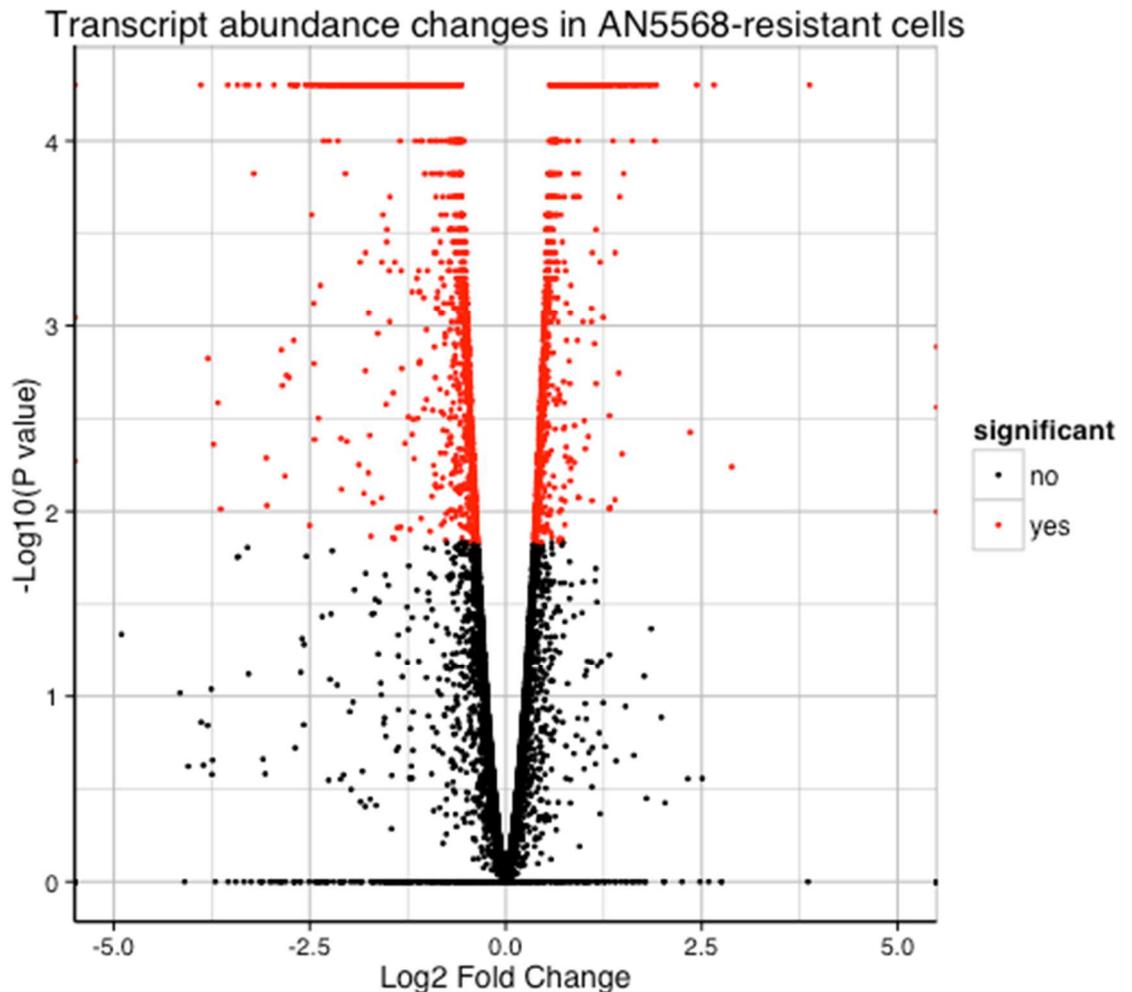
for each gene that had non-zero coverage). The mean percentage coverage of the transcriptome was  $81.33 \pm 0.32\%$ , although this included a large number of contigs and transcripts corresponding to VSGs, which are not usually mapped. In addition, it was important to ensure that reads mapped across the genome in a similar fashion in all samples. Using samtools, the numbers of reads mapped on each individual chromosome was assessed (fig. 4-4A). These data showed that each sample had similar numbers of reads mapped to each chromosome. The two largest chromosomes, 10 and 11 (4,144,375 bp and 10,821,667 bp respectively) had the highest numbers of mapped reads. Whilst the *T. brucei* genome is known to contain 11 chromosomes, the assembled genome thus far contains numerous extra unassembled contigs, which appear in datasets as chromosomes by themselves (Berriman *et al.*, 2005). In this experiment, these contigs are presented under the “other” genome sequences. As shown in figure 4-4A, these sequences only make up a small portion of the RNA-seq assembly. It was important to confirm that the number of reads matched to each chromosome was similar in each sample. Taken together, we concluded from these analyses that the rate of false positives would be limited and differential expression would be due to variation between samples, not RNA quality or sequencing errors.



**Figure 4-4: RNA-seq sequencing data analysis of the AN5568-resistant cell line.** A) The samtools package was used to observe the numbers of reads mapped to each chromosome in the *T. brucei* genome. For each sample, similar numbers of reads were mapped to each chromosome, with the highest number being identified as belonging to chromosomes 10 and 11. The fewest number of reads were mapped to extra-chromosomal contigs. B) A density plot of the Log<sub>10</sub> FPKM values was plotted for all 6 samples. This showed significant overlap of all samples, indicating high quality of reads.

A density plot, or smooth histogram, was generated using kernel density estimation to map the distributions of RNA-seq read counts in FPKM. The resulting figure showed a high amount of overlap between samples, indicative that the data were comparable in terms of their quality. In addition, the data distribution appeared to be unimodal.

The reads obtained by RNAseq mapped to 11,563 genes. Of these, 2422 genes showed a significant change in transcript abundance based on the P-value alone, with a threshold of  $P=0.05$  (fig. 4-5). This number was more than expected, mainly because this analysis also included changes in transcript abundance below 2-fold. Therefore, we set out to refine the dataset and identify the significantly changed genes.



**Figure 4-5:** Volcano plot of changes in transcript abundance in the AN5568<sup>R</sup> cell line. The RNAseq transcriptome dataset was analysed using the R package cummeRbund. Across the transcriptome, 2422 genes were found to exhibit significant changes in transcript abundance in the AN5568<sup>R</sup> line (red dots), compared to a combination of a parental, and high-passage controls. As can be seen here, many of the significant hits did not show large Log<sub>2</sub> fold changes. There were 11 genes that had a Log<sub>2</sub> fold change above 5 or below -5, and therefore these are not shown.

To refine the data set containing all significant hits further, we selected all genes that showed a Log<sub>2</sub> fold-change above 1, or below -1, equating to a fold change of 2 or 0.5 respectively. This led to datasets of 126 genes showing significant upregulation, and 371 genes showing significant down-regulation. These are shown in the supplementary data (S4). The top 15 hits for both upregulation and downregulation are presented in tables 4-1 and 4-2 respectively.

### 4.3.1 PCF-associated genes are upregulated

Strikingly, many of the genes shown to be significantly upregulated ( $P < 0.05$ ,  $\text{Log}_2$  fold change  $= > 1$ ) in the AN5568<sup>R</sup> line were PCF-form specific genes. In particular, GPEET and EP procyclins, both surface markers of early and late-differentiated PCFs respectively, were found at high levels of transcript abundance (table 4-1). The assembled transcriptome was analysed in detail to find the reads for these proteins, which were localised on chromosome 6 (fig. 4-6). Indeed, almost no reads were mapped to procyclins for the two control samples, whilst all 4 resistant clones showed a high number of mapped reads in this region.

GeneID	Product Description	Control FPKM	AN5568 <sup>R</sup> FPKM	$\text{Log}_2$ (fold_change)	p_value	q_value
Tb927.6.510	GPEET2 procyclin precursor	112.276	1648.920	3.8764	0.00005	0.00045
Tb927.5.5530	Variant Surface Glycoprotein, putative	0.339	2.505	2.8842	0.00575	0.02376
Tb927.5.4020	hypothetical protein	35.287	222.540	2.6569	0.00005	0.00045
Tb927.10.10260	EP1 procyclin (EP1)	25.538	138.305	2.4371	0.00005	0.00045
Tb927.10.10250	EP2 procyclin (EP2)	2.785	14.234	2.3534	0.00375	0.01693
Tb927.9.7470	purine nucleoside transporter (NT10)	45.384	172.052	1.9226	0.00005	0.00045
Tb927.10.10760	hypothetical protein	13.029	48.740	1.9033	0.00010	0.00082
Tb927.11.2410	hypothetical protein, conserved	43.069	159.920	1.8926	0.00005	0.00045
Tb927.11.6280	pyruvate phosphate dikinase (PPDK)	70.542	254.852	1.8531	0.00005	0.00045
Tb927.3.5790	expression site-associated gene (ESAG, pseudogene), putative,	36.384	129.655	1.8333	0.00005	0.00045
Tb927.8.6170	transketolase, putative (TK)	59.517	210.032	1.8192	0.00005	0.00045
Tb927.10.12780	ZC3H37	6.190	21.829	1.8183	0.00005	0.00045
Tb927.1.2820	pteridine transporter, putative	12.890	43.965	1.7701	0.00005	0.00045
Tb927.6.2890	single strand-specific nuclease, putative	42.781	143.279	1.7438	0.00005	0.00045
Tb927.4.3500	Amastin surface glycoprotein, putative	133.087	436.540	1.7137	0.00005	0.00045

Table 4-1: Top 15 genes significantly upregulated in the AN5568<sup>R</sup> line.

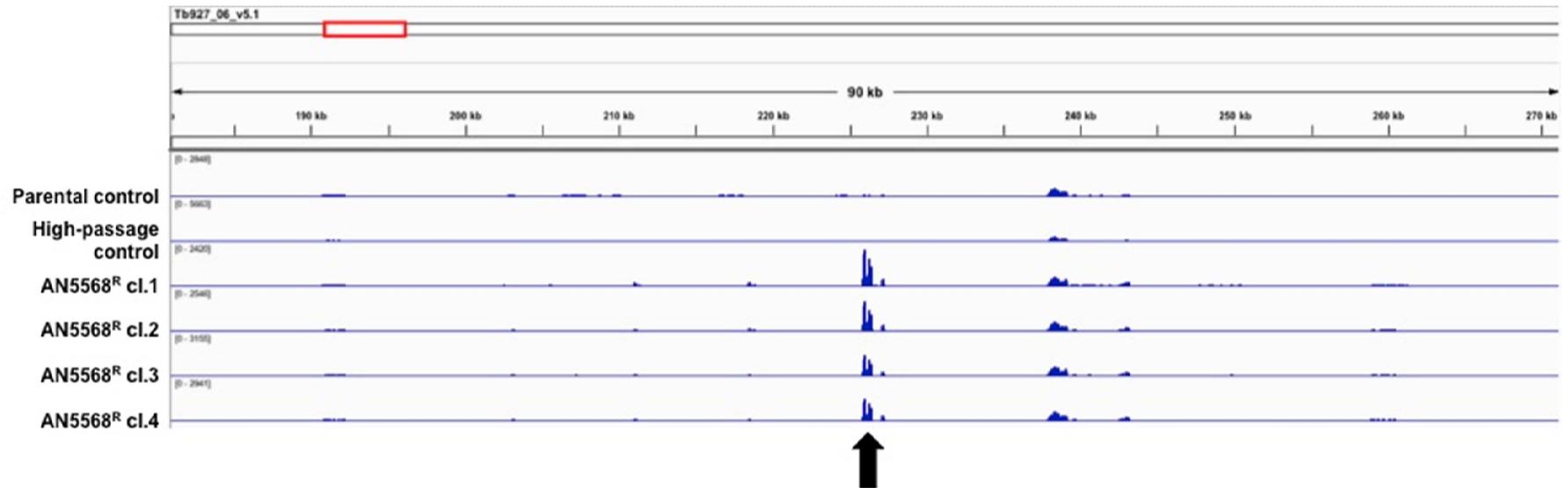
Transporter expression levels are known to change over the course of differentiation from BSF to stumpy-forms and subsequently to PCFs. Several analyses of transcriptomic changes that occur during the differentiation process have been published (Siegel *et al.*, 2010, Capewell *et al.*, 2013b). Both these studies reported increased expression of the purine transporter identified here (Tb927.9.7470) in SSFs and PCFs. Moreover there was increased expression of the pteridine transporter (Tb927.1.2820) in both these differentiated forms, as was found in the benzoxaborole-resistant cells.

A predicted CCCH-domain containing protein (Tb927.10.12780) was identified. Interestingly, this protein was not found to be differentially expressed in studies of *T. brucei* differentiation

(Siegel *et al.*, 2010, Capewell *et al.*, 2013b). It is unknown what increased expression of this gene could lead to, but its domains suggest it plays an important role in mRNA regulation.

Further increases that corresponded to differentiation phenotypes were found with a putative pyruvate phosphate dikinase (Tb927.11.6280) and a glycosomal transketolase (Tb927.8.6170) previously shown to be expressed exclusively in PCF *T. brucei* (Stoffel *et al.*, 2011). In addition, a nuclease (Tb927.6.2890) and an amastin surface glycoprotein (Tb927.4.3500) were also upregulated, although these were found not to be differentially expressed in the Siegel study (Siegel *et al.*, 2010).

The upregulated genes also included a VSG, which was in contrast to the main trend of PCF-specific genes. However, the FPKM values for both sample groups were low for this gene, and could be a result of improper alignment, given the high mutation rates in VSG genes. Three hypothetical proteins were identified and as such, were not followed up.



**Figure 4-6: RNAseq coverage of chromosome 6.** RNAseq data output from the cufflinks tool suggested increased abundance of both GPEET and EP procyclins, which are known to be upregulated in PCF trypanosomes. These genes are found in an array on chromosome 6 in the *T. brucei* genome. Coverage of the relevant portion of chromosome 6 was analysed using the IGV software (Robinson *et al.*, 2011), and is shown here. In this result, which mirrors that of the transcript analysis, coverage analysis showed a similar increase in transcript abundance around the procyclin loci (black arrow). Whilst both the controls showed minimal number of reads mapped to this portion of the genome, there were similarly high increases in all four AN5568<sup>R</sup> clones.

### 4.3.2 Downregulation of BSF-specific genes

A similar table was generated for the 15 genes showing highest amounts of downregulation in the AN5568<sup>R</sup> line (table 4-2). Here, we found widespread downregulation of VSGs and their associated expression site associated genes. In addition, a retrotransposon hot spot protein was also found to be downregulated.

Again, there were similarities between these data and the aforementioned transcriptomics datasets discussing differentiation in *T. brucei*. A nucleoside transporter (Tb927.9.15980) was shown to exhibit decreased transcript abundance in stumpy forms, but not in PCF trypanosomes. A putative potassium voltage-gated channel (Tb11.v5.0396) showed decreased transcript abundance, but was not reported in either the Siegel (Siegel *et al.*, 2010), or the Capewell (Capewell *et al.*, 2013b) study. Furthermore, a leucine-rich repeat protein was found to decrease significantly here, but not in stumpy forms, although it did have decrease abundance in PCF cells according to the Siegel study.

Overall, these data indicate significant cell surface remodelling, albeit at the transcript level, with the loss of VSG and related expression site genes. In addition, there is a corresponding increase in procyclins. In addition there is upregulation of transketolase which could suggest increased PPP metabolism. Finally, there are changes in transporters which could have a profound impact of uptake of metabolites, as well as drugs. Indeed, changes in these transporters could potentially be responsible for the increased sensitivity of the AN5568<sup>R</sup> line to diamidines such as pentamidine and diminazene, as is discussed in more detail in [section 4.3.3](#).

GeneID	Product Description	Control FPKM	AN5568 <sup>R</sup> FPKM	log <sub>2</sub> (fold_change)	p_value	q_value
Tb09.v4.0136	variant surface glycoprotein (VSG)	112.834	0.236	-8.8991	0.00975	0.03612
Tb927.2.1380	leucine-rich repeat protein (LRRP)	30.159	0.424	-6.1539	0.00010	0.00082
Tb927.9.7400	expression site-associated gene 11 (ESAG11)	35.798	2.416	-3.8892	0.00005	0.00045
Tb927.9.16010	expression site-associated gene 4 (ESAG4)	29.454	2.117	-3.7987	0.00150	0.00799
Tb927.2.3330	hypothetical protein	20.694	1.562	-3.7278	0.00435	0.01903
Tb927.1.520	variant surface glycoprotein (VSG)	9.346	0.733	-3.6720	0.00260	0.01249
Tb927.5.160	retrotransposon hot spot protein (RHS)	12.345	0.993	-3.6354	0.00970	0.03595
Tb927.6.170	receptor-type adenylate cyclase (GRESAG4)	6.239	0.534	-3.5455	0.00005	0.00045
Tb11.v5.0396	potassium voltage-gated channel	16.682	1.557	-3.4211	0.00005	0.00045
Tb927.3.5750	retrotransposon hot spot protein (RHS)	42.142	4.211	-3.3229	0.00005	0.00045
Tb927.2.3340	hypothetical protein	138.552	14.271	-3.2793	0.00005	0.00045
Tb09.v4.0018	expression site-associated gene (ESAG)	13.292	1.432	-3.2145	0.00015	0.00117
Tb927.9.7290	variant surface glycoprotein (VSG)	23.027	2.594	-3.1503	0.00005	0.00045
Tb927.9.15980	nucleoside transporter 1, putative	3.611	0.435	-3.0544	0.00515	0.02174
Tb05.5K5.290	hypothetical protein	6.141	0.743	-3.0463	0.00930	0.03485

**Table 4-2: Top 15 genes significantly downregulated in the AN5568<sup>R</sup> cell line.** A significant portion of downregulated genes were identified as BSF-specific proteins, including VSGs and VSG expression site associated genes.

### 4.3.3 Differential expression of *T. brucei* transporters

As mentioned previously, drug-resistance in *T. brucei* can be generated due to differential expression of transporters. We therefore analysed whether any transporters exhibited altered transcript abundance in the AN5568<sup>R</sup> cell line. The results are shown in table 4-3.

The gene set was chosen based on a GO-term search of “transport” on the TriTryp database. Genes annotated as “hypothetical” were removed from the search results, leaving a gene set comprising of 313 genes. Analysis of expression levels for these genes was carried out through manual filtering of the data, as well as analysis using the cummeRbund package for R (Trapnell *et al.*, 2012).

A total of 24 identified and putative transporters underwent significant changes in transcript abundance, based on both the *P* value as well as the fold change (log<sub>2</sub> fold change = >1 or <-1). Of these, 11 were found to have a higher transcript abundance in the AN5568<sup>R</sup> line, whilst the remaining 13 genes were shown to have decreased transcript abundances (table 4-3).

The transporter that underwent the highest increase, a purine nucleotide transporter (Tb927.9.7470), was previously shown to play a role in nifurtimox efficacy and resistance, as detected by RIT-seq analyses (Alsford *et al.*, 2012).

Gene ID	Product description	Control FPKM	AN5568 <sup>R</sup> FPKM	log <sub>2</sub> (fold_change)	p value	q value
Tb927.2.5410	ABC transporter family-like protein	33.7451	13.2269	-1.3512	0.00005	0.00045
Tb927.2.6150	adenosine transporter 2 (TbNT2/927)	43.8016	17.4699	-1.32611	0.00005	0.00045
Tb927.2.6200	adenosine transporter 2, putative (TbNT3)	13.023	5.11097	-1.34939	0.00005	0.00045
Tb927.2.6220	adenosine transporter 2, putative (TbNT4)	65.725	10.3236	-2.6705	0.00005	0.00045
Tb927.2.6320	adenosine transporter 2, putative (TbNT6)	47.2069	7.01062	-2.75138	0.00005	0.00045
Tb927.8.7670	amino acid transporter, putative	25.3989	56.7563	1.16001	0.00005	0.00045
Tb927.8.7740	amino acid transporter, putative	84.4543	226.075	1.42056	0.00005	0.00045
Tb927.8.8300	amino acid transporter, putative	39.0075	99.0958	1.34507	0.00005	0.00045
Tb927.4.4730	amino acid transporter, putative (AATP11)	109.314	314.637	1.52522	0.00005	0.00045
Tb927.10.14160	Aquaglyceroporin 3 (AQP3)	489.331	155.568	-1.65327	0.00005	0.00045
Tb927.5.3400	calcium-translocating P-type ATPase, calcium pump	176.261	78.7892	-1.16164	0.00005	0.00045
Tb927.6.2200	DJ-1 family protein, putative	237.118	115.203	-1.04142	0.00005	0.00045
Tb927.10.8460	glucose transporter, putative	56.438	11.3104	-2.31902	0.00005	0.00045
Tb927.6.440	haptoglobin-hemoglobin receptor (HpHbR)	160.353	68.9258	-1.21813	0.00005	0.00045
Tb927.11.3620	nucleobase/nucleoside transporter 8.1 (NT8.1)	19.8014	47.1702	1.25227	0.00005	0.00045
Tb927.9.15980	nucleoside transporter 1, putative	3.61119	0.434711	-3.05435	0.00515	0.02174
Tb09.v4.0106	nucleoside transporter 1, putative, chrIX additional, unordered contigs	4.64334	0.643949	-2.85015	0.00210	0.01052
Tb927.7.5930	Protein Associated with Differentiation (PAD1)	68.5347	148.767	1.11815	0.00005	0.00045
Tb927.7.5940	Protein Associated with Differentiation (PAD2)	92.4903	256.789	1.47321	0.00005	0.00045
Tb927.7.5950	protein associated with differentiation 3, putative (PAD3)	15.9397	31.8947	1.0007	0.00005	0.00045
Tb927.1.2820	pteridine transporter, putative	12.8901	43.9649	1.77009	0.00005	0.00045
Tb927.1.2880	pteridine transporter, putative	14.4442	43.9246	1.60453	0.00005	0.00045
Tb927.9.7470	purine nucleoside transporter (NT10)	45.3842	172.052	1.92259	0.00005	0.00045
Tb11.v5.0330	THT1 - hexose transporter, putative	48.3846	7.48095	-2.69326	0.00005	0.00045

**Table 4-3: Transporter proteins that exhibited significant differential expression in AN5568-resistant *T. brucei* cells, shown alphabetically.**

The NT8 nucleobase transporter that was observed to have a higher abundance (table 4-3, gene ID: Tb927.11.3620) was also recently shown to be upregulated in late log phase PCF cells (Fernandez-Moya *et al.*, 2014), thereby supporting the hypothesis that these cells were showing PCF phenotypes.

The most significant finding from the transporter dataset was undoubtedly the increased expression levels of 3 PAD proteins, which have been implicated in playing major roles in *T. brucei* differentiation (Dean *et al.*, 2009). Whilst PAD3 was not upregulated to the same degree as the other PAD receptors, the increase was still significant statistically. PAD1 and PAD2 showed much

higher transcript abundance, thereby concretely supporting our hypotheses that the AN5568<sup>R</sup> line had undergone, or was undergoing, a differentiation event.

There were several interesting transporters showing decreased abundance in AN5568<sup>R</sup> cells. An ATP-binding cassette (ABC) transporter family-like protein (Tb927.2.5410) was previously shown to localise to the outer mitochondrial membrane (Niemann *et al.*, 2013), and was present at lower levels in the resistant cell line. In addition, AQP3, a transporter implicated in suramin efficacy (Alsford *et al.*, 2012), was also present in decreased abundance. The AN5568<sup>R</sup> cell line exhibited a decrease in mean EC<sub>50</sub> against suramin (fig. 4-2), but this was found not to be significant by a *t*-test.

Further transcripts with decreased abundance in resistant cells included glucose transporters, as well as several other adenosine and nucleotide transporters (table 4-2). The glucose transporters in particular were an interesting finding, because they support the hypothesis that these cells are similar to PCF cells with regards to their transcriptome. The results were also in agreement with the previous studies that carried out comparative analyses of BSF, stumpy and PCF parasite transcriptomes (Siegel *et al.*, 2010, Capewell *et al.*, 2013b).

The data presented here, in addition to the differential expression data in general, indicate that the benzoxaborole-resistant cell line possess a transcriptome that resembles a PCF or stumpy *T.b. brucei* phenotype. Investigations into transporter expression in resistant cells requires molecular biology to confirm whether they are responsible for the PCF-like phenotype these cells possess, as well as whether the differential expression can explain any changes in the parasites' sensitivity to particular trypanocidals (fig. 4-2). In addition, these expression differences also suggest the AN5568<sup>R</sup> cell line might show altered sensitivity to drugs such as nifurtimox, and overexpression of the pteridine transporter could compensate for pteridine reductase inhibition by allowing for higher levels of uptake. Pteridine reductase is a validated drug target in *T. brucei* (Sienkiewicz *et al.*, 2010).

#### 4.3.4 Pathway analysis of RNA-seq data

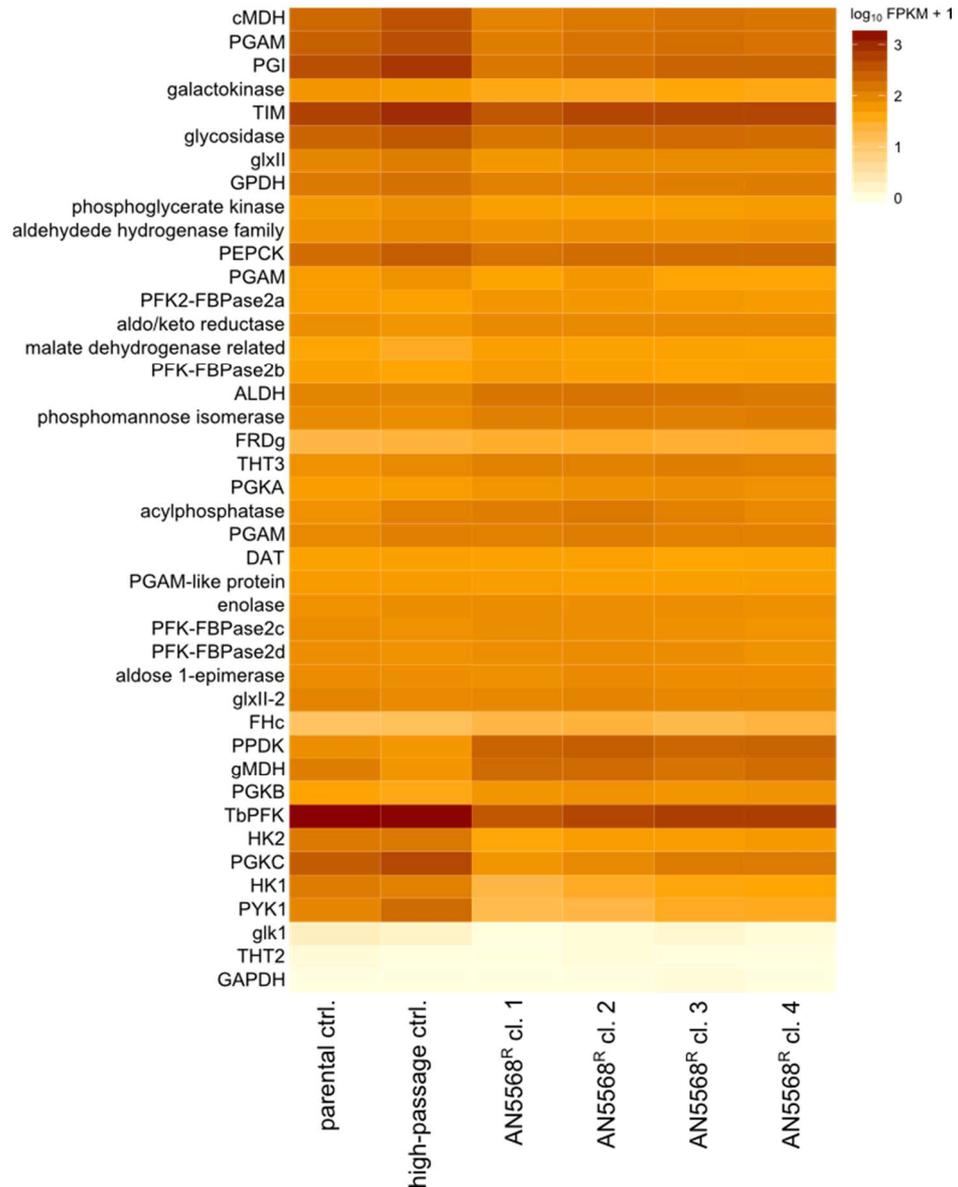
Using genome and transcriptome analyses published elsewhere, further analysis was done on the transcriptome dataset. In particular, specific pathways were analysed using the cummeRbund package (Trapnell *et al.*, 2012). A recent study from the Keith Matthews group provided many gene sets of various essential pathways in *T. brucei* (Capewell *et al.*, 2013b). Using these gene sets, we looked at individual pathways and areas of metabolism in the wild-type and resistant lines by plotting the Log<sub>10</sub>FPKM values for individual pathway components (fig. 4-7, 4-8 & S5) Firstly, glycolysis was analysed (fig. 4-7). Whilst there was no significant general trend in glycolysis between the sample groups, several genes showed decreased FPKM in the AN5568<sup>R</sup> lines (fig. 4-7). In particular, glucose 6-phosphate isomerase (PGI), phosphoglycerate kinase C (PGKC), pyruvate

kinase (PYR1) and the two hexokinases (HK1 & HK2) exhibited decreased  $\text{Log}_{10}\text{FPKM}$  values in the resistant clones (fig. 4-7). This is consistent with a switch to a procyclic-like transcriptome.

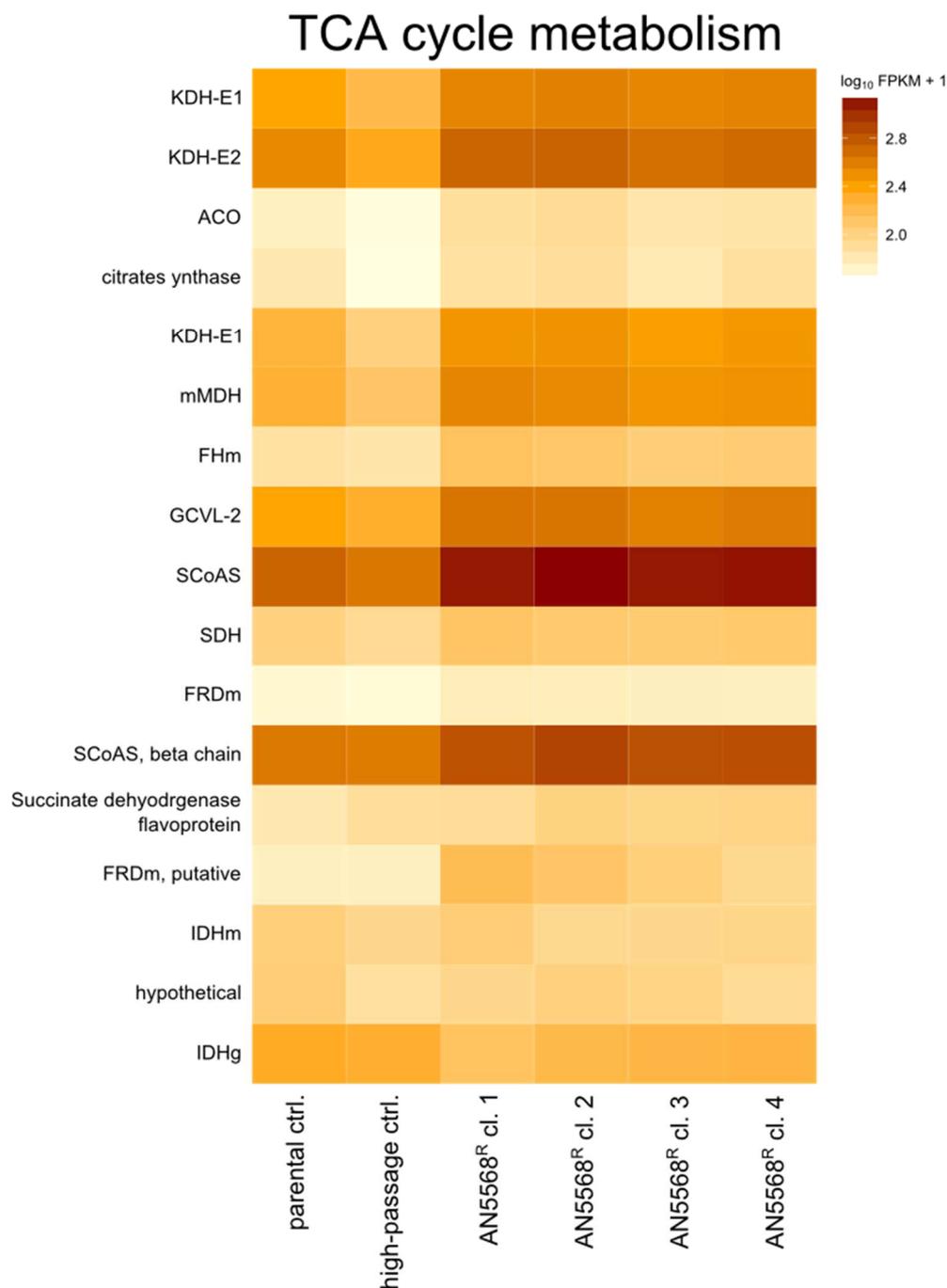
Central carbon metabolism in PCF trypanosomes contrasts widely with that in BSFs (Mazet *et al.*, 2013). In particular, whilst BSF parasites rely solely on glycolysis for ATP production, PCF cells import malate and pyruvate into the mitochondrion, and produce succinate, acetate and alanine. PCF metabolism involves an incomplete TCA (Krebs) cycle, and therefore, we analysed whether gene components of this pathway showed differential expression in the AN5568<sup>R</sup> line (fig. 4-8). Interestingly, there seemed to be a general trend of increased expression in resistant cells, as shown by higher  $\text{Log}_{10}\text{FPKM}$  values. Significantly increased transcript abundance was found in Tb927.10.2560 (mitochondrial malate dehydrogenase,  $\text{Log}_2$  fold change=1.04,  $P=5.0\text{e-}5$ ), Tb927.3.2230 (succinyl-CoA synthetase,  $\alpha$  sub-unit,  $\text{Log}_2$  fold change=1.45,  $P=5.0\text{e-}5$ ) and Tb927.5.940 (fumarate reductase,  $\text{Log}_2$  fold change=1.09,  $P=5.0\text{e-}5$ ). Whilst significance was reported for further genes in this cluster, the change in differential expression was below 2-fold and they were therefore omitted from further analyses.

These transcriptome changes were interesting, because they suggest there is increased activity in the mitochondrion in the benzoxaborole-resistant cell line. Combined with the transcriptome changes in surface protein expression, these data support the evidence that the AN5568<sup>R</sup> lines possess a PCF-like transcriptome. Whether these changes translated to proteomic or metabolic changes in the cell would be analysed later on using mass spectrometry ([section 4-6](#)).

## Glycolysis



**Figure 4-7: Heat map of abundances of transcripts involved in glycolytic metabolism.** Full names and gene IDs: **cMDH**: cytosolic malate dehydrogenase, Tb927.11.11250; **PGAM**: phosphoglycerate mutase protein, Tb927.10.7930; **PGI**: glucose-6-phosphate isomerase, Tb927.1.3830; **galactokinase**: Tb927.5.290; **TIM**: triosephosphate isomerase, Tb927.11.5520; **glycosidase**: Tb927.10.13630; **glxII**: glyoxalase II, Tb927.6.1080; **GPDH**: glyceraldehyde-3-phosphate dehydrogenase, Tb927.9.9820; **phosphoglycerate kinase**: Tb927.11.2380; **aldehyde hydrogenase**: Tb927.6.3050; **PEPCK**: phosphoenolpyruvate carboxykinase, Tb927.2.4210; **PGAM**: Tb927.5.3580; **PFK2-FBPase2a**: 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase, Tb927.8.1020; **aldo/keto reductase**: Tb927.11.5360; **malate dehydrogenase related**: Tb927.10.2550; **PFK-FBPase2b**: Tb927.3.2710; **ALDH**: aldehyde dehydrogenase, Tb927.6.4210; **phosphomannose isomerase**: Tb927.11.14780; **FRDg**: NADH-dependent fumarate reductase, Tb927.5.930; **THT3**: trypanosome hexose transporter 3, Tb927.4.2290; **PGKA**: phosphoglycerate kinase A, Tb927.1.720; **acylphosphatase**: Tb927.3.2030; **PGAM**: phosphoglycerate mutase, Tb927.11.10340; **DAT**: dihydroxyacetone phosphate acyltransferase, Tb927.4.3160; **PGAM-like**: phosphoglycerate mutase-like protein, Tb927.11.11490; **enolase**: Tb927.11.16410; **PFK-FBPase2c**: Tb927.7.1610; **PFK-FBPase2d**: Tb927.10.4520; **aldose 1-epimerase**: Tb927.4.1360; **glxII-2**: Tb927.4.1350; **Fhc**: fumarate hydratase, Tb927.3.4500; **PPDK**: pyruvate phosphate dikinase, Tb927.11.6280; **gMDH**: glycosomal malate dehydrogenase, Tb927.10.15410; **PGKB**: phosphoglycerate kinase B, Tb927.1.710; **PFK**: ATP-dependent phosphofructokinase, Tb927.3.3270; **HK2**: hexokinase 2, Tb927.10.2020; **PGKC**: phosphoglycerate kinase C, Tb927.1.700; **HK1**: hexokinase 1, Tb927.10.2010; **PYK1**: pyruvate kinase 1, Tb927.10.14140; **glk1**: glycerol kinase, Tb927.9.12610; **THT2**: trypanosome hexose transporter 2, Tb927.10.8440; **GAPDH**: glyceraldehyde 3-phosphate dehydrogenase, Tb927.10.6880.



**Figure 4-8: Heat map of abundances of transcripts involved in carboxylic acid metabolism.** Full names and gene IDs: **KDH-E1**: 2-oxoglutarate dehydrogenase E1 complement, Tb927.11.1450; **KDH-E2**: 2-oxoglutarate dehydrogenase E2 complement, Tb927.11.11680; **ACO**: aconitase, Tb927.10.14000; **citrate synthase**: Tb927.10.13430; **KDH-E1**: 2-oxoglutarate dehydrogenase E1 complement, Tb927.11.9980; **mMDH**: mitochondrial malate dehydrogenase, Tb927.10.2560; **FHm**: fumarate hydratase, Tb927.11.5050; **GCVL-2**: dihydrolipoyl dehydrogenase, Tb927.11.16730; **SCoAS beta chain**: succinyl-CoA ligase, Tb927.10.7410; succinate dehydrogenase flavoprotein: Tb927.8.6580; **FRDm**: NADH-dependent fumarate reductase, Tb927.10.3650; **IDHm**: isocitrate dehydrogenase, mitochondrial precursor, Tb927.8.3690; **hypothetical**: putative 2-oxoglutarate dehydrogenase, Tb927.11.9940; **IDHg**: isocitrate dehydrogenase, Tb927.11.900.

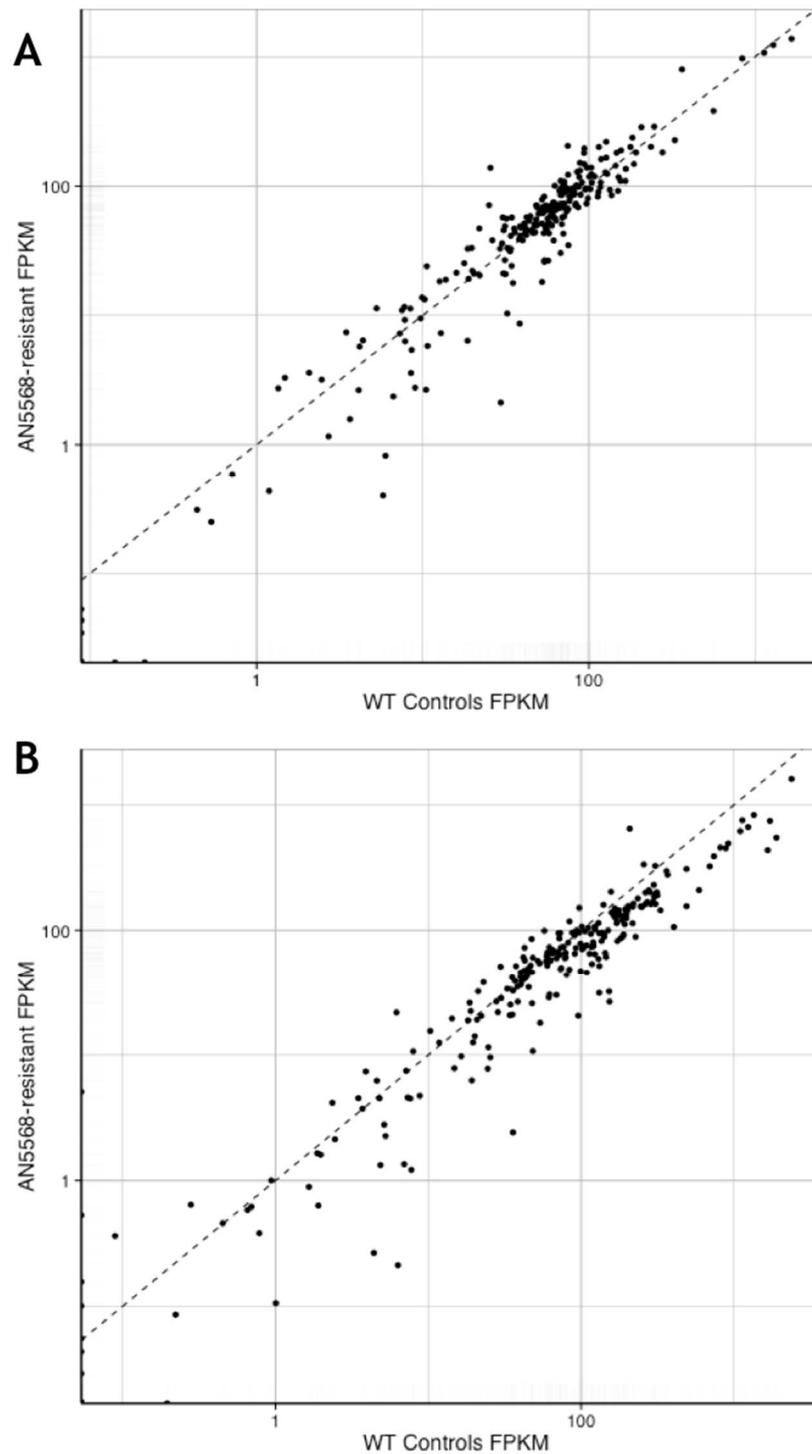
Further heat maps were generated to investigate transcript abundance changes in transcripts specific to mitochondrial metabolism (S5). This organelle is the site where the respiratory chain is located, generating another pool of ATP. In addition, the trypanosome alternative oxidase (AOX) is located in this compartment, which has previously been recognised as a potential drug target (Nihei *et al.*, 2002).

Within the genes that make up the respiratory chain, we could only identify two genes that had significant upregulation with a Log<sub>2</sub> fold change higher than 1. These were Tb927.3.1380 (ATP synthase  $\beta$  sub-unit, Log<sub>2</sub> fold change=1.18,  $P=5.0e-5$ ) and Tb927.3.1940 (TLP18.3, Log<sub>2</sub> fold change=1.09,  $P=5.0e-5$ ). All other genes did not show significant changes. There were also no major expression differences in mitochondrial carriers, which transport various small molecules across the outer and inner membrane. Whilst some genes showed a potential upregulation, these were not found to very high fold changes, and were thus disregarded.

Two components of acetate metabolism were found to be upregulated significantly. One of these, Tb927.3.2230, was already found as a component of the TCA cycle, whilst the other was Tb927.3.1970 (Log<sub>2</sub> fold change=1.13,  $P=5.0e-5$ ), a putative 4'-phosphopantothencycysteine synthase. There were decreasing trends in lipid metabolism as well as fatty acid metabolism, although none of these data sets contained changes of less than 0.5-fold in the resistant lines (S5).

One side effect of long-term *in vitro* culture is the differentiation of *T.b. brucei* cells into stumpy-form-like parasites that possess stumpy form characteristics, but still replicate (Breidbach *et al.*, 2002). Whilst stumpy-form cells do not normally possess the ability to proliferate, there have been reports of a "semi-differentiation" phenotype that involves formation of RNA granules containing PCF-specific mRNA (Kramer *et al.*, 2008). In the field, this is a natural occurrence, thought to be initiated by cascades related to quorum sensing as a result of increased cell density (Mony *et al.*, 2014). This process enables the parasite to pre-adapt to the environment of the tsetse fly midgut.

