DNA Interstrand Crosslink Repair in Trypanosoma brucei
Kumar, Ambika

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

For additional information about this publication click this link.
http://qmro.qmul.ac.uk/xmlui/handle/123456789/36675

Information about this research object was correct at the time of download; we occasionally make corrections to records, please therefore check the published record when citing. For more information contact scholarlycommunications@qmul.ac.uk
DNA Interstrand Crosslink Repair in *Trypanosoma brucei*

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

Ambika Kumar

**Supervisor:** Dr Shane Wilkinson

School of Biological & Chemical Sciences

Queen Mary University of London
Statement of originality

I, Ambika Kumar, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below, and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party’s copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature: 

Date: 24/01/2018

Details of publications:

Acknowledgements

I would like to thank Dr Shane Wilkinson for coming up with the initial ideas for this project as well as his advice and support along the way. I hope all the stress and ‘stupid’ questions were worth it! I am also very grateful to members of the John Kelly and Sam Alsford laboratories at the London School of Hygiene and Tropical Medicine for their guidance and resources which made it possible for me to carry out this project as well as Dr Emma Louise Meredith for teaching me basic molecular biology and cell culturing techniques. Special thanks to all other research laboratories which have provided antibodies, cell lines and compounds for this project, in particular the McCulloch lab at the University of Glasgow. This gratitude is extended out to Queen Mary University of London for their financial support.

Furthermore, a massive thank you to all members of the 4th floor, both present and past, for their positivity during all the highs and lows of this project and willingness to have tea (and biscuits) at a moment’s notice. I’m so glad I decided to do my PhD in such a great department with so many wonderful and talented people (too many to name!). I would also like to appreciate all the love and support from my family, including my parents, brothers and sisters who may not have understood what I was doing but always appreciated how committed I was to this project. Moreover, a huge thank you to everyone at Dongola, GradSoc and Cambridge House for being so understanding and making sure I always had something to laugh about at the end of a long, hard day (despite my constant nagging, unreasonable requests and mood swings). On that note, special thanks to Inva Hoti (soon to be Dr. Hoti) for making me realise exercise and gossiping can help deal with stress just as much as a good laugh!

Finally, I wouldn’t have got this far without my amazing, supportive fiancé Abhishek who has always been there for me, regardless of any challenges. I don’t know how to thank you for everything you’ve done for me, but I hope I’ve made you proud.
Abstract

Genomes are constantly challenged by agents that promote DNA damage, with interstrand crosslinks (ICLs) representing a particularly dangerous lesion. Ongoing work in the Wilkinson laboratory aimed at identifying novel agents that target *Trypanosoma brucei*, the causative agent of African trypanosomiasis, identified several prodrugs that once activated form ICLs in this protozoan parasite. To understand the complexity of ICL repair systems that *T. brucei* employs to resolve such damage, a variety of null mutant lines were generated that lack activities postulated to fix such lesions. Phenotypic screens using various DNA damaging agents revealed that TbMRE11, TbEXO1, TbCSB, TbCHL1, TbFAN1, TbBRCA2 and TbRAD51 all help to resolve ICLs, implicating components of the homologous recombination, nucleotide excision repair and mismatch repair pathways in resolving this form of damage: This approach demonstrated that components of the translesion synthesis pathway (TbREV2 and TbREV3) do not play a significant role in ICL repair. In many organisms, nucleases belonging to the SNM1/PSO2 family play a key and specific role in the repair of ICLs with this property extending to the *T. brucei* homologue, TbSNM1. To assess whether there is a functional linkage between the DNA repair factors noted above and TbSNM1, a series of double null mutants were constructed and the susceptibility of these lines to ICL inducing agents determined. Identification of their epistatic/non-epistatic interactions revealed that *T. brucei* expresses at least two ICL repair systems with one pathway involving the concerted activities of TbSNM1/TbCSB/TbEXO1, that we postulate functions to repair ICLs encountered by the transcriptional machinery, while the other is centred upon TbMRE11/TbFAN1/TbEXO1 that may help resolve lesions which cause stalling of DNA replication forks. By unravelling how *T. brucei* repairs ICLs, specific inhibitors against key components of these pathways could be developed and used in combination with DNA damaging agents to target trypanosomal infections.
Table of Contents

Statement of originality........................................................................................................................................2
Acknowledgements.............................................................................................................................................3
Abstract..........................................................................................................................................................4
List of figures...................................................................................................................................................8
List of tables....................................................................................................................................................10
Abbreviations..................................................................................................................................................12

1. Introduction..................................................................................................................................................14
1.1 Trypanosomatids.........................................................................................................................................14
1.2 Human African trypanosomiasis.............................................................................................................15
  1.2.1 Distribution of disease ......................................................................................................................16
1.3 Causative agent .........................................................................................................................................17
  1.3.1 Parasite life cycle .............................................................................................................................18
  1.3.2 Staging and pathologies ...................................................................................................................19
  1.3.3 Diagnosis ........................................................................................................................................20
1.4 Anti-trypanosomal therapy....................................................................................................................21
1.5 DNA damage and repair .......................................................................................................................24
  1.5.1 General overview .............................................................................................................................24
  1.5.2 Trypanosomal DNA repair pathways ...............................................................................................26
1.6 Interstrand crosslinks and their repair....................................................................................................29
  1.6.1 ICLs: definition, formation and significance ....................................................................................29
  1.6.2 ICL repair in E. coli ..........................................................................................................................32
  1.6.3 S. cerevisiae ICL repair systems .......................................................................................................34
  1.6.4 Mammalian ICL repair .....................................................................................................................36
1.7 ICL repair in T. brucei ............................................................................................................................41
1.8 Research Aims...........................................................................................................................................43

2. Material and Methods ..............................................................................................................................44
2.1 Cell culturing and storage ......................................................................................................................44
  2.1.1 Maintenance of bacterial strains .......................................................................................................44
  2.1.2 Maintenance of Trypanosoma brucei brucei ......................................................................................44
2.2 Trypanocidal compounds and DNA damaging agents .........................................................................45
  2.2.1 Plasmids ...........................................................................................................................................45
2.3 Nucleic Acid Extractions ......................................................................................................................46
2.4 Nucleic acid manipulation ....................................................................................................................46
  2.4.1 cDNA synthesis ...............................................................................................................................46
  2.4.2 DNA amplification ...........................................................................................................................46
  2.4.3 Reverse transcription quantitative PCR .............................................................................................47
  2.4.4 Site directed mutagenesis ...............................................................................................................48
  2.4.5 Fractionation of DNA by conventional agarose gel electrophoresis ...............................................51
  2.4.6 Restriction digestion .........................................................................................................................51
  2.4.7 DNA purification ............................................................................................................................52
  2.4.8 DNA Ligation ....................................................................................................................................52
  2.4.9 Construction of gene deletion/disruption vectors ...........................................................................52
  2.4.10 DNA Sequencing ............................................................................................................................53
2.5 Bioinformatics............................................................................................................................................53
2.6 Bacterial Transformation .......................................................................................................................55
2.7 T. brucei nucleofection ........................................................................................................................55
2.8 Analysis of recombinant parasites .......................................................................................................58
  2.8.1 Cumulative growth assays ...............................................................................................................58
8.1.2 SNM1-independent ICL repair system ................................................................. 152
8.1.3 Summary ........................................................................................................... 157
8.2 TbSNM1 structure and function .......................................................................... 158
8.3 Cytotoxic effects of trypanocidal compounds ..................................................... 162
8.4 An unexpected journey ......................................................................................... 162
9. Future work ........................................................................................................... 164
Thesis summary ........................................................................................................ 166
Appendix 1: Stocks .................................................................................................... 167
Appendix 2: Validation of single null mutant lines ..................................................... 169
Appendix 3: Validation of double null lines ............................................................... 173
Appendix 4: Validation of lines lacking a trypanosomal FA factor .............................. 180
Appendix 5: Structure and function studies on TbSNM1 .......................................... 184
Appendix 6: Validation of Tbntr expressing lines ...................................................... 188
Appendix 7: Published papers ................................................................................... 190
References ............................................................................................................... 228
# List of figures

<table>
<thead>
<tr>
<th>Figure number</th>
<th>Title of figure</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Typical structure of a trypanosomatid cell</td>
<td>15</td>
</tr>
<tr>
<td>1.2</td>
<td>Estimated number of human African trypanosomiasis cases</td>
<td>18</td>
</tr>
<tr>
<td>1.3</td>
<td><em>T. brucei</em> life cycle</td>
<td>20</td>
</tr>
<tr>
<td>1.4</td>
<td>Progression of HAT in a human host</td>
<td>21</td>
</tr>
<tr>
<td>1.5</td>
<td>Structures of different clinically relevant trypanocidal agents</td>
<td>23</td>
</tr>
<tr>
<td>1.6</td>
<td>DNA damage and repair pathways</td>
<td>26</td>
</tr>
<tr>
<td>1.7</td>
<td>Structures of different types of ICL inducing agents.</td>
<td>31</td>
</tr>
<tr>
<td>1.8</td>
<td>ICL repair systems in <em>E. coli</em></td>
<td>34</td>
</tr>
<tr>
<td>1.9</td>
<td>ICL repair systems in yeast</td>
<td>36</td>
</tr>
<tr>
<td>1.10</td>
<td>ICL repair system in transcriptionally active mammalian cells</td>
<td>38</td>
</tr>
<tr>
<td>1.11</td>
<td>ICL repair in dividing mammalian cells</td>
<td>41</td>
</tr>
<tr>
<td>1.12</td>
<td>Comparison of domain structure in SNM1/PSO2 nucleaseles.</td>
<td>42</td>
</tr>
<tr>
<td>1.13</td>
<td>Overall structure of SNM1A</td>
<td>43</td>
</tr>
</tbody>
</table>

## Chapter 2: Material and Methods

| 2.1           | Single Primer Reactions In Parallel Site Directed Mutagenesis (SPRINP-SDM)      | 50          |
| 2.2           | Cloning schematic to create disruption vectors                                 | 55          |

## Chapter 3: Evaluating the extent of the *T. brucei* ICL REPAIRtoire

| 3.1           | Generation of *T. brucei* line lacking TbCSB                                   | 66          |
| 3.2           | Cumulative growth of *T. brucei* null mutants                                   | 68          |
| 3.3           | Cell cycle analysis of *T. brucei* null mutants                                 | 70          |
| 3.4           | Susceptibility of *T. brucei* Tbmre11 null mutant towards various DNA damaging treatments | 71          |
| 3.5           | Comparison of *T. brucei* null mutant EC<sub>50</sub> values                   | 73          |
| 3.6           | Susceptibility of *T. brucei* null mutants towards mechloretamine and DFMO     | 75          |
| 3.7           | Comparison of *T. brucei* null mutant EC<sub>50</sub> values.                  | 77          |
| 3.8           | γH2A formation in mechloretamine treated wild type and null lines              | 79          |
| 3.9           | Comparison of γH2A signals in mechloretamine treated *T. brucei* null mutants  | 80          |
| 3.10          | Comparison of γH2A formation in mechloretamine treated *T. brucei* null mutants | 81          |

## Chapter 4: Studying the interplay of ICL repair enzymes

| 4.1           | Validating the *T. brucei* Tbsnm1Δ TbcshbΔ null mutant line.                   | 85          |
| 4.2           | Cumulative growth of *T. brucei* double mutant lines.                          | 87          |
| 4.3           | Susceptibility of *T. brucei* Tbsnm1Δ Tbmre11Δ null line towards various DNA damaging treatments | 88          |
| 4.4           | Comparison of *T. brucei* null mutant EC<sub>50</sub> values                  | 90          |
| 4.5           | Susceptibility of *T. brucei* null mutants towards mechloretamine and DFMO     | 92          |
| 4.6           | Comparison of *T. brucei* double null mutant EC<sub>50</sub> values            | 94          |
| 4.7           | Susceptibility of *T. brucei* double null mutants towards mechloretamine and DFMO | 95          |
| 4.8           | Comparison of *T. brucei* null mutant EC<sub>50</sub> values                  | 96          |
### Chapter 5: The trypanosomal Fanconi Anaemia system

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Identification of putative components of the trypanosomal FA pathway</td>
<td>100</td>
</tr>
<tr>
<td>5.2</td>
<td>Domain and sequence analysis of TbHEL1</td>
<td>101</td>
</tr>
<tr>
<td>5.3</td>
<td>Phylogenetic tree analysis of FANCM sequences</td>
<td>102</td>
</tr>
<tr>
<td>5.4</td>
<td>Susceptibility of <em>T. brucei</em> Tbhel1-deficient null lines towards various DNA damaging treatments</td>
<td>103</td>
</tr>
<tr>
<td>5.5</td>
<td>Domain and sequence analysis of TbHEL1</td>
<td>105</td>
</tr>
<tr>
<td>5.6</td>
<td>Phylogenetic tree analysis of TbHEL1 and DICER sequences</td>
<td>106</td>
</tr>
<tr>
<td>5.7</td>
<td>Confirming the deletion of Tbhel1</td>
<td>107</td>
</tr>
<tr>
<td>5.8</td>
<td>Analysing the role of TbHEL1 in RNA interference</td>
<td>108</td>
</tr>
<tr>
<td>5.9</td>
<td>Phylogenetic analysis of FANCJ sequences</td>
<td>109</td>
</tr>
<tr>
<td>5.10</td>
<td>Domain and sequence comparisons of TbCHL1</td>
<td>110</td>
</tr>
<tr>
<td>5.11</td>
<td>Susceptibility of Tbchl1 null lines towards DNA damaging treatments</td>
<td>112</td>
</tr>
<tr>
<td>5.12</td>
<td>Sequence domain analysis of FAN1 sequences</td>
<td>114</td>
</tr>
<tr>
<td>5.13</td>
<td>Rooted phylogenetic tree analysis of FAN1 sequences</td>
<td>115</td>
</tr>
<tr>
<td>5.14</td>
<td>Phenotypic analysis of <em>T. brucei</em> lines lacking <em>fan1</em></td>
<td>116</td>
</tr>
<tr>
<td>5.15</td>
<td>Susceptibility screening of <em>T. brucei</em> <em>Tbfan1</em> double null lines against mechlorethamine and DFMO.</td>
<td>118</td>
</tr>
<tr>
<td>5.16</td>
<td>Susceptibility screening of <em>Tbbrca2Δ</em> cells and <em>Tbrad51Δ</em> cells against mechlorethamine.</td>
<td>120</td>
</tr>
</tbody>
</table>

### Chapter 6: Structure and function analysis of TbSNM1

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Conserved residues of TbSNM1</td>
<td>122</td>
</tr>
<tr>
<td>6.2</td>
<td>Predicted model of TbSNM1</td>
<td>123</td>
</tr>
<tr>
<td>6.3</td>
<td>Residues in TbSNM1 postulated to bind the metal ion</td>
<td>124</td>
</tr>
<tr>
<td>6.4</td>
<td>Validating engineered <em>T. brucei</em> lines</td>
<td>126</td>
</tr>
<tr>
<td>6.5</td>
<td>Confirming expression of <em>gfp</em>-Tbsnm1</td>
<td>127</td>
</tr>
<tr>
<td>6.6</td>
<td>Localisation of mutant GFP-TbSNM1 in bloodstream form <em>T. brucei</em></td>
<td>128</td>
</tr>
<tr>
<td>6.7</td>
<td>Susceptibility of <em>T. brucei</em> lines towards mechlorethamine</td>
<td>129</td>
</tr>
<tr>
<td>6.8</td>
<td>Structure of wild type and mutated forms of TbSNM1</td>
<td>130</td>
</tr>
<tr>
<td>6.9</td>
<td>Location of V518 residue</td>
<td>131</td>
</tr>
<tr>
<td>6.10</td>
<td>Susceptibility of <em>T. brucei</em> lines towards mechlorethamine</td>
<td>131</td>
</tr>
<tr>
<td>6.11</td>
<td>Effect of the V518A mutation on TbSNM1</td>
<td>132</td>
</tr>
<tr>
<td>6.12</td>
<td>Predicted NLS of TbSNM1</td>
<td>133</td>
</tr>
<tr>
<td>6.13</td>
<td>Localisation of <em>T. brucei</em> mutant GFP-TbSNM1</td>
<td>134</td>
</tr>
<tr>
<td>6.14</td>
<td>Susceptibility of <em>T. brucei</em> conditional null lines against mechlorethamine.</td>
<td>135</td>
</tr>
</tbody>
</table>

### Chapter 7: DNA damage and trypanocidal compounds

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>Susceptibility screening of <em>T. brucei</em> against inhibitors of HaSNM1A</td>
<td>138</td>
</tr>
<tr>
<td>7.2</td>
<td>Susceptibility of <em>T. brucei</em> null lines towards anti-trypanosomal agents.</td>
<td>139</td>
</tr>
<tr>
<td>7.3</td>
<td>Susceptibility of <em>T. brucei</em> lines expressing altered levels of TbEXO1, TbMRE11 and/or TbNTR to benznidazole.</td>
<td>142</td>
</tr>
</tbody>
</table>

### Chapter 8: Discussion

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1</td>
<td>Current model for trypanosomal ICL repair.</td>
<td>158</td>
</tr>
</tbody>
</table>
List of tables

<table>
<thead>
<tr>
<th>Table number</th>
<th>Title of table</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Comparisons between <em>T.b rhodesiense</em> and <em>T.b gambiense</em> strains</td>
<td>19</td>
</tr>
<tr>
<td>1.2</td>
<td>FA Repair factors</td>
<td>39</td>
</tr>
<tr>
<td>2.1</td>
<td><em>Trypanosoma brucei</em> strains used in this project.</td>
<td>45</td>
</tr>
<tr>
<td>2.2</td>
<td>Primers used to create single point mutations in pRPα-GFP-TbSNM1 plasmid</td>
<td>51</td>
</tr>
<tr>
<td>2.3</td>
<td>Restriction enzymes, their restriction sites and corresponding buffer used in this study</td>
<td>52</td>
</tr>
<tr>
<td>2.4</td>
<td>Schematic for interruption/deletion of all genes analysed in this study</td>
<td>57</td>
</tr>
<tr>
<td>2.5</td>
<td>Parasite lines lacking one DNA repair enzyme</td>
<td>58</td>
</tr>
<tr>
<td>2.6</td>
<td>Parasites lacking two DNA repair enzymes</td>
<td>58</td>
</tr>
<tr>
<td>2.7</td>
<td>Parasite lines expressing mutated version of Tbsnm1 in a Tbsnm1-deficient background</td>
<td>59</td>
</tr>
<tr>
<td>2.8</td>
<td>Tbntr expressing lines generated in this study</td>
<td>59</td>
</tr>
<tr>
<td>3.1</td>
<td>Susceptibility of <em>T. brucei</em> null lines to DNA damaging treatments</td>
<td>72</td>
</tr>
<tr>
<td>3.2</td>
<td>Null mutants towards mechlorethamine and DFMO</td>
<td>76</td>
</tr>
<tr>
<td>4.1</td>
<td>Susceptibility of <em>T. brucei</em> double null lines to DNA damaging treatments</td>
<td>89</td>
</tr>
<tr>
<td>4.2</td>
<td>Susceptibility of double null mutants towards mechlorethamine and DFMO</td>
<td>93</td>
</tr>
<tr>
<td>4.3</td>
<td>Susceptibility of null mutants towards mechlorethamine and DFMO</td>
<td>96</td>
</tr>
<tr>
<td>5.1</td>
<td>Susceptibility of <em>T. brucei</em> Tbhel1 null lines to DNA damaging treatments.</td>
<td>104</td>
</tr>
<tr>
<td>5.2</td>
<td>Susceptibility of <em>T. brucei</em> Tbchl1 null lines to DNA damaging treatments.</td>
<td>113</td>
</tr>
<tr>
<td>5.3</td>
<td>Susceptibility of <em>T. brucei</em> Tbfan1 null line to DNA damaging treatments.</td>
<td>117</td>
</tr>
<tr>
<td>5.4</td>
<td>Susceptibility of <em>T. brucei</em> null lines to mechlorethamine and DFMO</td>
<td>119</td>
</tr>
<tr>
<td>5.5</td>
<td>Susceptibility of <em>T. brucei</em> null lines to mechlorethamine.</td>
<td>120</td>
</tr>
<tr>
<td>6.1</td>
<td>Susceptibility of <em>T. brucei</em> null lines to mechlorethamine</td>
<td>129</td>
</tr>
<tr>
<td>6.2</td>
<td>Susceptibility of <em>T. brucei</em> lines to mechloroethamine</td>
<td>132</td>
</tr>
<tr>
<td>6.3</td>
<td>Susceptibility of <em>T. brucei</em> null lines to mechloroethamine</td>
<td>135</td>
</tr>
<tr>
<td>7.1</td>
<td>Susceptibility of <em>T. brucei</em> lines to inhibitors of HaSNM1A</td>
<td>138</td>
</tr>
<tr>
<td>7.2</td>
<td>Susceptibility of <em>T. brucei</em> double null lines to anti-trypanosomal agents</td>
<td>140</td>
</tr>
</tbody>
</table>
7.3 Data represents EC₅₀ values of genetically engineered lines against benznidazole and DFMO

<table>
<thead>
<tr>
<th>Chapter 8: Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1 Summary of data presented for all trypanosomal DNA repair factors.</td>
</tr>
<tr>
<td>148</td>
</tr>
<tr>
<td>8.2 Summary of TbSNM1 mutational analysis.</td>
</tr>
<tr>
<td>159</td>
</tr>
</tbody>
</table>
Abbreviations

AQP | Aquaglyceroporin 2
BER | Base excision repair
BLA | Blasticidin–S-deaminase gene
BSF | Bloodstream form
BSF | Bloodstream form
CATT | Card agglutination test for trypanosomiasis
cDNA | Complementary DNA
CSF | Cerebral spinal fluid
DAPI | 4’,6-diamidino-2-phenylindole
DFMO | Difluoromethylornithine
DSB | Double strand break (dsDNA break)
dsDNA | Double stranded DNA
dsRNA | Double stranded RNA
EAT | East African trypanosomiasis
EC50 | Concentration of drug needed to kill 50% of cells
EP | Epimastigote form
FA | Fanconi anaemia
G418 | Neomycin phosphotransferase gene
gDNA | Genomic DNA
GFP | Green fluorescent protein
GG-NER | Global genome-nucleotide excision repair
GOI | Gene of interest
HAT | Human African trypanosomiasis
HN2 | Mechlorethamine
HR | Homologous recombination
HU | Hydroxyurea
HYG | Hygromycin phosphotransferase gene
ICL | Interstand Crosslink
K (DNA) | Kinetoplastid (DNA)
LAPT | Low affinity pentamidine transporter
LS-BSF | Long slender bloodstream form trypomastigotes
MGMT | Methylguanine methyltransferase
MMEJ | Microhomology-mediated end joining
MMR | Mismatch repair
MMS | Methyl methanesulfonate
MRN | MRE11-RAD50-NSB1
MSF | Médecins Sans Frontières
MT | Metacyclic trypomastigote
N (DNA) | Nuclear (DNA)
NECT | Nifurtimox eflornithine combinational therapy
NER | Nucleotide excision repair
NGOs | Non-governmental organisations
NHEJ | Non-homologous end joining
NLS | Nuclear localisation signal
NTD | Neglected tropical disease
NTR | Type I nitroreductase
P2 | Purine transporter
PAC | Puromycin-N-acetyltransferase gene
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCF</td>
<td>Procyclic form trypomastigotes</td>
</tr>
<tr>
<td>PCNA</td>
<td>Proliferating cell nuclear antigen</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PHLEO</td>
<td>Bleomycin resistance protein gene</td>
</tr>
<tr>
<td>PSO2</td>
<td>Psoralen sensitive 2</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>Reverse transcription-quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>SNM1</td>
<td>Sensitive to nitrogen mustard 1</td>
</tr>
<tr>
<td>SSB</td>
<td>Single strand break</td>
</tr>
<tr>
<td>SS-BSF</td>
<td>Short stumpy bloodstream form trypomastigotes</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single stranded DNA</td>
</tr>
<tr>
<td>TC-NER</td>
<td>Transcription coupled-nucleotide excision repair</td>
</tr>
<tr>
<td>TLS</td>
<td>Translesion synthesis</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet radiation</td>
</tr>
<tr>
<td>VSG</td>
<td>Variant surface glycoprotein</td>
</tr>
<tr>
<td>WAT</td>
<td>West African trypanosomiasis</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
<tr>
<td>WT</td>
<td>Wild type</td>
</tr>
</tbody>
</table>
1. Introduction

1.1 Trypanosomatids

Trypanosomatida represents an order of protozoan parasites belonging to the class Kinetoplastida. Most are monoxenous, primarily affecting insects, although several have evolved strategies to infect secondary hosts that as a grouping can effect a range of organisms including plants, invertebrates or vertebrates (Stevens et al., 2001). Characteristically, trypanosomatid cells in at least one of their life cycle stages have a long slender morphology and possess a single flagellum. Together, these combine to give these parasites a distinctive corkscrew-like motility resulting in the order name: “trypó” and “soma” are Greek for “drill” and “body”, respectively. Being eukaryotes, Trypanosomatida contain a number of ‘classical’ membrane-bound organelles such as the nucleus, mitochondria, endoplasmic reticulum but interestingly also harbour several ‘unusual’ structures including glycosomes, which represent modified peroxisomes that play a key role in glycolysis, a kinetoplast that contains the mitochondrion genome, and a small invagination of the plasma membrane called the flagellar pocket where the flagellum exits the cytoplasm (Figure 1.1) (Stevens et al., 2001; Barrett et al., 2003; Lopes et al., 2010).

![Figure 1.1](image.png)

Figure 1.1: Typical structure of trypanosomatid cell.
The image shown is of the trypomastigote form of *Trypanosoma cruzi* and illustrates the organelles found in this parasites life cycle stage (Teixeira et al., 2012).

Trypanosomatids are responsible for a number of insect transmitted diseases that are of medical and veterinary importance. The pathogens *Trypanosoma cruzi, Trypanosoma brucei*
and *Leishmania* are of particular note as they cause Chagas disease (also known as American trypanosomiasis), human African trypanosomiasis (HAT; also known as African sleeping sickness) and Leishmaniasis, respectively. Together, over 7 million people are infected by one of these agents resulting in around 55,000 deaths per year (‘WHO | World Health Organization’, 2018). In addition to their direct effect on human health, these infections have had a major impact on the economic development of poor, rural populations in Latin America, sub-Saharan Africa and south-east Asia. With a series of other viral, bacterial and parasitic diseases, trypanosomatid infections are recognised by the World Health Organisation (WHO) as 'neglected tropical diseases' (NTDs) that often trap individuals and their families in a cycle of poverty and disease (*CDC - Global Health - Neglected Tropical Diseases*, 2018). Breaking this sequence is a major aim of many non-government organizations (NGOs) such as the Drugs of Neglected Diseases initiative, Médecins Sans Frontières and Pan American Health Organization with such agencies playing key roles in the development and implementation of new preventative strategies, novel diagnostic procedures and treatment schedules to combat such deliberating conditions. In relation to human infectious trypanosomatids, these approaches have seen a dramatic fall in HAT and Chagas disease (Simarro *et al.*, 2011). Buoyed by the success of such strategies, Dr Jean Jannin, Coordinator of the Innovative and Intensified Disease Management unit of the WHO Department for Control of Neglected Tropical Diseases, commented that the target is to eliminate HAT as a public-health problem by 2020 by bringing down the incidence of the disease to less than 1 case per 10,000 of the population in at least 90% of the areas where cases exist (‘WHO | Trypanosomiasis, human African (sleeping sickness)’, 2017).

### 1.2 Human African trypanosomiasis

The first accounts of HAT are closely linked with the 18th Century slave trade with ship doctors and naval surgeons noting several symptoms associated with the captured people. The first accurate report of the neurological conditions of the disease were identified in 1734 by John Aktins with other characteristic signs such as localised lymphadenopathies subsequently described by the English physician Thomas Winterbottom (Steverding, 2008). It is now recognised that HAT is a problem across the whole of sub-Saharan Africa where two discernible forms of the disease are prevalent with each disorder caused by a different *T. brucei* sub-species. In modern times there have been three major HAT epidemics with the first of these lasting approximately one decade (1896-1906), primarily affecting Uganda, Kenya and the Congo Basin resulting in the death of an estimated 800,000 people (Steverding, 2008). This
resulted in a drive to find a cure for this devastating disease and lead to the development of the organo-arsenical compound tryparsamide and Bayer 205, now known as suramin. During the 1920’s through to the late 1940’s the second HAT outbreak occurred at a number of sites across sub-Saharan Africa with the above treatments being used to treat infected individuals. This promoted the then colonial powers to implement active measures based on mobile disease surveillance teams to identify and treat infected people, the application of vector/host reservoir control strategies aimed at targeting disease spread and continue the development of new therapies with the latter resulting in discovery of pentamidine and melarsoprol. The most recent epidemic occurred during the last three decades of the 20th Century (1970’s to the late 1990’s) with highest prevalence in Angola, the Democratic Republic of the Congo, South Sudan and Uganda. The reasons for this outbreak are numerous but are primarily associated with political instability. As such, NGO-mediated intervention and treatment strategies could not be implemented with the estimated number of new cases and deaths hitting a peak in the mid-1990’s with an estimated 450,000 new HAT cases and predicted 66,000 deaths in one year alone (Figure 1.2) (Ruiz-Postigo et al., 2012). As political tensions eased, the situation regarding HAT has drastically improved with only 2,804 cases reported in 2015 (Figure 1.2), a figure which grossly underestimates the true extent of the problem, estimated at 20,000 cases (‘WHO | Trypanosomiasis, human African (sleeping sickness)’, 2017). One major reason for this success has been the employment of the nifurtimox eflornithine combination therapy (NECT), the first co-administered treatment to target this infection. With sustained preventative and therapeutic strategies, the goal for the eradication of HAT as a medical problem from all of its endemic sites by 2030 has been envisaged by the WHO Department of Control of Neglected Tropical Diseases in Geneva (Qian and Zhou, 2016).

1.2.1 Distribution of disease

The distribution of HAT in sub-Saharan Africa is solely determined by the geographical range of its insect vector, the tsetse fly. Approximately 30 different Glossina species and subspecies found in riverine, savannah or forest habitats have been recorded to transmit the causative agent through the arthropods haematophagic behaviour. In contrast to some other insect transmitted parasitic diseases (e.g. malaria and leishmaniasis), both male and female tsetse flies are able to spread HAT via a salivarian route. One area of the insect’s life cycle that is amenable to vector control concerns their reproductive strategy. Unlike many other insects, female tsetse flies are viviparous, investing their energy into giving birth to one fully formed larvae at a time. This aspect of the arthropods life cycle allows vector control teams a window
of opportunity to target the fly, thus preventing disease transmission. Other routes for spread of HAT, such as sexual and congenital transmission, modern medical practices (e.g. blood transfusions) and lab-based injuries, have been reported for this condition but these are uncommon (Franco, 2014).

![Figure 1.2: Estimated number of human African trypanosomiasis cases.](image)


1.3 Causative agent

In relation to location, transmission, epidemiology and disease staging, HAT can be split into two distinct variants, both caused by different sub-species of *T. brucei*. In West and Central Africa, *T. brucei gambiense* is responsible for a chronic infection termed West African trypanosomiasis (WAT) which accounts for 98% of all recorded HAT cases. The parasite is predominantly transmitted anthropomorphically by riverine and forest dwelling tsetse flies with this disease form being commonly associated with rain forest, river or lake habitats (Table 1.1). Within a human host, *T. b. gambiense* parasitaemia is generally low with many individuals being asymptomatic and therefore acting as a reservoir for the disease. In eastern and southern regions of Africa the so-called East African trypanosomiasis (EAT) prevails. Caused by *T. brucei rhodesiense*, the acute form of the disease is associated with dry bush and woodland, and transmitted by savannah dwelling tsetse flies (Table 1.1). Commonly found in animal reservoirs, particularly ungulates, EAT is a zoonosis with high parasitaemia commonly noted in an infected individual. As a result, this form of HAT can kill a patient within a month of being infected (Brun *et al.*, 2010).
Table 1.1: Comparisons between different forms of human African trypanosomiasis

<table>
<thead>
<tr>
<th></th>
<th>East African trypanosomiasis</th>
<th>West African trypanosomiasis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Causative agent</strong></td>
<td><em>T. b rhodesiense</em></td>
<td><em>T. b gambiense</em></td>
</tr>
<tr>
<td><strong>Tsetse fly location</strong></td>
<td>Savannah</td>
<td>Riverine</td>
</tr>
<tr>
<td><strong>Ecology</strong></td>
<td>Dry bush and woodland</td>
<td>Rain forest, rivers, or lakes</td>
</tr>
<tr>
<td><strong>Geographical location</strong></td>
<td>East/Southern Africa</td>
<td>West/Central Africa</td>
</tr>
<tr>
<td><strong>Transmission</strong></td>
<td>Zoonotic</td>
<td>Anthropomorphic</td>
</tr>
<tr>
<td><strong>Non-human hosts</strong></td>
<td>Wild &amp; domestic animals</td>
<td>Rare</td>
</tr>
<tr>
<td><strong>Epidemiology</strong></td>
<td>Sporadic</td>
<td>Endemic with cases of epidemics</td>
</tr>
<tr>
<td><strong>Prognosis</strong></td>
<td>Rapid progression to death (weeks to months)</td>
<td>Slow progression (~1 year) leading to death</td>
</tr>
<tr>
<td><strong>Parasitaemia</strong></td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td><strong>Asymptomatic carriers</strong></td>
<td>Rare</td>
<td>Common</td>
</tr>
</tbody>
</table>

Table 1.1: Comparisons between different forms of human African trypanosomiasis

Table adapted from (Stevens et al., 2001).

1.3.1 Parasite life cycle

The two sub-species of *T. brucei* show similar life cycle strategies in their insect and human hosts (Figure 1.3). Transfer of the parasite-infected insect saliva occurs when the tsetse fly takes a blood meal from a mammalian host. Trypanosomes in the metacyclic trypomastigote (MT) form are injected into the victim’s blood and/or lymphatic systems where they transform, possibly in response to a temperature shift, into long slender bloodstream form trypomastigotes (LS-BSF). Here, the parasite multiplies by binary fission to colonise the above circulatory systems with fluctuating parasitaemia, and if left untreated, can eventually cross the blood brain barrier to infect the cerebrospinal fluid. The oscillation of *T. brucei* numbers in the bloodstream and lymphatic system occurs in response to antibody-mediated immune insults with this pathogen able to alter its surface coat and hence antigenic profile presented to the host. This process, known as antigenic variation, involves the periodic changing of their variant surface glycoprotein (VSG), which acts as an antigen, with the vast majority of cells in a given wave of parasitaemia expressing the same form of the surface protein. The chronic nature of a HAT infection can be maintained as the parasite is able to call upon more than 1000 encoding genes and pseudogenes (Horn and McCulloch, 2010). At the peak of parasitaemia in a given wave, LS-BSF can express a stumpy induction factor which triggers differentiation of some of these trypanosomes into short stumpy bloodstream form trypomastigotes (SS-BSF) (Seed and Wenck, 2003). This new form of the parasite is pre-adapted for transmission into the insect vector but if not transferred, can be readily cleared from the mammalian host.
When a tsetse fly takes a bloodmeal, both LS-BSF and SS-BSF are taken up into the insect’s midgut where the former trypanosome form dies. In response to a drop-in temperature and metabolites (e.g. citrate/cis-aconitate) present in the flies’ digestive tract, the latter trypomastigote then undergoes another round of differentiation leading to procyclic form trypomastigotes (PCF) (Czichos, Nonnengaesser and Overath, 1986). This parasite stage divides by binary fission and colonises this region of the arthropod. The parasite makes its way to the tsetse fly salivary gland, possibly via a mechanism involving conversion of the PCF to an epimastigote form (EP). Here, the EP attaches to epithelial cell lining of this insect organ via its flagellum, divides by binary fission and potentially undergoes genetic exchange (the latter step is not an obligatory part of the T. brucei life cycle). When the EP density increases to high levels, the cells transform into MT parasites, detaching from the epithelial lining of the salivary gland and entering its lumen ready for transmission into the next mammalian host (Stevens et al., 2001; Matthews, Ellis and Paterou, 2004).

**Figure 1.3: T. brucei life cycle.**

1.3.2 Staging and pathologies
The interaction of bloodstream form trypomastigotes with the human host causes defined pathological outcomes that can be divided into two distinct stages (Figure 1.4). The early or
haemolymphatic stage of HAT occurs while the parasite is found systemically within the bloodstream and lymphatic system. Towards the peak of parasitaemia, trypanosomes are recognised by host immune mechanisms which stimulates the production of pyrogenic cytokines such as tumour necrosis factor, interleukin-1, interleukin IL-6 and interferons. This results in a characteristic fever which declines following parasite clearance. The process is then repeated with each subsequent wave of parasites such that the patient presents with a so-called relapsing fever with a periodicity of 5 to 7 days. Other signs of HAT can be manifested at this point such as localised lymphadenopathies with 50% of patients infected with *T. b. rhodesiense* also displaying a chancre at the site of the tsetse fly bite. If left untreated, the parasite can penetrate the blood brain barrier resulting in a series of neurological conditions that represent the late or meningoencephalitic stage of HAT. Here, the patient can exhibit anorexia, apathy and lassitude, have severe sleep disturbance patterns (hence the alternative name for HAT, African sleeping sickness) and undergo personality changes. Towards the end of this phase of the disease, patients often start convulsing before becoming comatose and dying (Kennedy, 2013).

![Figure 1.4: Progression of HAT in a human host.](image)

*T. brucei rhodesiense* and *T. brucei gambiense* progression in a human host. Image adapted from Wiser (2011) ©2010 from Protozoa and Human Disease by Wiser.

### 1.3.3 Diagnosis

Currently, diagnosis of HAT is based on a combination of direct microscopic screening techniques for the presence of the parasite in blood, lymph and cerebral spinal fluids and/or
indirect serology (Kennedy, 2013). WHO guidelines are only available for WAT which can be applied to patients displaying the aforementioned disease signs and/or populations living in an endemic site. In both cases, undiluted blood samples are collected and analysed using the serology-based Card Agglutination Test for Trypanosomiasis (CATT). If positive, the above analysis is then repeated on a diluted blood sample. If the CATT is negative for either of the above, no further action is required. However, if the individual is positive after screening the diluted sample, the blood is microscopically analysed for the presence of the parasite: For *T.b. gambiense* a parasite enrichment process is often required due to the low parasitaemia associated with such infections. If this proves negative, an additional sample is taken a few days later to account for possible fluctuations in the individual. If either sample is parasite positive, the cerebral spinal fluid of that individual is then checked to determine disease staging. Based on these outcomes, appropriate chemotherapies can then be administered (‘WHO | Trypanosomiasis, human African (sleeping sickness)’, 2017).

As CATT is specific for *T.b. gambiense*, this element of the diagnostic procedure cannot be applied to EAT. Given the low incidence rate of EAT, it is not financially favourable to develop a diagnostic test for this form of the condition. Here, direct observation of the parasite in bodily fluids is assessed without the need for any form of enrichment step.

1.4 *Anti-trypanosomal therapy*

Treatment of HAT is comprised of three monotherapies, pentamidine, suramin and melarsoprol, and the combinational therapy NECT. These drugs are used differently according to 1. the disease stage an individual presents with, and 2. the parasite sub-species an individual is infected with.

Pentamidine, an aromatic diamidine (Figure 1.5), administered as daily intramuscular injections over the course of 7-10 days, is the preferred treatment for early stage infections of *T. brucei gambiense*. This drug has restricted usage as its positive charge at physiological pH means it is unable to cross the blood brain barrier to target late stages of HAT. A series of trypanosomal transporters (P2 - purine transporter; AQP - aquaglyceroporin 2; LAPT - low affinity pentamidine transporter) help concentrate the drug in the trypanosomal cell to millimolar concentrations which leads to its cytotoxic effects in the parasite (Bernhard *et al.*, 2007; ORTIZ *et al.*, 2009; Baker *et al.*, 2012; Graf *et al.*, 2013). It is postulated that this drug undergoes a multifactorial attack on the trypanosomes mitochondria by preventing
mitochondrial DNA synthesis and promoting the collapse of the mitochondria membrane potential (Wilkinson and Kelly, 2009; Alsford \textit{et al.}, 2012).

![Figure 1.5: Structures of different clinically relevant trypanocidal agents.](Image obtained from (D’Silva, 2007)).

Early stage EAT can be treated with the polysulphonated naphthylamine (Figure 1.5) drug suramin that is administered by weekly intravenous injections over the course of 5-6 weeks. Again, due to charge, this drug does not readily cross the blood brain barrier thus restricting its use. Uptake of suramin occurs through receptor-mediated endocytosis within the flagellar pocket. This process is triggered by binding of the drug to parasite surface proteins such as ISG75 and VSG isoforms with ligand/receptor complexes then transferred through the endocytic pathway to the lysosome (Leung \textit{et al.}, 2011). Through the action of peptidases and proteases, suramin is released and transferred from the vesicular lumen by a major facilitator superfamily transporter. The precise mechanism of suramin is unclear but by virtue of its negative charge, \textit{in vitro} biochemical studies have shown it could block endocytic and glycolytic activities with cell-based phenotypic assays, indicating it may interfere with polyamine synthesis and N-acetylglucosamine synthesis (Wilkinson and Kelly, 2009; Alsford \textit{et al.}, 2012).

The only drug able to target the late stage of both HAT forms is melarsoprol. This compound is a trivalent, melaminophenyl arsenical prodrug (Figure 1.5) that is administered as a very
painful intravenous injection: the adjuvant used to suspend melarsoprol is polyethylene glycol which causes the collapse of blood vessel epithelial layers that manifests as a burning sensation in the patient. Additionally, this therapy is highly toxic causing death in about 5% of the individuals it is administered too. By uncharacterised mechanisms that may result from host activities, melarsoprol is converted to its druggable form melarsen oxide. This (and melarsoprol) can be internalised by the P2 and AQP transporters where it can readily form complexes with thiol-containing molecules (Bernhard et al., 2007; Baker et al., 2012). The main free thiol in trypanosomes is trypanothione and it has been shown that this can readily conjugate with melarsen oxide, with the resultant complexes inhibiting trypanothione-dependent enzymes (Alsford et al., 2012). Resistance to melarsoprol can be readily achieved in the laboratory through mutation of the P2 transporter and loss or reduction of AQP activity with such lines also displaying cross resistance to pentamidine (Mäser et al., 1999; Graf et al., 2013). Worryingly, these phenotypes have been observed in the field with treatment failures associated with these aforementioned genetic alterations (Bernhard et al., 2007).

Therapeutic switching of both nifurtimox, originally used as an antichagasic treatment, and eflornithine, initially developed as an anticancer agent, has resulted in NECT, the only combinational therapy available against HAT. Previous work revealed that eflornithine was a highly effective anti-HAT agent, earning it the name of the resuscitation or miracle drug. However, it was also extremely expensive, thereby limiting its usage. The use of NECT has drastically reduced the cost of treatment such that it now represents WHO’s recommended therapy against late stage WAT. Following intravenous infusion, eflornithine, a variant of ornithine, is taken up by parasite via a specific amino acid transporter, AAT6, and then goes on to act as an irreversible inhibitor of ornithine decarboxylase, a key enzyme in polyamine and trypanothione synthesis (Baker, Alsford and Horn, 2011; Schumann Burkard, Jutzi and Roditi, 2011). The orally administered and hydrophobic 5-nitrofuran nifurtimox can readily pass through cell membranes. In the trypanosome, this prodrug enters the parasites single mitochondrion where it undergoes type I nitroreductase (NTR) mediated activation, an enzyme absent from human cells. This reaction goes on to generate a series of nitroso, hydroxylamine and open chain nitrile metabolites with the latter displaying high cytotoxicity. To date, there have been no NECT treatment failures although laboratory-based studies have shown that reduction of AAT6 or NTR activities can result in eflornithine or nifurtimox resistance, respectively (Wilkinson and Kelly, 2009; Alsford et al., 2012).
1.5 DNA damage and repair

DNA is a biological structure that is highly prone to different types of damage. If left unchecked, these can cause alterations to the encoding nucleotide sequence and/or promote chromosome instability. These insults arise from various endogenous sources such as errors in DNA processing and reaction of DNA with metabolic products, or as result of exogenous influences including exposure to chemicals or radiation. As a consequence, cells have evolved a series of distinct and complementary DNA repair pathways that function to fix the variety of lesions that may arise (Sancar et al., 2004; Boiteux and Jinks-Robertson, 2013; Kreuzer, 2013). A healthy cell requires a balance between DNA damage and repair mechanisms with disease or cell death occurring when appropriate DNA repair does not occur (Friedberg, 1995).

1.5.1 General overview

DNA damage can be broadly separated into 1. single base changes, 2. large structural distortions and 3. breakage events, with each type of damage requiring a series of specific repair mechanisms (Figure 1.6).

Single base changes can include nucleotide alkylation, hydrolysis, oxidation, depurination, depyrimidination and deamination events as well as tautomeric shifts which have mutagenic effects on the cellular genome (Waters et al., 2009). These toxic lesions can be neutralized by the direct reversal repair (DRR), base excision repair (BER) or nucleotide excision repair (NER) pathways with the invoked repair mechanism determined by the exact alteration event that has taken place. DNA alkylation events can be readily reversed by the DRR pathway with $O^6$-alkylguanine DNA alkyltransferase (MGMT) playing a critical role in the transfer of the mutagenic alkyl-group onto its own protein backbone. For other single base changes, the NER and BER operate with the latter pathway playing the major role. BER can be divided into short-patch BER that functions to repair only the affected nucleotide or long-patch BER where a wider region surrounding the damage is fixed. These pathways rely upon the unaltered complementary DNA strand to aid efficient repair (Hakem, 2008).
Structural DNA distortions, in particular adduct formation, photodimerism and DNA intercalation events, require more complex repair mechanisms due to the torsional stress created on the nucleic acid backbone (Waters et al., 2009). Photodimerism, most commonly cyclobutane pyrimidine dimers (CPD), is the UV-mediated bonding between up to 4 parallel or neighbouring nucleotides resulting in major helix distortion. DNA intercalation events usually involve the insertion of chemical moieties (e.g. ethidium bromide) between nucleotides causing a stretching of the DNA helix. These events are predominantly detected and repaired via the NER pathway with additional assistance from the mismatch repair (MMR) pathway. The NER pathway can be divided into transcription-coupled NER (TC-NER) and global-genome NER (GG-NER), with different pathways exploited at different points in the cell cycle.

The most lethal type of DNA damage are breakage events that result in the impediment of both transcription and replication. Exposure to ionizing radiation or cytotoxic agents can result in double strand breaks (DSBs) with potential loss of whole sections of genomic material dependent on 1. the extent of material lost during the damage, 2. the length of the overhanging strand after the breakage event and 3. introduction of errors in the overhanging strands (Jackson, 2002). Due to the lethality of this type of damage, three pathways play a role in repairing such lesions. Substantial damage to both strands of DNA can be repaired by the non-homologous end joining (NHEJ) or homologous recombination (HR) pathways. NHEJ

**Figure 1.6: DNA damage and repair pathways.**
Different types of damage to DNA (both endogenous and exogenous) and the respective cellular repair mechanisms. Image taken from (Genois et al., 2014).
involves the direct assembly of the 5’ and 3’ termini of the DNA double helix with the possible loss of 5’ and 3’ sequences and introduction of base errors. The HR pathway is employed for error-free DSB repair with sequences upstream from the damage used as templates for sister chromatid searches. The translesion synthesis (TLS) pathway can be employed to bypass lesions when repair is unfavourable e.g. during transcription/replication.

Single strand breaks are repaired by the single strand annealing (SSA) pathway which involves a simple ligation of the overhanging strand back to the DNA backbone.

In certain situations, DNA crosslinks can form involving reaction of the nucleic acid structure with another, usually nucleophilic chemical entity that contains at least two reactive centres. This can result in 1. intrastrand crosslinks, where the mutagen covalently bonds to bases within the same DNA strand, 2. interstrand crosslinks, where the mutagen covalently bonds to bases within complementary DNA strands or 3. formation of DNA-protein complexes. Any combination of the above ‘classical’ DNA repair pathways can be invoked as crosslinking can cause single base changes, large structural distortions and/or breakage events (Peter J McHugh, Spanswick and Hartley, 2001).

A comprehensive list of the components of each DNA repair pathway can be found at the Trypanosomatid DNA repair resource: https://sites.google.com/site/trypdnarepair/home.

1.5.2 Trypanosomal DNA repair pathways

Comparative studies relating to how T. brucei repairs DNA damage have been performed with many of components of the HR, MMR, BER, TLS, NER and MMEJ pathways identified. A comprehensive list of the components of each DNA repair pathway found at the Trypanosomatid DNA repair resource (reviewed in (Passos-Silva et al., 2010; Genois et al., 2014)). A key finding from such metadata analyses corroborated by experimental observations is that trypanosomes lack a functional NHEJ pathway despite expressing several components of this system (e.g. Ku70, Ku80 and ARTEMIS) (Glover, McCulloch and Horn, 2008; Genois et al., 2014).