To assess whether the transcriptome observations were due to the effects of long-term culture, we analysed expression levels of genes previously shown to be either up or downregulated during BSF to stumpy differentiation. These gene sets were previously collated by the Matthews group in Edinburgh (Capewell *et al.*, 2013b). Two gene sets were made for upregulated genes and downregulated genes, containing 263 and 330 genes respectively, and the FPKM values for these gene sets from each sample group were plotted against each other to ascertain the correlation between the controls and AN5568-resistant cells (fig. 4-9).



**Figure 4-9: Comparative analyses of the AN5568R line and “stumpy” expression.** The data analysed in the RNA-seq experiment was compared to data from a recent publication that investigated differential expression in slender BSF and “stumpy” form trypanosomes (Capewell *et al.*, 2013b). Two groups of genes were made: One containing all the genes significantly upregulated in stumpy forms (A), and one containing all genes significantly downregulated in SSFs, compared to slender form BSF cells (B). The correlations in FPKM values were then analysed for these gene sets in the RNA-seq data. A) Genes upregulated in stumpy forms show no significant changes between wild-type cells and AN5568<sup>R</sup> cells (correlation coefficient = 0.9626). B) Similarly, there are no significant changes in genes normally downregulated in stumpy-forms, although the correlation was lower than that between stumpy upregulated genes (correlation coefficient = 0.9197). Overall, the changes in transcript abundance seen in the AN5568<sup>R</sup> cell line appear to be a direct cause of resistance generated against the benzoxaborole, rather than differentiation of cells into stumpy forms.

In general, there was very good correlation between genes in the wild-type lines and the AN5568R lines with the “upregulated” and “downregulated” gene sets showing correlation coefficients of

0.9626 and 0.9197 respectively. This correlations suggest the overall expression of the majority of genes perceived to exhibit differential expression during differentiation are not altered between the Lister 427 wild-type cells and the AN5568<sup>R</sup> cell line. Hence, there is a real possibility that the “semi-differentiated” phenotype these cells possess are a result of their resistance to the benzoxaborole compound. However, with the large amount of differential expression occurring, it is very challenging to pin-point a single mechanism.

#### 4.4 SNP and Indel analysis of the AN5568<sup>R</sup> line

To further investigate potential genetic alterations in the AN5568-resistant cell line, analysis of single nucleotide polymorphisms (SNPs) and insertions or deletions (indels) was undertaken. This was carried out on the same RNA-seq assembly that was utilised in transcriptome analysis, and this alignment data was further processed using the Genome Analysis Toolkit (McKenna *et al.*, 2010). Annotation of the data set was done using the SnpEff and SnpSift packages (Cingolani *et al.*, 2012). Results of the SNP and indel counts are summarised in table 4-4.

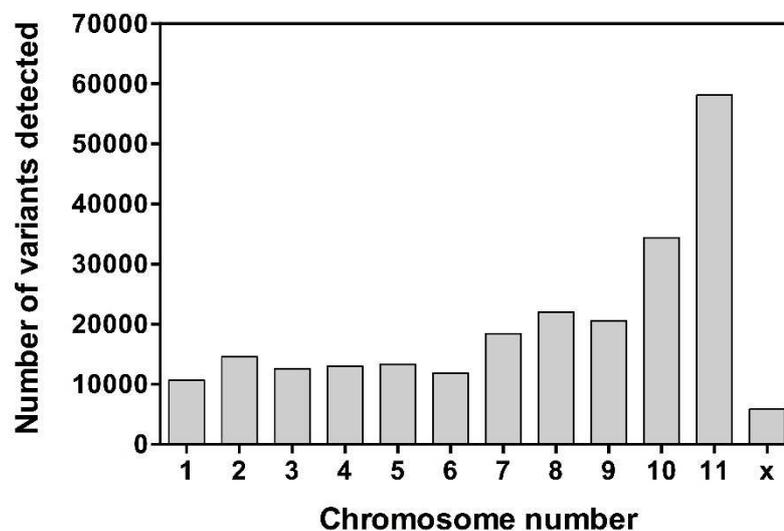
Type of SNP	Count	%
DOWNSTREAM	431806	38.31%
INTERGENIC	66072	5.86%
INTRAGENIC	2587	0.23%
INTRON	430	0.04%
NON_SYNONYMOUS_CODING	47068	4.18%
NON_SYNONYMOUS_START	4	0.00%
SPLICE_SITE_ACCEPTOR	1	0.00%
SPLICE_SITE_REGION	2	0.00%
START_GAINED	2157	0.19%
START_LOST	49	0.00%
STOP_GAINED	396	0.04%
STOP_LOST	230	0.02%
SYNONYMOUS_CODING	58364	5.18%
SYNONYMOUS_STOP	117	0.01%
UPSTREAM	453228	40.21%
UTR_3_PRIME	47703	4.23%
UTR_5_PRIME	16836	1.49%

**Table 4-4: SNP and indel counts for the complete RNA-seq data set.** Data was processed using GATK, following by SNP and indel annotation using the SnpEff package. Upstream and Downstream mutations accounted for the majority of polymorphisms by far, followed by mutations in intergenic and coding regions (both synonymous and non-synonymous).

One of the main challenges associated with SNP/indel analysis was the fact that the sequencing data had been aligned to the TREU 927 line, as mentioned previously. Whilst this did aid in more accurate alignments and fine read mapping, the result was many SNPs that were found in all samples analysed. The likely reasons for these are that they are not true SNPs that can account for

phenotypic difference. Instead, these are most likely a result of natural variation between the two strains of *T. brucei*. This was taken into account for the polymorphism analysis.

Further complications can arise from using RNA-seq to analyse polymorphisms, when compared to DNA-seq (Pareek *et al.*, 2011). Some of these problems have been well described (Piskol *et al.*, 2013). One of the challenges is the complexity of the transcriptome, which often involves processes such as splicing, leading to technical complications when calling SNPs (Piskol *et al.*, 2013). This challenge is somewhat eliminated in the study of *T. brucei*, because the only trans-splicing to occur is the addition of the spliced leader and poly-A tail to RNA transcripts, rather than intragenic splicing (Huang & van der Ploeg, 1991). However, another potential pitfall that does apply to the study of *T. brucei* is that the ability of analysis tools to detect variants such as SNPs and indels often depends on the expression levels of the gene in question (Piskol *et al.*, 2013, Quinn *et al.*, 2013). For example, if a mutation in a gene causes decreased levels of expression of that variant, it might not be detected with RNA-seq. In other words, there is a higher risk of both false positives and false negatives. However, given the discoveries that were made from transcript abundance analyses, the use of RNA-seq over DNA-seq was warranted. Nevertheless, we set out to analyse some of the variants in detail.



**Figure 4-10: Number of variants (SNPs and indels) detected on each chromosome.** At least 10,000 variants were detected on each chromosome using the SnpEff tool. This reflects the inherent differences between the Lister 427 and TREU 927 *T. brucei* strains. Many of the variants called were common to all cell lines that were sequenced, suggesting they played no role in benzoxaborole resistance. In addition, many variants were identified in only one sample, which suggests a high mutation rate in cell undergoing long-term *in vitro* culture. X = unassembled contigs.

Due to the high number of downstream and upstream mutations (fig. 4-10), many of which overlapped due to the nature of the poly-cistronic arrays that *T. brucei* possesses (i.e. downstream mutations for one gene were the same as upstream mutations for the next gene), these were disregarded in the analysis of SNPs. Instead, we focused our attention on non-synonymous coding mutations.

Using the SnpSift tool, we filtered both SNPs and indels very strictly, by ensuring both controls matched the reference genome, and the polymorphism resulted in a non-synonymous coding mutation. In addition we filtered out only SNPs and indels that appeared in all 4 AN5568<sup>R</sup> clones. The result of this filter still led to the calling of 277 SNPs, but no indels. Therefore, we further refined this SNP set by only filtering homozygous mutations leading to 90 SNPs in 82 genes.

The 82 genes found to contain homozygous SNPs were analysed with the TriTryp database. In total, 42 genes were annotated as hypothetical (including “hypothetical, unlikely”, and “hypothetical, conserved”). In addition, 2 genes were annotated as “unspecified product”. Whilst further analyses of these genes could be carried out in the form of domain and sequence analysis in the future, this was not done due to time constraints. Further genes identified were 7 BARP proteins, 6 VSG genes, two expression site-associated genes, a retrotransposon hot spot protein (Tb927.6.5160), a receptor-type adenylate cyclase (Tb927.6.780), and a protein of unknown function (Tb927.7.2330). These genes were not thought to be responsible for the resistance phenotype, and were therefore not studied further. In addition, several surface proteins such as VSGs, receptor adenylate cyclases and BARP proteins, are known to have high mutation rates (Saada *et al.*, 2014, Lopez *et al.*, 2015) and should be considered as false positives. This left 20 genes, which are shown in table 4-5.

We were interested to find several SNPs in proteins known to be important in regulation of protein expression. This included a SNP in a Zinc finger protein and 2 RNA binding proteins (RBPs). In addition, the dataset included a protein associated with differentiation (PAD2). This protein is known to be upregulated in stumpy and procyclic form parasites. Indeed, we found the protein to be upregulated in the AN5568<sup>R</sup> cell line as well ( $\text{Log}_2$  fold change=1.47,  $P=5.0\text{e-}5$ ). However, the SNP identified in PAD2 was outside any predicted domains, and thus we could not confirm its importance.

Interestingly, the zinc finger CCCH-domain containing protein, ZC3H37 (Tb927.10.12780), was the same protein identified as exhibiting significantly increased transcript abundance in the AN5568-resistant cell line. Whilst no domains of interest were found using the SMART and Pfam databases, this protein was recently shown to contain a CCCH domain by the Carrington group (Kramer *et al.*, 2010). However, by means of an Interpro search, we were able to conclude that the SNP once again did not localise to any predicted domains.

GeneID	Product Description	Reference	Mutation	Codon variant	Amino acid variant
Tb11.v5.0807	ribonuclease, putative	A	G	cAg/cGg	Q107R
Tb11.v5.1056	DNA polymerase kappa, putative	T	C	gTc/gCc	V13A
Tb927.1.880	Midasin, putative (MDN1)	C	A	Gcg/Tcg	A473S
Tb927.10.12780	Zinc finger CCCH domain-containing protein 37 (ZC3H37)	G	C	tCt/tGt	S136C
Tb927.10.13930	phosphatidic acid phosphatase, putative	G	A	Ccg/Tcg	P206S
Tb927.10.15290	tubulin binding cofactor c, putative	A	G	cAc/cGc	H64R
Tb927.10.9600	serine/threonine kinase, putative	G	A	Gcc/Acc	A538T
Tb927.11.540	ABC transporter, mitochondrial, putative, multidrug resistance protein, mitochondrial, putative (ABCT)	A	G	aTc/aCc	I5T
Tb927.11.980	Nuclear pore complex protein 158, serine peptidase, Clan SP, family S59, putative (Nup158)	A	G	Acg/Gcg	T227A
Tb927.2.4230	NUP-1 protein, putative	C	T	Ggt/AgT	G3292S
Tb927.3.710	RING-variant domain containing protein, putative	C	T	gGt/gAt	G424D
Tb927.4.4230	RNA-binding protein, putative (RBP31)	C	T	aGt/aAt	S360N
Tb927.5.1120	Phage tail fibre repeat, putative	C	G	gCt/gGt	A3059G
Tb927.5.4140	ZIP Zinc transporter, putative	A	G	aTc/aCc	I310T
Tb927.7.5940	Protein Associated with Differentiation (PAD2)	T	A	gaT/gaA	D312E
Tb927.8.6650	RNA-binding protein, putative (DRBD12)	C	T	Gta/Ata	V147I
Tb927.8.7600	amino acid transporter, putative	A	G	tAc/tGc	Y40C
Tb927.9.14000	60S ribosomal protein L12, putative	C	T	Gct/Act	A12T
Tb927.9.14660	SLACS reverse transcriptase, putative	G	A	Gac/Aac	D356N
Tb927.9.8250	atypical dual specificity phosphatase, putative	T	C	Aaa/Gaa	K23E

**Table 4-5: Homozygous SNPs unique to the AN5568<sup>R</sup> cell lines.** This dataset was generated using the SnpSift tool by filtering only SNPs where both controls were found to match the reference genotype. In addition, we searched for non-synonymous coding SNPs that were found in all 4 resistant clones.

Several other genes identified were followed up with domain analysis. For example, a SNP in RBP12 (Tb927.4.4230) was found to be located far downstream of the RNA recognition motif (RRM) domain, and therefore probably did not affect the protein's interactions with RNA. Another RBP, called DRBD12 (Tb927.8.6650), contained 2 RRM domains, but neither were predicted to be in the same position as the homozygous SNP identified in the protein.

Similarly, the SNP found in a serine/threonine kinase (Tb927.10.9600), was predicted to be located far downstream of the predicted protein kinase domain (pfam, position 6–311), and was also disregarded. Interestingly, the variant identified in a mitochondrial ABC transporter (Tb927.11.540) was predicted to lie in a 21-mer signal peptide. Whilst we could not confirm whether the protein was targeted to the mitochondrion, a previous study found it to be localised in the mitochondrial outer membrane (Niemann *et al.*, 2013).

Whether the variant discoveries are significant either in terms of the PCF-phenotype, or AN5568-resistance, would be important to analyse in the future. Due to time-constraints, we were not able to analyse these 20 genes in depth. However, domain analyses of the remaining genes to determine

where exactly all of these variants lie, would go some distance in determining whether any of the SNPs are significant. In addition, this study did not investigate heterozygous SNPs found in the AN5568<sup>R</sup> cell line. Using the same parameters with the SnpSift tool, a search for heterozygous instead of homozygous SNPs led to the discovery of 187 variants (100% SNPs, no indels). *T. brucei* is a diploid organism, and therefore, one gain of function or loss of function allele could have a significant impact on the cellular response to the benzoxaborole. Indeed, this mechanism has been shown to generate resistance in previous studies of other organisms (Anderson *et al.*, 2004).

There is, however, a possibility that the differentiation phenotype arose due to a loss-of-function of a repressor. There are at least two repressor of differentiation kinases, the loss of which leads to BSF differentiation (Jones *et al.*, 2014a). Whilst neither of these kinases contained SNPs in the benzoxaborole-resistant cell line, there are potentially other candidate genes that cause similar effects. Should a loss-of-function mutation occur leading to loss of expression, the variants would not be identified using RNA-seq, for reasons described above.

Whilst this method of searching for SNPs and indels certainly proved useful, we also used a targeted approach where subsets of genes were filtered from the datasets. This was done for three subsets of interest: a methyltransferase subset first presented in the previous chapter, a kinase subset collated by Jones and colleagues (Jones *et al.*, 2014a), and an RNA binding protein set assembled by Wurst and colleagues (Wurst *et al.*, 2009). With these searches we again filtered out only SNPs that were unique to the AN5568<sup>R</sup> lines and the results are shown in the supplementary data (S6). Currently, no leads have been identified from these 3 data-sets, but more in-depth analysis is ongoing. Further analyses could also be carried out on SNPs found in the UTRs, as mutations here can disrupt the binding of proteins such as transcription factors, thereby inhibiting, or stimulating gene expression.

## 4.5 The AN5568<sup>R</sup> line exhibits BSF morphology

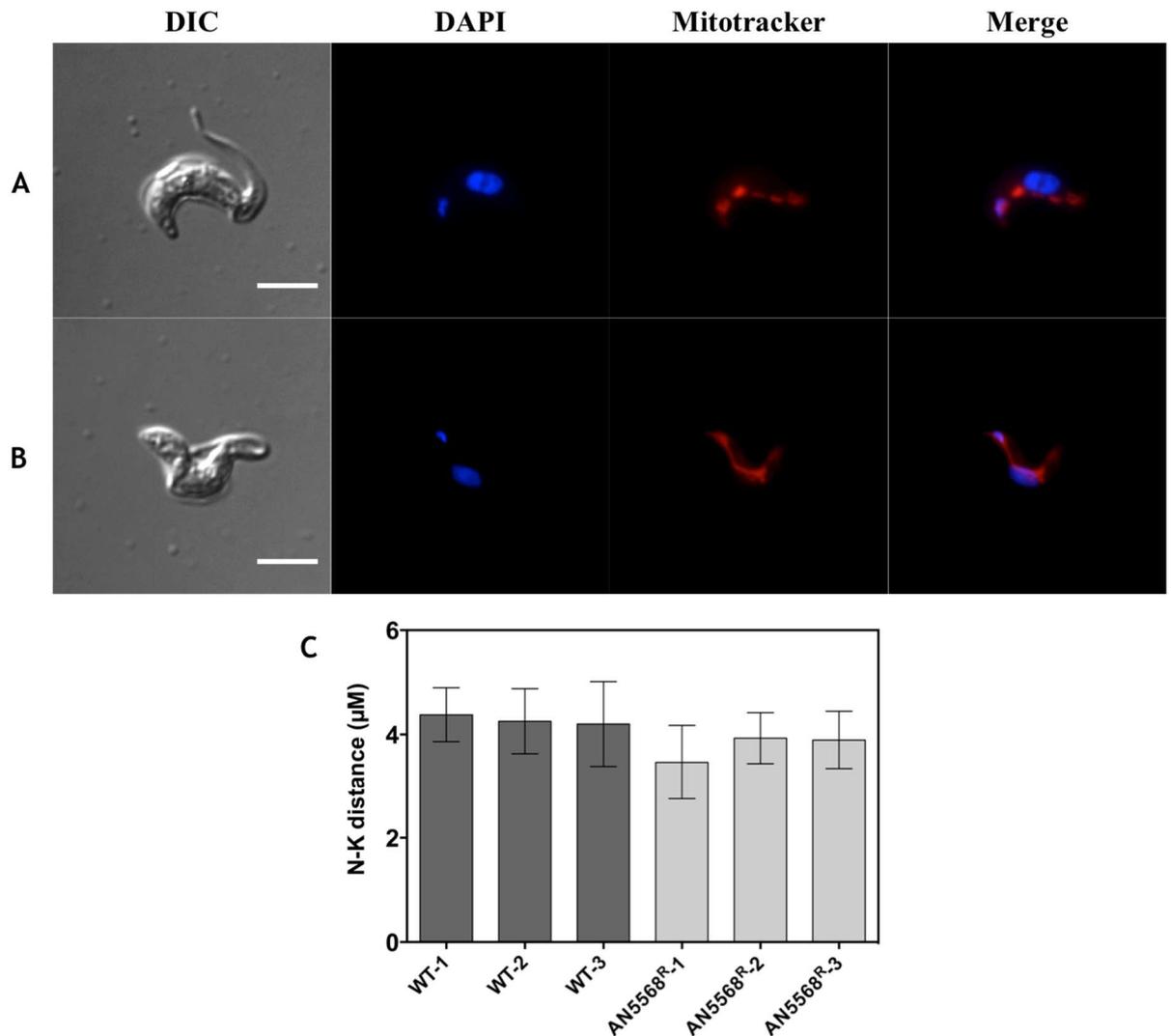
---

The transcriptomics results for the AN5568<sup>R</sup> cell line were quite surprising, especially given the fact that all *in vitro* culture to generate resistant cells was carried out in HMI-9, a culture medium designed for BSF parasites. Moreover, the parental Lister 427 cell line is perceived to be monomorphic, although recent work has shown that these cells can still exhibit differentiation phenotypes (K. Matthews, personal communication). To investigate whether the transcriptomics changes caused any differences to the cellular morphology, microscopy analysis was undertaken. Similar to the microscopy experiments shown in chapter 3, cells were stained with mitotracker and counter-stained with DAPI in order to observe the morphology of the nucleus, kinetoplast and mitochondrion.

Under direct light microscopy, the benzoxaborole resistant cell line appeared to exhibit normal wild-type BSF morphology (fig. 4-11). Cellular length did not appear to be altered, nor was the appearance of the flagella and cell body in general (fig. 4-11). Furthermore, the kinetoplasts and nuclei showed no adverse morphology. Cell cycle analysis was not carried out at this stage, but this should not be ruled out as an experiment in the future. However, for all intents and purposes, the drug-resistant parasites appeared to show no defects in cell cycle progression.

Interestingly, when visualising the mitochondria using Mitotracker Red, we found a mixed population of cells. Whilst many cells showed a normal mitochondrion, characterised as a long slender organelle that lies alongside the long-axis of the cell, other cells were found to contain a mitochondrion that seemed to be composed of several small organelles, rather akin to the “patchy” phenotype observed during AN5568 treatment of wild-type cells ([chapter 3](#)). Whilst this phenotype only occurred in approximately 50%, it could be a consequence of becoming resistant to the benzoxaborole. This phenomenon should be studied in more detail, in addition to quantification of wild-type morphology versus “patchy” morphology.

The RNA-seq analysis hinted at the possibility of the cells having undergone a differentiation event. Therefore, it was crucial to understand whether the morphology and localisation of intracellular organelles has also been altered. Procyclic trypanosomes are known to have different localisation of the kinetoplast and nucleus, compared to BSF parasites (Matthews, 2005). Whilst the mammalian form has a kinetoplast very close to the posterior limit of the cell body, it migrates during differentiation to a position equidistant between the posterior end and the nucleus (Matthews, 2005). To find out whether the AN5568R exhibited PCF-like, or BSF-like positioning of the kinetoplast and nuclei, the distance between these organelles was measured in both wild-type and resistant cells. The experiment was repeated three times and approximately 30-40 measurements were taken for each sample (fig. 4-12C).



**Figure 4-11: Microscopy analysis of the AN5568<sup>R</sup> cell line.** General morphology of the resistant cell line was analysed, and localisation of the nucleus, kinetoplast and mitochondria was studied. In general, cellular morphology was found to be normal, without significant deviation from that of BSF cells. A) In some cases, the mitochondrion showed a “patchy” phenotype. B) In contrast, many cells also exhibited wild-type mitochondrial morphology, with the entire mitochondrion running along the long axis of the cell. C) To investigate changes in nucleic and kinetoplast localisation, the distance between these two organelles was measured. This distance is known to decrease in PCF cells. However, in three separate experiments, no significant changes in this distance were found, even though the graphs are suggestive of a decreasing trend.

The average distance between nucleus and kinetoplast in wild-type cells was between 4 and 4.5 µm, as calculated using the Fiji software (Schindelin *et al.*, 2012). Similarly, the same distance in AN5568<sup>R</sup> cells showed a mean value of 3.5–4 µm, over three replicates. Whilst this did show a slight decrease in distance, this was calculated with a Student’s *t*-test to bear no significance. Given these results, we concluded that the AN5568<sup>R</sup> cell line exhibits mostly BSF-like morphology.

#### 4.6 Metabolomics analysis of the AN5568<sup>R</sup> line

To assess whether the transcriptome changes seen in the AN5568<sup>R</sup> line translated to proteomic or metabolic changes, we next carried out metabolomics analysis, maintaining the same methodologies as previously undertaken experiments. For this experiment, wild-type Lister 427

cells were grown in  $10\times$   $EC_{50}$  of the benzoxaborole for 6 hours, to ensure consistency with experimental data provided in the earlier chapter. Two AN5568<sup>R</sup> clones were incubated in the absence or presence of 4.8  $\mu$ M AN5568, equal to the concentration they were cultured in, and they were resistant to. Four biological replicates were prepared for each sample group, and metabolites were extracted at the 6-hour time-point as described in the methods (section 2.6.1). Samples were run on a Q-Exactive orbitrap mass spectrometer, and data was processed using the same mzMatch/Ideom pipeline used in all metabolomics analyses.

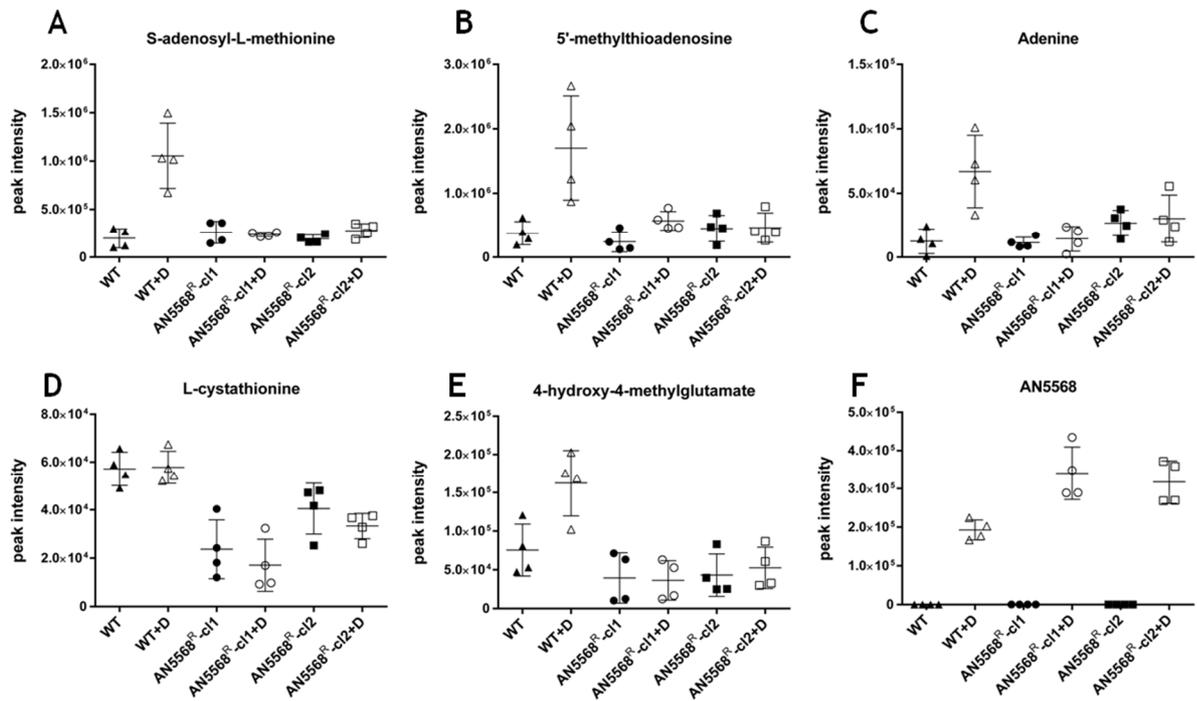
#### 4.6.1 The AdoMet/5'-MTA phenotype is abolished in AN5568<sup>R</sup> cells

The first metabolites analysed were those that showed major perturbations in earlier experiments using the benzoxaborole with wild-type cells (fig. 4-12). As shown previously, wild-type cells showed high levels of AdoMet, 5'-MTA, adenine, cystathionine and methylated lysines subsequent to AN5568-treatment. In addition there were elevated levels of 8-amino-7-oxononanoate, GDP-mannose and general decreases in lipid metabolism. However, when the resistant cells were analysed, these cells showed almost none of the aforementioned changes (fig. 4-12).

In particular, AdoMet ( $m/z$ : 398.1380, RT: 13.95 mins), 5'-MTA ( $m/z$ : 297.0896, RT: 7.73 mins) and adenine ( $m/z$ : 135.0546, RT: 10.12 mins), which thus far showed very reproducible increases in wild-type cells treated with AN5568, remained at normal levels in the AN5568<sup>R</sup> line when cells were treated with the benzoxaborole. In addition, high levels of adenine, characteristic of AN5568-treatment, had also been abolished, suggesting methionine metabolism in general had been completely downregulated in the benzoxaborole-resistant line, when in the presence of AN5568.

This finding was also intriguing in the context of AN5568-treated PCF parasites. As shown in chapter 3, these cells showed the same changes in methionine metabolism as BSF cells when treated with  $10\times$   $EC_{50}$  of the benzoxaborole. The AN5568<sup>R</sup> line exhibits a procyclic phenotype based on transcriptome analyses, yet they are more resistant to AN5568 treatment than wild-type PCF cells.

Methylation of hydroxyl-glutamate was shown to be carried out in an AdoMet-dependent fashion, using <sup>13</sup>C-(U)-methionine (chapter 3). However, as mentioned previously, the role of this metabolite is currently not understood in *T. brucei*. There is some indication it provides an alternate substrate to 4-hydroxyglutamate transaminase in the production of glutamate (Winter & Dekker, 1989). However, such an enzyme is not currently known to exist in *T. brucei*. The peak corresponding to 4-hydroxy-4-methylglutamate ( $m/z$ : 177.0634, RT: 12.02 mins), was again increased in intensity in AN5568-treated wild-type cells (fig. 4-12E). However, levels of this metabolite were decreased in AN5568<sup>R</sup> cells, even in the presence of the benzoxaborole. Indeed, the levels were even low compared to untreated wild-type cells. This suggests metabolism in this part of the metabolome has been significantly downregulated, potentially by a compensatory mechanism.



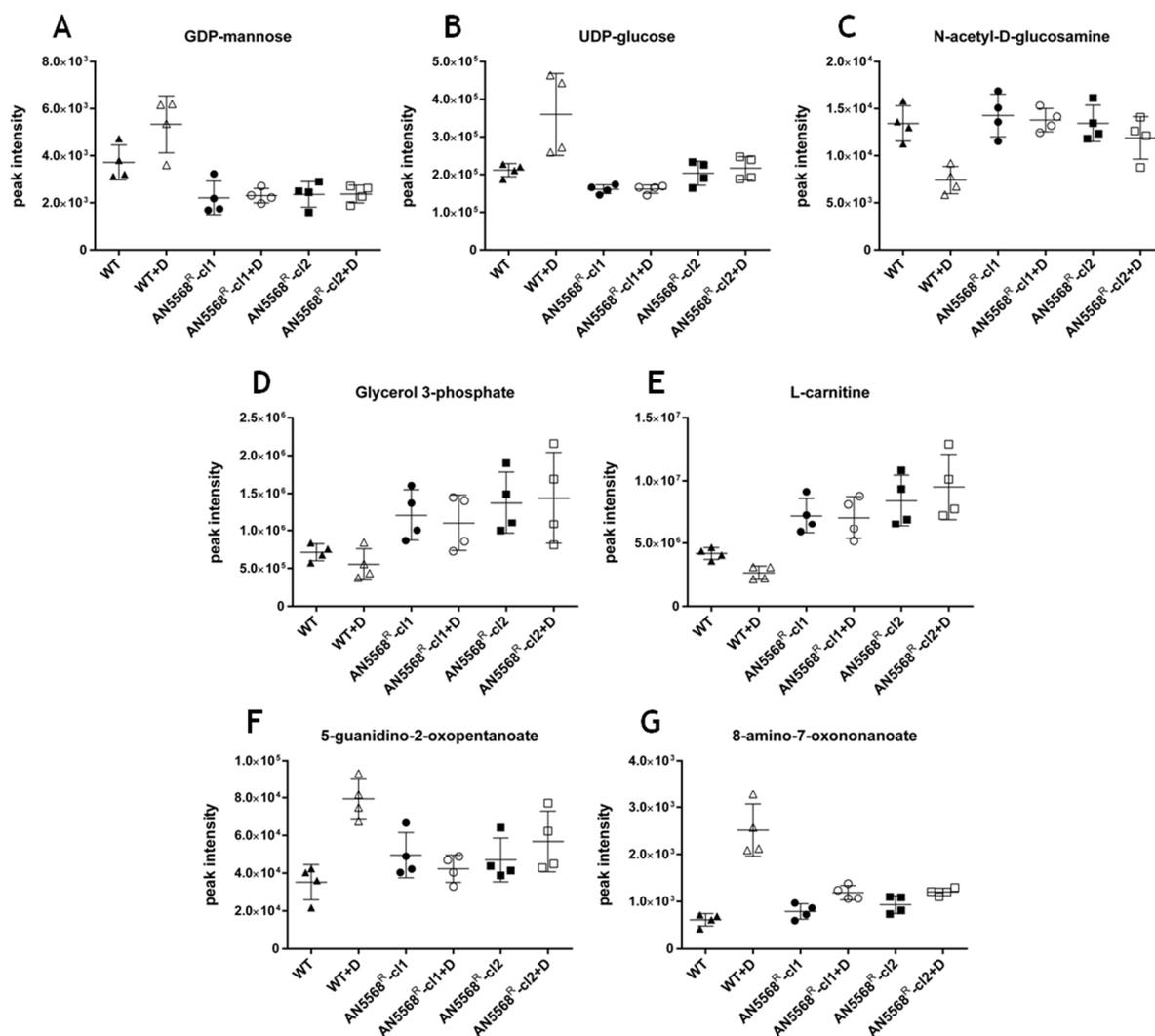
**Figure 4-12: Significant metabolic changes in methionine metabolism seen during AN5568 treatment are abolished in AN5568<sup>R</sup> cells.** The three most significant changes that usually occur in AN5568-treated cells were found in methionine metabolism. However, these changes were completely abolished in both benzoxaborole clones. There was no change in levels of AdoMet (A), 5'-MTA (B) or adenine (C), suggesting the resistant cell line had developed a mechanism to control these drug-induced effects. In addition, L-cystathionine peak intensities were less than half of that detected in WT and drug-treated WT cells (D). A similar phenotype was observed for 4-hydroxy-4-methylglutamate (E). Interestingly, AN5568 levels in treated cell pellets were still at high levels. Indeed, AN5568 levels were higher in the resistant cell line than WTs, possibly due to the fact a higher concentration of the drug was added to cultures during treatment incubations. This suggests the parasite was still taking up the benzoxaborole, even though it had managed to diminish its potency drastically. Key: WT: wild-type; WT+D: AN5568-treated wild-type; cl1: resistant line clone 1; cl1+D: AN5568-treated resistant line clone 1; cl2: resistant line clone 2; cl2+D: AN5568-treated resistant line clone 2;

Previous studies of drug resistance in trypanosomes showed that decreased transporter expression is often the cause, resulting in significantly reduced uptake of the active drug (Vincent *et al.*, 2010). Indeed, transcriptomics data did indicate differential expression of several transporters, as shown in previous sections. A search for the benzoxaborole amongst the mass spectrometry data showed that it was still present in the cell pellets at high levels, indicating that uptake was probably occurring (fig. 4-12F). Therefore, metabolic phenotypes observed indicated that the resistant line had become resistant against drug action, but not through the action of downregulated cell surface transporters. Importantly, this does not rule out loss of expression in transporters located on organelle membranes. Should the mitochondrion contain the AN5568 target, this could not be confirmed by metabolomics of whole cells. Instead, fractionation of cellular compartments followed by metabolomics analyses could allow more accurate information as to where the benzoxaborole localises. However, given the expression data for mitochondrial transporters, it is unlikely that any of the transporters investigated played a role in AN5568 uptake (S5).

#### 4.6.2 Lipid and carbohydrate metabolism in the AN5568<sup>R</sup> line

Previous experiments showed that levels of metabolites involved in GPI metabolism pathways increased in the presence of AN5568. However, we consistently found these changes to be nullified in the resistant cell line (fig. 4-13). GDP-mannose ( $m/z$ : 605.0754, RT: 14.49 mins) levels were lower than those seen in wild-type either treated or untreated, similar to UDP-glucose ( $m/z$ : 566.0542, RT: 13.56 mins) abundance. Interestingly, N-acetyl-D-glucosamine ( $m/z$ : 211.0898, RT: 11.56 mins), which was previously thought to increase in response to AN5568, was observed to decrease in this experiment, for unknown reasons. In addition, high levels of this metabolite were maintained in two resistant clones, both in the presence and absence of the benzoxaborole (fig. 4-13C).

Two metabolites related to lipid metabolism of interest, glycerol 3-phosphate ( $m/z$ : 172.0135, RT: 12.92) and L-carnitine ( $m/z$ : 161.105, RT: 12.49 mins), were again decreased in AN5568-treated cells (fig. 4-13D & E respectively). In contrast, these two metabolites were both present at higher levels in the resistant cell line, both in the presence and absence of AN5568. The reason for this is unknown, although one could hypothesize it is a requirement in a “PCF-like” cell which possesses higher mitochondrial activity compared to BSF cells. This would result in higher levels of fatty acid transport across the mitochondrial membranes.



**Figure 4-13: Further wild-type metabolic signatures that are altered in the AN5568<sup>R</sup> cell line.** Several metabolites involved in GPI metabolism were previously shown to be elevated in the presence of AN5568, including GDP-mannose (A) and UDP-glucose (B). Both these metabolites did not exhibit any changes in the AN5568<sup>R</sup> line, even in the presence of AN5568. For GDP-mannose in particular, this could indicate reduced GPI production in the PCF-like cells. Furthermore, N-acetyl-D-glucosamine (C), which was reduced in AN5568-treated WT cells, was observed to be at WT levels in the resistant cell line. Interestingly, glycerol 3-phosphate (D) and L-carnitine (E), the latter of which exhibits increased abundance in differentiating cells (J. Anderson, unpublished), were both increased more than 2-fold in the resistant cell line. Finally, 5-guanidino-2-oxopentanoate (F) and 8-amino-7-oxononanoate (G), both shown to increase for unexplained reasons in WT AN5568-treated cells, were reduced close to WT levels in benzoxaborole-resistant cells, although the former was slightly increased, even in untreated AN5568<sup>R</sup> cells. Key: WT: wild-type; WT+D: AN5568-treated wild-type; cl1: resistant line clone 1; cl1+D: AN5568-treated resistant line clone 1; cl2: resistant line clone 2; cl2+D: AN5568-treated resistant line clone 2

Two further metabolites of interest were identified and shown to have abolished the wild-type drug-treated metabolic phenotype. 5-guanidino-2-oxopentanoate ( $m/z$ : 173.0800, RT: 13.26 mins) is also known as 2-ketoarginine. Interestingly, this metabolite has been shown to be involved in an AdoMet-dependent MTase reaction (Braun *et al.*, 2010). This reaction results in the production of 5-guanidino-3-methyl-2-oxopentanoate and is part of the 3-methylarginine biosynthesis pathway. The MTase study presented in the previous chapter did not lead to the accurate prediction of a potential methyltransferase for this reaction, although its presence should not be ruled out. In addition, in all of the metabolomics experiments carried out with AN5568 so far, no significant

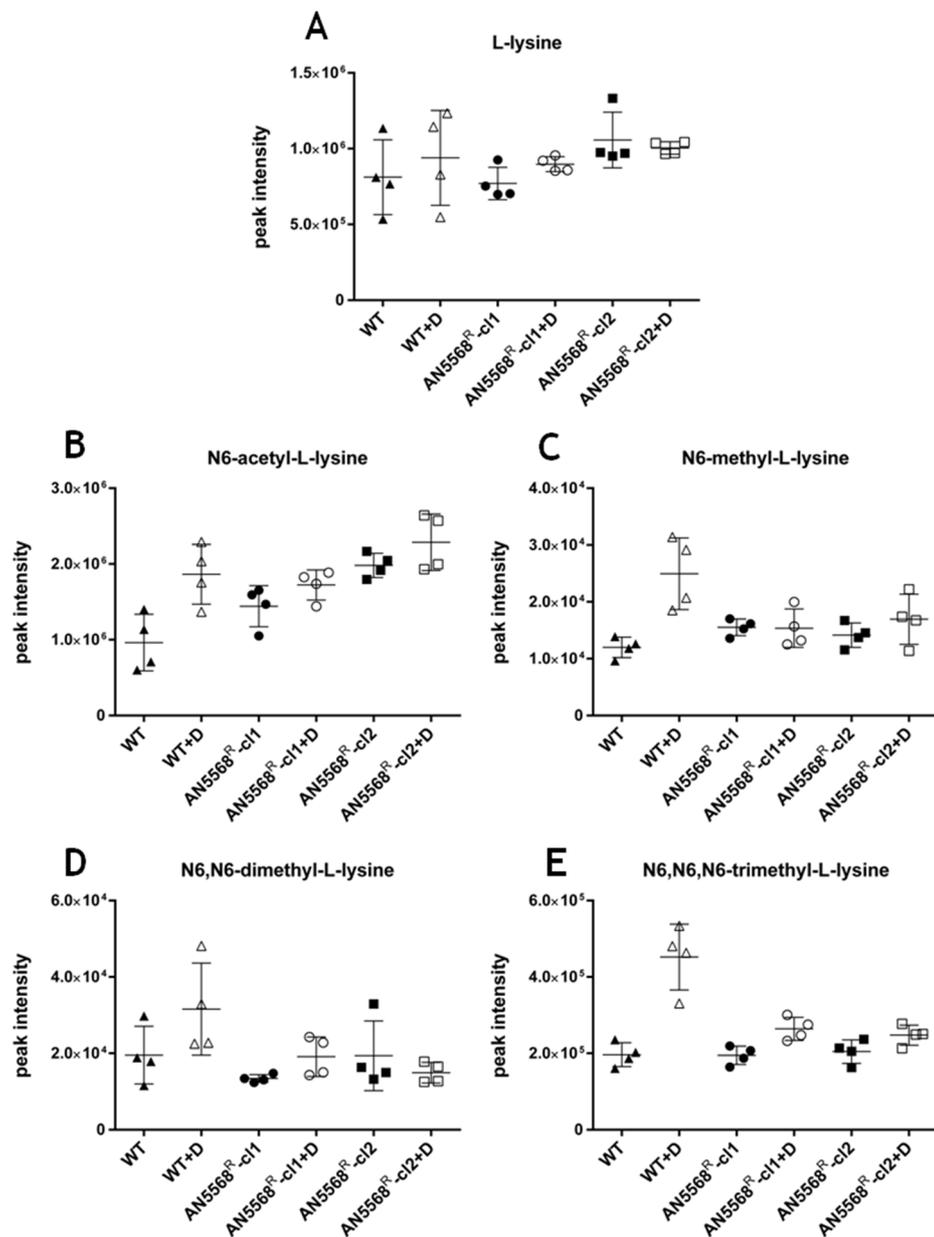
perturbations in L-arginine, N $\gamma$ -monomethyl-L-arginine or N $\gamma$ ,N $\gamma$ -dimethyl-L-arginine have been reported. There was a small increase seen in L-arginine phosphate (data not shown), but this was not thought to be significant, given the other changes observed. We did find a peak for homoarginine, which possesses the same monoisotopic mass as 3-methylarginine, but again, this metabolite showed no significant changes with either cell line in the presence or absence of AN5568.

Finally, 8-amino-7-oxononanoate ( $m/z$ : 187.1208, RT: 111.77 mins) was again found to increase significantly in drug-treated cells. However, this phenotype was again ablated in resistant cells, both in the presence or absence of AN5568 (fig. 4-13G), although there did appear to be slight increases in the presence of AN5568, even in the resistant cells. However, these increases were still vastly reduced compared to the wildtype peaks. As discussed previously, 8-amino-7-oxononanoate has been shown in other organisms to be involved in an AdoMet-dependent transaminase reaction, resulting in the production of 7,8-diaminononanoate and S-adenosyl-4-methylthio-2-oxobutanoate. This reaction is part of the biotin synthesis pathway. Unfortunately, we were not able to detect either metabolite, leaving the interpretation of this result open to discussion.

### 4.6.3 Lysine metabolism in the absence and presence of AN5568

Changes in lysine metabolism were previously discussed as they were highly reproducible and potentially crucial to the mode of action of AN5568 in the context of a stress response. Given the transcriptomics changes seen in the benzoxaborole-resistant cell line, changes in lysine metabolism and methyl lysine production in particular, could be due to changes in histone methylation, as well as protein activation/inhibition.

Concurrent with previous analyses, intracellular L-lysine ( $m/z$ : 146.1056, RT: 19.79 mins) levels did not undergo significant changes, either in wild-type cells or benzoxaborole-resistant cells (fig. 4-14A). An increase in acetyl-L-lysine ( $m/z$ : 188.1157, RT: 12.09 mins) was found in the first metabolomics experiment presented in chapter 3. interestingly, this metabolite's abundance was increased in the resistant clones in both the absence and presence of AN5568 (fig. 4-14B). In contrast, the significant increases first seen in the methylation of lysine were completely abolished (fig. 4-14C, D & E). All three methylated lysine metabolites, N6-methyl-L-lysine ( $m/z$ : 160.1212, RT: 19.16 mins), N6,N6-dimethyl-L-lysine ( $m/z$ : 174.1368, RT: 17.85 mins) and N6,N6,N6-trimethyl-L-lysine ( $m/z$ : 188.1527, RT: 18.29 mins), were identified and no increase was observed.



**Figure 4-14: Modified lysine metabolism in benzoxaborole-resistant cells.** As shown in chapter 3, changes in lysine metabolism were found in AN5568-treated cells. Whilst there were no changes in L-lysine (A) itself, there were again interesting changes in lysine metabolism in the AN5568<sup>R</sup> cell line. Acetyl-L-lysine (B) was found at varying levels in resistant cells ranging from a 1.5-fold increase compared to WT untreated cells, to 2.3-fold increased. Interestingly, this metabolite was detected at higher levels in differentiating cells (J. Anderson, unpublished). In contrast, most methylated lysines were observed at WT levels in the AN5568<sup>R</sup> lines. As shown previously, N6-methyl-L-lysine (C), N6,N6-dimethyl-L-lysine (D) and N6,N6,N6-trimethyl-L-lysine (E), were increased after AN5568 treatment under WT conditions and whilst the latter still showed a slight increase in drug-treated resistant cells, these metabolites were all unchanged, even after AN5568-treatment. Key: WT: wild-type; WT+D: AN5568-treated wild-type; c11: resistant line clone 1; c11+D: AN5568-treated resistant line clone 1; c12: resistant line clone 2; c12+D: AN5568-treated resistant line clone 2

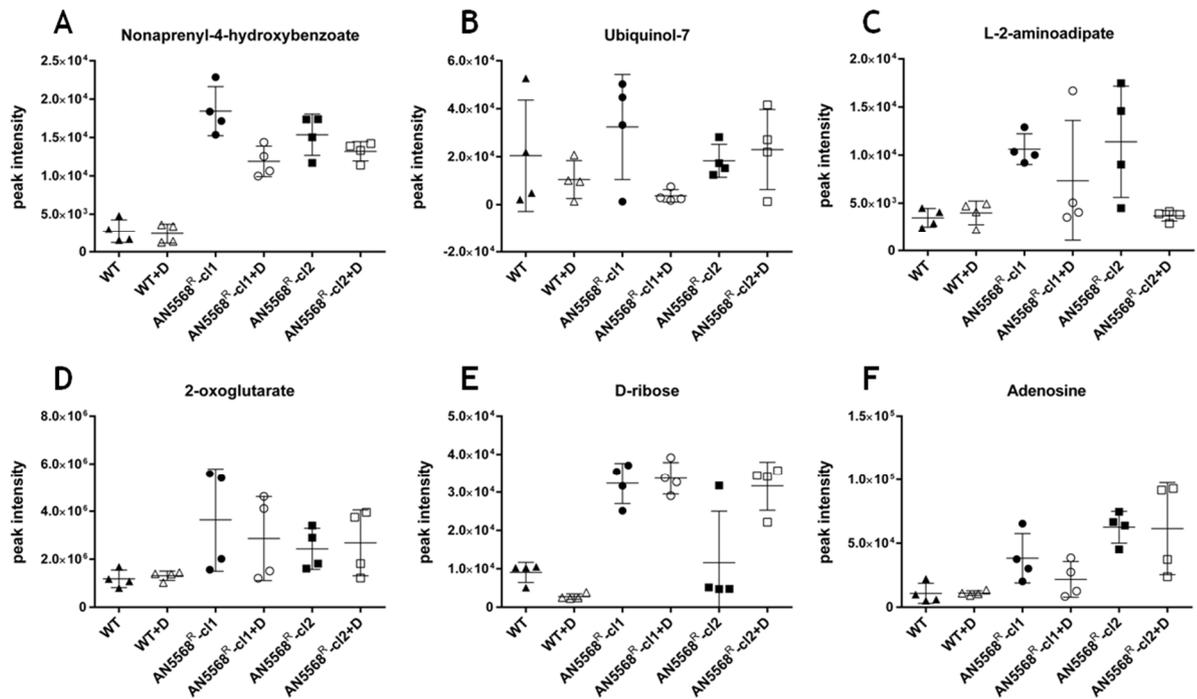
#### 4.6.4 Differences in metabolism between untreated wild-type and untreated AN5568<sup>R</sup> cells

Having discovered many metabolites to have reverted to wild-type levels in the benzoxaborole-resistant cell line, even in the presence of AN5568, we next turned our attention to changes in metabolism that were unique to this cell line compared to wild-type parasites.

The most significant increase found in the AN5568<sup>R</sup> line was putatively identified as nonaprenyl-4-hydroxybenzoate ( $m/z$ : 750.5959, RT: 3.65 mins) (fig. 4-15A), a molecule involved in the biosynthesis of ubiquinol, which is an important intermediate in the respiratory chain of the inner mitochondrial membrane. There was one isomer reported for this metabolite, 2-nonaprenyl-6-methoxy-1,4-benzoquinone. Interestingly, this molecule is found upstream, in the same ubiquinol biosynthesis pathway. Attempts were made to locate a peak corresponding to ubiquinol itself, and this led to the putative detection of ubiquinol-7 ( $m/z$ : 660.5116, RT: 4.64 mins) (fig. 4-15B). This peak was highly variable between individual replicates, but the changes this metabolite underwent were interesting because of the localisation of this pathway, in the electron transport chain of the mitochondrion.

Another interesting change that was not seen in wild-type cells was increased levels of L-2-aminoadipate ( $m/z$ : 161.0688, RT: 12.84 mins) (fig. 4-15C). Whilst the presence of the benzoxaborole caused a significant decrease of this metabolite back to wild-type levels, there were high levels in the absence of drug. This metabolite has been implicated in both the synthesis and degradation of lysine (Danhauser *et al.*, 2012). In addition, it has been identified as a marker of oxidative stress (Zeitoun-Ghandour *et al.*, 2011). However, its role in metabolism of *T. brucei* has not been discussed. Given the changes occurring in lysine metabolism, the increased levels of L-2-aminoadipate could be very important in the abolishment of methylated lysine increases.

Finally, higher levels of ribose ( $m/z$ : 150.0530, RT: 1.48 mins) (fig. 4-15E) in AN5568<sup>R</sup> cells could suggest increase pentose-phosphate metabolism. Indeed, the upregulation of transketolase found in transcriptome data supports this theory. In addition, resistant cells showed higher levels of adenosine ( $m/z$ : 267.0967, RT: 9.73 mins) (fig. 4-15F). This result was interesting because the transcriptomics data indicated that adenosine transporter mRNA transcripts were present at lower abundance in the resistant cell line



**Figure 4-15: Novel metabolic changes in the benzoxaborole-resistant cell line.** A putatively identified nonaprenyl-4-hydroxybenzoate (A) was found in high abundance in AN5568<sup>R</sup> cells. In the same ubiquinol biosynthesis pathway, ubiquinol-7 (B) was putatively identified, and shown to be present in high levels in one of the resistant clones. Varying levels of L-2-aminoadipate (C) were seen in the resistant line. Whilst drug-treatment caused a reduction in this metabolite, untreated cells exhibited a higher level of this metabolite compared to WT cells. Key: WT: wild-type; WT+D: AN5568-treated wild-type; c1: resistant line clone 1; c1+D: AN5568-treated resistant line clone 1; c2: resistant line clone 2; c1+D: AN5568-treated resistant line clone 2

#### 4.6.5 Lipid metabolism is almost completely restored in AN5568<sup>R</sup> cells

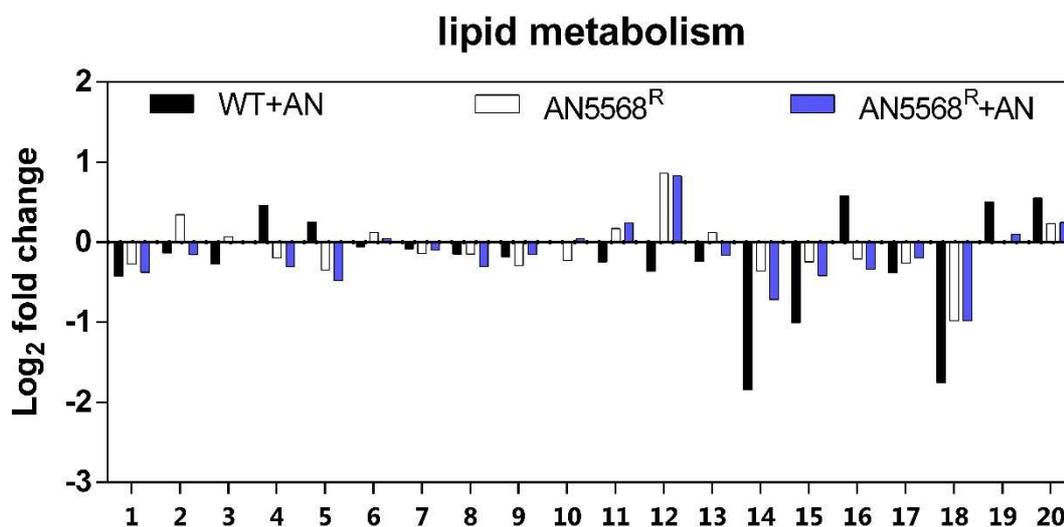
Treatment of wild-type *T. brucei* with AN5568 led to reduced abundance of several essential lipids (chapter 3). We therefore attempted to assess whether these lipids were restored to wild-type levels in the AN5568<sup>R</sup> line, and whether lipid metabolism in general was perturbed (fig. 4-16).

As mentioned previously, glycerol 3-phosphate, levels of which were found to decrease in AN5568-treated WT *T. brucei* cells, was observed at far higher abundance in AN5568<sup>R</sup> cells both treated and untreated (fig. 4-16).

Ethanolamine derivatives were generally decreased in AN5568-treated WT cells, and whilst the case was similar in AN5568<sup>R</sup> cells, the reduction was not as dramatic (fig. 4-16). Ethanolamine is used to generate phosphatidylethanolamine (PE) through CDP-ethanolamine branch of the so-called Kennedy pathway (Gibellini *et al.*, 2009). This is an essential pathway in both PCF and BSF *T. brucei*, and in general, PE is present at lower abundance in PCF cells (Richmond *et al.*, 2010). Therefore, the metabolomics findings presented here support the transcriptomics data described previously.

The other branch of the Kennedy pathway uses CDP-choline to generate phosphatidylcholine (PC) (Farine *et al.*, 2015). In addition to CDP-choline, several choline derivatives were also observed in the metabolomics dataset (fig. 4-16). For unknown reasons, CDP-choline was increased in WT

AN5568-treated cells. This contrasted with previous findings (chapter 3) that showed a reduction in CDP-choline after treatment. This could be due to a technical fault, or sensitivity of the MS, or incorrect annotation (both  $m/z$  ratio and RT were found to be equal in the datasets). CDP-glycero-3-phosphocholine was, like CDP-glycero-3-ethanolamine, decreased in AN5568-treated WT cells, and again the extent of these decreases was reduced in AN5568<sup>R</sup> cells, suggesting that the resistance mechanism had reduced the impact of AN5568 on the *T. brucei* lipidome (fig. 4-16)



**Figure 4-16: Lipid metabolism in AN5568-resistant cells.** Several putatively identified metabolites involved in lipid metabolism were altered in abundance in the different sample groups. A metabolite identified as taurine (2), a cysteine derivative, was found to increase in resistant cells in the absence of AN5568. Several fatty acids (docosanoic acid & tetracosanoic acid, 4 & 5 respectively) were increased in WT AN5568-treated cells, and in contrast, these were reduced in resistant cells, in the presence and absence of drug. As mentioned previously, glycerol 3-phosphate (12) was highly abundant in the resistant line, compared to WT cells. Furthermore, the extent of reduction in PC- and PE-derived metabolites was diminished in AN5568R cells, as exemplified by phosphodimethylethanolamine (14), sn-glycero-3-phosphocholine (15) and sn-glycero-3-phosphoethanolamine. Metabolites corresponding to each number are shown in table 4-6, below.

Lipid metabolism			
1	CMP-N-trimethyl-2-aminoethylphosphonate	11	Hexadecanal
2	Taurine	12	sn-Glycerol 3-phosphate
3	Glycodeoxycholate	13	Choline phosphate
4	Docosanoic acid	14	Phosphodimethylethanolamine
5	Tetracosanoic acid	15	sn-glycero-3-phosphocholine
6	Dodecanoic acid	16	CDP-choline
7	Tetradecanoic acid	17	Diethanolamine
8	Octadecanoic acid	18	sn-glycero-3-phosphoethanolamine
9	(9Z)-Hexadecenoic acid	19	Acetylcholine
10	Decanoic acid	20	Triethanolamine

**Table 4-6: Corresponding putative metabolite identifications for figure 4-16.**

For the most part, lipid metabolism was not significantly perturbed either in treated or un-treated benzoxaborole-resistant cells, as shown by the abundances of carboxylic acids such as dodecanoic acid and tetradecanoic acid (fig. 4-16). However, significant changes in the PE and PC branches of

lipid metabolism seen in AN5568-treated cells had been reduced, or as was shown by Richmond and colleagues, the lipid composition might have changed somewhat due to the procyclic phenotypes of the resistant line.

#### 4.6.6 Comparison of metabolomics and transcriptomics datasets

The ability to generate large data sets unravelling a complete biological system at the genome, transcriptome, proteome and metabolome levels can provide huge amounts of data. However, the integration of these datasets to give a more complete picture of a biological system is crucial. Systems approaches allow the mapping out of entire pathways of interest. Thus, the metabolomics data was analysed in the context of the changing transcriptome of benzoxaborole-resistant cells to find significant changes in *T. brucei* upon AN5568-resistance.

One area of metabolism that showed upregulation on a transcriptome levels was mitochondrial metabolism in the form of the TCA cycle. As mentioned previously, mitochondrial malate dehydrogenase was shown to be upregulated in the AN5568R. Indeed, the product of this enzyme, oxoglutarate, was found to have a higher abundance (fig. 4-15D).

Another interesting metabolic finding was increased levels of glycerol 3-phosphate (fig. 4-13 & 4-16). Interestingly, when looking back through the transcriptome dataset we found significantly decreased levels of glycerol 3-phosphate dehydrogenase (Tb11.v5.0516, S4). This suggests that glycerol 3-phosphate is not utilised at the same rate in the resistant cell line compared to WT cells. Furthermore, glycerol 3-phosphate dehydrogenase is a glycosomal enzyme (Bringaud *et al.*, 2006), and the aforementioned data support the theory that the resistant cell line possesses a PCF-like transcriptome as well as metabolome.

In the some cases, transcriptome did not match metabolome. For example, pyruvate, phosphate dikinase (PPDK) was shown to have increased transcript abundance in AN5568<sup>R</sup> cells (S4). However, metabolomics data indicated that there was not much change in the levels of substrate (pyruvate), nor product (PEP) (S7).

Unfortunately, time constraints meant that data integration of metabolomics and transcriptomics was not completed. There is currently no efficient pipeline for the integration of omics datasets in *T. brucei*, although work to this extent is ongoing at Glasgow Polyomics.

## 4.7 Discussion

---

We here report the generation of a 427-Lister derived *T. brucei* cell line that shows high levels of AN5568-resistance, with an EC<sub>50</sub> value almost 20-fold higher than WT *T. brucei*. This cell line also exhibited moderate resistance against the MTase inhibitor sinefungin, thereby supporting the theory that these two compounds act on similar pathways, or are taken up by similar transporters. The resistant line is hypersensitive to diminazene and pentamidine, but the known transporters for these compounds, which include TbHAPT1, TbLAPT and the P2 aminopurine transporter, were not differentially expressed. Several studies have suggested that pentamidine, like other diamidines, targets the mitochondrion (Ludewig *et al.*, 1994). There is a possibility that the mitochondrion is more active in the resistant cell line, based on the RNA-seq data discussed below. In theory, this could render the parasite hypersensitive to mitochondrion-targeting compounds.

Interestingly, this study shows that the AN5568<sup>R</sup> line, whilst exhibiting BSF-morphology, possesses a transcriptome that more closely resembles a differentiating, or SSF trypanosome. BSF-specific transcripts such as VSG expression sites and proteins involved in glycolysis were present at a lower abundance and in contrast, PCF-specific transcripts such as those encoding mitochondrial metabolism and EP procyclin proteins, were found at higher levels. Transcripts corresponding to three PAD receptors, known to be instrumental in *T. brucei* differentiation (MacGregor & Matthews, 2012), were significantly more abundant in the AN5568<sup>R</sup> line. Furthermore, there was a significant decrease in THT-1 glucose transporter transcription. This transporter is specific to BSF-stage trypanosomes (Bringaud & Baltz, 1993), and its downregulation supports the aforementioned hypothesis. Pathway-specific analyses hinted at a general downward trend in glycolytic component transcripts and increased transcription of TCA cycle component transcripts (fig. 4-7). In addition, acetate metabolism, which was until recently thought to be carried out in PCF-stage cells only, was also upregulated (fig. 4-8). Taken together, the above suggests that the *T. brucei* AN5568-resistant cell line had undergone a differentiation event, which coincided with increased resistance to the benzoxaborole.

The AN5568<sup>R</sup> RNA-seq data was also analysed for any SNPs of interest. Here, we found an unexpectedly high number of SNPs across the transcriptome. However, in this study, the RNA-seq data was aligned to the TREU 927 reference genome, as it was judged to have superior annotation. It would be beneficial to carry out this analysis by alignment to the updated (v9.0) genome of the Lister 427 cell line, from which this resistant cell line was derived. Nonetheless, homozygous SNPs were filtered out from the dataset, and whilst no indels of note were discovered, we found 20 SNPs that were present in all 4 clones of the AN5568<sup>R</sup> line. Interestingly, several SNPs were found in essential regulatory proteins such as the PAD2 receptor and a zinc-finger protein. However, domain analysis predicted these SNPs did not lie in regions of importance. Further analyses of heterozygous

SNPs are ongoing and will investigate in particular whether RNA-binding proteins or other master regulators of differentiation, such as zinc finger proteins, are mutated, which could lead to resistance.

The AN5568-resistant cell line was also subject to metabolomics analyses both in the presence and absence of the benzoxaborole. This dataset showed the resistant cell line to be almost unaffected by the presence of the benzoxaborole. In particular, there was a noticeable ablation of the AdoMet, 5'-MTA and adenine increases seen in WT cells. In addition, L-cystathionine levels were decreased, even when compared to WT cells in the absence of AN5568. Similarly, 5-guanidino-2-oxopentanoate and 8-amino-7-oxononanoate levels were comparable to WT cells, as were those of methylated lysine residues. These data suggest that the cell line has managed to control the drug effects, even though uptake is still occurring.

Interestingly, there were several metabolic changes that have also been observed in differentiating cells. The Barrett group has recently carried out LC-MS analyses of differentiating *T. brucei* (M. Barrett & J. Anderson, unpublished). In this experiment, the most significant metabolites identified were increases in acetyl-L-lysine and L-carnitine. Whilst L-carnitine levels were reduced in AN5568-treated WT cells, levels of this metabolite were increased in both treated and untreated AN5568<sup>R</sup> cells (fig. 4-13). Furthermore, acetyl-L-lysine was also increased in abundance, although only when compared to untreated WT cells (fig. 4-14). However, an increase in 5'-MTA, one of the main metabolic phenotypes of AN5568 treatment, was not observed in the resistant cell line. Whilst the cell line is able to compensate and abolish the severe upregulation of methionine metabolism, the metabolome, just like the transcriptome, carries the signature of a differentiating trypanosome.

One important drawback of the metabolomics experiment was the lack of a control sample group incubated with 10× EC<sub>50</sub> concentration of AN5568 on the resistant line. Instead, the concentration of drug the cells were cultured in (4.8 μM) was used. The EC<sub>50</sub> of the drug in resistant cells is close to 5 μM. Therefore, further work for this analysis might require a novel time course with addition of drug at a concentration closer to 10× EC<sub>50</sub> (50 μM), to allow for consistency, and to assess whether cells incubated with a high dose of drugs undergo metabolic changes that more closely resemble wild-type cells treated at 10× EC<sub>50</sub>. One problem that might be associated with this experiment is the high concentration of AN5568 could potentially lead to non-specific drug-induced effects that can drastically alter the resulting metabolic profiles of the parasites. Interestingly, previous work by Jacobs and colleagues found that brain exposure to AN5568 is high, with a C<sub>max</sub> of >10 μg/mL (equivalent to ~27 μM) and AUC<sub>0-24</sub> hr equal to > 100 μg\*hr/mL after treatment with 25 mg/kg of the benzoxaborole (Jacobs *et al.*, 2011). The resistant cell line generated here is unlikely to survive in concentrations of that magnitude.

Importantly, the metabolomics study presented here did not investigate changes in nutrient uptake or waste excretion by the resistant cell line. This can be done by carrying out metabolite extractions on spent medium of wild-type cells as well as resistant cells, running these through an LC-MS system, and carrying out subsequent analyses through the same bioinformatics pipeline. For example, the transcriptomics data indicated increased expression of cystathionine gamma-lyase. We therefore expected to see increased abundance of L-cysteine, one of the products of the reaction this enzyme catalyses. However, there was no such increase in intracellular L-cysteine levels (electronic supplementary data). There is a possibility that excess L-cysteine was excreted from the cells, which would be detectable through the analyses of spent medium.

The logical hypothesis that has arisen from this study, is that the benzoxaborole may target proteins involved in the *T. brucei* differentiation pathway. Indeed, we did investigate whether two recently discovered repressors of differentiation kinases could be targeted (Jones *et al.*, 2014a). The Mottram group possesses two Lister 427-derived cell lines that overexpress the two kinases, RDK-1 and RDK-2, individually. Alamar blue assays (two replicates only) were carried out to determine whether these cell lines showed increased resistance, but both showed no change when compared to a wild-type control. However, there are thought to be many uncharacterised regulators of differentiation, as shown by a recent RIT-seq study carried out by the K. Matthews group (Mony *et al.*, 2014).

Of utmost importance for this study is to replicate the generation of resistance multiple times *in vitro*. The data presented here was generated from one replicate which was cloned by dilution towards the end of the experiment. In most cases where this process has been successfully employed, several independent drug resistance lines are raised to investigate the compound in question, a good example being the study that investigated the mode of action of the benzoxaborole AN2690 (Rock *et al.*, 2007). Should the PCF phenotype appear in more replicates of this resistance generation, only then could we concretely conclude whether this differentiation phenotype is involved in the generation of AN5568 resistance.

It is important to note that several other phenomena could lead to a “semi-differentiating” transcriptome profile. Firstly, whilst RNA samples were only taken once cells had reached the end of log phase at a density of  $2 \times 10^6$  cells/mL, no work was done to investigate whether these cells were stressed. It has been observed that cells undergoing extreme stress events can alter their expression profiles to match those of differentiating cells (K. Matthews, personal communication). Furthermore, there are few significant hits that have a fold change higher than 2-fold or lower than -2-fold, and critical analysis of the dataset would suggest there is not a large amount of differential expression (C. Clayton, personal communication). To this extent, further molecular biology must be utilised to follow-up on several of the most significant changes in transcript abundance. For

example, western blots and immunofluorescent microscopy could help to understand the extent of VSG downregulation and EP as well as GPEET procyclin upregulation. The Lister 427 line exhibits stable expression of VSG221 (Lamont *et al.*, 1986). Antibodies for this VSG subtype exist, and were in fact successfully used earlier in this study (chapter 3). In addition, antibodies exist for EP procyclin. Western blots could indicate whether the changing surface transcripts are consistent at a protein level.

Further analyses should centre on investigating the extent to which the resistant cell-line proteome matches its transcriptome, and this work could involve Western blotting for several other proteins to analyse whether these also change in abundance. Confirmation of the transcriptome dataset findings should include Northern blot or real-time PCR analysis of mRNA transcript abundance, and this work would greatly compliment the proteomics analyses.

There are several major limitations in the genomic analyses of *T. brucei*. Firstly, whilst recent updates in the TriTrypDB database are aiding in the annotations of the complete genomes of several African trypanosome strains, the majority of the genome remains in the absence of annotations. For example, there were 35 and 134 “hypothetical” proteins identified in the upregulation and downregulation datasets respectively. This amounts to approximately one third of genes in either case, a significant number. Whilst domain analysis using databases such as pfam, COG, SCOP and SMART can help to identify proteins, it remains difficult to implicate hypothetical genes in mechanisms of resistance of drug actions, and further work should include selection of several important hits and further characterisation of these genes/proteins. This is a challenging task, especially if there are a significant number of hypothetical genes to investigate.

It would be interesting to observe the ability of the AN5568<sup>R</sup> line to differentiate. This was in fact attempted using a differentiating trypanosome medium (DTM) (Overath *et al.*, 1986). The recipe to this medium is provided in Appendix A. Unfortunately, the cell doubling time was very slow indeed, with passages carried out once every two weeks on average. Live cells were still visible more than a month after cis-aconitate-mediated induction of differentiation. In contrast, WT cells subjected to the same experiment did not survive past 72 hours. This suggests that the benzoxaborole-resistant cell line does possess pleomorphic capabilities to some extent. For this reason, these cells are interesting not solely from a drug-resistance point of view, but potentially for the study of the molecular mechanisms of differentiation.

## Chapter 5. Loss-of-heterozygosity in African trypanosomes

---

### 5.1 Introduction

---

Genetic diversity in all organisms is crucial for survival and lack of sexual recombination is known to play a vital role in evolution (Diogo *et al.*, 2009, Capewell *et al.*, 2015a, Weir *et al.*, 2016). However, organisms that reproduce asexually have been described. These are clonal populations, which in theory presents an interesting problem with regards to maintaining genetic diversity in the context of species evolution, as explained in [chapter 1](#).

The human-infective *T.b. gambiense* possesses a VSG-like gene called TgsGP, which confers resistance to human serum (Capewell *et al.*, 2013a). This species of African trypanosome can be divided further, based on other mechanisms of human serum resistance (Koffi *et al.*, 2009, Capewell *et al.*, 2011). Type I *T.b. gambiense* are also able to avoid uptake of TLF-1 and thus, Apo-L1, due to downregulation of TbHbHbR (Capewell *et al.*, 2011). In contrast, type II *T.b. gambiense* exhibit variable resistance to human serum, and no downregulation of TbHbHbR has been reported (Capewell *et al.*, 2011). Instead, type II *gambiense* still internalize TLF-1, relying solely on the poorly understood mechanisms of TgsGP-mediated serum resistance (Capewell *et al.*, 2011, Capewell *et al.*, 2013a).

The majority of *T. brucei* species so far described have been reported to reproduce sexually with Mendelian allelic segregation. This has been shown to occur in the Tsetse fly vector stage of the parasite's life cycle, prior to differentiation to mammalian-infective metacyclics (Turner *et al.*, 1990, MacLeod *et al.*, 2005). In contrast, sexual recombination has not been reported in type I *T.b. gambiense* and they are considered a clonal subpopulation of trypanosomes arising from a single ancestor (MacLeod *et al.*, 2001, Koffi *et al.*, 2009, Weir *et al.*, 2016). This *T. brucei* subspecies exhibits low genetic diversity within populations, and higher levels of diversity between different geographical foci (Morrison *et al.*, 2008, Koffi *et al.*, 2009).

One of the recently described mechanisms by which *T.b. gambiense* type I maintains genetic diversity is termed loss-of-heterozygosity (LOH) (Weir *et al.*, 2016). Genomic analyses of SNP distributions of several isolates showed large areas devoid of SNPs, suggesting LOH plays an important role in driving genetic diversity and evolution in this subspecies of African trypanosomes (Weir *et al.*, 2016). Interestingly, chromosome 10 appeared to be particularly affected by LOH. Weir

and colleagues also showed that SNPs are regained independently through the Meselson effect (Weir *et al.*, 2016), as described in [chapter 1](#).

As well as identifying LOH events occurring in natural populations in *T.b. gambiense*, the MacLeod group, through routine genotyping, uncovered several laboratory adapted *T.b. brucei* cell lines showing significant genomic alterations, mainly on chromosome 10, the second largest of the 11 classical *T. brucei* chromosomes (A. MacLeod, personal communication). A combination of genotyping studies, as well as Illumina sequencing, showed widespread LOH across the chromosome in two TREU 927-derived *T. brucei* PCF cell lines. In addition, LOH was found in cell lines derived from STIB 247 *T.b. brucei* and STIB 386 *T.b. gambiense* (Cooper, 2009). This suggests that LOH is not a rarity, and can occur independently *in vitro*.

Interestingly, in both TREU 927-derived PCF, as well as the STIB 247 BSF cell lines, a growth phenotype was observed, where LOH lines exhibited a significantly increased growth rate compared to the wild-type strains (Cooper, 2009). Indeed, it was the increased growth rates that first led the group to discover the genetic alterations (A. MacLeod, personal communication).

The reasons these types of mutations could lead to increased growth rates is poorly understood. Observations indicate that this process has the same consequences in BSF and PCF trypanosomes, suggesting that the underlying metabolic changes are consistent across life cycle stages. However, only *in vitro*, genomics analyses had been carried out up to this point. Therefore, a study was initiated to attempt to elucidate the metabolic perturbations that may have occurred as a result of LOH, and lead to increased growth rates in these mutated lines. The most significant phenotype had been observed in two LOH lines derived independently from a TREU 927 PCF wild-type cell line *in vitro*. These lines were chosen for further *in vitro* studies as well as metabolomics analysis.

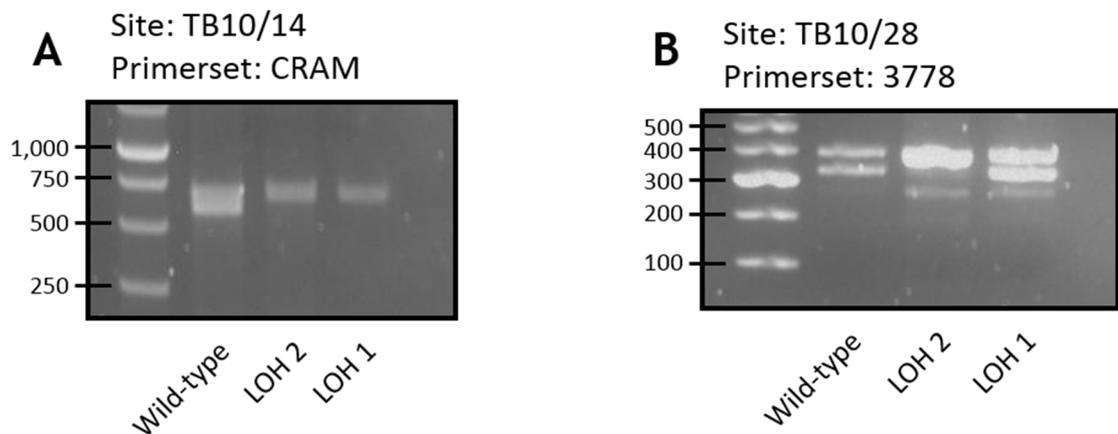
### 5.1.1 Aims & Objectives

The previous work presented in this thesis clearly demonstrates the advantages and potential offered by the use of various omics technologies to dissect *T. brucei* metabolism. We therefore attempted to use similar techniques in order to investigate the metabolic changes occurring in cells that undergo LOH under *in vitro* conditions. Whilst this would not necessarily reflect how LOH impacts genetic diversity and evolution in the field, it would allow clear demonstration in the ways genetic variation can impact cell metabolism. For this study, we decided to utilise two TREU 927 PCF LOH lines previously characterised by Cooper and colleagues in the MacLeod group, with the following objectives in mind:

- To what extent does LOH affect growth of *T. brucei* parasites?
- Can cell culture medium influence these changes?
- Does cell metabolism change in LOH parasites?
- Which genes are located in the LOH region and could mutations in these genes explain the LOH metabolotype?

## 5.2 Occurrence of LOH on chromosome 10

To ensure that the lines being used were identified correctly, we carried out genotyping PCRs of minisatellite markers using primers designed by the MacLeod group (MacLeod *et al.*, 2000). The two primer sets chosen were CRAM (fig. 5-1A), for which both LOH lines were homozygous, and 3778, for which only LOH 2 was homozygous (fig. 5-1B). In both cases, the resultant DNA bands that were seen concurred with previous studies of these cell lines (A. Cooper, personal communication). In all cases, the TREU 927 wild-type genome is heterozygous for these loci on chromosome 10 (A. MacLeod, personal communication).



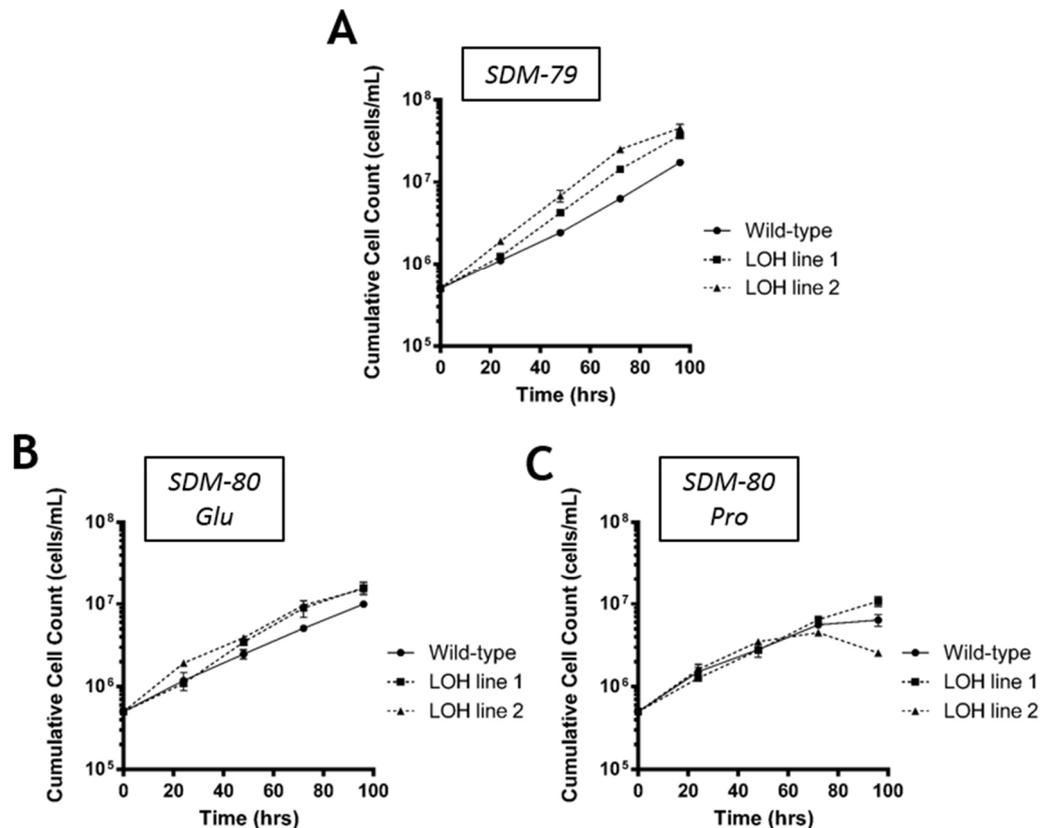
**Figure 5-1: Genotyping PCRs to identify the LOH 1 and LOH 2 cell lines.** A) Both LOH lines were homozygous for the CRAM primer set, which is located between positions 1,805,332 and 1,808,751 on chromosome 10. WT cells were heterozygous for this primer set. B) The 3778 primer set amplifies a region between position 2,404,358 and 2,404,930 on chromosome 10. Only the LOH 2 line was homozygous for this region, which shows that this line possesses a greater extent of LOH on this chromosome compared to the LOH 1 line.

## 5.3 Growth phenotype in LOH lines is glucose-dependent, but not related to chromosome 10 glucose transporter arrays

Growth curves carried out previously by the MacLeod group had shown that several LOH lines, including those derived from a TREU 927 line, exhibited significantly increased growth rates (Cooper, 2009). Growth assays were repeated to confirm these findings (fig. 5-2). Cells from a wild-type TREU 927 PCF line and two LOH lines used for this study were seeded in SDM-79 at a concentration of  $5 \times 10^5$  cells/mL. Cell density was monitored every 24 hours and cells were passaged back to the starting concentration once they had reached the end of log phase (estimated to be  $\sim 2 \times 10^7$  cells/mL).

Results showed that the original phenotype of these cells was reproducible in SDM-79, the standard PCF *T. brucei* culture medium. Here, both LOH lines growing at a significantly faster rate compared to a wild-type control (fig. 5-2A). Growth curves were subsequently set up in SDM-80, a culture medium derived from SDM-79, which allows the user more control over the concentrations of

glucose and proline, the main carbon sources used by PCF trypanosomes (Coustou *et al.*, 2008, Ebikeme *et al.*, 2008). The growth phenotype was maintained in glucose-rich SDM-80 (fig. 5-2B). However, LOH cells grown in proline-rich SDM-80, which contains only 100  $\mu$ M glucose, did not show any growth phenotype compared to wild-type cells. Instead, all three cell lines grew at a similar rate, with the LOH 2 line reaching the end of log phase before the end of the time course (fig. 5-2C). This suggested that in the absence of abundant sources of glucose, LOH cells returned to a wild-type growth phenotype and in addition did not grow to the same densities as those in glucose-rich conditions.



**Figure 5-2: Growth of WT and LOH *T. brucei* TREU 927 PCF cells in procyclic culture media.** A) LOH cells (dotted lines) cultured in the standard PCF medium, SDM-79, showed increased growth rate when compared to WT cells (unbroken line). B) In glucose-rich SDM-80, a growth phenotype was still apparent, although to a lesser degree than that seen in SDM-79. C) LOH cells grown in proline-rich SDM-80, containing only 100  $\mu$ M glucose, did not exhibit a growth phenotype compared to WT cells grown in the same media. Indeed, the LOH 2 line (broken line, triangles) reached the end of log phase more than 24 hours before the end of the time course.

These results suggest that the growth phenotype is dependent on the availability of glucose to the parasite. L-proline is the preferred source of carbon for PCF trypanosomes, as this is the most abundant source available in the tsetse fly haemolymph (Michalkova *et al.*, 2014). However, this life cycle stage is routinely cultured in high levels (>10 mM) of glucose, which has been shown to induce a metabolic shift that causes the semi-complete citric acid cycle to be inactivated in favour of glucose catabolism (Lamour *et al.*, 2005). We therefore hypothesised that the LOH lines had

undergone a metabolic alteration to more efficiently utilise the abundant glucose available from its extracellular environment in culture.

Genetic changes occurring due to LOH could result in significant alterations in gene expression, as well as leading to the production of mutated proteins. It was therefore crucial to understand which genes were affected by LOH. We hypothesised from genotyping experiments that LOH has occurred in laboratory strains from the start of chromosome 10 until approximately the halfway point. Specifically, it was determined from analysis of all LOH lines currently in possession by the MacLeod group, that the furthest occurrence was at approximately 2,400,00 bp, close to the centre of the 4 Mb chromosome. We therefore analysed the gene complement present in this locus using the TREU 927 reference genome (v9.0), which represents the most up-to-date release of this reference genome.

Interestingly, the current version of TriTrypDB predicts a total of 12,094 genes in *T. brucei* TREU 927. Of these, 1,762 predicted genes are found on chromosome 10, and a total of 918 genes were predicted to lie in the LOH region. This was a surprisingly large fraction (~7.6%) of the genome, and included 363 hypothetical proteins. In addition, there were 2 VSG genes (located at the telomeric end of the chromosome), 11 proteins of unknown function (which differ from hypotheticals as they are known to be expressed) and 27 ribosomal proteins. Genes of interest were filtered out from this dataset based on whether they were likely to contribute to energy metabolism. This area of metabolism was chosen with the phenotype of altered growth rate in mind. To further condense this dataset, we filtered genes with very strict parameters in that they should be directly involved in glucose metabolism, given the growth phenotype was glucose dependent, as well mitochondrial metabolism. This resulted in a greatly refined list of 40 genes (table 5-1).

Amongst the genes found in this dataset, were two sequential arrays of glucose transporters and two hexokinases. In addition, several genes encoding subunits of the mitochondrial ATP synthase were present in this part of the chromosome. Further genes involved in glycolytic metabolism included fructose-bisphosphate aldolase, 2-phosphofructo-2-kinase, cytosolic glyceraldehyde 3-phosphate dehydrogenase, and glucose-6-phosphate 1-dehydrogenase (table 5-1). The mitochondrial trypanosome alternative oxidase (TAO) is also located in the LOH region of chromosome 10, along with the mitochondrial malate dehydrogenase, pyruvate dehydrogenase complex E3 binding protein, cytochrome subunits, and the beta-chain of succinyl-CoA ligase. In the scope of this thesis, it was unlikely all of these candidates could be investigated in much detail, but combining these findings with genomics and metabolomics data could aid in pinpointing the underlying mechanisms for increased growth rate in the LOH lines.

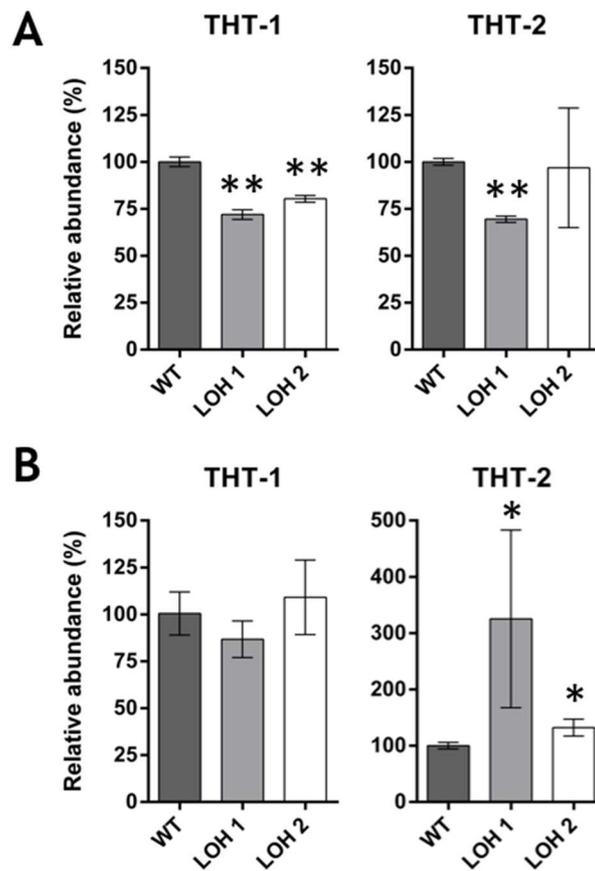
GENE ID	GENOMIC LOCATION (STRAND)	PRODUCT DESCRIPTION
Tb927.10.180	72,359 - 73,442 (+)	ATP synthase F1 subunit gamma protein, putative
Tb927.10.200	74,489 - 75,381 (+)	vacuolar ATP synthase, putative
Tb927.10.280	83,997 - 84,699 (+)	cytochrome oxidase subunit VI (COXVI)
Tb927.10.730	187,319 - 190,277 (+)	ATP synthase, putative
Tb927.10.1570	432,405 - 433,074 (-)	ATPase subunit 9, putative
Tb927.10.2010	526,631 - 529,526 (+)	hexokinase (HK1)
Tb927.10.2020	529,677 - 531,586 (+)	hexokinase (HK2)
Tb927.10.2230	575,971 - 578,176 (+)	NADPH:adrenodoxin oxidoreductase, mitochondrial, putative
Tb927.10.2350	604,810 - 605,707 (+)	pyruvate dehydrogenase complex E3 binding protein, putative
Tb927.10.2490	647,196 - 651,453 (+)	glucose-6-phosphate 1-dehydrogenase (G6PD)
Tb927.10.2550	661,714 - 663,255 (+)	malate dehydrogenase-related
Tb927.10.2560	663,121 - 664,733 (+)	mitochondrial malate dehydrogenase (mMDH)
Tb927.10.2760	706,132 - 708,326 (+)	ATPase family associated with various cellular activities (AAA), putative
Tb927.10.3100	810,236 - 812,441 (+)	glycerol-3-phosphate acyltransferase, putative
Tb927.10.3120	815,459 - 816,499 (+)	cytochrome c oxidase assembly protein, putative
Tb927.10.3260	841,483 - 844,883 (+)	Long-chain-fatty-acid--CoA ligase (Long-chain acyl-CoA synthetase 5) (LACS 5), putative
Tb927.10.3420	884,101 - 884,568 (+)	NADH-ubiquinone oxidoreductase c, putative
Tb927.10.3650	944,892 - 951,041 (-)	NADH-dependent fumarate reductase, putative
Tb927.10.3760	976,594 - 977,949 (-)	vacuolar ATP synthase subunit d, putative
Tb927.10.4520	1,159,602 - 1,161,066 (-)	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase, putative
Tb927.10.5050	1,249,111 - 1,249,681 (+)	Mitochondrial ATP synthase epsilon chain, putative
Tb927.10.5620	1,404,187 - 1,406,027 (+)	fructose-bisphosphate aldolase, glycosomal (ALD)
Tb927.10.5890	1,474,573 - 1,477,290 (-)	Galactose oxidase, central domain containing protein, putative
Tb927.10.6190	1,569,800 - 1,571,566 (-)	aldehyde dehydrogenase, putative (ALDH)
Tb927.10.6880	1,730,060 - 1,731,607 (+)	glyceraldehyde 3-phosphate dehydrogenase, cytosolic (GAP)
Tb927.10.7090	1,783,387 - 1,785,997 (+)	Alternative oxidase, mitochondrial, Trypanosome alternative oxidase, mitochondrial (AOX)
Tb927.10.7410	1,868,838 - 1,871,292 (-)	succinyl-CoA ligase [GDP-forming] beta-chain, putative
Tb927.10.8070	2,011,046 - 2,012,760 (-)	ATPase, putative
Tb927.10.8320	2,056,893 - 2,057,660 (-)	cytochrome oxidase subunit IX (COXIX)
Tb927.10.8440	2,092,756 - 2,095,694 (+)	glucose transporter 1B (THT1-)
Tb927.10.8450	2,095,852 - 2,098,083 (+)	glucose transporter 1E (THT1E)
Tb927.10.8460	2,098,241 - 2,100,472 (+)	glucose transporter, putative
Tb927.10.8470	2,100,630 - 2,102,861 (+)	glucose transporter, putative
Tb927.10.8480	2,103,019 - 2,105,423 (+)	glucose transporter, putative
Tb927.10.8490	2,105,581 - 2,107,989 (+)	glucose transporter, putative
Tb927.10.8500	2,108,147 - 2,110,554 (+)	glucose transporter, putative
Tb927.10.8510	2,110,665 - 2,113,118 (+)	glucose transporter, putative
Tb927.10.8520	2,113,229 - 2,115,680 (+)	glucose transporter, putative
Tb927.10.8530	2,115,838 - 2,118,125 (+)	glucose transporter 2A (THT2A)
Tb927.10.8640	2,139,619 - 2,140,922 (+)	Galactose oxidase, central domain containing protein, putative

Table 5-1: Genes of interest located in the LOH region of *T. brucei* TREU 927 chromosome 10. All annotated genes on chromosome 10 between position 1 and position 2,200,000 were downloaded from the TriTrypDB and filtered depending on the metabolic pathways they were involved with. The 40 genes in this table are all involved in glycolytic metabolism, mitochondrial metabolism, or ATP production.