Functional analysis of the trypanosomal NER pathway indicates that the transcription coupled pathway preferentially operates relative to the global-genome mechanism (Machado et al., 2014). This may reflect that trypanosomatid gene expression is polycistronic with
transcription of the vast majority of genes effectively being ‘on’ all the while with regulation of each cistron occurring post-transcriptionally. As such there is an absolute requirement in these parasites to have their TC-NER mechanism constantly functioning as a sentinel to identify and repair lesions encountered during transcription. In other organisms, stalled RNA polymerase acts to target CSA and CSB to the site of DNA damage. These enzymes displace the polymerase from the DNA and recruit TFIIH NER factors to promote DNA helix unwinding (e.g. XPB, XPD, etc) to create an open bubble structure that is maintained in this state by a preincision complex (e.g. RPA, XPA) prior to a dual incision event mediated by the ERCC1-XPF complex and XPG (Lagerwerf et al., 2011; Genois et al., 2014; Iyama et al., 2015). This results in single stranded DNA that is filled in by DNA polymerases using the undamaged, complementary strand as template followed by ligation of the sugar-phosphate backbone. The GG-NER system differs slightly from the TC-NER counterpart primarily in how the damage is detected which involves the damage recognition activities of DDB1/DDB2 or XPC/Rad23b complexes. These then recruit TFIIH NER factors and promote repair as described for TC-NER.

Several components of TC-NER (e.g. TbCSB, TbXPB, TbXPBz and TbXPG) and GG-NER (TbXPC) have been identified and characterised in T. brucei with their role in the corresponding pathway confirmed. Intriguingly, a number of elements such as CSA, XPA, DDB2 and components of the TFIIH complex are apparently missing (Machado et al., 2014). Based on such differences it is implicit that this parasites TC- and GG-NER pathways are distinct from model eukaryotic organisms, a feature that may be exploitable in terms of chemotherapeutic development.

Cells are able to bypass certain DNA lesions using the TLS pathway with different systems predominating at various points within the cell cycle. During the S-phase, the so-called polymerase switching model operates. Here, stalling of DNA replication complex at a site of DNA damage causes posttranslational modification of proliferating cell nuclear antigen (PCNA) that leads to replacement of the incumbent DNA polymerase within the replication complex with a TLS specific DNA polymerase (Waters et al., 2009). At other times within the cell cycle, the TLS pathway operates through the so-called gap-filling model (Shen et al., 2006; Waters et al., 2009). In this situation, detection of a stalled transcription complex by NER factors can result in recruitment of the TLS specific DNA polymerases (Enoiu, Jiricny and Schärer, 2012). Using ssDNA present in a transcription bubble as template, these protein
complexes guide the addition of nucleotides at the lesion site to generate a dsDNA molecule with DNA ligases then operating to maintain the integrity of the sugar-phosphate backbone. Several TLS specific DNA polymerases can operate in these two models with these differing in their fidelity: TLS DNA polymerases are said to be error-free (high fidelity) or error-prone (low fidelity). In the latter case, the polymerase activity can lead to ‘secondary’ base mismatch mutations that require downstream repair (e.g. by the MMR pathway). To date, little is known about the trypanosomal TLS pathways although these systems are known to operate with at least five TLS polymerases (polymerase nu (PolN), polymerase zeta (PolZ), polymerase theta (PolQ), polymerase eta (PolH) and polymerase kappa (PolK)) implicated in these activities (Machado et al., 2014; Zurita-Leal, 2016). Two trypanosomal TLS polymerases, PPL1 and PPL2, have been implicated with trypanosomal TLS activity (Rudd et al., 2013).

Many trypanosomal components of both HR pathways, single strand annealing and synthesis-dependent strand annealing pathway, have been functionally characterised and linked to one way these parasites undergo antigenic variation (Horn, 2014; Morrison, McCulloch and Hall, 2015). In other organisms, the synthesis-dependent HR pathway relies on detection of a break by the MRE11-RAD50-NBS1 (MRN) complex with DNA resection processing single stranded DNA (ssDNA). This DNA acts as the guide for the strand exchange reaction, carried out by RAD51 and maintained by the activities of BRCA2, PALB2, RAD52 and RAD51 paralogs. This complex carries out the homology search for an intact sister chromatid sequence with RAD51 and RAD54 mediating the cross-over. With strand synthesis complete, the invading strand dissociates from the complex and anneals to the complementary sequence with gaps filled and nicks ligated to complete repair. The single strand annealing pathway operates in a RAD51-independent complex with substantial cleavage of 5’ and 3’ termini at a DSB followed by single strand annealing and ligation. Several components of the single strand annealing and synthesis-dependent strand annealing pathway have been identified and characterised in T. brucei including the initial detection and resection activities of TbMRE11 and TbRAD50, recombinase activity of TbRAD51 and ancillary activities of TbBRCA2 and TbRAD51 paralogues (TbRAD51-1 to TbRAD51-4) (McCulloch and Barry, 1999; Robinson, McCulloch, Conway, Browitt and Barry, 2002; Hartley and McCulloch, 2008; Dobson et al., 2011). This work has shown that not all components of the HR pathway present in this parasite with absence of RAD52 and PALB2 noted in various in silico analyses of the T. brucei genome (Genois et al., 2014).
Mismatch repair, often referred to as the postreplicative repair pathway, is employed at mismatched base errors with damage recognition carried out by the MutS (MSH2, MSH6 and MSH3) heterodimer and lesions cleaved by the endonuclease activity of MutL (MLH1, MLH3 and PMS1) homodimers. This complex recruits EXO1, a 5’-3’ exonuclease, which carries out exonucleolytic degradation of the nucleotide. DNA synthesis is carried out by a DNA polymerase with nicks in both strands sealed by DNA ligases. Almost all components of the MMR machinery are conserved in trypanosomatids: MutS-like genes in the parasite (e.g. TbMSH2, TbMSH8 and TbMSH3) have been identified and characterised to play a role in both nuclear and mitochondrial DNA repair (Barnes and McCulloch, 2007), MutL-like genes, primarily PMS1 and MLH1, have also been functionally characterised however no characterisation of an EXO1 homologue has been studied (Szankasi and Smith, 1992; Bell et al., 2004).

1.6 Interstrand crosslinks and their repair

The covalent bonding of two complementary strands of DNA leads to formation of an interstrand crosslink (ICL). This form of DNA damage is highly detrimental to a cell as it can block any process that requires DNA strand separation and can result in stalling of the transcription and/or replication machinery, mismatch errors and loss of genomic material. As few as one ICL in a bacterial cell or 20 in a mammalian cell can result in the death of that cell (Dronkert and Kanaar, 2001). To survive such insults, a series of activities have been co-opted to generate multiple, complementary and overlapping repair systems. These function at different points in the cell cycle with the nature of the mechanisms varying between groups of organisms (P J McHugh, Spanswick and Hartley, 2001; Clauson, Schärer and Niedernhofer, 2013). As such, comparative studies involving humans don’t necessarily inform what the situation is in yeast, which in turn do not reflect mechanisms employed by trypanosomes. Therefore, each organism’s ICL REPAIRtore is distinct.

1.6.1 ICLs: definition, formation and significance

ICLs are predominantly but not exclusively formed by exogenous chemicals with many anticancer agents exploiting their detrimental effects. In rare cases, ICLs can be formed endogenously by ionizing radiation, oxidative stress and simple unsaturated aldehydes that are generated during lipid peroxidation and prostaglandin synthesis promoting such lesions (Noll, Mason and Miller, 2006; Clauson, Schärer and Niedernhofer, 2013). These ICL inducing agents have two reactive centres with one of these commonly becoming activated and reacting
with a guanine residue on one DNA strand (Noll, Mason and Miller, 2006). This interaction then promotes activation of the second reactive centre, facilitating binding to a nearby guanine. Many ICL forming chemicals favour interacting with this particular base due to the large electrostatic potential that occurs between two guanine residues resulting in preference of alkylation events at guanine-rich regions of DNA (Noll, Mason and Miller, 2006). The sequences targeted by different ICL promoting compounds vary e.g. nitrogen mustards preferentially reacting at 5′-GNC-3′ sites, mitomycin C favouring 5′-CG-3′ sequences, and platinum-based compounds promoting the binding at 5′-GC-3′. An exception to this pattern are the psoralen-based compounds as they show a marginal preference for 5′-TA-3’ and 5′-AT-3’ sequences (Figure 1.7).

Figure 1.7: Structures of different types of ICL inducing agents. Structures redrawn on ChemDoodle.

Nitrogen mustards are chemical structures that contain a chloroethylamine grouping in their backbone and function by promoting alkylation events primarily affecting nucleic acid targets.
For those that induce ICLs, e.g. mechloretamine and cyclophosphamide (Figure 1.7), at least two such groupings are present on the structures’ backbone, both of which can form highly reactive and nucleophilic centres. In the presence of DNA, particularly where 5′-GNC-3′ sequences are found, the terminal chloride found on one arm of the alkyl reactive centre is displaced to form an aziridinium ion that attacks the N7 position of a guanine resulting in the alkylation of that base. In 1 to 2% of cases the terminal chloride found on the second reactive centre undergoes a similar event leading to ICL formation (Dronkert and Kanaar, 2001). This damage results in structural distortion of the double helix (Rajski and Williams, 1998).

A number of ICL forming compounds contain at least one aziridinyl grouping in their structure with these being embedded within (‘trapped’) or as a side chain (‘free’) on the chemicals backbone. In all situations, an activation event is required to promote formation of a reactive centre. For compounds such as CB1954 and RH1 (Figure 1.7) that contain two or more ‘free’ aziridinyl moieties, the above leads to formation of an aziridinium ion that then promotes ICL formation analogous to that noted for the nitrogen mustards. In chemicals such as the mitomycins, where the azirinidyl is ‘trapped’, activation causes a redistribution of electrons through the indolequinone ring structure that promotes DNA binding to sites neighbouring the aziridinyl. Initially, a methoxy grouping adjacent to the aziridinyl is eliminated from the chemical backbone leading to formation of a quinone methide. In the presence of DNA, particular where 5′-CG-3′ sequences are found, the methide then functions as a nucleophile to promote formation of firstly a monoadduct at the N2 position of a guanine base on strand of DNA then a bisadduct by reaction with the N2 position of a guanine base on the complementary strand (Meredith et al., 2017).

Other ICL inducing agents of note include the platinum-based and furocoumarin-containing compounds (Figure 1.7). In the case of the former, as represented by cisplatin and carboplatin, moieties in the structure (e.g. the chloride ions in cisplatin) are slowly displaced by water to create aquo complex intermediates with the water subsequently supplanted by a purine nucleotide, particularly guanine, at the N7 position. Most damage via platinum-based compounds occurs within the same DNA strand (intrastrand crosslink) with a minor component (~10%) being interstrand in nature (Peter J McHugh, Spanswick and Hartley, 2001). In the latter case, 5′-CG-3′ sequences being particularly prone to covalent binding.
Furocoumarin-containing compounds such as psoralen and angelicin promote ICLs by a mechanism distinct from that described above. Analogous to UV-induced pyrimidine dimer formation, these are photoactivated such that their furan and pyrone rings form cyclobutane structures with thymidine residues found on opposite DNA strands leading to the crosslink formation and intercalation at $5'\text{-TA-3'}$ and $5'\text{-AT-3'}$ sequences (Friedberg, 1995). As with the platinum-based agents, ICL formation constitutes a minor component of any resultant DNA lesion.

To conserve genome integrity, all cells exploit various combinations of enzymes from the “classical” DNA repair pathways to resolve ICL damage. The precise makeup of these systems is unclear as different mechanisms predominate at different stages in the cell cycle while evolutionary diverse organisms employ distinct mechanisms to accomplish this task (Lehoczký, McHugh and Chovanec, 2007; McVey, 2010; Hashimoto, Anai and Hanada, 2016). Broadly, in a mechanism initiated by stalling of a DNA replication complex, the damaged DNA is processed through “unhooking” of the ICL from the nucleic acid backbone with the concomitant formation of a DSB (Hashimoto, Anai and Hanada, 2016). Using complementary DNA sequences (e.g. from a sister chromatid), enzymes from the HR pathway function to fix this break and, in conjunction with components from the TLS pathway, completely remove the “unhooked” ICL from the DNA backbone. In a DNA replication independent mechanism, components of the NER pathway recognise, cleave and “unhook” the ICL from one strand of the DNA backbone (Hashimoto, Anai and Hanada, 2016). This generates a single stranded gap within the nucleic acid molecule with the dsDNA structure restored through TLS activity. Another round of NER and TLS then completely removes the “unhooked” ICL from the DNA. As the above mechanisms involve TLS, which can be error prone, the MMR pathway operates as a sentinel to fix any damage that may arise. Due to the complexity of ICL repair, the focus of this literature review will exclusively centre on the ICL repair systems found in *Escherichia coli*, *Saccharomyces cerevisiae* and mammalian cells.

### 1.6.2 ICL repair in *E. coli*

Over the last 40 years, a picture of the *E. coli* ICL repair system, often referred to as the Cole Model of ICL repair (Cole, 1971) (Figure 1.8), has been elucidated and shown to involve the coordinated activity of components of the NER and HR pathways. Here, the UvrABC constituents of the bacterial NER pathway function in the 1. DNA damage recognition (UvrA and UvrB), 2. unwinding the DNA strand surrounding the lesion (UvrB), and 3. cleaving the
DNA backbone either side of the ICL (UvrC) (Houten et al., 1986). In certain cases where the torsion strain caused by the crosslinking event creates structural distortions to the DNA double helix, additional helicase activities (e.g. UvrD) can be employed to aid the unwinding process (Noll, Mason and Miller, 2006). The above generates an ‘unhooked crosslink’ with the ICL remaining covalently attached to one strand of the DNA backbone. Sequences immediately upstream of the initial incision are processed by the exonuclease activity of DNA Polymerase I (Pol I) which in turn recruits RecA from the HR pathway. By virtue of its recombinase activity, RecA then promotes strand invasion from a related DNA sequence, a process aided by the activities of RecB, RecC, and RecD (these act in strand recognition, binding of terminal ends and processing of breaks). The above results in a Holliday junction that can be resolved using the RuvABC complex (Houten et al., 1986). Following recombination, a second round of UvrABC-mediated incision is performed on the other strand, releasing the ‘unhooked crosslink’ from the DNA backbone to create a gap that is filled by Pol I and DNA ligase activities.

![Diagram](Figure 1.8: ICL repair systems in E. coli.)

E. coli expresses two distinct ICL repair systems one being HR-dependent and the other being a TLS-dependent mechanism. Crosslinked DNA (DNA and crosslink shown in black and red, respectively) is unhooked from the double helix by the activity of NER factors (purple arrows)
(part 1). The unhooked ICL remains attached to one of the DNA strands while a small formed in the other. This gap can then filled by a HR-mediated strand invasion process (part 2) or a TLS-dependent mechanism (part 3): The homologous template used in the former is shown in brown while the newly synthesised DNA noted in the latter is in blue. The ICL is then completely removed from the DNA structure by a second round of NER (part 4) followed by DNA polymerase and DNA ligase activities (part 5). Parts 1, 2, 4 and 5 constitute the so-called Cole Model for ICL repair. Imaged adapted from (Huang and Li, 2013).

In addition to the above Cole Model, an alternative recombination-independent ICL repair system has been also proposed. In this case, following the initial UvrABC incision, DNA Pol IV catalyses TLS activity past the still attached ‘unhooked crosslink’ with the integrity of the sugar-phosphate backbone maintained by downstream DNA ligase activity. As above, a second round of UvrABC incision completely removes the crosslink from the genome with Pol I and DNA ligase activities filling in the resultant gaps (Kumari et al., 2008).

1.6.3 S. cerevisiae ICL repair systems

Many of the features displayed in the E. coli ICL repair system are mirrored in simple eukaryote organisms such as budding yeast, S. cerevisiae, although the complexity and overlapping nature of these systems alters throughout the cell cycle (Figure 1.9). Throughout most of this cycle (G0-, G1-, early S- and G2-phase), a NER and TLS pathway similar to the bacterial mechanism predominates, which additionally incorporates novel ICL repair specific activities displayed by the nuclease PSO2 (also known as SNM1) (Part A of Figure 1.9). In this system, stalling of the transcriptional machinery at an ICL promotes ubiquitination of PCNA that then functions to recruit components, including nucleases, of the NER pathway. These proceed to mediate the unhooking of the ICL from one strand of the double helix (Waters et al., 2009). RAD1 (with RAD10) and RAD2 possess endonuclease activities that cleave sequences 5’ and 3’ around the damaged DNA, respectively, with RAD4/RAD14 postulated to aid in recognition of the adduct and the RAD3/RAD25 helicases acting to unwind the DNA helix (Sarkar et al., 2006). The PSO2 nuclease then processes the ’unhooked crosslink’ to remove all the bases up to and across the linkage present within this sequence (Bonatto et al., 2005). The resultant gap is then filled by damage-tolerant TLS polymerases (e.g. REV1 and REV3) with the sugar-phosphate backbone reformed by DNA ligases (Ho et al., 2011). As with the E. coli systems, a second round of NER coupled with DNA polymerase/ligase activities help to completely remove the ICL from the DNA template.
Cells that encounter ICLs while in the late S-phase utilize a series of factors from the NER, MMR and HR pathways that act in concert help resolve the lesion (Part C of Figure 1.9). In this system, the NER pathway is believed to carry out the initial incision events either side of the ICL. If not repaired by the PSO2-dependent pathway, an approaching replication fork can stall in the immediate vicinity of the lesion resulting in recruitment of MPH1, MUTSα and MGM101. MPH1 causes the replication fork to reverse and allows loading of EXO1 onto the ‘unhooked crosslink’ (Ward et al., 2012). This nuclease then removes the bases from this sequence up to but not beyond the ICL, with any resulting gap on the other strand filled by damage-tolerant TLS polymerases and DNA ligases. A second round of NER activity begins processing the DNA strand containing the ‘unhooked crosslink’ with the lesion being completely uncoupled from this backbone with any gap filled via DNA polymerase/ligase activities. Before DNA replication can recommence, the nascent leading or lagging (one not

![Diagram of ICL repair systems in yeast.](image)

**Figure 1.9: ICL repair systems in yeast.**
Yeast express at least three distinct ICL repair mechanisms whose activities appear to be cell cycle dependent. Initial incision of the damaged DNA is common to these systems and employs NER factors. For cells in the G0-, G1-, early S- or G2 phase, a PSO2/TLS-dependent mechanism functions (A) while cells in the S-phase employ HR-based activities (B). An alternative G2-phase system that functions independently of PSO2 and involves EXO1 in conjunction with TLS activities has also been reported (C). The final steps of mechanisms A,
B and C is common with the ICL completely removed from the DNA structure by a second round of NER followed by DNA polymerase and DNA ligase activities. Starting DNA and crosslink are shown in back and red, respectively while the purple arrows correspond to the incisions made by NER factors. Newly synthesised DNA as a result of DNA repair is shown in blue, the homologous DNA template used in the HR mechanism is shown in brown while newly synthesised leading and lagging DNA strands associated with replication forks is in green. Image adapted from (Barber et al., 2005).

both) DNA strand undergoes a DSB that is repaired by components of the single strand annealing HR pathway: RAD52 is recruited to the site of DNA damage through interaction with MGM101 with ancillary factors such as CHL1 and SLX4 aiding this event (Kenneth F. Grossmann et al., 2001; Ward et al., 2012; Munari et al., 2013).

In addition to the PSO2-dependent G2-phase ICL repair system, an alternative mechanism may also operate (Part B of Figure 1.9). Incision and processing events possibly involving NER components leads to unhooking of the ICL with sister chromatids then providing a template for MRE11- and RAD51-dependent HR (Kenneth F. Grossmann et al., 2001). It is hypothesised that a second round of NER coupled with DNA polymerase/DNA ligase activities completely removes the ICL from the genome.

1.6.4 Mammalian ICL repair

Two cell-cycle dependent ICL repair systems have been identified in mammalian cells. A transcription-coupled ICL repair system utilises components of the NER and TLS pathways as well as the ICL-specific activity of SNM1 (homolog of PSO2) for removal of crosslinks (Figure 1.10). The second and dominant ICL repair system termed the Fanconi Anaemia (FA) pathway is a replication-coupled mechanism (Sancar et al., 2004). This system is favoured due to a readily available sister chromatid sequence which enables error-free repair (mammalian cells are diploid organisms) (Hashimoto, Anai and Hanada, 2016) (Figure 1.11). This system was initially discovered through genetic profiling of patients with Fanconi Anaemia, a genetic condition associated with bone marrow failure, of which over 20 linkage groups and hence defective genes (termed FA factors) have been described (Moldovan and D’Andrea, 2009; Krishnan, Tay and Ito, 2015).

1.6.4.1 ICL repair in transcriptionally active cells

In transcriptionally active cells, an ICL can block progression of the transcription machinery leading to stalling of the RNA polymerase complex. This results in recruitment of factors such as CSB and UHRF1 (a RING finger protein) to the lesion site with these interacting directly or
through a partner(s) with other DNA repair enzymes such as SNM1, XPG and XPF-ERCC1 (Iyama et al., 2015; Hashimoto, Anai and Hanada, 2016). The XPF-ERCC1 and XPG endonucleases make the initial 5’ and 3’ incisions around the ICL with nucleolytic degradation of the resultant unhooked sequence occurring through SNM1 (PSO2 in yeast) processing (Al-Minawi et al., 2009; Wood, 2011; Zhang and Walter, 2014; Abdullah et al., 2017). The resultant gap on the strand not containing the unhooked ICL is then filled using damage tolerant TLS polymerases such as Pol ζ, Pol η, Pol ι, Pol Θ or Pol κ in conjunction with DNA ligase (Sarkar et al., 2006; Ho et al., 2011; Sharma and Canman, 2012). Once this strand has been repaired the ICL is completely removed by a second round of NER incision, DNA polymerase and DNA ligase activities.

Figure 1.10: ICL repair system in transcriptionally active mammalian cells. The mechanism outlined here is analogous to that of the Cole Model (see Figure 1.8). Image adapted from (Huang and Li, 2013).
1.6.4.2 ICL repair in dividing cells

ICL-induced collapse of replication forks are recognised and repaired in a mechanism that involves: 1. recognition of the damaged DNA and assembly of eight proteins that form the so-called FA core complex, 2. activation of the FANCD2/FANCI heterodimer that functions as a recruitment complex and 3. association of effector factors that operate to resolve the ICL. The activities of all these proteins constitute the FA pathway (Figure 1.11) with the components involved listed in Table 1.2: Note, the factors listed in this table are those that have been directly associated with FA but it is clear that other components are implicated in this repair system but loss of these activities do not result in the genetic condition (or have not yet been identified).

<table>
<thead>
<tr>
<th>Factor</th>
<th>Function/DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FA core complex</strong></td>
<td></td>
</tr>
<tr>
<td>FANCA</td>
<td>Creation of FA core complex in nucleus</td>
</tr>
<tr>
<td>FANCB</td>
<td>FA core complex structure dimerization.</td>
</tr>
<tr>
<td>FANCC</td>
<td>Adaptor for substrates in FA core complex</td>
</tr>
<tr>
<td>FANCE</td>
<td>Adaptor for substrates in FA core complex</td>
</tr>
<tr>
<td>FANCF</td>
<td>Anchor FANCM to FA core complex</td>
</tr>
<tr>
<td>FANCG (XRCC9)</td>
<td>Creation of FA core complex in nucleus</td>
</tr>
<tr>
<td>FANCL</td>
<td>Part of the FA core complex as a E3 ligase to alter ubiquitinated proteins</td>
</tr>
<tr>
<td>FANCM (MPH1)</td>
<td>Helicase activity to detect damage and upregulate FANCD2/FANCI for ubiquitination</td>
</tr>
<tr>
<td><strong>FA recruitment complex</strong></td>
<td></td>
</tr>
<tr>
<td>FANCD2</td>
<td>Ubiquitinated when DNA damage sensed by FA core complex</td>
</tr>
<tr>
<td>FANCI</td>
<td>Ubiquitinated when DNA damage sensed by FA core complex</td>
</tr>
<tr>
<td><strong>FA effector proteins</strong></td>
<td></td>
</tr>
<tr>
<td>FANCD1 (BRCA2)</td>
<td>Homologous recombination</td>
</tr>
<tr>
<td>FANCJ (CHL1)</td>
<td>Helicase downstream of FANCD2/FANCI ubiquitination</td>
</tr>
<tr>
<td>FANCN (PALB2)</td>
<td>Regulates and interacts with FANCD1 after ubiquitination of FANCD2</td>
</tr>
<tr>
<td>FANCO (RAD51C)</td>
<td>Homologous recombination</td>
</tr>
<tr>
<td>FANCP (SLX4)</td>
<td>Endonuclease to cleave ICL</td>
</tr>
<tr>
<td>FANCQ (XPF)</td>
<td>Exonuclease to cleave ICL</td>
</tr>
<tr>
<td>FANCR (RAD51)</td>
<td>Homologous recombination</td>
</tr>
<tr>
<td>FANCS (BRCA1)</td>
<td>Unloading of replication machinery at stalled replication fork</td>
</tr>
<tr>
<td>FANCT (UBE2T)</td>
<td>E2 enzyme for ubiquitin transfer to substrates</td>
</tr>
<tr>
<td>FANCU (XRCC2)</td>
<td>Homologous recombination</td>
</tr>
<tr>
<td>FANCV (MAD2)</td>
<td>Cell cycle checkpoint protein</td>
</tr>
</tbody>
</table>

**Table 1.2: FA Repair factors:**
The above table depicts the key proteins involved in the Fanconi Anaemic Repair Pathway including their aliases (in parenthesis) where available (Moldovan and D’Andrea, 2009; Genois et al., 2014).
ICL damage can lead to stalling of a replication fork(s) and alteration of the chromatin structure. Together, this alteration in the DNA configuration is recognised by the FANCM helicase that goes on to recruit other components of the FA core complex (FANCA, -B, -C, -E, -F, -G and -L) to the lesion site, a process aided by various ancillary factors such as FAAP24, MHF1 and MHF2 (Moldovan and D’Andrea, 2009; Tan and Deans, 2017). Within this complex, FANCF acting as a scaffold protein, FANCA and FANCG function to guide the complex to nucleus and to site of DNA damage within the genome while FANCC and FANCE act as adapters for various substrates required for downstream activities: Note that the displacement of the replication fork does not occur during the assembly of the FA core complex but is reversed back along the newly synthesised DNA strands under the guidance of FANCM. The key role of the FA complex is to act as a ubiquitin ligase, an activity mediated by FANCL and FANCT, resulting in the monoubiquitination and activation of lysine residues present on FANCD2 and FANCI (Moldovan and D’Andrea, 2009; Ward et al., 2012).

Activation of the FANCD2/FANCI heterodimeric results in recruitment of FAN1 and FANCP (SLX4) to the site of DNA damage which together with FANCQ (XPF)-ERCC1 and MUS81-EME1 cleaves the DNA backbone at sequences 5’ and 3’ to the DNA lesion ((Moldovan and D’Andrea, 2009; Siddiqui et al., 2017). This results in the unhooking of the ICL from one of the DNA strands creating a gap within the other strand adjacent to where the newly synthesised DNA strands associated with replication forks is found (Part A of Figure 1.11). TLS polymerases along with DNA ligase fill in the gap and links the parental ICL-containing DNA molecule to one of the newly synthesised DNA strands. The ICL is then completely removed by a second round of NER/DNA polymerase and DNA ligase activity. During the above gap filling, the other newly synthesised DNA strand forms a DSB that is presumably recognised by the MRN complex. This guides components of the HR pathway (e.g. FANCD1, FANCR, FANCO, FANCU) to mediate recombination between the broken dsDNA molecule with the newly repaired DNA structure generated from the second round of NER/DNA polymerase and DNA ligase activities. This results in reformation of the Y-sharked forked structure that can serve as a template for recommencement of DNA replication (Wang, 2007).
Figure 1.1: ICL repair in dividing mammalian cells.
FA factors (dark blue) recognize crosslinked DNA (red and black) through stalling of a single replication fork (A) or convergence of two replication forks (B) resulting in the recruitment of FA effector proteins. These proteins unhook and process the DNA surrounding the ICL to create a double strand break which is repaired by the HR pathway (blue). Given this pathway is activated by stalling of the replication fork, leading and lagging strands are shown in green. The activity of TLS polymerases fills any remaining gaps in the DNA strand (blue). The ICL is completely removed from DNA by a second round of incision by NER factors. Image adapted from (Huang and Li, 2013; Krishnan, Tay and Ito, 2015).
Due to the bidirectional nature of DNA replication, ICLs can actually cause the stalling of two converging DNA replication forks (I. Chaudhury, Stroik and Sobeck, 2014; Zhang and Walter, 2014). Such X-shaped structures (Part B of Figure 1.11) can be resolved using the mechanism outlined above although in this case two of the other newly synthesised DNA strands both form DSBs. In this situation, the above HR mediated repair systems are activated with the newly repaired DNA strand serving as template. The outcome of this event is formation of two dsDNA molecules and the disassociation of the replication machinery.

1.7 ICL repair in T. brucei

To date reports relating to how trypanosomes resolve ICLs are limited although some components of the NER pathways (e.g. TbXPC, TbCSB, TbXPG and TbXPBz) may potentially be involved in this process (Machado et al., 2014). Recent work has established that T. brucei expresses TbSNM1, a member of the SNM1/PSO2 family of nucleases (Sullivan et al., 2015). This group of enzymes are characterised by the presence of metallo-β-lactamase (MβL; pfam12706) and β-CASP (named after its representative members CPSF, Artemis, SNM1 and PSO2; pfam10996) domains that together function to bind zinc atom(s) that is central to their catalytic activity (Allerston et al., 2015) (Figure 1.12). Most eukaryotic organisms express one SNM1/PSO2 homologue with these playing a key and specific role in ICL repair (Li and Moses, 2003): They do not participate in the resolution of any other form of DNA damage. Mammals are unusual in that they possess three distinct orthologs belonging to this family of nucleases with SNM1A (or DRLCE1A) functioning in solely ICL repair, SNM1B (or Apollo) playing additional roles in the processing DNA hairpin structures and telomeric maintenance, while SNM1C (or Artemis) is involved in generate antibody diversity through VDJ recombination (Dominski, 2007; Hazrati et al., 2008).

![Figure 1.12: Comparison of domain structure in SNM1/PSO2 nucleases.](image)

Schematic diagrams showing the domain alignment of the yeast (ScPSO2), human (HsSNM1A, HsSNM1B and HsSNM1C) and trypanosomal (TbSNM1) SNM1/SPO2 homologues (Sengerová et al., 2012).
For those members of the SNM1/PSO2 family that have been characterised, a 5′-3′ exonuclease and flap endonuclease catalytic activity have been reported dependent upon the substrate present which play a key role in the processing of unhooked ICLs (Li, Hejna and Moses, 2005; Hazrati et al., 2008). Unusually, these enzymes mediate the nucleolytic degradation of such unhooked regions up to and beyond the crosslink with this resulting in them playing a key role in the resolution of such damage. The structure of human SNM1A and B has recently been solved and shown to contain an MβL fold consisting of 11 β-sheets with the β-CASP domain inserted between strands 10 and 11 (Figure 1.13) (Allerston et al., 2015). Several charged residues have been implicated in zinc co-factor coordination although the number of metals ions appears to vary between the two mammalian isoforms; SMM1A binds one zinc ion, facilitated by residues H732, H734, D736, H737, H793, D815, D838, E993, and H994 while SNM1B interacts with two metal ions using these same conserved residues Binding of the nuclease to the template is via a wide DNA binding groove located between the two domains with the size of this region governing the ability of the enzyme to read through a crosslink.

Figure 1.13: Overall structure of SNM1A.
SNM1A, PDB 4B87, is shown as a cartoon representation with α-helices shown in purple and β-sheets in yellow. The zinc ion is centrally depicted as a grey sphere. The β-CASP domain (bottom) is an insert between MBL (top) strands 10 and 11.
1.8 Research Aims

The goal of this project centres upon deciphering the repair systems used by *T. brucei* to resolve ICLs. By analysing these mechanisms, we aim to unravel their importance to the pathogen with the future intention of exploiting these systems in drug development. The specific aims are to:

1. Identify which *T. brucei* DNA repair proteins are involved in resolving ICLs within the parasite.
2. Evaluate for any epistatic or non-epistatic interaction between components of the trypanosomal ICL repair system(s)
3. Assess the relationship between TbSNM1 structure & function by analysing the relative importance of key amino acids postulated to be involved in catalytic activity, nucleic acid substrate specificity and localisation.
4. Inform on the mechanism of action of clinically used and novel trypanocidal prodrugs
2. Material and Methods

2.1 Cell culturing and storage

2.1.1 Maintenance of bacterial strains

The *E. coli* strain XL1-Blue (*recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17(rK-, mK+), supE44, relA1, lac, [F*, *proAB*, *lacIqZΔM15::Tn10(tetr)]) was grown in NZCYM (10 g l\(^{-1}\) enzymatic casein digest, 1 g l\(^{-1}\) Casamino acids, 5 g l\(^{-1}\) yeast extract, 5 g l\(^{-1}\) NaCl, 0.98 g l\(^{-1}\) MgSO\(_4\); Melford Laboratories Ltd) broth or on plates (NZCYM broth supplemented with 15g l\(^{-1}\) bacteriological agar no. 1; Oxoid) at 37°C. Recombinant *E. coli* were cultivated in or on this medium supplemented with 100 µg ml\(^{-1}\) ampicillin (Sigma-Aldrich).

For long-term storage, bacterial strains were deposited at -80°C in 1 ml aliquots of growth medium containing sterile 20% (w/v) glycerol (Sigma-Aldrich) in 1.2 ml cryogenic vials (Nunc). To revive such reserves, a small scrapping (~10-50 µl) of the frozen bacterial line was transferred into NZCYM broth (5 ml) and cultured at 37°C.

2.1.2 Maintenance of *Trypanosoma brucei brucei*

The *Trypanosoma brucei brucei* (referred to here as *Trypanosoma brucei/T. brucei*) bloodstream form (BSF) trypanomastigotes stocks used in this project are listed in Table 2.1.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>T. brucei</em> Lister 427, antigenic type MITat 1.2, clone 221a</td>
<td>(Doyle JJ, Hirumi H, Hirumi K, Lupton EN, 1980)</td>
</tr>
<tr>
<td><em>T. brucei</em> 221 Tbsnm1Δ</td>
<td>(Sullivan et al., 2015)</td>
</tr>
<tr>
<td><em>T. brucei</em> 2TAG1</td>
<td>(Alsford et al., 2005)</td>
</tr>
<tr>
<td><em>T. brucei</em> 2TAG1 Tbsnm1Δ gfp-Tbsnm1</td>
<td>(Sullivan et al., 2015)</td>
</tr>
<tr>
<td><em>T. brucei</em> 2TAG1 gfp-Tbsnm1</td>
<td>(Sullivan et al., 2015)</td>
</tr>
<tr>
<td><em>T. brucei</em> 2T1 pBSF-TbNTR-9e10</td>
<td>(Wilkinson et al., 2008)</td>
</tr>
<tr>
<td><em>T. brucei</em> 2T1 RNAi-TbTAO</td>
<td>Prof J. Kelly, LSHTM</td>
</tr>
<tr>
<td><em>T. brucei</em> SMB</td>
<td>(Wirtz et al., 1999)</td>
</tr>
<tr>
<td><em>T. brucei</em> SMB RNAi-TbSODB</td>
<td>(Wilkinson et al., 2006)</td>
</tr>
<tr>
<td><em>T. brucei</em> 221 RH1(^R)</td>
<td>(Meredith et al., 2017)</td>
</tr>
</tbody>
</table>

Table 2.1: *Trypanosoma brucei* strains used in this project.

*T. brucei* BSF trypanomastigotes were cultured at 37°C under a 5% (v/v) CO\(_2\) atmosphere in HMI-9 (Hirumi H, 1989) (Invitrogen) media supplemented with 3g l\(^{-1}\) sodium bicarbonate (Sigma-Aldrich), 0.014% (w/v) β-mercaptoethanol (Sigma-Aldrich) and 10% (w/v) heat-inactivated Foetal Bovine Serum (Pan Biotech) (Hirumi H, 1989). Genetically modified
parasite lines were maintained in this medium containing 2.5 μg ml\(^{-1}\) hygromycin, 10 μg ml\(^{-1}\) blasticidin, 2 μg ml\(^{-1}\) puromycin, 2 μg ml\(^{-1}\) G418 or 2 μg ml\(^{-1}\) phleomycin. In certain experiments where activation of an inducible promoter was required, 1 μg ml\(^{-1}\) tetracycline (Sigma-Aldrich) was added to the growth medium.

For long-term storage, trypanosomal cultures were frozen slowly (2-3 days) to -80°C as a 1 ml aliquot (growth medium containing 20% (w/v) glycerol (Sigma-Aldrich)) in 1.2 ml cryogenic vials (Nunc) using a Cryo 1°C Mr. Frosty™ Freezing Container (Nalgene) containing isopropanol. Samples were then transferred into liquid N\(_2\). To revive parasite stocks, the entire 1 ml stablate was quickly thawed, transferred to fresh parasite growth medium (10 ml) and cultured as described above.

2.2 Trypanocidal compounds and DNA damaging agents

A range of anti-parasitic agents and treatments, many of which promote DNA damage, were used in this project. These were purchased from Sigma-Aldrich (methyl methanesulfonate (MMS), hydroxyurea (HU), phleomycin, CB1954), Melford Laboratories Ltd (blasticidin, puromycin, G418, hygromycin) or Cambridge Biosciences (mechlorethamine). Other chemicals were supplied free of charge by Prof. Mike Barrett (University of Glasgow; difluoromethylornithine (DFMO), megazol), Prof. Simon Croft (London School of Hygiene and Tropical Medicine; benznidazole, nifurtimox), Prof. Longqin Hu (Rutgers, The State University of New Jersey, New Jersey, USA; LH34), and Prof. Chris Scofield & Prof. Peter McHugh (University of Oxford; ML302-HCl, ML302-F Na, Si4-33, LBE314, Ceftriaxone). UV irradiation was performed with a Stratalinker UV crosslinker (Stratagene).

2.2.1 Plasmids

The precursor plasmids used in this study to construct the trypanosomal gene deletion/disruption-expression vectors reported here are pKO-TcFHM-HYG, pKO-TcFHM–NEO, pKO-TcFHM–BLA, pKO-TcFHM–PAC (all de Padua et al., 2017), pKO-TbSNM1-BLA, pKO-TbSNM1-PAC and pRPA-GFP-TbSNM1 (all Sullivan et al., 2015) (Appendix 1a).
2.3 Nucleic Acid Extractions

Plasmid DNA was extracted from *E. coli* using the GeneJet™ Plasmid Miniprep kit (ThermoFischer Scientific) according to the manufacturer’s on-line protocol (https://tools.thermofisher.com/content/sfs/manuals/MAN0013117_GeneJET_Plasmid_Miniprep_UG.pdf; last accessed 26/12/2017). Purified plasmid DNA was stored at -20°C.

*T. brucei* genomic DNA (gDNA) was extracted from Phosphate Buffer Saline (PBS: Sigma Aldrich) washed parasites (approximately 5×10⁷) using the GeneJet™ Genomic DNA purification kit (ThermoFischer Scientific) according to the manufacturer’s on-line ‘Cultured Mammalian Cells Genomic DNA Purification Protocol’ (https://tools.thermofisher.com/content/sfs/manuals/MAN0012656_Genomic_DNA_Purification_UG.pdf; last accessed 26/12/2017). Purified gDNA was stored at 4°C.

*T. brucei* total RNA was extracted from PBS washed parasites (approximately 5×10⁷) using the RNeasy® Mini kit with the QIAshredder homogenizer and the RNase-free DNase hit add-ons (all Qiagen) according to the manufacturer’s ‘Purification of Total RNA from Animal Cells using Spin Technology’ on-line (http://www.bea.ki.se/documents/EN-RNeasy%20handbook.pdf; last accessed 26/12/2017). All RNA samples were stored at -20°C.

2.4 Nucleic acid manipulation

2.4.1 cDNA synthesis

Complementary DNA (cDNA) was synthesized from purified *T. brucei* total RNA using the qScript™ cDNA Synthesis Kit (Quantabio). Briefly, in a 0.2 ml thin walled micro-centrifuge tube, parasite RNA (10 μl), nuclease-free H₂O (5 μl), 5X qScript Reaction Mix (containing an optimized buffer, magnesium, oligo(dT) and random primers, and dNTPs) (4 μl) and qScript reverse transcriptase (1 μl) were combined to give a 20 μl total volume. The reaction was then placed in a TC-412 thermal cycler (Techne) and subjected to once cycle at 5 minutes at 22°C, 60 minutes at 42°C, 5 minutes at 85°C and then held at 4°C.

2.4.2 DNA amplification

A variety of DNA fragments were amplified from different DNA templates (parasite genomic DNA, parasite cDNA, plasmid DNA *etc*) using oligonucleotide primers synthesised by Sigma Aldrich (see Appendix 1b for primer information) and dNTP stocks purchased from ThermoFisher Scientific. Dependent upon the downstream usage of the DNA fragment, some
reactions were performed using the high-fidelity Vent® DNA polymerase (New England Biolabs) while others were conducted with a low fidelity Taq DNA Polymerase (New England Biolabs) and their respective x10 buffers.

A typical DNA amplification reaction with a final volume of 50 μl was set up and contained 1 x Taq Buffer with KCl-Mg²⁺ free (20 mM Tris HCl, 10 mM (NH₄)₂SO₄, 10 mM KCl), 200 μM dNTPs, 1.5 to 5 mM Mg²⁺, 2.5 % (v/v) DMSO (Sigma Aldrich), ~0.1 ng μl⁻¹ template DNA, 50 pmol of each forward/reverse primer and 1 U Vent® or Taq DNA polymerase. Using a peqSTAR thermal cycler (PEQLab Biotechnologien GmbH) or TC-412 thermal cycler (Techne), a standard PCR programme was performed. This consisted of an initial denaturation step at 96 °C for 2 minutes, followed by 30 cycles of 96 °C for 30 seconds (denaturation), 55 °C for 30 seconds (annealing) and 72 °C for 30 seconds (extension). At the end of the 30th cycle, a further extension step of 10 minutes at 72°C was added and the samples were then incubated at 4°C. Dependent upon the template and primer combination used, the denaturation/annealing/extension conditions (temperature and time) were adjusted to optimise each amplification.

In some cases, and using MyTaq™ Red Mix (Bioline), a multiplex PCR method was used to generate several amplicons in a single reaction. A typical DNA amplification with a final volume of 25 μl contained 1x MyTaq™ Red Mix (MgCl₂, dNTPs, MyTaq DNA Polymerase and enhancers concentrations optimized by manufacturer), T. brucei gDNA (~20 ng) and forward/reverse primers (usually 1 pmol µl⁻¹; see Appendix 1b). Using a peqSTAR thermal cycler (PEQLab Biotechnologien GmbH) or TC-412 thermal cycler (Techne), a standard PCR programme was performed (see above).

2.4.3 Reverse transcription quantitative PCR

Reverse transcription quantitative PCR (RT-qPCR) was performed on T. brucei cDNA using PerfeCTa® SYBR® Green FastMix® kit (Quantabio). A typical reaction contained 1x PerfeCTa SYBR Green FastMix (MgCl₂, dNTPs, AccuFast Tag DNA polymerase, SYBR Green I dye and stabilizers concentrations optimized by manufacturer), cDNA (100 to 1000 ng) and a pre-optimised concentration of forward/reverse primer (usually 1 pmol µl⁻¹). In some cases, additional MgCl₂ was added to the above mix to optimise the reaction. Samples were loaded into the wells of a clear 96 well plate and an adhesive cover (Bio-Rad) used to seal the plate. Using a CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad), an amplification
programme was performed. This consisted of a single initial denaturation step at 95 °C for 2 minutes, followed by 30 cycles of 95 °C for 30 seconds (denaturation), 55 °C for 30 seconds (annealing) and 68 °C for 30 seconds (extension). At completion of the amplification, a melt curve analysis was carried out to ensure target specificity. Fluorescence data collected during the amplification was analysed using CFX96 Touch Software (Bio-Rad). From the resultant sigmoidal curves, cycle threshold (CT) values were obtained and exported into Excel 2007 (Microsoft). Data was analysed by the comparative CT method (Schmittgen TD, 2008) by normalizing the sequence of interest against a standardized control (Tbtert) that was examined in parallel (Brenndörfer and Boshart, 2010).

2.4.4 Site directed mutagenesis

The Single Primer Reactions In Parallel-Site Directed Mutagenesis (SPRINP-SDM) approach (Edelheit, Hanukoglu and Hanukoglu, 2009), as outlined in Figure 2.1, was used to introduce specific mutations into Tbsnm1 using the template plasmid pRPa-GFP-Tbsnm1 (Sullivan et al., 2015). A total of fourteen mutations were attempted in order to alter the amino acid of interest to alanine (primers used are shown in Table 2.2). Four of the residues evaluated were hypothesised to play a role in protein localisation (R428, R429, R430 & H431), one in mediating nucleic acid specificity (V518) and nine in catalytic activity/zinc co-factor binding (D40, H78, H80, D82, H83, H159, H178, D221 & H496). These were selected based on the sequence alignment shown in (Sullivan et al., 2015).

A typical 25 μl SPRINP-SDM reaction was set up by combining 1× Q5 reaction buffer containing 2 mM MgSO$_4$ (New England Biolabs), 200 μM dNTPs (New England Biolabs), 5% (v/v) DMSO (New England Biolabs), 500 ng pRPa-GFP-TbSNM1 template DNA, 10 pmol μl$^{-1}$ primer (either the forward or reverse; Table 2.2) and 0.5 U Q5 high fidelity polymerase (New England Biolabs) in a thin-walled 0.5 ml microcentrifuge tube. The tube contents were consolidated by centrifugation and the sample subject to DNA amplification using a TC-412 thermal cycler (Techne). A standard DNA amplification programme consisted of a single initial denaturation cycle (98 °C for 30 seconds), followed by 30 cycles of 96 °C for 10 seconds (denaturing stage), 50 °C for 30 seconds (annealing stage), 72 °C for 600 seconds (extension stage) before a single, final extension cycle (72 °C for 120 seconds). Samples were then maintained at 4 °C.
Figure 2.1: Single Primer Reactions In Parallel Site Directed Mutagenesis (SPRINP-SDM).

A forward and reverse reaction, each containing a primer modified with the desire mutation, was carried out separately. Both forward and reverse reactions were combined for annealing. There were three possible combination during annealing: 1. forward parental strand with reverse parental strand, 2. forward PCR product strand annealing to reverse PCR product strand, 3. parental strand with PCR product strand with the second option as the desired combination. DpnI treatment was used to remove methylated parental DNA (methylation not shown).
<table>
<thead>
<tr>
<th>Residue role /primer name</th>
<th>Primer sequence (5’→3’)</th>
<th>Resultant plasmid</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Localisation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tbsnm1-R428A-F</td>
<td>GGCAGACCCAGTGAAGACAGATGGGATGAGAAGGC</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-R428A-R</td>
<td>GCCCTCTTCACTGCCTTGCTTCACTGTGGTCTGCC</td>
<td>Tbsnm1-R428A</td>
</tr>
<tr>
<td>Tbsnm1-R429A-F</td>
<td>GGCAGACCCAGTGAAGACAGATGGGATGAGAAGGC</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-R429A-R</td>
<td>GCCCTCTTCACTGCCTTGCTTCACTGTGGTCTGCC</td>
<td>Tbsnm1-R429A</td>
</tr>
<tr>
<td>Tbsnm1-R430A-F</td>
<td>AGACAGCTGTAGAAGACAGATGGGATGAGAAGGCACC</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-R430A-R</td>
<td>GGTTGCTTCACTGCCTTGCTTCACTGTGGTCTGCC</td>
<td>Tbsnm1-R430A</td>
</tr>
<tr>
<td>Tbsnm1-H431A-F</td>
<td>CCAGTGAGAAGAAGACACCGAGAAGAAGGACACC</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-H431A-R</td>
<td>TGCGGTGCTCTCTCTCTCGACTTG</td>
<td>Tbsnm1-R431A</td>
</tr>
<tr>
<td><strong>Substrate specificity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tbsnm1-V51A-F</td>
<td>CTCTGTGCGTCAACGCATCAGAAGAAGACGTTCC</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-V51A-R</td>
<td>GAACGCTCTCTTTGATGCCGTTGCGCCACGAG</td>
<td>Tbsnm1-V51A</td>
</tr>
<tr>
<td><strong>Enzyme activity/co-factor binding</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tbsnm1-D40A-F</td>
<td>GTGGCTATCTCTGTGCGACAGGCTTGTGGTTGCACACA</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-D40A-R</td>
<td>TGTTGAAACAACGCTGCTCAACAGGATGAGCAAC</td>
<td>Tbsnm1-D40A</td>
</tr>
<tr>
<td>Tbsnm1-H78A-F</td>
<td>TTGTTTCTTCTCTGGTACATTCAATTCAGGATCAT</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-H78A-R</td>
<td>ATGGTCTGCAATGGAAATGCGAGAGGAAAGAAAAACAA</td>
<td>Tbsnm1-H78A</td>
</tr>
<tr>
<td>Tbsnm1-H80A-F</td>
<td>TTTCTCTCTGCACTTGCATCTGGCATATTACAGGCGG</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-H80A-R</td>
<td>CCGCTGTAATGATCCTGAGCTGAAATGAGAAGAAAGA</td>
<td>Tbsnm1-H80A</td>
</tr>
<tr>
<td>Tbsnm1-D82A-F</td>
<td>TCGACCTTCTCTTGGACACATTACACGCGTGATT</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-D82A-R</td>
<td>AATACCAGCTGTAAATGTGCACGAAATGGAAATGCGA</td>
<td>Tbsnm1-D82A</td>
</tr>
<tr>
<td>Tbsnm1-H83A-F</td>
<td>CACTCTCTCTCTGACAGACCTACAGCCTATTACT</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-H83A-R</td>
<td>AGTACTTCGCTGTATGCTACGAGGAAATGAGAATG</td>
<td>Tbsnm1-H83A</td>
</tr>
<tr>
<td>Tbsnm1-H159A-F</td>
<td>TTGATACCCGGCTAATGCTGCTCGGTAGTT</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-H159A-R</td>
<td>AACTGCACAGACATACTGCTGCGAGGATGATCAA</td>
<td>Tbsnm1-H159A</td>
</tr>
<tr>
<td>Tbsnm1-H178A-F</td>
<td>TTCGCACAGGATCTCTGGACACGGGATTTCCGC</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-H178A-R</td>
<td>GCCAAGAATCCACGCTTGCCGAGAATCTGCGGAA</td>
<td>Tbsnm1-H178A</td>
</tr>
<tr>
<td>Tbsnm1-D221A-F</td>
<td>GATGTCGCTTGCTTGGACATACATAGCCAA</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-D221A-R</td>
<td>TGCGCGTTATGATATTGCAAGAGAGACACATC</td>
<td>Tbsnm1-D221A</td>
</tr>
<tr>
<td>Tbsnm1-H496A-F</td>
<td>ATCCCTACAGTGAGGCACTGTTCTTTTGGCAA</td>
<td>pRPa-GFP-</td>
</tr>
<tr>
<td>Tbsnm1-H496A-R</td>
<td>TCCAGAAAACAGAATGCTCACCTGAGGAAAGA</td>
<td>Tbsnm1-H496A</td>
</tr>
</tbody>
</table>

**Table 2.2:** Primers used to create single point mutations in pRPA-GFP-TbSNM1 plasmid. Primers to mutate D40, H78, H80, D82, H83, H159, H178, D221, H496, R428, R429, R430, H431 & V518 residues into alanines in the pRPa-GFP-Tbsnm1 plasmid with resultant plasmids are shown. Highlighted sequences correspond to the mutated residues to alanine residues.

The two DNA amplification reactions (forward and reverse) associated with each mutation were combined and subject to an annealing temperature gradient. Here, the samples were incubated at 95 °C for 300 seconds, before sequentially transferring to 90 °C for 60 seconds, 80 °C for 60 seconds, 70 °C for 30 seconds, 60 °C for 30 seconds, 50 °C for 30 seconds and
40 °C for 30 seconds. The DNA were then digested at overnight 37 °C with DpnI (New England Biolabs).

2.4.5 Fractionation of DNA by conventional agarose gel electrophoresis

Conventional agarose gel electrophoresis was used to separate DNA molecules between 100 bp to 10 kbp. Dependent on the fragment sizes, the gels used varied in regard to their size and agarose concentrations while the voltage and run times were altered to achieve the desired fractionation result. A standard 0.8-1.2% (w/v) agarose gel was made by dissolving (0.8 to 1.2 g) agarose powder (Bioline) in (100 ml) 1X TAE (40 mM Tris base (Sigma-Aldrich), 40 mM acetic acid (VWR), 1 mM EDTA (Sigma-Aldrich)) containing ethidium bromide (0.1 mg l⁻¹: Sigma-Aldrich) by boiling using a microwave. The molten agarose was allowed to cool to around 60°C, dispensed into a sealed gel tray containing a comb and the agarose allowed to solidify. Once cast, the gel was placed into a gel electrophoresis tank containing 1X TAE supplemented with 0.1 mg l⁻¹ ethidium bromide. DNA samples (containing DNA loading dye) and a 1 kb GeneRuler™ (Thermo Fischer Scientific) were loaded onto the gel and the samples fractionated through the gel by applying a constant voltage (70-100 V) across the gel for 40 minutes to 2 hours: in some cases overnight electrophoresis (18 hours) using a 15 V constant voltage were performed. Migration of the samples through the agarose matrix was followed by monitoring the bromophenol blue and xylene cyanol dye fronts present in the samples loading buffer. When the DNA had migrated the desired distance, the agarose gel was visualized on a UV transliminator and documented (Syngene).

2.4.6 Restriction digestion

The restriction enzymes (and the associated 10x restriction buffers) used in this project are listed in Table 2.3 and were sourced from New England Biolabs and/or ThermoFisher Scientific. A typical 20 μl restriction digestion was set up by combining the DNA sample (4 μl) sample, 10x restriction buffer (2 μl), sterile distilled water (13 μl) and restriction enzyme (1 μl: ~10 units) in a 1.5 ml microcentrifuge tube. The tube contents were consolidated by centrifugation and incubated at an appropriate incubation temperature, typically 37 °C, for 1 to 3 hours. The reaction was halted by addition of a 1/10th volume (2 μl) of loading dye (10 mM Tris-HCl pH 7.6, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol, 60 mM EDTA) (Fermentas) to the mixture.
### Table 2.3: Restriction enzymes, their restriction sites and corresponding buffer used in this study.

The above table details restriction enzymes, the corresponding restriction site and resultant buffer used in the cloning steps of this project. The “^” symbol denotes the DNA cleavage site on the single strand shown.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Restriction site</th>
<th>NEB Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asc1</td>
<td>GG^CGCGGC</td>
<td>4</td>
</tr>
<tr>
<td>ApaI</td>
<td>GGCC^C</td>
<td>4</td>
</tr>
<tr>
<td>BamHI</td>
<td>G^GATCC</td>
<td>3</td>
</tr>
<tr>
<td>BglII</td>
<td>GCCNNNN^NGGC</td>
<td>3</td>
</tr>
<tr>
<td>Cla1</td>
<td>AT^CGAT</td>
<td>4</td>
</tr>
<tr>
<td>EcoRI</td>
<td>G^AATTTC</td>
<td>EcoRI</td>
</tr>
<tr>
<td>EcoRV</td>
<td>GAT^ATC</td>
<td>3</td>
</tr>
<tr>
<td>HindIII</td>
<td>A^AGCTT</td>
<td>2</td>
</tr>
<tr>
<td>KpnI</td>
<td>GGTAC^C</td>
<td>1</td>
</tr>
<tr>
<td>NotI</td>
<td>GC^GGCCGC</td>
<td>3</td>
</tr>
<tr>
<td>SacI</td>
<td>GAGCT^C</td>
<td>1</td>
</tr>
<tr>
<td>SacII</td>
<td>CCGC^GG</td>
<td>4</td>
</tr>
<tr>
<td>XbaI</td>
<td>T^CTAGA</td>
<td>4</td>
</tr>
</tbody>
</table>

2.4.7 DNA purification

DNA fragments (~200 bp to 10 kb) were purified in solution or from an agarose gel slice using the GeneJet™ Gel Extraction kit (ThermoFischer Scientific) according to the manufacturer’s on-line protocol (https://tools.thermofisher.com/content/sfs/manuals/MAN0012661_GeneJET_Gel_Extraction_UG.pdf; last accessed 26/12/2017). Purified DNA was stored at -20°C.

2.4.8 DNA Ligation

A typical ligation reaction (final volume of 30 µl) was set up by combining purified, digested vector (5 µl) and insert (21 µl) DNAs, 10x ligation buffer (50 mM Tris-HCl, 10 mM MgCl₂, 10 mM dithiothreitol, 1 mM ATP, pH 7.5) (3 µl) (New England Biolabs) and T4 DNA ligase (1 µl; 400 units) (New England Biolabs) in a 1.5 ml microcentrifuge tube. The tube contents were consolidated by centrifugation and incubated at room temperature for a minimum of 2 hrs. In parallel, a control ligation was performed where the insert DNA had been replaced with sterile distilled H₂O.

2.4.9 Construction of gene deletion/disruption vectors

A flow diagram illustrating the cloning steps used to generate the gene deletion/disruption vectors used in this project is shown in Figure 2.2. In this process, DNA fragments containing the 5’ or 3’ untranslated region (UTR)/coding sequence of the gene of interest (Tbcsb, TbmreII, Tbexo1, Tbrev2, Tbrev3, Tbfan1, Tbchl1, Tbhell1) were amplified from T. brucei gDNA and digested with SacI (or SacII where appropriate)/XbaI or ApaI/KpnI, respectively. These fragments were gel purified and sequentially cloned either side of a drug resistance
cassette containing genes that encode for hygromycin B phosphotransferase (*hyg*), neomycin phosphotransferase (*neo*), blasticidin-S deaminase (*bla*) or puromycin N-acetyltransferase (*pac*), plus ancillary sequences needed for the correct processing of the transcript. These above enzymes confer resistance to the selective agents hygromycin B, G418, blasticidin and puromycin, respectively. The primer combinations, band sizes and the resultant gene deletion/interruption vectors are listed in Table 2.4.

2.4.10 DNA Sequencing

The concentration of the template DNA was determined using a NanoDrop™ UV Spectrophotometer (ThermoFisher Scientific). An aliquot of this sample (~1 µg) along with an appropriate primer (20-30 pM) was then commercially sequenced by Eurofins Genomics. The resultant outputs were analysed on Chromas (http://technelysium.com.au/wp/chromas/) and BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.5 Bioinformatics

DNA sequences encoding for or having potential to encode for TbCSB (Tb927.7.4080), TbmRE11 (Tb927.2.4390), TbxEXO1 (Tb927.8.3220), TbvREV2 (Tb927.7.1090), TbvREV3 (Tb927.8.3290), TbvCHL1 (Tb927.10.3550), TbvHEL1 (Tb927.7.2970) and TbvFAN1 (Tb927.10.7540) along with their flanking regions were identified from the TriTryp database (http://tritrypdb.org/tritrypdb/) with homologues identified using a reciprocal best Blast hit analysis. The deduced protein sequences were aligned to the counterparts expressed by other organisms using CLUSTALΩ (http://www.ebi.ac.uk/Tools/msa/clustalo/) and domain structures evaluated using HMMR (http://www.ebi.ac.uk/Tools/hmmer/). To determine potential cellular localisation and possible amino acid targeting motifs, a variety of on-line tools were used including TrypTag (Dean, Sunter and Wheeler, 2017), PSORT II (https://psort.hgc.jp/), NucPred (http://www.sbc.su.se/~maccallr/nucpred/), Nuc-PLoc (http://www.csbio.sjtu.edu.cn/bioinf/) Nuc-PLoc() and cNLS Mapper (http://nlsmapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi). Phylogenetic relationships of DNA sequences were evaluated using the Phylogeny.fr “One Click” mode tool (http://www.phylogeny.fr/index.cgi) (Dereeper et al., 2008): sequence alignments were performed using MUSCLE (Edgar, 2004), curation using GBLOCKS (Castresana, 2000), phylogeny using PhyML 3.0 (Guindon et al., 2010) and tree rendering using TreeDyn (http://www.treedyn.org/). Molecular modelling was carried out
Figure 2.2: Cloning schematic to create disruption vectors.
A. Illustration depicting insertion of a PCR-amplified 5’ untranslated region (UTR)/coding sequence of the gene of interest (GOI) into a vector containing a drug resistance cassette (DRC) to facilitate disruption of genes (see Appendix 1a). Both DNA pieces were digested with SacI/XbaI and ligated together. B. Insertion of the 3’ GOI/UTR into the previously generated vector. These DNA fragments were digested with ApaI/KpnI, purified and ligated together to create the final disruption/deletion vector.
using Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) (Kelley et al., 2015) with data analysed (and in silico mutagenesis) carried out on PyMOL (https://pymol.org/2/).

2.6 Bacterial Transformation

Competent E. coli were generated using a CaCl₂-based technique (M. Mandel, 1970). Briefly, a bacterial stock culture was diluted 1:10 into fresh NZCYM medium (generally 1 ml stock in 10 ml fresh medium) and the new culture grown with aeration at 37°C for two hours: at this point the optical density at 600 nm of the culture was between 0.6-1.0. The cells were harvested by centrifugation (3,555 g for 5 minutes), suspended in an equal volume (generally 10 ml) ice cold 0.1 M CaCl₂ and incubated on ice for 10 minutes. The cells were re-harvested by centrifugation (3,555 g for 5 minutes) and the pellet suspended in a 1/10th volume (generally 1 ml) of ice cold 0.1 M CaCl₂. The DNA to be transformed (generally 30 µl of a ligation reaction or 1 µl purified plasmid DNA) was added to an aliquot (100 µl) of the CaCl₂ suspended competent cells, the mixture held on ice for 10-30 minutes, heat shocked at 42 °C for 2 minutes then immediately returned to ice for a further 2 minutes. The transformed bacteria were plated onto NZCYM agar plates containing an appropriate antibiotic for plasmid selection and left to incubate overnight at 37 °C to allow growth of colonies.

2.7 T. brucei nucleofection

T. brucei was transformed using the Amaxa® nucleofection system (Lonza) (Burkard, Fragoso and Roditi, 2007). Parasites (5 to 7 x10⁷ per transfection) in the exponential phase of growth were harvested by centrifugation (1,600 g for 10 minutes at room temperature), residual growth medium removed and the pellet resuspended in 100 µl Human T-cell Nucleofector® solution (Lonza): this solution was made no more than 5 minutes before use by combining 82 µl of reagent A with 18 µl reagent B in the Human T-cell Nucleofector® kit. The cell suspension was immediately transferred to a nucleofection cuvette containing ~10 µg DNA (in 10 µl) to be transformed and parasite electroporation performed using the Amaxa Nucleofector® device set to programme X-001. The cells were transferred into pre-warmed HMI-11 (generally 48 ml) medium and left at 37 °C. After at least 6 hours, selective antibiotics were added to the transformed culture, aliquots (generally 2 ml) transferred into the wells of a 24 well plate and the parasites incubated at 37 °C under a 5 % (v/v) CO₂ atmosphere. In some cases, a 1:10 dilution of the transformed culture were set up prior to addition of the selective antibiotics.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Nature of integration</th>
<th>5’ fragment region</th>
<th>3’ fragment region</th>
<th>Vectors generated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tb<em>mr</em>11</td>
<td>gene disruption (34%)</td>
<td>670 bp - 5’ region of Tb<em>mr</em>11 (Tb<em>mr</em>11-KO1/ Tb<em>mr</em>11-KO2)</td>
<td>768 bp - 3’ region of Tb<em>mr</em>11 (Tb<em>mr</em>11-KO3/ Tb<em>mr</em>11-KO4)</td>
<td>pKO-Tb<em>mr</em>11-hyg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>mr</em>11-neo</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>mr</em>11-bla</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>mr</em>11-pac</td>
</tr>
<tr>
<td>Tb<em>ex</em>o1</td>
<td>gene disruption (50%)</td>
<td>600 bp - 5’ region of Tb<em>ex</em>o1 (Tb<em>ex</em>o1-KO1/ Tb<em>ex</em>o1-KO2)</td>
<td>594 bp - 3’ region of Tb<em>ex</em>o1 (Tb<em>ex</em>o1-KO3/ Tb<em>ex</em>o1-KO4)</td>
<td>pKO-Tb<em>ex</em>o1-hyg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>ex</em>o1-neo</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>ex</em>o1-bla</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>ex</em>o1-pac</td>
</tr>
<tr>
<td>Tb<em>cs</em>b</td>
<td>gene deletion (100%)</td>
<td>425 bp - 5’ UTR of Tb<em>cs</em>b (Tb<em>cs</em>b-KO1/ Tb<em>cs</em>b-KO2)</td>
<td>538 bp - 3’ UTR of Tb<em>cs</em>b (Tb<em>cs</em>b-KO3/ Tb<em>cs</em>b-KO4)</td>
<td>pKO-Tb<em>cs</em>b-hyg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>cs</em>b-neo</td>
</tr>
<tr>
<td>Tb<em>rev</em>2</td>
<td>gene disruption (67%)</td>
<td>806 bp - 5’ region of Tb<em>rev</em>2 (Tb<em>rev</em>2-KO1/ Tb<em>rev</em>2-KO2)</td>
<td>809 bp - 3’ region of Tb<em>rev</em>2 (Tb<em>rev</em>2-KO3/ Tb<em>rev</em>2-KO4)</td>
<td>pKO-Tb<em>rev</em>2-hyg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>rev</em>2-neo</td>
</tr>
<tr>
<td>Tb<em>rev</em>3</td>
<td>gene disruption (80%)</td>
<td>344 bp - 5’ region of Tb<em>rev</em>3 (Tb<em>rev</em>3-KO1/ Tb<em>rev</em>3-KO2)</td>
<td>787 bp - 3’ region of Tb<em>rev</em>3 (Tb<em>rev</em>3-KO3/ Tb<em>rev</em>3-KO4)</td>
<td>pKO-Tb<em>rev</em>3-hyg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>rev</em>3-neo</td>
</tr>
<tr>
<td>Tb<em>he</em>l1</td>
<td>gene disruption (60%)</td>
<td>495 bp - 5’ UTR of Tb<em>he</em>l1 (Tb<em>he</em>l1-KO1/ Tb<em>he</em>l1-KO2)</td>
<td>482 bp - 3’ region of Tb<em>he</em>l1 (Tb<em>he</em>l1-KO3/ Tb<em>he</em>l1-KO4)</td>
<td>pKO-Tb<em>he</em>l1-hyg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>he</em>l1-neo</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>he</em>l1-bla</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>he</em>l1-pac</td>
</tr>
<tr>
<td>Tb<em>ch</em>l1</td>
<td>gene disruption (65%)</td>
<td>557 bp - 5’ region of Tb<em>ch</em>l1 (Tb<em>ch</em>l1-KO1/ Tb<em>ch</em>l1-KO2)</td>
<td>546 bp - 3’ region of Tb<em>ch</em>l1 (Tb<em>ch</em>l1-KO3/ Tb<em>ch</em>l1-KO4)</td>
<td>pKO-Tb<em>ch</em>l1-hyg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>ch</em>l1-neo</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>ch</em>l1-bla</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>ch</em>l1-pac</td>
</tr>
<tr>
<td>Tb<em>fa</em>n1</td>
<td>gene disruption (72%)</td>
<td>415 bp - 5’ region of Tb<em>fa</em>n1 (Tb<em>fa</em>n1-KO1/ Tb<em>fa</em>n1-KO2)</td>
<td>400 bp - 3’ region of Tb<em>fa</em>n1 (Tb<em>fa</em>n1-KO3/ Tb<em>fa</em>n1-KO4)</td>
<td>pKO-Tb<em>fa</em>n1-hyg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pKO-Tb<em>fa</em>n1-neo</td>
</tr>
</tbody>
</table>

**Table 2.4: Schematic for interruption/deletion of all genes analysed in this study.**

The above table lays out percentage disruption/deletion of every gene studied in this report, with percentage deletion/disruption of said gene calculated according to percentage of gene removed using this technique. Also shown are the primers (sequences can be found in Appendix 1b) used to amplify 5’ and 3’ UTRs/GOIs with resulting band sizes for each amplification reaction. This enabled the creation of all vectors listed in the last column.
Drug-resistant *T. brucei* were apparent 5-7 days post-transfection.

Using the above approach, a series of genetically engineered *T. brucei* lines were generated. The lines created were divided into those that were:

1. null for a single DNA repair enzyme (Table 2.5)
2. null for two DNA repair enzymes (Table 2.6)
3. expressing mutated version of Tbsnm1 in a Tbsnm1-deficient background (Table 2.7)
4. expressing Tbntr in a Tbmre11- or Tbexo1-deficient background (Table 2.8)

<table>
<thead>
<tr>
<th><em>T. brucei</em> line</th>
<th>Selectable markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>221 Tbsnm1Δ</td>
<td>Hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbsnm1Δ TbcsbΔ</td>
<td>Hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbsnm1Δ Tbexo1Δ</td>
<td>Hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbsnm1Δ Tbrev2Δ</td>
<td>Hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbsnm1Δ Tbrev3Δ</td>
<td>Hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbhel1Δ</td>
<td>Hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbchl1Δ</td>
<td>Hygromycin, puromycin</td>
</tr>
<tr>
<td>221 Tbfan1Δ</td>
<td>Hygromycin, G418</td>
</tr>
</tbody>
</table>

**Table 2.5: Parasite lines lacking one DNA repair enzyme.**

Each single null parasite line generated in this study with corresponding selectable markers used to maintain the engineered line.

<table>
<thead>
<tr>
<th><em>T. brucei</em> line</th>
<th>Selectable markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>221 Tbsnm1Δ Tbmre11Δ</td>
<td>Blasticidin, puromycin, hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbsnm1Δ TbcsbΔ</td>
<td>Blasticidin, puromycin, hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbsnm1Δ Tbexo1Δ</td>
<td>Blasticidin, puromycin, hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbsnm1Δ Tbrev2Δ</td>
<td>Blasticidin, puromycin, hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbsnm1Δ Tbrev3Δ</td>
<td>Blasticidin, puromycin, hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbsnm1Δ Tbhel1Δ</td>
<td>Blasticidin, puromycin, hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbsnm1Δ Tbchl1Δ</td>
<td>Blasticidin, puromycin, hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbsnm1Δ Tbfan1Δ</td>
<td>Blasticidin, puromycin, hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbmre11Δ TbcsbΔ</td>
<td>Blasticidin, puromycin, hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbmre11Δ Tbfan1Δ</td>
<td>Blasticidin, puromycin, hygromycin, G418</td>
</tr>
<tr>
<td>221 Tbmre11Δ Tbexo1Δ</td>
<td>Blasticidin, puromycin, hygromycin, G418</td>
</tr>
<tr>
<td>221 TbcsbΔ Tbexo1Δ</td>
<td>Blasticidin, puromycin, hygromycin, G418</td>
</tr>
</tbody>
</table>

**Table 2.6: Parasites lacking two DNA repair enzymes.**

Shown in this table are all double null lines generated for this report, including each selectable marker used to maintain the null line.
Table 2.7: Parasite lines expressing mutated version of Tbsnm1 in a Tbsnm1-deficient background.

Conditional cell lines generated for TbSNM1 structure and function assays in this study with lines null for Tbsnm1 and expressing a mutated GFP-tagged TbSNM1. Also shown are the selectable markers used to maintain all cell lines detailed. The parental 2TAG1 (2T1) is phleomycin resistant.

<table>
<thead>
<tr>
<th>T. brucei line</th>
<th>Selectable markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-D40A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-H78A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-H80A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-D82A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-H83A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-H159A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-H178A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-D221A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-R428A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-R429A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-R430A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-R431A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-V518A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbsnm1Δ gfp-Tbsnm1-H496A</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
</tbody>
</table>

Table 2.8: Tbntr expressing lines generated in this study.

Tbexo1 and Tbmre11 were both deleted in a background overexpressing TbNTR with all selectable markers used to maintain said lines shown above. The parental 2TAG1 (2T1) is phleomycin resistant.

<table>
<thead>
<tr>
<th>T. brucei line</th>
<th>Selectable markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>2TAG1 Tbmre11Δ Tbntr++</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
<tr>
<td>2TAG1 Tbexo1Δ Tbntr++</td>
<td>Blasticidin, puromycin, hygromycin</td>
</tr>
</tbody>
</table>

2.8 Analysis of recombinant parasites

2.8.1 Cumulative growth assays

T. brucei BSF parasites in the logarithmic phase of growth were seeded at $1 \times 10^4$ parasites ml$^{-1}$ in 5 ml HMI-11 medium in a well of a 6-well plate and then incubated at 37 °C in a 5 % (v/v) CO$_2$ atmosphere. Each day, the cell density of each culture was measured using a Neubauer haemocytometer. When the number of parasites reached ~1 to $10^6$ ml$^{-1}$, a new culture seeded at $1 \times 10^4$ parasites ml$^{-1}$ was set up. This analysis was carried out over a 8 to 14 day period. Growth curves were generated using GraphPad Prism (GraphPad Software Inc.). All growth assays were performed in triplicate and each count at each time point expressed as a mean ± standard deviation. Mean generation times were calculated using http://www.doubling-time.com/compute.php.
2.8.2 Anti-proliferation assays

All growth inhibition assays were carried out in 96-well plates. *T. brucei* BSF parasites in the logarithmic phase of growth were seeded at $1 \times 10^4$ cells ml$^{-1}$ in 200 µl HMI-11 medium containing various concentrations of the compound under study. Control assays performed in parallel included culturing cells grown in the absence of drug (100 % growth) and a medium background control. After incubation at 37 °C for 3 days, resazurin (Sigma Aldrich) was added to each well at a final concentration of 12.5 mg l$^{-1}$ (or 2.5 µg resazurin per well). The plates were further incubated for 6-8 hours at 37 °C before measuring the fluorescence of each culture using a Gemini Fluorescent Plate reader (Molecular Devices) at $\lambda_{EX} = 530$ nM and $\lambda_{EM} = 585$ nM with a filter cut off at 550 nM. The colour change (and subsequent fluorescence) of resazurin (blue) to resorufin (reddish pink) stems from the reduction of the input dye by viable cells with the change in fluorescence being proportional to the amount of cells present. Therefore, the highest fluorescence values were observed in untreated cultures, whereas the lowest was observed in the absence of cells (e.g. background control).

To calculate the fluorescence value at each compound concentration ($f_{lv}$), the following calculation was used:

$$f_{lv} = f_{raw} - f_{background}$$  \hspace{1cm} \text{equation 1}

where $f_{raw}$ is the raw fluorescence value at a given compound concentration and $f_{background}$ is the average background fluorescence value of the medium control.

The fluorescence value at a given compound concentration ($f_{lv}$) was then expressed as a % growth relative to the average fluorescence value for untreated cultures ($f_{drug=0}$):

$$\% \text{ growth at given [compound]} = \frac{f_{lv}}{f_{drug=0}} \times 100$$  \hspace{1cm} \text{equation 2}

Data expressed as % growth was used to construct dose response curves from which the compound concentration that inhibits cell growth by 50 % ($EC_{50}$) was extrapolated using the non-linear regression tool on GraphPad Prism (GraphPad Software Inc.). The statistical significance of any differences in parasite susceptibilities was assessed using the Student’s *t* test calculator (GraphPad Software Inc.).

2.8.3 Cell imaging

*T. brucei* BSF parasites in the logarithmic phase of growth and cultured with (or without) 1 µg ml$^{-1}$ tetracycline for 36 hours were pelleted (800 g for 10 minutes), washed twice in PBS before being suspended in 2% (w/v) paraformaldehyde in PBS at a density of $\sim 1 \times 10^7$ cells ml$^{-1}$. 

59
Following incubation at room temperature for 20 minutes, an aliquot (20 μl) of the parasite suspension (~1 x 10^5 cells) was air dried onto a single well of a 10-well printed microscope slide (Hendley-Essex). The slides were then incubated in 100 % (v/v) methanol at -20 °C for 30 minutes, washed once in PBS before being allowed to dry. Vectashield Mounting Medium containing 4',6-diamidino-2-phenylindole (DAPI) (Vectorshield Laboratories) was dispensed across the surface of the slide, a cover slip placed over the wells of that slide and a seal (clear nail varnish) applied around the edge of the coverslip. The seal was allowed to harden overnight at room temperature in the dark and the slides stored at 4 °C until use. Florescence, DAPI and phase images were captured using a Leica DMRA2 upright epi-fluorescent microscope fitted with a digital camera (Hamamatsu Photonics).

For cell cycle arrest assays, *T. brucei* were treated as described above except that the parasites were incubated for varying time periods in different concentrations of mechlorethamine prior to fixation, their immobilization onto a slide and the DAPI/phase images captured. The DAPI and phase images were then merged using ImageJ. To determine whether mechloethamine treatment affected the trypanosome, parasites at various stages in cell cycle were identified in asynchronous cultures by staining their nuclear (N) and mitochondrial (known as the kinetoplast (K)) genomes with DAPI. The ratio of these two DNA-containing structures within a single cell represents an excellent marker for the trypanosomal cell cycle, with *T. brucei* in the G1/S phase having a 1N1K arrangement, those in G2/M phase possessing a 1N2K ratio while cells displaying a 2N2K profile are in the post-M stage (Woodward and Gull, 1990; T.N. Siegel, D.R. Hekstra, 2008; Glover and Horn, 2012). For each cell line at each mechlorethamine treatment, the DNA staining pattern observed for a minimum of 600 cells was analysed and the number of cells at each cell cycle stage expressed as a % relative to the total parasite count.

### 2.8.4 Immuno-fluorescence assay

*T. brucei* BSF parasites in the logarithmic phase of growth and cultured in the presence of 30μm mechlorethamine for varying time periods, were pelleted (800 g for 10 minutes), washed twice in PBS before being suspended in 2% (w/v) paraformaldehyde in PBS at a density of ~1 x 10^7 cells ml^{-1}. Following incubation at room temperature for 20 minutes, an aliquot (20 μl) of the parasite suspension (~1 x 10^5 cells) was air dried onto a single well of a 10-well printed microscope slide (Hendley-Essex). The slides were then washed in PBS for 5 minutes, treated with 0.5 % Triton-X (Sigma) in PBS for 15 minutes followed by an additional wash of the slide.
in PBS (5 minutes). Slides were then incubated in blocking solution (50 % [v/v] Fetal Bovine Serum (FBS) in PBS) for 45 minutes followed by incubation in blocking solution (3 % (v/v) FBS in PBS) containing primary antibody (rabbit anti-γH2AX (supplied by Dr R. McCulloch, University of Glasgow (Glover and Horn, 2012) at a 1:250 dilution) for 45 minutes. Slides were then washed twice in PBS for 5 minutes with subsequent incubation of slides in blocking solution (3 % (v/v) FBS in PBS) containing secondary antibody (Goat anti-Mouse IgG (H+L) Alexa Fluor 488 at a dilution of 1:1000) for 45 minutes. Slides were finally washed three times in PBS, 5 minutes per wash. Vectashield Mounting Medium containing 4′,6-diamidino-2-phenylindole (DAPI) (Vectorshield Laboratories) was dispensed across the surface of the slide, a cover slip placed over the wells of that slide and a seal (clear nail varnish) applied around the edge of the coverslip. The seal was allowed to harden overnight at room temperature in the dark and the slides stored at 4 °C until use. Florescence, DAPI and phase images were captured using a Leica SP5 confocal microscope (Leica Microsystems (UK) Ltd. Images were subsequently measured for intensity levels using Fiji (https://fiji.sc/) with additional collation of images on Microsoft Powerpoint.

2.8.5 Parasite extracts

Parasites (5 to 8 x 10^7) in the logarithmic phase of growth were pelleted by centrifugation (1,640 g for 10 minutes), suspended in sterile PBS (1 ml) and transferred to a 1.5 ml microcentrifuge tube. The cells were harvested (16,000 g for 1 minute), the supernatant removed and the parasites lysed in Laemmli buffer (10 % (v/v) glycerol, 2.3 % (w/v) SDS, 5 % (v/v) β-mercaptoethanol, 62 mM Tris.Cl pH 6.8, 0.002 % (w/v) bromophenol blue) at a density of 1 x 10^6 cells μl^-1. Aliquots were denatured (96°C for 5 minutes) prior to separation alongside a PageRuler prestained Protein Ladder (ThermoFischer Scientific) on a 12 % SDS-PAGE in a 1x SDS-PAGE running buffer (25 mM Tris-HCl, 200 mM glycine, 0.1% (w/v) SDS) using a Mini-PROTEAN® II cell system (Bio-Rad) set at a constant voltage 120 V for 90-120 minutes at 4 °C. Gels were either stained for 1 hour in Coomassie Blue solution (0.05 % (w/v) Coomassie brilliant blue, 10 % (v/v) acetic acid, 50 % (v/v) methanol) followed by destaining in boiling distilled water (gels scanned using a Xerox Flatbed Scanner), or transferred to 0.45 μm Protan nitrocellulose membrane (GE Healthcare).
2.8.6 Protein blotting and immunodetection

Protein blotting was carried out using a Trans-Blot Semi-Dry Transfer Cell (Bio-Rad), in accordance with the manufacturer’s instructions. A piece of 0.45 µm Protan nitrocellulose membrane (GE Healthcare) and 8 pieces of standard thickness blotting paper (Whatman), all cut to the size of the SDS-PAGE gel, and the SDS-PAGE gel itself were soaked in transfer buffer (25 mM Tris, 192 mM glycine, 0.5 % (w/v) SDS, 20 % (v/v) methanol) for 15 minutes. Four piece of wetted blotting paper were laid on the anode of the transfer cell on to which was placed the nitrocellulose membrane. The SDS-PAGE gel was then sandwiched between the membrane and four additional pieces of wetted filter paper, with the cathode plate placed on top of the wetted filter paper, completing the transfer cell. A constant voltage of 15 volts for 1 hour was then applied to the apparatus and protein transfer confirmed by observing the presence of prestained markers on the membrane.

A near infrared fluorescence system was used for protein detection. Protein containing membranes were washed in blocking solution (5 % (w/v) marvel powdered milk, PBS, 0.1% (v/v) Tween 20) for one hour at room temperature. The membrane was transferred to a 50 ml centrifuge tube and challenged with primary antibody diluted in blocking buffer (20 ml) for 1 hour at room temperature with gentle rotation. Following copious washes (4 x 5 minutes washed) in PBS containing 0.1% (v/v) Tween 20 the membrane was then challenged with the secondary antibody, diluted in blocking buffer, for 1 hour at room temperature with gentle rotation. After four washes in PBS containing 0.1% (v/v) Tween 20 and two in PBS alone, detection of the near infrared fluorescence signal was monitored using an Odyssey® CLx infrared imaging system (LI-COR). Band signal intensity measurements were carried out using the Analysis Tool in Image Studio™ Lite (Li-COR Biosciences).

The primary antibodies used during this project were a rabbit anti-γH2AX (supplied by Dr R. McCulloch, University of Glasgow) at a 1:1,000 dilution and a rabbit anti-enolase (Prof J. Lukes, University of South Bohemia) at a 1:150,000 dilution. The secondary antibodies used were a goat anti-rabbit IRDye™ 800CW antibody (LI-COR) and an goat anti-mouse IRDye™ 680LT antibody (LI-COR), both diluted 1:5,000.
3. **Evaluating the extent of the *T. brucei* ICL REPAIRtoire**

3.1 **Identifying trypanosomal ICL repair proteins**

In many model organisms, ICL repair is a complex process. It involves the concerted action of factors from various DNA repair pathways (e.g. HR, NER, and FA etc.) that function to resolve this lesion in a cell cycle dependent manner. Work from the McHugh, Moses, Hashimoto, and Symington laboratories (K F Grossmann *et al.*, 2001; Lam, Krogh and Symington, 2008; Iyama *et al.*, 2015; Hashimoto, Anai and Hanada, 2016) has shown that enzymes such as SNM1, MRE11, CSB, EXO1, REV2 and REV3 all play a role in repairing ICLs. This chapter will involve the construction and characterisation of *T. brucei* null mutant lines lacking TbMRE11, TbCSB, TbEXO1, TbREV2 and TbREV3 with the goal of identifying which of these constitute components of the parasites ICL REPAIRtoire.

3.2 **Creation and validation of single null lines**

3.2.1 **Generation of gene disruption vectors**

To assess whether a *T. brucei* DNA repair enzyme plays a role in the repair of ICLs, a classical genetic-based approach was taken. This involved the construction of parasite lines that lack a specific activity by deleting or interrupting the encoding gene from the trypanosome genome. This can then be followed up using phenotypic screens to characterise such recombinant cells. To initiate this, a series of DNA constructs were designed to facilitate replacement (for *Tbcsb* (Tb927.7.4080)) or disruption (for *Tbmre11* (Tb927.2.4390), *Tbexo1* (Tb927.8.3220), *Tbrev2* (Tb927.7.1090), *Tbrev3* (Tb927.8.3290)) of each allelic copy of the gene target with an appropriate drug resistance marker. A schematic diagram outlining how such plasmids were made is shown in Figure 2.2. Briefly, DNA fragments containing the 5’ or 3’ flanking or coding sequences from the gene of interest were amplified from wild type *T. brucei* genomic DNA. These were then digested with restriction enzymes (5’ sequences were processed with SacI (or SacII, where appropriate) and XbaI while 3’ sequences were treated with ApaI and KpnI) and cloned either side of a DNA cassette containing a gene that encodes for the hygromycin B phosphotransferase (*hyg*), neomycin phosphotransferase (*neo*), blasticidin-S deaminase (*bla*) or puromycin N-acetyltransferase (*pac*) and supplementary sequences (*T. brucei* βα tubulin and αβ tubulin intergenic regions) needed for the correct processing of the transcript. The cloned fragments in the resultant gene interruption plasmids were confirmed using a combination of restriction mapping and DNA sequencing.
3.2.2 Development of null mutant lines

To generate parasite null mutant lines, the gene deletion/disruption plasmids (see Table 2.4) were digested to completion with SacI/KpnI, for vectors containing the neo, hyg, or pac resistance cassettes, or SacII/KpnI, for vectors containing the bla resistance cassette. The resultant fragments were purified, heat-treated at 70°C for 30 minutes to minimise contamination and then introduced into wild type *T. brucei* by nucleofection. To remove both allelic copies of the target gene, two rounds of nucleofection were performed to firstly create heterozygous and then null mutant lines. The latter cells were subsequently validated by DNA amplification approaches using genomic DNA or cDNA as template. Below, an example of the above strategy is given detailing how trypanosomes lacking Tbsb were made.

The plasmids pKO-Tbsb-neo and pKO-Tbsb-hyg, which facilitate deletion of Tbsb and are based around the neo and hyg resistance cassettes, were digested with SacI/KpnI and the resultant DNA fragments purified. The SacI/KpnI-treated pKO-Tbsb-hyg DNAs were introduced into wild type *T. brucei* with hygromycin resistant putative Tbsb heterozygote parasite clones obtained 5 to 7 days post-transformation. These cells were then transformed with SacI/KpnI-treated pKO-Tbsb-neo DNAs with hygromycin and G418 resistant putative Tbsb Δ null mutant trypanosomes selected 5 to 7 days post transformation.

To confirm that integration of the introduced DNA fragment had occurred at the correct genetic loci, DNA amplification reactions using gDNA extracted from *T. brucei* wild type and putative Tbsb Δ as template were performed. Such PCRs used primer combinations that generate biomarker fragments specific for the intact Tbsb gene or the hyg and neo disrupted alleles (Tbsb –hyg and Tbsb –neo, respectively). When using a primer combination designed to detect intact Tbsb (Tbsb-KO1/Tbsb-q2), DNA amplification generated a band of the expected size (~1.4 kb) from gDNA extracted from wild type parasites with no band(s) observed in DNA isolated from *T. brucei* Tbsb Δ cells (Figure 3.1b). In contrast, primer combinations that detect the Tbsb-hyg (Tbsb-KO1/hyg-2) or Tbsb-neo (Tbsb-KO1/neo-1) alleles generated amplicons of the predicted size (~1.7 and ~1.5 kb, respectively) but only from the gDNA purified from *T. brucei* Tbsb Δ null mutant line: No band(s) was observed when using *T. brucei* wild type gDNA as template.
Figure 3.1: Generation of *T. brucei* line lacking TbCSB.

A. Schematic representation of the Tbcsb allele and the effects of gene disruption using genes encoding for neomycin (*neo*) and hygromycin (*hyg*) resistance. P1 (Tbcsb-KO1), P2 (Tbcsb-q2), P3 (neo-2) and P4 (hyg-2) correspond to the regions where the primers anneal to in the appropriate allele.

B. DNA fragments obtained after amplification on gDNA extracted from *T. brucei* wild type (lane 1) and *T. brucei* TbcsbΔ (lane 2) using primer combinations that specifically amplify biomarkers corresponding to the Tbcsb (Tbcsb-KO1/Tbcsb-q2), Tbcsb-hyg (Tbcsb-KO1/hyg-2) or Tbcsb-neo (Tbcsb-KO1/neo-2) alleles.

C. DNA fragments obtained after amplification on cDNA derived from RNA extracted from *T. brucei* wild type (lane 1) or *T. brucei* TbcsbΔ (lane 2) using primer combinations that specifically amplify intact Tbcsb (Tbcsb-q1/Tbcsb-q2). In both cases, the integrity of the cDNA (and hence RNA) was evaluated by amplification of a control fragment, Tbtert.

To show that the *T. brucei* TbcsbΔ null mutant line was not expressing the interrupted gene, PCRs were performed using a primer combination (Tbcsb-q1/Tbcsb-q2) designed to generate a Tbcsb specific amplicon against cDNA generated from total RNA (Figure 3.1c). When using wild type cDNA as template, a single band of the expected size (~140 bp) was observed with no band(s) detected in material derived from *T. brucei* TbcsbΔ. To confirm that RNA had been extracted from both cell lines and that cDNA had indeed been made, control reactions amplifying Tbtert were conducted in parallel. For all tested samples a band of the expected size (~100 bp) was observed showing that the template nucleic acid was suitable for these studies.
Parasites lines null for other *T. brucei* genes (Tbmre11, Tbxo1, Tbre2 or Tbre3) were generated as described above. Validation of the resultant recombinant trypanosomes was performed using similar PCR strategies with this showing that integration of the input DNA fragments had successfully occurred into the appropriate loci in the parasite’s genome and confirmed that the Tbmre11, Tbxo1, Tbre2 or Tbre3 transcripts were not expressed by the corresponding null mutant. All the confirmatory data relating to the validation of *T. brucei* Tbmre11Δ, Tbxo1Δ, Tbre2Δ and Tbre3Δ is presented in Appendices 2a-2d.

### 3.3 Phenotypic analysis of null lines

#### 3.3.1 Growth analysis

To assess whether lack of a given DNA repair activity affected parasite growth, the growth properties of the null mutants generated in Section 3.2 (TbcsbΔ Tbmre11Δ, Tbxo1Δ, Tbre2Δ and Tbre3Δ) was determined and compared against wild type *T. brucei* grown in parallel (Figure 3.2). For all lines tested, three independent cultures containing 1 x 10⁴ parasites ml⁻¹ were established. Every day over a period of 8 days, the number of cells in each culture was determined using a haemocytometer with the cells in these cultures maintained in the exponential phase of growth: Cultures with a density approaching ~1x10⁶ cells ml⁻¹ were used to establish a new culture containing 1x10⁴ cells ml⁻¹. From the resultant cumulative cell density curves, the Tbxo1Δ, TbcsbΔ, Tbre2Δ and Tbre3Δ lines all exhibited a profile equivalent to wild type. From this data the growth rate, expressed as the mean generation or doubling time, of wild type cells was calculated to be 6.2 hours as compared to 6.5, 6.4, 6.1 and 6.3 hours displayed by the Tbxo1Δ, TbcsbΔ, Tbre2Δ, and Tbre3Δ lines, respectively. In contrast, cells lacking TbMRE11 exhibited a slightly longer mean generation time relative to wild type: Tbmre11Δ and wild type parasites had a mean generation time of 7.5 and 6.2 hours, respectively. This is in keeping with previous findings from the McCulloch and Cross Laboratories (Robinson, McCulloch, Conway, Browitt and David Barry, 2002; Tan, Leal and Cross, 2002).
Figure 3.2: Cumulative growth of *T. brucei* null mutants.
The cumulative cell density of *T. brucei* Tbexo1Δ, Tbmre11Δ, TbcshΔ, Tbrev2Δ and Tbrev3Δ (dashed line) cultures was followed for 8 days and compared against wild type *T. brucei* (solid line) cultures grown in parallel. Each data point represents the mean cell density ± standard deviation from three independent cultures.
3.3.2 Cell cycle analysis

To further evaluate the above growth characteristics, the cell cycle progression of Tbmre11Δ and TbcspΔ was evaluated relative to wild type and Tbsnm1Δ lines. Asynchronous cultures of bloodstream form parasites in the exponential phase of growth (cultures contained ~5 to 10 x 10⁵ trypanosomes ml⁻¹) were fixed, permeabilised and their nuclear (N) and mitochondrial (known as the kinetoplast (K)) genomes stained with DAPI. By determining the ratio of these two DNA-containing structures in a single trypanosome, the stage that each parasite is in within the cell cycle can be determined. T. brucei in the G1/S phase of the cell cycle have a 1K1N arrangement, those that possess a 2K1N ratio are said to be in the G2/M phase while trypanosomes displaying a 2K2N profile are in the post M phase (Figure 3.3a) (Woodward and Gull, 1990; T.N. Siegel, D.R. Hekstra, 2008; Glover and Horn, 2012). For wild type T. brucei, Tbsnm1Δ and TbcspΔ, most cells (~75%) in the asynchronous population were in G1/S phase, with ~20% in the G2/M phase and ~5% in the post M phase of the cell cycle (Figure 3.3b). In contrast, fewer Tbmre11Δ cells (~55%) were in the G1/S phase, with a concomitant increase (~40%) of cells in the G2/M phase and 5% in the post M phase of the cell cycle. This indicates that the increased mean generation time exhibited by Tbmre11Δ cells may be due the stalling of the replication forks. This causes a delay at this point in the cell cycle, an event probably driven by a reduced ability of this null mutant line to repair DNA damage via HR pathways (Grossmann, Ward and Moses, 2000; Robinson, McCulloch, Conway, Browitt and Barry, 2002; Tan, Leal and Cross, 2002). This cell cycle checkpoint checks the integrity of DNA before the final commitment of cell division, therefore, preventing cells containing damaged DNA progressing into mitosis. In contrast to our data, other researchers have not observed the above trait in relation to TbMRE11-deficient cells (Robinson, McCulloch, Conway, Browitt and Barry, 2002). Potentially this difference could be accounted for by the different assay systems used.
Figure 3.3: Cell cycle analysis of *T. brucei* null mutants.

A. The ratio of DAPI-stain kinetoplast (K) and one nuclear (N) genomes in one trypanosomes represents a frequently used marker for the *T. brucei* cell cycle. Images show cells with a 1K1N, 2K1N and 2K2N arrangement which is indicative of the G1/S, G2/M and post M stages, respectively.

B. The relative number of nuclear and mitochondrial genomes structures from asynchronous cultures of wild type, *Tbmre11Δ, Tbsnm1Δ* and *TbcshΔ* parasites was evaluated. The number of cells analysed per cell line is given above each bar.

3.3.3 Susceptibility of *T. brucei* null mutants towards DNA damaging agents

To evaluate whether lack of a given DNA repair activity affects parasite susceptibility to various types of DNA damaging agent, a series of growth inhibition assays were performed. The treatments used were hydroxyurea (which promotes base oxidation), phleomycin (causes DNA DSBs), methyl methanesulfonate (MMS; alkylation leading to DNA DSBs) and ultraviolet radiation (UV; pyrimidine dimer formation). In many organisms, specific pathways operate to repair the type of DNA damage these treatments induce e.g. the NER pathway functions in the repair of UV induced lesions while HR plays a central role in resolving DNA DSBs. In these inhibition assays, parasites (*T. brucei* wild type, *Tbmre11Δ, TbcshΔ, TbxolΔ, Tbv2Δ and Tbv3Δ; *Tbsnm1Δ* as analysed in parallel) were grown for 3 days in the presence...
of different concentrations of compound (hydroxyurea, phleomycin, MMS) or exposed to a single bolus prior to culturing for 3 days (UV) with the number of cells in each sample then determined using the vital dye resazurin. The resultant data was then plotted to generate dose response curves from which the treatment EC\textsubscript{50} value was calculated (see Materials and Methods). Figure 3.4 illustrates the type of dose response curves generated using \textit{T. brucei} wild type and \textit{Tbmre11Δ} against the four previously noted treatments with all the calculated EC\textsubscript{50} values shown in Table 3.1. A bar chart emphasising the fold differences displayed by each treatment against each cell lines is also shown (Figure 3.5).

Figure 3.4: Susceptibility of \textit{T. brucei} \textit{Tbmre11} null mutant towards various DNA damaging treatments.
Dose response curves of \textit{T. brucei} wild type (solid line) and \textit{Tbmre11Δ} (dotted line) to hydroxyurea, UV, phleomycin and MMS. All data points are mean values ± standard deviations from experiments performed in quadruplicate.
Table 3.1: Susceptibility of *Trypanosoma brucei* null lines to DNA damaging treatments.

Data represents EC$_{50}$ values of *T. brucei* wild type, *Tb*snm1Δ, *Tb*exo1Δ, *Tb*mre11Δ, *Tb*csbΔ, *Tb*rev2Δ and *Tb*rev3Δ against hydroxyurea, phleomycin, MMS and UV. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC$_{50}$ values between wild type and null mutant lines is given as fold difference. *Indicates significant differences in susceptibility (P < 0.0001 between wild type and null lines, as assessed by Student’s t test (GraphPad Software).
Figure 3.5: Comparison of *T. brucei* null mutant EC$_{50}$ values.
The susceptibility of *T. brucei* wild type and the various null mutant lines against hydroxyurea (HU), phleomycin (Phleo), methyl methanesulfonate (MMS) and ultraviolet radiation (UV), as judged by their EC$_{50}$ values, was compared and expressed as a fold difference.