One of the candidates we chose to pursue further were the trypanosome hexose transporter (THT)-1 and THT-2 arrays (table 5-1). These two arrays differ in their expression in a life-cycle dependent stage, with THT-1 expression occurring only in BSF trypanosomes, whereas both PCF and BSF cells express THT-2 in a glucose-dependent manner (Bringaud & Baltz, 1993). Previous work has shown that these arrays can vary widely in copy number in the field (Barrett *et al.*, 1996). Mitotic recombination could in theory also affect copy number. Increased expression of these transporters could lead to increase glucose uptake, which could lead to the glucose-dependent growth effect observed in the LOH lines.

Due to the poly-cistronic nature of the trypanosome genome, one would hypothesise that increased copy numbers will lead to increased expression, allowing quicker uptake of glucose and increased flux through the glycolytic pathway.

To test this theory, real-time PCR (RT-PCR) analysis was performed on both total DNA and cDNA of all three cell lines (fig. 5-3). Whilst southern blots are generally used to investigate copy number, real-time PCR experiments have been shown to be effective at revealing copy number differences, given that increased copy number will lead to a fluorescence detection at a lower Ct value (D'haene *et al.*, 2010). Real-time PCRs were run using primer sets for both the THT-1 and THT-2 transporters individually. The genetic materials for copy number and expression analyses (total DNA and total RNA respectively), were isolated from cells in log phase of growth as described in the methods.



**Figure 5-3: Comparative Analysis of the THT-1 and THT-2 glucose transporters from total genomic DNA and cDNA in WT and LOH *T. brucei* parasites.** Genetic material was isolated from all three cells as described in the methods, and real-time PCR analyses were carried out by comparing relative abundance of the THT-1 and THT-2 transcripts as well as DNA. A) RT-PCR analysis of total genomic DNA suggested that there was a lower abundance of THT-1 genes in both LOH cell lines compared to WT cells, which could be consistent with a lower copy number. Whilst the same was true for THT-2 transporters in the LOH 1 line, the LOH 2 line showed an identical abundance of THT-2 DNA compared to WT cells. \*= $P < 0.05$ , \*\*= $P < 0.001$ , Student's *t*-test.

Our analyses showed that THT-1 transporters seemed to have a lower copy number in both LOH lines. This decrease was significant in both LOH lines (LOH 1,  $P = 2.3 \times 10^{-6}$ ; LOH 2,  $P = 1.8 \times 10^{-5}$ ) using a Student's *t*-test. Further analysis of RT-PCR data using the THT-2 primer set showed that there was again a lower abundance of DNA in the LOH 1 line ( $P = 1.6 \times 10^{-7}$ ). However, the LOH 2 cell line showed no variation compared to WT cells ( $P = 0.43$ ).

Analysis of cDNA isolates from all three cell lines showed no significant changes in THT-1 transcript abundance (LOH 1,  $P = 0.06$ ; LOH 2,  $P = 0.24$ ). However, THT-2, a transporter expressed in all life cycle stages, was found to exhibit increased abundance in both LOH lines (LOH 1,  $P = 0.03$ ; LOH 2,  $P = 0.003$ ). For the LOH 1 line, this increase was almost 3 fold, albeit with a high amount of variability.

Given these data, it is unlikely that changes in THT-1 glucose transporter were responsible for the growth phenotype in both LOH lines. In LOH 1 cells, THT-1 transcript abundance and copy number were decreased and unchanged respectively. In contrast, there is a possibility that THT-2 expression could have an impact in the LOH lines. Further protein analyses would be needed to confirm this as transcript abundance only provides weak correlation to protein expression in trypanosomatids.

Given the similarities in both the LOH genomes, and their growth phenotypes, there is a significant possibility that the underlying mechanisms are identical.

Other alterations in glycolysis should by no means be ruled out. As shown in table 5-1, there were several other glycolytic proteins of interest found on the LOH region of chromosome 10, including two hexokinases, a glucose 6-phosphate 1-dehydrogenase, a 6-phosphofructo-2-kinase and a fructose-bisphosphate aldolase. LOH across all these genes could in theory cause significant changes in gene regulation or expression, leading to widespread alterations in metabolism not due to one protein in particular, but rather, the combination of many. This phenomenon has been shown in other organisms such as *Candida albicans* (Diogo *et al.*, 2009, Forche *et al.*, 2011).

#### 5.4 Comparative genomics analyses of independent LOH lines

---

Previous work by the MacLeod group identified LOH in STIB 247 and TREU 927, two different strains of *T.b. brucei*. In both cases, there were similar increases in growth rate (A. Cooper, personal communication). STIB 247 is a strain of *T.b. brucei* isolated from a Hartebeest in the Serengeti (Turner *et al.*, 1990), and in contrast to TREU 927, this strain is predominantly homozygous, with up to 94% of commonly used genotyping markers showing homozygosity (Morrison *et al.*, 2009). Several analyses have already shown that this strain exhibits differing phenotypes with regards to virulence and growth, when compared to TREU 927 (Morrison *et al.*, 2009). Furthermore, genetic crosses of the aforementioned *T.b. brucei* strains have been carried out to generate a library of progeny with differing genetics and varied phenotypes, in a bid to uncover quantitative trait loci (QTL) associated with common phenotypes. In addition, a cell line of a *T.b. gambiense* type II strain commonly used in laboratory cultures, STIB 386, has been shown to undergo LOH localised to chromosome 10, in similar fashion to the aforementioned *T.b. brucei* strains (A. MacLeod, personal communication).

A collaboration with the Sanger Institute recently investigated SNPs found in all three of these strains (A. MacLeod, personal communication). Using this dataset, we set out to analyse SNPs along chromosome 10 in both the TREU 927 and STIB 247 strains. Any of these SNPs falling in the LOH region are likely to become homozygous, and previous sequencing data analyses of the LOH lines have indicated that the homozygous alleles remaining are often non-reference (Cooper, 2009). To our surprise, only 27 SNPs were predicted to exist in both strains along the entirety of chromosome 10 (S8). Moreover, only one of these SNPs fell in the region designated as the LOH region responsible for the growth phenotype. Using TriTrypDB, we determined this SNP to lie in a predicted mitochondrial malate dehydrogenase (mMDH) gene, Tb927.10.2560.

This mutation was located at E293, a glutamate residue towards the C-terminal end of mMDH (fig. 5-4A). Using domain analysis tools SMART and Pfam, two domains characteristic of an MDH were

found in the protein. Firstly, a lactate/malate dehydrogenase N-terminal domain (pfam: PF00056, start: 9, end: 151,  $E=1.2e-46$ ), thought to be a Rossmann NAD-binding fold, was present. Secondly, a lactate/malate dehydrogenase C-terminal domain (pfam: PF02866, start: 153, end: 315,  $E=3.9e-46$ ) was also present, and was the domain thought to harbour the E293G variant present in both STIB 247 and TREU 927.

To further probe the potential effect of this SNP, we attempted to predict the 3-dimensional structure of both forms of the mMDH protein. For this, the 3D structure prediction tool Phyre2 was utilised (Pettersen *et al.*, 2004). Protein prediction algorithms were run for both versions of the mMDH protein sequence, using the default parameters. As expected, the Phyre2 package predicted the protein to possess a structure similar to other mMDH proteins. Using the Chimera software (Pettersen *et al.*, 2004), the two proteins were superimposed to analyse whether the SNP caused any predictable structure changes (fig. 5-4B).

In general, no significant change was seen in the protein structure, which maintained its characteristic Rossmann fold and its C-terminal domain (fig. 5-4 & 5-5). We next looked at several residue interactions around the E293G mutation site. The main finding of interest was the presence of several predicted hydrogen bonds between the E293 residue and a serine residue found at position 193 (fig 5-5A). Hydrogen bonds are integral to the secondary structure of a protein (Martin & Derewenda, 1999). In the mutated mMDH protein, 3 of these predicted bonds were absent (fig. 5-5B). It is unknown what effect this would have on the overall stability and folding of the protein, although the Phyre2 software did not predict any significant differences in the two structures when overlaid (fig. 5-4).

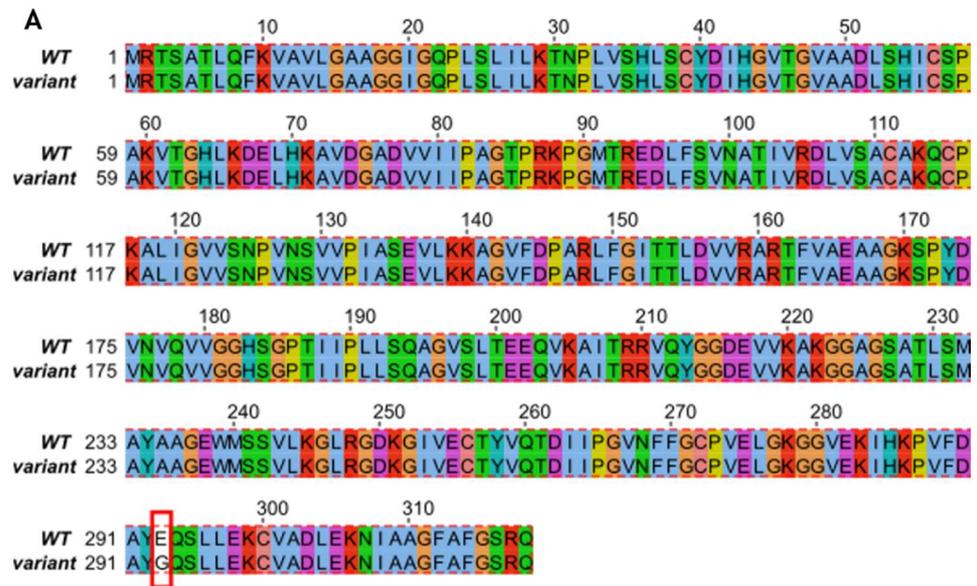
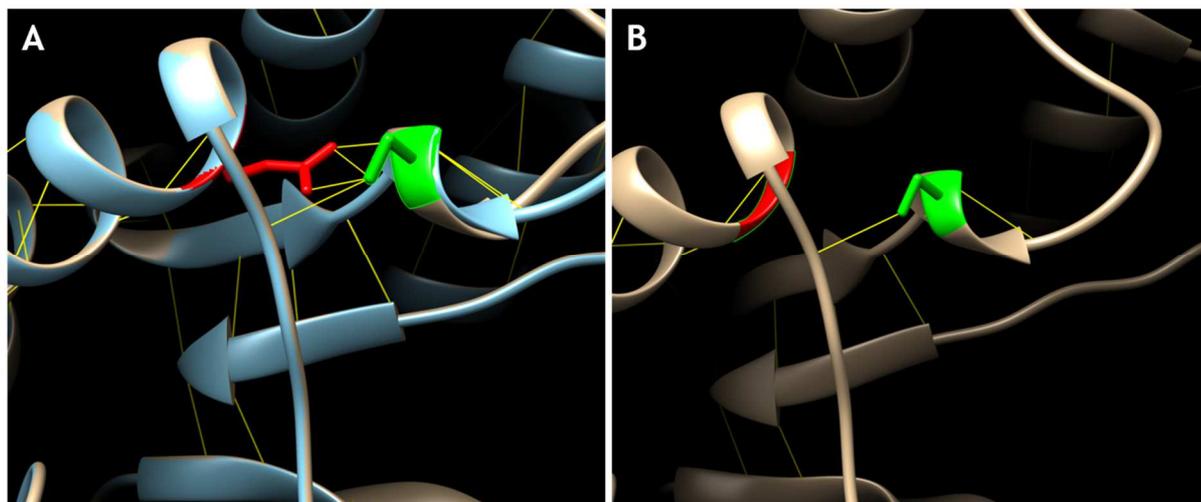


Figure 5-4: Predicted sequence and structure of mitochondrial malate dehydrogenase (mMDH, Tb927.10.2560) in the *T.b. brucei* TREU 927 LOH lines. A) mMDH consists of 319 amino acids. The predicted SNP is found on E293, leading to an E293G mutation. B) The Phyre2 software was used to superimpose the two mMDH variants on top of one another, with the E293 residue highlighted in red. By superimposing the two structures, it was clear that no significant structural alterations are predicted to occur as a result of the E293G mutation.



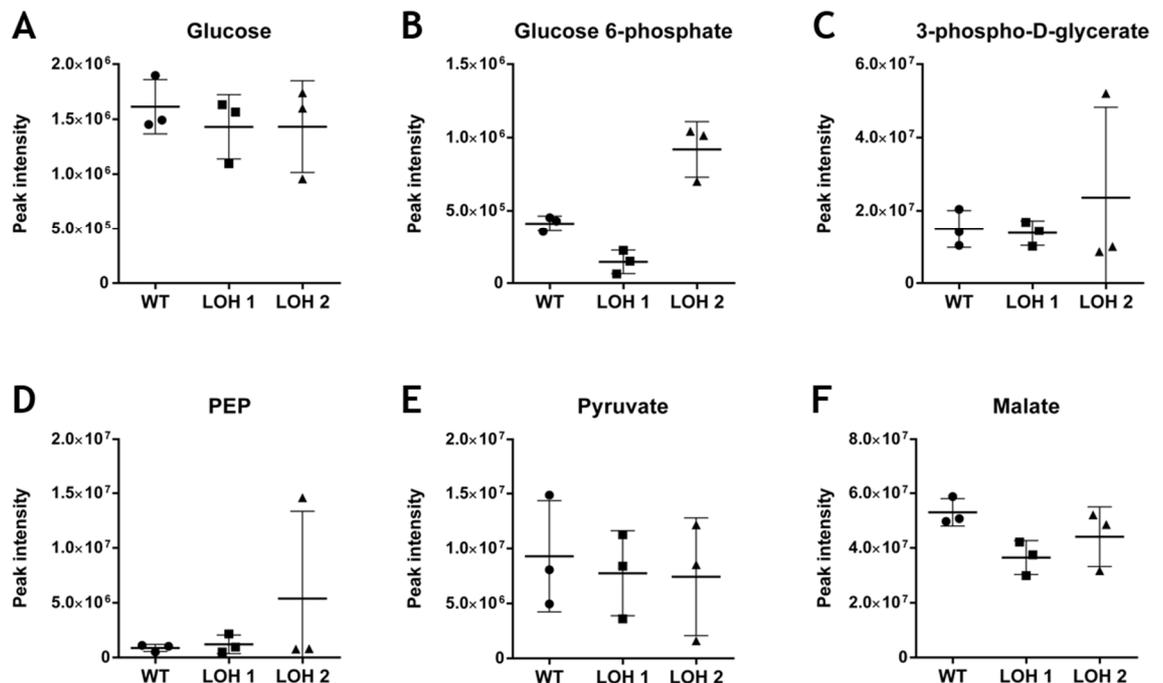
**Figure 5-5: Hydrogen bonding around residue E293 of mMDH.** A) The glutamate residue that harbours the SNP was predicted to form 3 stabilising hydrogen bonds with S193, when structural predictions were carried out using Phyre2. B) By replacing the glutamate residue with the glycine residue as was predicted in the LOH lines removes these hydrogen bonds. However, this did not affect the overall predicted structure of the protein as the serine residue (S193) maintained further hydrogen bonds.

## 5.5 Metabolomics analyses of the TREU 927 LOH lines

To further probe differences in metabolism between wild-type TREU 927 cells and the LOH lines, we carried out LC-MS-based metabolomics analyses. This experiment was untargeted, without the addition of stable isotopes. Samples were cultured in SDM-79 and kept between  $1 \times 10^6$  and  $2 \times 10^7$  cells/mL. The differences in growth rates were taken into account (i.e. wild-type cells were seeded at a higher density) and all samples were extracted in triplicate when cells were nearing the end of log phase growth at  $2 \times 10^7$  cells/mL. Metabolite extractions for this LC-MS experiment were carried out as described in the methods. The samples were run on a Q-Exactive Orbitrap mass spectrometer by Glasgow Polyomics and raw data was processed and analysed using the mzMatch/Ideom pipeline described in [chapter 2](#).

The glycolytic components did not appear to change significantly in the LOH lines, compared to wild-type cells (fig. 5-6). Glucose ( $m/z$ : 180.0635, RT: 10.58 mins) was found at a lower abundance, but not significantly (fig. 5-6A). Pyruvate ( $m/z$ : 88.0160, RT: 6.99 mins), the final component of classical glycolysis, showed no significant change in either of the LOH samples (fig. 5-6E). Interestingly, glucose 6-phosphate ( $m/z$ : 260.0297, RT: 10.57 mins) was decreased and increased in LOH 1 and LOH 2 respectively compared to WT cells (fig. 5-6B), for reasons we were not able to determine. In addition, one replicate of the LOH 2 line exhibited very high levels of both 3-phospho-D-glycerate ( $m/z$ : 185.9929, RT: 10.56 mins) and phosphoenolpyruvate (PEP,  $m/z$ : 167.9823, RT: 10.76 mins), whilst all other replicates for both LOH lines were comparable to the wild-type samples for both metabolites (fig. 5-6C & 5-6D).

Amongst the metabolites showing a similar level in both LOH lines, malate ( $m/z$ : 134.0215, RT: 10.27 mins) was decreased, albeit not significantly, compared to WT cells. Further work could be carried out to identify whether this malate is produced in the glycosome, or in the mitochondrion, as it is produced and consumed in both organelles (Creek *et al.*, 2015). This work could include both stable isotope labelling of glucose or proline, as well as fractionation experiments prior to sample preparation for LC-MS analysis (Fly *et al.*, 2015).

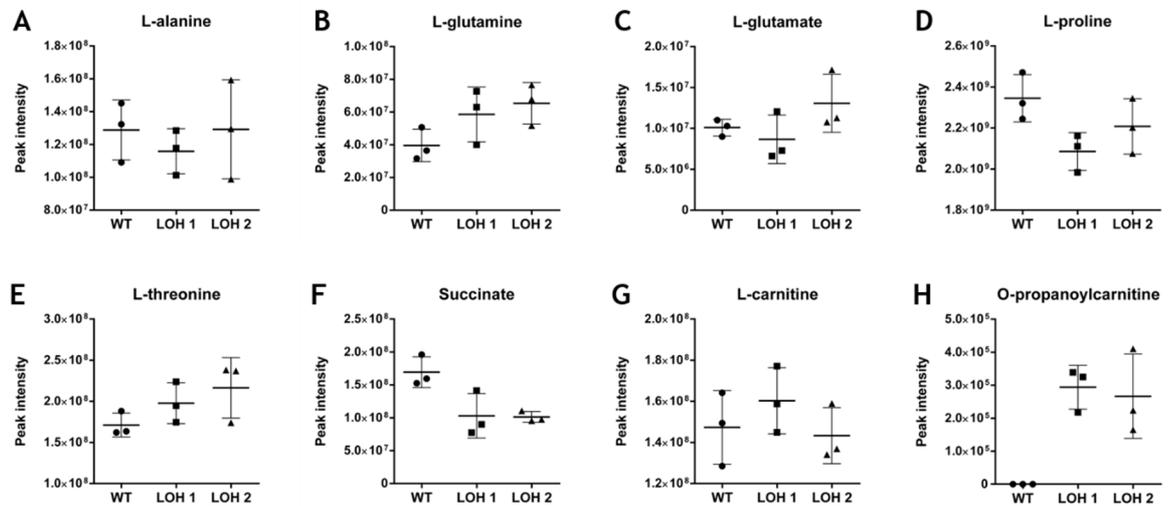


**Figure 5-6: Glycolytic metabolism in WT and LOH cells.** Whilst the growth phenotype observed in LOH cells was glucose-dependent, there were no changes in intracellular glucose (A) levels. Glucose 6-phosphate was less abundant in LOH 1 cells compared to WT (B). However, there was an increase in this metabolite's abundance in LOH 2 cells. Levels of 3-phospho-D-glycerate (C) and PEP (D) did not change in the LOH lines, although 1 LOH sample exhibited high levels of both these metabolites. Pyruvate (E), the end product of glycolysis, was unchanged in LOH cells. However, there was a downward trend in malate (F) levels in both LOH lines.

We next analysed metabolites from other areas of central carbon metabolism (fig. 5-7). Whilst BSF cells primarily utilise glucose as a carbon source, PCF cells are able to switch to proline metabolism, as explained in [chapter 1](#). In particular, components of the semi-complete citric acid cycle, which uses proline to generate ATP, with succinate as an end product (Besteiro *et al.*, 2005), were studied. Here, we found succinate ( $m/z$ : 118.0266, RT: 9.72 mins) to exhibit lower abundance in both LOH lines compared to the wild-type cells (fig. 5-7F). Importantly, this metabolite is produced in varying amounts in both glucose- and proline-rich conditions. Whilst it is the major excretory product of the semi-complete citric acid cycle, it is one of several excretory products in PCF cells where glycolysis is the main central carbon metabolic pathway (Bringaud *et al.*, 2006).

Alanine is generated by an aminotransferase reaction that utilises pyruvate (Spitznagel *et al.*, 2009), and it has been hypothesised that this shunt can function at high rates to remove pyruvate from

the intracellular environment (M. Barrett, personal communication). In the LOH metabolomics dataset, there was no change in alanine ( $m/z$ : 89.0477, RT: 10.63 mins) levels, which suggests this pathway is not more active in LOH cells than in WT cells (fig. 5-7A).



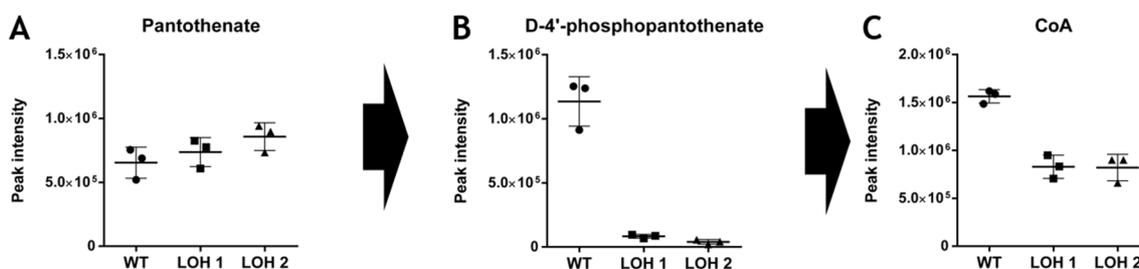
**Figure 5-7: Levels of metabolites involved in PCF central carbon metabolism.** Metabolites involved in mitochondrial metabolism were identified from the metabolomics dataset. L-alanine, a product of an aminotransferase reaction that consumes pyruvate, was unchanged in LOH cells (A). In contrast, L-glutamine levels were increased in LOH cells (B), whilst L-glutamate was only increased in the LOH 2 line (C). Proline levels were decreased in LOH cells (D), an interesting result when compared to the changes in glutamine metabolism. However, this could also reflect the LOH cells' dependency on glucose, as shown by growth analyses. Further changes were seen in L-threonine (E), which was more abundant in LOH cells. Succinate (F), an end product in both glycolysis and TCA metabolism, was reduced in the LOH lines. Whilst no drastic changes were observed in L-carnitine (G), a metabolite derived from it, O-propanoyl-L-carnitine (H), was found only in LOH cells, suggesting it might play a role in generating the growth phenotype seen in LOH cells.

Glutamine can also be fed into the TCA cycle, where it is converted to glutamate and subsequently,  $\alpha$ -ketoglutarate (Besteiro *et al.*, 2005). Interestingly, intracellular L-glutamine ( $m/z$ : 146.0691, RT: 10.53 mins) levels were higher in both LOH lines (fig. 5-7B), and whilst L-glutamate ( $m/z$ : 147.0532, RT: 9.76 mins) levels did not appear to change (fig. 5-7C), there was a decrease observed in L-proline ( $m/z$ : 115.0633, RT: 9.75 mins) levels. This could suggest that the LOH lines take up more L-glutamine and less L-proline than the WT counterparts.

Threonine is an important amino acid that *T. brucei* incorporates into fatty acids (Millerioux *et al.*, 2013). This amino acid was shown to be present in higher intracellular levels in the LOH lines ( $m/z$ : 119.0582, RT: 10.30 mins), compared to WT cells (fig. 5-7E), suggesting more uptake. This can be understood as a need for higher levels of fatty acid production in a rapidly dividing cell.

Carnitine, as discussed in [chapter 3](#) and [chapter 4](#), is crucial in the transport of long-chain acyl groups from fatty acids across the mitochondrial membrane (Klein *et al.*, 1982). L-carnitine ( $m/z$ : 161.1051, RT: 9.81 mins) levels did not change in the LOH lines (fig. 5-7G), but an acyl carnitine derivative, O-propanoyl-L-carnitine ( $m/z$ : 217.1314, RT: 8.10 mins), which was not detected in WT cells, was present in high abundance in both LOH lines.

One of the most significant and surprising findings was observed in the coenzyme A (CoA) biosynthesis pathway which involves phosphorylation of pantothenate, a vitamin (Spry *et al.*, 2008). Whilst pantothenate ( $m/z$ : 219.1107, RT: 7.08 mins) levels were unchanged between the three sample groups (fig. 5-8A), D-4'-phosphopantothenate ( $m/z$ : 299.0771, RT: 9.46 mins) was almost completely absent from either LOH line (fig. 5-8B).



**Figure 5-8: Phosphorylated pantothenate is absent from the LOH lines.** Pantothenate is an essential vitamin required for CoA biosynthesis. Whilst levels of this vitamin were unchanged in the LOH lines compared to WT cells (A), D-4'-phosphopantothenate, the product of the pantothenate kinase reaction was absent in LOH cells (B). Furthermore, CoA levels were reduced in LOH cells compared to WT cells.

In this pathway, D-4'-phosphopantothenate is converted to 3'-dephospho-CoA through R-4'-phosphopantothenoil-L-cysteine and 4'-phosphopantetheine, before CoA is generated (Spry *et al.*, 2008). Whilst the intermediates of this pathway were not detected by LC-MS, CoA ( $m/z$ : 383.5577, RT: 8.66 mins) was found to be decreased in the LOH lines. This decrease, although not as dramatic as that seen in D-4'-phosphopantothenate, was consistent between the two LOH lines studied. These data suggest that the CoA synthesis pathway is much less active in the LOH lines, which could have a significant impact on mitochondrial metabolism.

An attempt was made to search for genes related to pantothenate kinase in the *T. brucei* genome, using the TriTrypDB tool. There is currently one pantothenate kinase subunit (Tb927.11.7290) predicted in the genome and there are homologues of this subunit in all the trypanosomatids. In addition nothing was found in the literature concerning pantothenate kinase in *T. brucei*, indicating it might not have been characterised in this organism. Therefore, further bioinformatics analyses are required to identify other pantothenate kinase subunits in the *T. brucei* genome.

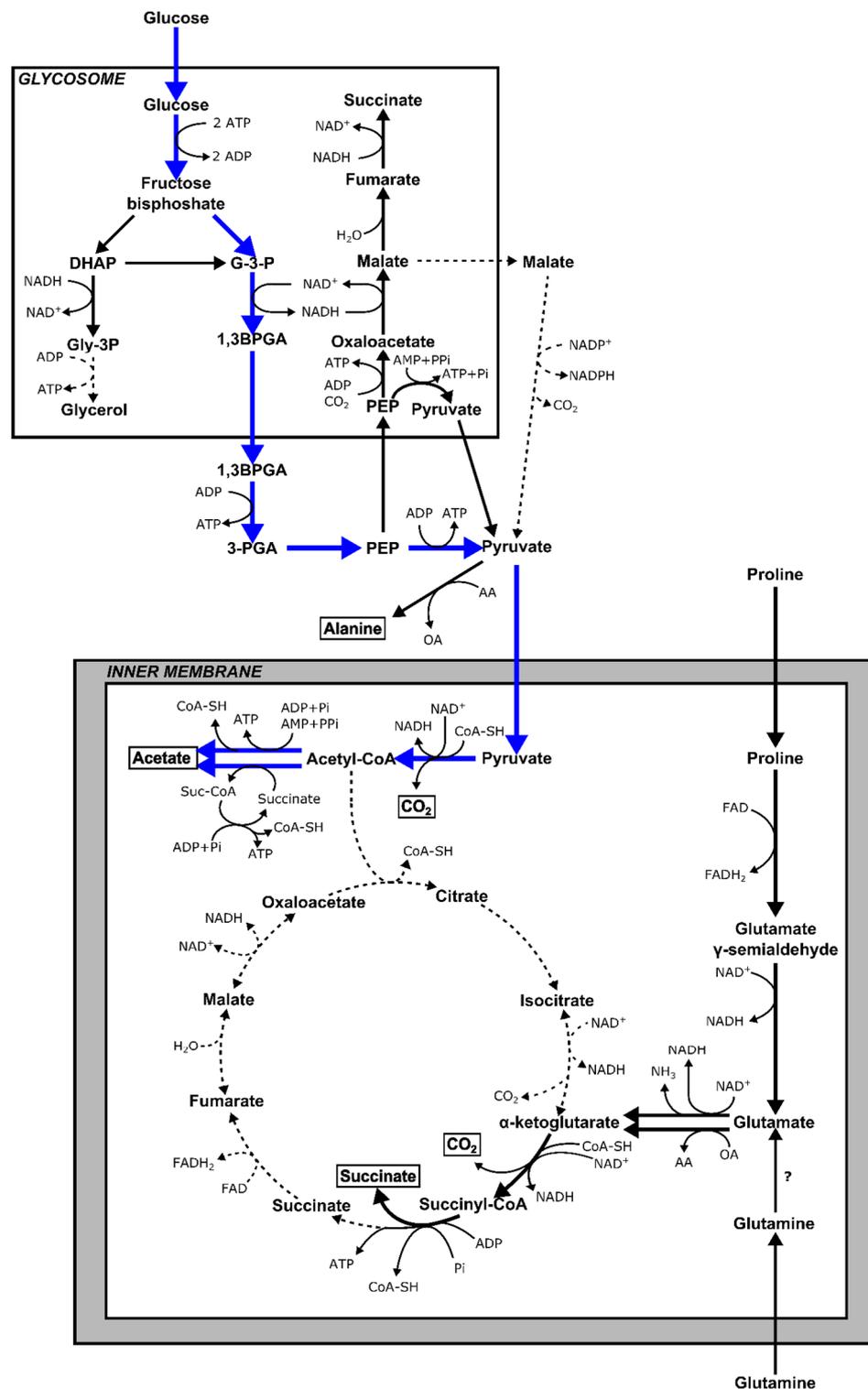
## 5.6 Discussion

---

Recent evidence has emerged suggesting that African trypanosomes can undergo loss-of-heterozygosity (LOH), both in the field and in a laboratory environment (Weir *et al.*, 2016). During routine genotyping experiments, the MacLeod group uncovered significant LOH on chromosome 10, occurring multiple times independently in laboratory-adapted *T.b. brucei* and *T.b. gambiense* strains, both BSF and PCF (Cooper, 2009). Interestingly, these lab-adapted LOH cells exhibited significantly increased growth rates, as shown by this study (fig. 5.2).

In this study, further work was carried out on two PCF LOH TREU 927 *T.b. brucei* cell lines. Subsequent to confirmation of both the previously generated genotyping data and differential growth rates in SDM-79 culture medium, we sought to understand how the genomic alterations led to increased growth rates *in vitro*. Further growth analyses showed that the increased growth rates are glucose-dependent (fig. 5-2), suggesting that the LOH lines had streamlined their metabolism to suit the high levels of glucose present in laboratory culture media. Current evidence shows that PCF *T. brucei* can switch between utilising glucose, or L-proline as its primary carbon source. This enables it to survive in the tsetse fly haemolymph in the absence of glucose (Michalkova *et al.*, 2014). However, it can switch to glycolysis during tsetse fly feeding, utilising glucose from the host bloodmeal (Besteiro *et al.*, 2005).

Genomic analyses of chromosome 10 showed the presence of 918 genes in the LOH region. This list was refined to include only those genes thought to be involved in energy metabolism, given the growth phenotype (table 5-1). This gene set contained the THT-1 and THT-2 glucose transporter arrays, and it was hypothesised that differential expression of these transporters could impact glycolytic metabolism through increased uptake of glucose. Analysis of gene expression in the glucose transporter arrays on chromosome 10 showed increased expression of the THT-2 transporter, which is known to be expressed in a glucose-dependent fashion in PCF trypanosomes in culture (Bringaud & Baltz, 1993). No changes were detected in THT-1 expression in LOH cells compared to wild-type cells. This transporter is normally not expressed in PCF trypanosomes (Bringaud & Baltz, 1993). We were unable to determine any changes in glycolytic flux. One method that could be used is uptake assays involving 2-deoxy-D-glucose (Barrett *et al.*, 1995). This experiment could assess whether uptake of glucose from the extracellular environment was altered. Indeed, one replicate of this experiment was done, with no change detected. However, at least 5 replicates would be required before any conclusions could be drawn from this experiment. Other experiments such as overexpression of the THT transporters could also establish whether increased expression levels can lead to increased growth rates.



**Figure 5-9: Putative changes in central carbon metabolism occurring in LOH cell line, based on metabolomics data.** Lower levels of glycolytic end products such as succinate and malate led us to hypothesize that this pathway of metabolism is utilised less in the LOH lines. In PCF cells, and to a lesser extent BSF cells, pyruvate is also imported to mitochondria, where it is utilised in acetate production, an important pathway for fatty acid biosynthesis in situ. Whilst acetate was not detected by LC-MS, further investigations in collaboration with the Bringaud group are ongoing to determine the fate of pyruvate in LOH cells.

The MacLeod group has utilised F1 progeny from several *T. brucei* crosses to investigate the genetic determinants of several phenotypes including virulence and growth (Morrison *et al.*, 2009, Capewell *et al.*, 2015b). Analysis of the latter phenotype uncovered several QTLs, with the most

significant one located on chromosome 10 (MacLeod, unpublished). This finding supports the theory that loci on this chromosome can influence growth. Further work combining data from the QTL study with the data presented here could lead to novel candidate genes and loci underlying *T. brucei* growth.

Analysis of SNPs on chromosome 10 in both TREU 927 and STIB 247 led to the discovery of one shared SNP located towards the 3' end of the mMDH gene (supp. fig. 8). Bioinformatics approaches were employed in order to investigate the effect of this SNP on the mMDH structure. However, we were unable to conclude whether the SNP could impede enzyme function (fig. 5-4 & 5-5). Molecular biology approaches such as allele replacement strategies could ultimately determine whether this SNP can impact *T. brucei* growth. The hypothesis that a SNP possessed by both strains might be responsible for the growth phenotype is questionable, and it is perhaps more likely that several SNPs in similar regions of chromosome 10 might determine the LOH phenotype.

Metabolomics experiments were carried out on the two LOH lines, using a wild-type TREU 927 line as a control. This experiment, carried out in SDM-79, showed several interesting metabolic phenotypes in the LOH lines. Whilst glycolytic metabolism was not altered to a great extent, a reduction in malate and succinate was found (fig. 5-6 & 5-7). Given that SDM-79 contains high levels of glucose, and the growth phenotype is glucose-dependent, we hypothesise that glycolysis is carried out in these cells in favour of L-proline uptake and TCA metabolism. Therefore, it is likely that the succinate and malate detected in this experiment were formed through glycolytic intermediates. Indeed, stable isotope labelling experiments could allow us to confirm this hypothesis. Reduced malate and succinate could suggest reduced production of these metabolites during glycolysis, with pyruvate instead being transported into the mitochondria for acetate production (fig. 5-9).

Interestingly, a previous study found that pyruvate kinase activity is stimulated by the presence of L-carnitine (Klein *et al.*, 1982). This enzyme presents a key metabolic step as it generates ATP (Coustou *et al.*, 2003). Gilbert and Klein argued in their paper that L-carnitine removes acetyl-CoA, an inhibitory modulator, through the action of carnitine acyl transferases (Klein *et al.*, 1982). Production of acyl carnitines could therefore indicate elevated production of ATP in the cell. Indeed, our metabolomics analyses showed high levels of O-propanoylcarnitine, which was not detected in WT cells. In addition downregulation of the CoA biosynthesis pathway could also impact the generation of inhibitory acetyl-CoA, and our findings regarding D-4'-phosphopantothenate and CoA levels suggest downregulation has occurred in the LOH lines. More work must be carried out to confirm whether perturbations in acyl carnitine production could lead to increased growth rates. A collaboration with the Bringaud group in Bordeaux will determine the fate of pyruvate through the use of NMR, which can detect acetate and succinate (Ebikeme *et al.*, 2010). Further work involving

the use of pantothenate kinase inhibitors is also ongoing, in an attempt to replicate the LOH growth phenotype through downregulation of the CoA synthesis pathway.

Further experiments can also be carried out regarding the genomics of the LOH lines. In particular, RNA-seq analysis could help to narrow down genes of interest on chromosome 10. Lower or higher expression levels of particular genes could implicate them in the altered growth observed in the LOH cells. Due to time constraints, this analysis was not carried out during the study, although previous transcriptome work in the form of digital SAGE has been done on a melarsoprol-resistant STIB 247 line that was retrospectively shown to have undergone LOH in the same area of chromosome 10 as the LOH PCF cell lines used in this study (MacLeod, personal communication). Combining these data could enable the identification of further candidate genes. In particular, genes involved in the regulation of protein synthesis, such as RNA-binding proteins, could be investigated.

The underlying mechanisms leading to LOH remain poorly understood, although current data indicates that mitotic recombination is likely to be responsible (Cooper, 2009). There are several mechanisms of mitotic recombination, including reciprocal exchange and non-reciprocal break-induced repair, both of which have been observed in *T. brucei* in the context of VSG switching (Boothroyd *et al.*, 2009). As discussed in the introduction to this chapter, the occurrence of LOH has implications for the field. Whilst the cell lines studied here are thought to have undergone LOH during *in vitro* culture, recent genomics analyses by the MacLeod group have shown that LOH occurs in *T.b. gambiense* group 1 isolates. Here, it is hypothesised that LOH provides a crucial mechanism to maintain genetic diversity in a clonal population of parasites (Weir *et al.*, 2016). In addition, whilst chromosome 10 in particular undergoes significant amounts of LOH, all chromosomes are affected. This raises the intriguing possibility that LOH occurs across the *T. brucei* genome, even when cultured *in vitro*. Increased growth rates led to the identification of the LOH lines analysed here, but in theory, LOH could occur in the absence of visible phenotypes.

Routine genotyping is important to ensure consistency in the cell lines used for biological research. However, this requires much time, and normally only several markers are used during genotyping assays. In addition, studies in other organisms such as *Candida albicans*, have found that environmental stresses such as oxidative stress, drug pressure and even temperature alterations can influence the formation of DNA strand breakage that leads to LOH (Diogo *et al.*, 2009, Forche *et al.*, 2011). This is important, because if the same is true in *T. brucei* and related kinetoplastids, it highlights the crucial need for standard practices during *in vitro* culturing, in order to maintain genetically identical cell lines. Furthermore, it could implicate LOH as an important driver of drug resistance.

## Chapter 6. Discussion & final thoughts

---

### 6.1 The mode of action of AN5568 – Where do we stand?

---

This study used multiple omics-based approaches combined with generation of resistant cell lines to attempt to elucidate the mechanism of action of AN5568, a novel benzoxaborole compound currently in clinical trials as a therapeutic against HAT. Many of the findings from chapter 3 and 4 were described in detail in the respective discussions. Here, I discuss the current hypotheses on the MoA taking into account this study as well as other research published or carried out since this project was initiated. There are several hypotheses which should be studied in further detail:

1. **AN5568 targets the mitochondrion**
2. **AN5568 causes defects in cytokinesis**
3. **AN5568 inhibits an essential methyltransferase**
4. **AN5568 targets repressors of *T. brucei* differentiation**
5. **AN5568 targets a protein that is more essential in BSF *T. brucei* than PCF**

A collaboration with several groups at the University of Dundee that pioneered the high-throughput RITseq approach has led to several interesting findings regarding AN5568 (M. Barrett, personal communication). One of the proteins found to influence drug resistance was a hypothetical protein (Tb927.9.9030). Further sequence analysis showed this protein to be a predicted mitochondrial transporter with similarities to TIM54. This supports the hypothesis that the mitochondrion could be targeted.

Further work by the Barrett group also investigated drug-drug interactions between AN5568 and drugs known to target the mitochondrion. An isobologram with DB-75, a fluorescent diamidine that localises to the mitochondrion (Lanteri *et al.*, 2006), showed significant antagonism, suggesting similarities between the two MoAs. Work is underway to repeat this experiment (F. Giordani, personal communication). Investigations of AN5568 treatment on akinetoplast trypanosomes, of which several species exist (Carnes *et al.*, 2015), could also provide useful information as to the MoA of AN5568.

There have been many reports of mitochondrial defects resulting in arrest of cytokinesis. For example, depletion of the mitochondrial acyl carrier protein was shown to lead to defects in kinetoplast segregation (Clayton *et al.*, 2011). In addition, RNAi-mediated knock-down of the serine palmitoyltransferase enzyme involved in sphingolipid synthesis also leads to aberrant kinetoplast

division (Fridberg *et al.*, 2008). *T. brucei* possess a single dynamin-like protein (DLP) which is involved in mitochondrial fission (Chanez *et al.*, 2006). One study has shown that a DLP KO leads to arrest of cytokinesis in the 2N2K stage of cell division (Chanez *et al.*, 2006). In contrast, AN5568 led to enrichment of 2N1K and MNMK cells. However, the aforementioned studies do confirm the involvement of mitochondrial proteins in cell division.

During the final writing stages of this thesis, a paper was published on the investigation of the AN5568 target (Jones *et al.*, 2015). This study used omics approaches, in particular genomic and proteomics, to attempt to unravel the mode of action of AN5568. Interestingly, this study observed the same cell cycle inhibition phenotype that was observed in this study. In addition, Jones and colleagues study generated an AN5568-resistant cell line, although no transcriptomic studies of this cell line were undertaken. The resistance phenotype was found to be reversible, similar to our observations (Jones *et al.*, 2015).

The comparative metabolomics analyses between AN5568 and various other drugs, provided a detailed insight into the mechanisms of benzoxaborole action. In particular, the similarities observed between AN5568 and sinefungin led to a hypothesis that a methyltransferase is targeted. There is further work that could be done with these datasets. For example, in their 2007 study on AN2690, Rock and colleagues noted that the leucyl-tRNA synthetase-targeting anti-fungal formed an adduct with AMP, which was mediated by the boronic acid binding hydroxide groups on the adenosine (Rock *et al.*, 2007). We did not discover any adducts in the metabolomics analyses, and work should be carried out to identify whether this phenomenon occurs during AN5568 treatment with, for example, the hydroxide groups on AdoMet, which is highly increased in drug-treated cells.

AdoMet is the second most utilised substrate, after ATP (Fontecave *et al.*, 2004). Whilst its importance as a methyl group donor has long been established, it has been reported as a donor for several other reactions. In one of these, known as a radical SAM reaction, AdoMet acts as a source of 5'-deoxyadenosyl radicals (Sofia *et al.*, 2001, Fontecave *et al.*, 2004). This reaction is catalysed by enzymes with an iron-sulphur (Fe-S) cluster and generates a highly reactive radical as well as L-methionine. Interestingly, our metabolomics analyses showed high levels of deoxyadenosine (S1), which could result from the formation of radicals. In addition, adenosine is highly increased in the AN5568<sup>R</sup> cell line (S7), which could be a compensatory mechanism of some sort. Finally, the mitochondrial enzyme identified by collaborators in Dundee was also thought to contain a radical SAM domain. Therefore, other processes that involve AdoMet as a substrate should not be excluded from hypotheses regarding the AN5568 MoA.

There are 145 MTases predicted in the *T. brucei* genome and a RIT-seq approach similar to that carried out with the kinome (Jones *et al.*, 2014a) could provide a significant insight into the roles

these enzymes play in the biology of *T. brucei*, as well as whether any of these are targeted by AN5568. Indeed, MTases that regulate differentiation could exist, but are as-of-yet unidentified. Histone methylases with important roles in antigenic variation have already been described (Figueiredo *et al.*, 2008). Furthermore, the DOT1 MTases, of which *T. brucei* possesses two, have been implicated in differentiation and cell cycle regulation in other organisms (Kim *et al.*, 2014). Finally, there are several further predicted mRNA capping MTases in the *T. brucei* genome, as evidenced in [chapter 3](#). Therefore, the hypothesis that AN5568 targets the spliced leader should not be completely ruled out.

The generation of an AN5568-resistant cell line led to the hypothesis that the benzoxaborole targets repressors, or regulators, of differentiation. The procyclic phenotype observed in the transcriptomic and metabolomics datasets allows the parasite to “escape” sensitivity to the compound. This process was reversible, although the parameters of unstable resistance were not quantified in this study. Interestingly, a collaboration with the Clayton group led to the finding that the PCF-specific PIP39 protein was upregulated in WT BSF *T. brucei* treated with another benzoxaborole (D. Begolo, personal communication). This finding supports the theory that differentiation is targeted. To further explore this hypothesis, a cell line expressing GFP with the 3'-UTR of a differentiation marker such as EP procyclic or PAD could be utilised to look for activation of differentiation cascades during benzoxaborole treatment. In addition, transcript and protein analyses such as real-time PCR and western blotting respectively, could further probe upregulation of SSF or PCF specific proteins.

One of the top priorities is to assess the infectivity of the AN5568<sup>R</sup> cell line *in vivo*. Given the downregulation of VSG transcripts, it is likely this cell line would quickly succumb to the host immune response. In addition, further studies of the differentiation phenotype are planned to identify abnormal regulation of RBPs and other master regulators that could activate differentiation, whilst simultaneously repress morphological changes such as those normally occurring in the transformation of long slender BSF cells to SSF and PCF.

Work on other benzoxaborole scaffolds has shown that the targets of these compounds vary widely. Aside from AN2690 (Tavaborole), which is undergoing FDA review as a cream to treat onychomycosis (FDA, 2014, Markham, 2014), several other benzoxaborole drug targets have been identified, mostly through work carried out by researchers at Anacor Pharmaceuticals Inc. One study developed a novel hinge-motif in the oxaborole scaffold which allowed strong inhibition of rho kinases (Akama *et al.*, 2013b). Further work on 6-(benzoylamino)benzoxaboroles has shown these modified scaffolds can also act as anti-inflammatories and might specifically target Toll-like receptor-mediated inflammation (Akama *et al.*, 2013a, Dong *et al.*, 2013). Optimisation of

benzoxaboroles targeting tRNA synthetases has also continued since the discovery of AN2690 (Baker *et al.*, 2011, Goldstein *et al.*, 2013).

These publications, in addition to the knowledge that benzoxaboroles form strong –diol entities with their target (Liu *et al.*, 2014), suggest that these compounds could bind to more than one protein in their host system. Whilst Rock and colleagues identified the oxaborole tRNA trapping (OBORT) mechanism, this might not apply to all benzoxaboroles, and additional mechanisms of action might exist.

Unfortunately, this study did not pinpoint a mechanism by which AN5568 targets the *T. brucei* parasite. However, significant strides have been made to identify this mechanism, and with follow-up work planned, as well as plenty of scientifically valid hypotheses to explore, it is expected that the MoA of this exciting new drug against HAT will not remain elusive for long.

## 6.2 The relationship between metabolism and its genetic determinants – LOH as a case study

---

The integration of genomics and metabolomics, which has been referred to as genetical metabolomics (Keurentjes, 2009), enables a very detailed insight into the mechanisms by which genomic alterations can impact metabolic traits. This concept is well established and has been explored through the use of genetically modified cell lines and subsequent metabolomics analysis (Creek *et al.*, 2015). However, the study described here has important implications for the field, given the LOH phenotype is a natural phenomenon that can impact cellular growth and metabolism.

With recently emerged high-throughput technologies, one can now phenotype metabolic traits across a species in a similar way to genotypic traits, where instead of QTLs, the metabolomic (m)QTLs are investigated, with regards to phenotypes of interest such as growth and drug resistance (Nicholson *et al.*, 2011, Lewis *et al.*, 2014, Hill *et al.*, 2015, Kraus *et al.*, 2015). This approach has proved highly successful in recent years, in particular in the field of plant biology, where large scale crosses are relatively simple to carry out (Carreno-Quintero *et al.*, 2013). With the *T. brucei* genetic crosses carried out by the MacLeod group several years ago (MacLeod *et al.*, 2005, Cooper *et al.*, 2008), this type of approach could give rise to valuable insights into trypanosome metabolism.

One significant question that remains for the LOH study presented here is whether the natural infectivity of LOH cell lines is altered subsequent to the chromosomal alteration. In PCF lines, we were able to show that increased growth was glucose-dependent. However, PCF cells naturally reside in the midgut of the tsetse fly where, aside from bloodmeal contents, the major carbon sources are proline and glutamine, both components of the tsetse haemolymph (Besteiro *et al.*, 2005, Michalkova *et al.*, 2014). It is not known whether LOH cells would outcompete WT cells in this situation. However, in the bloodstream, where glucose is the primary carbon source, LOH lines could in theory maintain a growth advantage over WT cells, should the phenotype be identical to that observed in PCF strains *in vitro*. The priority is therefore to carry out *in vivo* experiments with LOH lines, to determine whether genetic determinants of metabolism can influence natural infections.

The recent genomics study carried out on the clonal *T.b. gambiense* type 1 highlights the mechanisms by which genetic diversity can be generated, even in clonal trypanosomes (Weir *et al.*, 2016). This knowledge is valuable as it will aid in the prediction of drug resistance and potentially affect vaccine efficacy, should these become available in the future (Weir *et al.*, 2016). It also underlines the fact that trypanosome genetics might be far more complex than thought.

### 6.3 Omics in the study of biological organisms – final thoughts

---

Biological research has increasingly profited from rapidly modernising computer science which has enabled the generation, as well as analyses, of large datasets (Berger *et al.*, 2013). Metabolomics alone is supported by a myriad of computer tools designed specifically for the analysis of these types of data (Misra & van der Hooft, 2016). The integration of laboratory science and bioinformatics, especially with regards to building and developing computer models, has aided in many areas of science, including drug discovery (Haanstra & Bakker, 2015).

One aspect of omics studies that was not covered during this project is the study of individual organelles in biological systems. Metabolomics studies of fractionated cells have been carried out in plant biology, with particular emphasis on the chloroplast (Geigenberger *et al.*, 2011, Kueger *et al.*, 2012). Cellular fractionation could be used to further understand the biology of the trypanosome, both in the context of general metabolism (e.g. mitochondrial malate and succinate versus that present in the glycosome), and in the study of drug target deconvolution (e.g. confirmation that reduced transporter expression is correlated to decreased uptake of trypanocidals). Whilst these techniques are in early developmental stages, they will provide new opportunities to investigate metabolism.

Although the benefits (and some disadvantages) of omics research have been openly discussed throughout this thesis, there are caveats associated with these techniques. For the MoA study presented here, 7 metabolomics, 3 proteomics and a transcriptomic dataset were generated, each containing vast amounts of information. For the purpose of this study, data was mined to search for clues relating to the MoA of AN5568. However, there were also further discoveries that are of great interest in their own right. For example, knock-down of the TbCgm1 cap MTase leads to a build-up of free nucleotides, which might be expected during arrest in translation, but also suggests high levels of mRNA degradation. The point being made, is that the inferences made from these datasets must be well-informed, and researchers must be well equipped to deal with, and understand the generation of, the datasets in question.

It is fundamental to piece together the biologically relevant information from omics datasets, and to interpret and present them in a logical manner, as this enables the generation of new hypotheses that can be explored. In his book “La Science et l’Hypothèse”, the French philosopher Henri Poincaré wrote “*Science is built up with facts, as a house is with stones. But a collection of facts is no more a science than a heap of stones is a house*” (Poincaré, 1905). As long as omics research adheres to this principle, scientists will benefit from it for generations to come.

## References

- Acestor, N., Panigrahi, A. K., Ogata, Y., Anupama, A. & Stuart, K. D. (2009) Protein composition of *Trypanosoma brucei* mitochondrial membranes. *Proteomics*, **9**, 5497-508.
- Achcar, F., Fadda, A., Haanstra, J. R., Kerkhoven, E. J., Kim, D. H., Leroux, A. E., Papamarkou, T., Rojas, F., Bakker, B. M., Barrett, M. P., Clayton, C., Girolami, M., Krauth-Siegel, R. L., Matthews, K. R. & Breitling, R. (2014a) The silicon trypanosome: a test case of iterative model extension in systems biology. *Advances in Microbial Physiology*, **64**, 115-43.
- Achcar, F., Kerkhoven, E. J., Bakker, B. M., Barrett, M. P. & Breitling, R. (2012) Dynamic modelling under uncertainty: the case of *Trypanosoma brucei* energy metabolism. *PLoS Computational Biology*, **8**, e1002352.
- Achcar, F., Kerkhoven, E. J. & Barrett, M. P. (2014b) *Trypanosoma brucei*: meet the system. *Current Opinion in Microbiology*, **20**, 162-9.
- Adams, D. O. & Yang, S. F. (1977) Methionine metabolism in apple tissue: implication of s-adenosylmethionine as an intermediate in the conversion of methionine to ethylene. *Plant Physiology*, **60**, 892-6.
- Aitcheson, N., Talbot, S., Shapiro, J., Hughes, K., Adkin, C., Butt, T., Shearer, K. & Rudenko, G. (2005) VSG switching in *Trypanosoma brucei*: antigenic variation analysed using RNAi in the absence of immune selection. *Molecular Microbiology*, **57**, 1608-22.
- Akama, T., Dong, C., Virtucio, C., Freund, Y. R., Chen, D., Orr, M. D., Jacobs, R. T., Zhang, Y. K., Hernandez, V., Liu, Y., Wu, A., Bu, W., Liu, L., Jarnagin, K. & Plattner, J. J. (2013a) Discovery and structure-activity relationships of 6-(benzoylamino)benzoxaboroles as orally active anti-inflammatory agents. *Bioorganic and Medicinal Chemistry Letters*, **23**, 5870-3.
- Akama, T., Dong, C., Virtucio, C., Sullivan, D., Zhou, Y., Zhang, Y. K., Rock, F., Freund, Y., Liu, L., Bu, W., Wu, A., Fan, X. Q. & Jarnagin, K. (2013b) Linking phenotype to kinase: identification of a novel benzoxaborole hinge-binding motif for kinase inhibition and development of high-potency rho kinase inhibitors. *The Journal of Pharmacology and Experimental Therapeutics*, **347**, 615-25.
- Ali, J. A., Creek, D. J., Burgess, K., Allison, H. C., Field, M. C., Maser, P. & De Koning, H. P. (2013) Pyrimidine salvage in *Trypanosoma brucei* bloodstream forms and the trypanocidal action of halogenated pyrimidines. *Molecular Pharmacology*, **83**, 439-53.
- Alsford, S., Eckert, S., Baker, N., Glover, L., Sanchez-Flores, A., Leung, K. F., Turner, D. J., Field, M. C., Berriman, M. & Horn, D. (2012) High-throughput decoding of antitrypanosomal drug efficacy and resistance. *Nature*, **482**, 232-6.
- Alsford, S., Horn, D. & Glover, L. (2009) DNA breaks as triggers for antigenic variation in African trypanosomes. *Genome Biology*, **10**, 223.
- Alsford, S., Turner, D. J., Obado, S. O., Sanchez-Flores, A., Glover, L., Berriman, M., Hertz-Fowler, C. & Horn, D. (2011) High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Research*, **21**, 915-24.
- Ammerman, M. L., Tomasello, D. L., Faktorova, D., Kafkova, L., Hashimi, H., Lukes, J. & Read, L. K. (2013) A core MRB1 complex component is indispensable for RNA editing in insect and human infective stages of *Trypanosoma brucei*. *PLoS One*, **8**, e78015.
- Anacor. 2015. *Pipeline > Neglected Diseases* [Online]. Available: [http://www.anacor.com/neglected\\_diseases.php](http://www.anacor.com/neglected_diseases.php) [Accessed 16th January 2016].
- Andersen, M. P., Nelson, Z. W., Hetrick, E. D. & Gottschling, D. E. (2008) A genetic screen for increased loss of heterozygosity in *Saccharomyces cerevisiae*. *Genetics*, **179**, 1179-95.
- Anderson, J., Phan, L. & Hinnebusch, A. G. (2000) The Gcd10p/Gcd14p complex is the essential two-subunit tRNA(1-methyladenosine) methyltransferase of *Saccharomyces cerevisiae*.

- Proceedings of the National Academy of Sciences of the United States of America*, **97**, 5173-8.
- Anderson, J. B., Sirjusingh, C. & Ricker, N. (2004) Haploidy, diploidy and evolution of antifungal drug resistance in *Saccharomyces cerevisiae*. *Genetics*, **168**, 1915-23.
- Andrews, S. 2010. *FastQC: a quality control tool for high throughput sequence data* [Online]. Available: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> [Accessed 12th December 2015].
- Annis, D. A., Nickbarg, E., Yang, X., Ziebell, M. R. & Whitehurst, C. E. (2007) Affinity selection-mass spectrometry screening techniques for small molecule drug discovery. *Current Opinion in Chemical Biology*, **11**, 518-26.
- Aoyama, Y. & Motokawa, Y. (1981) L-Threonine dehydrogenase of chicken liver. Purification, characterization, and physiological significance. *The Journal of Biological Chemistry*, **256**, 12367-73.
- Archer, S. K., Inchaustegui, D., Queiroz, R. & Clayton, C. (2011) The cell cycle regulated transcriptome of *Trypanosoma brucei*. *PLoS One*, **6**, e18425.
- Arhin, G. K., Li, H., Ullu, E. & Tschudi, C. (2006a) A protein related to the vaccinia virus cap-specific methyltransferase VP39 is involved in cap 4 modification in *Trypanosoma brucei*. *RNA*, **12**, 53-62.
- Arhin, G. K., Ullu, E. & Tschudi, C. (2006b) 2'-O-methylation of position 2 of the trypanosome spliced leader cap 4 is mediated by a 48 kDa protein related to vaccinia virus VP39. *Molecular and Biochemical Parasitology*, **147**, 137-9.
- Ariyanayagam, M. R. & Fairlamb, A. H. (2001) Ovothiol and trypanothione as antioxidants in trypanosomatids. *Molecular and Biochemical Parasitology*, **115**, 189-98.
- Aslett, M., Aurecochea, 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. & Wang, H. (2010) TriTrypDB: a functional genomic resource for the Trypanosomatidae. *Nucleic Acids Research*, **38**, D457-62.
- Auty, H., Torr, S. J., Michoel, T., Jayaraman, S. & Morrison, L. J. (2015) Cattle trypanosomosis: the diversity of trypanosomes and implications for disease epidemiology and control. *Revue Scientifique et Technique*, **34**, 587-98.
- Babokhov, P., Sanyaolu, A. O., Oyibo, W. A., Fagbenro-Beyioku, A. F. & Iriemenam, N. C. (2013) A current analysis of chemotherapy strategies for the treatment of human African trypanosomiasis. *Pathogens and Global Health*, **107**, 242-52.
- Bacchi, C. J., Sufrin, J. R., Nathan, H. C., Spiess, A. J., Hannan, T., Garofalo, J., Alecia, K., Katz, L. & Yarlett, N. (1991) 5'-Alkyl-substituted analogs of 5'-methylthioadenosine as trypanocides. *Antimicrobial Agents and Chemotherapy*, **35**, 1315-20.
- Baker, J. R. (1995) The subspecific taxonomy of *Trypanosoma brucei*. *Parasite*, **2**, 3-12.
- Baker, M. (2013) Big biology: The 'omes puzzle. *Nature*, **494**, 416-9.
- Baker, N., Glover, L., Munday, J. C., Aguinaga Andres, D., Barrett, M. P., De Koning, H. P. & Horn, D. (2012) Aquaglyceroporin 2 controls susceptibility to melarsoprol and pentamidine in African trypanosomes. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 10996-1001.
- Baker, S. J., Tomsho, J. W. & Benkovic, S. J. (2011) Boron-containing inhibitors of synthetases. *Chemical Society Reviews*, **40**, 4279-85.
- Balasegaram, M., Harris, S., Checchi, F., Hamel, C. & Karunakara, U. (2006) Treatment outcomes and risk factors for relapse in patients with early-stage human African trypanosomiasis (HAT) in the Republic of the Congo. *Bulletin of the World Health Organization*, **84**, 777-82.
- Balmer, O., Beadell, J. S., Gibson, W. & Caccone, A. (2011) Phylogeography and taxonomy of *Trypanosoma brucei*. *PLoS Neglected Tropical Diseases*, **5**, e961.

- Bangs, J. D., Crain, P. F., Hashizume, T., McCloskey, J. A. & Boothroyd, J. C. (1992) Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. *The Journal of Biological Chemistry*, **267**, 9805-15.
- Barker, R. H., Jr., Liu, H., Hirth, B., Celatka, C. A., Fitzpatrick, R., Xiang, Y., Willert, E. K., Phillips, M. A., Kaiser, M., Bacchi, C. J., Rodriguez, A., Yarlett, N., Klinger, J. D. & Sybertz, E. (2009) Novel S-adenosylmethionine decarboxylase inhibitors for the treatment of human African trypanosomiasis. *Antimicrobial Agents and Chemotherapy*, **53**, 2052-8.
- Barrett, M. P. (2010) Potential new drugs for human African trypanosomiasis: some progress at last. *Current Opinion in Infectious Diseases*, **23**, 603-8.
- Barrett, M. P. (2015). Why aren't there more scientists in the National Portrait Gallery? *New Statesman*. London, UK.
- Barrett, M. P., Boykin, D. W., Brun, R. & Tidwell, R. R. (2007) Human African trypanosomiasis: pharmacological re-engagement with a neglected disease. *British Journal of Pharmacology*, **152**, 1155-71.
- Barrett, M. P., Bringaud, F., Doua, F., Melville, S. E. & Baltz, T. (1996) Hypervariability in gene copy number for the glucose transporter genes in trypanosomes. *Journal of Eukaryotic Microbiology*, **43**, 244-249.
- Barrett, M. P. & Croft, S. L. (2012) Management of trypanosomiasis and leishmaniasis. *British Medical Bulletin*, **104**, 175-96.
- Barrett, M. P., Tetaud, E., Seyfang, A., Bringaud, F. & Baltz, T. (1995) Functional expression and characterization of the *Trypanosoma brucei* procyclic glucose transporter, THT2. *The Biochemical Journal*, **312**, 687-91.
- Bauer, S., Morris, J. C. & Morris, M. T. (2013) Environmentally regulated glycosome protein composition in the African trypanosome. *Eukaryotic Cell*, **12**, 1072-9.
- Begolo, D., Erben, E. & Clayton, C. (2014) Drug target identification using a trypanosome overexpression library. *Antimicrobial Agents and Chemotherapy*, **58**, 6260-4.
- Benarroch, D., Jankowska-Anyszka, M., Stepinski, J., Darzynkiewicz, E. & Shuman, S. (2010) Cap analog substrates reveal three clades of cap guanine-N2 methyltransferases with distinct methyl acceptor specificities. *RNA*, **16**, 211-20.
- Bennett, R. J., Forche, A. & Berman, J. (2014) Rapid mechanisms for generating genome diversity: whole ploidy shifts, aneuploidy, and loss of heterozygosity. *Cold Spring Harbor Perspectives in Medicine*, **4**.
- Berg, M., Garcia-Hernandez, R., Cuypers, B., Vanaerschot, M., Manzano, J. I., Poveda, J. A., Ferragut, J. A., Castanys, S., Dujardin, J. C. & Gamarro, F. (2015) Experimental resistance to drug combinations in *Leishmania donovani*: metabolic and phenotypic adaptations. *Antimicrobial Agents and Chemotherapy*, **59**, 2242-55.
- Berg, S. S., Brown, K. N., Hill, J. & Wragg, W. R. (1961) A new prophylactic trypanocidal drug, 2,7-di-(m-aminophenyldiazoamino)-10-ethyl-9-phenyl-phenanthridinium chloride dihydrochloride (M & B 4596). *Nature*, **192**, 367-8.
- Berger, B., Peng, J. & Singh, M. (2013) Computational solutions for omics data. *Nature Reviews Genetics*, **14**, 333-46.
- Berger, B. J., Carter, N. S. & Fairlamb, A. H. (1993) Polyamine and pentamidine metabolism in African trypanosomes. *Acta Tropica*, **54**, 215-24.
- 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., Rabinowitsch, E., Rajandream, M. A., Reitter, C., Salzberg, S. L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A. J., Tallon,

- L., Turner, C. M., Tait, A., Tivey, A. R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M. D., Embley, T. M., Gull, K., Ullu, E., Barry, J. D., Fairlamb, A. H., Opperdoes, F., Barrell, B. G., Donelson, J. E., Hall, N., Fraser, C. M., *et al.* (2005) The genome of the African trypanosome *Trypanosoma brucei*. *Science*, **309**, 416-22.
- Besteiro, S., Barrett, M. P., Riviere, L. & Bringaud, F. (2005) Energy generation in insect stages of *Trypanosoma brucei*: metabolism in flux. *Trends in Parasitology*, **21**, 185-91.
- Biagini, G. A., Fisher, N., Shone, A. E., Mubarak, M. A., Srivastava, A., Hill, A., Antoine, T., Warman, A. J., Davies, J., Pidathala, C., Amewu, R. K., Leung, S. C., Sharma, R., Gibbons, P., Hong, D. W., Pacorel, B., Lawrenson, A. S., Charoensutthivarakul, S., Taylor, L., Berger, O., Mbekeani, A., Stocks, P. A., Nixon, G. L., Chadwick, J., Hemingway, J., Delves, M. J., Sinden, R. E., Zeeman, A. M., Kocken, C. H., Berry, N. G., O'neill, P. M. & Ward, S. A. (2012) Generation of quinolone antimalarials targeting the *Plasmodium falciparum* mitochondrial respiratory chain for the treatment and prophylaxis of malaria. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 8298-303.
- Bitonti, A. J., Kelly, S. E. & Mccann, P. P. (1984) Characterization of spermidine synthase from *Trypanosoma brucei brucei*. *Molecular and Biochemical Parasitology*, **13**, 21-8.
- Black, S. J., Sendashonga, C. N., O'brien, C., Borowy, N. K., Naessens, M., Webster, P. & Murray, M. (1985) Regulation of parasitaemia in mice infected with *Trypanosoma brucei*. *Current Topics in Microbiology and Immunology*, **117**, 93-118.
- Blum, J., Schmid, C. & Burri, C. (2006) Clinical aspects of 2541 patients with second stage human African trypanosomiasis. *Acta Tropica*, **97**, 55-64.
- Boothroyd, C. E., Dreesen, O., Leonova, T., Ly, K. I., Figueiredo, L. M., Cross, G. A. & Papavasiliou, F. N. (2009) A yeast-endonuclease-generated DNA break induces antigenic switching in *Trypanosoma brucei*. *Nature*, **459**, 278-81.
- Boothroyd, J. C. & Cross, G. A. (1982) Transcripts coding for variant surface glycoproteins of *Trypanosoma brucei* have a short, identical exon at their 5' end. *Gene*, **20**, 281-9.
- Borchardt, R. T., Eiden, L. E., Wu, B. & Rutledge, C. O. (1979) Sinefungin, a potent inhibitor of S-adenosylmethionine: protein O-methyltransferase. *Biochemical and Biophysical Research Communications*, **89**, 919-24.
- Braun, S. D., Hofmann, J., Wensing, A., Ullrich, M. S., Weingart, H., Volksch, B. & Spiteller, D. (2010) Identification of the biosynthetic gene cluster for 3-methylarginine, a toxin produced by *Pseudomonas syringae* pv. *syringae* 22d/93. *Applied and Environmental Microbiology*, **76**, 2500-8.
- Breidbach, T., Ngazoa, E. & Steverding, D. (2002) *Trypanosoma brucei*: in vitro slender-to-stumpy differentiation of culture-adapted, monomorphic bloodstream forms. *Experimental Parasitology*, **101**, 223-30.
- Brems, S., Guilbride, D. L., Gundlesdodjir-Planck, D., Busold, C., Luu, V. D., Schanne, M., Hoheisel, J. & Clayton, C. (2005) The transcriptomes of *Trypanosoma brucei* Lister 427 and TREU927 bloodstream and procyclic trypomastigotes. *Molecular and Biochemical Parasitology*, **139**, 163-72.
- Bridges, D. J., Gould, M. K., Nerima, B., Maser, P., Burchmore, R. J. & De Koning, H. P. (2007) Loss of the high-affinity pentamidine transporter is responsible for high levels of cross-resistance between arsenical and diamidine drugs in African trypanosomes. *Molecular Pharmacology*, **71**, 1098-108.
- Bringaud, F. & Baltz, T. (1993) Differential regulation of two distinct families of glucose transporter genes in *Trypanosoma brucei*. *Molecular and Cellular Biology*, **13**, 1146-54.
- Bringaud, F., Barrett, M. P. & Zilberstein, D. (2012) Multiple roles of proline transport and metabolism in trypanosomatids. *Frontiers in Bioscience*, **17**, 349-74.
- Bringaud, F., Biran, M., Millerioux, Y., Wargnies, M., Allmann, S. & Mazet, M. (2015) Combining reverse genetics and nuclear magnetic resonance-based metabolomics unravels trypanosome-specific metabolic pathways. *Molecular Microbiology*, **96**, 917-26.

- Bringaud, F., Peyruchaud, S., Baltz, D., Giroud, C., Simpson, L. & Baltz, T. (1995) Molecular characterization of the mitochondrial heat shock protein 60 gene from *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, **74**, 119-23.
- Bringaud, F., Riviere, L. & Coustou, V. (2006) Energy metabolism of trypanosomatids: adaptation to available carbon sources. *Molecular and Biochemical Parasitology*, **149**, 1-9.
- Brun, R., Blum, J., Chappuis, F. & Burri, C. (2010) Human African trypanosomiasis. *Lancet*, **375**, 148-59.
- Brun, R., Don, R., Jacobs, R. T., Wang, M. Z. & Barrett, M. P. (2011) Development of novel drugs for human African trypanosomiasis. *Future Microbiology*, **6**, 677-91.
- Buckner, F. S., Kateete, D. P., Lubega, G. W., Van Voorhis, W. C. & Yokoyama, K. (2002) *Trypanosoma brucei* prenylated-protein carboxyl methyltransferase prefers farnesylated substrates. *The Biochemical Journal*, **367**, 809-16.
- Bustin, S. A., Benes, V., Garson, J. A., Hellems, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J. & Wittwer, C. T. (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, **55**, 611-22.
- Butikofer, P., Ruepp, S., Boschung, M. & Roditi, I. (1997) 'GPEET' procyclin is the major surface protein of procyclic culture forms of *Trypanosoma brucei brucei* strain 427. *The Biochemical Journal*, **326**, 415-23.
- Buxbaum, L. U., Milne, K. G., Werbovetz, K. A. & Englund, P. T. (1996) Myristate exchange on the *Trypanosoma brucei* variant surface glycoprotein. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 1178-83.
- Caligiuri, M., Molz, L., Liu, Q., Kaplan, F., Xu, J. P., Majeti, J. Z., Ramos-Kelsey, R., Murthi, K., Lievens, S., Tavernier, J. & Kley, N. (2006) MASPIT: three-hybrid trap for quantitative proteome fingerprinting of small molecule-protein interactions in mammalian cells. *Chemistry and Biology*, **13**, 711-22.
- Campbell, G. H., Esser, K. M. & Weinbaum, F. I. (1977) *Trypanosoma rhodesiense* infection in B-cell-deficient mice. *Infection and Immunity*, **18**, 434-8.
- Campos, M. C., Castro-Pinto, D. B., Ribeiro, G. A., Berredo-Pinho, M. M., Gomes, L. H., Da Silva Bellieny, M. S., Goulart, C. M., Echevarria, A. & Leon, L. L. (2013) P-glycoprotein efflux pump plays an important role in *Trypanosoma cruzi* drug resistance. *Parasitology Research*, **112**, 2341-51.
- Canuto, G. A., Castilho-Martins, E. A., Tavares, M., Lopez-Gonzalvez, A., Rivas, L. & Barbas, C. (2012) CE-ESI-MS metabolic fingerprinting of *Leishmania* resistance to antimony treatment. *Electrophoresis*, **33**, 1901-10.
- Canuto, G. A., Castilho-Martins, E. A., Tavares, M. F., Rivas, L., Barbas, C. & Lopez-Gonzalvez, A. (2014) Multi-analytical platform metabolomic approach to study miltefosine mechanism of action and resistance in *Leishmania*. *Analytical and Bioanalytical Chemistry*, **406**, 3459-76.
- Capewell, P., Clucas, C., Dejesus, E., Kieft, R., Hajduk, S., Veitch, N., Steketee, P. C., Cooper, A., Weir, W. & Macleod, A. (2013a) The TgsGP gene is essential for resistance to human serum in *Trypanosoma brucei gambiense*. *PLoS Pathogens*, **9**, e1003686.
- Capewell, P., Cooper, A., Clucas, C., Weir, W. & Macleod, A. (2015a) A co-evolutionary arms race: trypanosomes shaping the human genome, humans shaping the trypanosome genome. *Parasitology*, **142**, S108-19.
- Capewell, P., Cooper, A., Clucas, C., Weir, W., Vaikkinen, H., Morrison, L., Tait, A. & Macleod, A. (2015b) Exploiting genetic variation to discover genes involved in important disease phenotypes. *Methods in Molecular Biology*, **1201**, 91-107.
- Capewell, P., Monk, S., Ivens, A., Macgregor, P., Fenn, K., Walrad, P., Bringaud, F., Smith, T. K. & Matthews, K. (2013b) Regulation of *Trypanosoma brucei* Total and Polysomal mRNA during Development within Its Mammalian Host. *PLoS One*, **8**, e67069.
- Capewell, P., Veitch, N. J., Turner, C. M., Raper, J., Berriman, M., Hajduk, S. L. & Macleod, A. (2011) Differences between *Trypanosoma brucei gambiense* groups 1 and 2 in their resistance to killing by trypanolytic factor 1. *PLoS Neglected Tropical Diseases*, **5**, e1287.

- Carnes, J., Anupama, A., Balmer, O., Jackson, A., Lewis, M., Brown, R., Cestari, I., Desquesnes, M., Gendrin, C., Hertz-Fowler, C., Imamura, H., Ivens, A., Koreny, L., Lai, D. H., Macleod, A., McDermott, S. M., Merritt, C., Monnerat, S., Moon, W., Myler, P., Phan, I., Ramasamy, G., Sivam, D., Lun, Z. R., Lukes, J., Stuart, K. & Schnauffer, A. (2015) Genome and phylogenetic analyses of *Trypanosoma evansi* reveal extensive similarity to *T. brucei* and multiple independent origins for dyskinetoplasty. *PLoS Neglected Tropical Diseases*, **9**, e3404.
- Carpenter, L. R. & Englund, P. T. (1995) Kinetoplast maxicircle DNA replication in *Crithidia fasciculata* and *Trypanosoma brucei*. *Molecular and Cellular Biology*, **15**, 6794-803.
- Carreno-Quintero, N., Bouwmeester, H. J. & Keurentjes, J. J. (2013) Genetic analysis of metabolome-phenotype interactions: from model to crop species. *Trends in Genetics*, **29**, 41-50.
- Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C. A., Keseler, I. M., Kothari, A., Krummenacker, M., Latendresse, M., Mueller, L. A., Ong, Q., Paley, S., Subhraveti, P., Weaver, D. S. & Karp, P. D. (2015) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Research*.
- Caspi, R., Foerster, H., Fulcher, C. A., Hopkinson, R., Ingraham, J., Kaipa, P., Krummenacker, M., Paley, S., Pick, J., Rhee, S. Y., Tissier, C., Zhang, P. & Karp, P. D. (2006) MetaCyc: a multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Research*, **34**, D511-6.
- Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M. Y., Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S. L., Nuwaysir, L. M., Lefebvre, B., Kuhlmann, F., Roark, J., Rainer, P., Detlev, S., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B., Chadick, T., Holly, K., Eckels, J., Deutsch, E. W., Moritz, R. L., Katz, J. E., Agus, D. B., Maccoss, M., Tabb, D. L. & Mallick, P. (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, **30**, 918-20.
- Chanez, A. L., Hehl, A. B., Engstler, M. & Schneider, A. (2006) Ablation of the single dynamin of *T. brucei* blocks mitochondrial fission and endocytosis and leads to a precise cytokinesis arrest. *Journal of Cell Science*, **119**, 2968-74.
- Chappuis, F., Loutan, L., Simarro, P., Lejon, V. & Buscher, P. (2005) Options for field diagnosis of human african trypanosomiasis. *Clinical Microbiology Reviews*, **18**, 133-46.
- Chaudhuri, M., Ajayi, W., Temple, S. & Hill, G. C. (1995) Identification and partial purification of a stage-specific 33 kDa mitochondrial protein as the alternative oxidase of the *Trypanosoma brucei brucei* bloodstream trypomastigotes. *The Journal of Eukaryotic Microbiology*, **42**, 467-72.
- Chaudhuri, M., Ott, R. D. & Hill, G. C. (2006) Trypanosome alternative oxidase: from molecule to function. *Trends in Parasitology*, **22**, 484-91.
- Chin, R. M., Fu, X., Pai, M. Y., Vergnes, L., Hwang, H., Deng, G., Diep, S., Lomenick, B., Meli, V. S., Monsalve, G. C., Hu, E., Whelan, S. A., Wang, J. X., Jung, G., Solis, G. M., Fazlollahi, F., Kaweeteerawat, C., Quach, A., Nili, M., Krall, A. S., Godwin, H. A., Chang, H. R., Faull, K. F., Guo, F., Jiang, M., Trauger, S. A., Saghatelian, A., Braas, D., Christofk, H. R., Clarke, C. F., Teitell, M. A., Petrascheck, M., Reue, K., Jung, M. E., Frand, A. R. & Huang, J. (2014) The metabolite alpha-ketoglutarate extends lifespan by inhibiting ATP synthase and TOR. *Nature*, **510**, 397-401.
- Chokkathukalam, A., Jankevics, A., Creek, D. J., Achcar, F., Barrett, M. P. & Breitling, R. (2013) mzMatch-ISO: an R tool for the annotation and relative quantification of isotope-labelled mass spectrometry data. *Bioinformatics*, **29**, 281-3.
- Chokkathukalam, A., Kim, D. H., Barrett, M. P., Breitling, R. & Creek, D. J. (2014) Stable isotope-labeling studies in metabolomics: new insights into structure and dynamics of metabolic networks. *Bioanalysis*, **6**, 511-524.
- Chrebet, G. L., Wisniewski, D., Perkins, A. L., Deng, Q., Kurtz, M. B., Marcy, A. & Parent, S. A. (2005) Cell-based assays to detect inhibitors of fungal mRNA capping enzymes and

- characterization of sinefungin as a cap methyltransferase inhibitor. *Journal of Biomolecular Screening*, **10**, 355-64.
- Cingolani, P., Platts, A., Wang Le, L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X. & Ruden, D. M. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, **6**, 80-92.
- Clayton, A. M., Guler, J. L., Povelones, M. L., Gluenz, E., Gull, K., Smith, T. K., Jensen, R. E. & Englund, P. T. (2011) Depletion of mitochondrial acyl carrier protein in bloodstream-form *Trypanosoma brucei* causes a kinetoplast segregation defect. *Eukaryotic Cell*, **10**, 286-92.
- 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. & Et Al. (1998) Genetic nomenclature for *Trypanosoma* and *Leishmania*. *Molecular and Biochemical Parasitology*, **97**, 221-4.
- Clayton, C. E. (1987) Import of fructose bisphosphate aldolase into the glycosomes of *Trypanosoma brucei*. *The Journal of Cell Biology*, **105**, 2649-54.
- Clayton, C. E. (1988) Most proteins, including fructose bisphosphate aldolase, are stable in the procyclic trypomastigote form of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, **28**, 43-6.
- Clayton, C. E. (2014) Networks of gene expression regulation in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, **195**, 96-106.
- Cobbold, S. A., Chua, H. H., Nijagal, B., Creek, D. J., Ralph, S. A. & Mcconville, M. J. (2015) Metabolic Dysregulation Induced in *Plasmodium falciparum* by Dihydroartemisinin and Other Front-Line Antimalarial Drugs. *The Journal of infectious diseases*.
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. & Zhang, F. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819-23.
- Cooper, A. 2009. *Linkage mapping and genetic analysis of Trypanosoma brucei*. Doctor of Philosophy, University of Glasgow.
- Cooper, A., Tait, A., Sweeney, L., Tweedie, A., Morrison, L., Turner, C. M. & Macleod, A. (2008) Genetic analysis of the human infective trypanosome *Trypanosoma brucei gambiense*: chromosomal segregation, crossing over, and the construction of a genetic map. *Genome Biology*, **9**, R103.
- Coulter, C., Hamilton, J. T. & Harper, D. B. (1993) Evidence for the existence of independent chloromethane- and S-adenosylmethionine-utilizing systems for methylation in *Phanerochaete chrysosporium*. *Applied and Environmental Microbiology*, **59**, 1461-6.
- Coustou, V., Besteiro, S., Biran, M., Dioloz, P., Bouchaud, V., Voisin, P., Michels, P. A., Canioni, P., Baltz, T. & Bringaud, F. (2003) ATP generation in the *Trypanosoma brucei* procyclic form: cytosolic substrate level is essential, but not oxidative phosphorylation. *The Journal of Biological Chemistry*, **278**, 49625-35.
- Coustou, V., Biran, M., Breton, M., Guegan, F., Riviere, L., Plazolles, N., Nolan, D., Barrett, M. P., Franconi, J. M. & Bringaud, F. (2008) Glucose-induced remodeling of intermediary and energy metabolism in procyclic *Trypanosoma brucei*. *The Journal of Biological Chemistry*, **283**, 16342-54.
- Couture, J. F., Collazo, E., Hauk, G. & Trievel, R. C. (2006) Structural basis for the methylation site specificity of SET7/9. *Nature Structural & Molecular Biology*, **13**, 140-6.
- Cox, F. E. (2002) History of human parasitology. *Clinical Microbiology Reviews*, **15**, 595-612.
- Creek, D. J., Anderson, J., Mcconville, M. J. & Barrett, M. P. (2012a) Metabolomic analysis of trypanosomatid protozoa. *Molecular and Biochemical Parasitology*, **181**, 73-84.
- Creek, D. J. & Barrett, M. P. (2014) Determination of antiprotozoal drug mechanisms by metabolomics approaches. *Parasitology*, **141**, 83-92.
- Creek, D. J., Jankevics, A., Burgess, K. E. V., Breitling, R. & Barrett, M. P. (2012b) IDEOM: an Excel interface for analysis of LC-MS-based metabolomics data. *Bioinformatics*, **28**, 1048-1049.

- Creek, D. J., Mazet, M., Achcar, F., Anderson, J., Kim, D. H., Kamour, R., Morand, P., Millerioux, Y., Biran, M., Kerkhoven, E. J., Chokkathukalam, A., Weidt, S. K., Burgess, K. E., Breitling, R., Watson, D. G., Bringaud, F. & Barrett, M. P. (2015) Probing the metabolic network in bloodstream-form *Trypanosoma brucei* using untargeted metabolomics with stable isotope labelled glucose. *PLoS Pathogens*, **11**, e1004689.
- Creek, D. J., Nijagal, B., Kim, D. H., Rojas, F., Matthews, K. R. & Barrett, M. P. (2013) Metabolomics guides rational development of a simplified cell culture medium for drug screening against *Trypanosoma brucei*. *Antimicrobial Agents and Chemotherapy*, **57**, 2768-79.
- Crossman, A., Jr., Paterson, M. J., Ferguson, M. A., Smith, T. K. & Brimacombe, J. S. (2002) Further probing of the substrate specificities and inhibition of enzymes involved at an early stage of glycosylphosphatidylinositol (GPI) biosynthesis. *Carbohydrate Research*, **337**, 2049-59.
- D'haene, B., Vandesompele, J. & Hellems, J. (2010) Accurate and objective copy number profiling using real-time quantitative PCR. *Methods*, **50**, 262-270.
- Danhauser, K., Sauer, S. W., Haack, T. B., Wieland, T., Staufner, C., Graf, E., Zschocke, J., Strom, T. M., Traub, T., Okun, J. G., Meitinger, T., Hoffmann, G. F., Prokisch, H. & Kolker, S. (2012) DHTKD1 mutations cause 2-amino adipic and 2-oxoadipic aciduria. *American Journal of Human Genetics*, **91**, 1082-7.
- De Gaudenzi, J. G., Noe, G., Campo, V. A., Frasch, A. C. & Cassola, A. (2011) Gene expression regulation in trypanosomatids. *Essays in Biochemistry*, **51**, 31-46.
- De Gee, A. L., Mccann, P. P. & Mansfield, J. M. (1983) Role of antibody in the elimination of trypanosomes after DL-alpha-difluoromethylornithine chemotherapy. *The Journal of Parasitology*, **69**, 818-22.
- De Greef, C. & Hamers, R. (1994) The serum resistance-associated (SRA) gene of *Trypanosoma brucei rhodesiense* encodes a variant surface glycoprotein-like protein. *Molecular and Biochemical Parasitology*, **68**, 277-84.
- De Koning, H. P. (2001) Uptake of pentamidine in *Trypanosoma brucei brucei* is mediated by three distinct transporters: implications for cross-resistance with arsenicals. *Molecular Pharmacology*, **59**, 586-92.
- De Koning, H. P., Anderson, L. F., Stewart, M., Burchmore, R. J., Wallace, L. J. & Barrett, M. P. (2004) The trypanocide diminazene aceturate is accumulated predominantly through the TbAT1 purine transporter: additional insights on diamidine resistance in african trypanosomes. *Antimicrobial Agents and Chemotherapy*, **48**, 1515-9.
- De Koning, H. P. & Jarvis, S. M. (2001) Uptake of pentamidine in *Trypanosoma brucei brucei* is mediated by the P2 adenosine transporter and at least one novel, unrelated transporter. *Acta Tropica*, **80**, 245-50.
- Dean, S., Marchetti, R., Kirk, K. & Matthews, K. R. (2009) A surface transporter family conveys the trypanosome differentiation signal. *Nature*, **459**, 213-7.
- Denise, H., Giroud, C., Barrett, M. P. & Baltz, T. (1999) Affinity chromatography using trypanocidal arsenical drugs identifies a specific interaction between glycerol-3-phosphate dehydrogenase from *Trypanosoma brucei* and Cymelarsan. *European Journal of Biochemistry*, **259**, 339-46.
- Deschamps, J. R. (2005) The role of crystallography in drug design. *AAPS Journal*, **7**, E813-9.
- Desquesnes, M., Holzmüller, P., Lai, D. H., Dargantes, A., Lun, Z. R. & Jittaplapong, S. (2013) *Trypanosoma evansi* and *surra*: a review and perspectives on origin, history, distribution, taxonomy, morphology, hosts, and pathogenic effects. *Biomedical Research International*, **2013**, 194176.
- Devkota, K., Lohse, B., Liu, Q., Wang, M. W., Staerk, D., Berthelsen, J. & Clausen, R. P. (2014) Analogues of the Natural Product Sinefungin as Inhibitors of EHMT1 and EHMT2. *ACS Medicinal Chemistry Letters*, **5**, 293-7.
- Ding, D., Zhao, Y., Meng, Q., Xie, D., Nare, B., Chen, D., Bacchi, C. J., Yarlett, N., Zhang, Y. K., Hernandez, V., Xia, Y., Freund, Y., Abdulla, M., Ang, K. H., Ratnam, J., Mckerrow, J. H., Jacobs, R. T., Zhou, H. & Plattner, J. J. (2010) Discovery of novel benzoxaborole-based potent antitrypanosomal agents. *ACS Medicinal Chemistry Letters*, **1**, 165-9.