Based on dose response curves and extrapolated EC$_{50}$ values, trypanosomes lacking TbSNM1, TbREV2 and TbREV3 are as equally sensitive to the tested DNA damaging treatments as wild type indicating that they do not play a front line role in resolving the resultant lesions. For TbSNM1, this confirms previously published data (Sullivan *et al.*, 2015) while the TbREV2 and TbREV3 data indicates that the different DNA polymerases involved in TLS may be able complement for each other: *T. brucei* expresses at least 5 such enzymes (Zurita-Leal, 2016).

In the case of the TbcsbΔ line, a wild type response was noted towards hydroxyurea, phleomycin and methyl methanesulfonate. In contrast, cells lacking this helicase were 5-fold more sensitive to UV relative to controls. This data confirmed previous findings that TbCSB is only involved in the parasites NER pathway (Machado *et al.*, 2014).

Interestingly, TbEXO1-deficient cells are only slightly more susceptible (2-fold) to phleomycin than wild type while exhibiting wild type sensitivities towards the other 3 treatments. It had initially been envisaged that this particular mutant should also be more sensitive to other dsDNA generating agents such as MMS but our observations contradict our initial hypothesis. This may be due to how the dsDNA break arises, for example due to base alkylation (MMS) or DNA intercalation (phleomycin), or which trypanosomal repair pathway predominants to fix such lesions. The data presented here indicates that the initial repair of phleomycin induced DNA damage results in base mismatches that are subsequently resolved.
using the MMR pathway, of which TbEXO1 is part. This indicates that MMS does not cause such secondary alterations.

For the final null line, Tbmre11Δ cells displayed increased sensitivity to 3 out of the 4 treatments relative to wild type. As expected and in keeping with the published literature, TbMRE11 deficient cells were more susceptible to agents that promote dsDNA breaks (phleomycin and MMS) with the HR pathway well known to play a role in repairing this type of damage (Robinson, McCulloch, Conway, Browitt and David Barry, 2002; Tan, Leal and Cross, 2002). Intriguingly, these mutants were also 5-fold more sensitive to UV, not previously documented in relation to trypanosomes. One reason to account for this latter observation may stem from the formation of double-strand DNA breaks caused by the arrest of replication forks at UV-induced damaged sites (Limoli et al., 2002).

### 3.3.4 Susceptibility of T. brucei null mutants towards mechlorethamine

The above growth inhibition assays were extended to investigate the phenotype displayed by T. brucei null mutants towards mechlorethamine, an archetypal ICL inducing agent. As noted previously, the resultant cell number data was plotted as dose response curves from which EC50 values were extrapolated (Figure 3.6: Table 3.2). Again, the fold differences displayed by each mutant cell line against the ICL inducing compound is emphasised in Figure 3.7.

Previous work has shown that TbSNM1 plays an important role in resolving the damage caused by mechlorethamine with our data confirming this earlier finding (Sullivan et al., 2015): Cells lacking TbSNM1 are >8-fold more susceptible to this compound as compared to controls. A similar alteration in sensitivity was also observed in cells deficient in TbMRE11, TbCSB or TbEXO1. Intriguingly, the difference in EC50 values displayed by these 3 mutants relative to wild type was not as great as that noted for Tbsnm1Δ cells; Tbmre11Δ, Tbcseq1Δ and Tbxol1Δ cells were 3-, 4- and 2-fold more susceptible to mechlorethamine than wild type. This could suggest the relative importance of each enzyme in the so-called T. brucei ICL REPAIRtoire. In contrast, cells lacking TbREV2 or TbREV3 displayed wild type sensitivities towards the ICL inducing agent indicating they play no significant role in resolving this type of damage.
Figure 3.6: Susceptibility of *T. brucei* null mutants towards mechlorethamine and DFMO.
Dose response curves of *T. brucei* wild type (solid line) and the various null mutants (dotted line) towards mechlorethamine and DFMO. All data points are mean values ± standard deviations from experiments performed in quadruplicate.
### Table 3.2: Null mutants towards mechlorethamine and DFMO.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>$T. brucei$ EC$_{50}$</th>
<th>Fold difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wild type</td>
<td>$Tbsnm1\Delta$</td>
</tr>
<tr>
<td>mechlorethamine</td>
<td>50.47 ± 4.39</td>
<td>6.43 ± 0.49*</td>
</tr>
<tr>
<td>DFMO</td>
<td>48.77 ± 2.19</td>
<td>38.61 ± 2.18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>$T. brucei$ EC$_{50}$</th>
<th>Fold difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wild type</td>
<td>$Tbexo1\Delta$</td>
</tr>
<tr>
<td>mechlorethamine</td>
<td>50.47 ± 4.39</td>
<td>24.94 ± 2.87*</td>
</tr>
<tr>
<td>DFMO</td>
<td>48.77 ± 2.19</td>
<td>46.41 ± 0.32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>$T. brucei$ EC$_{50}$</th>
<th>Fold difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wild type</td>
<td>$Tbmre11\Delta$</td>
</tr>
<tr>
<td>mechlorethamine</td>
<td>51.82 ± 0.32</td>
<td>18.76 ± 1.43*</td>
</tr>
<tr>
<td>DFMO</td>
<td>34.57 ± 0.87</td>
<td>42.95 ± 1.50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>$T. brucei$ EC$_{50}$</th>
<th>Fold difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wild type</td>
<td>$Tbcbsb\Delta$</td>
</tr>
<tr>
<td>mechlorethamine</td>
<td>70.08 ± 2.55</td>
<td>18.68 ± 1.64*</td>
</tr>
<tr>
<td>DFMO</td>
<td>57.14 ± 0.48</td>
<td>57.40 ± 5.44</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>$T. brucei$ EC$_{50}$</th>
<th>Fold difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wild type</td>
<td>$Tbrev2\Delta$</td>
</tr>
<tr>
<td>mechlorethamine</td>
<td>70.94 ± 1.25</td>
<td>60.17 ± 0.90</td>
</tr>
<tr>
<td>DFMO</td>
<td>36.75 ± 0.57</td>
<td>31.40 ± 0.54</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>$T. brucei$ EC$_{50}$</th>
<th>Fold difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wild type</td>
<td>$Tbrev3\Delta$</td>
</tr>
<tr>
<td>mechlorethamine</td>
<td>61.94 ± 1.90</td>
<td>75.20 ± 2.22</td>
</tr>
<tr>
<td>DFMO</td>
<td>36.75 ± 0.57</td>
<td>42.97 ± 0.64</td>
</tr>
</tbody>
</table>

EC$_{50}$ values of $T. brucei$ wild type, $Tbsnm1\Delta$, $Tbexo1\Delta$, $Tbmre11\Delta$, $Tbcbsb\Delta$, $Tbrev2\Delta$ and $Tbrev3\Delta$ against mechlorethamine and DFMO. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC$_{50}$ values between wild type and null mutant lines is given as fold difference. *Indicates significant differences in susceptibility (P < 0.0001) between wild type and null lines, as assessed by Student’s $t$ test (GraphPad Software).
The susceptibility of T. brucei wild type and the various null mutant lines against mechlorethamine and DFMO, as judged by their EC$_{50}$ values, was compared and expressed as a fold difference.

To confirm that the above phenotyping was due to the engineered gene deletion/disruption events and not due to off target effects, the susceptibility of all mutants towards the non-DNA damaging trypanocidal agent DFMO was analysed. In all cases the null lines were as equally sensitive to this compound as each other and the wild type, providing evidence that the above susceptibility profiles are specific for each DNA damaging treatment. The above data implicates TbSNM1, TbCSB, TbMRE11 and TbEXO1 in the bloodstream form T. brucei ICL repair system and thus constituting a part of this parasites ICL REPAIRtoire.

3.3.5 Production of DNA double strand breaks by mechlorethamine

Previous work on the formation of ICLs has shown that crosslinked DNA can be further processed to form double strand breaks (Li and Moses, 2003). In many cases, nuclear damage (some of which may include DSBs), results in the phosphorylation of histone H2A to form γH2A, with this post-translational modification representing a useful biomarker for detecting DNA damage (Fernandez-Capetillo, Allis and Nussenzweig, 2004). This response is well conserved in eukaryotes and has been shown to take place in T. brucei (Glover and Horn, 2012).

To demonstrate whether mechlorethamine promotes nuclear damage which is associated with this phosphorylation event, wild type parasites were treated with mechlorethamine (30, 10 and 3 µM) with samples taken from these cultures at various time points over an 8-hour period. Protein extracts from these cells were generated and probed for γH2A formation by
western blot analysis using antiserum against this modified histone and enolase as a loading control (Figure 3.8). The band intensities of γH2A and enolase for each treatment on each cell line were determined using Image Studio™ Lite (Li-COR Biosciences) with the signals normalised against untreated controls (Figure 3.9).

For all mechlorethamine concentrations tested, a gradual increase in the γH2A signal is observed peaking around 4 hours after initiating treatment. The marker then began to decline over the next 4 hours indicating resolution of the DNA damage. When these studies were extended to the null lines, various outcomes were observed. TbSNM1- and TbEXO1-deficient cells behaved in a similar way to wild type cells. It is tempting to speculate that this nuclear damage is due to ICL-induced DSBs and that these enzymes play no role in processing the such lesions. In TbcshbΔ cells, the formation of γH2A was delayed, peaking around 6 hours after initiating treatment and staying at the elevated level for the remainder of the experiment. This indicates that these null cells are unable to effectively recognise and resolve the ICL-induced DSBs: In other organisms, stalling of the transcriptional machinery at such sites leads to post-translational modification of the RNA polymerase which then acts to recruit CSB to the DSB (Batenburg et al., 2015). Intriguingly, when following γH2A in cells lacking TbMRE11, no alteration in signal intensity was observed suggesting this enzyme plays an important role in detecting nuclear damage, which leads to phosphorylation of histone H2A (Lavin, 2008; Zha, Boboila and Alt, 2009; Panier and Durocher, 2013).

To further evaluate mechlorethamine and DSB formation in mutant lines, cells treated for 8 hours in the presence of the ICL inducing agent (30 µM) were analysed using immunofluorescence microscopy to detect γH2A (Figure 3.10). For wild type, Tbsnm1Δ, TbcshbΔ and Tbeoxo1Δ, a signal throughout the whole nucleus was observed showing that nuclear damage formation occurs throughout the whole genome; no discrete foci were observed, even when analysing earlier time points (data not shown). In contrast, and confirming the western blot data, no signal was observed in mechlorethamine treated TbMRE11 null parasites.
Figure 3.8: γH2A formation in mechlorethamine treated wild type and null lines.
*T. brucei* wild type, Tbsnm1Δ, Tbmre11Δ, TbcshΔ and Tbexo1Δ were treated with 30, 10 or 3μM concentrations of mechlorethamine. Cell lysates were generated at time intervals (0, 1, 2, 4, 6, and 8 hours) and analysed by western blot using anti-*T. brucei* enolase (loading control) and anti-*T. brucei* γH2A antiserum.
Figure 3.9: Comparison of γH2A signals in mechlorethamine treated T. brucei null mutants. The γH2A and enolase signal intensities obtained from western blot analysis were determined using Image Studio™ Lite (Li-COR Biosciences) with the signals normalised against untreated controls.
**Figure 3.10: Comparison of $\gamma$H2A formation in mechlorethamine treated *T. brucei* null mutants.**

*T. brucei* wild type, Tbsnm1Δ, Tbmre11Δ, TbsbΔ and Tbxo1Δ were treated with 30µM of mechlorethamine for 8 hours and analysed by immunofluorescence assay using anti-*T. brucei* $\gamma$H2A antisera. Florescence, DAPI and phase images were captured using a Leica SP5 confocal microscope (Leica Microsystems (UK) Ltd.) and were subsequently processed using Fiji.
3.4 Chapter Summary

In this chapter we have:

1. Successfully deleted Tbsb and disrupted Tbxo1, Tbmre11, Tbrev2 and Tbrev3 from the T. brucei genome to create null mutant lines
2. Demonstrated that, in agreement with the published literature, cells lacking TbMRE11 display a growth defect due to a delay in the progression through the G2/M phase of the cell cycle
3. Shown that cells lacking TbMRE11 are more susceptible to chemical (in agreement with the published literature) and UV-induced dsDNA breaks relative to controls
4. Confirmed that TbCSB deficient cells are more susceptible to UV DNA damage
5. Established that TbMRE11, TbCSB and TbEXO1 represent key components of the T. brucei ICL REPAIRtoire
6. Shown that TbREV2 and TbREV3 do not function in ICL repair in the parasite
7. Revealed that mechlorethamine induced ICLs in trypanosomes promote nuclear damage, most likely in the form of dsDNA breaks
8. Revealed that TbMRE11 plays a crucial role in the formation of γH2A at sites of nuclear damage
4. Studying the interplay of ICL repair proteins

The interplay between components of the *S. cerevisiae* / human ICL repair system has been evaluated with studies showing these organisms express multiple, complementary and overlapping mechanisms to mediate repair of this type of DNA damage (Hashimoto, Anai and Hanada, 2016). However, the precise molecular basis of these pathways in a given organism remains unclear. Broadly, these ICL repair mechanisms operate in a cell cycle dependent manner with HR based pathways predominating during the S-phase, while through the remainder of the cell cycle factors from NER, TLS, and MMR prevail (Zhang and Walter, 2014). In yeast, the assignment of a DNA repair activity to a given system was established using a genetics-based approach through the generation of fungal lines lacking two or more activities. The susceptibility of such cells to ICL inducing agents was determined to reveal whether the components under investigation function in an epistatic or non-epistatic manner. Here, we have extended this approach to evaluate the interplay of the trypanosomal ICL repair factors identified previously (TbSNM1; (Sullivan et al., 2015)) and in the earlier chapter (TbMRE11, TbCSB and TbEXO1).

4.1 Evaluating the central role of TbSNM1 in ICL repair

4.1.1 Creation and validation of *T. brucei* null mutant lines

As an initial step to analysing the relationship between components of the trypanosomal ICL repair system, the genes encoding for TbEXO1, TbMRE11 or TbCSB were deleted or disrupted in a *Tbsnm1*-deficient *T. brucei* line, thereby creating a series of double null mutant lines (Table 2.6). Although not primarily involved in resolving ICLs, TbREV2 and TbREV3 were also taken forward in this analysis to determine whether their activities in ICL repair become apparent in the absence of TbSNM1: Some yeast ICL factors are only apparent when the activity of another DNA repair protein is missing (Ward *et al.*, 2012). The *hyg*- and *neo*-based gene integration vectors used in the previous chapter were digested to completion with SacI/KpnI, the fragments purified and then heat-treated (70°C for 30 minutes) to minimise contamination. The digested DNAs were sequentially introduced into *T. brucei Tbsnm1Δ* cells by nucleofection (in this mutant line, *Tbsnm1* has been disrupted using gene integration vectors based around *pac* and *bla*; (Sullivan *et al.*, 2015)) to generate heterozygous and subsequent null mutant lines. The recombinant cell lines were then validated by DNA amplification-based approaches using gDNA or cDNA templates. Below, an example of the above strategy is given detailing how Tbsnm1-deficient trypanosomes also lacking Tbscb were made.
Purified SacI/KpnI-digested pKO-Tbcsb-hyg DNA fragments were introduced into *T. brucei* Tbsnm1Δ cells with hygromycin/blasticidin/puromycin resistant parasite clones (putatively designated as Tbsnm1Δ TbsnbΔc) obtained 5 to 7 days post-transformation. These cells were subsequently transformed with purified SacI/KpnI-digested pKO-Tbcsb-neo DNA fragments with selection carried out using blasticidin/puromycin/hygromycin/G418. Clones resistant to the four selective agents (putatively designated as Tbsnm1Δ TbsnbΔcΔnull mutant trypanosomes) were obtained 5 to 7 days post-transformation.

Confirmation that the input DNA fragments had integrated at the correct genetic loci in the trypanosome genome was carried out by PCR (Figure 4.1). Here, gDNA extracted from four *T. brucei* lines (wild type, Tbsnm1Δ, TbsnbΔc and Tbsnm1Δ TbsnbΔc) was probed for the presence or absence of intact Tbsnm1 and Tbsnb, and their disrupted alleles (Tbsnm1–bla; Tbsnm1–pac; Tbsnb–hyg; Tbsnb–neo). When using a primer combination to detect intact Tbsnm1 (Tbsnm1-KO1/Tbsnm1-seq1), an amplicon of ~1.3 kb was detected only in gDNA extracted from *T. brucei* wild type and TbsnbΔ (no bands observed using Tbsnm1Δ and Tbsnm1Δ TbsnbΔ gDNA) (Figure 4.1c). In contrast, when probing for the Tbsnm1–bla (Tbsnm1-KO1/bla-1) or Tbsnm1–pac (Tbsnm1-KO1/pac-1) alleles, DNA fragments of the predicted size (~1.1 and ~1.7 kb, respectively) were generated only by gDNA purified from *T. brucei* Tbsnm1Δ and Tbsnm1Δ TbsnbΔ lines. Analysis of the Tbsnb loci using primers to detect intact Tbsnb (Tbsnb-KO1/Tbsnb-q2) amplified a band of the expected size (~1.4 kb) only from gDNA extracted from *T. brucei* wild type and TbsnbΔ (no band noted when using gDNA derived from TbsnbΔ and Tbsnm1Δ TbsnbΔ cells). When probed with Tbsnb-KO1/hyg-2 or Tbsnb-KO1/neo-2, primer combinations used to detect for the Tbsnb–hyg and Tbsnb–neo alleles respectively, bands of the expected size (~1.7 and ~1.5 kb) were detected only in gDNA derived from TbsnbΔ and Tbsnm1Δ TbsnbΔ.

Additional validation of the double null mutant line was performed aimed at assessing whether Tbsnm1 and Tbsnb were expressed by *T. brucei* Tbsnm1Δ TbsnbΔ cells. DNA amplification reactions were performed on cDNA templates generated from total RNA using primer combinations that detect Tbsnm1 (Tbsnm1-q1/Tbsnm1-q2) or Tbsnb (Tbsnb-q1/Tbsnb-q2). For RNA extracted from *T. brucei* wild type and TbsnbΔ a single, correctly sized (~110 bp) amplicon corresponding to Tbsnm1 was observed with no band(s) detected in material derived from Tbsnm1Δ and Tbsnm1Δ TbsnbΔ cells (Figure 4.1d). Similarly, a single, correctly sized (~140 bp) amplicon corresponding to Tbsnb was only observed in reactions using cDNA.
derived from *T. brucei* wild type and Tbsnm1Δ cDNA (and hence RNA) integrity used in these reactions was evaluated by probing the template for Tbert (Tbert-F/Tbert-R). In all cases, an amplicon of ~100 bp was noted showing that the template nucleic acid was appropriate for these studies.

Together, the above DNA amplification reactions validate the *T. brucei* Tbsnm1Δ TcsbΔ double null mutant line, confirming that Tcsb had indeed been successfully disrupted in Tbsnm1Δ parasites leading to the associated loss in gene expression.

**Figure 4.1: Validating the *T. brucei* Tbsnm1Δ TcsbΔ null mutant line.**
A. Diagram illustrating the Tbsnm1 allele and resultant changes upon disruption of this locus using genes encoding for blasticidin (bla) and puromycin (pac) resistance. P1 (Tbsnm1-KO1), P2 (Tbsnm1-seq1), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.

B. Schematic representation of the Tcsb allele and effects of disrupting this allele with genes encoding for neomycin (neo) and hygromycin (hyg) resistance. P1 (Tcsb-KO1), P2 (Tcsb-q2), P3 (neo-2) and P4 (hyg-2) correspond to regions where the primers anneal at the specific allele.

C. Amplicons corresponding to the Tbsnm1 (Tbsnm1-KO1/Tbsnm1-seq1), Tbsnm1-bla (Tbsnm1-KO1/bla-1), Tbsnm1-pac (Tbsnm1-KO1/pac-1), Tcsb (Tcsb-KO1/Tcsb-q2), Tcsb-hyg (Tcsb-KO1/hyg-2) or Tcsb-neo (Tcsb-KO1.neo-2) were generated from template gDNA extracted from *T. brucei* wild type (lane 1), *T. brucei* Tbsnm1Δ (lane 2), *T. brucei* TcsbΔ (lane 3) and *T. brucei* Tbsnm1Δ TcsbΔ (lane 4).

D. DNA fragments amplified from cDNA derived from RNA extracted from *T. brucei* wild type (lane 1), Tbsnm1Δ (lane 2), TcsbΔ (lane 3) and Tbsnm1Δ TcsbΔ (lane 4) using primer combinations that specifically amplify intact Tbsnm1 (Tbsnm1-q1/Tbsnm1-seq1), Tcsb (Tcsb-q1/Tcsb-q2) or Tbert (Tbert-F/Tbert-R).
Double null mutant parasites lines (Tbsnm1Δ Tbmre11Δ, Tbsnm1Δ TbcshΔ, Tbsnm1Δ Tbrev2Δ and Tbsnm1Δ Tbrev3Δ) were generated and validated as described above. All the confirmatory data relating to validation is shown in Appendix 3a-3d.

4.1.2 Phenotypic analysis of T. brucei double null lines

4.1.2.1 Growth analysis of T. brucei double null lines

To evaluate whether lack of TbSNM1 and TbCSB, TbMRE11, TbEXO1, TbREV2 or TbREV3 affects parasite growth, the growth properties of the double null mutants was determined and compared against wild type T. brucei (Figure 4.2). From the resulting cumulative cell density curves, Tbsnm1Δ Tbexo1Δ, Tbsnm1Δ TbcshΔ, Tbsnm1Δ Tbrev2Δ and Tbsnm1Δ Tbrev3Δ all grew at roughly the same rate as wild type, with mean generation times of 7.0, 6.7, 7.0 and 6.6 hours respectively, in comparison to 6.6 hours for wild type cells. In contrast, Tbsnm1Δ Tbmre11Δ cells exhibited a slightly longer mean generation time relative to wild type (7.8 hours vs 6.6 hours) which is comparable to the growth properties previously noted for Tbmre11Δ cells (See Chapter 3; Robinson et al., 2002; Tan, Leal and Cross, 2002).

4.1.2.2 Susceptibility of T. brucei double null lines towards DNA damaging agents

Growth inhibition assays were carried out on the double null mutant lines to determine whether these cells exhibit altered susceptibility to different DNA damaging agents relative to wild type and single null mutant controls. Assays were performed as described previously (see Materials and Methods and Chapter 3) with the data used to create dose response curves from which EC₅₀ values were derived. Figure 4.3 illustrates dose response curves generated from T. brucei wild type, Tbsnm1Δ, Tbmre11Δ, and Tbsnm1Δ Tbmre11Δ cells against the treatments used with concomitant EC₅₀ values given in Table 4.1. A bar chart emphasising the fold differences displayed by each treatment against each cell line is also shown (Figure 4.4).
Figure 4.2: Cumulative growth of *T. brucei* double mutant lines.
The cumulative cell density of *T. brucei* Tbsnm1Δ Tbeso1Δ, Tbsnm1Δ Tbmre11Δ, Tbsnm1Δ TbscsbΔ, Tbsnm1Δ Tbrev2Δ, and Tbsnm1Δ Tbrev3Δ (▲) cultures was followed for 8 days and compared against *T. brucei* wild type (●) cells, grown in parallel. Each data point represents the mean cell density ± standard deviation from three independent cultures.
Figure 4.3: Susceptibility of *T. brucei* Tbsnm1Δ Tbmre11Δ null line towards various DNA damaging treatments.
Dose response curves of *T. brucei* wild type (●), Tbsnm1Δ (■), Tbmre11Δ (♦), and Tbsnm1Δ Tbmre11Δ (ο) cell lines to hydroxyurea, UV, phleomycin and MMS. All data points are mean values ± standard deviations from experiments performed in quadruplicate.

For most of the DNA damaging treatments tested, the susceptibility displayed by the double null mutant lines was equivalent to the more sensitive phenotype shown by the single null for each gene pairing. For example, all Tbscsb-deficient cell types are approximately 5-fold more susceptible to UV than wild type or Tbsnm1Δ lines such that the Tbsnm1Δ TbscsbΔ line have an EC50 similar to TbscsbΔ parasites. As such, lack of TbSNM1 and TbCSB activities in the same cell does not lead to an increase in susceptibility to UV. The only situation where a combinatorial effect on susceptibility was observed in cells lacking TbSNM1 and TbMRE11. Intriguingly, such double null mutant parasites were more than 20-fold more sensitive to UV-induced lesions relative to *T. brucei* wild type and Tbsnm1Δ and 4-fold more sensitive relative to Tbmre11Δ. This difference may be an experimental artefact possibly due to how the UV dosage was applied or could reflect that TbSNM1 does plays a secondary role in resolving UV-induced damage, a phenotype that is only apparent in cells comprised in the HR pathway.
Table 4.1: Susceptibility of *T. brucei* double null lines to DNA damaging treatments.

Data represents EC\(_{50}\) values of various *T. brucei* lines against hydroxyurea, phleomycin, MMS and UV. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC\(_{50}\) values between wild type and single/double null mutant lines is shown as fold difference, in parentheses. *Indicates significant differences in susceptibility (P < 0.0001 between wild type and null lines, as assessed by Student’s t test (GraphPad Software).

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Wild type</th>
<th><em>T. brucei</em> EC(_{50})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Tbsnm1A</em></td>
<td><em>Tbmre1A</em></td>
</tr>
<tr>
<td>Hydroxyurea (µM)</td>
<td>53.37 ± 3.68</td>
<td>40.62 ± 2.26 (1.3)</td>
</tr>
<tr>
<td>Phleomycin (µM)</td>
<td>0.12 ± 0.01</td>
<td>0.12 ± 0.00 (1.0)</td>
</tr>
<tr>
<td>MMS (µM)</td>
<td>33.02 ± 0.28</td>
<td>35.82 ± 0.29 (0.9)</td>
</tr>
<tr>
<td>UV (kJ/cm)</td>
<td>2.58 ± 0.39</td>
<td>2.66 ± 0.18 (1.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Wild type</th>
<th><em>T. brucei</em> EC(_{50})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Tbsnm1A</em></td>
<td><em>Tbexo1A</em></td>
</tr>
<tr>
<td>Hydroxyurea (µM)</td>
<td>47.56 ± 1.90</td>
<td>38.62 ± 2.71 (1.2)</td>
</tr>
<tr>
<td>Phleomycin (µM)</td>
<td>0.12 ± 0.00</td>
<td>0.12 ± 0.00 (1.0)</td>
</tr>
<tr>
<td>MMS (µM)</td>
<td>27.89 ± 0.63</td>
<td>42.28 ± 2.89 (0.7)</td>
</tr>
<tr>
<td>UV (kJ/cm)</td>
<td>3.54 ± 0.20</td>
<td>3.90 ± 0.32 (0.9)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Wild type</th>
<th><em>T. brucei</em> EC(_{50})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Tbsnm1A</em></td>
<td><em>Tbcsb1A</em></td>
</tr>
<tr>
<td>Hydroxyurea (µM)</td>
<td>71.88 ± 0.89</td>
<td>67.91 ± 1.80 (1.1)</td>
</tr>
<tr>
<td>Phleomycin (µM)</td>
<td>0.04 ± 0.00</td>
<td>0.03 ± 0.00 (1.3)</td>
</tr>
<tr>
<td>MMS (µM)</td>
<td>27.78 ± 5.17</td>
<td>30.94 ± 13.45 (0.9)</td>
</tr>
<tr>
<td>UV (kJ/cm)</td>
<td>2.12 ± 0.18</td>
<td>2.47 ± 1.49 (0.9)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Wild type</th>
<th><em>T. brucei</em> EC(_{50})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Tbsnm1A</em></td>
<td><em>Tbrev2A</em></td>
</tr>
<tr>
<td>Hydroxyurea (µM)</td>
<td>67.39 ± 0.65</td>
<td>50.21 ± 1.72 (1.3)</td>
</tr>
<tr>
<td>Phleomycin (µM)</td>
<td>0.062 ± 0.00</td>
<td>0.046 ± 0.00 (1.4)</td>
</tr>
<tr>
<td>MMS (µM)</td>
<td>41.41 ± 3.56</td>
<td>45.76 ± 0.65 (0.9)</td>
</tr>
<tr>
<td>UV (kJ/cm)</td>
<td>3.61 ± 0.29</td>
<td>3.17 ± 0.34 (1.1)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Wild type</th>
<th><em>T. brucei</em> EC(_{50})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Tbsnm1A</em></td>
<td><em>Tbrev3A</em></td>
</tr>
<tr>
<td>Hydroxyurea (µM)</td>
<td>67.39 ± 0.65</td>
<td>50.21 ± 1.72 (1.3)</td>
</tr>
<tr>
<td>Phleomycin (µM)</td>
<td>0.062 ± 0.00</td>
<td>0.046 ± 0.00 (1.4)</td>
</tr>
<tr>
<td>MMS (µM)</td>
<td>41.41 ± 3.56</td>
<td>45.76 ± 0.65 (0.9)</td>
</tr>
<tr>
<td>UV (kJ/cm)</td>
<td>3.46 ± 0.45</td>
<td>3.92 ± 1.56 (0.9)</td>
</tr>
</tbody>
</table>
The susceptibility of *T. brucei* various single and double null mutant lines against hydroxyurea (HU), phleomycin (Phleo), methyl methanesulfonate (MMS) and ultraviolet radiation (UV), as judged by their EC$_{50}$ values, was compared and expressed as a fold difference relative to wild type.

### 4.1.2.3 Susceptibility of *T. brucei* double null lines towards mechlorethamine

As with the single null mutant lines, a series of phenotypic screens were performed using mechlorethamine as the selective agent. The resultant cell density data was analysed, plotted as dose response curves from which EC$_{50}$ values were determined (Figure 4.5; Table 4.2). Again, the fold differences for each single and double null mutant line against the ICL inducing agent was displayed as a bar chart (Figure 4.6).

The growth profiles and EC$_{50}$ values show a number of interesting features relating to how components of the *T. brucei* ICL repair interact. Firstly, cells lacking TbSNM1 and TbMRE11 display a hypersensitivity phenotype towards mechloethamine as compared to wild type cells and Tb$snm1$ or Tb$mre11$ single null mutant lines: Tb$snm1\Delta$ Tb$mr11\Delta$ parasites are approximately 40-, 6- and 15-fold more susceptible to this agent relative to *T. brucei* wild type, Tb$snm1\Delta$ and Tb$mr11\Delta$, respectively. This increase in sensitivity shown by the double null trypanosomes is indicative that the DNA repair enzymes TbSNM1 and TbMRE11 function in a non-epistatic fashion and do not operate in the same ICL repair system. Secondly, cells lacking TbSNM1 and TbCSB exhibit a mechlorethamine sensitivity similar to Tb$snm1\Delta$ parasites with both lines being more susceptible to this compound than wild type and Tb$csb\Delta$ cells. As the double null cells do not show an additional increase in sensitivity to the ICL-
inducing agent, it is plausible that TbSNM1 and TbCSB are epistatic and as such do operate in the same ICL repair system. This epistatic interaction appears to extend across to TbEXO1 with the Tbsnm1Δ Tbexo1Δ double null cells showing a similar trait as the Tbsnm1Δ TbcshΔ line. Finally, TbREV2 and TbREV3 do not play a discernible role in trypanosomal ICL repair, even in the absence of TbSNM1, confirming the observations reported in Chapter 3.

Control growth inhibition assays using DFMO revealed that the above susceptibility profiles displayed by each double mutant line were not due to off target effects as all lines showed similar sensitivities to this trypanocidal agent.
Figure 4.5: Susceptibility of *T. brucei* null mutants towards mechlorethamine and DFMO.

Dose response curves of *T. brucei* wild type (●), Tbsnm1Δ (■), single null mutant (♦) and double null mutant line (○) towards mechlorethamine (A) and DFMO (B). All data points are mean values ± standard deviations from experiments performed in quadruplicate.
### Table 4.2: Susceptibility of double null mutants towards mechlorethamine and DFMO.

EC$_{50}$ values of *T. brucei* wild type and null lines against mechlorethamine and DFMO (in $\mu$M). All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC$_{50}$ values between wild type and null mutant lines is given as fold difference in parentheses. *Indicates significant differences in susceptibility (P < 0.0001) between wild type and null lines, as assessed by Student’s $t$ test (GraphPad Software).

<table>
<thead>
<tr>
<th>Treatment</th>
<th><em>T. brucei</em> EC$_{50}$</th>
<th>Wild type</th>
<th>Tbsnm1Δ</th>
<th>Tbmre11Δ</th>
<th>Tbsnm1ΔTbmre11Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>mechlorethamine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFMO</td>
<td>34.57 ± 0.87</td>
<td>34.14 ± 0.37 (1.0)</td>
<td>42.95 ± 1.50 (0.8)</td>
<td>37.18 ± 3.38 (0.9)</td>
<td></td>
</tr>
<tr>
<td>mechlorethamine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFMO</td>
<td>57.14 ± 0.48</td>
<td>42.65 ± 2.72 (1.3)</td>
<td>57.40 ± 5.44 (1.0)</td>
<td>40.60 ± 1.96 (1.4)</td>
<td></td>
</tr>
<tr>
<td>mechlorethamine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFMO</td>
<td>48.77 ± 2.19</td>
<td>38.61 ± 2.18 (1.3)</td>
<td>46.41 ± 0.32 (1.1)</td>
<td>44.30 ± 2.04 (1.1)</td>
<td></td>
</tr>
<tr>
<td>mechlorethamine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFMO</td>
<td>36.75 ± 0.57</td>
<td>33.81 ± 2.20 (1.0)</td>
<td>42.97 ± 0.64 (0.9)</td>
<td>29.31 ± 0.40 (0.9)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.6: Comparison of *T. brucei* double null mutant EC\(_{50}\) values.
The susceptibility of *T. brucei* various single and double null mutant lines against mechlorethamine, as judged by their EC\(_{50}\) values, was compared and expressed as a fold difference relative to wild type.

### 4.2 Confirming the interplay of HR, NER, and MMR pathways in ICL repair

Based on our initial findings, we hypothesised that *T. brucei* expresses two distinct ICL repair systems, one involving the activities of TbSNM1/TbCSB/TbEXO1 with the other centred on TbMRE11. To explore this further a series of additional double null lines were generated to investigate the interplay between TbMRE11, TbCSB and TbEXO1.

Using the previously described gene interruption plasmids in conjunction with nucleofection, *T. brucei* lines lacking TbCSB and TbMRE11 (Tbcsb\(\Delta\) Tbmre11\(\Delta\)), TbEXO1 and TbMRE11 (Tbexo1\(\Delta\) Tbmre11\(\Delta\)) or TbCSB and TbEXO1 (Tbcsb\(\Delta\) Tbexo1\(\Delta\)) activities were made and subsequently validated (see Appendices 3e-3g). The susceptibility of these lines to mechloretamine was then tested with dose response curves drawn and EC\(_{50}\) values determined and compared (Figure 4.7 & Table 4.3). Fold differences are shown in bar chart form in Figure 4.8.
Figure 4.7: Susceptibility of *T. brucei* double null mutants towards mechlorethamine and DFMO.

A. Dose response curves of *T. brucei* wild type (●), TbcspΔ (■), Tbexo1Δ (■) and TbcspΔ Tbexo1Δ (○) towards mechlorethamine (A) and DFMO.

B. Dose response curves of *T. brucei* wild type (●), TbcspΔ (■), Tbmr11Δ (▲) and TbcspΔ Tbmr11Δ (○) towards mechlorethamine (A) and DFMO.

C. Dose response curves of *T. brucei* wild type (●), Tbexo1Δ (■), Tbmr11Δ (○) and Tbexo1Δ Tbmr11Δ (□) towards mechlorethamine (A) and DFMO.

All data points are mean values ± standard deviations from experiments performed in quadruplicate.
Table 4.3: Susceptibility of null mutants towards mechlorethamine and DFMO.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Wild type</th>
<th>TbsbΔ</th>
<th>Tbexo1Δ</th>
<th>TbsbΔTbexo1Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mechlorethamine</td>
<td>204.25 ± 38.44</td>
<td>62.19 ± 3.30</td>
<td>159.35 ± 3.81 (1.6)*</td>
<td>44.48 ± 5.86 (4.5)*</td>
</tr>
<tr>
<td>DFMO</td>
<td>41.85 ± 0.68</td>
<td>48.14 ± 1.00 (1.1)</td>
<td>46.41 ± 0.32 (1.1)</td>
<td>45.70 ± 4.12 (1.1)</td>
</tr>
<tr>
<td>Mechlorethamine</td>
<td>140.05 ± 5.10</td>
<td>37.08 ± 3.03 (3.8)*</td>
<td>41.40 ± 3.62 (3.4)*</td>
<td>10.30 ± 1.90 (13.6)*</td>
</tr>
<tr>
<td>DFMO</td>
<td>41.85 ± 0.68</td>
<td>48.14 ± 1.00 (1.1)</td>
<td>42.95 ± 1.50 (1.0)</td>
<td>45.88 ± 2.16 (1.1)</td>
</tr>
<tr>
<td>Mechlorethamine</td>
<td>223.43 ± 2.83</td>
<td>81.27 ± 1.83</td>
<td>60.57 ± 3.60 (3.7)*</td>
<td>50.31 ± 5.91 (4.4)*</td>
</tr>
<tr>
<td>DFMO</td>
<td>48.77 ± 2.16</td>
<td>46.41 ± 0.32 (0.9)</td>
<td>40.31 ± 1.74 (0.8)</td>
<td>34.92 ± 1.90 (0.7)</td>
</tr>
</tbody>
</table>

Table 4.3: Susceptibility of null mutants towards mechlorethamine and DFMO.

EC$_{50}$ values of *T. brucei* wild type and null lines against mechlorethamine and DFMO. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC$_{50}$ values between wild type and null mutant lines is given as fold difference in parentheses. *Indicates significant differences in susceptibility (P < 0.0001) between wild type and null lines, as assessed by Student’s $t$ test (GraphPad Software).

Figure 4.8: Comparison of *T. brucei* double null mutant EC$_{50}$ values.

The susceptibility of *T. brucei* various single and double null mutant lines against mechlorethamine, as judged by their EC$_{50}$ values, was compared and expressed as a fold difference relative to wild type.

As before, observations relating to the mechlorethamine growth inhibition data confirms our initial hypothesis. Cells lacking TbCSB and TbEXO1 reveal that these two DNA repair
enzymes, one involved in TC-NER and the other in MMR, function epistatically in ICL repair: The TbcshΔ Tbxo1Δ parasites exhibit a mechloretamine sensitivity similar to TbcshΔ trypanosomes with both the single and double mutants being more susceptible to the ICL inducing compound than wild type. In contrast and analogous to observations made using the Tbsnm1Δ Tbmre11Δ line, T. brucei TbcshΔ Tbmre11Δ exhibit hypersensitivity towards mechloretamine relative to wild type cells and either of the single null mutant lines. This increase in sensitivity is indicative that TbcSB and TbmRE11 function in a non-epistatic fashion and do not operate in the same ICL repair system. Finally, cells lacking TbxO1 and TbmRE11 exhibit an epistatic relationship with the fold difference in mechloretamine susceptibility between the Tbxo1Δ Tbmre11Δ and Tbmre11Δ lines being approximately the same (4-fold). This indicates that TbxO1 can also operate in the MRE11-dependent pathway.

All of the double null mutant data involving Tbsnm1, TbcSB, TbxO1 and TbmRE11 clearly shows that T. brucei expresses at least two distinct ICL repair systems with one involving the concerted action of Tbsnm1, TbcSB and TbxO1 with the other using TbmRE11 and TbxO1. It is tempting to speculate that given TbcSB plays a role in TC-NER, that the former SNM1-dependent pathway functions to resolve ICLs encountered during DNA transcription while the MRE11-dependent system helps remove ICLs encountered during DNA replication.
4.3 Chapter Summary
This chapter has focused on understanding the relationship between various DNA repair proteins that function in different DNA repair pathways to help resolve ICLs. Here, we have shown that:

1. *T. brucei* possesses at least two distinct ICL repair systems
   a. An SNM1-dependent ICL repair system where TbSNM1 co-operates with TC-NER (TbCSB) and MMR (TbEXO1) enzymes to repair the encountered damaged, possibly triggered as a result of RNA polymerase stalling
   b. An SNM1-independent ICL repair systems which utilises components the HR (TbMRE11) and MMR (TbEXO1) pathways to repair the encountered damaged, possibly triggered as a result of DNA polymerase stalling
2. The TbREV2 DNA helicase and TbREV3 DNA polymerase appear to play no role in ICL repair
5. **The trypanosomal Fanconi Anaemia System**

The Fanconi Anaemia (FA) repair pathway is the major ICL repair system in mammalian cells. This mechanism, first discovered through genetic profiling of patients with the disorder Fanconi Anaemia, involves at least 30 proteins which can be divided into four groupings comprising of 1. The FA core complex, 2. Recruitment factors, 3. Effector proteins and 4. FA-associated proteins (Figure 5.1). It is now recognised that other eukaryotes including fish, plants, insects and fungi all have their own version of the FA pathway, with such organisms often expressing a subset of homologues related to the human FA proteins (Moldovan and D’Andrea, 2009; Jones and Huang, 2012; McHugh, Ward and Chovanec, 2012; Zhang and Walter, 2014).

This chapter focuses on the identification of FA factors present in *T. brucei* and further unravelling of their function in relation to ICL repair. To determine whether the parasite possesses components of this system, human, zebrafish (*D. rerio*), fruit fly (*D. melanogaster*), plant (*A. thaliana*) and yeast (*S. cerevisiae*) FA protein sequences (complete or a specific domain) were used to search for trypanosome homologues held on the TriTryp and NCBI databases using a reciprocal Blast analysis. This revealed that *T. brucei* has the potential to express a minimal FA repair pathway, analogous to the situation noted for yeast (Ward *et al.*, 2012). Out of the 31 sequences examined, *T. brucei* has discernible homologues for 9 activities (Figure 5.1). Most of these are direct homologues of FA effector proteins that play or are postulated to function in the parasite’s HR pathway (FANCD1 (known as BRCA2; (Hartley and McCulloch, 2008)), FANCJ (referred to as CHL1), FANCO (or RAD51C; (Dobson *et al.*, 2011)), and FANCR (commonly known as RAD51; (Proudfoot and McCulloch, 2005))), NER pathway (FANCQ (or XPF; (Machado *et al.*, 2014))) or as an essential spindle checkpoint protein (FANCV (known as MAD2)). Of the remaining trypanosomal sequences, only one showed homology to a component of the FA core complex (FANCM (also known as MPH1)) and two related to FA-associated proteins (FAN1 and USP1). Interestingly, *T. brucei* appears to have no detectable versions of FANCD2 or FANCI, the two FA recruitment factors proteins.
Figure 5.1: Identification of putative components of the trypanosomal FA pathway. Comparison of the FA pathway components expressed by *Danio rerio* (*D. rerio*; fish), *Drosophila melanogaster* (*D. melanogaster*; insect), *Arabidopsis thaliana* (*A. thaliana*; plant), *Saccharomyces cerevisiae* (*S. cerevisiae*; yeast) and *T. brucei* in relation to the system possessed by humans (*H. sapiens*). Trypanosomal FA factors were identified using a reciprocal Blast analysis against human FA factors. A coloured circle indicates the possibility for the presence of a given factor whereas no circle indicates that that component is most probably absent. In mammalian cells, the FA ICL repair pathway can be subdivided the FA core complex (red), recruitment factors (blue), FA effectors (further split as to whether it involves homologous recombination (HR) (purple) or other pathways (green)) and FA-associated proteins (orange). Figure layout derived from (Zhang *et al.*, 2009; Dong *et al.*, 2015).

5.1 Analysis of the *T. brucei* FANCM homologue

Analysis of the *T. brucei* genome database indicated that this parasite has the potential to express only one component of the FA core complex, namely TbFANCM (Tb927.7.2970; referred from now on as TbHEL1 due to presence of a helicase domain in the protein sequence). The deduced trypanosomal amino acid sequence showed identity to the amino termini of human FANCM (23 %) and DICER (21 %) with this homology centring on the DEAD (PF00270) and Helicase C (PF00271) domains, regions found in members of the helicase superfamily 2 (SF2) RigI group (Figure 5.2). Intriguingly, the equivalent HEL1 sequence was found in other *Trypanosoma* species that are predominantly found in the bloodstream of their mammalian host (*e.g.* *T. congolense*, *T. evansi*, *T. vivax*), some other trypanosomatids (*e.g.* *Crithidia fasciculate*, *Leptomonas pyrrhocoris*) but absent from intracellular *Trypanosoma* and *Leishmania* species (*e.g.* *T. cruzi*, *L. major*, *L. donovani*). Phylogenetic analysis of the trypanosomatid *hel1* genes with their plant, fungus and vertebrate *fancm* counterparts using the Phylogeny.fr algorithm revealed that the parasite sequences are evolutionarily distinct from
homologues expressed in the other organisms, forming their own distinct clade (Figure 5.3). In relation to higher eukaryotic organisms, this difference was unsurprising given that the encoded proteins generally possess additional domains in their backbone (Figure 5.2). For the human FANCM, these extra regions facilitate 1. binding of the helicase to double stranded DNA via a FANCM-MHF (PF16783) domain, 2. function to activate downstream cell cycle kinases via a mediator of replication checkpoint 1 (MRC1; PF09444) domain and 3. cleave the target substrate through a ERCC4 nuclease (PF02732) domain. Given that equivalent regions are apparently absent from TbHEL1 then the trypanosomal enzyme, if it does function as a DNA helicase, does so by a mechanism distinct from its mammalian counterpart. The yeast FANCM homologue which has been shown to function in ICL repair (Ward et al., 2012) also lacks many of these additional domains but in our phylogenetic analysis the fungal sequences tend to be more closely rooted to the helicases expressed by vertebrates.

Figure 5.2: Domain and sequence analysis of TbHEL1.
Schematic diagram showing the organization of conserved domains found in H. sapiens FANCM (HsFANCM; BC140776), S. cerevisiae MPH1 (ScMPH1; KI893982) and TbHEL1 (Tb927.7.2970). The regions highlighted include the DEAD (green; PF00270), Helicase C (red; PF00271), FANCM-MHF (purple; PF16783), MRC1 (light blue; PF09444) and ERCC4 (yellow; PF02732) domains. The DEAD (top alignment) and Helicase C (bottom alignment) domains of TbHEL1 was compared with the equivalent regions from T. evansi (TevSTIB805.7.3080), S. cerevisiae MPH1, Arabidopsis thaliana FANCM (NM_001198212), Drosophila melanogaster FANCM (EF088195), H. sapiens FANCM and Xenopus laevis (XM_012969053). Conserved residues (in black or grey) are highlighted with arrows marking the conserved amino acids that form the HMM logo.
Figure 5.3: Phylogenetic tree analysis of FANCM sequences. Phylogenetic analysis of the trypanosomatid hel1 genes (green) forms a different clade to plant (yellow), fungal (pink), and vertebrate genes (purple) homologues. The sequences analysed using the Phylogeny.fr tool were *T. brucei* hel1 (Tbhel1, Tb927.7.2970), *T.b.gambiense* hel1 (Tbg972.7.3290), *T. congoles* hel1 (TcIL3000_0_22580), *T. evansi* hel1 (TevSTIB805.7.3080), *T. vivax* hel1 (TvY486_0702830), *T. grayi* hel1 (DQ04_01821030), *C. fasiculata* hel1 (CFAC1_240016600), *L. pyrrhocoris* hel1 (LpyrH10_01_0930), *B. saltans* hel (CYKH01000268), *H. sapiens* fancm (BC140776), *P. troglodytes* fancm (XM_003314302), *M. mulatta* fancm (XM_015147628), *Drosophilia melanogaster* fancm (EF088195), *S. cerevisiae* mph1 (KI893982), *Xenopus fancm* (XM_012969053), *Arabidopsis thaliana* fancm (NM_001198212), *A. fumigatus* mph1 (XM_744962), *B. rapa* mph1 (XM_013864549), *N. crassa* mph1 (XM_958929), *O.s. Japonica*.fml1 (XM_015760033), and *N. tomentosiformis* fancm (XM_018776163).

5.1.1 Construction and validation of null mutant lines

Using the previously described gene deletion approach from Section 3, DNA constructs based around the hygromycin and neomycin resistance cassettes (pKO-Tbhel1-hyg and pKO-Tbhel1-neo DNA) were generated, the DNAs processed by restriction enzyme digestion and fragments purified. These were sequentially transformed into *T. brucei* wild type and *Tbsnm1*-deficient cells, creating the Tbhel1Δ and Tbsnm1Δ Tbhel1Δ lines. The gene interruption events and resultant loss of gene expression was validated using PCR-based approaches described earlier, with this data presented in Appendix 4a.

5.1.2 Phenotypic analysis of *T. brucei* lines lacking *hel\*\*

Following confirmation of the *T. brucei* single and double null Tb*hel1* null mutant lines, their growth properties were evaluated in comparison to wild type cells. This demonstrates that the recombinant cell lines grew at an equivalent rate as the controls (data not shown). The susceptibility of Tb*hel1*-deficient cells to various DNA damaging agents was then tested to evaluate whether the encoded enzyme does function in DNA repair. For each treatment, dose
response curves were generated (Figure 5.4), EC$_{50}$ values calculated (Table 5.1) and the fold difference of EC$_{50}$ values between wild type and null lines determined (in parentheses of Table 5.1). Together, this data strongly indicates that TbHEL1 has no role in any DNA repair pathway tested here, including any mechanism used to resolve ICLs.