- Diogo, D., Bouchier, C., D'enfert, C. & Bounoux, M. E. (2009) Loss of heterozygosity in commensal isolates of the asexual diploid yeast *Candida albicans*. *Fungal Genetics and Biology*, **46**, 159-68.
- Docampo, R., De Souza, W., Miranda, K., Rohloff, P. & Moreno, S. N. (2005) Acidocalcisomes - conserved from bacteria to man. *Nature Reviews Microbiology*, **3**, 251-61.
- Doll, A. & Grzeschik, K. H. (2001) Characterization of two novel genes, WBSCR20 and WBSCR22, deleted in Williams-Beuren syndrome. *Cytogenetics and Cell Genetics*, **95**, 20-7.
- Dong, C., Sexton, H., Gertrudes, A., Akama, T., Martin, S., Virtucio, C., Chen, C. W., Fan, X., Wu, A., Bu, W., Liu, L., Feng, L., Jarnagin, K. & Freund, Y. R. (2013) Inhibition of Toll-like receptor-mediated inflammation in vitro and in vivo by a novel benzoxaborole. *The Journal of Pharmacology and Experimental Therapeutics*, **344**, 436-46.
- Douris, V., Telford, M. J. & Averof, M. (2010) Evidence for multiple independent origins of trans-splicing in Metazoa. *Molecular Biology and Evolution*, **27**, 684-93.
- Drain, J., Bishop, J. R. & Hajduk, S. L. (2001) Haptoglobin-related protein mediates trypanosome lytic factor binding to trypanosomes. *The Journal of Biological Chemistry*, **276**, 30254-60.
- Droll, D., Minia, I., Fadda, A., Singh, A., Stewart, M., Queiroz, R. & Clayton, C. (2013) Post-transcriptional regulation of the trypanosome heat shock response by a zinc finger protein. *PLoS Pathogens*, **9**, e1003286.
- Dukes, P., Gibson, W. C., Gashumba, J. K., Hudson, K. M., Bromidge, T. J., Kaukus, A., Asonganyi, T. & Magnus, E. (1992) Absence of the LiTat 1.3 (CATT antigen) gene in *Trypanosoma brucei* gambiense stocks from Cameroon. *Acta Tropica*, **51**, 123-34.
- Dunn, W. B., Erban, A., Weber, R. J. M., Creek, D. J., Brown, M., Breitling, R., Hankemeier, T., Goodacre, R., Neumann, S., Kopka, J. & Viant, M. R. (2013) Mass appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics. *Metabolomics*, **9**, S44-S66.
- Dutton, J. E. (1902) Preliminary note upon a trypanosome occurring in the blood of man. *Thompson Yates Laboratory Reports*, **4**, 455-468.
- Ebikeme, C., Hubert, J., Biran, M., Gouspillou, G., Morand, P., Plazolles, N., Guegan, F., Diolez, P., Franconi, J. M., Portais, J. C. & Bringaud, F. (2010) Ablation of succinate production from glucose metabolism in the procyclic trypanosomes induces metabolic switches to the glycerol 3-phosphate/dihydroxyacetone phosphate shuttle and to proline metabolism. *The Journal of Biological Chemistry*, **285**, 32312-24.
- Ebikeme, C. E., Peacock, L., Coustou, V., Riviere, L., Bringaud, F., Gibson, W. C. & Barrett, M. P. (2008) N-acetyl D-glucosamine stimulates growth in procyclic forms of *Trypanosoma brucei* by inducing a metabolic shift. *Parasitology*, **135**, 585-94.
- Eisen, J. A. (2012) Badomics words and the power and peril of the ome-meme. *Gigascience*, **1**, 6.
- Eknoyan, G. (1999) Santorio Sanctorius(1561-1636) - Founding father of metabolic balance studies. *American Journal of Nephrology*, **19**, 226-233.
- Eknoyan, G. & Nagy, J. (2005) A history of diabetes mellitus or how a disease of the kidneys evolved into a kidney disease. *Advances in Chronic Kidney Disease*, **12**, 223-229.
- El-Sayed, N. M., Myler, P. J., Bartholomeu, D. C., Nilsson, D., Aggarwal, G., Tran, A. N., Ghedin, E., Worthey, E. A., Delcher, A. L., Blandin, G., Westenberger, S. J., Caler, E., Cerqueira, G. C., Branche, C., Haas, B., Anupama, A., Arner, E., Aslund, L., Attipoe, P., Bontempi, E., Bringaud, F., Burton, P., Cadag, E., Campbell, D. A., Carrington, M., Crabtree, J., Darban, H., Da Silveira, J. F., De Jong, P., Edwards, K., Englund, P. T., Fazelina, G., Feldblyum, T., Ferella, M., Frasch, A. C., Gull, K., Horn, D., Hou, L., Huang, Y., Kindlund, E., Klingbeil, M., Kluge, S., Koo, H., Lacerda, D., Levin, M. J., Lorenzi, H., Louie, T., Machado, C. R., McCulloch, R., McKenna, A., Mizuno, Y., Mottram, J. C., Nelson, S., Ochaya, S., Osoegawa, K., Pai, G., Parsons, M., Pentony, M., Pettersson, U., Pop, M., Ramirez, J. L., Rinta, J., Robertson, L., Salzberg, S. L., Sanchez, D. O., Seyler, A., Sharma, R., Shetty, J., Simpson, A. J., Sisk, E., Tammi, M. T., Tarleton, R., Teixeira, S., Van Aken, S., Vogt, C., Ward, P. N., Wickstead, B., Wortman, J., White, O., Fraser, C. M., Stuart, K. D. & Andersson, B. (2005a) The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science*, **309**, 409-15.

- 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. & Hall, N. (2005b) Comparative genomics of trypanosomatid parasitic protozoa. *Science*, **309**, 404-9.
- Emmer, B. T., Daniels, M. D., Taylor, J. M., Epting, C. L. & Engman, D. M. (2010) Calflagin inhibition prolongs host survival and suppresses parasitemia in *Trypanosoma brucei* infection. *Eukaryotic Cell*, **9**, 934-42.
- Fairlamb, A. H., Blackburn, P., Ulrich, P., Chait, B. T. & Cerami, A. (1985) Trypanothione: a novel bis(glutathionyl)spermidine cofactor for glutathione reductase in trypanosomatids. *Science*, **227**, 1485-7.
- Fairlamb, A. H. & Bowman, I. B. (1977) *Trypanosoma brucei*: suramin and other trypanocidal compounds' effects on sn-glycerol-3-phosphate oxidase. *Experimental Parasitology*, **43**, 353-61.
- Fairlamb, A. H., Henderson, G. B. & Cerami, A. (1989) Trypanothione is the primary target for arsenical drugs against African trypanosomes. *Proceedings of the National Academy of Sciences of the United States of America*, **86**, 2607-11.
- Fairlamb, A. H., Smith, K. & Hunter, K. J. (1992) The interaction of arsenical drugs with dihydrolipoamide and dihydrolipoamide dehydrogenase from arsenical resistant and sensitive strains of *Trypanosoma brucei brucei*. *Molecular and Biochemical Parasitology*, **53**, 223-31.
- Fang, J. & Beattie, D. S. (2003) Alternative oxidase present in procyclic *Trypanosoma brucei* may act to lower the mitochondrial production of superoxide. *Archives of Biochemistry and Biophysics*, **414**, 294-302.
- Farine, L., Niemann, M., Schneider, A. & Butikofer, P. (2015) Phosphatidylethanolamine and phosphatidylcholine biosynthesis by the Kennedy pathway occurs at different sites in *Trypanosoma brucei*. *Scientific Reports*, **5**, 16787.
- FDA (2014). Kerydin Tavaborole topical solution, Initial U.S. approval.
- Feagin, J. E., Abraham, J. M. & Stuart, K. (1988) Extensive editing of the cytochrome c oxidase III transcript in *Trypanosoma brucei*. *Cell*, **53**, 413-22.
- Fenn, K. & Matthews, K. R. (2007) The cell biology of *Trypanosoma brucei* differentiation. *Current Opinion in Microbiology*, **10**, 539-46.
- Ferguson, M. A., Homans, S. W., Dwek, R. A. & Rademacher, T. W. (1988) Glycosyl-phosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane. *Science*, **239**, 753-9.
- Ferguson, M. L., Torri, A. F., Perez-Morga, D., Ward, D. C. & Englund, P. T. (1994) Kinetoplast DNA replication: mechanistic differences between *Trypanosoma brucei* and *Crithidia fasciculata*. *The Journal of Cell Biology*, **126**, 631-9.
- Fernandez-Moya, S. M., Carrington, M. & Estevez, A. M. (2014) A short RNA stem-loop is necessary and sufficient for repression of gene expression during early logarithmic phase in trypanosomes. *Nucleic Acids Research*, **42**, 7201-9.
- Ferrante, A., Ljungstrom, I., Huldt, G. & Lederer, E. (1984) Amoebicidal activity of the antifungal antibiotic sinefungin against *Entamoeba histolytica*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **78**, 837-8.
- Ferrante, A., Ljungstrom, I. & Lederer, E. (1988) Antitoxoplasmosis properties of sinefungin in mice. *Comptes rendus de l'Academie des sciences. Serie III, Sciences de la vie*, **306**, 109-13.
- Field, M. C. & Carrington, M. (2009) The trypanosome flagellar pocket. *Nature Reviews Microbiology*, **7**, 775-86.
- Figueiredo, L. M., Janzen, C. J. & Cross, G. A. (2008) A histone methyltransferase modulates antigenic variation in African trypanosomes. *PLoS Biology*, **6**, e161.

- Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L., Tate, J. & Punta, M. (2014) Pfam: the protein families database. *Nucleic Acids Research*, **42**, D222-30.
- Fisk, J. C., Sayegh, J., Zurita-Lopez, C., Menon, S., Presnyak, V., Clarke, S. G. & Read, L. K. (2009) A type III protein arginine methyltransferase from the protozoan parasite *Trypanosoma brucei*. *The Journal of Biological Chemistry*, **284**, 11590-600.
- Fisk, J. C., Zurita-Lopez, C., Sayegh, J., Tomasello, D. L., Clarke, S. G. & Read, L. K. (2010) TbPRMT6 is a type I protein arginine methyltransferase that contributes to cytokinesis in *Trypanosoma brucei*. *Eukaryotic Cell*, **9**, 866-77.
- Fivelman, Q. L., Adagu, I. S. & Warhurst, D. C. (2004) Modified fixed-ratio isobologram method for studying in vitro interactions between atovaquone and proguanil or dihydroartemisinin against drug-resistant strains of *Plasmodium falciparum*. *Antimicrobial Agents and Chemotherapy*, **48**, 4097-102.
- Fly, R., Lloyd, J., Krueger, S., Fernie, A. & Van Der Merwe, M. J. (2015) Improvements to Define Mitochondrial Metabolomics Using Nonaqueous Fractionation. *Methods in Molecular Biology*, **1305**, 197-210.
- Fontecave, M., Atta, M. & Mulliez, E. (2004) S-adenosylmethionine: nothing goes to waste. *Trends in Biochemical Sciences*, **29**, 243-9.
- Forche, A., Abbey, D., Pisithkul, T., Weinzierl, M. A., Ringstrom, T., Bruck, D., Petersen, K. & Berman, J. (2011) Stress alters rates and types of loss of heterozygosity in *Candida albicans*. *MBio*, **2**.
- Forsburg, S. L. (2004) Eukaryotic MCM proteins: beyond replication initiation. *Microbiology and Molecular Biology Reviews*, **68**, 109-31.
- Franco, J. R., Simarro, P. P., Diarra, A. & Jannin, J. G. (2014) Epidemiology of human African trypanosomiasis. *Clinical Epidemiology*, **6**, 257-75.
- Frederiks, F., Van Welsem, T., Oudgenoeg, G., Heck, A. J., Janzen, C. J. & Van Leeuwen, F. (2010) Heterologous expression reveals distinct enzymatic activities of two DOT1 histone methyltransferases of *Trypanosoma brucei*. *Journal of Cell Science*, **123**, 4019-23.
- Fridberg, A., Olson, C. L., Nakayasu, E. S., Tyler, K. M., Almeida, I. C. & Engman, D. M. (2008) Sphingolipid synthesis is necessary for kinetoplast segregation and cytokinesis in *Trypanosoma brucei*. *Journal of Cell Science*, **121**, 522-35.
- Friedheim, E. A. (1949) Mel B in the treatment of human trypanosomiasis. *The American Journal of Tropical Medicine and Hygiene*, **29**, 173-80.
- Funfschilling, U. & Rospert, S. (1999) Nascent polypeptide-associated complex stimulates protein import into yeast mitochondria. *Molecular Biology of the Cell*, **10**, 3289-99.
- Furuya, T., Kessler, P., Jardim, A., Schnauffer, A., Crudder, C. & Parsons, M. (2002) Glucose is toxic to glycosome-deficient trypanosomes. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 14177-82.
- Gao, W., Kim, J. Y., Anderson, J. R., Akopian, T., Hong, S., Jin, Y. Y., Kandror, O., Kim, J. W., Lee, I. A., Lee, S. Y., McAlpine, J. B., Mulugeta, S., Sunoqrot, S., Wang, Y., Yang, S. H., Yoon, T. M., Goldberg, A. L., Pauli, G. F., Suh, J. W., Franzblau, S. G. & Cho, S. (2015) The cyclic peptide ecumicin targeting ClpC1 is active against *Mycobacterium tuberculosis* in vivo. *Antimicrobial Agents and Chemotherapy*, **59**, 880-9.
- Gassen, A., Brechtefeld, D., Schandry, N., Arteaga-Salas, J. M., Israel, L., Imhof, A. & Janzen, C. J. (2012) DOT1A-dependent H3K76 methylation is required for replication regulation in *Trypanosoma brucei*. *Nucleic Acids Research*, **40**, 10302-11.
- Geerts, S., Holmes, P. H., Eisler, M. C. & Diall, O. (2001) African bovine trypanosomiasis: the problem of drug resistance. *Trends in Parasitology*, **17**, 25-8.
- Geigenberger, P., Tiessen, A. & Meurer, J. (2011) Use of non-aqueous fractionation and metabolomics to study chloroplast function in *Arabidopsis*. *Methods in Molecular Biology*, **775**, 135-60.
- George, D. M., Breinlinger, E. C., Friedman, M., Zhang, Y., Wang, J., Argiriadi, M., Bansal-Pakala, P., Barth, M., Duignan, D. B., Honore, P., Lang, Q., Mittelstadt, S., Potin, D., Rundell, L. &

- Edmunds, J. J. (2015) Discovery of selective and orally bioavailable protein kinase C $\theta$  (PKC $\theta$ ) inhibitors from a fragment hit. *Journal of Medicinal Chemistry*, **58**, 222-36.
- Ghosh, S. & Chan, C. K. (2016) Analysis of RNA-Seq Data Using TopHat and Cufflinks. *Methods in Molecular Biology*, **1374**, 339-61.
- Giaever, G., Shoemaker, D. D., Jones, T. W., Liang, H., Winzeler, E. A., Astromoff, A. & Davis, R. W. (1999) Genomic profiling of drug sensitivities via induced haploinsufficiency. *Nature Genetics*, **21**, 278-83.
- Giannotti, E., Longo, M., Messina, M., Ghiglieri, A. & Launay, D. (2014) SCYX-7158: Reproductive toxicity studies. *Toxicology Letters*, **229**, S90-S91.
- Gibellini, F., Hunter, W. N. & Smith, T. K. (2009) The ethanolamine branch of the Kennedy pathway is essential in the bloodstream form of *Trypanosoma brucei*. *Molecular Microbiology*, **73**, 826-43.
- Gibson, W. & Stevens, J. (1999) Genetic exchange in the trypanosomatidae. *Advances in Parasitology*, **43**, 1-46.
- Gilbert, R. & Klein, R. A. (1982) Carnitine stimulates ATP synthesis in *Trypanosoma brucei brucei*. *FEBS Letters*, **141**, 271-4.
- Goldshmidt, H., Matas, D., Kabi, A., Carmi, S., Hope, R. & Michaeli, S. (2010) Persistent ER stress induces the spliced leader RNA silencing pathway (SLS), leading to programmed cell death in *Trypanosoma brucei*. *PLoS Pathogens*, **6**, e1000731.
- Goldstein, E. J., Citron, D. M., Tyrrell, K. L. & Merriam, C. V. (2013) Comparative in vitro activities of GSK2251052, a novel boron-containing leucyl-tRNA synthetase inhibitor, against 916 anaerobic organisms. *Antimicrobial Agents and Chemotherapy*, **57**, 2401-4.
- Goulah, C. C., Pelletier, M. & Read, L. K. (2006) Arginine methylation regulates mitochondrial gene expression in *Trypanosoma brucei* through multiple effector proteins. *RNA*, **12**, 1545-55.
- Graf, F. E., Ludin, P., Wenzler, T., Kaiser, M., Brun, R., Pyana, P. P., Buscher, P., De Koning, H. P., Horn, D. & Maser, P. (2013) Aquaporin 2 mutations in *Trypanosoma brucei gambiense* field isolates correlate with decreased susceptibility to pentamidine and melarsoprol. *PLoS Neglected Tropical Diseases*, **7**, e2475.
- Grogan, D. W. & Cronan, J. E., Jr. (1997) Cyclopropane ring formation in membrane lipids of bacteria. *Microbiology and Molecular Biology Reviews*, **61**, 429-41.
- Grover, M. P., Ballouz, S., Mohanasundaram, K. A., George, R. A., Sherman, C. D., Crowley, T. M. & Wouters, M. A. (2014) Identification of novel therapeutics for complex diseases from genome-wide association data. *BMC Medical Genomics*, **7**, S8.
- Gubbels, M. J., Vaishnava, S., Boot, N., Dubremetz, J. F. & Striepen, B. (2006) A MORN-repeat protein is a dynamic component of the *Toxoplasma gondii* cell division apparatus. *Journal of Cell Science*, **119**, 2236-45.
- Guerra-Giraldez, C., Quijada, L. & Clayton, C. E. (2002) Compartmentation of enzymes in a microbody, the glycosome, is essential in *Trypanosoma brucei*. *Journal of Cell Science*, **115**, 2651-8.
- Gunasekera, K., Wuthrich, D., Braga-Lagache, S., Heller, M. & Ochsenreiter, T. (2012) Proteome remodelling during development from blood to insect-form *Trypanosoma brucei* quantified by SILAC and mass spectrometry. *BMC Genomics*, **13**, 556.
- Gunzl, A. (2010) The pre-mRNA splicing machinery of trypanosomes: complex or simplified? *Eukaryotic Cell*, **9**, 1159-70.
- Guther, M. L., Urbaniak, M. D., Tavendale, A., Prescott, A. & Ferguson, M. A. (2014) High-confidence glycosome proteome for procyclic form *Trypanosoma brucei* by epitope-tag organelle enrichment and SILAC proteomics. *Journal of Proteome Research*, **13**, 2796-806.
- Haag, J., O'hugin, C. & Overath, P. (1998) The molecular phylogeny of trypanosomes: evidence for an early divergence of the Salivaria. *Molecular and Biochemical Parasitology*, **91**, 37-49.
- Haanstra, J. R. & Bakker, B. M. (2015) Drug target identification through systems biology. *Drug discovery today*, **15**, 17-22.
- Haanstra, J. R., Gonzalez-Marcano, E. B., Gualdron-Lopez, M. & Michels, P. A. (2015) Biogenesis, maintenance and dynamics of glycosomes in trypanosomatid parasites. *Biochimica et Biophysica Acta*.

- Haanstra, J. R., Kerkhoven, E. J., Van Tuijl, A., Blits, M., Wurst, M., Van Nuland, R., Albert, M. A., Michels, P. A., Bouwman, J., Clayton, C., Westerhoff, H. V. & Bakker, B. M. (2011) A domino effect in drug action: from metabolic assault towards parasite differentiation. *Molecular Microbiology*, **79**, 94-108.
- Hall, J. P., Wang, H. & Barry, J. D. (2013) Mosaic VSGs and the scale of *Trypanosoma brucei* antigenic variation. *PLoS Pathogens*, **9**, e1003502.
- Hall, M. J., Middleton, R. F. & Westmacott, D. (1983) The fractional inhibitory concentration (FIC) index as a measure of synergy. *The Journal of antimicrobial chemotherapy*, **11**, 427-33.
- Hall, M. P. & Ho, C. K. (2006a) Characterization of a *Trypanosoma brucei* RNA cap (guanine N-7) methyltransferase. *RNA*, **12**, 488-97.
- Hall, M. P. & Ho, C. K. (2006b) Functional characterization of a 48 kDa *Trypanosoma brucei* cap 2 RNA methyltransferase. *Nucleic Acids Research*, **34**, 5594-602.
- Hamamoto, R. & Nakamura, Y. (2016) Dysregulation of protein methyltransferases in human cancer: an emerging target class for anticancer therapy. *Cancer Science*.
- Hamidou Soumana, I., Klopp, C., Ravel, S., Nabihoudine, I., Tchicaya, B., Parrinello, H., Abate, L., Rialle, S. & Geiger, A. (2015) RNA-seq de novo Assembly Reveals Differential Gene Expression in *Glossina palpalis gambiense* Infected with *Trypanosoma brucei gambiense* vs. Non-Infected and Self-Cured Flies. *Frontiers in Microbiology*, **6**, 1259.
- Hannaert, V., Albert, M. A., Rigden, D. J., Da Silva Giotto, M. T., Thiemann, O., Garratt, R. C., Van Roy, J., Opperdoes, F. R. & Michels, P. A. (2003) Kinetic characterization, structure modelling studies and crystallization of *Trypanosoma brucei* enolase. *European Journal of Biochemistry*, **270**, 3205-13.
- Haqqani, A. S., Kelly, J. F. & Stanimirovic, D. B. (2008) Quantitative protein profiling by mass spectrometry using label-free proteomics. *Methods in Molecular Biology*, **439**, 241-56.
- Hasne, M. P. & Barrett, M. P. (2000) Transport of methionine in *Trypanosoma brucei brucei*. *Molecular and Biochemical Parasitology*, **111**, 299-307.
- Heddergott, N., Kruger, T., Babu, S. B., Wei, A., Stellamanns, E., Uppaluri, S., Pfohl, T., Stark, H. & Engstler, M. (2012) Trypanosome motion represents an adaptation to the crowded environment of the vertebrate bloodstream. *PLoS Pathogens*, **8**, e1003023.
- Helliwell, C. A., Wesley, S. V., Wielopolska, A. J. & Waterhouse, P. M. (2002) High-throughput vectors for efficient gene silencing in plants. *Functional Plant Biology*, **29**, 1217-1225.
- Henderson, G. B., Yamaguchi, M., Novoa, L., Fairlamb, A. H. & Cerami, A. (1990) Biosynthesis of the trypanosomatid metabolite trypanothione: purification and characterization of trypanothione synthetase from *Crithidia fasciculata*. *Biochemistry*, **29**, 3924-9.
- Herman, M., Perez-Mora, D., Shtickzelle, N. & Michels, P. a. M. (2008) Turnover of glycosomes during life-cycle differentiation of *Trypanosoma brucei*. *Autophagy*, **4**, 294-308.
- Hernandez, V., Crepin, T., Palencia, A., Cusack, S., Akama, T., Baker, S. J., Bu, W., Feng, L., Freund, Y. R., Liu, L., Meewan, M., Mohan, M., Mao, W., Rock, F. L., Sexton, H., Sheoran, A., Zhang, Y., Zhang, Y. K., Zhou, Y., Nieman, J. A., Anugula, M. R., Keramane El, M., Savariraj, K., Reddy, D. S., Sharma, R., Subedi, R., Singh, R., O'leary, A., Simon, N. L., De Marsh, P. L., Mushtaq, S., Warner, M., Livermore, D. M., Alley, M. R. & Plattner, J. J. (2013) Discovery of a novel class of boron-based antibacterials with activity against gram-negative bacteria. *Antimicrobial Agents and Chemotherapy*, **57**, 1394-403.
- Herz, H. M., Garruss, A. & Shilatifard, A. (2013) SET for life: biochemical activities and biological functions of SET domain-containing proteins. *Trends in Biochemical Sciences*, **38**, 621-39.
- Hide, G., Cattand, P., Leray, D., Barry, J. D. & Tait, A. (1990) The identification of *Trypanosoma brucei* subspecies using repetitive DNA sequences. *Molecular and Biochemical Parasitology*, **39**, 213-25.
- Hill, C. B., Taylor, J. D., Edwards, J., Mather, D., Langridge, P., Bacic, A. & Roessner, U. (2015) Detection of QTL for metabolic and agronomic traits in wheat with adjustments for variation at genetic loci that affect plant phenology. *Plant Science*, **233**, 143-54.
- Hines, J. C. & Ray, D. S. (2010) A mitochondrial DNA primase is essential for cell growth and kinetoplast DNA replication in *Trypanosoma brucei*. *Molecular and Cellular Biology*, **30**, 1319-28.

- Hines, J. C. & Ray, D. S. (2011) A second mitochondrial DNA primase is essential for cell growth and kinetoplast minicircle DNA replication in *Trypanosoma brucei*. *Eukaryotic Cell*, **10**, 445-54.
- Hiraishi, H., Shimada, T., Ohtsu, I., Sato, T. A. & Takagi, H. (2009) The yeast ubiquitin ligase Rsp5 downregulates the alpha subunit of nascent polypeptide-associated complex Egd2 under stress conditions. *FEBS Journal*, **276**, 5287-97.
- Hirosawa, M., Hoshida, M., Ishikawa, M. & Toya, T. (1993) MASCOT: multiple alignment system for protein sequences based on three-way dynamic programming. *Computer Applications in the Biosciences*, **9**, 161-7.
- Hirth, B., Barker, R. H., Jr., Celatka, C. A., Klinger, J. D., Liu, H., Nare, B., Nijjar, A., Phillips, M. A., Sybertz, E., Willert, E. K. & Xiang, Y. (2009) Discovery of new S-adenosylmethionine decarboxylase inhibitors for the treatment of Human African Trypanosomiasis (HAT). *Bioorganic and Medicinal Chemistry Letters*, **19**, 2916-9.
- Hirumi, H. & Hirumi, K. (1989) Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *The Journal of Parasitology*, **75**, 985-9.
- Hoppel, C. L., Cox, R. A. & Novak, R. F. (1980) N6-Trimethyl-lysine metabolism. 3-Hydroxy-N6-trimethyl-lysine and carnitine biosynthesis. *The Biochemical journal*, **188**, 509-19.
- Horn, D. (2014) Antigenic variation in African trypanosomes. *Molecular and Biochemical Parasitology*, **195**, 123-9.
- Horn, D. & Duraisingh, M. T. (2014) Antiparasitic chemotherapy: from genomes to mechanisms. *Annual Review of Pharmacology and Toxicology*, **54**, 71-94.
- Hotz, R. L. (2012) Here's an omical tale: scientists discover spreading suffix. *The Wall Street Journal*.
- Huang, J. & Van Der Ploeg, L. H. (1991) Maturation of polycistronic pre-mRNA in *Trypanosoma brucei*: analysis of trans splicing and poly(A) addition at nascent RNA transcripts from the hsp70 locus. *Molecular and Cellular Biology*, **11**, 3180-90.
- Hughes, J. P., Rees, S., Kalindjian, S. B. & Philpott, K. L. (2011) Principles of early drug discovery. *British Journal of Pharmacology*, **162**, 1239-49.
- Hunter, R. G., Mccarthy, K. J., Milne, T. A., Pfaff, D. W. & Mcewen, B. S. (2009) Regulation of hippocampal H3 histone methylation by acute and chronic stress. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 20912-7.
- Huynh, T. T., Huynh, V. T., Harmon, M. A. & Phillips, M. A. (2003) Gene knockdown of gamma-glutamylcysteine synthetase by RNAi in the parasitic protozoa *Trypanosoma brucei* demonstrates that it is an essential enzyme. *The Journal of Biological Chemistry*, **278**, 39794-800.
- Ignatochkina, A. V., Takagi, Y., Liu, Y., Nagata, K. & Ho, C. K. (2015) The messenger RNA decapping and recapping pathway in *Trypanosoma*. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 6967-72.
- Ilboudo, H., Bras-Goncalves, R., Camara, M., Flori, L., Camara, O., Sakande, H., Leno, M., Petitdidier, E., Jamonneau, V. & Bucheton, B. (2014) Unravelling human trypanotolerance: IL8 is associated with infection control whereas IL10 and TNFalpha are associated with subsequent disease development. *PLoS Pathogens*, **10**, e1004469.
- Imboden, M. A., Laird, P. W., Affolter, M. & Seebeck, T. (1987) Transcription of the intergenic regions of the tubulin gene cluster of *Trypanosoma brucei*: evidence for a polycistronic transcription unit in a eukaryote. *Nucleic Acids Research*, **15**, 7357-68.
- Irmer, H. & Clayton, C. (2001) Degradation of the unstable EP1 mRNA in *Trypanosoma brucei* involves initial destruction of the 3'-untranslated region. *Nucleic Acids Research*, **29**, 4707-15.
- Iten, M., Mett, H., Evans, A., Enyaru, J. C., Brun, R. & Kaminsky, R. (1997) Alterations in ornithine decarboxylase characteristics account for tolerance of *Trypanosoma brucei* rhodesiense to D,L-alpha-difluoromethylornithine. *Antimicrobial Agents and Chemotherapy*, **41**, 1922-5.

- Ivens, A. C., Peacock, C. S., Worthey, 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., Rabinowitsch, 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., *et al.* (2005) The genome of the kinetoplastid parasite, *Leishmania major*. *Science*, **309**, 436-42.
- Izquierdo, L., Atrih, A., Rodrigues, J. A., Jones, D. C. & Ferguson, M. A. (2009) *Trypanosoma brucei* UDP-glucose:glycoprotein glucosyltransferase has unusual substrate specificity and protects the parasite from stress. *Eukaryotic Cell*, **8**, 230-40.
- Jackson, A. P., Allison, H. C., Barry, J. D., Field, M. C., Hertz-Fowler, C. & Berriman, M. (2013) A cell-surface phylome for African trypanosomes. *PLoS Neglected Tropical Diseases*, **7**, e2121.
- Jackson, A. P., Goyard, S., Xia, D., Foth, B. J., Sanders, M., Wastling, J. M., Minoprio, P. & Berriman, M. (2015) Global Gene Expression Profiling through the Complete Life Cycle of *Trypanosoma vivax*. *PLoS Neglected Tropical Diseases*, **9**, e0003975.
- 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. & Hertz-Fowler, C. (2010) The genome sequence of *Trypanosoma brucei gambiense*, causative agent of chronic human african trypanosomiasis. *PLoS Neglected Tropical Diseases*, **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. & Don, R. (2011) SCYX-7158, an orally-active benzoxaborole for the treatment of stage 2 human African trypanosomiasis. *PLoS Neglected Tropical Diseases*, **5**, e1151.
- Jafari, R., Almquist, H., Axelsson, H., Ignatushchenko, M., Lundback, T., Nordlund, P. & Martinez Molina, D. (2014) The cellular thermal shift assay for evaluating drug target interactions in cells. *Nature Protocols*, **9**, 2100-22.
- Jamonneau, V., Ilboudo, H., Kabore, J., Kaba, D., Koffi, M., Solano, P., Garcia, A., Courtin, D., Laveissiere, C., Lingue, K., Buscher, P. & Bucheton, B. (2012) Untreated human infections by *Trypanosoma brucei gambiense* are not 100% fatal. *PLoS Neglected Tropical Diseases*, **6**, e1691.
- Janzen, C. J., Hake, S. B., Lowell, J. E. & Cross, G. A. (2006) Selective di- or trimethylation of histone H3 lysine 76 by two DOT1 homologs is important for cell cycle regulation in *Trypanosoma brucei*. *Molecular Cell*, **23**, 497-507.
- Jarrett, J. T. (2003) The generation of 5'-deoxyadenosyl radicals by adenosylmethionine-dependent radical enzymes. *Current Opinion in Chemical Biology*, **7**, 174-82.
- Jennings, F. W. & 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. *Zeitschrift für Parasitenkunde*, **69**, 577-81.
- Jensen, B. C., Sivam, D., Kifer, C. T., Myler, P. J. & Parsons, M. (2009) Widespread variation in transcript abundance within and across developmental stages of *Trypanosoma brucei*. *BMC Genomics*, **10**, 482.

- Jones, D. C., Foth, B. J., Urbaniak, M. D., Patterson, S., Ong, H. B., Berriman, M. & Fairlamb, A. H. (2015) Genomic and Proteomic Studies on the Mode of Action of Oxaboroles against the African Trypanosome. *PLoS Neglected Tropical Diseases*, **9**, e0004299.
- Jones, N. G., Thomas, E. B., Brown, E., Dickens, N. J., Hammarton, T. C. & Mottram, J. C. (2014a) Regulators of *Trypanosoma brucei* cell cycle progression and differentiation identified using a kinome-wide RNAi screen. *PLoS Pathogens*, **10**, e1003886.
- Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., Mcanulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S. Y., Lopez, R. & Hunter, S. (2014b) InterProScan 5: genome-scale protein function classification. *Bioinformatics*, **30**, 1236-40.
- Jones, T. W. & Davila, A. M. (2001) *Trypanosoma vivax*--out of Africa. *Trends in Parasitology*, **17**, 99-101.
- Josephson, K., Ricardo, A. & Szostak, J. W. (2014) mRNA display: from basic principles to macrocycle drug discovery. *Drug Discovery Today*, **19**, 388-99.
- Kabore, J., Koffi, M., Bucheton, B., Macleod, A., Duffy, C., Ilboudo, H., Camara, M., De Meeus, T., Belem, A. M. & Jamonneau, V. (2011) First evidence that parasite infecting apparent aparasitemic serological suspects in human African trypanosomiasis are *Trypanosoma brucei gambiense* and are similar to those found in patients. *Infection, Genetics and Evolution*, **11**, 1250-5.
- Kagnoff, M. F. & Eckmann, L. (2001) Analysis of host responses to microbial infection using gene expression profiling. *Current Opinion in Microbiology*, **4**, 246-50.
- Kaiser, M., Bray, M. A., Cal, M., Bourdin Trunz, B., Torrelee, E. & Brun, R. (2011) Antitrypanosomal activity of fexinidazole, a new oral nitroimidazole drug candidate for treatment of sleeping sickness. *Antimicrobial Agents and Chemotherapy*, **55**, 5602-8.
- Kakule, T. B., Zhang, S., Zhan, J. & Schmidt, E. W. (2015) Biosynthesis of the tetramic acids Sch210971 and Sch210972. *Organic Letters*, **17**, 2295-7.
- Kalhor, H. R., Penjwini, M. & Clarke, S. (2005) A novel methyltransferase required for the formation of the hypermodified nucleoside wybutosine in eucaryotic tRNA. *Biochemical and Biophysical Research Communications*, **334**, 433-40.
- Kamleh, A., Barrett, M. P., Wildridge, D., Burchmore, R. J., Scheltema, R. A. & Watson, D. G. (2008) Metabolomic profiling using Orbitrap Fourier transform mass spectrometry with hydrophilic interaction chromatography: a method with wide applicability to analysis of biomolecules. *Rapid Communications in Mass Spectrometry*, **22**, 1912-8.
- Kandoi, G., Acencio, M. L. & Lemke, N. (2015) Prediction of Druggable Proteins Using Machine Learning and Systems Biology: A Mini-Review. *Frontiers in Physiology*, **6**, 366.
- Katz, J. E., Dlakic, M. & Clarke, S. (2003) Automated identification of putative methyltransferases from genomic open reading frames. *Molecular & Cellular Proteomics : MCP*, **2**, 525-40.
- Kennedy, P. G. (2004) Human African trypanosomiasis of the CNS: current issues and challenges. *The Journal of Clinical Investigation*, **113**, 496-504.
- Kennedy, P. G. (2013) Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). *The Lancet. Neurology*, **12**, 186-94.
- Kerkhoven, E. J., Achcar, F., Alibu, V. P., Burchmore, R. J., Gilbert, I. H., Trybilo, M., Driessen, N. N., Gilbert, D., Breitling, R., Bakker, B. M. & Barrett, M. P. (2013) Handling uncertainty in dynamic models: the pentose phosphate pathway in *Trypanosoma brucei*. *PLoS Computational Biology*, **9**, e1003371.
- Keun, H. C. & Athersuch, T. J. (2007) Application of metabonomics in drug development. *Pharmacogenomics*, **8**, 731-41.
- Keurentjes, J. J. (2009) Genetical metabolomics: closing in on phenotypes. *Current Opinion in Plant Biology*, **12**, 223-30.
- Khalaf, A. I., Huggan, J. K., Suckling, C. J., Gibson, C. L., Stewart, K., Giordani, F., Barrett, M. P., Wong, P. E., Barrack, K. L. & Hunter, W. N. (2014) Structure-based design and synthesis of antiparasitic pyrrolopyrimidines targeting pteridine reductase 1. *Journal of Medicinal Chemistry*, **57**, 6479-94.

- Kieft, R., Capewell, P., Turner, C. M., Veitch, N. J., Macleod, A. & Hajduk, S. (2010) Mechanism of *Trypanosoma brucei gambiense* (group 1) resistance to human trypanosome lytic factor. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 16137-41.
- Kim, D. H. & Sim, T. (2010) Chemical kinomics: a powerful strategy for target deconvolution. *BMB Reports*, **43**, 711-9.
- Kim, H. S., Park, S. H., Gunzl, A. & Cross, G. A. (2013) MCM-BP is required for repression of life-cycle specific genes transcribed by RNA polymerase I in the mammalian infectious form of *Trypanosoma brucei*. *PLoS One*, **8**, e57001.
- Kim, W., Choi, M. & Kim, J. E. (2014) The histone methyltransferase Dot1/DOT1L as a critical regulator of the cell cycle. *Cell Cycle*, **13**, 726-38.
- Klein, R. A., Angus, J. M. & Waterhouse, A. E. (1982) Carnitine in *Trypanosoma brucei brucei*. *Molecular and Biochemical Parasitology*, **6**, 93-110.
- Kleine, F. K. (1909) Positiv Infektionversuche mit *Trypanosoma brucei* durch *Glossina palpalis*. *Deutsche Medizinische Wochenschrift*, **35**, 469-470.
- Koffi, M., De Meeus, T., Bucheton, B., Solano, P., Camara, M., Kaba, D., Cuny, G., Ayala, F. J. & Jamonneau, V. (2009) Population genetics of *Trypanosoma brucei gambiense*, the agent of sleeping sickness in Western Africa. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 209-14.
- Koike-Yusa, H., Li, Y., Tan, E. P., Velasco-Herrera Mdel, C. & Yusa, K. (2014) Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nature Biotechnology*, **32**, 267-73.
- Kolev, N. G., Franklin, J. B., Carmi, S., Shi, H., Michaeli, S. & Tschudi, C. (2010) The transcriptome of the human pathogen *Trypanosoma brucei* at single-nucleotide resolution. *PLoS Pathogens*, **6**, e1001090.
- Kost, G. C., Yang, M. Y., Li, L., Zhang, Y., Liu, C. Y., Kim, D. J., Ahn, C. H., Lee, Y. B. & Liu, Z. R. (2015) A Novel Anti-Cancer Agent, 1-(3,5-Dimethoxyphenyl)-4-[(6-Fluoro-2-Methoxyquinoxalin-3-yl)Aminocarbonyl] Piperazine (RX-5902), Interferes With beta-Catenin Function Through Y593 Phospho-p68 RNA Helicase. *Journal of Cellular Biochemistry*, **116**, 1595-601.
- Koumandou, V. L., Natesan, S. K., Sergeenko, T. & Field, M. C. (2008) The trypanosome transcriptome is remodelled during differentiation but displays limited responsiveness within life stages. *BMC Genomics*, **9**, 298.
- Kramer, S., Kimblin, N. C. & Carrington, M. (2010) Genome-wide in silico screen for CCCH-type zinc finger proteins of *Trypanosoma brucei*, *Trypanosoma cruzi* and *Leishmania major*. *BMC Genomics*, **11**, 283.
- Kramer, S., Queiroz, R., Ellis, L., Webb, H., Hoheisel, J. D., Clayton, C. & Carrington, M. (2008) Heat shock causes a decrease in polysomes and the appearance of stress granules in trypanosomes independently of eIF2(alpha) phosphorylation at Thr169. *Journal of Cell Science*, **121**, 3002-14.
- Kraus, W. E., Muoio, D. M., Stevens, R., Craig, D., Bain, J. R., Grass, E., Haynes, C., Kwee, L., Qin, X., Slentz, D. H., Krupp, D., Muehlbauer, M., Hauser, E. R., Gregory, S. G., Newgard, C. B. & Shah, S. H. (2015) Metabolomic Quantitative Trait Loci (mQTL) Mapping Implicates the Ubiquitin Proteasome System in Cardiovascular Disease Pathogenesis. *PLoS Genetics*, **11**, e1005553.
- Kueger, S., Steinhäuser, D., Willmitzer, L. & Giavalisco, P. (2012) High-resolution plant metabolomics: from mass spectral features to metabolites and from whole-cell analysis to subcellular metabolite distributions. *The Plant Journal*, **70**, 39-50.
- Kuettel, S., Wadum, M. C., Guthrie, M. L., Marino, K., Riemer, C. & Ferguson, M. A. (2012) The de novo and salvage pathways of GDP-mannose biosynthesis are both sufficient for the growth of bloodstream-form *Trypanosoma brucei*. *Molecular Microbiology*, **84**, 340-51.
- Lai, D. H., Hashimi, H., Lun, Z. R., Ayala, F. J. & Lukes, J. (2008) Adaptations of *Trypanosoma brucei* to gradual loss of kinetoplast DNA: *Trypanosoma equiperdum* and *Trypanosoma evansi* are petite mutants of *T. brucei*. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 1999-2004.

- Lamont, G. S., Tucker, R. S. & Cross, G. A. (1986) Analysis of antigen switching rates in *Trypanosoma brucei*. *Parasitology*, **92**, 355-67.
- Lamour, N., Riviere, L., Coustou, V., Coombs, G. H., Barrett, M. P. & Bringaud, F. (2005) Proline metabolism in procyclic *Trypanosoma brucei* is down-regulated in the presence of glucose. *The Journal of Biological Chemistry*, **280**, 11902-10.
- Lander, N., Li, Z. H., Niyogi, S. & Docampo, R. (2015) CRISPR/Cas9-Induced Disruption of Paraflagellar Rod Protein 1 and 2 Genes in *Trypanosoma cruzi* Reveals Their Role in Flagellar Attachment. *MBio*, **6**, e01012.
- Langousis, G. & Hill, K. L. (2014) Motility and more: the flagellum of *Trypanosoma brucei*. *Nature Reviews Microbiology*, **12**, 505-18.
- Lanteri, C. A., Stewart, M. L., Brock, J. M., Alibu, V. P., Meshnick, S. R., Tidwell, R. R. & Barrett, M. P. (2006) Roles for the *Trypanosoma brucei* P2 transporter in DB75 uptake and resistance. *Molecular Pharmacology*, **70**, 1585-92.
- Lee, J. & Bogoy, M. (2013) Target deconvolution techniques in modern phenotypic profiling. *Current Opinion in Chemical Biology*, **17**, 118-26.
- Legros, D., Ollivier, G., Gastellu-Etchegorry, M., Paquet, C., Burri, C., Jannin, J. & Buscher, P. (2002) Treatment of human African trypanosomiasis--present situation and needs for research and development. *The Lancet. Infectious diseases*, **2**, 437-40.
- Lejon, V., Bentivoglio, M. & Franco, J. R. (2013) Human African trypanosomiasis. *Handbook of Clinical Neurology*, **114**, 169-81.
- Letunic, I., Doerks, T. & Bork, P. (2015) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Research*, **43**, D257-60.
- Leulliot, N., Bohnsack, M. T., Graille, M., Tollervey, D. & Van Tilbeurgh, H. (2008) The yeast ribosome synthesis factor Emg1 is a novel member of the superfamily of alpha/beta knot fold methyltransferases. *Nucleic Acids Research*, **36**, 629-39.
- Lewis, I. A., Wacker, M., Olszewski, K. L., Cobbold, S. A., Baska, K. S., Tan, A., Ferdig, M. T. & Llinas, M. (2014) Metabolic QTL analysis links chloroquine resistance in *Plasmodium falciparum* to impaired hemoglobin catabolism. *PLoS Genetics*, **10**, e1004085.
- Li, F., Hua, S. B., Wang, C. C. & Gottesdiener, K. M. (1998) *Trypanosoma brucei brucei*: characterization of an ODC null bloodstream form mutant and the action of alpha-difluoromethylornithine. *Experimental Parasitology*, **88**, 255-7.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & Genome Project Data Processing, S. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078-9.
- Li, H., Wang, H., Liu, Y. & Liu, Z. (2012) A benzoboroxole-functionalized monolithic column for the selective enrichment and separation of cis-diol containing biomolecules. *Chemical Communications*, **48**, 4115-7.
- Li, J. V., Saric, J., Wang, Y. L., Utzinger, J., Holmes, E. & Balmer, O. (2011) Metabonomic Investigation of Single and Multiple Strain *Trypanosoma brucei brucei* Infections. *American Journal of Tropical Medicine and Hygiene*, **84**, 91-98.
- Lima, C. D., Wang, L. K. & Shuman, S. (1999) Structure and mechanism of yeast RNA triphosphatase: an essential component of the mRNA capping apparatus. *Cell*, **99**, 533-43.
- Linstead, D. J., Klein, R. A. & Cross, G. A. (1977) Threonine catabolism in *Trypanosoma brucei*. *Journal of General Microbiology*, **101**, 243-51.
- Liow, L. H., Van Valen, L. & Stenseth, N. C. (2011) Red Queen: from populations to taxa and communities. *Trends in Ecology & Evolution*, **26**, 349-58.
- Liu, B., Yildirim, G., Wang, J., Tolun, G., Griffith, J. D. & Englund, P. T. (2010) TbPIF1, a *Trypanosoma brucei* mitochondrial DNA helicase, is essential for kinetoplast minicircle replication. *The Journal of Biological Chemistry*, **285**, 7056-66.
- Liu, C. T., Tomsho, J. W. & Benkovic, S. J. (2014) The unique chemistry of benzoxaboroles: current and emerging applications in biotechnology and therapeutic treatments. *Bioorganic and Medicinal Chemistry*, **22**, 4462-73.
- Llorente, B., Smith, C. E. & Symington, L. S. (2008) Break-induced replication: what is it and what is it for? *Cell Cycle*, **7**, 859-64.

- Lomenick, B., Hao, R., Jonai, N., Chin, R. M., Aghajan, M., Warburton, S., Wang, J., Wu, R. P., Gomez, F., Loo, J. A., Wohlschlegel, J. A., Vondriska, T. M., Pelletier, J., Herschman, H. R., Clardy, J., Clarke, C. F. & Huang, J. (2009) Target identification using drug affinity responsive target stability (DARTS). *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 21984-9.
- Lomenick, B., Jung, G., Wohlschlegel, J. A. & Huang, J. (2011a) Target identification using drug affinity responsive target stability (DARTS). *Current Protocols in Chemical Biology*, **3**, 163-180.
- Lomenick, B., Olsen, R. W. & Huang, J. (2011b) Identification of direct protein targets of small molecules. *ACS Chemical Biology*, **6**, 34-46.
- Lopez, M. A., Saada, E. A. & Hill, K. L. (2015) Insect stage-specific adenylate cyclases regulate social motility in African trypanosomes. *Eukaryotic Cell*, **14**, 104-12.
- Lott, K., Zhu, L., Fisk, J. C., Tomasello, D. L. & Read, L. K. (2014) Functional interplay between protein arginine methyltransferases in *Trypanosoma brucei*. *Microbiology*, **3**, 595-609.
- Ludewig, G., Williams, J. M., Li, Y. & Staben, C. (1994) Effects of pentamidine isethionate on *Saccharomyces cerevisiae*. *Antimicrobial Agents and Chemotherapy*, **38**, 1123-8.
- Lukes, J., Hashimi, H. & Zikova, A. (2005) Unexplained complexity of the mitochondrial genome and transcriptome in kinetoplastid flagellates. *Current Genetics*, **48**, 277-99.
- Macaulay, I. C. & Voet, T. (2014) Single cell genomics: advances and future perspectives. *PLoS Genetics*, **10**, e1004126.
- Macgregor, P. & Matthews, K. R. (2012) Identification of the regulatory elements controlling the transmission stage-specific gene expression of PAD1 in *Trypanosoma brucei*. *Nucleic Acids Research*, **40**, 7705-17.
- Macleod, A., Tweedie, A., Mclellan, S., Hope, M., Taylor, S., Cooper, A., Sweeney, L., Turner, C. M. & Tait, A. (2005) Allelic segregation and independent assortment in *T. brucei* crosses: proof that the genetic system is Mendelian and involves meiosis. *Molecular and Biochemical Parasitology*, **143**, 12-9.
- Macleod, A., Tweedie, A., Welburn, S. C., Maudlin, I., Turner, C. M. & Tait, A. (2000) Minisatellite marker analysis of *Trypanosoma brucei*: reconciliation of clonal, panmictic, and epidemic population genetic structures. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 13442-7.
- Macleod, A., Welburn, S., Maudlin, I., Turner, C. M. & Tait, A. (2001) Evidence for multiple origins of human infectivity in *Trypanosoma brucei* revealed by minisatellite variant repeat mapping. *Journal of Molecular Evolution*, **52**, 290-301.
- Magnus, E., Vervoort, T. & Van Meirvenne, N. (1978) A card-agglutination test with stained trypanosomes (C.A.T.T.) for the serological diagnosis of *T. B. gambiense* trypanosomiasis. *Annales de la Société Belge de Médecine Tropicale*, **58**, 169-76.
- Mahalingam, A., Geonnotti, A. R., Balzarini, J. & Kiser, P. F. (2011) Activity and safety of synthetic lectins based on benzoboroxole-functionalized polymers for inhibition of HIV entry. *Molecular Pharmaceutics*, **8**, 2465-75.
- Mair, G., Shi, H., Li, H., Djikeng, A., Aviles, H. O., Bishop, J. R., Falcone, F. H., Gavrilescu, C., Montgomery, J. L., Santori, M. I., Stern, L. S., Wang, Z., Ullu, E. & Tschudi, C. (2000) A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. *RNA*, **6**, 163-9.
- Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., Dicarlo, J. E., Norville, J. E. & Church, G. M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823-6.
- Marcello, L. & Barry, J. D. (2007) Analysis of the VSG gene silent archive in *Trypanosoma brucei* reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. *Genome Research*, **17**, 1344-52.
- Markham, A. (2014) Tavaborole: first global approval. *Drugs*, **74**, 1555-8.
- Martin, J. L. & Mcmillan, F. M. (2002) SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Current Opinion in Structural Biology*, **12**, 783-93.
- Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. **17**, 10-12.

- Martin, T. W. & Derewenda, Z. S. (1999) The name is bond - H bond. *Nature Structural Biology*, **6**, 403-406.
- Martinez Molina, D., Jafari, R., Ignatushchenko, M., Seki, T., Larsson, E. A., Dan, C., Sreekumar, L., Cao, Y. & Nordlund, P. (2013) Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science*, **341**, 84-7.
- Masocha, W., Rottenberg, M. E. & Kristensson, K. (2007) Migration of African trypanosomes across the blood-brain barrier. *Physiology and Behavior*, **92**, 110-4.
- Matthews, K. R. (2005) The developmental cell biology of *Trypanosoma brucei*. *Journal of Cell Science*, **118**, 283-90.
- Matthews, K. R., Ellis, J. R. & Paterou, A. (2004) Molecular regulation of the life cycle of African trypanosomes. *Trends in Parasitology*, **20**, 40-7.
- Mazet, M., Morand, P., Biran, M., Bouyssou, G., Courtois, P., Daulouede, S., Millerioux, Y., Franconi, J. M., Vincendeau, P., Moreau, P. & Bringaud, F. (2013) Revisiting the central metabolism of the bloodstream forms of *Trypanosoma brucei*: production of acetate in the mitochondrion is essential for parasite viability. *PLoS Neglected Tropical Diseases*, **7**, e2587.
- Mckean, P. G. (2003) Coordination of cell cycle and cytokinesis in *Trypanosoma brucei*. *Current Opinion in Microbiology*, **6**, 600-7.
- Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytzky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. & Depristo, M. A. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**, 1297-303.
- Mercer, L., Bowling, T., Perales, J., Freeman, J., Nguyen, T., Bacchi, C., Yarlett, N., Don, R., Jacobs, R. & Nare, B. (2011) 2,4-Diaminopyrimidines as potent inhibitors of *Trypanosoma brucei* and identification of molecular targets by a chemical proteomics approach. *PLoS Neglected Tropical Diseases*, **5**, e956.
- Michaeli, S. (2012) Spliced leader RNA silencing (SLS) - a programmed cell death pathway in *Trypanosoma brucei* that is induced upon ER stress. *Parasites & Vectors*, **5**, 107.
- Michaeli, S. (2015) The response of trypanosomes and other eukaryotes to ER stress and the spliced leader RNA silencing (SLS) pathway in *Trypanosoma brucei*. *Critical Reviews in Biochemistry and Molecular Biology*, **50**, 256-67.
- Michalkova, V., Benoit, J. B., Weiss, B. L., Attardo, G. M. & Aksoy, S. (2014) Vitamin B-6 Generated by Obligate Symbionts Is Critical for Maintaining Proline Homeostasis and Fecundity in Tsetse Flies. *Applied & Environmental Microbiology*, **80**, 5844-5853.
- Michels, P. A., Bringaud, F., Herman, M. & Hannaert, V. (2006) Metabolic functions of glycosomes in trypanosomatids. *Biochimica et Biophysica Acta*, **1763**, 1463-77.
- Millerioux, Y., Ebikeme, C., Biran, M., Morand, P., Bouyssou, G., Vincent, I. M., Mazet, M., Riviere, L., Franconi, J. M., Burchmore, R. J., Moreau, P., Barrett, M. P. & Bringaud, F. (2013) The threonine degradation pathway of the *Trypanosoma brucei* procyclic form: the main carbon source for lipid biosynthesis is under metabolic control. *Molecular Microbiology*, **90**, 114-29.
- Millerioux, Y., Morand, P., Biran, M., Mazet, M., Moreau, P., Wagnies, M., Ebikeme, C., Deramchia, K., Gales, L., Portais, J. C., Boshart, M., Franconi, J. M. & Bringaud, F. (2012) ATP synthesis-coupled and -uncoupled acetate production from acetyl-CoA by mitochondrial acetate:succinate CoA-transferase and acetyl-CoA thioesterase in *Trypanosoma*. *The Journal of Biological Chemistry*, **287**, 17186-97.
- Mininno, M., Brugiére, S., Pautre, V., Gilgen, A., Ma, S., Ferro, M., Tardif, M., Alban, C. & Ravel, S. (2012) Characterization of chloroplastic fructose 1,6-bisphosphate aldolases as lysine-methylated proteins in plants. *The Journal of Biological Chemistry*, **287**, 21034-44.
- Mirza, S. P., Halligan, B. D., Greene, A. S. & Olivier, M. (2007) Improved method for the analysis of membrane proteins by mass spectrometry. *Physiological Genomics*, **30**, 89-94.
- Misra, B. B. & Van Der Hooft, J. J. (2016) Updates in metabolomics tools and resources: 2014-2015. *Electrophoresis*, **37**, 86-110.

- Misset, O., Bos, O. J. & Opperdoes, F. R. (1986) Glycolytic enzymes of *Trypanosoma brucei*. Simultaneous purification, intraglycosomal concentrations and physical properties. *European Journal of Biochemistry*, **157**, 441-53.
- Mitashi, P., Hasker, E., Lejon, V., Kande, V., Muyembe, J. J., Lutumba, P. & Boelaert, M. (2012) Human african trypanosomiasis diagnosis in first-line health services of endemic countries, a systematic review. *PLoS Neglected Tropical Diseases*, **6**, e1919.
- Mittra, B., Zamudio, J. R., Bujnicki, J. M., Stepinski, J., Darzynkiewicz, E., Campbell, D. A. & Sturm, N. R. (2008) The TbMTr1 spliced leader RNA cap 1 2'-O-ribose methyltransferase from *Trypanosoma brucei* acts with substrate specificity. *The Journal of Biological Chemistry*, **283**, 3161-72.
- Moller, W., Schipper, A. & Amons, R. (1987) A conserved amino acid sequence around Arg-68 of *Artemia* elongation factor 1 alpha is involved in the binding of guanine nucleotides and aminoacyl transfer RNAs. *Biochimie*, **69**, 983-9.
- Monnerat, S., Clucas, C., Brown, E., Mottram, J. C. & Hammarton, T. C. (2009) Searching for novel cell cycle regulators in *Trypanosoma brucei* with an RNA interference screen. *BMC Research Notes*, **2**, 46.
- Mony, B. M., Macgregor, P., Ivens, A., Rojas, F., Cowton, A., Young, J., Horn, D. & Matthews, K. (2014) Genome-wide dissection of the quorum sensing signalling pathway in *Trypanosoma brucei*. *Nature*, **505**, 681-5.
- Moore, J. D. (2015) The impact of CRISPR-Cas9 on target identification and validation. *Drug Discovery Today*, **20**, 450-7.
- Morris, J. C., Wang, Z., Drew, M. E. & Englund, P. T. (2002) Glycolysis modulates trypanosome glycoprotein expression as revealed by an RNAi library. *The EMBO journal*, **21**, 4429-38.
- Morrison, L. J., Tait, A., McCormack, G., Sweeney, L., Black, A., Truc, P., Likeufack, A. C., Turner, C. M. & Macleod, A. (2008) *Trypanosoma brucei gambiense* Type 1 populations from human patients are clonal and display geographical genetic differentiation. *Infection, Genetics and Evolution*, **8**, 847-54.
- Morrison, L. J., Tait, A., McLellan, S., Sweeney, L., Turner, C. M. & Macleod, A. (2009) A major genetic locus in *Trypanosoma brucei* is a determinant of host pathology. *PLoS Neglected Tropical Diseases*, **3**, e557.
- Mortazavi, A., Williams, B. A., Mccue, K., Schaeffer, L. & Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, **5**, 621-8.
- Moss, C. X., Brown, E., Hamilton, A., Van Der Veken, P., Augustyns, K. & Mottram, J. C. (2015) An essential signal peptide peptidase identified in an RNAi screen of serine peptidases of *Trypanosoma brucei*. *PLoS One*, **10**, e0123241.
- Mugasa, C. M., Adams, E. R., Boer, K. R., Dyserinck, H. C., Buscher, P., Schallig, H. D. & Leeflang, M. M. (2012) Diagnostic accuracy of molecular amplification tests for human African trypanosomiasis—systematic review. *PLoS Neglected Tropical Diseases*, **6**, e1438.
- Mukadi, P., Gillet, P., Lukuka, A., Atua, B., Sheshe, N., Kanza, A., Mayunda, J. B., Mongita, B., Senga, R., Ngoyi, J., Muyembe, J. J., Jacobs, J. & Lejon, V. (2013) External quality assessment of Giemsa-stained blood film microscopy for the diagnosis of malaria and sleeping sickness in the Democratic Republic of the Congo. *Bulletin of the World Health Organization*, **91**, 441-8.
- Mulenga, C., Mhlanga, J. D., Kristensson, K. & Robertson, B. (2001) *Trypanosoma brucei brucei* crosses the blood-brain barrier while tight junction proteins are preserved in a rat chronic disease model. *Neuropathology and Applied Neurobiology*, **27**, 77-85.
- Munday, J. C., Eze, A. A., Baker, N., Glover, L., Clucas, C., Aguinaga Andres, D., Natto, M. J., Teka, 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. & 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. *The Journal of Antimicrobial Chemotherapy*, **69**, 651-63.
- Murray, M., Morrison, W. I. & Whitelaw, D. D. (1982) Host susceptibility to African trypanosomiasis: trypanotolerance. *Advances in Parasitology*, **21**, 1-68.

- Mwanakasale, V., Songolo, P. & Daka, V. (2013) Challenges in the control of human African trypanosomiasis in the Mpika district of Zambia. *BMC Research Notes*, **6**, 180.
- Narayan, M., Seeley, K. W. & Jinwal, U. K. (2015) Identification and quantitative analysis of cellular proteins affected by treatment with withaferin A using a SILAC-based proteomics approach. *Journal of Ethnopharmacology*, **175**, 86-92.
- 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. & 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. *Antimicrobial Agents and Chemotherapy*, **54**, 4379-88.
- Nicholson, G., Rantalainen, M., Li, J. V., Maher, A. D., Malmodin, D., Ahmadi, K. R., Faber, J. H., Barrett, A., Min, J. L., Rayner, N. W., Toft, H., Krestyaninova, M., Viksna, J., Neogi, S. G., Dumas, M. E., Sarkans, U., Mol, P. C., Donnelly, P., Illig, T., Adamski, J., Suhre, K., Allen, M., Zondervan, K. T., Spector, T. D., Nicholson, J. K., Lindon, J. C., Baunsgaard, D., Holmes, E., Mccarthy, M. I. & Holmes, C. C. (2011) A genome-wide metabolic QTL analysis in Europeans implicates two loci shaped by recent positive selection. *PLoS Genetics*, **7**, e1002270.
- Niemann, M., Wiese, S., Mani, J., Chanfon, A., Jackson, C., Meisinger, C., Warscheid, B. & Schneider, A. (2013) Mitochondrial outer membrane proteome of *Trypanosoma brucei* reveals novel factors required to maintain mitochondrial morphology. *Molecular & Cellular Proteomics : MCP*, **12**, 515-28.
- Nihei, C., Fukai, Y. & Kita, K. (2002) Trypanosome alternative oxidase as a target of chemotherapy. *Biochimica et Biophysica Acta*, **1587**, 234-9.
- Njiokou, F., Nimpaye, H., Simo, G., Njitchouang, G. R., Asonganyi, T., Cuny, G. & Herder, S. (2010) Domestic animals as potential reservoir hosts of *Trypanosoma brucei gambiense* in sleeping sickness foci in Cameroon. *Parasite*, **17**, 61-6.
- Noma, A., Kirino, Y., Ikeuchi, Y. & Suzuki, T. (2006) Biosynthesis of wybutosine, a hyper-modified nucleoside in eukaryotic phenylalanine tRNA. *The EMBO journal*, **25**, 2142-54.
- Nugent, B. M., Wright, C. L., Shetty, A. C., Hodes, G. E., Lenz, K. M., Mahurkar, A., Russo, S. J., Devine, S. E. & Mccarthy, M. M. (2015) Brain feminization requires active repression of masculinization via DNA methylation. *Nature Neuroscience*, **18**, 690-7.
- Nzila, A. & Mwai, L. (2010) In vitro selection of *Plasmodium falciparum* drug-resistant parasite lines. *The Journal of Antimicrobial Chemotherapy*, **65**, 390-8.
- Okuda, S., Yamada, T., Hamajima, M., Itoh, M., Katayama, T., Bork, P., Goto, S. & Kanehisa, M. (2008) KEGG Atlas mapping for global analysis of metabolic pathways. *Nucleic Acids Research*, **36**, W423-6.
- Olin-Sandoval, V., Gonzalez-Chavez, Z., Berzunza-Cruz, M., Martinez, I., Jasso-Chavez, R., Becker, I., Espinoza, B., Moreno-Sanchez, R. & Saavedra, E. (2012) Drug target validation of the trypanothione pathway enzymes through metabolic modelling. *FEBS J*, **279**, 1811-33.
- Oliver, S. G., Winson, M. K., Kell, D. B. & Baganz, F. (1998) Systematic functional analysis of the yeast genome. *Trends in Biotechnology*, **16**, 373-8.
- Olsen, J. B., Cao, X. J., Han, B., Chen, L. H., Horvath, A., Richardson, T. I., Campbell, R. M., Garcia, B. A. & Nguyen, H. (2016) Quantitative profiling of the activity of protein lysine methyltransferase SMYD2 using SILAC-based proteomics. *Molecular & Cellular Proteomics : MCP*.
- Omidfar, K. & Daneshpour, M. (2015) Advances in phage display technology for drug discovery. *Expert Opinion on Drug Discovery*, **10**, 651-69.
- Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A. & Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & Cellular Proteomics : MCP*, **1**, 376-86.
- Ong, S. E., Schenone, M., Margolin, A. A., Li, X., Do, K., Doud, M. K., Mani, D. R., Kuai, L., Wang, X., Wood, J. L., Tolliday, N. J., Koehler, A. N., Marcaurelle, L. A., Golub, T. R., Gould, R. J., Schreiber, S. L. & Carr, S. A. (2009) Identifying the proteins to which small-molecule

- probes and drugs bind in cells. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 4617-22.
- Oppendoes, F. R., Borst, P., Bakker, S. & Leene, W. (1977) Localization of glycerol-3-phosphate oxidase in the mitochondrion and particulate NAD<sup>+</sup>-linked glycerol-3-phosphate dehydrogenase in the microbodies of the bloodstream form of *Trypanosoma brucei*. *European Journal of Biochemistry*, **76**, 29-39.
- Overath, P., Czichos, J. & Haas, C. (1986) The effect of citrate/cis-aconitate on oxidative metabolism during transformation of *Trypanosoma brucei*. *European Journal of Biochemistry*, **160**, 175-82.
- Overath, P. & Engstler, M. (2004) Endocytosis, membrane recycling and sorting of GPI-anchored proteins: *Trypanosoma brucei* as a model system. *Molecular Microbiology*, **53**, 735-44.
- Ozsolak, F., Platt, A. R., Jones, D. R., Reifengerger, J. G., Sass, L. E., Mcinerney, P., Thompson, J. F., Bowers, J., Jarosz, M. & Milos, P. M. (2009) Direct RNA sequencing. *Nature*, **461**, 814-8.
- Pai, M. Y., Lomenick, B., Hwang, H., Schiestl, R., McBride, W., Loo, J. A. & Huang, J. (2015) Drug affinity responsive target stability (DARTS) for small-molecule target identification. *Methods in Molecular Biology*, **1263**, 287-98.
- Pareek, C. S., Smoczynski, R. & Tretyn, A. (2011) Sequencing technologies and genome sequencing. *Journal of Applied Genetics*, **52**, 413-35.
- Paris, Z., Horakova, E., Rubio, M. A., Sample, P., Fleming, I. M., Armocida, S., Lukes, J. & Alfonzo, J. D. (2013) The *T. brucei* TRM5 methyltransferase plays an essential role in mitochondrial protein synthesis and function. *RNA*, **19**, 649-58.
- Parsons, M., Nelson, R. G., Watkins, K. P. & Agabian, N. (1984) Trypanosome mRNAs share a common 5' spliced leader sequence. *Cell*, **38**, 309-16.
- Parsons, M., Worthey, E. A., Ward, P. N. & Mottram, J. C. (2005) Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*. *BMC Genomics*, **6**, 127.
- Pasternack, D. A., Sayegh, J., Clarke, S. & Read, L. K. (2007) Evolutionarily divergent type II protein arginine methyltransferase in *Trypanosoma brucei*. *Eukaryotic Cell*, **6**, 1665-81.
- Patnaik, P. K., Field, M. C., Menon, A. K., Cross, G. A., Yee, M. C. & Butikofer, P. (1993) Molecular species analysis of phospholipids from *Trypanosoma brucei* bloodstream and procyclic forms. *Molecular and Biochemical Parasitology*, **58**, 97-105.
- Pelletier, M., Pasternack, D. A. & Read, L. K. (2005) In vitro and in vivo analysis of the major type I protein arginine methyltransferase from *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, **144**, 206-17.
- Peng, D., Kurup, S. P., Yao, P. Y., Minning, T. A. & Tarleton, R. L. (2015) CRISPR-Cas9-mediated single-gene and gene family disruption in *Trypanosoma cruzi*. *MBio*, **6**, e02097-14.
- Pepin, J. & Milord, F. (1994) The treatment of human African trypanosomiasis. *Advances in Parasitology*, **33**, 1-47.
- Perez-Moreno, G., Sealey-Cardona, M., Rodrigues-Poveda, C., Gelb, M. H., Ruiz-Perez, L. M., Castillo-Acosta, V., Urbina, J. A. & Gonzalez-Pacanowska, D. (2012) Endogenous sterol biosynthesis is important for mitochondrial function and cell morphology in procyclic forms of *Trypanosoma brucei*. *International Journal for Parasitology*, **42**, 975-89.
- Perez-Morga, D., Vanhollebeke, B., Paturiaux-Hanocq, F., Nolan, D. P., Lins, L., Hombler, F., Vanhamme, L., Tebabi, P., Pays, A., Poelvoorde, P., Jacquet, A., Brasseur, R. & Pays, E. (2005) Apolipoprotein L-I promotes trypanosome lysis by forming pores in lysosomal membranes. *Science*, **309**, 469-72.
- Perry, K. L., Watkins, K. P. & Agabian, N. (1987) Trypanosome mRNAs have unusual "cap 4" structures acquired by addition of a spliced leader. *Proceedings of the National Academy of Sciences of the United States of America*, **84**, 8190-4.
- Peterson, J. R., Lebensohn, A. M., Pelish, H. E. & Kirschner, M. W. (2006) Biochemical suppression of small-molecule inhibitors: a strategy to identify inhibitor targets and signaling pathway components. *Chemistry and Biology*, **13**, 443-52.
- Petrossian, T. & Clarke, S. (2009) Bioinformatic Identification of Novel Methyltransferases. *Epigenomics*, **1**, 163-175.

- Petrossian, T. C. & Clarke, S. G. (2011) Uncovering the human methyltransferasome. *Molecular & Cellular Proteomics : MCP*, **10**, M110 000976.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, **25**, 1605-12.
- Phillips, M. A., Coffino, P. & Wang, C. C. (1987) Cloning and sequencing of the ornithine decarboxylase gene from *Trypanosoma brucei*. Implications for enzyme turnover and selective difluoromethylornithine inhibition. *The Journal of Biological Chemistry*, **262**, 8721-7.
- Phillips, M. A. & Wang, C. C. (1987) A *Trypanosoma brucei* mutant resistant to alpha-difluoromethylornithine. *Molecular and Biochemical Parasitology*, **22**, 9-17.
- Piskol, R., Ramaswami, G. & Li, J. B. (2013) Reliable identification of genomic variants from RNA-seq data. *American Journal of Human Genetics*, **93**, 641-51.
- Poincaré, H. 1905. *Science and Hypothesis*, London, UK, The Walter Scott Publishing Co., Ltd.
- Pratt, C., Nguyen, S. & Phillips, M. A. (2014) Genetic validation of *Trypanosoma brucei* glutathione synthetase as an essential enzyme. *Eukaryotic Cell*, **13**, 614-24.
- Priotto, G., Kasparian, S., Mutombo, W., Ngouama, D., Ghorashian, S., Arnold, U., Ghabri, S., Baudin, E., Buard, V., Kazadi-Kyanza, S., Ilunga, M., Mutangala, W., Pohlig, G., Schmid, C., Karunakara, U., Torreele, E. & Kande, V. (2009) Nifurtimox-eflornithine combination therapy for second-stage African *Trypanosoma brucei* gambiense trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. *Lancet*, **374**, 56-64.
- Quinn, E. M., Cormican, P., Kenny, E. M., Hill, M., Anney, R., Gill, M., Corvin, A. P. & Morris, D. W. (2013) Development of strategies for SNP detection in RNA-seq data: application to lymphoblastoid cell lines and evaluation using 1000 Genomes data. *PLoS One*, **8**, e58815.
- R Core Team. 2013. *R: A language and environment for statistical computing*. [Online]. Available: <http://www.R-project.org/>.
- Rabinowitz, J. D., Purdy, J. G., Vastag, L., Shenk, T. & Koyuncu, E. (2011) Metabolomics in drug target discovery. *Cold Spring Harbor Symposia on Quantitative Biology*, **76**, 235-46.
- Ragsdale, S. W. (2008) Catalysis of methyl group transfers involving tetrahydrofolate and B(12). *Vitamins and Hormones*, **79**, 293-324.
- Ranade, R. M., Gillespie, J. R., Shibata, S., Verlinde, C. L., Fan, E., Hol, W. G. & Buckner, F. S. (2013) Induced resistance to methionyl-tRNA synthetase inhibitors in *Trypanosoma brucei* is due to overexpression of the target. *Antimicrobial Agents and Chemotherapy*, **57**, 3021-8.
- Raper, J., Fung, R., Ghiso, J., Nussenzweig, V. & Tomlinson, S. (1999) Characterization of a novel trypanosome lytic factor from human serum. *Infection and Immunity*, **67**, 1910-6.
- Rassi, A., Jr., Rassi, A. & Marin-Neto, J. A. (2010) Chagas disease. *Lancet*, **375**, 1388-402.
- Raz, B., Iten, M., Grether-Buhler, Y., Kaminsky, R. & Brun, R. (1997) The Alamar Blue assay to determine drug sensitivity of African trypanosomes (*T.b. rhodesiense* and *T.b. gambiense*) in vitro. *Acta Tropica*, **68**, 139-47.
- Reguera, R. M., Calvo-Alvarez, E., Alvarez-Velilla, R. & Balana-Fouce, R. (2014) Target-based vs. phenotypic screenings in *Leishmania* drug discovery: A marriage of convenience or a dialogue of the deaf? *International Journal for Parasitology: Drugs and Drug Resistance*, **4**, 355-7.
- Reinitz, D. M. & Mansfield, J. M. (1990) T-cell-independent and T-cell-dependent B-cell responses to exposed variant surface glycoprotein epitopes in trypanosome-infected mice. *Infection and Immunity*, **58**, 2337-42.
- Rezwan, M. & Auerbach, D. (2012) Yeast "N"-hybrid systems for protein-protein and drug-protein interaction discovery. *Methods*, **57**, 423-9.
- Richardson, J. P., Beecroft, R. P., Tolson, D. L., Liu, M. K. & Pearson, T. W. (1988) Procyclin: an unusual immunodominant glycoprotein surface antigen from the procyclic stage of African trypanosomes. *Molecular and Biochemical Parasitology*, **31**, 203-16.
- Richmond, G. S., Gibellini, F., Young, S. A., Major, L., Denton, H., Lilley, A. & Smith, T. K. (2010) Lipidomic analysis of bloodstream and procyclic form *Trypanosoma brucei*. *Parasitology*, **137**, 1357-92.

- Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. & Mesirov, J. P. (2011) Integrative genomics viewer. *Nature Biotechnology*, **29**, 24-6.
- Robles, A. & Clayton, C. (2008) Regulation of an amino acid transporter mRNA in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, **157**, 102-6.
- Rock, F. L., Mao, W., Yaremchuk, A., Tuko, M., Crepin, T., Zhou, H., Zhang, Y. K., Hernandez, V., Akama, T., Baker, S. J., Plattner, J. J., Shapiro, L., Martinis, S. A., Benkovic, S. J., Cusack, S. & Alley, M. R. (2007) An antifungal agent inhibits an aminoacyl-tRNA synthetase by trapping tRNA in the editing site. *Science*, **316**, 1759-61.
- Roditi, I. & Lehane, M. J. (2008) Interactions between trypanosomes and tsetse flies. *Current Opinion in Microbiology*, **11**, 345-51.
- Roditi, I., Schwarz, H., Pearson, T. W., Beecroft, R. P., Liu, M. K., Richardson, J. P., Buhning, H. J., Pleiss, J., Bulow, R., Williams, R. O. & Et Al. (1989) Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of *Trypanosoma brucei*. *The Journal of Cell Biology*, **108**, 737-46.
- Rodriguez-Contreras, D., Feng, X., Keeney, K. M., Bouwer, H. G. & Landfear, S. M. (2007) Phenotypic characterization of a glucose transporter null mutant in *Leishmania mexicana*. *Molecular and Biochemical Parasitology*, **153**, 9-18.
- Rojo, D., Canuto, G. A., Castilho-Martins, E. A., Tavares, M. F., Barbas, C., Lopez-Gonzalez, A. & Rivas, L. (2015) A Multiplatform Metabolomic Approach to the Basis of Antimonial Action and Resistance in *Leishmania infantum*. *PLoS One*, **10**, e0130675.
- Ruan, J. P., Shen, S., Ullu, E. & Tschudi, C. (2007a) Evidence for a capping enzyme with specificity for the trypanosome spliced leader RNA. *Molecular and Biochemical Parasitology*, **156**, 246-54.
- Ruan, J. P., Ullu, E. & Tschudi, C. (2007b) Characterization of the *Trypanosoma brucei* cap hypermethylase Tgs1. *Molecular and Biochemical Parasitology*, **155**, 66-9.
- Russell, C., Rahman, A. & Mohammed, A. R. (2013) Application of genomics, proteomics and metabolomics in drug discovery, development and clinic. *Therapeutic Delivery*, **4**, 395-413.
- Ryland, G. L., Doyle, M. A., Goode, D., Boyle, S. E., Choong, D. Y., Rowley, S. M., Li, J., Australian Ovarian Cancer Study, G., Bowtell, D. D., Tothill, R. W., Campbell, I. G. & Goringe, K. L. (2015) Loss of heterozygosity: what is it good for? *BMC Medical Genomics*, **8**, 45.
- Saada, E. A., Kabututu, Z. P., Lopez, M., Shimogawa, M. M., Langousis, G., Oberholzer, M., Riestra, A., Jonsson, Z. O., Wohlschlegel, J. A. & Hill, K. L. (2014) Insect stage-specific receptor adenylate cyclases are localized to distinct subdomains of the *Trypanosoma brucei* flagellar membrane. *Eukaryotic Cell*, **13**, 1064-76.
- Salavati, R. & Najafabadi, H. S. (2010) Sequence-based functional annotation: what if most of the genes are unique to a genome? *Trends in Parasitology*, **26**, 225-9.
- Santos, C. C., Coombs, G. H., Lima, A. P. & Mottram, J. C. (2007) Role of the *Trypanosoma brucei* natural cysteine peptidase inhibitor ICP in differentiation and virulence. *Molecular Microbiology*, **66**, 991-1002.
- Sartori, A., Garay-Malpartida, H. M., Forni, M. F., Schumacher, R. I., Dutra, F., Sogayar, M. C. & Bechara, E. J. (2008) Aminoacetone, a putative endogenous source of methylglyoxal, causes oxidative stress and death to insulin-producing RINm5f cells. *Chemical Research in Toxicology*, **21**, 1841-50.
- Savitski, M. M., Reinhard, F. B., Franken, H., Werner, T., Savitski, M. F., Eberhard, D., Martinez Molina, D., Jafari, R., Dovega, R. B., Klaeger, S., Kuster, B., Nordlund, P., Bantscheff, M. & Drewes, G. (2014) Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science*, **346**, 1255784.
- Schax, E., Walter, J. G., Marzhauser, H., Stahl, F., Scheper, T., Agard, D. A., Eichner, S., Kirschning, A. & Zeilinger, C. (2014) Microarray-based screening of heat shock protein inhibitors. *Journal of Biotechnology*, **180**, 1-9.
- Scheltema, R. A., Jankevics, A., Jansen, R. C., Swertz, M. A. & Breitling, R. (2011) PeakML/mzMatch: A File Format, Java Library, R Library, and Tool-Chain for Mass Spectrometry Data Analysis. *Analytical Chemistry*, **83**, 2786-2793.

- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. & Cardona, A. (2012) Fiji: an open-source platform for biological-image analysis. *Nature Methods*, **9**, 676-82.
- Schubert, H. L., Blumenthal, R. M. & Cheng, X. (2003) Many paths to methyltransfer: a chronicle of convergence. *Trends in Biochemical Sciences*, **28**, 329-35.
- Schumann Burkard, G., Jutzi, P. & Roditi, I. (2011) Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters. *Molecular and Biochemical Parasitology*, **175**, 91-4.
- Scott, A. G., Tait, A. & Turner, C. M. (1996) Characterisation of cloned lines of *Trypanosoma brucei* expressing stable resistance to MelCy and suramin. *Acta Tropica*, **60**, 251-62.
- Scott, J. M. & Pegram, R. G. (1974) A high incidence of *Trypanosoma congolense* strains resistant to homidium bromide in Ethiopia. *Tropical Animal Health and Production*, **6**, 215-21.
- Seifert, K., Munday, J., Syeda, T. & Croft, S. L. (2011) In vitro interactions between sitamaquine and amphotericin B, sodium stibogluconate, miltefosine, paromomycin and pentamidine against *Leishmania donovani*. *The Journal of Antimicrobial Chemotherapy*, **66**, 850-4.
- Serricchio, M. & Butikofer, P. (2012) An essential bacterial-type cardiolipin synthase mediates cardiolipin formation in a eukaryote. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, E954-61.
- Shak, S., Davitz, M. A., Wolinsky, M. L., Nussenzweig, V., Turner, M. J. & Gurnett, A. (1988) Partial characterization of the cross-reacting determinant, a carbohydrate epitope shared by decay accelerating factor and the variant surface glycoprotein of the African *Trypanosoma brucei*. *Journal of Immunology*, **140**, 2046-50.
- Shateri Najafabadi, H. & Salavati, R. (2010) Functional genome annotation by combined analysis across microarray studies of *Trypanosoma brucei*. *PLoS Neglected Tropical Diseases*, **4**.
- Shea, C., Lee, M. G. & Van Der Ploeg, L. H. (1987) VSG gene 118 is transcribed from a cotransposed pol I-like promoter. *Cell*, **50**, 603-12.
- Shi, H., Barnes, R. L., Carriero, N., Atayde, V. D., Tschudi, C. & Ullu, E. (2014) Role of the *Trypanosoma brucei* HEN1 family methyltransferase in small interfering RNA modification. *Eukaryotic Cell*, **13**, 77-86.
- Shlomai, J. (2004) The structure and replication of kinetoplast DNA. *Current Molecular Medicine*, **4**, 623-47.
- Siegel, T. N., Hekstra, D. R. & Cross, G. A. (2008) Analysis of the *Trypanosoma brucei* cell cycle by quantitative DAPI imaging. *Molecular and Biochemical Parasitology*, **160**, 171-4.
- Siegel, T. N., Hekstra, D. R., Wang, X., Dewell, S. & Cross, G. A. (2010) Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. *Nucleic Acids Research*, **38**, 4946-57.
- Sienkiewicz, N., Ong, H. B. & Fairlamb, A. H. (2010) *Trypanosoma brucei* pteridine reductase 1 is essential for survival in vitro and for virulence in mice. *Molecular Microbiology*, **77**, 658-71.
- Simarro, P. P., Diarra, A., Ruiz Postigo, J. A., Franco, J. R. & 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.
- Singh, A., Minia, I., Droll, D., Fadda, A., Clayton, C. & Erben, E. (2014) Trypanosome MKT1 and the RNA-binding protein ZC3H11: interactions and potential roles in post-transcriptional regulatory networks. *Nucleic Acids Research*, **42**, 4652-68.
- Slon-Usakiewicz, J. J., Dai, J. R., Ng, W., Foster, J. E., Deretey, E., Toledo-Sherman, L., Redden, P. R., Pasternak, A. & Reid, N. (2005) Global kinase screening. Applications of frontal affinity chromatography coupled to mass spectrometry in drug discovery. *Analytical Chemistry*, **77**, 1268-74.
- Smith, T. K. & Butikofer, P. (2010) Lipid metabolism in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, **172**, 66-79.
- Smith, T. K., Cottaz, S., Brimacombe, J. S. & Ferguson, M. A. (1996) Substrate specificity of the dolichol phosphate mannose: glucosaminyl phosphatidylinositol alpha1-4-

- mannosyltransferase of the glycosylphosphatidylinositol biosynthetic pathway of African trypanosomes. *The Journal of Biological Chemistry*, **271**, 6476-82.
- Smith, T. K., Vasileva, N., Gluenz, E., Terry, S., Portman, N., Kramer, S., Carrington, M., Michaeli, S., Gull, K. & Rudenko, G. (2009) Blocking variant surface glycoprotein synthesis in *Trypanosoma brucei* triggers a general arrest in translation initiation. *PLoS One*, **4**, e7532.
- Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F. & Miller, N. E. (2001) Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Research*, **29**, 1097-106.
- Sokolova, A. Y., Wyllie, S., Patterson, S., Oza, S. L., Read, K. D. & Fairlamb, A. H. (2010) Cross-resistance to nitro drugs and implications for treatment of human African trypanosomiasis. *Antimicrobial Agents and Chemotherapy*, **54**, 2893-900.
- Sollelis, L., Ghorbal, M., Macpherson, C. R., Martins, R. M., Kuk, N., Crobu, L., Bastien, P., Scherf, A., Lopez-Rubio, J. J. & Sterkers, Y. (2015) First efficient CRISPR-Cas9-mediated genome editing in *Leishmania* parasites. *Cellular Microbiology*, **17**, 1405-12.
- Spitznagel, D., Ebikeme, C., Biran, M., Nic a' Bhaird, N., Bringaud, F., Henehan, G. T. & Nolan, D. P. (2009) Alanine aminotransferase of *Trypanosoma brucei*--a key role in proline metabolism in procyclic life forms. *FEBS Journal*, **276**, 7187-99.
- Spry, C., Kirk, K. & Saliba, K. J. (2008) Coenzyme A biosynthesis: an antimicrobial drug target. *FEMS Microbiology Reviews*, **32**, 56-106.
- Stephens, J. W. W. & Fantham, H. B. (1910) On the peculiar morphology of a trypanosome from a case of sleeping sickness and the possibility of its being a new species (*T. rhodesiense*). **83**, 28-33.
- Stephens, N. A., Kieft, R., Macleod, A. & Hajduk, S. L. (2012) Trypanosome resistance to human innate immunity: targeting Achilles' heel. *Trends in Parasitology*, **28**, 539-45.
- Sternberg, J. M., Gierlinski, M., Bieler, S., Ferguson, M. A. & Ndung'u, J. M. (2014) Evaluation of the diagnostic accuracy of prototype rapid tests for human African trypanosomiasis. *PLoS Neglected Tropical Diseases*, **8**, e3373.
- Steverding, D. (2008) The history of African trypanosomiasis. *Parasit Vectors*, **1**, 3.
- Stoffel, S. A., Alibu, V. P., Hubert, J., Ebikeme, C., Portais, J. C., Bringaud, F., Schweingruber, M. E. & Barrett, M. P. (2011) Transketolase in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, **179**, 1-7.
- Stoffel, S. A., Rodenko, B., Schweingruber, A. M., Maser, P., De Koning, H. P. & Schweingruber, M. E. (2006) Biosynthesis and uptake of thiamine (vitamin B1) in bloodstream form *Trypanosoma brucei brucei* and interference of the vitamin with melarsen oxide activity. *International Journal for Parasitology*, **36**, 229-36.
- Sturm, N. R. & Simpson, L. (1990) Kinetoplast DNA minicircles encode guide RNAs for editing of cytochrome oxidase subunit III mRNA. *Cell*, **61**, 879-84.
- Subramaniam, D., Thombre, R., Dhar, A. & Anant, S. (2014) DNA methyltransferases: a novel target for prevention and therapy. *Frontiers in Oncology*, **4**, 80.
- Swinehart, W. E. & Jackman, J. E. (2015) Diversity in mechanism and function of tRNA methyltransferases. *RNA Biology*, **12**, 398-411.
- Sykes, M. L. & Avery, V. M. (2009a) Development of an Alamar Blue viability assay in 384-well format for high throughput whole cell screening of *Trypanosoma brucei brucei* bloodstream form strain 427. *The American Journal of Tropical Medicine and Hygiene*, **81**, 665-74.
- Sykes, M. L. & Avery, V. M. (2009b) A luciferase based viability assay for ATP detection in 384-well format for high throughput whole cell screening of *Trypanosoma brucei brucei* bloodstream form strain 427. *Parasite Vectors*, **2**, 54.
- Tagoe, D. N. A., Kalejaiye, T. D. & De Koning, H. P. (2015) The ever unfolding story of cAMP signaling in trypanosomatids: vive la difference! *Frontiers in Pharmacology*, **6**.
- Tait, A., Babiker, E. A. & Le Ray, D. (1984) Enzyme variation in *Trypanosoma brucei* spp. I. Evidence for the sub-speciation of *Trypanosoma brucei gambiense*. *Parasitology*, **89**, 311-26.