Figure 5.4: Susceptibility of *T. brucei* Tbhel1-deficient null lines towards various DNA damaging treatments.
Dose response curves of *T. brucei* wild type (●), Tbsnm1Δ (■), Tbhel1Δ (♦), and Tbsnm1Δ Tbhel1Δ (○) cell lines to hydroxyurea, UV, phleomycin, MMS, mechlorethamine and DFMO. All data points are mean values ± standard deviations from experiments performed in quadruplicate.
<table>
<thead>
<tr>
<th>Treatment</th>
<th>Wild type</th>
<th>Tbsnm1Δ</th>
<th>Tbhel1Δ</th>
<th>Tbsnm1ΔTbhel1Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>hydroxyurea</td>
<td>57.76 ± 2.20</td>
<td>48.97 ± 2.03 (1.2)</td>
<td>42.46 ± 0.34 (1.4)</td>
<td>41.12 ± 0.98 (1.4)</td>
</tr>
<tr>
<td>phleomycin</td>
<td>0.12 ± 0.00</td>
<td>0.08 ± 0.01 (1.5)</td>
<td>0.10 ± 0.00 (1.2)</td>
<td>0.09 ± 0.01 (1.3)</td>
</tr>
<tr>
<td>MMS</td>
<td>32.70 ± 0.12</td>
<td>36.35 ± 0.62 (0.9)</td>
<td>38.88 ± 0.53 (0.8)</td>
<td>29.90 ± 1.22 (1.1)</td>
</tr>
<tr>
<td>UV (kJ/cm)</td>
<td>0.52 ± 0.15</td>
<td>0.84 ± 0.11 (0.6)</td>
<td>0.76 ± 0.17 (0.7)</td>
<td>0.57 ± 0.03 (0.9)</td>
</tr>
<tr>
<td>mechlorethamine</td>
<td>56.85 ± 2.26</td>
<td>5.55 ± 0.07 (10.2)*</td>
<td>55.37 ± 2.08 (1.0)</td>
<td>3.99 ± 0.15 (14.3)*</td>
</tr>
<tr>
<td>DFMO</td>
<td>25.74 ± 1.90</td>
<td>34.14 ± 0.37 (0.8)</td>
<td>33.38 ± 0.15 (0.8)</td>
<td>32.73 ± 0.48 (0.8)</td>
</tr>
</tbody>
</table>

Table 5.1: Susceptibility of *T. brucei* Tbhel1 null lines to DNA damaging treatments.

Data represents EC₅₀ values of *T. brucei* wild type, Tbsnm1Δ, Tbhel1Δ and Tbsnm1ΔTbhel1Δ against hydroxyurea, phleomycin, MMS, UV, mechlorethamine and DFMO. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC₅₀ values between wild type and null mutant lines is given as a fold difference, in parentheses. *Indicates significant differences in susceptibility (P < 0.0001 between wild type and null lines, as assessed by Student’s t test (GraphPad Software).

### 5.1.3 TbHEL1 and its role in RNA interference

Given that TbHEL1 appears to play no role in DNA repair and also has significant identity to the helicase domain of HsDICER, we next evaluated whether it could function in the trypanosomal RNA interference pathway. As with the FANCM alignment, comparison of the trypanosomatid HEL1s with DICERs and DRHs (DICER related helicases) from various other organisms showed that sequence homology was most pronounced in the DEAD (PF00270) and Helicase C (PF00271) domains (Figure 5.5). Again, phylogenetic analysis revealed that the protozoan parasites gene sequences were evolutionarily distinct from their counterparts possessed by other organisms, reflecting the early branching nature of trypanosomatid on the tree of life and potentially indicating that these pathogens have an unconventional DICER activity (Figure 5.6).
Figure 5.5: Domain and sequence analysis of TbHEL1.
Comparisons between the domain structure of *T. brucei* HEL1 (TbHEL1; Tb927.7.2970), *T. b. gambiense* HEL1 (TbgHEL1; Tbg972.7.3290), *C. fasiculata* HEL1 (CfHEL1; CFAC1_240016600), *H. sapiens* DICER (HsDICER; NM_0012916281), *D. rerio* DICER (DrDICER; NM_001161453), *C. elegans* DCR1 (CeDCR1; NM_068617), *A. thaliana* DICER (AtDICER; JN661702), *H. sapiens* DDX58 (HsDDX58; NM_014314), and *C. elegans* DRH1 (CeDRH1; NM_068617) with further comparisons made on conserved residues (emphasised using arrows). Domains shown include DEAD domain (green; PF00270), Helicase C domain (red; PF00271), RIG-1 domain (light purple; PF11648), Dicer dimerization domain (dark green; PF03368), PAZ domain (dark purple; PF02170), ribonuclease domain (blue; PF00636), CARD_2 (grey; PF16739), Sec34 (burgundy; PF04136), dsrm (yellow; PF00035), and DND1_DSRM (orange; PF14709).
Figure 5.6: Phylogenetic tree analysis of TbHEL1 and DICER sequences.

The phylogenetic tree outlines the trypanosomatid HEL genes (green) in a different clade to that of the plant (yellow) and vertebrate genes (purple) with plant FANCM genes outlined as evolutionarily distinct from that of vertebrate. Gene names and sequence identities are given on the end of each clade: Lpyr (L. pyrrhocoris; LpyrH10_01_0930), Cfas (C. fasciculate; CFAC1_240016600), Tgra (T. grayi; DQ04_01821030), Tviv (T. vivax; TvY486_0702830), Tcong (T. congolense; TcL3000_0_22580), Tbrucei (T. brucei; Tb927.7.2970), Teva (T. evansi; TevSTIB805.7.3080), Tbgam (T. brucei gambiense; Tbg972.7.3290), Bsal (B. saltans; CYKH01000268), Cedrh1 (C. elegans drh1; NM_068617), BmalDRH1 (B. malayi DRH1; LN855620), BmalDRH3 (B. malayi DRH3; LN856998), DrerDDX58 (D. rerio DDX58; NM_001306095), XlaeDDX58 (X. laevis DDX58; XM_018257225), GjaDDX58 (G. japonicus DDX58; XM_015414538), Ptro (P. troglodytes DDX58; XM_001156662), HsapDDX58 (H. sapiens; NM_014314), Mmul dicer (M. mulatta dicer; NM_001257872), Ptro dicer (P. troglodytes dicer; XM_001154369), Hsapdicer (H. sapiens dicer; NM_0012916281), CcanDICER (C. canadensis DICER1; GFFV01000155), Pvm dicer (P. vampyrus dicer; XM_011379597), Gjap dicer (G. japonicus dicer; XM_015416341), XlaeDICER (X. laevis DICER; XM_018230462), Dreddicer (D. rerio dicer; NM_001161453), Bmal (B. malayi DCR1; LN857010), Cedcr1 (C. elegans dcr-1; NM_066360), AfumDICER (A. fumigatus; XM_748378), CimmCAF (C. immitis CAF; DS017001), Spom dicer (S. pombe dicer; NM_001023205), SjapCAF (S. japonicus CAF; XM_002174899), BnapDICER (B. napus DICER; XM_013841052), AthaDicer (A. thaliana Dicer; JN661702), OsDICER (O. sativa DICER; CM000140) and AoffDICER (A. officinalis DICER; XM_020390741).

To see whether TbHEL1 functions in the T. brucei RNAi pathway, gene integration vectors based around puromycin and blasticidin resistance cassettes (pKO-TbHEL1-bla and pKO-TbHEL1-pac) were constructed, processed by restriction enzyme digestion and the resultant
fragments purified. These were then sequentially transformed into *T. brucei* lines that could be induced to undergo RNAi against *Tbtao* or *TbsodB*. These two lines were used because 1. the targeted activities, namely trypanosomal alternative oxidase or superoxide dismutase isoform b, respectively, are essential for the growth of bloodstream form trypanosomes (Helfert *et al.*, 2001; Wilkinson *et al.*, 2006) and 2. the dsRNA that promotes down regulation of the *Tbtao* or *TbsodB* transcript are formed by a stem loop or linear RNA molecule, respectively. The resultant null parasites, designated as RNAi *Tbtao* TbhellΔ and RNAi *TbsodB* TbhellΔ, were then subject to PCR validation that confirmed that Tbhell had been successfully interrupted in both lines with an accompanying loss of Tbhell expression (Figure 5.7).

**Figure 5.7: Confirming the deletion of Tbhell.**
A. Schematic for the representation of the Tbhell allele and effects of disrupting this allele with genes encoding for blasticidin (bla) and puromycin (pac) resistance. P1 (Tbhel1-KO1), P2 (Tbhel1-q2), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.

B. DNA fragments obtained after amplification on gDNA extracted from *T. brucei* wild type (lane 1), *T. brucei* RNAi-*Tbtao* (lane 2), *T. brucei* RNAi-*Tbsodb* (lane 3), RNAi-*Tbtao* TbhellΔ (lane 4), and RNAi-*Tbsodb* TbhellΔ (lane 5). Extractions from all lines were tested using primer combinations that specifically amplify biomarkers corresponding to the Tbhell (Tbhel1-KO1/Tbhel1-q2), Tbhell-pac (Tbhel1-KO1/pac-1) or Tbhell-bla (Tbhel1-KO1/bla-1) alleles.

C. DNA fragments obtained after amplification on cDNA derived from RNA extracted from *T. brucei* wild type (lane 1), TbhellΔ (lane 2), RNAi-*Tbtao* (lane 3), RNAi-*TbxobB* (lane 4), RNAi *Tbtao* TbhellΔ (lane 5), and RNAi *Tbsodb* TbhellΔ (lane 6) using primer combinations that specifically amplify intact Tbhell (Tbhel1-q1/Tbhel1-q2) and Tbtert (Tbtert-F/Tbtert-R).

Following validation, the growth properties of the various Tbhell null mutants that can undergo RNAi was determined in the presence/absence of the inducer (tetracycline) which promotes down regulation of the targeted transcript. Upon induction of RNAi in RNAi-*Tbtao*
and RNAi-Tb\textit{sobB} cells (i.e. that contain a functional \textit{Tb}\textit{hel1}), parasite growth was significantly reduced in the former and halted in the latter in agreement with previous reports (Helfert \textit{et al.}, 2001; Wilkinson \textit{et al.}, 2006). In contrast, \textit{Tb}\textit{hel1}-deficient cells undergoing RNAi targeting these two transcripts continued to grow at an equivalent rate to non-induced controls (Part A and B of Figure 5.8). This suggests that \textit{TbHEL1} plays a crucial role in the \textit{T. brucei} RNAi machinery.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.8.png}
\caption{Analysing the role of \textit{TbHEL1} in RNA interference.}
\end{figure}

\textbf{A.} Cumulative growth of non-induced RNAi-\textit{Tbtao} (○), induced RNAi \textit{Tbtao} (●), non-induced RNAi \textit{Tbtao} \textit{Tb}\textit{hel1Δ} (●) and induced RNAi \textit{Tbtao} \textit{Tb}\textit{hel1Δ} (▲).

\textbf{B.} Cumulative growth density curves of non-induced RNAi \textit{TbsodB} (○), induced RNAi \textit{TbsodB} (■), non-induced RNAi \textit{TbsodB} \textit{Tb}\textit{hel1Δ} (●) and induced RNAi \textit{TbsodB} \textit{Tb}\textit{hel1Δ} (▲).

\textbf{C & D.} The mRNA levels of \textit{Tbtao} (C) or \textit{TbsodB} (D) in various induced (+tet) and non-induced parasite lines was evaluated by RT-qPCR and compared against the expression level of a standardized control (\textit{Tb}\textit{tert}). The average fold difference, as judged by \textit{2}−(ΔΔ\textit{C}_T) from reactions performed in triplicate standard deviation, plotted as a measure of the relative expression level.

To confirm that down-regulation had been blocked, RT-qPCR to detect the targeted mRNA (\textit{Tbtao}, \textit{TbsodB}) was performed. In tetracycline treated \textit{Tb}\textit{hel1}-proficient RNAi-\textit{Tbtao} and RNAi-\textit{TbsodB} cells, \textit{Tbtao} or \textit{TbsodB} expression was shown to be ~50-fold and ~3-fold lower than in the non-induced trypanosomes, respectively (Part C and D of Figure 5.8). When \textit{Tb}\textit{hel1}-deficient parasites undergoing RNAi were analysed, transcript levels were equivalent to that observed in wild type and non-induced controls. This confirms that \textit{TbHEL1} does play a key
role in the \textit{T. brucei} RNAi machinery and can interact with both linear and stem loop derived dsRNA templates. The precise interactions of this trypanosomal helicase with other RNAi components was not determined here although it is tempting to speculate that TbHEL1 does interact with TbDCL1 or TbDCL2, atypical DICER enzymes that usually only possess a nuclease domain (Patrick \textit{et al.}, 2009).

5.2 \textbf{Analysis of the} \textit{T. brucei} FANCJ homologue

FANCJ, also known as CHL1 or BRIP1, is one of the many FA effector proteins which are upregulated upon ubiquitination of the FANCD2/I complex to help unwind damaged DNA (Wu and Brosh, 2009; Daee and Myung, 2012; Brosh and Cantor, 2014; Wilson \textit{et al.}, 2017). Searches of the \textit{T. brucei} database identified a putative TbCHL1 homologue (Tb927.10.3550) that shares ~20\% identity to counterpart sequences expressed by other organisms with this homology centering on the DEAD (PF00270), DEAD_2 (PF06733) and Helicase C2 (PF00271) domains, regions found in some members of the helicases of the SF2 RAD3 group.

![Figure 5.9: Phylogenetic analysis of FANCJ sequences.](image)

A rooted phylogenetic tree of trypanosomatid CHL1 (purple), fungal CHL1 (pink) and vertebrate and plant FANCJ sequences, including humans and Arabopisis (yellow). Gene names and sequence identities are given on the end of each clade. \textit{T. brucei} chl1 (Tb927.10.3550), \textit{T.b.gambiense} chl1 (Tbg972.10.4420), \textit{T. cruzi} chl1 (TcCLB.510311.130), \textit{L. Mexicana} chl1 (LmjF.03.0590), \textit{L. panamensis} chl1 (LBRM2903_15001800), \textit{C. fasciculate} chl1 (CFAC1_060012500), \textit{L. pyrrhocoris} chl1 (LpyrH10_34_0770), \textit{H. sapiens} fancj (NM_032043), \textit{R. norvegicus} fancj (XM_340869), \textit{P. troglodytes} fancj (XM_511607), \textit{X. laevis} brip (XM_018244911), \textit{Drosophila} fancj (XM_009291678), \textit{Aradopsis} rad3-like (AT1G79950), \textit{B. napus} fancj-like (LK032238), \textit{O.s.Japonica} fancj (XM_015757026), \textit{S. cerevisiae} chl1 (Z71255), \textit{S. pombe} chl1 (NM_001020408), \textit{A. fumigatus} chl1 (JZDT01000837), \textit{N. crassa} chl1 (XM_001728469), and \textit{P. evpansum} chl1 (LHQQ01000074).
Figure 5.10: Domain and sequence comparisons of TbCHL1.
Schematic diagram showing the organization of conserved domains found *T. brucei* CHL1 (TbCHL1; Tb927.10.3550), *H. sapiens* FANCJ (HsFANCJ; NM_032043) and *S. cerevisiae* CHL1 (ScCHL1; Z71255). The regions highlighted include the DEAD (green; PF00270), DEAD_2 (blue; PF06733) and Helicase C2 domain (red; PF00271) domains. Alignments of the DEAD, DEAD_2 and Helicase C2 domains of the above sequence with *T. cruzi* CHL1 (TcCHL1; TCCLL.510311.130), *Arabidopsis thaliana* FANCJ (AtFANCJ; AT1G79950), *Xenopus laevis* BRIP (XlBRIP; XM_018244911), and *Danio rerio* FANCJ (DrFANCJ; XM_009291678). Conserved residues (in black or grey) are highlighted with arrows marking the conserved amino acids that form the HMM logo.
Phylogenetic analysis of trypanosomal, fungal, vertebrate and plant *fancj*-like sequences revealed that the trypanosomatid *chl1*’s are evolutionarily related to their fungal counterparts that together form a clade distinct from vertebrate and plant homologues (Figure 5.9).

### 5.2.1 Construction and validation of null mutant lines

Plasmids based around the hygromycin and puromycin resistance cassettes (pKO-*Tbchl1*-hyg and pKO-*Tbchl1*-pac DNA) were made and processed by restriction enzyme digestion. The resultant fragments were purified and sequentially transformed into *T. brucei* wild type cells, creating the *Tbchl1*Δ line. These recombinant parasites were then taken and *Tbsnm1* deleted using gene integration fragments containing neomycin and blasticidin resistance cassettes. Confirmation of the cell lines at a gDNA and mRNA level is given in Appendix 4b.

### 5.2.2 Phenotypic analysis of *T. brucei* lines lacking *chl1*

Following confirmation that *T. brucei* single and double null mutant lines lack *Tbchl1*, their growth properties were evaluated relative to wild type cells. This demonstrated that the recombinant cell lines grew at an equivalent rate as the controls (data not shown). The susceptibility of *Tbchl1*-deficient cells to various DNA damaging agents was then tested to evaluate whether the encoded enzyme does function in DNA repair. For each treatment dose response curves were generated (Figure 5.11), EC$_{50}$ values calculated (Table 5.2) and the fold difference of EC$_{50}$ values between wild type and null lines determined (Table 5.2).

Interestingly, no significant change in susceptibility to hydroxyurea, phleomycin, MMS, and UV was observed for *Tbchl1Δ* and *Tbsnm1Δ* *Tbchl1Δ* cells as compared against wild type *T. brucei* and *Tbsnm1Δ* (Figure 5.11). This indicates that TbCHL1 does not play a major role in the repair of base errors, double strand breaks and the formation of pyrimidine dimers. In contrast, *Tbchl1Δ* cells treated with mechlorethamine were 2-fold more susceptible to the ICL inducing compound indicating that this putative helicase plays a (minor) role in repair of such lesions. When these studies were extended to the *Tbsnm1Δ* *Tbchl1Δ* line, the level of sensitive to mechlorethamine of the double null was similar to that of *Tbchl1Δ* cells and not to the *Tbsnm1*-deficient line analysed in parallel. This phenotype is rather perplexing as it does not fit in with the “classical” epistatic/non-epistatic interactions observed with other components of the trypanosomal ICL REPAIRtoire. However, it may suggest that TbCHL1 and a
component in the TbSNM1-dependent pathway haves some sort of co-regulatory role, possibly suppressing the activity of other DNA repair factors. In the absence of both TbCHL1 and TbSNM1 these silenced/reduced activities are de-repressed and are able participate in ICL repair. The result observed using this Tbsnm1Δ Tbchl1Δ double null mutant line was observed using multiple clones generated from independent transformations.

Figure 5.11: Susceptibility of Tbcfl1 null lines towards DNA damaging treatments. Dose response curves of T. brucei wild type (●), Tbsnm1Δ (■), Tbcfl1Δ (♦), and Tbsnm1Δ Tbcfl1Δ (○) cell lines to hydroxyurea, UV, phleomycin, MMS, mechlorethamine and DFMO. All data points are mean values ± standard deviations from experiments performed in quadruplicate.
Table 5.2: Susceptibility of *T. brucei* Tbchl1 null lines to DNA damaging treatments.
Data represents EC$_{50}$ values of *T. brucei* wild type, Tbsnm1Δ, Tbchl1Δ and Tbchn1Δ Tbchl1Δ against hydroxyurea, phleomycin, MMS, UV, mechlorethamine and DFMO. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC$_{50}$ values between wild type and null mutant lines is given as fold difference, in parentheses. *Indicates significant differences in susceptibility (P < 0.0001 between wild type and null lines, as assessed by Student's t test (GraphPad Software)).

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Wild type</th>
<th>Tbchn1Δ</th>
<th>Tbchl1Δ</th>
<th>Tbchn1Δ Tbchl1Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydroxyurea</td>
<td>67.39 ± 0.65</td>
<td>50.21 ± 1.72 (1.3)</td>
<td>53.90 ± 0.85 (1.2)</td>
<td>66.67 ± 2.27 (1.0)</td>
</tr>
<tr>
<td>Phleomycin</td>
<td>0.04 ± 0.00</td>
<td>0.04 ± 0.00 (1.0)</td>
<td>0.03 ± 0.00 (1.3)</td>
<td>0.06 ± 0.00 (0.7)</td>
</tr>
<tr>
<td>MMS</td>
<td>32.70 ± 0.12</td>
<td>36.35 ± 0.62 (0.9)</td>
<td>41.20 ± 0.77 (1.2)</td>
<td>33.56 ± 1.37 (1.0)</td>
</tr>
<tr>
<td>UV (kJ/cm)</td>
<td>3.46 ± 0.46</td>
<td>4.58 ± 1.04 (0.8)</td>
<td>4.48 ± 2.18 (0.8)</td>
<td>2.98 ± 0.25 (1.2)</td>
</tr>
<tr>
<td>Mechlorethamine</td>
<td>70.94 ± 1.25</td>
<td>5.76 ± 0.41 (12.2)*</td>
<td>46.3 ± 0.75 (1.5)</td>
<td>37.92 ± 1.02 (1.8)</td>
</tr>
<tr>
<td>DFMO</td>
<td>41.32 ± 0.81</td>
<td>34.24 ± 0.97 (1.2)</td>
<td>41.30 ± 0.46</td>
<td>44.01 ± 0.76 (0.9)</td>
</tr>
</tbody>
</table>

5.3 *Analysis of the T. brucei* FAN1 homologue

Following ubiquitination of the FANCD2/FANCI complex, FAN1 (Fanconi-Associated Nuclease 1) is recruited to an ICL site where it functions as a nuclease to cleave the DNA backbone helping to excise the DNA lesion (Takahashi *et al.*, 2015; Jin and Cho, 2017). Searches of *T. brucei* databases identified a single sequence (TbFAN1; Tb927.10.7540) that shares 17 – 20 % identity with FAN1’s from a range of different organisms. These enzymes are unique in that they represent the only proteins that contain a VRR-NUC (PF08774) domain found towards the carboxyl terminal (Figure 5.12). This region is centred upon a conserved “PDX(D/E)XK” motif that coordinates binding of divalent metal ions (either manganese or magnesium) and is essential for their dual nuclease activities (Fontebasso *et al.*, 2013). Additionally, these enzymes often contain additional domains (*e.g.* TPR_12 (PF13424), HIRAN (PF08797), SAP (PF02037)) that mediate protein:protein interactions with TbFAN1 following this pattern, containing the so-called TPR_2 (Tetratricopeptide repeat 2; PF07719) motifs. As with TbHEL1, phylogenetic analysis of FAN1 sequences indicates that the trypanosomatid homologues forms an outlining group and again is evolutionary distinct from vertebrate, plant, nematode and fungal clades (Figure 5.13). This difference could reflect that for at least *T. brucei*, the factors that recruit FAN1 to the ICL lesion, namely the
FANCD2/FANCI complex, are absent and as such if it does play a role in ICL repair, how it is targeted to the site of DNA damage will be distinct from the reported systems.

Figure 5.12 Sequence domain analysis of FAN1 sequences.
Schematic diagram showing the domain structure of FAN1 sequences from *Trypanosoma brucei* (TbFAN1; accession number EAN78160), *Homo sapiens* (HsFAN1; NP_055782), *Arabidopsis thaliana* (AtFAN1; Q5XVJ4) and *Schizosaccharomyces pombe* (SpFAN1; Q9Y804). The regions corresponding to VRR_NUC (PF08774), SAP (PF02037), HIRAN (PF08797), TPR_2 (PF07719), TPR_12 (PF13424) and NatB_MDM20 (PF09797) domains are highlighted as red, purple, orange, dark blue, light blue and brown boxes, respectively. An alignment of the VRR_NUC domain of TbFAN1 with other FAN1 proteins is shown. Residues that are common with the TbFAN1 sequence are represented by dots with differences shown. Dashes have been inserted to optimise alignment.
Figure 5.13: Rooted phylogenetic tree analysis of FAN1 sequences.
A rooted phylogenetic tree outlines how trypanosomatid FAN1 (purple) sequences are evolutionarily dissimilar to plant/fungal FAN1 sequences (green), nematode FAN1 sequences (red) and vertebrate FAN1 sequences (yellow). Gene names and sequence identities are given on the end of each clade: T. brucei fan1 (Tb927.10.7540), T. cruzi fan1 (TeCLB.510105.90), L. major fan1 (LmjF.36.2620), C. fasciculate fan1 (CFAC1_250039400), L. braziliensis fan1 (LbrM.35.2830), L. pyrrhocoris fan1 (LpyrH10.06.0080), T. brucei gambiense fan1 (Tbg972.10.9240), H. sapiens fan1 (NM_014967), P. troglodytes fan1 (XM_510266), A. fumigatus fan1 (XM_743888), S. pombe fan1 (NM_001021302), B. napus fan1 (XM_013800152), O.s. Japonica fan1 (XM_015788183), A. thaliana fan1 (NM_103732), X. laevis fan1 (XM_018253000), D. rerio fan1 (NM_001045081), C. elegans fan1 (NM_068596), T. canis fan1 (JPKZ01003248), B. malayi fan1 (LN856244), P. vampyurus fan1 (XM_011379121), R. norvegicus fan1 (NM_001191633) and A. officinalis fan1 (XM_020397757).
5.3.1 TbFAN1 and ICL Repair

5.3.1.1 Creating and confirming single null mutants
As described elsewhere, Tbfan1 single null cells were created by the sequential transformation of processed pKO-Tbfan1-hyg and pKO-Tbfan1-neo DNA fragments into T. brucei wild type cells. Recombinant parasites were validated using PCR on gDNA or mRNA templates (Appendix 4c).

5.3.1.2 Phenotypic analysis of T. brucei lines lacking fan1
Following confirmation, the growth properties of the T. brucei Tbfan1Δ was assessed and compared against wild type cells cultivated in parallel. This demonstrate that the null line grew at an equivalent rate as the controls (data not shown). The susceptibility of Tbfan1-deficient cells to various DNA damaging agents was then tested to evaluate whether the encoded enzyme does function in DNA repair. For each treatment dose response curves were generated (Figure 5.14), EC50 values calculated (Table 5.3) and the fold difference of EC50 values between wild type and null lines determined (Table 5.3).

Figure 5.14: Susceptibility of T. brucei Tbfan1 null mutant towards various DNA damaging treatments.
Dose response curves of T. brucei wild type (solid line) and Tbfan1Δ (dotted line) to hydroxyurea, UV, phleomycin, MMS, mechlorethamine and DFMO. All data points are mean values ± standard deviations from experiments performed in quadruplicate.
Table 5.3: Susceptibility of *T. brucei* Tbfan1 null line to DNA damaging treatments.

Data represents EC50 values of *T. brucei* wild type and Tbfan1Δ against hydroxyurea, phleomycin, MMS, UV, mechlorethamine and DFMO. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC50 values between wild type and null mutant lines is given as fold difference, in parentheses. *Indicates significant differences in susceptibility (P < 0.0001 between wild type and null lines, as assessed by Student’s *t* test (GraphPad Software).

<table>
<thead>
<tr>
<th>Treatment</th>
<th><em>T. brucei</em> EC50</th>
<th>Tbfan1Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydroxyurea</td>
<td>75.15 ± 1.49</td>
<td>74.57 ± 1.97 (1.0)</td>
</tr>
<tr>
<td>Phleomycin</td>
<td>0.14 ± 0.01</td>
<td>0.11 ± 0.01 (1.3)</td>
</tr>
<tr>
<td>MMS</td>
<td>39.24 ± 1.64</td>
<td>34.68 ± 1.74 (1.1)</td>
</tr>
<tr>
<td>UV (kJ/cm)</td>
<td>0.79 ± 0.11</td>
<td>0.78 ± 0.08 (1.0)</td>
</tr>
<tr>
<td>Mechlorethamine</td>
<td>76.28 ± 1.29</td>
<td>37.92 ± 6.03 (2.0)*</td>
</tr>
<tr>
<td>DFMO</td>
<td>28.36 ± 0.54</td>
<td>35.64 ± 2.29 (0.8)</td>
</tr>
</tbody>
</table>

Analysis of the dose response curve data demonstrates that the Tbfan1Δ parasites exhibit a wild type-like response to hydroxyurea, phleomycin, MMS and UV indicating this putative nuclease plays no role in the repair of damage induced by these treatments. When the susceptibility of the null line to mechlorethamine was explored, recombinant cells were slightly (~2-fold) more sensitive to the ICL inducing agent than wild type indicating that this enzyme plays a role in the trypanosomal ICL REPAIRtoire.

### 5.3.2 Investigating the role of TbFAN1 in the trypanosomal ICL repair system

Our previous data indicates that *T. brucei* possess at least two ICL repair systems with one being dependent on TbSNM1 and the other on TbMRE11. To assess which, if any, of these TbFAN1 operates within, double null mutant lines lacking TbFAN1 and TbSNM1 (Tb*snm1*Δ Tb*fan1*Δ) or TbFAN1 and TbMRE11 (Tb*fan1*Δ Tbmre11Δ) were constructed and then validated as described (see Appendix 4d). Culturing of Tb*snm1*Δ Tb*fan1*Δ parasites revealed that they grew at the same rate at wild type and the single null mutant cells (Tb*fan1*Δ and Tbsnm1Δ) while Tb*fan1*Δ Tbmre11Δ trypanosomes exhibited the growth defect displayed by the Tbmre11Δ line (data not shown). Further phenotyping of these lines using mechlorethamine as selective agent generated dose response curves (Figure 5.15) and EC50 values (Table 5.4) indicting that TbSNM1 and TbFAN1 function non-epistatically while TbMRE11 and TbFAN1 exhibit an epistatic interaction. Cells lacking TbSNM1 and TbFAN1 are hypersensitive to mechlorethamine relative to the single null lines which in turn are more susceptible to the ICL.
inducing agent than wild type. In contrast, TbMRE11 and TbFAN1-deficient parasites display the same sensitivity to mechlorethamine as Tbmre11Δ.

Figure 5.15: Susceptibility screening of T. brucei Tbfan1 double null lines against mechlorethamine and DFMO.
A. Dose response curves of T. brucei wild type (●), Tbsnm1Δ (▲), Tbfan1Δ (♦), and Tbsnm1Δ Tbfan1Δ (○) lines to mechlorethamine and DFMO.
B. Dose response curves of T. brucei wild type wild type (●), Tbmre11Δ (▲), Tbfan1Δ (♦), and Tbmre11Δ Tbfan1Δ (○) to mechlorethamine and DFMO.
All data points are mean values ± standard deviations from experiments performed in quadruplicate.
<table>
<thead>
<tr>
<th>Treatment</th>
<th>T. brucei EC50 (µM)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wild type</td>
<td>Tbsnm1Δ</td>
<td>Tbfan1Δ</td>
<td>Tbsnm1Δ Tbfan1Δ</td>
</tr>
<tr>
<td>mechlorethamine</td>
<td>76.28 ± 1.29</td>
<td>10.50 ± 0.92 (7.2)*</td>
<td>37.92 ± 6.03 (2.0)*</td>
<td>3.67 ± 0.34 (20.5)*</td>
</tr>
<tr>
<td>DFMO</td>
<td>28.36 ± 0.54</td>
<td>31.41 ± 0.61 (1.2)</td>
<td>35.64 ± 2.29 (0.8)</td>
<td>31.20 ± 4.79 (0.9)</td>
</tr>
<tr>
<td>mechlorethamine</td>
<td>76.28 ± 1.29</td>
<td>10.38 ± 1.05 (7.6)*</td>
<td>37.92 ± 6.03 (2.0)*</td>
<td>9.63 ± 1.67 (7.6)*</td>
</tr>
<tr>
<td>DFMO</td>
<td>41.32 ± 0.81</td>
<td>47.35 ± 5.16 (0.9)</td>
<td>39.78 ± 1.61 (1.1)</td>
<td>43.54 ± 3.75 (1.0)</td>
</tr>
</tbody>
</table>

Table 5.4: Susceptibility of T. brucei null lines to mechlorethamine and DFMO.

Data represents EC50 values of T. brucei wild type, Tbsnm1Δ, Tbfan1Δ, Tbmre11Δ, Tbsnm1Δ Tbfan1Δ and Tbmre11Δ cells treated with mechlorethamine and DFMO. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC50 values between wild type and null mutant lines is given as fold difference. *Indicates significant differences in susceptibility (P < 0.0001 between wild type and null lines, as assessed by Student’s t test (GraphPad Software).

5.4 Role of TbFANCD1 and TbFANCR in trypanosomal ICL repair

To evaluate other T. brucei FA repair proteins, parasite lines lacking TbFANCD1 (also known as TbBRCA2; (Hartley and McCulloch, 2008) and TbFANCR (also known as TbRAD51; (Dobson et al., 2011) were obtained from Dr Richard McCulloch, University of Glasgow. The susceptibility of these lines towards mechlorethamine was assessed using a modified version of the growth assay reported elsewhere in this thesis: Here the initial cell density used was 10-fold higher (10^5 parasites ml^{-1}) than employed elsewhere due to the slow growth rate of these parasites (Hartley and McCulloch, 2008; Dobson et al., 2011). The resultant dose response curves (Figure 5.16) and EC50 values (Table 5.4) indicate that both enzymes function in the T. brucei ICL REPAIRtoire. Given that TbBRCA2 and TbRAD51 are known components of the HR pathway, we hypothesise that these function in the TbMRE11-dependent ICL repair system although this may be difficult to confirm given the slow growth rates of all the single null lines.
Figure 5.16: Susceptibility screening of Tbbra2Δ cells and Tbrad51Δ cells against mechlorethamine.

A. Dose response curves of T. brucei wild type (●) and Tbbra2Δ (♦) cells to mechlorethamine.
B. Dose response curves of T. brucei wild type (●) and Tbrad51Δ (♦) cells to mechlorethamine.
All data points are mean values ± standard deviations from experiments performed in quadruplicate.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>T. brucei EC50</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wild type</td>
</tr>
<tr>
<td>Mechlorethamine</td>
<td>138.00 ± 3.0</td>
</tr>
</tbody>
</table>

Table 5.5: Susceptibility of T. brucei null lines to mechlorethamine.
Data represents EC50 values of T. brucei wild type, Tbrad51Δ and Tbbra2Δ cells treated with mechlorethamine. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC50 values between wild type and null mutant lines is given as fold difference. *Indicates significant differences in susceptibility (P < 0.0001 between wild type and null lines, as assessed by Student’s t test (GraphPad Software).
5.5 Chapter Summary

This chapter has focused on characterising *T. brucei* components the Fanconi Anaemia repair system. This work has shown:

1. *T. brucei* has potential to express a minimal FA pathway, containing 9 putative FA homologues
2. TbHEL1 (initially believed to be a FANCM homologue) plays no role in DNA or ICL repair.
3. TbHEL1 plays a crucial role in the *T. brucei* RNA interference machinery, potentially functioning as the RNA helicase that is absent from the *T. brucei* DICER sequence but present in the counterparts expressed by higher eukaryotes
4. TbCHL1 (a FANCJ homologue) appears to play an ancillary role in the *T. brucei* ICL REPAIRtoire and may have a regulatory function
5. TbFAN1 plays an ancillary role in the *T. brucei* ICL REPAIRtoire and functions in the MRE11-dependent ICL repair system
6. TbBRCA2 (FANCD1) and TbRAD51 (FANCO) play roles in the *T. brucei* ICL REPAIRtoire predicted to be in the MRE11-dependent ICL repair system
6. **Structure and function analysis of TbSNM1**

6.1 **Understanding the role of conserved residues within TbSNM1**

Previous work characterising TbSNM1 identified several conserved residues that were hypothesised to play roles in 1. zinc co-factor binding, 2. localisation or 3. nucleic acid specificity (Figure 6.1) (Sullivan *et al.*, 2015). The amino acids involved in metal ion binding (either histidine or aspartic acid) are mostly found in motifs that comprise the MβL domain while the presence of a monopartite ‘RRRH’ sequence present in the β-CASP domain has potential to target the enzyme to the nucleus. Additionally, a conserved valine residue found towards the carboxyl end of β-CASP could determine the substrate specificity of TbSNM1: SNM1/PSO2 proteins containing this particular residue tend to bind DNA while the presence of histidine results in enzymes interacting with RNA (Callebaut *et al.*, 2002).

![Figure 6.1: Conserved residues of TbSNM1.](image)

TbSNM1 contains two key regions; the metallo-β-lactamase (MβL; red) and β-CASP (purple) domains. Conserved residues within these domains are emphasised in blue (hypothesised to play a role in zinc cofactor binding), green (theorised to play a role in TbSNM1 nuclear localisation) and orange (TbSNM1 specificity to binding DNA).

Functional analysis of TbSNM1 in BSF *T. brucei* has shown that cells lacking this activity are highly susceptible to ICL inducing agents, a phenotype that can be readily reversed by ectopic expression of an amino terminally GFP tagged version of the enzyme (Sullivan *et al.*, 2015). Using pRPa-GFP-Tbsnm1, the complementation DNA vector used in the above study as template, the SPRINP site-directed mutagenesis technique (Edelheit, Hanukoglu and Hanukoglu, 2009) was used to mutate the residues highlighted in Figure 6.1 to alanine. Following the introduction of the mutated GFP constructs into *T. brucei* Tbsnm1 heterozygotes (heterozygote line referred to as *Tbsnm1*+/−), the remaining Tbsnm1 allele was interrupted thereby creating the desired null mutant line. Comparison of the susceptibility profiles displayed by these lines against wild type, Tbsnm1Δ and Tbsnm1Δ parasites engineered to express GFP tagged Tbsnm1 enabled us to test whether a given residue does play a role in
TbSNM1 1. catalytic activity, 2. localisation and/or 2. substrate specificity. This functional analysis was augmented by in silico studies where the TbSNM1 protein sequence was modelled onto the recently described HsSNM1A (PDB: 4B87) and HsSNM1B (PDB: 5AHO) structures (Figure 6.2) (Allerston et al., 2015) using Phyre 2 and PyMOL programs. Analysis at the gross level indicated that TbSNM1 contained 12 β-pleated sheets that fold to form the MβL domain with the β-CASP domain inserted between β-sheets 10 and 11, typical of members of the SNM1/PSO2 family. Further predictions indicate that the T. brucei enzyme binds one zinc ion similar to HsSNM1A and distinct from HsSNM1B which binds two, with analysis of interactions at specific sites showing the molecular effects of each individual mutation.

Figure 6.2: Predicted model of TbSNM1.
Molecular model of TbSNM1, as compared against HsSNM1A and HsSNM1B. Created using Phyre 2 and Pymol. The α-helices in TbSNM1 are shown in blue (also blue in HsSNM1A and HsSNM1B) with β-pleated sheets shown in red (purple in HsSNM1A and HsSNM1B).
6.2 Analysis of putative TbSNM1 zinc binding residues

Comparison of the TbSNM1 sequence with other SNM1/PSO2 proteins identified several residues, including D40, H78, H80, D82, H83, H159, H178, D221 and H496 (numbering for TbSNM1), as potentially playing a role in zinc binding. Using the plasmid pRPa-gfp-Tbsnm1 as template, the codons that translate for each of these amino acids were mutated to alanine using the SPRINP approach (Figure 6.3).

Figure 6.3: Residues in TbSNM1 postulated to bind the metal ion. Nine conserved residues across both TbSNM1 domains (D40, H78, H80, D82, H83, H159, H178, D221 and H496) are hypothesised to play a role in coordination of the zinc ion. These were independently mutated to alanine residues.

Following confirmation of their sequence, the mutated plasmids were digested with AscI with the purified fragments introduced into a T. brucei 2T1 Tbsnm1+/− line (this line is blasticidin and puromycin resistant). Cultures were placed under selective pressure and hygromycin/blasticidin resistance cultures selected 5 to 7 days post-transformation: integration of RPa-GFP-Tbsnm1 is selected for using hygromycin while blasticidin is used to maintain selection for the Tbsnm1 disruption. Cells were then tested for puromycin sensitivity (2T1, the parental line, is puromycin resistant, a phenotype that should be lost upon integration of the ectopic construct) with susceptible clones then transformed with DNA fragments containing a puromycin resistant cassette that facilitates interruption of the remaining Tbsnm1 allele. As before, cultures were placed under selective pressure and hygromycin/blasticidin/puromycin resistance cultures selected 5 to 7 days post-transformation.

To confirm that Tbsnm1 had been successfully interrupted, that mutated gfp-Tbsnm1 DNA fragment had integrated into the T. brucei genome and that the recombinant lines expressed the fusion gene, DNA amplification-based validation strategies were performed. To assess for appropriate integration events, PCR on gDNA extracted from wild type and manipulated cell
lines was carried out using primer combinations to amplify biomarkers corresponding to 1. intact Tbsnm1 (Tbsnm1-KO1/Tbsnm1-seq1), 2. the Tbsnm1-bla and Tbsnm1-pac disrupted alleles (Tbsnm1-KO1/bla-1 and Tbsnm1-KO1/pac-1, respectively) and 3. the gfp-Tbsnm1 allele (GFP/Tbsnm1-seq1) (Figure 6.4). Amplification of intact Tbsnm1 generated a band of 1.1 kb (the size predicted from in silico mapping) in templates purified from wild type and Tbsnm1+/− cells with no band(s) seen in genetic material purified from the other lines tested. In contrast, primer combinations that detect the Tbsnm1-bla and Tbsnm1-pac alleles generated amplicons of the predicted size (~1.3 and ~1.7 kb, respectively) but only from the gDNA purified from T. brucei Tbsnm1Δ, T. brucei 2T1 Tbsnm1Δ gfp-Tbsnm1 and all the Tbsnm1Δ lines engineered to express mutated versions of the GFP-TbSNM1 fusion protein. In contrast, no bands were observed in wild type T. brucei while only the Tbsnm1-bla allele was detected in the Tbsnm1+/− line. Finally, a DNA band of 0.4 kb (the size predicted from in silico mapping) that spans the gfp-Tbsnm1 fusion was detected only from parasites containing wild type or mutated Tbsnm1 in the ectopic expression vector. These strategies indicate that the appropriate cell lines have been made:

- T. brucei 2T1 Tbsnm1Δ gfp-Tbsnm1-D40A → line now referred to as TbSNM1-D40A
- T. brucei 2T1 Tbsnm1Δ gfp-Tbsnm1-H78A → line now referred to as TbSNM1-H78A
- T. brucei 2T1 Tbsnm1Δ gfp-Tbsnm1-H80A → line now referred to as TbSNM1-H80A
- T. brucei 2T1 Tbsnm1Δ gfp-Tbsnm1-H82A → line now referred to as TbSNM1-H82A
- T. brucei 2T1 Tbsnm1Δ gfp-Tbsnm1-D83A → line now referred to as TbSNM1-D83A
- T. brucei 2T1 Tbsnm1Δ gfp-Tbsnm1-H159A → line now referred to as TbSNM1-H159A
- T. brucei 2T1 Tbsnm1Δ gfp-Tbsnm1-H178A → line now referred to as TbSNM1-H178A
- T. brucei 2T1 Tbsnm1Δ gfp-Tbsnm1-D221A → line now referred to as TbSNM1-D221A
- T. brucei 2T1 Tbsnm1Δ gfp-Tbsnm1-H496A → line now referred to as TbSNM1-H496A
To assess whether the above recombinant lines express the GFP-TbSNM1 fusion, DNA amplification reactions were carried out on cDNA (generated from total RNA extracted from tetracycline induced parasites) using primers GFP/Tbsnm1-seq1 (Figure 6.5): ectopic expression of the wild type or mutated GFP-TbSNM1 protein in the 2T1 genetic background is under control of a tetracycline induced T7 promoter system (Alsford et al., 2005). When using material generated from wild type T. brucei, Tbsnm1+/− and Tbsnm1Δ cells, no band was observed indicating these lines do not express a version of gfp-Tbsnm1. In contrast, a band of the expected size (0.4 kb) was noted in all lines engineered to express variants of the fusion protein. The integrity of all cDNA (and hence RNA) samples was confirmed as control reactions amplifying Tbert (primer combinations Tbert-F/Tbert-R) carried out in parallel yielded a band of the expected size (~100 bp).
Figure 6.5: Confirming expression of \textit{gfp}-\textit{Tbsnm1}.
A. Schematic of integrated \textit{gfp}-\textit{Tbsnm1} analysing using primer combination P2 (Tbsnm1-seq1) and P5 (GFP) which amplified a band of 0.4kb across the GFP and TbSNM1 fusion.
B. DNA fragments obtained from amplification on cDNA derived from RNA extracted from \textit{T. brucei} wild type (lane 1), \textit{Tbsnm1}\textsuperscript{+/-} (lane 2), \textit{Tbsnm1}Δ (lane 3), \textit{Tbsnm1}Δ GFP-\textit{Tbsnm1} (lane 4), \textit{TbSNM1}-D40A (lane 5), \textit{TbSNM1}-H78A (lane 6), \textit{TbSNM1}-H80A (lane 7), \textit{TbSNM1}-D82A (lane 8), \textit{TbSNM1}-H83A (lane 9), \textit{TbSNM1}-H159A (lane 10), \textit{TbSNM1}-H178A (lane 11), \textit{TbSNM1}-D221A (lane 12), and finally \textit{TbSNM1}-H496A (lane 13) cells using primer combinations that specifically amplify \textit{gfp}-\textit{Tbsnm1} (GFP/Tbsnm1-seq1) and \textit{Tbtert} (Tbtert-F/Tbtert-R).

To evaluate whether a given TbSNM1 mutation effects localisation of the enzyme, the DNA of fixed parasites from tetracycline induced and non-induced cultures was stained with DAPI then imaged by epifluorescent microscopy (Figure 6.6). For most (8) TbSNM1 mutants, tetracycline induced expression of the GFP tagged protein within the trypanosome resulted in a pattern that coincided with the larger and more intense of the DAPI signals. As DAPI stains the trypanosomal nuclear and single mitochondrial genomes, with the latter presenting as a small, faint band relative to the former, the observed pattern reveals that most fusion proteins can be translocated to the \textit{T. brucei} nucleus. It is therefore implicit that mutation of most putative zinc binding amino acids does not affect protein localisation and presumably folding. The only exception to this pattern related to the TbSNM1-D221A line where a non-specific GFP signal was noted throughout the cell. This suggests that this particular amino acid may play a key role in TbSNM1 folding with the metal co-factor potentially affecting this process.
Figure 6.6: Localisation of mutant GFP-TbSNM1 in bloodstream form *T. brucei*.

*T. brucei* *Tbsnm1Δ* cells expressing GFP-TbSNM1 (wild type or mutant) were fixed, stained with DAPI and imaged by epifluorescent microscopy under a x40 objective. Examples of the phase, DAPI and GFP signal images are shown for selected cell lines: *Tbsnm1Δ* GFP-*Tbsnm1* (A), TbSNM1-D40A (B), TbSNM1-H80A (C), TbSNM1-D221A (D) and TbSNM1-H496A (E). The cells analysed were derived from tetracycline treated (induced) cultures and compared to parasites grown in the absence of this antibiotic (uninduced).

To assess if ectopic expression of a given TbSNM1 mutation could complement for the susceptibility phenotype displayed by *Tbsnm1Δ*, the growth of recombinant parasite lines in the presence/absence of tetracycline and mechlorethamine was determined. The data was plotted as dose response curves (Figure 6.7) from which EC₅₀ values were determined (Table 6.1). A ratio between uninduced and induced EC₅₀ values is shown as a fold differences (Table 6.1).
Figure 6.7: Susceptibility of *T. brucei* lines towards mechlorethamine.

Dose response curves of tetracycline induced (▲) and uninduced (▼) *T. brucei TbsnmlΔ gfp-Tbsnml* (A), TbsNM1-D40A (B), TbsNM1-H80A (C) and TbsNM1-H496A (D). All data points are mean values ± standard deviations from experiments performed in quadruplicate.

<table>
<thead>
<tr>
<th>Cell line</th>
<th><em>T. brucei</em> EC50 (µM)</th>
<th>Fold difference</th>
<th>Complements?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>uninduced</td>
<td>induced</td>
<td></td>
</tr>
<tr>
<td>TbsnlΔ gfp-TbsnlΔ</td>
<td>4.52 ± 0.41</td>
<td>14.69 ± 1.78</td>
<td>3.3*</td>
</tr>
<tr>
<td>TbsNM1-D40A</td>
<td>5.27 ± 0.43</td>
<td>6.45 ± 0.59</td>
<td>1.2</td>
</tr>
<tr>
<td>TbsNM1-H78A</td>
<td>5.40 ± 0.35</td>
<td>16.67 ± 0.92</td>
<td>3.1*</td>
</tr>
<tr>
<td>TbsNM1-H80A</td>
<td>5.52 ± 3.68</td>
<td>5.27 ± 0.79</td>
<td>1.0</td>
</tr>
<tr>
<td>TbsNM1-D82A</td>
<td>4.50 ± 1.84</td>
<td>4.26 ± 0.80</td>
<td>0.9</td>
</tr>
<tr>
<td>TbsNM1-H83A</td>
<td>2.00 ± 0.89</td>
<td>10.39 ± 0.37</td>
<td>5.2*</td>
</tr>
<tr>
<td>TbsNM1-H159A</td>
<td>4.70 ± 0.25</td>
<td>5.00 ± 0.10</td>
<td>1.1</td>
</tr>
<tr>
<td>TbsNM1-H178A</td>
<td>1.73 ± 0.58</td>
<td>12.91 ± 0.55</td>
<td>7.5*</td>
</tr>
<tr>
<td>TbsNM1-D221A</td>
<td>3.06 ± 0.44</td>
<td>4.48 ± 1.67</td>
<td>1.5</td>
</tr>
<tr>
<td>TbsNM1-H496A</td>
<td>5.19 ± 0.20</td>
<td>22.81 ± 1.80</td>
<td>4.4*</td>
</tr>
</tbody>
</table>

Table 6.1: Susceptibility of *T. brucei* null lines to mechlorethamine.

Data represents EC50 values of induced and uninduced null lines towards mechlorethamine. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC50 values between induced and uninduced null lines is given as a fold difference. *Indicates significant differences in susceptibility (P < 0.0001 between wild type and null lines, as assessed by Student’s *t* test (GraphPad Software).
The above data demonstrates that expression of H78, H83 H178 and H496 mutant forms of TbSNM1 can readily compensate for the mechloretamine sensitivity phenotype displayed by the TbSNM1 null line with these enzymes behaving similarly to wild type TbSNM1 analysed in parallel. It is therefore implicit that these residues do not play a fundamental role in the activity of this enzyme. In comparison, alteration of D40, H80, D82 and H159 results in impaired TbSNM1 activity such that expression of these mutants in the Tbsnm1Δ line fails to complement for the susceptibility phenotype. Molecular modelling of TbSNM1 wild type and mutants (Figure 6.8) structures indicates that these four residues all cluster near to the site where the zinc ion is predicted to sit. It is tempting to speculate that conversion of D40, H80, D82 and H159 to alanine affects coordination of the metal ion possibly affecting electron distribution in this region which can manifest itself as a reduction in the DNA cleaving activity. Whether the above amino acids interact directly with the zinc ion itself, function in an ancillary role to help with in this interaction or facilitate binding and/or cleavage of DNA backbone could not be determined due to the lack of an experimentally derived structure.

Figure 6.8: Structure of wild type and mutated forms of TbSNM1.
Illustration showing the effect of TbSNM1 D40A, H80A, D82A and H159A mutations (post mutation) in relation to wild type TbSNM1 (pre-mutation). Images were generated using the PyMOL in silico mutagenesis tool. The blue sphere corresponds to where the zinc ion is predicted to sit and the red arrows represent the mutation under analysis.
6.3 Evaluating the specificity of TbSNM1 to DNA

TbSNM1 possesses a conserved PTV motif in the carboxyl region of its β-CASP domain with the valine at this site (V518) possibly involved in DNA binding (Callebaut et al., 2002). To test this hypothesis, the codon for this amino acid was mutated to the base triplet for alanine (Figure 6.9) using pRPA-gfp-Tbsnm1 as template. The confirmed ‘mutant’ plasmid was used to generate the *T. brucei* 2T1 Tbsnm1Δ gfp-Tbsnm1-V518A (referred to as TbSNM1-V518A) line and subsequently validated as described in section 6.2 (validation shown in Appendix 5).

![Figure 6.9: Location of V518 residue.](image)

TbSNM1 V518 (located in the β-CASP domain of the protein) was mutated into an alanine due to non-bulky, chemically inert nature of this residue using site directed mutagenesis on the pRPA-gfp-Tbsnm1 plasmid.

Characterisation of the above cell line revealed that the mutated TbSNM1 was targeted to the nucleus (Appendix 5) while mechlorethamine susceptibility screens revealed that the modified enzyme could reverse the sensitivity phenotype that stems from lack of TbSNM1 expression from the endogenous loci (Figure 6.10; Table 6.2).

![Figure 6.10: Susceptibility of *T. brucei* lines towards mechlorethamine.](image)

Dose response curves of induced (▲) and non-induced (▼) *T. brucei* Tbsnm1Δ gfp-Tbsnm1 (A) and TbSNM1-V518A (B). All data points are mean values ± standard deviations from experiments performed in quadruplicate.
Table 6.2: Susceptibility of *T. brucei* lines to mechlorethamine.

Data represents EC\(_{50}\) values of induced and uninduced null lines towards mechlorethamine. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC\(_{50}\) values between induced and uninduced null lines is given as a fold difference. *Indicates significant differences in susceptibility (P < 0.0001 between wild type and null lines, as assessed by Student’s t test (GraphPad Software).

<table>
<thead>
<tr>
<th>Cell line</th>
<th><em>T. brucei</em> EC(_{50})</th>
<th>Fold difference</th>
<th>Complements?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>uninduced</td>
<td>induced</td>
<td></td>
</tr>
<tr>
<td>Tbsnm1Δ gfp-Tbsnm1</td>
<td>4.52 ± 0.41</td>
<td>14.69 ± 1.78</td>
<td>3.3*</td>
</tr>
<tr>
<td>TbSNM1-V518A</td>
<td>3.70 ± 0.34</td>
<td>15.37 ± 0.68</td>
<td>4.2*</td>
</tr>
</tbody>
</table>

Together, this implies that the valine at position 518 in TbSNM1 plays no role in determining substrate specificity. However, molecular modelling (Figure 6.11) at this site shows that because of the relatedness of the valine and alanine structures, the latter amino acid may, at least partially, be still able to promote DNA binding. As such additional experiments to analyse this particular residue could involve mutation of the valine to more diverse amino acids with one potential alteration involving conversion to histidine, the residue that may confer RNA interacting activity (Callebaut *et al.*, 2002).

**Figure 6.11: Effect of the V518A mutation on TbSNM1.**

Structural effects of mutating V518 into an alanine (mutated residue emphasised with red arrow).
6.4 Dissecting the TbSNM1 nuclear localisation signal

*In silico* analysis of TbSNM1 (NucPred software) predicted that the RRRH motif starting at position 428 could function as a nuclear localisation signal (NLS) (Sullivan *et al.*, 2015). The combination of amino acids at this site (Figure 6.12) forms a classical, monopartite NLS with the basic amino acids in this region guiding the protein to the neutrally-charged nucleus (Lange *et al.*, 2007). This type of signal has been observed to function in trypanosomatids (Marchetti *et al.*, 2000).

![Figure 6.12: Predicted NLS of TbSNM1.](image)

Four key residues (R428, R429, R430, and H431A; located in the β-CASP domain of the protein) were identified *in silico* to function in TbSNM1 localisation to the nucleus. Each residue was mutated to alanine to study this hypothesised function.

Using the approaches described above, variants of pRPA-GFP-Tbsnm1 containing appropriate mutations were generated, used to construct the corresponding recombinant *T. brucei* lines (*Tbsnm1Δ gfp-Tbsnm1-R428A* (referred to as TbSNM1-R428A), *Tbsnm1Δ gfp-Tbsnm1-R429A* (referred to as TbSNM1-R429A), *Tbsnm1Δ gfp-Tbsnm1-R430A* (referred to as TbSNM1-R430A) and *Tbsnm1Δ gfp-Tbsnm1-H431A* (referred to as TbSNM1-H431A)) with these parasites then validated using the aforementioned PCR-based strategy (validation shown in Appendix 5).

To assess the effect of each mutation on TbSNM1 localisation, the DNA of trypanosomes grown in the absence or presence of tetracycline was stained, the cells visualized by epifluorescent microscopy and the GFP pattern analysed (Figure 6.13). In situations where the outer two amino acids of the RRRH tetrapeptide had been mutated (ARRH or RRRA) a fluorescence signal within the nucleus was observed but only in tetracycline induced cultures. In contrast, alteration of each of the middle amino acids (RARH or RRAH) resulted in a non-nuclear arrangement thus indicating that the arginine residues at position 429 or 430 are both
key for targeting TbSNM1 to this organelle. However, it must also be noted that the pattern observed when expressing the R429 or 430 mutants in *T. brucei* could also reflect that these proteins are incorrectly folded and therefore non-functional.

**Figure 6.13: Localisation of *T. brucei* mutant GFP-TbSNM1.**
Parasites expressing GFP-TbSNM1 (mutant or wild type) were stained with DAPI and examined by epifluorescent microscopy under a x40 objective lens with phase, DAPI and GFP signal images shown: Tb*snm1*Δ GFP-Tb*snm1* (A), TbSNM1-R428A (B), TbSNM1-R429A (C), TbSNM1-R430A (D) and TbSNM1-H431A (E).

To determine whether there was an association between the above localisation patterns and mechloretamine susceptibility, Tb*snm1*Δ parasites that also express the mutant TbSNM1s cultured in medium containing or lacking tetracycline were treated with the ICL inducing agent and growth inhibition determined. The resulting cell number data was analysed and plotted as dose response curves (Figure 6.14), generating EC$_{50}$ and fold difference values (Table 6.3).
Figure 6.14: Susceptibility of *T. brucei* conditional null lines against mechlorethamine.
Dose response curves of induced (▲) and non-induced (▼) *T. brucei* Tbsnm1Δ gfp-Tbsnm1 (A), Tbsnm1-R428A (B), Tbsnm1-R429A (C), Tbsnm1-R430A (D) and Tbsnm1-H431A (E). All data points are mean values ± standard deviations from experiments performed in quadruplicate.

<table>
<thead>
<tr>
<th>Cell line</th>
<th><em>T. brucei</em> EC50 uninduced</th>
<th><em>T. brucei</em> EC50 induced</th>
<th>Fold difference</th>
<th>Complements?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tbsnm1Δ gfp-Tbsnm1</td>
<td>4.52 ± 0.41</td>
<td>14.69 ± 1.78</td>
<td>3.3*</td>
<td>Y</td>
</tr>
<tr>
<td>Tbsnm1-R428A</td>
<td>8.44 ± 0.47</td>
<td>54.70 ± 1.32</td>
<td>6.5*</td>
<td>Y</td>
</tr>
<tr>
<td>Tbsnm1-R429A</td>
<td>3.97 ± 0.14</td>
<td>4.14 ± 0.12</td>
<td>1.0</td>
<td>N</td>
</tr>
<tr>
<td>Tbsnm1-R430A</td>
<td>7.99 ± 0.26</td>
<td>7.81 ± 0.20</td>
<td>1.0</td>
<td>N</td>
</tr>
<tr>
<td>Tbsnm1-H432A</td>
<td>4.91 ± 0.32</td>
<td>46.63 ± 0.56</td>
<td>9.5*</td>
<td>Y</td>
</tr>
</tbody>
</table>

Table 6.3: Susceptibility of *T. brucei* null lines to mechlorethamine.
Data represents EC50 values of induced and uninduced null lines towards mechlorethamine. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC50 values between induced and uninduced null lines is given as a fold difference.

*Indicates significant differences in susceptibility (P < 0.0001 between wild type and null lines, as assessed by Student’s t test (GraphPad Software).

The above complementation data mirrors the localisation observations. Expression of Tbsnm1 mutated in either of the outer two amino acids (R428 and H431) readily complements for the mechlorethamine sensitivity displayed by cells that lack Tbsnm1 expressed from an endogenous loci. In contrast, alteration of R429 and 430 residues in Tbsnm1 to alanine generates variant proteins that are unable function in the same way.
6.5 Chapter Summary
This chapter has focused on understanding the structure and function of TbSNM1. The amalgamation of alanine scanning, conditional null mutant lines, cellular localisation assays and susceptibility screening demonstrated that:

1. D40, H80, D82, and H159 play a key role in TbSNM1 activity, with comparative studies and in silico modelling suggesting they may be involved in zinc co-factor binding
2. R429 and R430, but not R428 and H431, play a crucial role in the localisation of the protein to the nucleus
3. Conversion of the valine at position 518 to alanine has no effect on TbSNM1 localisation and activity
7. DNA damage and trypanocidal compounds

A major challenge in the field of infectious disease biology is the development of new therapies to target a given pathogen. In the case of HAT, the last treatment to be added to the WHO List of Essential Medicines was NECT in 2009. The anti-microbial effects of this therapy uses two partner drugs that have been in clinical use to target trypanosomal infections for more than 30 years. In this chapter we have utilised the cell lines reported earlier in the thesis to assess whether: 1. TbSNM1 can be exploited as a drug target using compounds supplied by Prof Chris Scoffield (University of Oxford), 2. novel prodrugs activated by a trypanosome specific nitroreductase mediate their downstream effects by promoting DNA damage and 3. existing nitroheterocyclic-based therapies cause DNA damage within the parasite.

7.1 Evaluating trypanocidal effects of HaSNM1A inhibitors

ICL inducing agents are extensively used in medicine to treat a range of conditions with compounds including mitomycin C, cisplatin, cyclophosphamide and mechlorethamine being used to target certain forms of cancer (Huang and Li, 2013). It is envisaged that cells compromised in their ICL repair systems should be more susceptible to such therapies. This has led to the identification of several cephalosporin-based SNM1 inhibitors that are believed to interact with this nuclease via their MβL fold (Lee et al., 2016): Cephalosporins are β-lactam antibiotics target bacterial cell wall biosynthesis with resistance mediated by β-lactamases, enzymes that contain a MβL domain.

To examine whether these SNM1 inhibitors display activity against BSF T. brucei, the growth inhibition of Ceftriaxone (a cephalosporin) and four other compounds (ML302-HCl, ML302-F, LBE314, Si4-33) against the BSF of the parasites was assessed (Figure 7.1; Table 7.1). This revealed that two of the structures (Ceftriaxone and ML302-F) do not have any trypanocidal effects at concentrations up to 100 μM. In contrast, the remaining three compounds displayed moderate, unimpressive, activity against this form of the parasite (EC50 values between 15 to 30 μM).

To elucidate whether this growth inhibition targets TbSNM1, the susceptibility of T. brucei lacking or over expressing this activity was determined (Sullivan et al., 2015) (Figure 7.1; Table 7.1). In all cases, the recombinant lines were as sensitive as wild type to each of the three chemicals tested. Therefore, it is implicit that the trypanocidal activity displayed by these
compounds does not involve TbSNM1 inhibition. Analysis of the TriTryp genome database indicates that in addition to TbSNM1, *T. brucei* expresses at least 8 other sequences that contain a β-lactamase like domain with potentially one (or more) of these being inhibited by the compounds tested here.

Figure 7.1: Susceptibility screening of *T. brucei* against inhibitors of HaSNM1A. Dose response curves of wild type *T. brucei* (●), Tbsnm1Δ (■), and Tbsnm1++ (♦) lines against ML302-HCl, Si4-33, LBE314, ML302-F and ceftriaxone. All data points are mean values ± standard deviations from experiments performed in quadruplicate.

<table>
<thead>
<tr>
<th></th>
<th>Wild type</th>
<th>Tbsnm1Δ</th>
<th>Tbsnm1++</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML302-HCl</td>
<td>26.57 ± 0.99</td>
<td>24.28 ± 1.75</td>
<td>28.45 ± 0.90</td>
</tr>
<tr>
<td>ML302-F Na</td>
<td>&gt;100</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LBE314</td>
<td>13.51 ± 0.25</td>
<td>16.33 ± 1.57</td>
<td>13.02 ± 0.69</td>
</tr>
<tr>
<td>Si4-33</td>
<td>15.41 ± 1.33</td>
<td>12.03 ± 0.80</td>
<td>11.16 ± 0.76</td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td>&gt;100</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 7.1: Susceptibility of *T. brucei* lines to inhibitors of HaSNM1A. Data represents EC$_{50}$ values of wild type *T. brucei*, Tbsnm1Δ and Tbsnm1++ lines towards ML302-HCl, Si4-33, LBE314, ML302-F and ceftriaxone. All values are means ± standard deviations from experiments performed in quadruplicate.
7.2 Unravelling the mode of action of nitro-based prodrugs

Ongoing work in the Wilkinson lab involves determining how nitro-based trypanosomal agents mediate their trypanocidal mode of action (Hall et al., 2010; Bot et al., 2010, Hall, Meredith and Wilkinson, 2012; Meredith et al., 2017). One theme arising from these studies is that the most potent agents with the best selectivity contain chemical motifs that in other organisms promote DNA (specifically ICL) damage. To determine whether this was the case, various trypanocidal nitroaromatic compounds, including benznidazole, nifurtimox, CB1954 and LH34, were screened against a number of single and double null mutant lines engineered to lack different DNA repair proteins (see section 3.2 and 4.1). The data was expressed as dose response curves (example shown in Figure 7.2) from which EC\textsubscript{50} values were extrapolated (Table 7.2).

![Figure 7.2](image)

**Figure 7.2: Susceptibility of *T. brucei* null lines towards anti-trypanosomal agents.**
Dose response curves of *T. brucei* wild type (●), Tbsnm1\textDelta (■), TbcshΔ (♦) and Tbsnm1\textDelta TbcshΔ (ο) to benznidazole, nifurtimox, CB1954 and LH34. All data points are mean values ± standard deviations from experiments performed in quadruplicate.
Table 7.2: Susceptibility of T. brucei double null lines to anti-trypanosomal agents.

Data represents EC<sub>50</sub> values of various T. brucei lines against benznidazole, nifurtimox, CB1954 (all in μM) and LH34 (in nM). All values are means ± standard deviations from experiments performed in quadruplicate. *Indicates significant differences in susceptibility (P < 0.0001 between wild type and null lines, as assessed by Student’s t test (GraphPad Software).

<table>
<thead>
<tr>
<th>Treatment</th>
<th>wild type</th>
<th>Tbsnm1Δ</th>
<th>Tbmre11Δ</th>
<th>Tbsnm1ΔTbmre11Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benznidazole</td>
<td>57.28 ± 4.82</td>
<td>44.84 ± 5.50</td>
<td>22.88 ± 3.50</td>
<td>24.19 ± 2.76 (2.4)*</td>
</tr>
<tr>
<td>Nifurtimox</td>
<td>3.90 ± 0.02</td>
<td>2.86 ± 0.05 (1.3 )</td>
<td>4.67 ± 0.22 (0.9)</td>
<td></td>
</tr>
<tr>
<td>CB1954</td>
<td>3.67 ± 0.24</td>
<td>0.91 ± 0.03 (4.1)*</td>
<td>0.42 ± 0.04 (9.3)*</td>
<td></td>
</tr>
<tr>
<td>LH34</td>
<td>197.85 ± 8.92</td>
<td>9.71 ± 0.88 (20)*</td>
<td>2.55 ± 0.68 (66)*</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wild type</th>
<th>Tbsnm1Δ</th>
<th>Tbexo1Δ</th>
<th>Tbsnm1ΔTbexo1Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benznidazole</td>
<td>57.28 ± 4.82</td>
<td>45.57 ± 4.70 (1.3)</td>
<td>50.29 ± 2.76 (1.1)</td>
</tr>
<tr>
<td>Nifurtimox</td>
<td>3.90 ± 0.02</td>
<td>3.50 ± 0.05 (1.1)</td>
<td>6.07 ± 0.26 (0.7)</td>
</tr>
<tr>
<td>CB1954</td>
<td>2.68 ± 0.04</td>
<td>2.88 ± 0.08 (0.9)</td>
<td>0.68 ± 0.05 (3.9)*</td>
</tr>
<tr>
<td>LH34</td>
<td>188.23 ± 8.94</td>
<td>18.07 ± 1.28 (10)*</td>
<td>22.25 ± 3.82 (8.5)*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wild type</th>
<th>Tbsnm1Δ</th>
<th>TbcshΔ</th>
<th>Tbsnm1ΔTbcshΔ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benznidazole</td>
<td>66.24 ± 0.84</td>
<td>76.39 ± 1.01 (0.9)</td>
<td>60.19 ± 1.61 (1.1)</td>
</tr>
<tr>
<td>Nifurtimox</td>
<td>3.28 ± 0.10</td>
<td>3.15 ± 0.23 (1.0)</td>
<td>3.62 ± 0.19 (0.9)</td>
</tr>
<tr>
<td>CB1954</td>
<td>3.67 ± 0.24</td>
<td>0.72 ± 0.05 (5.3)*</td>
<td>0.07 ± 0.02 (53)*</td>
</tr>
<tr>
<td>LH34</td>
<td>280.45 ± 13.47</td>
<td>15.36 ± 1.41 (18.7)*</td>
<td>11.98 ± 0.78 (23.3)*</td>
</tr>
</tbody>
</table>

The above data reveals that TbEXO1 plays no role in the trypanocidal mechanism of action of any of the nitroaromatic compounds tested: Cells lacking TbEXO1 display sensitivities to benznidazole, nifurtimox, CB1954 and LH34 similar to that observed with wild type cells while T. brucei lacking both TbEXO1 and TbSNM1 exhibit susceptibilities equivalent to those that are deficient only for TbSNM1. This implies that these agents do not promote a form of DNA damage that involves the MMR pathway.

Cells that lack TbCSB, either in a wild type or Tbsnm1-deficient background, are as equally susceptible to benznidazole and nifurtimox as controls. This indicates that these compounds do
not cause a type of DNA damage that involves resolution using the NER pathway. In contrast, TbcsbΔ cells are up to 3-fold more susceptible to CB1954 or LH34 relative to wild type. When examining the sensitivity of cells lacking both TbSNM1 and TbCSB towards these two agents, certain differences were noted. For the LH compound, an epistatic interaction between TbSNM1 and TbCSB was noted with the EC\textsubscript{50} value for the Tbsnm1Δ line being similar to that observed using Tbsnm1Δ TbcsbΔ cells. This suggests that this nitrogen mustard containing compound mediates most of its trypanocidal activity through the formation of ICLs. In the case of CB1954, the double null line (Tbsnm1Δ TbcsbΔ) was hypersensitive to this aziridinyl in relation to the Tbsnm1Δ or TbcsbΔ nulls indicating that the two encoded enzymes function non-epistatically to resolve any DNA damage that may ensue. This may reflect that CB1954 can promote various types of lesion within \textit{T. brucei} that require different systems to repair. It is implicit that as TbSNM1 deficient cells are susceptible to CB1954 then one form of DNA damage caused by this agent are ICLs: this enzyme can only function to repair this type of lesion. Other forms of damage that this aziridinyl-based structure could generate include intrastrand crosslinks and monoalkylation products (Nicolini, 1988) that can be resolved using the NER system, a pathway that does not require SNM1 activity.

Cells lacking TbMRE11 (Tb\textit{mre11}Δ and Tbsnm1Δ Tb\textit{mre11}Δ) are as sensitive to nifurtimox as wild type showing that this nitrofuran does not cause a form of DNA damage that can be repaired by the HR pathway. With benznidazole, parasites deficient in TbMRE11 (Tb\textit{mre11}Δ and Tbsnm1Δ Tb\textit{mre11}Δ) are ~2-fold more sensitive to this compound than the parental line suggesting that these agents do cause a type of DNA damage that can be repaired \textit{via} the HR pathway. This mechanism of action involving TbMRE11 was shown to extend across to megazol (data not shown), a nitroimidazole that has previously been shown to generate HR repaired lesions (Carvalho \textit{et al.}, 2014). When screening CB1954 and LH34 against the above lines, cells lacking just TbMRE11 were between 4- and 6-fold more susceptible to these agents respectively relative to wild type indicating that the HR pathway does play a role in repairing the damage caused by either compound. This repair activity is non-epistatic with TbSNM1 as cells lacking both enzymes are hypersensitive to the treatments.

7.3 \textbf{Activation of benznidazole promotes DNA damage}

It is now established that activation of benznidazole involves a parasite specific type I nitroreductase (TbNTR) (Wilkinson \textit{et al.}, 2008). To determine whether the DNA damaging activity noted above is linked to this trigger mechanism, parasite lines over expressing TbNTR
while null for TbMRE11 were constructed and validated (Appendix 6b). As a control, a TbEXO1 deficient, TbNTR expressing control line was generated in parallel (Appendix 6a). The susceptibility of the resultant lines towards benznidazole was then tested: The effect of DFMO was also analysed as drug control. The growth inhibition data, expressed as dose response curves (Figure 7.3) and as EC$_{50}$ values (Table 7.3), revealed that cells lacking only TbEXO1 were as equally susceptible to benznidazole as wild type, confirming our previous findings, while cells additionally engineered to over express TbNTR were as sensitive to this nitroimidazole as parasite expressing only elevated levels of the nitroreductase.

**Figure 7.3:** Susceptibility of *T. brucei* lines expressing altered levels of TbEXO1, TbMRE11 and/or TbNTR to benznidazole.

Dose response curves of wild type (●), Tbexo1Δ (♦), TbNTR over expression (■), and Tbexo1Δ TbNTR over expression (○) *T. brucei* to benznidazole (A) or DFMO (B).

Dose response curves of wild type (●), Tbmre11Δ (♦), TbNTR overexpression (■), and Tbmre11Δ TbNTR overexpressing (○) *T. brucei* to benznidazole (C) or DFMO (D). All data points are mean values ± standard deviations from experiments performed in quadruplicate.

When benznidazole was tested against *T. brucei* cells expressing altered levels of TbNTR and/or TbMRE11, a different outcome was observed. For TbMRE11-deficient cells or parasites expressing elevated levels of just TbNTR, treatment with the nitroimidazole resulted in increased susceptibility when compared against controls, with the TbNTR overexpressing cells being more sensitive than the null line: Tbmre11Δ cells and *T. brucei* expressing an ectopic
copy of Tbntr were ~1.6- and ~6.5-fold more sensitive to benznidazole, respectively than wild type. For TbMRE11 null parasites overexpressing Tbntr, this increase in potency was enhanced with these cells showing a 19-fold increase in susceptibility indicating a link between activation of benznidazole by Tbntr and downstream DNA lesion formation.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>T. brucei EC₅₀ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wild type</td>
</tr>
<tr>
<td>Benznidazole</td>
<td>48.73 ± 2.30</td>
</tr>
<tr>
<td>DFMO</td>
<td>23.93 ± 1.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>T. brucei EC₅₀ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wild type</td>
</tr>
<tr>
<td>Benznidazole</td>
<td>65.49 ± 2.90*</td>
</tr>
<tr>
<td>DFMO</td>
<td>18.79 ± 2.08</td>
</tr>
</tbody>
</table>

Table 7.3: Data represents EC₅₀ values of genetically engineered lines against benznidazole and DFMO. All values are means ± standard deviations from experiments performed in quadruplicate. The ratio of EC₅₀ values between wild type and null mutant lines is given as fold difference, in parentheses. *Indicates significant differences in susceptibility (P < 0.0001 between wild type and null lines, as assessed by Student’s t test (GraphPad Software).
7.4 Chapter Summary

This chapter has focused on evaluating whether TbSNM1 can be exploited as a drug target and determining if various trypanocidal agents mediate at least part of the mode of action by promoting DNA damage. We have demonstrated that:

1. Some of the SNM1 inhibitors tested displayed trypanocidal activity but this did not involve targeting TbSNM1.
2. Nifurtimox does not appear to promote DNA damage that is repaired by selected components of the NER, MMR, HR or ICL pathways/systems.
3. The nitrogen mustard LH34 promotes formation of ICLs that are resolved using SNM1-dependent and SNM1-independent repair systems.
4. The aziridyl agent CB1954 promotes formation of ICLs that are resolved using the SNM1-dependent and SNM1-independent repair systems.
5. The aziridyl agents CB1954 also promote formation of additional DNA lesions that are repaired by components of the NER pathway.
6. Activation of benznidazole results in cytotoxic moieties that promote a form of DNA damage that can be resolved using components of the HR pathway.
8. Discussion

The plan for the eradication of HAT by 2030 can only take place by the development of new, active preventative/curative therapies. Work carried out in the Wilkinson laboratory has studied clinically used HAT treatment regimes to identify and study the mode of action of novel trypanocidal agents. These studies have broadened our repository of trypanocidal agents and our understanding of trypanosomal cell and molecular biology (Wilkinson and Kelly, 2009; Meredith et al., 2017). These structure–activity relationship studies have focused on nitrogen- and quinone-containing chemicals which have revealed that many of these compounds promote mutagenic and clastogenic effects in the cell. These prodrugs, activated by a type 1 nitroreductase (TbNTR), are processed in the parasite to form an open aziridine structure which has the ability to form adducts with negatively charged DNA, predominantly at guanine bases, with a repeat of this event resulting in the formation of interstrand crosslinks (Hall, Bot and Wilkinson, 2011; Bot et al., 2013; Meredith et al., 2017).

This is the first study to comparatively analyse how trypanosomes repair crosslinks in DNA, with previous work carried out by the Machado & McCulloch laboratories speculating a role for the TC-NER pathway in trypanosomal intrastrand crosslink repair (Machado et al., 2014). Using an in vitro genetics-based approach, a series of null mutant lines were generated to delete/disrupt key components of ‘classical’ DNA repair pathways. Growth inhibition assays were carried out on these cell lines using various agents which promote DNA damage to assess the relative importance of each DNA repair protein in trypanosomal DNA and ICL repair mechanisms. This enabled us to 1. understand whether the DNA repair protein has a role in removing lesions introduced in the genome, 2. appreciate if the enzyme has a role(s) in removing ICLs and 3. unravel the epistatic/non-epistatic relationship between trypanosomal ICL repair factors to create the first model of trypanosomal ICL repair systems. The same cell lines generated for this project were used to evaluate how clinically relevant and novel trypanocidal agents mediate their cytotoxic activities by analysing the prodrug mode of action, especially in relation to ICL formation.

The construction of null mutant lines was carried out using a traditional plasmid- and homologous recombination-based approach. This project could have exploited the newly-developed CRISPR-Cas9 system created for T. cruzi and Leishmania organisms in T. brucei (Peng et al., 2014; Beneke et al., 2017) but given the ease of creating genetically engineered
null lines in *T. brucei* and readily available resources to carry this out, the development and exploitation of a CRISPR-Cas9 system in this parasite was not deemed necessary. However, it would be interesting to use this now well-established tool to confirm observations made in this thesis and to expand these studies into analysis of ICL repair systems in other protozoan parasites.

Phenotypic screening was carried out on all null lines after a PCR-based approach was exploited to validate the genetically engineered lines: Confirmation at the protein level using western blotting-based techniques was not deemed necessary as absent RNA expression can be correlated to a lack of protein expression. This susceptibility screening was not extended out to heterozygote null mutants (which would give us an understanding into haploinsufficiency) due to time constraints with multiple, independently-generated clones analysed to confirm observations made. These phenotypic screens used mechlorethamine as the ICL inducing agent of choice to enable valid comparisons to be made against similar studies carried out in fungal lines (Lam, Krogh and Symington, 2008) and because mechlorethamine binds DNA with high affinity to produce stable ICLs (Bauer and Povirk, 1997). The use of other ICL inducing agents (*e.g.* psoralen-based compounds) would have confirmed conclusions made in this thesis but due to time constraints, this was not deemed a priority.

### 8.1 Trypanosomal ICL repair systems

The afore-mentioned approach has revealed that trypanosomes utilise the activities of TbSNM1, TbMRE11 (HR pathway), TbCSB (NER pathway) and TbEXO1 (MMR pathway) alongside factors of the FA pathway (TbRAD51, TbBRCA2, TbFAN1 and TbCHL1) for repair of ICLs with all data summarised in Table 8.1. Intriguingly, no evidence for a role of TbREV2 and TbREV3 (TLS pathway) was identified. Further studies using double null mutant lines lacking two DNA repair proteins revealed that trypanosomes possess at least two distinct pathways for the repair of ICLs; an SNM1-dependent ICL repair pathway which employs the functions of TbSNM1, TbCSB and TbEXO1 and an SNM1-independent ICL repair pathway, which utilised the functions of TbMRE11, TbFAN1, TbEXO1, and possibly the activities of TbRAD51 and TbBRCA2. Which system predominates in trypanosomes is yet to be answered but it is clear that similar pathways are being exploited by this parasite as seen in *E. coli*, yeast and mammalian cells (Houten *et al.*, 1986; Barber *et al.*, 2005; Hashimoto, Anai and Hanada, 2016).
Alongside this work on dissecting trypanosomal ICL repair systems, studies on how crosslinked DNA is processed to form DSBs were undertaken and demonstrated a clear time-dependent formation of γH2A, a marker for DSBs, in cells treated with the mechlorethamine. It was clear from these results that this ICL inducing agent forms damage in the trypanosome which can be processed to form DSBs, in keeping with previous observations on the processing of crosslinked DNA with the exact mechanism behind this still to be unravelled (Bessho, 2003; Sczepanski et al., 2009). Accumulation of DSBs in mechlorethamine treated cells peaked at 4 hours incubation time which possibly reflects the time taken for the ICL inducing agent to crosslink DNA and produce the DSB. If this is the case, it would be interesting to understand the limiting factor for cleaving ICLs as well as how this relates to the different ICL repair systems. This could be evaluated using PFGE DNA fragmentation studies, the comet assay or the I-SceI meganuclease system (Singh et al., 1988; Lam, Krogh and Symington, 2008; Chiurillo et al., 2016).

8.1.1 SNM1-dependent ICL repair system

Studies on the trypanosomal SNM1-dependent ICL repair system has hypothesised that this mechanism functions by 1. stalling of RNA polymerase for recruitment of TbCSB, which unwinds damaged and neighbouring nucleotides for the recruitment of nucleases, 2. nucleolytic degradation of the unhooked DNA by TbSNM1 up to and past the crosslinked bases, and 3. TbEXO1 nuclease activity on crosslinked DNA as well as processing of base errors introduced into the genome during ICL repair. If this hypothesis is true, the SNM1-dependent ICL repair system operates like that seen in the transcription-coupled ICL repair system identified in mammalian cells as well as the dominant, PSO2-dependent mechanism employed by yeast (K F Grossmann et al., 2001; Enoiu, Jiricny and Schärer, 2012). Given these similarities, the enzymes employed to carry out the initial ICL unhooking event needs to be evaluated to fully understand this pathway (Clauson, Schärer and Niedernhofer, 2013).

8.1.1.1 Nuclease activity of TbSNM1

Studies on the SNM1/PSO2 family of proteins has centered around the key and specific activity of this metallonuclease towards ICLs, as demonstrated in mammalian cells (Munari et al., 2013), yeast (Hazrati et al., 2008), C. elegans (Wilson et al., 2017), A. thaliana (Johnson, Hellens and Love, 2011) and O. sativa (rice) (Kimura et al., 2005). This activity has been extended out to T. brucei with parasites lacking TbSNM1 displaying an increased sensitivity
towards ICL-inducing agents, as shown in both this study and published literature (Sullivan et al., 2015). This project has also shown that cells lacking TbSNM1 do not display any growth defects or arrest at any point in the cell cycle, with no growth defects also noted in yeast and C. elegans (Li, Hejna and Moses, 2005; Wilson et al., 2017). This is possibly linked to TbSNM1 functioning independently of the homologous recombination-coupled ICL repair system.

Moreover, data represented here has not directly implicated TbSNM1 in the repair of pyrimidine dimers, double strand breaks or base oxidation, in keeping with previously published work (Sullivan et al., 2015). However, work shown in this thesis has indicated that TbSNM1 plays a compensatory role for the repair of UV-induced damage in the absence of TbMRE11, indicating that both proteins have common substrates for this form of damage. This observation has not been noted in any previous reports of SNM1/PSO2 activity other than expression of PSO2 is inducible by UV (Wolter, Siede and Brendel, 1996). Both SNM1 and MRE11 possess endonuclease and exonuclease activity (Paull and Gellert, 1998; Li, Hejna and Moses, 2005) with the exonuclease activity being strand specific, *i.e.* SNM1 cleaves strands 5’-3’ and MRE11 cleaves DNA 3’-5’. This suggests that the endonuclease activity of both proteins is required for processing of dimerised DNA at a collapsed replication fork (Paull and Gellert, 1998; Limoli et al., 2002; Li, Hejna and Moses, 2005). It is possible that this endonuclease activity is critical for the creation of DSBs which initiates firing of a new replication origin as well as HR-mediated repair (Jones and Petermann, 2012).

Taking into consideration the wild type-like γH2A formation displayed by cells absent for TbSNM1, it is unlikely that this nuclease has a role in the incision of unhooked ICLs for DSB formation, which has also been observed in yeast where PSO2 nuclease activity does not lead to DSB formation (Dudás et al., 2007). It is most likely that TbSNM1 functions in the cleavage of the DNA backbone, up to and beyond the ICL, revealing the damaged site for further processing (Tiefenbach et al., 2007) and possibly in the processing of the double strand breaks formed from the interstrand cross link (Li and Moses, 2003). It would be interesting to see how this differs in cells treated with UV, given earlier observations.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Growth properties</th>
<th>Growth inhibition assays</th>
<th>DSB formation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Growth</td>
<td>Cell cycle</td>
<td>Phleomycin</td>
</tr>
<tr>
<td><strong>Growth</strong></td>
<td>Cell cycle</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>WT</strong></td>
<td>WT</td>
<td></td>
<td>WT</td>
</tr>
<tr>
<td><strong>WT</strong></td>
<td>WT</td>
<td></td>
<td>WT</td>
</tr>
<tr>
<td><strong>↓</strong></td>
<td>↑1N2K</td>
<td>↓1K1N</td>
<td>↓</td>
</tr>
<tr>
<td><strong>TbSNM1</strong></td>
<td><strong>WT</strong></td>
<td>WT</td>
<td>WT</td>
</tr>
<tr>
<td><strong>TbCSB</strong></td>
<td><strong>WT</strong></td>
<td>WT</td>
<td>WT</td>
</tr>
<tr>
<td><strong>↓</strong></td>
<td>↑1N2K</td>
<td>↓1K1N</td>
<td>↓</td>
</tr>
<tr>
<td><strong>TbMRE11</strong></td>
<td><strong>↓</strong></td>
<td>↑1N2K</td>
<td>↓1K1N</td>
</tr>
<tr>
<td><strong>TbEXO1</strong></td>
<td><strong>WT</strong></td>
<td>ND</td>
<td>↓</td>
</tr>
<tr>
<td><strong>TbREV2</strong></td>
<td><strong>WT</strong></td>
<td>ND</td>
<td>WT</td>
</tr>
<tr>
<td><strong>TbREV3</strong></td>
<td><strong>WT</strong></td>
<td>ND</td>
<td>WT</td>
</tr>
<tr>
<td><strong>TbFAN1</strong></td>
<td><strong>WT</strong></td>
<td>WT</td>
<td>WT</td>
</tr>
<tr>
<td><strong>TbHEL1</strong></td>
<td><strong>WT</strong></td>
<td>ND</td>
<td>WT</td>
</tr>
<tr>
<td><strong>TbCHL1</strong></td>
<td><strong>WT</strong></td>
<td>WT</td>
<td>WT</td>
</tr>
<tr>
<td><strong>TbRAD51</strong></td>
<td><strong>↓↓</strong></td>
<td>WT*</td>
<td>↓*</td>
</tr>
<tr>
<td><strong>TbBRCA2</strong></td>
<td><strong>↓</strong></td>
<td>↑1K1N</td>
<td>↓**</td>
</tr>
</tbody>
</table>

| **Table 8.1: Summary of data presented for all trypanosomal DNA repair factors.** |
| Arrows indicate an increase (↑) or a decrease (↓) in phenotypic response when compared to wild type (WT) cells with multiple arrows shown for larger increases/decreases. ND stands for not determined for when an experiment was not performed. NIL refers to no increase or decrease seen. Corresponding data relating to TbBRCA2 was obtained from (Hartley and McCulloch, 2008) (denoted by *). Data obtained from (McCulloch and Barry, 1999; Proudfoot and McCulloch, 2005) relating to responses of TbRAD51 are denoted by **. |
8.1.1.2 Critical role of TbCSB in damage recognition

CSB (Tijsterman et al., 1997; Stevnsner et al., 2008) has repeatedly been shown to play a pivotal role in transcription-associated damage recognition, an activity which has been extended out to trypanosomal transcription-coupled repair (Machado et al., 2014; Wilson et al., 2017). TbCSB is a nuclear protein (Dean, Sunter and Wheeler, 2017) which repairs UV-induced damage with no evidence for a direct role in repair of DSBs or base oxidation events, similar to observations noted by the McCulloch laboratory (Machado et al., 2014). An intriguing observation from these studies was a delayed accumulation of DSBs in Tbcsb-deficient cells, as seen by a late accumulation of γH2A in mechlorethamine treated null cells. This could be explained by stalling of the transcriptional machinery which causes recruitment of CSB to the site of damage and downstream processing of such damage into double strand breaks for repair via the HR pathway. CSB has been associated with regulation of double strand breaks at sites of transcription which indicates that this helicase remains at the site of the lesion during the repair process, possibly for activation of cell cycle checkpoints (Batenburg et al., 2015). Given that CSB is one of the first factors to be called upon at a stalled RNA polymerase, this key role in transcription-associated recognition of multiple types of damage is plausible (Tijsterman et al., 1997).

The role of TbCSB in repair of crosslinked DNA has been previously studied using cells expressing reduced levels of the helicase coupled with phenotypic screening using the crosslinking agents cisplatin and cyclophosphamide (Machado et al., 2014). This project has extended this initial work by screening of Tbcsb-deficient cells against the archetypal ICL inducing agent, mechlorethamine which has confirmed the role of TbCSB in ICL repair, specifically via the SNM1-dependent ICL repair system. The epistatic interaction between SNM1 and CSB has been extensively studied in mammalian cells (Iyama et al., 2015) with no discernible role for the yeast CSB homolog, RAD26, shown for ICL repair (Lambert et al., 2003; Barber et al., 2005). This work on mammalian cells has revealed that HsCSB stimulates the activity of HsSNM1A to help promote ICL processing in a transcription-associated, replication-independent repair system (Iyama et al., 2015). The process of HsCSB upregulation occurs through halting of the RNA polymerase which leads to a hyperphosphorylation event of the polymerase and immediate upregulation of CSB (Luo et al., 2001; Lagerwerf et al., 2011). HsCSB enhances the substrate specificity of HsSNM1A at the lesion site for processing of the unhooked ICL (carried out by the activities of XPF-ERCC1). It would be convenient to
assume that a similar post-translational modification of trypanosomal RNA polymerases could lead to TbCSB recruitment to site of DNA damage with further studies needed to evaluate this (Deaconescu, 2013). Biochemical in vitro assays are also needed to prove that the epistatic interaction of TbCSB and TbSNM1 is extended out to a physical, protein-protein interaction. With this information, the non-epistatic interaction between TbCSB and TbMRE11 was unsurprising as there is no currently evidence for a role of CSB in recombination-coupled ICL repair (Enoiu, Jiricny and Schärer, 2012).

It is important to note that TbCSB has also been linked to promoting TLS polymerase activity to bypass damaged DNA (Machado et al., 2014). If this is the case, TbCSB may play a critical role in this repair mechanism, acting as a regulator for all steps during transcription-associated repair. Whether CSA plays a similar role in ICL repair is yet to be studied (Iyama, Wilson and III, 2016).

8.1.1.3 The combinatorial role of TbEXO1 in trypanosomal ICL repair

This is the first study to look into the role of TbEXO1, an important nuclease of the MMR pathway, in trypanosomal DNA repair with other factors of the MMR previously studied elsewhere (Bell et al., 2004). This project has shown absence of TbEXO1 does not cause any growth defects, cell cycle defects or enhanced sensitivity to agents which promote base oxidation or pyrimidine dimers. Intriguingly, enhanced sensitivity of this null mutant line was seen towards phleomycin but not MMS, despite both agents promoting DSB formation. This may reflect the differences in how DSBs are generated by both agents with the length of the flap overhang (Pizzolato et al., 2015) affecting the efficiency of the 5'-3' exonuclease activity of TbEXO1 (Keijzers, Bohr and Rasmussen, 2015). DSBs formed from stalling of the replication fork (as seen with MMS (Ensminger et al., 2014)) produce large overhangs, with these substrates displaying a reduced processing efficiency towards EXO1. This differs from DSBs formed from phleomycin-induced oxygen free radicals which create DSBs with overhangs of only a few nucleotides. In this latter case, efficiency/binding of EXO1 to the damaged DNA is higher (Keijzers, Bohr and Rasmussen, 2015). Given the MMR pathway repairs single/few damaged bases, this low processing power for EXO1 is understandable and could be extended out to the activity of TbEXO1 in relation to ICL and MMR repair pathways.
EXO1 has been implicated in the fungal replication-coupled, PSO2-independent ICL repair system for the cleavage of unhooked DNA prior to strand annealing (Lam, Krogh and Symington, 2008; Ward et al., 2012; Kato, Kawasoe, Williams, Coates, Roy, Shi, Lorena S. Beese, et al., 2017). It has been shown that EXO1 acts as the alternative 5'-3' exonuclease in the absence of PSO2 (alongside factors of the MMR factors e.g. MutSa) (Ward et al., 2012). This differs from the situation seen in trypanosomes as TbEXO1 functions in both the SNM1-dependent and -independent ICL repair systems, epistatically interacting with TbSNM1, TbCSB and TbMRE11. Trypanosomes may utilise the activity of TbEXO1 in both systems to 1. function as the alternative nuclease in the SNM1-independent ICL repair pathway (as seen in yeast (Ward et al., 2012; Kato, Kawasoe, Williams, Coates, Roy, Shi, Lorena S Beese, et al., 2017)) and 2. repair errors introduced by the error-prone, SNM1-dependent ICL repair system (this mechanism is called ‘error-prone’ as the activity of TLS polymerases promotes base errors). To further evaluate the latter hypothesis, similar epistatic/non-epistatic studies should be carried out on cells lacking multiple factors of the MMR pathway, with the best candidate for this study being MSH2 given its well-characterised role in ICL repair (Williams et al., 2011; Ward et al., 2012; Kato, Kawasoe, Williams, Coates, Roy, Shi, Lorena S. Beese, et al., 2017).

8.1.1.4 Absent role for TLS factors TbREV2 & TbREV3

The TLS pathway plays a crucial role in ICL repair, filling gaps created by the cleavage activity of nucleases (Ho et al., 2011). Work in yeast has shown us that not all TLS polymerases play a role in ICL repair, where only five out of the fifteen polymerases are shown to function in this system (Kozmin and Jinks-Robertson, 2013). Absence of a growth, DNA or ICL repair phenotype for both TbREV2 and TbREV3, trypanosomal TLS enzymes, was peculiar but given that only two out of five of the TLS polymerases present in T. brucei were studied, it is still plausible that the TLS pathway plays a key role in trypanosomal ICL repair. Evidence for this can be seen when analysing the role of the TLS factor polymerase theta (Θ) of the protozoan parasite Leishmania infantum which has been shown to play a key role in repair of interstrand crosslinks (Fernández-Orgiler et al., 2016). Furthermore, trypanosomal TLS polymerases could be acting in a compensatory manner, functioning on behalf of another trypanosomal TLS polymerase when one is absent. With this in mind, a different approach may have to be undertaken to understand whether these TLS polymerases function in ICL repair. It is
hypothesised that the TLS pathway functions in both the SNM1-dependent and -independent pathway, given how these ICL repair systems operate in yeast and mammalian cells.

8.1.2 SNM1-independent ICL repair system

It is clear that trypanosomes possess a separate ICL repair system which utilises the activities of components of the homologous recombination and mismatch repair pathways, functioning independently of TbSNM1. It is hypothesised that this pathway involves 1. stalling of the replication fork resulting in the recruitment of nuclease to unhook the ICL, 2. cleavage of crosslinked DNA by the nuclease activity of TbMRE11, TbEXO1 and TbFAN1 (possibly resulting in DSB formation), and 3. error-free repair by the HR pathway utilising the activities of the TbRAD51, TbBRCA2 and TbCHL1 proteins. Trypanosomal FA homologues have been implicated in this system as many of the components identified in the parasites genome have additional roles in the HR pathway, a mechanism which has already been associated with the SNM1-independent ICL repair pathway by the activity of TbMRE11.

8.1.2.1 Concerted actions of TbMRE11 & TbEXO1

Studies carried out here on the role of TbMRE11 in trypanosomes has brought to light many interesting observations. The growth defect and G2/M cell cycle arrest displayed by the Tbmre11 null lines suggests this nuclease has a role in nuclear DNA replication. Work in mammalian cells has shown that ATM (ATM serine/threonine kinase) and ATR (Ataxia Telangiectasia) acts in the recruitment of MRE11 at the site of a collapsed replication fork (Trenz et al., 2006). This recruitment of MRE11 initiates the restart of the replication fork, thus preventing accumulation of DSBs and chromosomal abnormalities. T. brucei possesses homologues for both ATM (Tb927.2.2260) and ATR (Tb927.11.14680), which could make up the replication-coupled, DNA damage signalling network in trypanosomatids, alongside the recognition and repair activities of TbMRE11. This agrees with observations made from assays looking at γH2A formation, an indicator for DSBs, where Tbmre11 null cells show no phosphorylation event for this histone in mechlorethamine treated cells, indicating a key role for this nuclease in detection of DSBs and warrants further investigation into the regulation and checkpoint signalling of trypanosomal repair systems.