- Tait, A., Macleod, A., Tweedie, A., Masiga, D. & Turner, C. M. (2007) Genetic exchange in *Trypanosoma brucei*: evidence for mating prior to metacyclic stage development. *Molecular and Biochemical Parasitology*, **151**, 133-6.
- Tait, A., Masiga, D., Ouma, J., Macleod, A., Sasse, J., Melville, S., Lindegard, G., Mcintosh, A. & Turner, M. (2002) Genetic analysis of phenotype in *Trypanosoma brucei*: a classical approach to potentially complex traits. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **357**, 89-99.
- Tait, A., Morrison, L. J., Duffy, C. W., Cooper, A., Turner, C. M. & Macleod, A. (2011) Trypanosome genetics: populations, phenotypes and diversity. *Veterinary Parasitology*, **181**, 61-8.
- Takagi, Y., Sindkar, S., Ekonomidis, D., Hall, M. P. & Ho, C. K. (2007) *Trypanosoma brucei* encodes a bifunctional capping enzyme essential for cap 4 formation on the spliced leader RNA. *The Journal of Biological Chemistry*, **282**, 15995-6005.
- Tarral, A., Blesson, S., Mordt, O. V., Torreele, E., Sassella, D., Bray, M. A., Hovsepian, L., Evene, E., Gualano, V., Felices, M. & Strub-Wourgaft, N. (2014) Determination of an optimal dosing regimen for fexinidazole, a novel oral drug for the treatment of human African trypanosomiasis: first-in-human studies. *Clinical Pharmacokinetics*, **53**, 565-80.
- Teka, I. A., Kazibwe, A. J., El-Sabbagh, N., Al-Salabi, M. I., Ward, C. P., Eze, A. A., Munday, J. C., Maser, P., Matovu, E., Barrett, M. P. & De Koning, H. P. (2011) The diamidine diminazene aceturate is a substrate for the high-affinity pentamidine transporter: implications for the development of high resistance levels in trypanosomes. *Molecular Pharmacology*, **80**, 110-6.
- Tetley, L. & Vickerman, K. (1991) The glycosomes of trypanosomes: number and distribution as revealed by electron spectroscopic imaging and 3-D reconstruction. *Journal of Microscopy*, **162**, 83-90.
- Tiengwe, C., Brown, A. E. & Bangs, J. D. (2015) Unfolded Protein Response Pathways in Bloodstream-Form *Trypanosoma brucei*? *Eukaryotic Cell*, **14**, 1094-101.
- Timms, J. F. & Cramer, R. (2008) Difference gel electrophoresis. *Proteomics*, **8**, 4886-97.
- Tirados, I., Esterhuizen, J., Kovacic, V., Mangwiro, T. N., Vale, G. A., Hastings, I., Solano, P., Lehane, M. J. & Torr, S. J. (2015) Tsetse Control and Gambian Sleeping Sickness; Implications for Control Strategy. *PLoS Neglected Tropical Diseases*, **9**, e0003822.
- Torreele, E., Bourdin Trunz, B., Tweats, D., Kaiser, M., Brun, R., Mazue, G., Bray, M. A. & Pecoul, B. (2010) Fexinidazole--a new oral nitroimidazole drug candidate entering clinical development for the treatment of sleeping sickness. *PLoS Neglected Tropical Diseases*, **4**, e923.
- Trapnell, C., Hendrickson, D. G., Sauvageau, M., Goff, L., Rinn, J. L. & Pachter, L. (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnology*, **31**, 46-53.
- Trapnell, C., Pachter, L. & Salzberg, S. L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, **25**, 1105-11.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L. & Pachter, L. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, **7**, 562-78.
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., Salzberg, S. L., Wold, B. J. & Pachter, L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, **28**, 511-5.
- Trenaman, A., Hartley, C., Prorocic, M., Passos-Silva, D. G., Van Den Hoek, M., Nechyporuk-Zloy, V., Machado, C. R. & McCulloch, R. (2013) *Trypanosoma brucei* BRCA2 acts in a life cycle-specific genome stability process and dictates BRC repeat number-dependent RAD51 subnuclear dynamics. *Nucleic Acids Research*, **41**, 943-60.
- Trochine, A., Creek, D. J., Faral-Tello, P., Barrett, M. P. & Robello, C. (2014) Benznidazole biotransformation and multiple targets in *Trypanosoma cruzi* revealed by metabolomics. *PLoS Neglected Tropical Diseases*, **8**, e2844.

- Trochine, A., Creek, D. J., Faral-Tello, P., Barrett, M. P. & Robello, C. (2015) Bestatin induces specific changes in *Trypanosoma cruzi* dipeptide pool. *Antimicrobial Agents and Chemotherapy*, **59**, 2921-5.
- Troeberg, L., Morty, R. E., Pike, R. N., Lonsdale-Eccles, J. D., Palmer, J. T., Mckerrow, J. H. & Coetzer, T. H. (1999) Cysteine proteinase inhibitors kill cultured bloodstream forms of *Trypanosoma brucei brucei*. *Experimental Parasitology*, **91**, 349-55.
- Turner, C. M. R., Sternberg, J., Buchanan, N., Smith, E., Hide, G. & Tait, A. (1990) Evidence That the Mechanism of Gene Exchange in *Trypanosoma-Brucei* Involves Meiosis and Syngamy. *Parasitology*, **101**, 377-386.
- Ummanni, R., Mundt, F., Pospisil, H., Venz, S., Scharf, C., Baret, C., Falth, M., Kollermann, J., Walther, R., Schlomm, T., Sauter, G., Bokemeyer, C., Sultmann, H., Schuppert, A., Brummendorf, T. H. & Balabanov, S. (2011) Identification of Clinically Relevant Protein Targets in Prostate Cancer with 2D-DIGE Coupled Mass Spectrometry and Systems Biology Network Platform. *PLoS One*, **6**.
- Urbaniak, M. D., Guther, M. L. & Ferguson, M. A. (2012a) Comparative SILAC proteomic analysis of *Trypanosoma brucei* bloodstream and procyclic lifecycle stages. *PLoS One*, **7**, e36619.
- Urbaniak, M. D., Martin, D. M. & Ferguson, M. A. (2013) Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of *Trypanosoma brucei*. *Journal of Proteome Research*, **12**, 2233-44.
- Urbaniak, M. D., Mathieson, T., Bantscheff, M., Eberhard, D., Grimaldi, R., Miranda-Saavedra, D., Wyatt, P., Ferguson, M. A., Frearson, J. & Drewes, G. (2012b) Chemical proteomic analysis reveals the drugability of the kinome of *Trypanosoma brucei*. *ACS Chemical Biology*, **7**, 1858-65.
- Urbaniak, M. D., Turnock, D. C. & Ferguson, M. A. (2006) Galactose starvation in a bloodstream form *Trypanosoma brucei* UDP-glucose 4'-epimerase conditional null mutant. *Eukaryotic Cell*, **5**, 1906-13.
- Uzureau, P., Uzureau, S., Lecordier, L., Fontaine, F., Tebabi, P., Homble, F., Grelard, A., Zhendre, V., Nolan, D. P., Lins, L., Crowet, J. M., Pays, A., Felu, C., Poelvoorde, P., Vanhollebeke, B., Moestrup, S. K., Lyngso, J., Pedersen, J. S., Mottram, J. C., Dufourc, E. J., Perez-Morga, D. & Pays, E. (2013) Mechanism of *Trypanosoma brucei* gambiense resistance to human serum. *Nature*, **501**, 430-4.
- Vale, G. A., Hargrove, J. W., Cullis, N. A., Chamisa, A. & Torr, S. J. (2015) Efficacy of Electrocuting Devices to Catch Tsetse Flies (Glossinidae) and Other Diptera. *PLoS Neglected Tropical Diseases*, **9**, e0004169.
- Van Brummelen, A. C., Olszewski, K. L., Wilinski, D., Llinas, M., Louw, A. I. & Birkholtz, L. M. (2009) Co-inhibition of *Plasmodium falciparum* S-adenosylmethionine decarboxylase/ornithine decarboxylase reveals perturbation-specific compensatory mechanisms by transcriptome, proteome, and metabolome analyses. *The Journal of Biological Chemistry*, **284**, 4635-46.
- Van Hellemond, J. J., Opperdoes, F. R. & Tielens, A. G. (2005) The extraordinary mitochondrion and unusual citric acid cycle in *Trypanosoma brucei*. *Biochemical Society Transactions*, **33**, 967-71.
- Van Lanen, S. G., Kinzie, S. D., Matthieu, S., Link, T., Culp, J. & Iwata-Reuyl, D. (2003) tRNA modification by S-adenosylmethionine:tRNA ribosyltransferase-isomerase. Assay development and characterization of the recombinant enzyme. *The Journal of Biological Chemistry*, **278**, 10491-9.
- Van Valen, L. (1973) A new evolutionary law. **1**, 1-30.
- Vanhamme, L., Paturiaux-Hanocq, F., Poelvoorde, P., Nolan, D. P., Lins, L., Van Den Abbeele, J., Pays, A., Tebabi, P., Van Xong, H., Jacquet, A., Moguevsky, N., Dieu, M., Kane, J. P., De Baetselier, P., Brasseur, R. & Pays, E. (2003) Apolipoprotein L-I is the trypanosome lytic factor of human serum. *Nature*, **422**, 83-7.
- Vanhollebeke, B., De Muylder, G., Nielsen, M. J., Pays, A., Tebabi, P., Dieu, M., Raes, M., Moestrup, S. K. & Pays, E. (2008) A haptoglobin-hemoglobin receptor conveys innate immunity to *Trypanosoma brucei* in humans. *Science*, **320**, 677-81.

- Vansterkenburg, E. L., Coppens, I., Wilting, J., Bos, O. J., Fischer, M. J., Janssen, L. H. & Opperdoes, F. R. (1993) The uptake of the trypanocidal drug suramin in combination with low-density lipoproteins by *Trypanosoma brucei* and its possible mode of action. *Acta Tropica*, **54**, 237-50.
- Vassella, E., Acosta-Serrano, A., Studer, E., Lee, S. H., Englund, P. T. & Roditi, I. (2001) Multiple procyclin isoforms are expressed differentially during the development of insect forms of *Trypanosoma brucei*. *Journal of Molecular Biology*, **312**, 597-607.
- Vassella, E., Probst, M., Schneider, A., Studer, E., Renggli, C. K. & Roditi, I. (2004) Expression of a major surface protein of *Trypanosoma brucei* insect forms is controlled by the activity of mitochondrial enzymes. *Molecular Biology of the Cell*, **15**, 3986-93.
- Vassella, E., Reuner, B., Yutzy, B. & Boshart, M. (1997) Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway. *Journal of Cell Science*, **110 ( Pt 21)**, 2661-71.
- Vaughan, S. & Gull, K. (2003) The trypanosome flagellum. *Journal of Cell Science*, **116**, 757-9.
- Vedel, M., Lawrence, F., Robert-Gero, M. & Lederer, E. (1978) The antifungal antibiotic sinefungin as a very active inhibitor of methyltransferases and of the transformation of chick embryo fibroblasts by Rous sarcoma virus. *Biochemical and Biophysical Research Communications*, **85**, 371-6.
- Veenstra, T. D. (2012) Metabolomics: the final frontier? *Genome Medicine*, **4**, 40.
- Vickerman, K. (1985) Developmental cycles and biology of pathogenic trypanosomes. *British Medical Bulletin*, **41**, 105-14.
- Vickerman, K. & Luckins, A. G. (1969) Localization of variable antigens in the surface coat of *Trypanosoma brucei* using ferritin conjugated antibody. *Nature*, **224**, 1125-6.
- Vincent, I. M., Creek, D., Watson, D. G., Kamleh, M. A., Woods, D. J., Wong, P. E., Burchmore, R. J. & Barrett, M. P. (2010) A molecular mechanism for eflornithine resistance in African trypanosomes. *PLoS Pathogens*, **6**, e1001204.
- Vincent, I. M., Creek, D. J., Burgess, K., Woods, D. J., Burchmore, R. J. & Barrett, M. P. (2012) Untargeted metabolomics reveals a lack of synergy between nifurtimox and eflornithine against *Trypanosoma brucei*. *PLoS Neglected Tropical Diseases*, **6**, e1618.
- Visser, N., Opperdoes, F. R. & Borst, P. (1981) Subcellular compartmentation of glycolytic intermediates in *Trypanosoma brucei*. *European Journal of Biochemistry*, **118**, 521-6.
- Wang, C., Zhu, Y., Chen, J., Li, X., Peng, J., Chen, J., Zou, Y., Zhang, Z., Jin, H., Yang, P., Wu, J., Niu, L., Gong, Q., Teng, M. & Shi, Y. (2014a) Crystal structure of arginine methyltransferase 6 from *Trypanosoma brucei*. *PLoS One*, **9**, e87267.
- Wang, F., Ekiert, D. C., Ahmad, I., Yu, W., Zhang, Y., Bazirgan, O., Torkamani, A., Raudsepp, T., Mwangi, W., Criscitiello, M. F., Wilson, I. A., Schultz, P. G. & Smider, V. V. (2013) Reshaping antibody diversity. *Cell*, **153**, 1379-93.
- Wang, K., Zhou, Y. J., Liu, H., Cheng, K., Mao, J., Wang, F., Liu, W., Ye, M., Zhao, Z. K. & Zou, H. (2015) Proteomic analysis of protein methylation in the yeast *Saccharomyces cerevisiae*. *Journal of Proteomics*, **114**, 226-33.
- Wang, Y., Kavran, J. M., Chen, Z., Karukurichi, K. R., Leahy, D. J. & Cole, P. A. (2014b) Regulation of S-adenosylhomocysteine hydrolase by lysine acetylation. *The Journal of Biological Chemistry*, **289**, 31361-72.
- Watson, M. 2015. Opiniomics. Available from: <https://biomickwatson.wordpress.com/> [Accessed 19th January 2016].
- Webb, K. J., Lipson, R. S., Al-Hadid, Q., Whitelegge, J. P. & Clarke, S. G. (2010) Identification of protein N-terminal methyltransferases in yeast and humans. *Biochemistry*, **49**, 5225-35.
- Weir, W., Capewell, P., Foth, B. J., Clucas, C., Pountain, A., Steketee, P. C., Veitch, N., Koffi, M., De Meeus, T., Kabore, J., Camara, M., Cooper, A., Tait, A., Jamonneau, V., Bucheton, B., Berriman, M. & Macleod, A. (2016) Population genomics reveals the origin and asexual evolution of human infective trypanosomes. *eLife*, **5**, e11473.
- Weisse, S., Heddergott, N., Heydt, M., Pflasterer, D., Maier, T., Haraszti, T., Grunze, M., Engstler, M. & Rosenhahn, A. (2012) A quantitative 3D motility analysis of *Trypanosoma brucei* by use of digital in-line holographic microscopy. *PLoS One*, **7**, e37296.

- Welburn, S. C., Picozzi, K., Fevre, E. M., Coleman, P. G., Odiit, M., Carrington, M. & Maudlin, I. (2001) Identification of human-infective trypanosomes in animal reservoir of sleeping sickness in Uganda by means of serum-resistance-associated (SRA) gene. *Lancet*, **358**, 2017-9.
- Werbovetz, K. A. & Englund, P. T. (1997) Glycosyl phosphatidylinositol myristoylation in African trypanosomes. *Molecular and Biochemical Parasitology*, **85**, 1-7.
- Wertheim, H. F. L., Horby, P. & Woodall, J. P. 2012. *Atlas of Human Infectious Diseases*, Chichester, West Sussex, Wiley-Blackwell.
- Wheeler, R. J., Scheumann, N., Wickstead, B., Gull, K. & Vaughan, S. (2013) Cytokinesis in *Trypanosoma brucei* differs between bloodstream and tsetse trypomastigote forms: implications for microtubule-based morphogenesis and mutant analysis. *Molecular Microbiology*, **90**, 1339-55.
- Who (2012). The London Declaration on Neglected Tropical Diseases. London, UK.
- Who (2013). Report of a WHO meeting on elimination of African trypanosomiasis (*Trypanosoma brucei gambiense*). Geneva: World Health Organisation.
- Who. 2015a. *Global Health Observatory data repository: Human African Trypanosomiasis* [Online]. Available: <http://apps.who.int/gho/data/node.main.A1635> [Accessed 11th January, 2016].
- Who. 2015b. *Media Centre: Trypanosomiasis Fact Sheet* [Online]. Available: <http://www.who.int/mediacentre/factsheets/fs259/en/> [Accessed 11th January 2016].
- Willert, E. & Phillips, M. A. (2012) Regulation and function of polyamines in African trypanosomes. *Trends in Parasitology*, **28**, 66-72.
- Willert, E. K., Fitzpatrick, R. & Phillips, M. A. (2007) Allosteric regulation of an essential trypanosome polyamine biosynthetic enzyme by a catalytically dead homolog. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 8275-80.
- Willert, E. K. & Phillips, M. A. (2008) Regulated expression of an essential allosteric activator of polyamine biosynthesis in African trypanosomes. *PLoS Pathogens*, **4**, e1000183.
- Williams, R. A., Mottram, J. C. & Coombs, G. H. (2013) Distinct roles in autophagy and importance in infectivity of the two ATG4 cysteine peptidases of *Leishmania major*. *The Journal of Biological Chemistry*, **288**, 3678-90.
- Williams, R. A., Tetley, L., Mottram, J. C. & Coombs, G. H. (2006) Cysteine peptidases CPA and CPB are vital for autophagy and differentiation in *Leishmania mexicana*. *Molecular Microbiology*, **61**, 655-74.
- Winter, H. C. & Dekker, E. E. (1989) Specificity of aspartate aminotransferases from leguminous plants for 4-substituted glutamic acids. *Plant Physiology*, **89**, 1122-8.
- Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M. A., Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P., Amegbey, G., Block, D., Hau, D. D., Wagner, J., Miniaci, J., Clements, M., Gebremedhin, M., Guo, N., Zhang, Y., Duggan, G. E., Macinnis, G. D., Weljie, A. M., Dowlatabadi, R., Bamforth, F., Clive, D., Greiner, R., Li, L., Marrie, T., Sykes, B. D., Vogel, H. J. & Querengesser, L. (2007) HMDB: the Human Metabolome Database. *Nucleic Acids Research*, **35**, D521-6.
- Wlodarski, T., Kutner, J., Towpik, J., Knizewski, L., Rychlewski, L., Kudlicki, A., Rowicka, M., Dziembowski, A. & Ginalski, K. (2011) Comprehensive structural and substrate specificity classification of the *Saccharomyces cerevisiae* methyltransferase. *PLoS One*, **6**, e23168.
- Woodward, R. & Gull, K. (1990) Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*. *Journal of Cell Science*, **95**, 49-57.
- Wring, S., Gaukel, E., Nare, B., Jacobs, R., Bacchi, C., Beaudet, B., Bowling, T., Chen, D. T., Freund, Y., Jenks, M., Mercer, L., Noe, A., Orr, M., Parham, R., Plattner, J., Randolph, R., Rewerts, C., Sligar, J., Yarlett, N. & Don, R. (2010) Development of PK-Pd Models to Predict the Therapeutic Dose and CNS Disposition of Scyx-7158 in the Treatment of Stage 2 Human

- African Trypanosomiasis. *American Journal of Tropical Medicine and Hygiene*, **83**, 204-204.
- Wring, S., Gaukel, E., Nare, B., Jacobs, R., Beaudet, B., Bowling, T., Mercer, L., Bacchi, C., Yarlett, N., Randolph, R., Parham, R., Rewerts, C., Platner, J. & Don, R. (2014) Pharmacokinetics and pharmacodynamics utilizing unbound target tissue exposure as part of a disposition-based rationale for lead optimization of benzoxaboroles in the treatment of Stage 2 Human African Trypanosomiasis. *Parasitology*, **141**, 104-18.
- Wurst, M., Robles, A., Po, J., Luu, V. D., Brems, S., Marentije, M., Stoitsova, S., Quijada, L., Hoheisel, J., Stewart, M., Hartmann, C. & Clayton, C. (2009) An RNAi screen of the RRM-domain proteins of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, **163**, 61-5.
- Wyllie, S., Foth, B. J., Kelner, A., Sokolova, A. Y., Berriman, M. & Fairlamb, A. H. (2015) Nitroheterocyclic drug resistance mechanisms in *Trypanosoma brucei*. *The Journal of Antimicrobial Chemotherapy*.
- Xia, J. G., Sinelnikov, I. V., Han, B. & Wishart, D. S. (2015) MetaboAnalyst 3.0-making metabolomics more meaningful. *Nucleic Acids Research*, **43**, W251-W257.
- Xiao, Y., McCloskey, D. E. & Phillips, M. A. (2009) RNA interference-mediated silencing of ornithine decarboxylase and spermidine synthase genes in *Trypanosoma brucei* provides insight into regulation of polyamine biosynthesis. *Eukaryotic Cell*, **8**, 747-55.
- Xong, H. V., Vanhamme, L., Chamekh, M., Chimfwembe, C. E., Van Den Abbeele, J., Pays, A., Van Meirvenne, N., Hamers, R., De Baetselier, P. & Pays, E. (1998) A VSG expression site-associated gene confers resistance to human serum in *Trypanosoma rhodesiense*. *Cell*, **95**, 839-46.
- Xu, C. W., Hines, J. C., Engel, M. L., Russell, D. G. & Ray, D. S. (1996) Nucleus-encoded histone H1-like proteins are associated with kinetoplast DNA in the trypanosomatid *Crithidia fasciculata*. *Molecular and Cellular Biology*, **16**, 564-76.
- Yadav, M. K., Park, S. W., Chae, S. W. & Song, J. J. (2014) Sinefungin, a natural nucleoside analogue of S-adenosylmethionine, inhibits *Streptococcus pneumoniae* biofilm growth. *Biomed Research International*, **2014**, 156987.
- Yan, W., Hwang, D. & Aebersold, R. (2008) Quantitative proteomic analysis to profile dynamic changes in the spatial distribution of cellular proteins. *Methods in Molecular Biology*, **432**, 389-401.
- Yung, K. H., Yang, S. F. & Schlenk, F. (1982) Methionine synthesis from 3-methylthioribose in apple tissue. *Biochemical and Biophysical Research Communications*, **104**, 771-7.
- Zamudio, J. R., Mitra, B., Campbell, D. A. & Sturm, N. R. (2009) Hypermethylated cap 4 maximizes *Trypanosoma brucei* translation. *Molecular Microbiology*, **72**, 1100-10.
- Zamudio, J. R., Mitra, B., Foldynova-Trantirkova, S., Zeiner, G. M., Lukes, J., Bujnicki, J. M., Sturm, N. R. & Campbell, D. A. (2007) The 2'-O-ribose methyltransferase for cap 1 of spliced leader RNA and U1 small nuclear RNA in *Trypanosoma brucei*. *Molecular and Cellular Biology*, **27**, 6084-92.
- Zamudio, J. R., Mitra, B., Zeiner, G. M., Feder, M., Bujnicki, J. M., Sturm, N. R. & Campbell, D. A. (2006) Complete cap 4 formation is not required for viability in *Trypanosoma brucei*. *Eukaryotic Cell*, **5**, 905-15.
- Zeiner, G. M., Sturm, N. R. & Campbell, D. A. (2003) The *Leishmania tarentolae* spliced leader contains determinants for association with polysomes. *The Journal of Biological Chemistry*, **278**, 38269-75.
- Zeitoun-Ghandour, S., Leszczyszyn, O. I., Blindauer, C. A., Geier, F. M., Bundy, J. G. & Sturzenbaum, S. R. (2011) *C. elegans* metallothioneins: response to and defence against ROS toxicity. *Molecular Biosystems*, **7**, 2397-406.
- Zhang, B., Watts, K. M., Hodge, D., Kemp, L. M., Hunstad, D. A., Hicks, L. M. & Odom, A. R. (2011a) A second target of the antimalarial and antibacterial agent fosmidomycin revealed by cellular metabolic profiling. *Biochemistry*, **50**, 3570-7.
- Zhang, W., Li, F. & Nie, L. (2010) Integrating multiple 'omics' analysis for microbial biology: application and methodologies. *Microbiology*, **156**, 287-301.

- Zhang, W. W. & Matlashewski, G. (2015) CRISPR-Cas9-Mediated Genome Editing in *Leishmania donovani*. *MBio*, **6**, e00861.
- Zhang, Y. K., Plattner, J. J., Easom, E. E., Zhou, Y., Akama, T., Bu, W., White, W. H., Defauw, J. M., Winkle, J. R., Balko, T. W., Guo, S., Xue, J., Cao, J. & Zou, W. (2015) Discovery of an orally bioavailable isoxazoline benzoxaborole (AN8030) as a long acting animal ectoparasiticide. *Bioorganic and Medicinal Chemistry Letters*, **25**, 5589-93.
- Zhang, Y. K., Plattner, J. J., Freund, Y. R., Easom, E. E., Zhou, Y., Gut, J., Rosenthal, P. J., Waterson, D., Gamo, F. J., Angulo-Barturen, I., Ge, M., Li, Z., Li, L., Jian, Y., Cui, H., Wang, H. & Yang, J. (2011b) Synthesis and structure-activity relationships of novel benzoxaboroles as a new class of antimalarial agents. *Bioorganic and Medicinal Chemistry Letters*, **21**, 644-51.
- Zheng, L. L., Wen, Y. Z., Yang, J. H., Liao, J. Y., Shao, P., Xu, H., Zhou, H., Wen, J. Z., Lun, Z. R., Ayala, F. J. & Qu, L. H. (2013) Comparative transcriptome analysis of small noncoding RNAs in different stages of *Trypanosoma brucei*. *RNA*, **19**, 863-75.
- Zhu, K., Zhao, J., Lubman, D. M., Miller, F. R. & Barder, T. J. (2005) Protein pI shifts due to posttranslational modifications in the separation and characterization of proteins. *Analytical Chemistry*, **77**, 2745-55.
- Zomerdijk, J. C., Kieft, R. & Borst, P. (1991) Efficient production of functional mRNA mediated by RNA polymerase I in *Trypanosoma brucei*. *Nature*, **353**, 772-5.

## Appendix A. Differentiating trypanosome medium (DTM)

Component	Molarity (mM)
KCl	5.40
CaCl <sub>2</sub>	1.80
NaH <sub>2</sub> PO <sub>4</sub> .H <sub>2</sub> O	1.00
MgSO <sub>4</sub> .7H <sub>2</sub> O	0.80
Sodium pyruvate	1.00
Phenol red	0.03
Hypoxanthine	0.10
Alanine	0.10
Arginine.HCl	0.60
Asparagine	0.11
Aspartic acid	0.11
Cysteine	0.20
Glutamic acid	1.70
Glutamine	11.00
Glycine	0.11
Histidine.HCl.H <sub>2</sub> O	0.20
Isoleucine	0.40
Leucine	0.40
Lycine.HCl	0.40
Methionine	0.10
Phenylalanine	0.19
Proline	5.60
Serine	0.11
Threonine	0.40
Tryptophan	0.05
Tyrosine	0.20
Valine	0.39
NaCl	116.00
Hepes	32.00
NaHCO <sub>3</sub>	26.00
Hemin	0.01
β-mercaptoethanol	0.20
Sodium citrate	3.00
Cis-aconitate	3.00
Glycerol	0.08%
Vitamin solution (Sigma)	1%

## Appendix B. SDM-80

Component	Molarity (mM)
NaH <sub>2</sub> PO <sub>4</sub>	1.00
NaCl	116.00
MgSO <sub>4</sub>	0.80
KCl	5.40
CaCl <sub>2</sub>	1.80
NaHCO <sub>3</sub>	26.20
HEPES	30.70
MOPS	23.90
Pyruvate	4.00
L-Arginine	1.10
Glycine	0.10
L-Alanine	2.25
L-Asparagine	0.10
L-Aspartate	0.10
L-Cystine	0.10
L-Glutamate	0.09
L-Glutamine	0.46
L-Histidine	0.20
L-Isoleucine	0.40
L-Leucine	0.76
L-Lysine	0.40
L-Methionine	0.58
L-Phenylalanine	0.68
L-Serine	0.49
L-tryptophan	0.05
L-Tyrosine	0.75
L-Valine	0.40
Taurine	1.28
Mercaptoethanol	0.20
Hypoxanthine	0.10
Thymidine	0.02
Kanamycin	0.10
Hemin	0.01
Glucose	0.15
Proline	5.20
Threonine	5.90
Vitamins 100 X (Invitrogen)	1%

## Appendix C. Creek's minimal medium (CMM)

---

Component	Molarity (mM)
D-glucose	10.000
L-glutamine	1.000
L-cysteine	1.000
NaCl	77.590
CaCl <sub>2</sub> ·2H <sub>2</sub> O	1.490
KCl	4.400
MgSO <sub>4</sub> ·7H <sub>2</sub> O	0.814
NaHCO <sub>3</sub>	35.950
HEPES	25.032
Phenol Red	0.042
Bathocuproinedisulfonic acid	0.052
β-mercaptoethanol	0.192

## Appendix D. Buffers and reagents

---

### TAE buffer, 50×

- 2 M Tris
- 6% Acetic acid
- 50 mM EDTA

### PBS, 20×

- 2.74 M NaCl
- 200 mM Na<sub>2</sub>HPO<sub>4</sub>
- 54 mM KCl
- 36 mM KH<sub>2</sub>PO<sub>4</sub>

### PBS-T

- 1× PBS
- 0.05% Tween

### *T. brucei* transfection buffer

- 90 mM Sodium phosphate
- 5 mM Potassium chloride
- 0.15 mM Calcium chloride
- 50 mM HEPES
- pH = 7.3

### Protein extraction lysis buffer

- 2% SDS
- 50 mM Tris, pH 7.5
- 1 mM EDTA
- cOmplete™ ULTRA protease inhibitor cocktail (Roche)

### Hypotonic lysis buffer

- 10 mM Tris base, pH 7.5

### Protein sample loading buffer, 4×

- 200 mM Tris-HCl, pH 6.8
- 8% SDS
- 40% glycerol
- 4% β-mercaptoethanol
- 50 mM EDTA
- 0.08% bromophenol blue

**SDS-PAGE gel for 2D-DiGE**

- 12.5% acrylamide
- 0.4 M Tris, pH 8.8
- 0.1 SDS
- 0.04% APS
- 0.05% TEMED

**DiGE lysis buffer**

- 6 M Urea
- 2 M Thiourea
- 4% CHAPS
- 25 mM Tris-base

**DiGE rehydration buffer**

- 6 M Urea
- 2 M Thiourea
- 4% CHAPS
- 0.5% IPG buffer
- 65 mM DTT
- 0.001% bromophenol blue

**DARTS lysis buffer**

- 150 mM NaCl
- 0.5% Triton X-100
- 1× PBS

**Coomassie staining solution**

- 1 mM Coomassie brilliant blue
- 50% methanol
- 10% glacial acetic acid
- 40% ddH<sub>2</sub>O

**Colloidal Coomassie staining solution**

- 10% ammonium sulfate
- 0.1% Coomassie G-250
- 3% ortho-phosphoric acid
- 20% ethanol

**Protein gel destaining solution**

- 50% methanol
- 10% glacial acetic acid
- 40% ddH<sub>2</sub>O

**SDS running buffer, 20× (Life Technologies)**

- 50 mM MOPS
- 50 mM Tris-base
- 0.1% SDS
- 1 mM EDTA
- pH 7.7

**Western blot transfer buffer, 10×**

- 1.92 M Glycine
- 250 mM Tris-Base

**Western blot transfer buffer, 1×**

- 192 mM Glycine
- 25 mM Tris-base
- 10% methanol

**Western blot blocking solution**

- PBS-T
- 5% powdered milk

**Metabolite extraction solvent**

- 60% methanol
- 20% chloroform
- 20% ddH<sub>2</sub>O

**DEPC-treated H<sub>2</sub>O**

- 0.1% diethylpyrocarbonate in ddH<sub>2</sub>O
- Overnight incubation, followed by autoclaving

## Appendix E. Primers & Plasmids

GeneID	Description	Sequence	Restriction sites	Identifier
Tb427.10.3080	pURAN overexpression, fwd	GCGAATTCATGGGGATAGTCCGAGTACA	EcoRI	MB0867
"	pURAN overexpression, rev	GCGAATTCCTTATTCCGGCTCCTGTAAAC	EcoRI	MB0931
Tb427.10.7560	pURAN overexpression, fwd	GCGAATTCATGAGTGCCCTACTGAAAG	EcoRI	MB0869
"	pURAN overexpression, rev	GCGAATTCCTACTGTTCAACCCTCTTAC	EcoRI	MB0870
Tb927.10.9020	pURAN overexpression, fwd	GCGAATTCATGATGGGTGATGCTGGTAG	EcoRI	MB0871
"	pURAN overexpression, rev	GCGAATTCCTACGAGTCTGAAAGATTC	EcoRI	MB0872
Tb427tmp.01.4830	pURAN overexpression, fwd	GCGAATTCATGAGCGAGAATGACATTGC	EcoRI	MB0873
"	pURAN overexpression, rev	GCGAATTCCTAATTCAGTTTCACTGGCA	EcoRI	MB0874
Tb427tmp.01.1460	pURAN overexpression, fwd	GCGAATTCATGCTAGACAAAGTTGTAGG	EcoRI	MB0875
"	pURAN overexpression, rev	GCGAATTCCTACCCGGACAGTAAGAGGAA	EcoRI	MB0876
Tb427.03.2890	pURAN overexpression, fwd	GCGAATTCATGTGCGAAGCAAAAGAATGT	EcoRI	MB0877
"	pURAN overexpression, rev	GCGAATTCCTAATCGGTCCACGTGTCATGC	EcoRI	MB0878
Tb427.04.1900	pURAN overexpression, fwd	GCGAATTCATGCCCAACACGTCCCGA	EcoRI	MB0879
"	pURAN overexpression, rev	GCGAATTCCTAAAAACGGGGTGCCTGCC	EcoRI	MB0880

Table E-1: Primers used in this study

GeneID	Description	Sequence	Restriction sites	Identifier
Tb427.05.2050	pURAN overexpression, fwd	G <u>CGAATTC</u> ATGCTCCAATGGGTCGTGAA	EcoRI	MB0881
“	pURAN overexpression, rev	G <u>CGAATTC</u> CTATCGAGTCCGCTCTCGCT	EcoRI	MB0882
“	Mutagenesis primer, fwd	GAATATTGAGTTAAATAAGGAGAGGATACTCTCACGAAGCAATGG	n/a	n/a
“	Mutagenesis primer, rev	CCATTGCTTCGTGAGAGTATCCTCTCCTTATTTAACTCAATATTC	n/a	n/a
Tb427.07.4320	pURAN overexpression, fwd	G <u>CGAATTC</u> ATGCTGCAACGTATTCCACC	EcoRI	MB0883
“	pURAN overexpression, rev	G <u>CGAATTC</u> CTACTTCACCGTCAGGAACT	EcoRI	MB0884
Tb427.08.5040	pURAN overexpression, fwd	G <u>CGAATTC</u> ATGGATTACCGACAGAAAGC	EcoRI	MB0885
“	pURAN overexpression, rev	G <u>CGAATTC</u> CTAACACGCCTCGTTATCCT	EcoRI	MB0886
Tb427.06.2270	pURAN overexpression, fwd	G <u>CGAATTC</u> ATGCCCGCACGTGGCTGTGC	EcoRI	MB0887
“	pURAN overexpression, rev	G <u>CGAATTC</u> TTAGCGTAGTGCGTACATCT	EcoRI	MB0888
Tb427.10.7850	pURAN overexpression, fwd	G <u>CGAATTC</u> ATGGCATGGACATGTTCCGG	EcoRI	MB0927
“	pURAN overexpression, rev	G <u>CGAATTC</u> CTAAATACCTCCCCTTTGGT	EcoRI	MB0928
Tb427tmp.01.5300 (ODC)	pURAN overexpression, fwd	G <u>CGAATTC</u> ATGACCACCAATCAACCCC	EcoRI	MB0929
“	pURAN overexpression, rev	G <u>CGAATTC</u> TTATGATTTTTGACTTTTCA	EcoRI	MB0930
n/a	pURAN construct integration 5’	GTCAATACAACACACAATAGG	n/a	n/a
n/a	pRM482 construct integration 5’	GAGCTAGTGAGATCAACAGTAC	n/a	MB0925

Table E-1: (continued)

GeneID	Description	Sequence	Restriction sites	Identifier
Tb927.7.2080 (TbCgm1)	RNAi fwd	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCAGTTCCTCAGAAGCAAG	n/a	MB0783
“	RNAi rev	GGGGACCACTTTGTACAAGAAAGCTGGGTGCGAAGTGACCGAAAGAAAG	n/a	MB0784
Tb927.10.7940 (TbMtr1)	RNAi fwd	GGGGACAAGTTTGTACAAAAAAGCAGGCTAGGGCACCTTTACGTTTGTG	n/a	MB0785
“	RNAi rev	GGGGACCACTTTGTACAAGAAAGCTGGGTGCGGAGAAGCAGTCGAATAG	n/a	MB0786

**Table E-2: RNAi primers used in this study**

Marker	Name	Sequence	Chromosome location (TREU 927 v5.1)	PCR product size
TB10/14	CRAM, fwd	AACTCCCTCCGATCGATCACAAC	1,805,332 – 1,805,355	3,319 bp
	CRAM, rev	CTGCTGATGCCGTACATGATGATTC	1,808,726 – 1,808,751	
TB10/28	3778, fwd	CAACAGTGCCCGGAGGAGGTACG	2,404,358 – 2,404,380	509 bp
	3778, rev	GCGGTGACTCGTCCGTCTTCTTGC	2,404,907 – 2,404,930	

**Table E-3: Genotyping primers used in chapter 5 in this study**

GeneID	Description	Sequence	Identifier
Tb927.7.2080	TbCgm1 – Fwd primer	AGCAAGACGCAGCTCAAGTG	n/a
“	TbCgm 1 – Rev primer	GGCTCCAGTCGACGGAAA	n/a
Tb427.10.3080	MT 1 – Fwd primer	GCTGTTGGGTGCCGTGTT	n/a
“	MT 1 – Rev primer	CTGCATCGGCGCTGTCT	n/a
Tb427.10.7560	MT 2 – Fwd primer	CGGTCGCACCGAACTTCT	MB 0992
“	MT 2 – Rev primer	CCGTTGCGAAACGCTGTA	MB 0993
Tb927.10.9020	MT 3 – Fwd primer	CGGCGCAACCCAACAT	MB 0994
“	MT 3 – Rev primer	TTCCCGGACGTTCTCTATGC	MB 0995
Tb427tmp.01.4830	MT 4 – Fwd primer	CCGGGCGTTAACGTTGAA	MB 0996
“	MT 4 – Rev primer	CCACCTGGTATTGCGTATTGC	MB 0997
Tb427tmp.01.1460	MT 5 – Fwd primer	ACCCCGAGGGTGACTGGTA	MB 0998
“	MT 5 – Rev primer	CCCCAGCCGTGGTGTTT	MB 0999
Tb427.03.2890	MT 6 – Fwd primer	GACTACGTGCGCGTAAGC	MB 1000
“	MT 6 – Rev primer	ACACCGTACCGTCTTTGTTCTTC	MB 1001
Tb427.04.1900	MT 7 – Fwd primer	TTGCGAGAGTGTGGCCATAC	MB 1002
“	MT 7 – Rev primer	TGCCATTCCGAGCATATCC	MB 1003
Tb427.05.2050	MT 8 – Fwd primer	GGGACCAGAGGGAAACGAA	MB 1004
“	MT 8 – Rev primer	CCCCGCATCCGCAAT	MB 1005
Tb427.07.4320	MT 9 – Fwd primer	CCCTTCGCCACCAATACG	MB 1006
“	MT 9 – Rev primer	TGCCGACGCTGCTGTGT	MB 1007
Tb427.08.5040	MT 10 – Fwd primer	CAGGATCCCAGCGGAATG	MB 1008
“	MT 10 – Rev primer	TCAAGTGAAGCAACGCTTGGT	MB 1009
Tb427.06.2270	MT 11 – Fwd primer	AATGGGCGGCGATCTATCT	MB 1010
“	MT 11 – Rev primer	TGCAGCGTACCAAGAAAGCA	MB 1011
Tb427.10.7850	MT 12 – Fwd primer	GGCCTTTCGTCGTGTTGATC	MB 1012
“	MT 12 – Rev primer	TTCGGGTGGCGAATGC	MB 1013
Tb427tmp.01.5300	TbODC – Fwd primer	TCTCGGCACGGATTGTA	MB 1014
“	TbODC – Rev primer	CCAATGCCTCTCACACGTTGT	MB 1015
Tb927.10.8440	THT-1 – Fwd primer	TCGAGAATCCCAAATGTAGTGAAG	MB 913
“	THT-1 – Rev primer	GGCCACCTCGTTTTTGA	MB 914
Tb927.10.8490	THT-2 – Fwd primer	ACGGATTTTCCATCGGCTTT	MB 917
“	THT-2 – Rev primer	TGAACAATTCGTGCGCACTTC	MB 918
TERT	Endogenous ctrl, fwd	GAGCGTGTGACTTCCGAAGG	MB 0864
“	Endogenous ctrl, rev	AGGAACTGTCACGGAGTTTGC	MB 0865

Table E-4: real-time (RT) PCR primers used in this study.

GeneID	Description	Parent
Tb427.10.3080	Over-expressor for $\Delta 3080::NEO$	pURAN
Tb427.10.7560	Re-expressor for $\Delta 7560::NEO$	pRM482
Tb927.10.9020	Over-expressor for $\Delta 9020::NEO$	pURAN
Tb427tmp.01.4830	Over-expressor for $\Delta 4830::NEO$	pURAN
Tb427tmp.01.1460	Over-expressor for $\Delta 1460::NEO$	pURAN
Tb427.03.2890	Over-expressor for $\Delta 2890::NEO$	pURAN
Tb427.04.1900	Over-expressor for $\Delta 1900::NEO$	pURAN
Tb427.05.2050	Re-expressor for $\Delta 2050::NEO$	pRM482
Tb427.07.4320	Over-expressor for $\Delta 4320::NEO$	pURAN
Tb427.08.5040	Over-expressor for $\Delta 5040::NEO$	pURAN
Tb427.06.2270	Over-expressor for $\Delta 2270::NEO$	pURAN
Tb427.10.7850	Over-expressor for $\Delta 7850::NEO$	pURAN
Tb427tmp.01.5300	Over-expressor for $\Delta ODC::NEO$	pURAN
Tb927.7.2080 - TbCgm1	TbCgm1 <sup>RNAi</sup>	pGL2084
Tb927.10.7940 - TbMtr1	TbMtr1 <sup>RNAi</sup>	pGL2084

**Table E-5: Plasmids used in this study**

## Appendix F. PCR and IEF cycling conditions

### High-fidelity polymerase PCR (HF Phusion, New England Biolabs)

	Temp.	Time	
1	98°C	0:30	
2	98°C	0:15	34x
	65°C	0:30	
	72°C	1:00 per kb	
	72°C	5:00	
3	72°C	5:00	
4	10°C	HOLD	

### Colony screen/construct integration PCR (GoTaq polymerase, promega)

	Temp.	Time	
1	95°C	2:00	
2	95°C	0:30	32x
	55°C	0:30	
	72°C	1:00 per kb	
	72°C	5:00	
3	72°C	5:00	
4	10°C	HOLD	

### Genotyping PCR – 3778/CRAM

	Temp.	Time	
1	95°C	0:50	30x
	64°C	0:50	
	65°C	0:50 - 3778	
		3:00 - CRAM	

### Real-time PCR (ABI 7500 system)

	Temp.	Time	
1	95°C	10:00	
2	95°C	0:15	40x
	60°C	1:00	

### IEF cycling conditions

	Voltage (V)	Time (hours)	Notes
1	30	10	Step 'n' hold
2	300	2	Step 'n' hold
3	600	2	Gradient
4	1000	2	Gradient
5	8000	1	Gradient
6	8000	9	Step 'n' hold
7	1000	∞	Step 'n' hold