Phenotypic screening of Tbmre11-deficient cells revealed a clear role for this nuclease in repair of lesions created by phleomycin, MMS, UV and mechlorethamine. Growth inhibition
data relating to MMS and phleomycin agrees with the McCulloch and Cross laboratories which have shown that cells lacking TbMRE11 are more sensitive to phleomycin and MMS, respectively, implicating this nuclease in the HR pathway (Robinson, McCulloch, Conway, Browitt and Barry, 2002; Tan, Leal and Cross, 2002). Our results have also implicated a role for TbMRE11 in repair of UV-induced lesions with work in mammalian cells relating this observation to stalling of the replication fork at the site of dimerised DNA. This event leads to recruitment of the MRN complex for replication fork restart (Limoli et al., 2002), connecting observations made in both trypanosomal and mammalian cell lines.

The role of TbMRE11 in trypanosomal ICL repair was unsurprising due to the well-documented role of MRE11 in both mammalian and yeast cells, functioning independently of SNM1/PSO2 (Pichierri, Averbeck and Rosselli, 2002; Lam, Krogh and Symington, 2008). The epistatic interaction of TbEXO1 and TbMRE11 has been linked with processing of unhooked ICLs as MRE11 possesses 3’-5’ exonuclease activity and EXO1 has 5’-3’ exonuclease activity, ideal for cleavage at this site (Barber et al., 2005). Whether this processing leads to DSB formation is yet to be studied. This data could also explain why trypanosomes do not require the activity of TbSNM1 in the MRE11-dependent system: TbEXO1 is able to compensate for the activity of SNM1 (Pizzolato et al., 2015). It is tempting to speculate that this system is coupled with ICLs encountered during replication due to the correlation between factors of the homologous recombination pathway and DNA damage at a DNA replication fork (S-phase of the cell cycle). This also explains the non-epistatic interaction of TbMRE11 and TbCSB; unwinding of the double helix is not required due to the presence of the large replication bubble.

8.1.2.2 Trypanosomal FA factors and the SNM1-independent ICL repair system

This study has identified and characterised trypanosomal homologues of FA factors with these results demonstrating that trypanosomes employ a similar system to yeast where these FA proteins function alongside the classical DNA repair pathways to resolve such lesions (Daee and Myung, 2012; McHugh, Ward and Chovanec, 2012). The early branching nature of the trypanosomatid tree of life has resulted in trypanosomes exploiting components of classical DNA repair pathways for the removal of crosslinks as opposed to evolving a dedicated ICL repair pathway. Many organisms including A. thaliana (Molinier, Stamm and Hohn, 2004) and S. cerevisiae function in a similar mechanism with these factors predominantly acting in SNM1/PSO2 independent DNA repair pathways. Zebrafish are one of the few organisms, other
than humans, which possess all homologues of the FA core complex with these factors also playing a role in sex determination and genomic stability (Rodríguez-Marf and Postlethwait, 2011).

The FA pathway consists of 1. recognition of the damaged DNA and assembly of eight proteins that form the so-called FA core complex, 2. activation of the FANCD2/FANCI heterodimer that functions as a recruitment complex, and 3. association of effector proteins that operate to resolve the ICL. The absence of all components of the FA core complex (except FANCM) in and homologues of FANCD2 and FANCI suggests that trypanosomes employ a separate mechanism for ICL recognition and factor recruitment, possibly via stalling of the replication fork leading to the recruitment of ATM and/or ATR, which when phosphorylated, are able to upregulate DNA repair factors e.g. EXO1/MRE11 (Moldovan and D’Andrea, 2009). This also explains the absence of UHRF1 and UAF1, both of which are involved in recruitment/deubiquitination of FANCD2, respectively (Moldovan and D’Andrea, 2009; Liang et al., 2015).

Many FA effector proteins are also putatively absent from the trypanosome genome. FANCP, SLX4, is a key endonuclease which plays a role in the initial incision step for unhooking of the ICL and resolution of the Holliday junction (Cybulski and Howlett, 2011). Its absence in trypanosomes is intriguing given various endonucleases are exploited for cleavage of ICLs in mammalian cells. Studies in this thesis have implicated endonucleases TbSNM1, TbMRE11 and TbFAN1 in ICL repair and given that trypanosomes are simple eukaryotes, these parasites may only require a minimal number of nucleases for ICL processing, possibly explaining the absence of FANCP. Furthermore, absence of FANCN (PALB2), FANCS and FANCU (XRCC2) can be explained by TbBRCA2 and TbRAD51 paralogues carrying out their given roles in the FA pathway (Simhadri et al., 2014; Park et al., 2016). The nuclear localisation of FA effector proteins and RAD51-heterodimer formation for strand exchange, which is associated with the activities of these FA factors, can be carried out by TbBRCA2 and TbRAD51 paralogues, explaining their absence from the trypanosomal genome (Dobson et al., 2011; Genois et al., 2014)

Despite the absence of many FA factors, trypanosomes express some FA effector proteins, primarily belonging to the HR pathway. Studies in this thesis have shown a key role for TbRAD51 (or TbFANCR), TbBRCA2 (or TbFANCD1), TbCHL1 (or TbFANCJ) and
8.1.2.3 Functions of TbRAD51 and TbBRCA2 in trypanosomal ICL repair

It is unsurprising that TbRAD51 and TbBRCA2 play a role in ICL repair as both proteins have been extensively characterised in fungal and mammalian ICL repair systems (K F Grossmann et al., 2001; Noll, Mason and Miller, 2006; Ward et al., 2012). TbRAD51 and TbBRCA2 possess recombinase/helicase activity, respectively, to search for homologous sequences on sister chromatid DNA for error-free ICL repair. It would be interesting to understand the role of the four RAD51 paralogs identified within the T. brucei genome (RAD51-3 (also known as FANCO), RAD51-4, RAD51-5 and RAD51-6) and how they influence this activity and ICL repair (Dobson et al., 2011). It would also be interesting to identify whether a single residue in these proteins aids ICL repair, as seen with mammalian RAD51 (Wang et al., 2015), a feature that may be exploitable in terms of chemotherapeutic development.

It would have been interesting to carry out epistatic/non-epistatic relationship studies between TbMRE11 and these enzymes but due to the growth defects observed by the single null lines and the need for the HR pathway to generate these genetically engineered lines, this was not possible. A new approach is needed to fully understand the interaction of these HR factors with one possible method exploiting tagged-TbMRE11 lines generated in the Wilkinson laboratory for co-immunoprecipitation studies with this approach being an excellent tool for understanding this ICL repair at a protein-protein level by revealing the physically interacting partners of TbMRE11.

8.1.2.4 Nuclease activities of TbFAN1 on crosslinked DNA

Recent work has highlighted the importance of FAN1 in this repair system with many research groups demonstrating its function as comparable to SNM1 with activities carried out by a highly conserved and FAN1-specific VRR-NUC domain (O’Donnell and Durocher, 2010; Jin and Cho, 2017). TbFAN1, a nuclear DNA repair protein (Dean, Sunter and Wheeler, 2017), possesses a similar VRR-NUC domain and “PDX\(_d(D/E)XK\)” motif which was sufficient evidence to state this protein as true TbFAN1 homologue (Fontebasso et al., 2013). Phenotypic susceptibility screening of TbFAN1-deficient cells revealed no altered sensitivity to hydroxyurea, UV, MMS and phleomycin with a clear increase in susceptibility to
mechlorethamine. It would be interesting to analyse the sensitivity of this null line to mitomycin C, the archetypal DNA damaging agent for characterisation of FA factors (Jin and Cho, 2017).

This clear role for FAN1 in trypanosomes has further developed our understanding on how damaged DNA is processed for the BRCA2-RAD51 complex. Previous work on the role of FAN1 in ICL repair has shown that upon recruitment by FANCD2, this nuclease cleaves around the damaged DNA and acts as a replication fork recovery factor. Mammalian cells absent for FANCD2 cause MRE11-dependent access for FAN1 at a stalled replication fork to degrade damaged DNA strands (Indrajit Chaudhury, Stroik and Sobeck, 2014). Given that no trypanosomal homologue of FANCD2 has been identified and TbFAN1 and TbMRE11 function epistatically, a similar FANCD2-independent incision event may occur in trypanosomes. Upon stalling of the replication fork, it is hypothesised that these nucleases (TbMRE11 and TbFAN1) carry out extensive nucleolytic degradation of the DNA strands to create substrates which RAD51 is able to bind, thus promoting HR repair. Alongside this, they function in creating a new replication origin for fork restart to help promote DNA replication. New replication fork origins are needed to repair DNA without affecting the speed of DNA replication (Stanojcic et al., 2016). As FAN1 homologues incise DNA three or four nucleotides near the ICL, transversing the crosslink, it is predicted that FAN1 can unhook the ICL and incise flapped DNA surrounding the ICL which may explain how ICL-induced DSBs are formed in trypanosomes (Pizzolato et al., 2015).

8.1.2.5 The baffling function of TbCHL1

TbCHL1, a nuclear ATP-dependent helicase (Dean, Sunter and Wheeler, 2017), has been implicated as a trypanosomal ICL repair factor and taken together with information from literature, TbCHL1 is hypothesised to function in the SNM1-independent (Brosh and Cantor, 2014). As seen with TbFAN1-deficient cells, Tbchl1 null lines display no altered susceptibility to phleomycin, MMS, UV or hydroxyurea with additional work needed to look at susceptibility to mitomycin C to confirm a role in the FA pathway.

Analysis into the role of TbCHL1 in trypanosomal ICL repair has revealed some interesting findings with dose response curves against mechlorethamine demonstrating a regulatory role for TbCHL1 in ICL repair. It is hypothesised that TbCHL1 is acting as a co-regulator for the
activities of ICL repair factors with these factors acting to compensate for the activity of TbSNM1. Absence of both TbSNM1 and TbCHL1 releases these activities, restoring the cells ability to repair ICLs. Typically, FANCJ acts as a unique helicase which unwinds 5’ssDNA forked duplex substrates which contains at least 15 nucleotides (Wu and Brosh, 2009). In yeast cells, FANCJ plays a role in the Pso2-independent ICL repair pathway, functioning alongside the Mph1-Mgm101-MutSα complex for the activation and recruitment of Exo1 to the damage site (Ward et al., 2012). This contrasts observations made in this thesis as a classical, non-epistatic interaction would have been predicted if TbSNM1 and TbCHL1 functioned in independent ICL repair systems, as seen in yeast. Further studies are needed to explore the role of TbCHL1 in detail.

8.1.3 Summary

Our data and the work carried out on yeast and mammalian cells suggests that there is more than one system for the repair of ICLs, dependent on the temporal and spatial nature of ICL detection, with each system differing in relation to enzymes employed to the site of damage (Shen and Li, 2010). Whether the SNM1-independent or -dependent system is the central ICL repair mechanism is yet to be evaluated: ICL repair studies carried out on mammalian cells indicates that the FA, replication-coupled system is the dominant ICL repair mechanism. Given the multigenic nature of transcription in this parasite, we hypothesise the SNM1-dependent ICL repair pathway is the preferred repair system as regulation at a transcriptional level is of great importance in the parasite. It is tempting to speculate the fold differences observed between wild type and null lines reflects the relative importance of the ICL repair protein being tested however it is more likely that the variation observed here reflects the role of that enzyme in repair associated with transcription or replication: an asynchronous culture of trypanosomes contains more transcriptionally active cells than dividing cells. A summary of both trypanosomal ICL repair systems is shown in Figure 8.1.
8.2 TbSNM1 structure and function

Initial *in silico* analyses carried out on TbSNM1 were extended to understand the proteins 1. catalytic activity, 2. substrate specificity and 3. nuclear localisation (Sullivan *et al.*, 2015). This work was carried out by ectopic, inducible expression of either wild type or mutated GFP-tagged TbSNM1 in a genetic background absent for TbSNM1 which has proven to be an excellent tool to develop a clear understanding of which residues play a key role in this enzymes activity. Results from this analysis have demonstrated a role for residues D40, H80, D82 and H159 in TbSNM1 catalytic activity, residues R429 and R430 in protein localisation and a possible structural role for the D221 residue (Table 8.2). Structural consequences of mutating TbSNM1 were analysed *in silico* using a predicted TbSNM1 model created from both human SNM1 proteins as the structural templates. Ideally the residues hypothesised to be responsible for the coordination of the zinc metal ion in TbSNM1 would have been confirmed biochemically (e.g. using isothermal titration calorimetry) but given time constraints and lack of appropriate resources, this was not carried out. This model predicted TbSNM1 to bind one
zinc ion, possibly in an octahedral arrangement, similar to that observed by HaSNM1A, which may indicate that TbSNM1 solely functions in ICL repair, as seen with HaSNM1A (Sengerová et al., 2012).

<table>
<thead>
<tr>
<th>Putative function</th>
<th>TbSNM1 Mutation</th>
<th>Localisation</th>
<th>Complementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-factor binding</td>
<td>D40A</td>
<td>Nuclear</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>H78A</td>
<td>Nuclear</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>H80A</td>
<td>Nuclear</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>D82A</td>
<td>Nuclear</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>H83A</td>
<td>Nuclear</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>H159A</td>
<td>Nuclear</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>H178A</td>
<td>Nuclear</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>D221A</td>
<td>Non-specific</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>H496A</td>
<td>Nuclear</td>
<td>Yes</td>
</tr>
<tr>
<td>Localisation</td>
<td>R428A</td>
<td>Nuclear</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>R429A</td>
<td>Non-specific</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>R430A</td>
<td>Non-specific</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>H431A</td>
<td>Nuclear</td>
<td>Yes</td>
</tr>
<tr>
<td>Substrate specificity</td>
<td>V518A</td>
<td>Nuclear</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 8.2: Summary of TbSNM1 mutational analysis.

8.2.1 Key residues in TbSNM1 nuclease activity

Four key residues have been implicated in TbSNM1 nuclease activity with one of these residues being the arginine at position 40. The analogous residue in HsSNM1B, D14, has an important structural role for formation of hydrogen bonds to coordinate H36 (motif 2) and H99 (motif 3) residues, with both of these motifs having a direct role in enzyme catalytic activity (Allerston et al., 2015). Molecular modelling places the D40 residue in close proximity to residues of motif 2 (78-HxHxDH-83), suggesting that this residue may be playing a critical role in the coordination of this motif to the zinc ion as opposed to a direct role in zinc binding.

Furthermore, it was identified that residues H80 and D82 of the ‘HxHxDH’ motif play an important role in TbSNM1 catalytic activity with both residues in close proximity to the zinc metal ion. Given all residues of this motif in HsSNM1A play a role in coordination of the metal ion, it was predicted that mutation of H78 and H83 would have also displayed a similar phenotype as mutation of H80 and D82 residues (Hazrati et al., 2008). It is likely that the exact activity of these residues can only be revealed by biochemical analysis of the TbSNM1 structure (Sengerová et al., 2012).
Studies on the H159 residue of motif 3 have shown a key role in protein enzyme function. In HaSNM1A, this residue also plays a critical role in enzyme function to coordinate the third bond with the zinc ion, with the other two bonds coordinated by motifs 1 and 2, alongside motif 4 (Allerston et al., 2015). It is hypothesised that the residue H159 also functions in a similar manner as molecular modelling shows a clear interaction of this motif with all other TbSNM1 motifs. The conserved residue of motif 4, H178, did not display any evidence for a role in TbSNM1 catalytic activity, despite this motif playing a clear role in HaSNM1A activity. Furthermore, residue H496 did not show any indication for a role in TbSNM1 nuclease activity with the respective conserved residue in HsSNM1A (H994) also not displaying any evidence for a role in nucleotidase hydrolysis. However, the corresponding residue in PSO2 (H611) does show evidence for mediating the catalytic reaction with damaged DNA (Sengerová et al., 2012) which confirms that TbSNM1 functions in a similar manner to HsSNM1A and not HsSNM1B or PSO2. This data demonstrates that conserved residues found in motifs 1, 2 and 3 all play a key role in TbSNM1 nuclease activity which differs from other SNM1 homologs where all 4 motifs have catalytic functions. It is hypothesised that this histidine-rich coordination of the zinc metal ion aids redistribution of electrons around the aromatic imidazole ring of the amino acid to carry out the nucleophilic attack on DNA.

Further to this, it has been hypothesised that the D221 residue of motif 4 plays an important structural role in TbSNM1. Motif 4 in other SNM1 homologs contributes to the bridge between zinc ions and residues (Allerston et al., 2015), playing a critical process in maintaining this structure. This could explain why mutation of D221 into alanine resulted in a loss of TbSNM1 structural integrity; this residue plays a key role in maintaining this structure, potentially aiding the connection between the zinc binding pocket and the rest of the protein. Inhibitors of this residue could prove to be great therapeutic targets against TbSNM1.

8.2.2 Evaluating TbSNM1 substrate specificity

TbSNM1 substrate specificity has been associated with a conserved valine residue at position 518 with mutation of this residue displaying no evidence for altered TbSNM1 activity. This result has also been observed with HsSNM1A when mutation of the equivalent residue to either alanine or histidine maintained the nucleases enzymatic activity (Sengerová et al., 2012). It is plausible that many residues confer substrate specificity in the SNM1 family of proteins. Residues E213 and H254 in HsSNM1A have been hypothesised to aid this DNA specificity
which suggests that a triple mutation of all three residues may alter the enzymes ability to bind DNA efficiently (de Villartay et al., 2009). Although equivalent and conserved residues to E213 and H254 have not been identified in TbSNM1, further work on identifying residues responsible for TbSNM1 DNA binding is important to gain a full understanding of the proteins activity.

8.2.3 Dissecting the TbSNM1 nuclear localisation signal

This study has implicated two key residues, R429 and R430, in TbSNM1 nuclear localisation with no analogous role for R428 and R431 residues seen. With no comparable work carried out in mammalian and yeast homologues of this nuclease, it is hard to determine whether this trait is conserved within other organisms. It is tempting to speculate that residues R428 and R431 in TbSNM1 are required for protein localisation given that all four amino acids interact directly on the model of TbSNM1 and are all exposed on the surface of the protein, an attribute which is vital when targeting proteins to the nucleus. To test this, mutation of multiple residues may be required. Regardless, this data could be exploited for the development of inhibitors against TbSNM1.

8.2.4 Summary

The results observed here merit further investigation into the structural properties of TbSNM1. These same cell lines can be used to carry out pull-down assays to identify if these mutations affect the physical interacting partners of TbSNM1. The plasmids used for this assay can also be used for protein purification and in vitro activity assays to develop our knowledge on this structure-function relationship. Studies could also be broadened out to understand how mutation of two residues affects TbSNM1 catalytic activity. It would be interesting to undertake a biochemical analysis of TbSNM1 to analyse the enzymes binding specificity towards single stranded DNA and double stranded DNA. This could reveal how TbSNM1 hydrolyses DNA and whether it acts in a similar manner to HsSNM1A where small activity towards RNA is also seen. Structural, biophysical and fluorometric assays could be used to understand the processing power of TbSNM1 in relation to its efficiency in digesting past the ICL, as seen with HsSNM1A (Sengerová et al., 2012). This high processing power of HsSNM1A has been related to residues K904-K906 with mutations of these residues reducing the processing power of the enzyme but not affecting the enzymes catalytic activity (Allerston et al., 2015). This mutation resulted in a reduced capacity to digest substrate DNA with
premature termination of digestion. These lysine residues are semi-conserved in TbSNM1 with arginine residues present in the trypanosomal sequence, both of which are positively charged amino acids. Given that mutation of these residues does not affect the enzymes nuclease activity, studies carried out in this study cannot be used to evaluate the role of these residues with biochemical studies needed to dissect the function of these conserved arginine residues.

8.3 Cytotoxic effects of trypanocidal compounds

The cytotoxic effects towards DNA of clinically used and novel trypanocidal agents were evaluated to reveal that nifurtimox does not cause the formation of any DNA damage on the parasites genome, benznidazole promotes DNA damage which is repaired by the HR pathway (Campos et al., 2017), LH34 promotes ICL formation and CB1954 promotes both adduct and ICL formation in the parasite. The damage incurred by benznidazole involves prodrug activation which creates electrophilic drug metabolites which can lead to oxidative stress and adduct formation. Nitroreduction of CB1954 can result in the production of 4-hydroxylamine which can be processed to act as an electrophilic intermediate, alkylating cellular nucleophiles such as DNA (Meredith et al., 2017). This nitroreduction event can also result in the formation of 2-nitroso and 2-amine metabolites, both of which are able to alklyate DNA to create monoadducts (Mitchell and Minchin, 2008). Arizidinyl compounds vary in their structure which is reflected in the type of lesions formed on DNA. This could explain why LH34 only displays evidence for promotion of ICLs and no other adducts. This can help broaden our understanding of how arizidinyl-containing compounds function intracellularly, potentially extending the number of compounds which can be used in anti-tumour therapy. Many ICL-inducing agents are exploited to target cancer, with arizidine-containing compounds no exception.

8.4 An unexpected journey

TbHEL1, initially presumed to be a homologue of FANCM, displayed no evidence for a role in DNA repair, despite in silico analyses indicating a typical FANCM-like conserved protein domain and sequence structure. Given this, it was clear why trypanosomes are also absent for some ancillary FA factors (e.g. FAAP100, FAAP24, FAAP16 and FAAP20) as these proteins aid in FANCM localisation to chromatin and recognition of ICLs. Further studies on this enzyme has indicated a key role in trypanosomal RNA interference, possibly acting as the
missing helicase from DICER (Patrick et al., 2009). It is postulated that TbHEL1 directly interacts with TbDCL1 and TbDCL2 for the cleavage of double stranded RNA (dsRNA) molecules to produce siRNA molecules which are degraded by the RISC complex (Hammond, 2005). TbHEL1 could operate to determine which RNA molecules undergo dicer-mediated cleavage and process hairpin precursor microRNAs. Given that most DICER proteins possess both incision and helicase activities, this arrangement is rather strange but has been noted in other species e.g. Sporotrichum thermophile where dicer related helicases operate in a separate system to the DICER enzyme (Kidwell, Chan and Doudna, 2014). Other protozoan parasites are also able to undergo RNA interference but by different mechanisms which indicates that this silencing event in eukaryotes occurred at a very early point in evolution. Gene silencing in Plasmodium and Giardia species is hypothesised to operate in a classical RNA interference system, as seen in mammalian cells, with other trypanosomal RNA interference mechanisms similar to the above, multifactorial system (Ullu, Tschudi and Chakraborty, 2004). It is plausible that this complex RNA interference pathway requires the aid of additional helicases and ancillary proteins which are yet to be identified and studied in this parasite. It would be interesting to analyse the interaction of dsRNA and TbHEL1, including conserved residues responsible for this interaction and how the length of the dsRNA molecule affects the helicase activity of TbHEL1.
9. Future work

This is the first comprehensive study to look at how trypanosomes repair ICLs with many unanswered questions still remaining. Which other DNA repair factors are employed to remove ICLs? Are there any additional ICL repair systems employed by the parasite? Does *T. brucei* possess additional ICL repair specific factors currently not identified in prokaryotes or eukaryotes? How do these results compare with ICL repair systems in other protozoan parasites e.g. *T. cruzi* or *Leishmania*?

The next step for this project is confirming the dominant ICL repair system, particularly in relation to the different stages of the cell cycle. This could be carried out by stalling null lines at distinct stages of the cell cycle with further evaluation of their response to mechlorethamine (Nair et al., 2014). Evaluating whether these ICL repair systems are coupled to stalling of the transcription/replication machinery can be tested using inhibitors of RNA/DNA polymerases to evaluate how this affects sensitivity of null lines to ICL inducing agents. Furthermore, SMARD can be used to look at replication changes, both rate and direction, in mechlorethamine treated null cells (Calderano et al., 2015) with the hypothesis that any factors involved in replication-coupled ICL repair will display a prolonged rate of replication due to the lack of ICL repair. A similar study can be used to study transcription-coupled ICL repair factors with RT-qPCR carried out on a single polycistron to look at transcriptional changes upon treatment with mechlorethamine.

Identification of other trypanosomal ICL repair factors should focus on understanding how the initial unhooking event occurs in the parasite. A similar genetics-based approach (either using plasmid-based knockouts or the CRISPR-Cas9 system) coupled with susceptibility screening can be used to look at the ICL repair activities of TbXPG (Tb927.9.11760), TbERCC1 (Tb927.7.2060), TbXPF (Tb927.5.3670), TbMUS81 (Tb927.8.6740), TbSLX1 (Tb927.3.1220) with EME1 and SLX4 absent from the trypanosome genome (Genois et al., 2014; Machado et al., 2014). TbXPG has been studied for a role in the TC-NER pathway and is hypothesised to carry out these nuclease activities in the SNM1-dependent repair pathway (Machado et al., 2014). The activity of TbXPF was not assessed in the aforementioned study but given its activity in mammalian cells, it can be hypothesised to function alongside TbERCC1 in the SNM1-dependent repair pathway. The endonuclease activities of both these nucleases in trypanosomes can be tested by incubation of purified proteins with crosslinked,
circular DNA with cleavage activities confirmed by linearization of DNA (Wood, 2011; Klein Douwel et al., 2017). TbMUS81 has been functionally characterised by the McCulloch laboratory in relation to its role in recombination repair (Devlin, 2015) with its role in ICL repair currently unstudied. Given the interacting partner of TbMUS81 is absent, EME1, it is highly unlikely that this complex plays a role in removal of ICL lesions.

Identifying the role of the remaining FA factors identified in the trypanosome genome (FANCO, FANCQ, FANCV and USP1) would complete this analysis into the role of the trypanosomal FA pathway. FANCQ (XPF) is hypothesised to play a nuclease function for the unhooking of damaged DNA. FANCO, RAD51-C, has been studied for its auxiliary role in the HR pathway and may function alongside TbRAD51 and TbBRCA2 for strand invasion (Dobson et al., 2011). FANCV, MAD2, functions as a cell cycle checkpoint protein in spindle assembly during the metaphase-anaphase transition and has been hypothesised to play a role in the SNM1-independent ICL repair pathway for regulation of the cell cycle (Actis et al., 2016). Finally, USP1 functions in the deubiquitination of FANCD2 with several potential homologues of USP1 identified in the parasite genome (Kim and Kim, 2016). With the absence of FANDC2 in the parasite, this hydrolase may function in other regulatory mechanisms e.g. the post-translational modification of TbBRCA2 for protein activation (Schoenfeld et al., 2004). Studies identifying new ICL repair factors should also be carried out using a genome wide RNAi screen, RIT-seq (RNAi target sequencing) screen with mechlorethamine as the selective agent of choice (Stortz et al., 2017). This would reveal any novel DNA repair proteins involved in the trypanosomal ICL repair system as well as the ICL damage signalling pathways.

An extension of the above-mentioned biochemical analysis for TbSNM1 will reveal any potential targetable sites for the development of inhibitors of this protein. This can involve carrying out pull down assays on the ectopic expression of an amino terminally GFP tagged version of the enzyme to evaluate the physically interacting partners of TbSNM1, potentially identifying trypanosome-specific ICL repair factors. Further to this, purification of TbSNM1 and crystallisation studies would help advance the structure and function analysis shown in this thesis. This work would confirm the exo- and endonuclease activity of TbSNM1, substrate specificity, and interactions with the zinc co-factor. With components of the trypanosomal ICL repair systems identified, it would be interesting to develop the understanding of their helicase/nuclease activities using in vitro enzyme activity assays carried out on purified proteins and plasmid DNA containing crosslinked DNA.
**Thesis summary**

1. *T. brucei* possesses at least two distinct ICL repair systems
   a. An SNM1-dependent ICL repair system where TbSNM1 co-operates with TC-NER (TbCSB) and MMR (TbEXO1) enzymes to repair the encountered damaged, possibly triggered as a result of RNA polymerase stalling
   b. An SNM1-independent ICL repair systems which utilises components of the HR (TbMRE11 and potentially also TbRAD51 and TbBRCA2), MMR (TbEXO1) and FA (TbFAN1) pathways to repair the encountered damaged, possibly triggered as a result of DNA polymerase stalling
2. TbHEL1 plays a crucial role in the *T. brucei* RNA interference machinery, potentially functioning as the RNA helicase that is absent from the *T. brucei* DICER sequence but present in the counterparts expressed by higher eukaryotes
3. TbCHL1 (a FANCJ homologue) appears to play an ancillary role in the *T. brucei* ICL REPAIRtoire and may have a regulatory function
4. The TbREV2 and TbREV3 TLS enzymes appear to play no role in ICL repair suggesting that other TLS factors may function in trypanosomal ICL repair systems
5. D40, H80, D82, and H159 residues play a key role in TbSNM1 activity, with comparative studies and *in silico* modelling suggesting they may be involved in zinc co-factor binding
6. R429 and R430, but not R428 and H431, residues play a crucial role in the localisation of the protein to the nucleus
7. Activation of benznidazole results in cytotoxic moities that promote a form of DNA damage that can be resolved using components of the HR pathway
8. The nitrogen mustard LH34 promotes formation of ICLs that are resolved using SNM1-dependent and SNM1-independent repair systems
9. The aziridinyl agent CB1954 promotes formation of ICLs that are resolved using the SNM1-dependent and SNM1-independent repair systems and additional DNA lesions that are repaired by components of the NER pathway
were resuspended in 1mL of sterile water, roughly equating to 50
All primers used in this study are listed and were sourced from Sigma Aldrich. All primers
Appendix 1: Starting plasmids used in this study.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pKO-TcFHM-HYG</td>
<td>(de Padua et al., 2017)</td>
</tr>
<tr>
<td>pKO-TcFHM -NEO</td>
<td>(de Padua et al., 2017)</td>
</tr>
<tr>
<td>pKO-TcFHM -BLA</td>
<td>(de Padua et al., 2017)</td>
</tr>
<tr>
<td>pKO-TcFHM -PAC</td>
<td>(de Padua et al., 2017)</td>
</tr>
<tr>
<td>pKO-TbSNM1-BLA</td>
<td>(Sullivan et al., 2015)</td>
</tr>
<tr>
<td>pKO-TbSNM1-PAC</td>
<td>(Sullivan et al., 2015)</td>
</tr>
<tr>
<td>pRPa-GFP-TbSNM1</td>
<td>(Sullivan et al., 2015)</td>
</tr>
</tbody>
</table>

Appendix 1a: Starting plasmids used in this study.

All primers used in this study are listed and were sourced from Sigma Aldrich. All primers were resuspended in 1mL of sterile water, roughly equating to 50-100 pmol µl⁻¹.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tbexo1-KO1</td>
<td>5’-ggggagctccgctggtATGGGATTTAAGGGTTATGG-3’</td>
</tr>
<tr>
<td>Tbexo1-KO2</td>
<td>5’-gggtctagaCTCCGTATTACCGCTGCCAC-3’</td>
</tr>
<tr>
<td>Tbexo1-KO3</td>
<td>5’-aaagggcccGCTGACGCTGCCACGCTGCAA-3’</td>
</tr>
<tr>
<td>Tbexo1-KO4</td>
<td>5’-aaaggtaccCTACCGTCGCTGATAAGATAA-3’</td>
</tr>
<tr>
<td>Tbmre11-KO1</td>
<td>5’-gggaggtccggtgTCAAGTTCTTTGTGACCTTC-3’</td>
</tr>
<tr>
<td>Tbmre11-KO2</td>
<td>5’-gggtctagaATTGTTTGTAATGCCCACACT-3’</td>
</tr>
<tr>
<td>Tbmre11-KO3</td>
<td>5’-aaagggcccTGGCAAGCTTTGGAATTATG-3’</td>
</tr>
<tr>
<td>Tbcsh-KO1</td>
<td>5’-ggggagctccggtgTCATCCGTGCAACTCCAC-3’</td>
</tr>
<tr>
<td>Tbcsh-KO2</td>
<td>5’-gggtctagaCAACTTAACTACGATAATGACCTAA-3’</td>
</tr>
<tr>
<td>Tbcsh-KO3</td>
<td>5’-aaagggcccGATGGATCGTAAAGTGACT -3’</td>
</tr>
<tr>
<td>Tbcsh-KO4</td>
<td>5’-aaaggtaccCAATTCTATAGCTAACACAG -3’</td>
</tr>
<tr>
<td>Tbcshl-KO1</td>
<td>5’-ggggagctccggtgCATATCCGTGCAACTCCAC-3’</td>
</tr>
<tr>
<td>Tbcshl-KO2</td>
<td>5’-gggtctagaCAACTTAACTACGATAATGACCTAA-3’</td>
</tr>
<tr>
<td>Tbcshl-KO3</td>
<td>5’-aaagggcccGATGGATCGTAAAGTGACT -3’</td>
</tr>
<tr>
<td>Tbcshl-KO4</td>
<td>5’-aaaggtaccCAATTCTATAGCTAACACAG -3’</td>
</tr>
<tr>
<td>Tbhel1-KO1</td>
<td>5’-ggggagctccggtgCTCCTTCCTTACTCTAGTCA-3’</td>
</tr>
<tr>
<td>Tbhel1-KO2</td>
<td>5’-gggtctagaTATACATAAGATATGTCCCCG-3’</td>
</tr>
<tr>
<td>Tbhel1-KO3</td>
<td>5’-aaagggcccTTATCGTATCCTGGACTGCAAA-3’</td>
</tr>
<tr>
<td>Tbhel1-KO4</td>
<td>5’-aaaggtaccTAAGACATACAGCAGCAGGAC-3’</td>
</tr>
<tr>
<td>Tbfan1-KO1</td>
<td>5’-aaagagctccggtgATGCAAGGTGACGTCTGCCCT -3’</td>
</tr>
<tr>
<td>Tbfan1-KO2</td>
<td>5’-aaatctagaGCCGAAAGATCAGCCTT ACCAT -3’</td>
</tr>
<tr>
<td>Tbfan1-KO3</td>
<td>5’-aaagggcccTTTGAGCCTGAGGTGTTTC -3’</td>
</tr>
<tr>
<td>Tbfan1-KO4</td>
<td>5’-aaaggtaccTTATACAGGAGGTGCAGTTCC -3’</td>
</tr>
<tr>
<td>Tbrev2-KO1</td>
<td>5’-aaagagctccggtgATGCAAGGTGACGTCTGCCCT -3’</td>
</tr>
<tr>
<td>Tbrev2-KO2</td>
<td>5’-aaatctagaGTCAGAATCAACAGCAGCAGAAG-3’</td>
</tr>
<tr>
<td>Tbrev2-KO3</td>
<td>5’-aaagggcccTTTCAGAAGAATCAGCAGCAGAAG-3’</td>
</tr>
<tr>
<td>Tbrev2-KO4</td>
<td>5’-aaaggtaccTCAGGAGGTGCTTACGAGATGCAGTC-3’</td>
</tr>
<tr>
<td>Tbrev3-KO1</td>
<td>5’-aaagagctccggtgATGCAAGGTGACGTCTGCCCT -3’</td>
</tr>
<tr>
<td>Tbrev3-KO2</td>
<td>5’-aaatctagaGTCAGAATCAACAGCAGCAGAAG-3’</td>
</tr>
<tr>
<td>Tbrev3-KO3</td>
<td>5’-aaagggcccATTGTTGCTTACGAGGTCAGAG-3’</td>
</tr>
<tr>
<td>Tbrev3-KO4</td>
<td>5’-aaaggtaccTCAGGAGGTGCTTACGAGATGCAGTC-3’</td>
</tr>
<tr>
<td>NEO-2</td>
<td>5’-CTAGAAGAATCCGTAAGAAG-3’</td>
</tr>
</tbody>
</table>
HYG-2 5’-CTATTCCCTTTGCCCTCGGACG-3’
BLA1 5’-CTGGCTAATGTAAAGGTTCAT-3’
PAC1 5’-CGGGCTTGGGTGATGCATCA-3’
GFP 5’-GGCATGGAGAGCTGTCAGAAG-3’
Tbsnm1-KO1 5’-aaagagctccgccggGCAAGTTGGAAGCTGAGTAAAG-3’
Tbsnm1-seq1 5’-CACCAGATCACCTGATCCCGAG-3’
Tbsnm1-q1 5’-AGATGCTAAACAAGAAGAGTC-3’
Tbsno1-q1 5’-TCGGTGCCGAGTCTTCCGG-3’
Tbsno1-q2 5’-CAACGCACGTTTCACCGGTC-3’
Tbmc11-q1 5’-GGGTTTGCACATCCATCCAGC-3’
Tbmc11-q2 5’-CCGGACGACAGACGAATAC-3’
Tbcsb-q1 5’-GGTCTAGCAGGTTAGAAGCT-3’
Tbcsb-q2 5’-CGCATACCTCTTGTCTTCTCTT-3’
Tbc1l1-q1 5’-TATGTACATGTGCGCCGGGA-3’
Tbc1l1-q2 5’-TTTCTGTTTCGCGCTGCTT-3’
Tbh1l1-q1 5’-TTTGCTTGTTCCCGGCTGCTT-3’
Tbh1l1-q2 5’-TTTGCTTGCTCCGAGGCTTCT-3’
Tbfan1-q1 5’-GTGCGACCTTCTTGTCTCTTCC-3’
Tbfan1-q2 5’-ATTCCAACACAAATATACAGC-3’
Tbrev2-q1 5’-TGGGGCCAGAAGACAGTCTTCCAAA-3’
Tbrev2-q2 5’-AAAGGAAGGCGGTTGGAAGAAGA-3’
Tbrev3-q1 5’-CGTGCGGCTTACACAAACC-3’
Tbrev3-q2 5’-CAGACTGGGAGTCTTGAGTCTG-3’
Tbtert-F 5’-GAGCGTGACTTCCGAGGGA-3’
Tbtert-R 5’-AGGAACACTGTCACGGAGT-3’

Appendix 1b: Primer names and sequences used in this study. Lowercase, italicised sequences correspond to integrated restriction sites.
Appendix 2: Validation of single null mutant lines

Appendix 2a: Generation of *T. brucei* line lacking TbMRE11.

A. Representation of the *Tbmre11* gene and disrupted alleles using genes encoding for neomycin (*neo*) and hygromycin (*hyg*) resistance. P1 (Tbmre11-KO1), P2 (Tbmre11-q2), P3 (neo-2) and P4 (hyg-2) correspond to the regions where the primers anneal too in the appropriate allele.

B. Amplified products after multiplex PCR on gDNA extracted from *T. brucei* wild type (lane 1) and *T. brucei* *Tbmre11Δ* (lane 2) using primer combinations that specifically amplify alleles corresponding to the *Tbmre11* (Tbmre11-KO1/Tbmre11-q2), *Tbmre11-hyg* (Tbmre11-KO1/hyg-2) or *Tbmre11-neo* (Tbmre11-KO1/neo-2) alleles.

C. Amplified products obtained after PCR on cDNA derived from RNA extracted from *T. brucei* wild type (lane 1) or *T. brucei* *Tbmre11Δ* (lane 2) using primer combinations that specifically amplify intact *Tbmre11* (Tbmre11-q1/Tbmre11-q2). In both cases, the integrity of the cDNA (and hence RNA) was evaluated by amplification of a control fragment, *Tbtert*. 
Appendix 2b: Generation of *T. brucei* line lacking TbEXO1.

A. Diagram of the Tbexo1 gene and disrupted alleles using genes encoding for neomycin (*neo*) and hygromycin (*hyg*) resistance. P1 (Tbexo1-KO1), P2 (Tbexo1-q2), P3 (neo-2) and P4 (hyg-2) correspond to the regions where the primers anneal too in the appropriate allele.

B. DNA products created after amplification reactions carried out on gDNA extracted from *T. brucei* wild type (lane 1) and *T. brucei* Tbexo1Δ (lane 2) using primer combinations that specifically amplify alleles corresponding to the Tbexo1 (Tbexo1-KO1/Tbexo1-q2), Tbexo1-hyg (Tbexo1-KO1/hyg-2) or Tbexo1-neo (Tbexo1-KO1/neo-2) alleles.

C. DNA products amplified by PCR on cDNA derived from RNA extracted from *T. brucei* wild type (lane 1) or *T. brucei* Tbexo1Δ (lane 2) using primer combinations that specifically amplify intact Tbexo1 (Tbexo1-q1/Tbexo1-q2). In both cases, the integrity of the cDNA (and hence RNA) was evaluated by amplification of a control fragment, Tbtett.
Appendix 2c: Generation of *T. brucei* line lacking TbREV2.

A. Diagram of the *Tbrev2* gene and disrupted alleles using genes encoding for neomycin (*neo*) and hygromycin (*hyg*) resistance. P1 (Tbrev2-KO1), P2 (Tbrev2-q2), P3 (neo-2) and P4 (hyg-2) correspond to the regions where the primers anneal too in the appropriate allele.

B. DNA products created after amplification reactions carried out on gDNA extracted from *T. brucei* wild type (lane 1) and *T. brucei* Tbrev2Δ (lane 2) using primer combinations that specifically amplify alleles corresponding to the *Tbrev2* (Tbrev2-KO1/Tbrev2-q2), *Tbrev2-hyg* (Tbrev2-KO1/hyg-2) or *Tbrev2-neo* (Tbrev2-KO1/neo-2) alleles.

C. DNA products amplified by PCR on cDNA derived from RNA extracted from *T. brucei* wild type (lane 1) or *T. brucei* Tbrev2Δ (lane 2) using primer combinations that specifically amplify intact *Tbrev2* (Tbrev2-q1/Tbrev2-q2). In both cases, the integrity of the cDNA (and hence RNA) was evaluated by amplification of a control fragment, *Tbtert.*
Appendix 2d: Generation of *T. brucei* line lacking TbREV3.

A. Schematic representation of the *Tbrev3* allele and the effects of gene disruption using genes encoding for neomycin (*neo*) and hygromycin (*hyg*) resistance. P1 (*Tbrev3*-KO1), P2 (*Tbrev3*-q2), P3 (neo-2) and P4 (hyg-2) correspond to the regions where the primers anneal too in the appropriate allele.

B. DNA fragments obtained after amplification on gDNA extracted from *T. brucei* wild type (lane 1) and *T. brucei* *Tbrev3*Δ (lane 2) using primer combinations that specifically amplify biomarkers corresponding to the *Tbcsb* (*Tbrev3*-KO1/*Tbrev3*-q2), *Tbrev3*-hyg (*Tbrev3*-KO1/*hyg*-2) or *Tbrev3*-neo (*Tbrev3*-KO1/neo-2) alleles.

C. DNA fragments obtained after amplification on cDNA derived from RNA extracted from *T. brucei* wild type (lane 1) or *T. brucei* *Tbrev3*Δ (lane 2) using primer combinations that specifically amplify intact *Tbrev3* (*Tbrev3*-q1/*Tbrev3*-q2). In both cases, the integrity of the cDNA (and hence RNA) was evaluated by amplification of a control fragment, *Tbtert.*
Appendix 3: Validation of double null lines

Appendix 3a: Disruption and confirmation of Tbsnm1Δ Tbmre11Δ.

A. Diagram illustrating the Tbsnm1 allele and resultant changes upon disruption of this loci using genes encoding for blasticidin (bla) and puromycin (pac) resistance. P1 (Tbsnm1-KO1), P2 (Tbsnm1-seq1), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.

B. Schematic for the representation of the Tbmre11 allele and effects of disrupting this allele with genes encoding for neomycin (neo) and hygromycin (hyg) resistance. P1 (Tbmre11-KO1), P2 (Tbmre11-q2), P3 (neo-2) and P4 (hyg-2) correspond to regions where the primers anneal at the specific allele.

C. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1), T. brucei Tbsnm1Δ (lane 2), T. brucei Tbmre11Δ (lane 3), and T. brucei Tbsnm1Δ Tbmre11Δ (lane 4). Extractions from all 4 lines were tested using primer combinations that specifically amplify biomarkers corresponding to the Tbsnm1 (Tbsnm1-KO1/Tbsnm1-seq1), Tbsnm1-bla (Tbsnm1-KO1/bla-1) or Tbsnm1-pac (Tbsnm1-KO1/pac-1), Tbmre11 (Tbmre11-KO1/Tbmre11-q2), Tbmre11-hyg (Tbmre11-KO1/hyg-2) or Tbmre11-neo (Tbmre11-KO1/neo-2) alleles.

D. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1), Tbsnm1Δ (lane 2), Tbmre11Δ (lane 3) and Tbsnm1Δ Tbmre11Δ (lane 4) using primer combinations that specifically amplify intact Tbsnm1 (Tbsnm1-q1/Tbsnm1-seq1), Tbmre11 (Tbmre11-q1/Tbmre11-q2) and Tbert (Tbert-F/Tbert-R).
Appendix 3b: Disruption and confirmation of Tbsnm1Δ Tbexo1Δ null line.
A. Diagram illustrating the Tbsnm1 allele and resultant changes upon disruption of this loci using genes encoding for blasticidin (bla) and puromycin (pac) resistance. P1 (Tbsnm1-KO1), P2 (Tbsnm1-seq1), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.
B. Schematic for the representation of the Tbexo1 allele and effects of disrupting this allele with genes encoding for neomycin (neo) and hygromycin (hyg) resistance. P1 (Tbxo1-KO1), P2 (Tbxo1-q2), P3 (neo-2) and P4 (hyg-2) correspond to regions where the primers anneal at the specific allele.
C. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1), T. brucei Tbsnm1Δ (lane 2), T. brucei Tbexo1Δ (lane 3), and T. brucei Tbsnm1Δ Tbexo1Δ (lane 4). Extractions from all 4 lines were tested using primer combinations that specifically amplify biomarkers corresponding to the Tbsnm1 (Tbsnm1-KO1/Tbsnm1-seq1), Tbsnm1-bla (Tbsnm1-KO1/bla-1) or Tbsnm1-pac (Tbsnm1-KO1/pac-1), Tbexo1 (Tbxo1-KO1/Tbxo1-q2), Tbxo1-hyg (Tbxo1-KO1/hyg-2) or Tbxo1-neo (Tbxo1-KO1/neo-2) alleles.
D. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1), Tbsnm1Δ (lane 2), Tbexo1Δ (lane 3) and Tbsnm1Δ Tbexo1Δ (lane 4) using primer combinations that specifically amplify intact Tbsnm1 (Tbsnm1-q1/Tbsnm1-seq1), Tbexo1 (Tbxo1-q1/Tbxo1-q2) and Tbtert (Tbtert-F/Tbtert-R).
Appendix 3c: Disruption and confirmation of Tbsnm1Δ Tbrev2Δ.

A. Diagram illustrating the Tbsnm1 allele and resultant changes upon disruption of this loci using genes encoding for blasticidin (bla) and puromycin (pac) resistance. P1 (Tbsnm1-KO1), P2 (Tbsnm1-seq1), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.

B. Schematic for the representation of the Tbrev2 allele and effects of disrupting this allele with genes encoding for neomycin (neo) and hygromycin (hyg) resistance. P1 (Tbrev2-KO1), P2 (Tbrev2-q2), P3 (neo-2) and P4 (hyg-2) correspond to regions where the primers anneal at the specific allele.

C. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1), T. brucei Tbsnm1Δ (lane 2), T. brucei Tbrev2Δ (lane 3), and T. brucei Tbsnm1Δ Tbrev2Δ (lane 4). Extractions from all 4 lines were tested using primer combinations that specifically amplify biomarkers corresponding to the Tbsnm1 (Tbsnm1-KO1/Tbsnm1-seq1), Tbsnm1-bla (Tbsnm1-KO1/bla-1) or Tbsnm1-pac (Tbsnm1-KO1/pac-1), Tbrev2 (Tbrev2-KO1/Tbrev2-q2), Tbrev2-hyg (Tbrev2-KO1/hyg-2) or Tbrev2-neo (Tbrev2-KO1/neo-2) alleles.

D. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1), Tbsnm1Δ (lane 2), Tbrev2Δ (lane 3) and Tbsnm1Δ Tbrev2Δ (lane 4) using primer combinations that specifically amplify intact Tbsnm1 (Tbsnm1-q1/Tbsnm1-seq1), Tbrev2 (Tbrev2-q1/Tbrev2-q2) and Tbtert (Tbtert-F/Tbtert-R).
Appendix 3d: Disruption and confirmation of Tbsnm1Δ Tbrev3Δ.

A. Diagram illustrating the Tbsnm1 allele and resultant changes upon disruption of this loci using genes encoding for blasticidin (bla) and puromycin (pac) resistance. P1 (Tbsnm1-KO1), P2 (Tbsnm1-seq1), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.

B. Schematic for the representation of the Tbrev3 allele and effects of disrupting this allele with genes encoding for neomycin (neo) and hygromycin (hyg) resistance. P1 (Tbrev3-KO1), P2 (Tbrev3-q2), P3 (neo-2) and P4 (hyg-2) correspond to regions where the primers anneal at the specific allele.

C. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1), T. brucei Tbsnm1Δ (lane 2), T. brucei Tbrev3Δ (lane 3), and T. brucei Tbsnm1Δ Tbrev3Δ (lane 4). Extractions from all 4 lines were tested using primer combinations that specifically amplify biomarkers corresponding to the Tbsnm1 (Tbsnm1-KO1/Tbsnm1-seq1), Tbsnm1-bla (Tbsnm1-KO1/bla-1) or Tbsnm1-pac (Tbsnm1-KO1/pac-1), Tbrev3 (Tbrev3-KO1/Tbrev3-q2), Tbrev3-hyg (Tbrev3-KO1/hyg-2) or Tbrev3-neo (Tbrev3-KO1/neo-2) alleles.

D. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1), Tbsnm1Δ (lane 2), Tbrev3Δ (lane 3) and Tbsnm1Δ Tbrev3Δ (lane 4) using primer combinations that specifically amplify intact Tbsnm1 (Tbsnm1-q1/Tbsnm1-seq1), Tbrev3 (Tbrev3-q1/Tbrev3-q2) and Tbert (Tbert-F/Tbert-R).
Appendix 3e: Disruption and confirmation of TbsbΔ Tbexo1Δ.

A. Schematic for the representation of the Tbsb allele and effects of disrupting this allele with genes encoding for neomycin (neo) and hygromycin (hyg) resistance. P1 (Tbsb-KO1), P2 (Tbsb-q2), P3 (neo-2) and P4 (hyg-2) correspond to regions where the primers anneal at the specific allele.

B. Schematic for the representation of the Tbexo1 allele and effects of disrupting this allele with genes encoding for blastidicin (bla) and puromycin (pac) resistance. P1 (Tbexo1-KO1), P2 (Tbexo1-q2), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.

C. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1), T. brucei TbsbΔ (lane 2), T. brucei Tbexo1Δ (lane 3), and T. brucei TbsbΔ Tbexo1Δ (lane 4). Extractions from all 4 lines were tested using primer combinations that specifically amplify biomarkers corresponding to Tbsb (Tbsb-KO1/Tbsb-q2), Tbsb-hyg (Tbsb-KO1/hyg-2) or Tbsb-neo (Tbsb-KO1.neo-2), Tbexo1 (Tbexo1-KO1/Tbexo1-q2), Tbexo1-pac (Tbexo1-KO1/pac-1) or Tbexo1-bla (Tbexo1-KO1/bla-1) alleles.

D. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1), TbsbΔ (lane 2), Tbexo1Δ (lane 3) and TbsbΔ Tbexo1Δ (lane 4) using primer combinations that specifically amplify intact Tbsb (Tbsb-q1/Tbsb-q2), Tbexo1 (Tbexo1-q1/Tbexo1-q2) and Tbert (Tbert-F/Tbert-R).
Appendix 3f: Disruption and confirmation of TbsbΔ Tbmre11Δ.

A. Schematic for the representation of the Tbsb allele and effects of disrupting this allele with genes encoding for neomycin (neo) and hygromycin (hyg) resistance. P1 (Tbsb-KO1), P2 (Tbsb-q2), P3 (neo-2) and P4 (hyg-2) correspond to regions where the primers anneal at the specific allele.

B. Schematic for the representation of the Tbmre11 allele and effects of disrupting this allele with genes encoding for blasticidin (bla) and puromycin (pac) resistance. P1 (Tbmre11-KO1), P2 (Tbmre11-q2), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.

C. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1), T. brucei TbsbΔ (lane 2) and T. brucei TbsbΔ Tbmre11Δ (lane 3). Extractions from all 3 lines were tested using primer combinations that specifically amplify biomarkers corresponding to Tbsb (Tbsb-KO1/Tbsb-q2), Tbsb-hyg (Tbsb-KO1/hyg-2) or Tbsb-neo (Tbsb-KO1/neo-2), Tbmre11 (Tbmre11-KO1/Tbmre11-q2), Tbmre11-pac (Tbmre11-KO1/pac-1) or Tbmre11-bla (Tbmre11-KO1/bla-1) alleles.

D. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1), TbsbΔ (lane 2) and TbsbΔ Tbmre11Δ (lane 3) using primer combinations that specifically amplify intact Tbsb (Tbsb-q1/Tbsb-q2), Tbmre11 (Tbmre11-q1/Tbmre11-q2) and Tbert (Tbert-F/Tbert-R).
Appendix 3g: Disruption and confirmation of Tbexo1Δ Tbmre11Δ.

A. Schematic for the representation of the Tbexo1 allele and effects of disrupting this allele with genes encoding for blastidicin (bla) and puromycin (pac) resistance. P1 (Tbexo1-KO1), P2 (Tbexo1-q2), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.

B. Schematic for the representation of the Tbmre11 allele and effects of disrupting this allele with genes encoding for neomycin (neo) and hygromycin (hyg) resistance. P1 (Tbmre11-KO1), P2 (Tbmre11-q2), P3 (neo-2) and P4 (hyg-2) correspond to regions where the primers anneal at the specific allele.

C. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1), T. brucei Tbexo1Δ (lane 2), T. brucei Tbmre11Δ (lane 3) and T. brucei Tbexo1Δ Tbmre11Δ (lane 4). Extractions from all 4 lines were tested using primer combinations that specifically amplify biomarkers corresponding to Tbexo1 (Tbexo1-KO1/Tbexo1-q2), Tbexo1-pac (Tbexo1-KO1/pac-1) or Tbexo1-bla (Tbexo1-KO1/bla-1), Tbmre11 (Tbmre11-KO1/Tbmre11-q2), Tbmre11-hyg (Tbmre11-KO1/hyg-2) or Tbmre11-neo (Tbmre11-KO1/neo-2) alleles.

D. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1), T. brucei Tbexo1Δ (lane 2), T. brucei Tbmre11Δ (lane 3) and T. brucei Tbexo1Δ Tbmre11Δ (lane 4) using primer combinations that specifically amplify intact Tbexo1 (Tbexo1-q1/Tbexo1-q2), Tbmre11 (Tbmre11-q1/Tbmre11-q2) and Tbtert (Tbtert-F/Tbtert-R).
Appendix 4: Validation of lines lacking a trypanosomol FA factor

Appendix 4a: Disruption and confirmation of Tbhel1Δ and Tbsnm1Δ Tbhel1Δ null line.

A. Diagram illustrating the Tbsnm1 allele and resultant changes upon disruption of this loci using genes encoding for blasticidin (bla) and puromycin (pac) resistance. P1 (Tbsnm1-KO1), P2 (Tbsnm1-seq1), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.

B. Schematic for the representation of the Tbhel1 allele and effects of disrupting this allele with genes encoding for neomycin (neo) and hygromycin (hyg) resistance. P1 (Tbhel1-KO1), P2 (Tbhel1-q2), P3 (neo-2) and P4 (hyg-2) correspond to regions where the primers anneal at the specific allele.

C. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1), T. brucei Tbsnm1Δ (lane 2), T. brucei Tbhel1Δ (lane 3), and T. brucei Tbsnm1Δ Tbhel1Δ (lane 4). Extractions from all 4 lines were tested using primer combinations that specifically amplify biomarkers corresponding to the Tbsnm1 (Tbsnm1-KO1/Tbsnm1-seq1), Tbsnm1-bla (Tbsnm1-KO1/bla-1) or Tbsnm1-pac (Tbsnm1-KO1/pac-1), Tbhel1 (Tbhel1-KO1/Tbhel1-q2), Tbhel1-hyg (Tbhel1-KO1/hyg-2) or Tbhel1-neo (Tbhel1-KO1/neo-2) alleles.

D. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1), Tbsnm1Δ (lane 2), Tbhel1Δ (lane 3) and Tbsnm1Δ Tbhel1Δ (lane 4) using primer combinations that specifically amplify intact Tbsnm1 (Tbsnm1-q1/Tbsnm1-seq1), Tbhel1 (Tbhel1-q1/Tbhel1-q2) and Tbert (Tbert-F/Tbert-R).
Appendix 4b: Disruption and confirmation of Tbchl1-deficient null lines.

A. Diagram illustrating the Tbsnm1 allele and resultant changes upon disruption of this loci using genes encoding for blasticidin (bla) and neomycin (neo) resistance. P1 (Tbsnm1-KO1), P2 (Tbsnm1-seq1), P3 (bla-1) and P4 (neo-2) correspond to regions where the primers anneal at the specific allele.

B. Schematic for the representation of the Tbchl1 allele and effects of disrupting this allele with genes encoding for puromycin (pac) and hygromycin (hyg) resistance. P1 (Tbchl1-KO1), P2 (Tbchl1-q2), P3 (pac-1) and P4 (hyg-2) correspond to regions where the primers anneal at the specific allele.

C. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1), T. brucei Tbchl1Δ (lane 2) and T. brucei Tbsnm1Δ Tbchl1Δ (lane 3). Extractions from all 4 lines were tested using primer combinations that specifically amplify biomarkers corresponding to the Tbsnm1 (Tbsnm1-KO1/Tbsnm1-seq1), Tbsnm1-bla (Tbsnm1-KO1/bla-1) or Tbsnm1-neo (Tbsnm1-KO1/neo-2), Tbchl1 (Tbchl1-KO1/Tbchl1-q2), Tbchl1-hyg (Tbchl1-KO1/hyg-2) or Tbchl1-pac (Tbchl1-KO1/pac-1) alleles.

D. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1), Tbchl1Δ (lane 2), Tbchl1Δ (lane 3) and Tbsnm1Δ Tbchl1Δ (lane 4) using primer combinations that specifically amplify intact Tbsnm1 (Tbsnm1-q1/Tbsnm1-seq1), Tbchl1 (Tbchl1-q1/Tbchl1-q2) and Tbert (Tbert-F/Tbert-R).
Appendix 4c: Confirmation of Tbfan1 null line.

A. Schematic representation of the Tbfan1 allele and the effects of gene disruption using genes encoding for neomycin (neo) and hygromycin (hyg) resistance. P1 (Tbfan1-KO1), P2 (Tbfan1-q2), P3 (neo-2) and P4 (hyg-2) correspond to the regions where the primers anneal too in the appropriate allele.

B. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1) and T. brucei Tbfan1Δ (lane 2) using primer combinations that specifically amplify biomarkers corresponding to the Tbfan1 (Tbfan1-KO1/Tbfan1-q2), Tbfan1-hyg (Tbfan1-KO1/hyg-2) or Tbfan1-neo (Tbfan1-KO1/neo-2) alleles.

C. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1) or T. brucei Tbfan1Δ (lane 2) using primer combinations that specifically amplify intact Tbfan1 (Tbfan1-q1/Tbfan1-q2). In both cases, the integrity of the cDNA (and hence RNA) was evaluated by amplification of a control fragment, Tbtert.
Figure 4d: Validating Tbfan1 double null lines.

A. Diagram illustrating the Tbfan1 allele and resultant changes upon disruption of this loci using genes encoding for neomycin (neo) and hygromycin (hyg) resistance. P1 (Tbfan1-KO1), P2 (Tbfan1-q2), P3 (neo-2) and P4 (hyg-2) correspond to regions where the primers anneal at the specific allele.

B. Schematic for the representation of the Tbmre11 allele and effects of disrupting this allele with genes encoding for blasticidin (bla) and puromycin (pac) resistance. P1 (Tbmre11-KO1), P2 (Tbmre11-q2), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.

C. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1), T. brucei Tbsnm1Δ (lane 2), T. brucei Tbfan1Δ (lane 3), and T. brucei Tbsnm1Δ Tbfan1Δ (lane 4). Extractions from all 4 lines were tested using primer combinations that specifically amplify biomarkers corresponding to Tbfan1 (Tbfan1-KO1/Tbfan1-q2), Tbfan1-hyg (Tbfan1-KO1/hyg-2) or Tbfan1-neo (Tbfan1-KO1/neo-2) alleles.

D. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1), Tbmsnm1Δ (lane 2), Tbfan1Δ (lane 3) and Tbsnm1Δ Tbfan1Δ (lane 4) using primer combinations that specifically amplify intact Tbmsnm1 (Tbmsnm1-q1/Tbmsnm1-seq1), Tbfan1 (Tbfan1-q1/Tbfan1-q2) and Tbtert (Tbtert-F/Tbtert-R).

E. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1), Tbmre11Δ (lane 2), Tbfan1Δ (lane 3) and Tbmre11Δ Tbfan1Δ (lane 4). Extractions from all 4 lines were tested using primer combinations that specifically amplify biomarkers corresponding to intact Tbmre11 (Tbmre11-KO1/Tbmre11-q2), Tbmre11-pac (Tbmre11-KO1/pac-1) or Tbmre11-bla (Tbmre11-KO1/bla-1) alleles.

F. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1), Tbmre11Δ (lane 2), Tbfan1Δ (lane 3) and Tbmre11Δ Tbfan1Δ (lane 4) using primer combinations that specifically amplify intact Tbfan1 (Tbfan1-q1/Tbfan1-q2), Tbmre11 (Tbmre11-q1/Tbmre11-q2) and Tbtert (Tbtert-F/Tbtert-R).
Appendix 5: Structure and function studies on TbSNM1

Appendix 4.1: Validating the \textit{Tb}snm1Δ \textit{gfp}-\textit{Tb}snm1-V518A line.

A. Diagram representing the \textit{Tb}snm1 allele and resultant changes upon disruption of this loci using genes encoding for blasticidin (\textit{bla}) and puromycin (\textit{pac}) resistance. Alongside this, changes to the landing pad locus after integration of \textit{gfp}-\textit{Tb}snm1 was also analysed. P1 (\textit{Tb}snm1-KO1), P2 (\textit{Tb}snm1-seq1), P3 (bla-1), P4 (pac-1) and P5 (GFP) correspond to regions where the primers anneal at the specific allele.

B. DNA fragments obtained from amplification reactions on genomic DNA extracted from \textit{T. brucei} wild type (lane 1), \textit{Tb}snm1\textsuperscript{+/-} (lane 2), \textit{Tb}snm1Δ (lane 3), \textit{Tb}snm1Δ \textit{gfp}-\textit{Tb}snm1 (lane 4), \textit{Tb}SNM1-V518A (lane 5) cells. Primer combinations amplified intact \textit{Tb}snm1 (\textit{Tb}snm1-KO1/\textit{Tb}snm1-seq1), \textit{Tb}snm1-\textit{bla} (\textit{Tb}snm1-KO1/bla-1), \textit{Tb}snm1-\textit{pac} (\textit{Tb}snm1-KO1/pac-1), or \textit{gfp}-\textit{Tb}snm1 (GFP/\textit{Tb}snm1-seq1) alleles.
Appendix 4.2: Confirming expression of gfp-Tbsnm1.
A. Schematic of integrated gfp-Tbsnm1 analysing using primer combination P2 (Tbsnm1-seq1) and P5 (GFP) which amplified a band of 0.4kb across the GFP and TbSNM1 fusion.
B. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1), Tbsnm1+/− (lane 2), Tbsnm1Δ (lane 3), Tbsnm1Δ gfp-Tbsnm1 (lane 4), TbSNM1-V518A (lane 5) cells using primer combinations that specifically amplify gfp-Tbsnm1 (GFP/Tbsnm1-seq1) and Tbert (Tbert-F/Tbert-R).

Appendix 4.3: Localisation of mutant GFP-TbSNM1 in bloodstream form T. brucei. Parasites expressing GFP-TbSNM1 (mutant or wild type) were stained with DAPI and examined by epifluorescent microscopy under a x40 objective lens with phase, DAPI and GFP signal images shown: Tbsnm1Δ gfp-Tbsnm1 (A) and TbSNM1-V518A (B).
Appendix 4.4: Validating conditional null lines.

A. Diagram representing the Tbsnm1 allele and resultant changes upon disruption of this loci using genes encoding for blasticidin (bla) and puromycin (pac) resistance. Alongside this, changes to the landing pad locus after integration of *gfp*-Tbsnm1 was also analysed. P1 (Tbsnm1-KO1), P2 (Tbsnm1-seq1), P3 (bla-1), P4 (pac-1) and P5 (GFP) correspond to regions where the primers anneal at the specific allele.

B. DNA fragments obtained from amplification reactions on genomic DNA extracted from *T. brucei* wild type (lane 1), Tbsnm1+/− (lane 2), Tbsnm1Δ (lane 3), Tbsnm1Δ gfp-Tbsnm1 (lane 4), TbSNM1-R428A (lane 5), TbSNM1-R429A (lane 6), TbSNM1-R430A (lane 7) and TbSNM1-H431A (lane 8) cells. Primer combinations amplified intact Tbsnm1 (Tbsnm1-KO1/Tbsnm1-seq1), Tbsnm1-bla (Tbsnm1-KO1/bla-1), Tbsnm1-pac (Tbsnm1-KO1/pac-1), or *gfp*-Tbsnm1 (GFP/Tbsnm1-seq1) alleles.
Appendix 4.5: Confirming expression of *gfp*-Tbsnm1.
A. Schematic of integrated *gfp*-Tbsnm1 analysing using primer combination P2 (Tbsnm1-seq1) and P5 (GFP) which amplified a band of 0.4kb across the GFP and Tbsnm1 fusion.
B. DNA fragments obtained after amplification on cDNA derived from RNA extracted from *T. brucei* wild type (lane 1), Tbsnm1+/− (lane 2), Tbsnm1Δ (lane 3), Tbsnm1Δ *gfp*-Tbsnm1 (lane 4), Tbsnm1-R428A (lane 5), Tbsnm1-R429A (lane 6), Tbsnm1-R430A (lane 7) and Tbsnm1-H431A (lane 8) cells using primer combinations that specifically amplify *gfp*-Tbsnm1 (GFP/Tbsnm1-seq1) and Tbert (Tbert-F/Tbert-R).
Appendix 6: Validation of Tbntr expressing lines

Appendix 5a: Validating the interruption of Tbexo1 in T. brucei.
A. Representative diagram of the Tbexo1 allele and effects of disrupting this allele with genes encoding for blastidicin (bla) and puromycin (pac) resistance. P1 (Tbexo1-KO1), P2 (Tbexo1-q2), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.
B. DNA fragments obtained after amplification on gDNA extracted from T. brucei wild type (lane 1), 2T1 BSF-TbNTR 9e10 (lane 2) and 2T1 BSF-TbNTR 9e10 Tbexo1Δ (lane 3). Extractions from all 3 lines were tested using primer combinations that specifically amplify biomarkers corresponding to Tbexo1 (Tbexo1-KO1/Tbexo1-q2), Tbexo1-pac (Tbexo1-KO1/pac-1) or Tbexo1-bla (Tbexo1-KO1/bla-1) alleles.
C. DNA fragments obtained after amplification on cDNA derived from RNA extracted from T. brucei wild type (lane 1) and 2T1 BSF-TbNTR 9e10 Tbexo1Δ (lane 2) using primer combinations that specifically amplify intact Tbexo1 (Tbexo1-q1/Tbexo1-q2) and Tbert (Tbert-F/Tbert-R).
Appendix 5b: Validating the interruption of \textit{Tbmre11} in \textit{T. brucei}.

A. Representative diagram of the \textit{Tbmre11} allele and effects of disrupting this allele with genes encoding for blastidicin (\textit{bla}) and puromycin (\textit{pac}) resistance. P1 (Tbmre11-KO1), P2 (Tbmre11-q2), P3 (bla-1) and P4 (pac-1) correspond to regions where the primers anneal at the specific allele.

B. DNA fragments obtained after amplification on gDNA extracted from \textit{T. brucei} wild type (lane 1), 2T1 BSF-TbNTR 9e10 (lane 2) and 2T1 BSF-TbNTR 9e10 \textit{Tbmre11}\textgreek{Delta} (lane 3). Extractions from all 3 lines were tested using primer combinations that specifically amplify biomarkers corresponding to \textit{Tbmre11} (Tbmre11-KO1/Tbmre11-q2), \textit{Tbmre11-pac} (Tbmre11-KO1/pac-1) or \textit{Tbmre11-bla} (Tbmre11-KO1/bla-1) alleles.

C. DNA fragments obtained after amplification on cDNA derived from RNA extracted from \textit{T. brucei} wild type (lane 1) and 2T1 BSF-TbNTR 9e10 \textit{Tbmre11}\textgreek{Delta} (lane 2) using primer combinations that specifically amplify intact \textit{Tbmre11} (Tbmre11-q1/Tbmre11-q2) and \textit{Tbtert} (Tbtert-F/Tbtert-R).
Appendix 7: Published papers

Unravelling the role of SNM1 in the DNA repair system of Trypanosoma brucei

James A. Sullivan,1 Jie Lun Tong,1 Martin Wong,1 Ambika Kumar,1 Hajrah Sarkar,1 Sarah Ali,1 Ikran Hussein,1 Iqra Zaman,1 Emma Louise Meredith,1 Nuala A. Helsby,2 Longqin Hu3 and Shane R. Wilkinson1*
1School of Biological & Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK.
2Department of Molecular Medicine and Pathology, University of Auckland, Private Bag 92019, Auckland, New Zealand.
3Department of Medicinal Chemistry, Ernest Mario School of Pharmacy, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA.

Summary

All living cells are subject to agents that promote DNA damage. A particularly lethal lesion are interstrand cross-links (ICL), a property exploited by several anti-cancer chemotherapies. In yeast and humans, an enzyme that plays a key role in repairing such damage are the PSO2/SNM1 nucleases. Here, we report that Trypanosoma brucei, the causative agent of African trypanosomiasis, possesses a bona fide member of this family (called Tbsnm1) with expression of the parasite enzyme able to suppress the sensitivity yeast psq2 mutants display towards mechlorethamine, an ICL-inducing compound. By disrupting the Tbsnm1 gene, we demonstrate that Tbsnm1 activity is non-essential to the medically relevant T. brucei life cycle stage. However, trypanosomes lacking this enzyme are more susceptible to bi- and tri-functional DNA alkylating agents with this phenotype readily comple-mented by ectopic expression of Tbsnm1. Genetically modified variants of the null mutant line were subse-quenty used to establish the anti-parasitic mecha-nism of action of nitrobenzylphosphoramidate mustard and aziridinyl nitrobenzamide prodrugs, compounds previously shown to possess potent trypanocidal properties while exhibiting limited toxicity to mammalian cells. This established that these agents, following activation by a parasite specific type I nitroreductase, produce metabolites that promote formation of ICLs leading to inhibition of trypanosomal growth.

Introduction

The socioeconomic development of sub-Saharan Africa has been hindered by a group of medical and veterinary infections collectively known as African trypanosomiasis. The causative agents of many of these diseases are protozoan parasites belonging to the species Trypanosoma brucei, organisms that live and multiply extracellularly in the bloodstream and tissue fluids of their mammalian hosts. Transmission occurs via the blood-feeding habits of the insect vector, the tsetse fly. Over the last 15 years implementation of improved health surveil-lance programmes combined with new treatment regimens has led to a dramatic fall in the estimated number of new cases of the human form of the disease, known as human African trypanosomiasis (HAT) from around 450 000 in 1997 to about 20 000 in 2012 (Barrett, 2006; WHO, 2014). This situation has resulted in World Health Organization aiming to eliminate HAT as a public health problem by 2020. In contrast, animal African trypanosomiasis, particu-larly in domesticated livestock, remains a major problem with these infections killing around 3 million head of cattle each year and causing an annual loss of income estimated to be about US$4.75 billion (UNFAO, 2004).

With no immediate prospect of a vaccine or chemopro-phyaxis and with vector control being problematic, drug treatment represents the only option available to combat HAT. However, the current chemotherapies used are few in number and their use is controversial, as they can be costly, often require medical supervision for administration, some have limited efficacy and may cause adverse side effects, with drug resistance becoming more widespread (Wilkinson and Kelly, 2009; Alsford et al., 2013). One way to facilitate the development of new drugs targeting HAT is to better understand the mechanism of action of existing treatments with the properties that underlie parasite selec-tivity incorporated into the development of new trypano-cidal agents. For example, melamine rings have been incorporated into several compounds to exploit the sub-strate specificity displayed by the P2 adenosine trans-
porter, a permease implicated in the uptake of pentamidine and melarsoprol (Stewart et al., 2004; Baliani et al., 2005; Chollet et al., 2009; Klee et al., 2010; Capes et al., 2012; Giordani et al., 2014). Similarly, a parasite nitroreductase (NTR), an enzyme responsible for the activation of nifurtimox (Wilkinson et al., 2008; Hall et al., 2011), has been used to screen nitroaromatic libraries for anti-T. brucei properties (Bot et al., 2010; 2013; Hall et al., 2010; 2012; Hu et al., 2011; Papadopoulou et al., 2011; 2012; 2013). In the latter case, several NTR-activated chemicals containing nitrogen mustard or aziridine functional groups that promote DNA damage via formation of cross-linkages have been identified as having significant anti-parasitic activities and low mammalian cell toxicity (Bot et al., 2010; Hall et al., 2010; Hu et al., 2011).

Genomes are constantly challenged by endogenous and exogenous agents that promote DNA damage, with interstrand cross-links (ICLs) representing a particularly dangerous lesion (O’Connor and Kohn, 1990). Formed when the two complementary strands within the DNA double helix become covalently linked, ICLs block essen-tial-tial cellular process that require DNA strand separation including DNA replication and transcription, leading to chromosomal breakage, rearrangements, or cell death (Dronkert and Kanaar, 2001; McHugh et al., 2001; Deans and West, 2011; Sengerova et al., 2011). Estimates indi-cate that a single ICL can kill a unicellular microbe with as few as 20 being fatal to a mammalian cell (Magana-Schwencke et al., 1982; Lawley and Phillips, 1996). In order to preserve the integrity and functionality of DNA, eukaryotic cells have evolved a series of complementary and overlapping pathways to repair ICLs, although the precise mechanisms involved in these systems are not fully understood (Deans and West, 2011). In Saccharomyces cerevisiae, many of the major DNA repair pathways [nucleotide excision repair (NER), mismatch repair, post-replication repair/translesion synthesis and homologous recombination] have been implicated in fixing ICL damage, although only a few proteins specifically involved in ICL lesion repair have been identified (Barber et al., 2005; Lehoczky et al., 2007; Dae et al., 2012; Ward et al., 2012). Of these, Pso2p (also known as Snm1) is of great interest as cells lacking this activity are cells able to repair ICLs within and highly susceptible to ICL-forming agents including pso-ralen, cisplatin and mechloretamine but not to any other forms of DNA damage (Henriques and Moustacchi, 1980; Ruhland et al., 1981a,b). The precise role played by Pso2p in this repair system remains unknown, although biochemi-cal studies have shown that it displays a 5’ exonuclease activity (Li et al., 2005). This coupled with the observation that pso2 cells exposed to ICL-inducing compounds tend to accumulate DNA double stranded breaks indicates that Pso2p does not function in the initial incision event, which in yeast is primarily controlled by NER, but may be involved in the processing of DNA ends created during the gener-ga-tion of ICL-associated DNA double stranded breaks (Li and Moses, 2003; Barber et al., 2005; Dudas et al., 2007). Intriguingly, Pso2p also displays a structure-specific DNA hairpin opening endonuclease activity providing evidence that it may have other functions outside ICL repair (Tiefenbach and Junop, 2012).

Here, we report that T. brucei expresses a Pso2/Snm1 homologue that can readily complement for the susceptibility phenotype exhibited by pso2 yeast cells towards an ICL forming agent. Deletion of the gene, designated Tbsnm1, from the parasite genome revealed that although the encoded enzyme is not essential for viability and growth of bloodstream form (BSF) trypanosomes, cells lacking this activity were more susceptible to bifunctional nitrogen mustard- and aziridine-based ICL-inducing agents. Using recombinant T. brucei expressing altered levels of Tbsnm1, we establish that the trypanocidal mechanism of several potent nitroaromatic-based agents that contain ICL-promoting grouping are dependent on an initial activation catalysed by a parasite specific type I NTR that generates metabolites that then promote DNA damage.

Results

Identifying trypanocidal chemical tools for studying DNA repair

Previous screening studies have identified nitroaromatic-based azirindyl/nitrogen mustard compounds to be effec-tive trypanocidal agents (Bot et al., 2010; Hall et al., 2010; Hu et al., 2011). The antimicrobial activity of these involves a parasite specific activation step catalysed by a type I NTR that leads to metabolites postulated to promote DNA damage. To determine if the above compounds do function via this pathway, a range of anti-cancer compounds known to mediate their cytotoxicity by promoting DNA cross-linkages were screened for trypanocidal activity against BSF T. brucei. The structures tested included non-nitroaromatic-based aziridines and nitrogen mustards, nitrosoureas, platinum complexes, an alkyl sulfonate and non-classical DNA cross-linking agents.

Out of the non-nitroaromatic anti-cancer compounds assessed, 17 had no effect on parasite growth at concentrations of up to 30 μM, including busulfan, the only alkyl sulfonate analysed here, and all five non-classical DNA cross-linking agents (Table 1). These were not analysed further. For the remaining compounds, the concentration that inhibits parasite growth by 50% (IC50’s) was deter-mined (Table 1). For all the remaining classes of DNA cross-linking agents, two or more compounds displayed trypanocidal activities with IC50 values ranging from 13 nM for mitomycin C, the most potent agent identified here, to approximately 35 μM for mechloretamine and ThioTEPA.
Characterising a trypanosomal DNA repair enzyme

In other eukaryotes, the SNM1/PSO2 family of nucleases play an important role in repairing damage caused by DNA cross-linking agents (Cattell et al., 2010). Analysis of the T. b. brucei genome database (Aslett et al., 2010) identified a single hypothetical gene (designated as Tbsnm1) of 2163 bp located on chromosome 4 with potential to encode for a 79.5 kDa enzyme (Tbsnm1; Gene ID: Tb927.4.1480) related to this family of enzymes. Full-length Tbsnm1 is 42% identical to the T. cruzi homologue (GenBank accession no. XP_816034) and has 27–32% identity to the leishmanial enzymes LmSNM1 (XP_001866430) and LdSNM1 (XP_003864463). When compared with yeast, plant and mammalian counterparts sequence identity ranged from 15% to 24%. Based on sequence, Tbsnm1 can be divided into two regions (Fig. 1). The amino terminal section (residues 36–182) constitutes a non-canonical metallo-β-lactamase (MBL; pfam12706) domain containing four motifs (motifs 1–4), including a characteristic HxHxH signature (motif 2), that in other SNM1/PSO2 proteins cooperate to mediate zinc cofactor binding. The second section represents a β-CASP (named after its representative members CPSF, Artemis, SNM1 and PSO2; pfam10996) region (residues 213–519) that contains within it a stretch of 31 amino acids comprising a DRMBL (DNA repair metallo-β-lactamase; pfam07522) domain (residues 488–519). The β-CASP region contains a fifth zinc binding motif (motif 5) but as with other SNM1/PSO2 sequences the precise location of this has yet to be defined: in silico analysis of Tbsnm1 indicates that D220 or H497 (motifs 5' and 5" respectively) may fulfil this role with H497 being the most likely of the two candidate residues (Callebaut et al., 2002). The β-CASP domain of Tbsnm1 also contains a diagnostic valine residue (position 519) that indicates that the parasite enzyme is involved in DNA processing: DNA processing MBLs contain a valine residue at the equivalent site, whereas RNA processing MBLs contain a histidine (Callebaut et al., 2002).

To investigate whether the T. brucei enzyme is a SNM1/PSO2 homologue, Tbsnm1 minus its ATG initiation codon was amplified and cloned into a version of the yeast expression vector pYC1ac111 that contains a DNA sequence encoding for the FLAG-tag epitope. The result-ant plasmid was transformed into the S. cerevisiae wild type and pso2 strains and expression of recombinant Tbsnm1 confirmed by western blot analysis (Fig. 2A). The susceptibility of the fungal lines to mechlorethamine, a DNA cross-linking agent, was then determined and from the resultant dose-response curves the IC50 value for each strain calculated (Fig. 2B and C). Yeast lacking pso2 were clearly more susceptible to the nitrogen mustard than wild type with the null mutant displaying an IC50 value approximately 40% that of the control strain. When Tbsnm1 was expressed in wild type yeast a slight (1.4-fold) resistance was noted. This phenotype was also observed in the pso2 strain expressing Tbsnm1 correlating with an increase in the IC50 value from 1.3 μM in cells lacking Pso2p to 5.7 μM in pso2 yeast expressing FLAG-Tbsnm1. These data

Table 1. Susceptibility of T. brucei lines to DNA damaging agents.

<table>
<thead>
<tr>
<th>Compound</th>
<th>T. brucei IC50</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nitrogen mustards</strong></td>
<td></td>
</tr>
<tr>
<td>Chlorambucil, cyclophosphamide, uramustine, trofosfamide, ifosfamide, bendamustine</td>
<td></td>
</tr>
<tr>
<td>Mechloretamine</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>Melphalan</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>Estramustine</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>Prednimustine</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>LH7</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>LH17</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>LH32</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>LH33</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>LH34</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>LH37</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td><strong>Aziridines</strong></td>
<td></td>
</tr>
<tr>
<td>ThioTEPA</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>Triethylenemelamine</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>Mitomycin C</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>CB1954</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>NH1</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td><strong>Nitrosoureas</strong></td>
<td></td>
</tr>
<tr>
<td>Carmustine, nimustine, NSCC270516</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>Lomustine</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>Streptozotocin</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>Semustine</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td><strong>Alky sulfonate</strong></td>
<td></td>
</tr>
<tr>
<td>Busulfan</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td><strong>Non-classical DNA</strong></td>
<td></td>
</tr>
<tr>
<td>cross-linking agents</td>
<td></td>
</tr>
<tr>
<td>Alitretamine, pipobroman, dacarbazine, temozolomide, mitobronitol</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td><strong>Platinum-based</strong></td>
<td></td>
</tr>
<tr>
<td>Oxaliplatin, nedaplatin</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>Cisplatin</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>Carboplatin</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td><strong>Other agents</strong></td>
<td></td>
</tr>
<tr>
<td>Hydroxyurea</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>H2O2</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>UV irradiation</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>MMS</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>DMFO</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>Nitfurimox</td>
<td>&gt;30.000</td>
</tr>
<tr>
<td>Benznidazol</td>
<td>&gt;30.000</td>
</tr>
</tbody>
</table>

The cell lines analysed were T. brucei (wild type) and T. brucei null mutants (Tbsnm1–). IC50 values are given in μM.

Identification of the DNA repair enzyme Tbsnm1

except for UV irradiation, which is in J m−2. LH7, LH17, LH32-34, LH37, CB1954 and NH1 represent structures previously identified as trypanocidal agents (Bot et al., 2010; Hall et al., 2010; Hu et al., 2011). nd is not determined.
clearly show that TbSNM1 can complement for the pso2 mutation and that the trypanosomal enzyme is a bona fide SNM1/PSO2 homologue.

**TbSNM1 is targeted to the T. brucei nucleus**

When analysed using the PSORTII and WoLFPSORT algorithms, TbSNM1 was predicted to be targeted to the nucleus via a ‘four pattern’ RRRH (residues 428–431) nuclear localisation signal. To confirm this, the full-length Tbsnm1 gene minus its ATG initiation codon was amplified and ligated in-frame and downstream of the gene encoding for the enhanced green fluorescence protein (GFP) in a trypanosomal vector that facilitated tetracycline inducible gene expression (Alsford et al., 2005). The resultant construct was used to transform BSF T. brucei, and parasite clones were selected.

To induce expression of the tagged protein, cells were incubated in the presence of tetracycline for 48 hours. Recombinant parasites were examined by Western blotting using a monoclonal antibody against GFP (Fig. 3A), with extracts derived from these cells containing a band of the expected size (~ 105 kDa), or were fixed and examined by confocal microscopy (Fig. 3B). For parasites expressing GFP-TbSNM1, GFP fluorescence was restricted to a large single spot, a pattern reported for trypanosomal proteins localised to nucleus (Fig. 3B). To confirm this, cells were co-stained with the DNA dye, DAPI. When the images were compared, the pattern of localisation indicated that GFP-TbSNM1 was located in the larger of two compartments (the nucleus) where DAPI is found with the smaller, faint spot corresponding to the kinetoplast, the genome found in the parasites’ single mitochondrion.

**Functional analysis of TbSNM1 in T. brucei**

To assess whether TbSNM1 was essential to BSF T. brucei, an RNAi-based approach was initially employed. A DNA fragment corresponding to an internal region of Tbsnm1 was cloned into p2T7Ti (Wilkinson et al., 2003) and the construct transformed into BSF T. brucei. In the absence of tetracycline, recombinant clones were found to grow at approximately the same rate as the parental cells. Addition of tetracycline, recombinant clones were found to grow at approximately the same rate as the parental cells. Addition of tetracycline to parasites harbouring the RNAi construct did not affect the growth rate suggesting that TbSNM1 is not essential to BSF T. brucei. To confirm this, DNA fragments corresponding to the 5′ flank of Tbsnm1 and the 3′ region of the Tbsnm1 gene were cloned either side of a cassette containing blasticidin or puromycin resistance markers. The integration constructs were trans-
Characterising a trypanosomal DNA repair enzyme

Fig. 2. Complementation of the yeast pso2Δ mutation.
A. Western blot analysis was carried out using a monoclonal antibody to the FLAG-tag epitope on cell extracts made from S. cerevisiae BY4742 (wild type) and pso2Δ strains expressing the FLAG epitope (control) or FLAG-TbSNM1. A band of ∼ 80 kDa (indicated by **) was observed in lysates derived from cell expressing the recombinant trypanosomal protein. A cross-reactive epitope (*) and Ponceau S staining of the membrane (not shown) were used as loading controls.
B and C. The susceptibility of wild type (a and b) and pso2Δ (c and d) yeast strains expressing FLAG (a and c) or FLAG-TbSNM1 (b and d) to different concentrations of mechlorethamine. All data are mean values ± standard deviations from experiments performed in triplicate. In C, the values given in parenthesis represent the fold difference in IC_{50} values (in μM) relative to wild-type controls.

Cells lacking TbSNM1 were more susceptible to a range of nitrogen mustard and aziridinyl compounds, including several of the trypanocidal nitroaromatic structures previously identified (Bot et al., 2010; Hall et al., 2010; Hu et al., 2011). Intriguingly, Tbsnm1Δ−/− cells exhibited a larger difference in their sensitivities to the nitrogen mustards screened than that observed when using the aziridinyl compounds. When these growth assays were extended to look at other DNA damaging agents including mitomycin C, semustine, cisplatin, methyl methanesulphonate (MMS), H_2O_2, hydroxyurea and UV light, and to the clinically used trypanocidal drugs nifurtimox, benzimidazole or difluoromethylornithine (DFMO), no difference in IC_{50} was observed.

In order to demonstrate conclusively that the altered susceptibility phenotypes were specifically due to lack of TbSNM1, a complementation strategy was used. In
Fig. 3. Localisation of TbSNM1 in bloodstream form *T. brucei*.
A. Expression of GFP-TbSNM1 was examined by probing a blot containing cell lysates from *T. brucei* wild type (lane 1) and GFP-TbSNM1 expressing cells (lane 2) using an anti-GFP antibody (upper panel). Protein from $1.5 \times 10^7$ cells was loaded in each track and a cross-reactive epitope (lower panel) and by Coomassie staining (not shown) were used as loading controls.
B. Parasites expressing GFP-TbSNM1 were co-stained with DAPI (DNA) and the cells examined by confocal microscopy. The TbSNM1 signal is coincidental with the nucleus (n; large DAPI spot); the smaller DAPI spot corresponds to the kinetoplast (k), the trypanosome mitochondrial genome.

during these experiments, *Tbsnm1*−/− cells were transformed with a vector that facilitates constitutive expression of an ectopic copy of *Tbsnm1* integrated into one of the parasite’s tubulin arrays: wild-type cells also expressing this vector were also generated. The IC$_{50}$ of these parasites towards selected nitrogen mustard and aziridinyl compounds was determined and compared with values obtained using wild type and *Tbsnm1* null mutant lines (Fig. 5A). When the susceptibility of the complemented line to the nitrobenzyl-containing nitrogen mustard (mechlorethamine) and aziridinyl (triethylenemelamine) compounds, a resistance phenotype was noted, with the dose response curves (and associated IC$_{50}$ values) being distinct from the *Tbsnm1*+/- cells, which displayed increased sensitivity to both agents, and equivalent to the plots observed using *Tbsnm1* expressing parasites (wild type and wild-type cells engineered to express elevated levels of TbSNM1) (Fig. 5A). When the screens were extended to investigate the complemented line’s susceptibility to non-nitroaromatic nitrogen mustard (mechlorethamine) and aziridinyl (triethylenemelamine) compounds, a resistance phenotype was noted, with the dose response curves (and associated IC$_{50}$ values) in the complemented line mirroring that obtained for wild-type parasites expressing elevated levels of TbSNM1 (Fig. 5A): trypanosomes (wild type and *Tbsnm1* null mutants) expressing an ectopic copy of *Tbsnm1* were up

Fig. 4. Disruption of *Tbsnm1* in *T. brucei*.
A. Diagram of the *Tbsnm1* alleles and the effects of gene disruption. A 5′ *Tbsnm1* flanking sequence and a 3′ *Tbsnm1* coding region were amplified and cloned sequentially either side of a puromycin (pac) or blasticidin (bla) cassette (plus *T. brucei* tubulin intergenic elements required for processing of mRNA (hashed boxes). The dotted lines correspond to the probe used to check integration. The position of the predicted MluI sites plus the band sizes (in kbp) obtained after hybridisation are shown.
B. Southern blot analysis of MluI digested genomic DNA from *T. brucei* (lane 1), *Tbsnm1*+/- *bla* and *Tbsnm1*+/- *pac* heterozygous clones (lanes 2 and 3 respectively) and a *Tbsnm1*−/− null mutant line (lane 4). Blots were hybridised with labelled 3′ region of sequences. Sizes given are in kbp.
Characterising a trypanosomal DNA repair enzyme

Fig. 5. Susceptibility of T. brucei lines expressing altered levels of TbSNM1 to DNA damaging agents.

A. Growth inhibitory effects (expressed as IC₅₀ values in μM or nM) of the T. brucei wild type, Tbsnm1⁻/⁻ null mutant, Tbsnm1⁻/⁻ expressing an ectopic copy of Tbsnm1 (Tbsnm1⁻/⁻ + Tbsnm1⁺⁺) and T. brucei expressing elevated levels of Tbsnm1 (wild type + Tbsnm1⁺⁺) lines towards DNA damaging agents. Integration of the Tbsnm1 expression vector into a single tubulin array was confirmed by Southern hybridisation and expression evaluated through qPCR (data not shown).

B. Growth inhibitory effects (expressed as IC₅₀ values in μM or nM) of T. brucei wild type, Tbsnm1⁻/⁻ null mutant, T. brucei expressing an ectopic copy of Tbntr (wild type + Tbntr9e10⁺⁺ and Tbsnm1⁻/⁻ expressing elevated levels of Tbntr (Tb snm1⁻/⁻ + Tbntr9e10⁺⁺)) towards DNA damaging agents. Expression of Tbntr was evaluated through qPCR (data not shown). Data in panels A and B are mean values ± standard deviations from experiments performed in quadruplicate.

© 2015 John Wiley & Sons Ltd, Molecular Microbiology

the above complementation studies indicate that parasites (wild type or Tb snm1⁻/⁻) ectopically expressing Tb snm1 are resistant to non-nitroaromatic DNA cross-linking agents but not to the nitroaromatic-containing compounds. One reason for this could reflect that the latter structures function as prodrugs and must undergo an NTR catalysed activation step before mediating their trypanocidal DNA damaging activities.

Linking prodrug activation with DNA damage

To identify any link between the DNA damaging and the NTR-activating pathways, both copies of the Tb snm1 gene were deleted from T. brucei cells, expressing an ectopic copy of Tbntr and the susceptibilities of these recombinant cells towards selected nitrogen mustards and aziridinyl compounds determined (Fig. 5B). When treated with mechlorethamine or triethylenemelamine, both Tb snm1 expressing cell lines (wild type and trypanosomes expressing elevated levels of Tbntr) displayed similar dose response curves and therefore had similar IC₅₀S to either agent (Fig. 5B). When these studies were expanded to investigate the susceptibility of parasites lacking TbSNM1, the Tb snm1⁻/⁻ null mutant line and Tb snm1⁻/⁻ cells expressing the ectopic copy of Tbntr displayed equivalent IC₅₀ values, with both being more sensitive to mechlorethamine and triethylenemelamine than wild type (Fig. 5B). Importantly, no difference in IC₅₀ was observed using either of the lines lacking Tb snm1 indicating that Tbntr plays no role in metabolising either mechlorethamine or triethylenemelamine.

When the nitrobenzyl-containing DNA cross-linking agents LH34 and NH1 were tested against the parasite lines expressing altered levels of Tb snm1 and/or Tbntr, a different outcome was observed (Fig. 5B). For Tb snm1⁻/⁻ parasites or wild-type cells expressing an ectopic copy of Tbntr, treatment with either damaging agents resulted in increased susceptibility when compared against controls, with Tbntr overexpressing trypanosomes being more sen-sitve to LH34 and NH1 than the null mutant lines: Tb snm1⁻/⁻ cells and T. brucei overexpressing Tbntr were 18.0- and 24.0-fold more susceptible to LH34, respectively, with a 2.7- and 31.1-fold increase in sensitivity noted towards NH1. For Tb snm1⁻/⁻ null parasites express-
ing an ectopic copy of Tbntr, this increase in potency was magnified further with these cells showing a 80.0- and 38.7-fold increase in susceptibility towards LH34 and NH1, respectively, when as compared against wild type.

Discussion

Currently, very little is known about the mechanisms of *T. brucei* employs to repair ICL damage even though this parasite is exposed to such deleterious insults throughout its cell and life cycles. In other unicellular eukaryotes such as budding and fission yeast, the processing of ICLs occurs through the concerted activities of several major DNA repair pathways with one enzyme, Pso2p, playing a central and specific role in fixing such lesions. Although non-essential for yeast viability, the importance of Pso2p is only apparent in its absence on exposure to ICL-inducing, bifunctional alkylating agents but not to monofunctional alkylating agents, ionising radiation or ultraviolet light (Henriques and Moustacchi, 1980; Ruhland et al., 1981a,b). Here, we report the characterisation of TbSNM1, a trypanosomal Pso2p homologue, and demonstrate that this enzyme plays a key role in processing ICL lesions when generated by bifunctional nitrogen mustard and aziridine compounds including several nitroaromatic-based agents previously shown to have potent anti-trypanosomal properties with low toxicity to mammalian cells (Bot et al., 2010; Hall et al., 2010; Hu et al., 2011).

In terms of its sequence, TbSNM1 displays the main characteristics found in other Pso2/SNM1 proteins, possessing adjacent MBL and β-CASP domains that together form the enzyme’s zinc-binding central catalytic core (Cattell et al., 2010). To confirm the *in silico* identification, a complementation approach was undertaken. This involved ectopically expressing the trypanosomal enzyme in a yeast *pso2* line and then evaluating the susceptibility of the resultant cells to mechlorethamine, a bifunctional alkylating agent routinely used as an ICL-inducing agent. In this genetic background, the parasite protein was able to revert the susceptibility phenotype displayed by the *pso2* line resulting in an additional slight (approximately twofold) resistance towards this nitrogen mustard. This confirmed that the trypanosomal enzyme is a genuine Pso2p homologue and that it plays a role in the processing of ICL lesions. Furthermore, as TbSNM1 can complement for the *pso2* mutation then the parasite enzyme may interact with the same partner proteins as its yeast counterpart. For example, Pso2p contains an ubiquitin binding zinc finger (UBZ) C2HC motif upstream of its catalytic core (Yang et al., 2010). By analogy with hSNM1A, the only human PSO2/SNM1 homologue able to complement the yeast *pso2* mutation (Hazratli et al., 2008), this signature sequence is able to facilitate binding to monoubiquinated PCNA leading to recruitment of this repair enzyme to

ICL-stalled replication forks (Yang et al., 2010). Interestingly, *in silico* searches failed to identify any known UBZ C2HC domain or any other type of ubiquitin interaction motifs (UIMs) in the parasite protein sequence. Therefore, if formation of PSO2/SNM1-containing DNA repair complexes at the site of ICL damage does involve PCNA ubiquitylation, then the molecular mechanisms underlying TbSNM1 recruitment to such lesions occurs through an as yet uncharacterised UIM or via interactions involving a conserved adapter protein. Recently, it has been shown that the β-CASP domain of Pso2p can be phosphorylated leading to the suggestion that this event may play a role in modulating the enzyme’s exo- or endo-nucleolytic activity (Munari et al., 2014). Whether TbSNM1 undergoes a similar post-translational modification, and how this effects its nuclease activity has yet to be established.

The endogenous function of TbSNM1 is non-essential to replicating *T. brucei*; both copies of *Tbsnm1* could be deleted from the genome of BSF trypomastigote parasites. However, the importance of this enzyme to the trypanosome, with some only became evident following exposure to ICL-inducing compounds: null mutant cells were more susceptible to bi- and tri-functional alkylating agents as compared with controls, whereas these recombinant cells display an equivalent sensitivity to wild type when exposed to MMS, UV irradiation and *H*2*O*2, treatments normally repaired by homologous recombination, nucleotide excision repair or base excise repair pathways. This trait was solely due to loss of TbSNM1 activity as expression of an ectopic copy of *Tbsnm1* in the null mutant genetic back-ground restored the recombinant parasites IC50 near to wild-type levels. Intriguingly, the range of compounds that elicits the change in susceptibility in the *Tbsnm1*+/− tryposomes although similar to that noted for the yeast *pso2* line does have some notable differences (Henriques and Moustacchi, 1980; Ruhland et al., 1981a,b). For example, yeast *pso2* mutants are reported to be more susceptible to cisplatin and mitomycin C, whereas *T. brucei* lacking TbSNM1 displays sensitivities similar to that exhibited by wild-type parasites. This may be because that although both compounds can function as ICL-inducing agents they can also mediate their cytotoxic activities via other mechanisms including promoting formation of intrastrand cross-links, activating signal transduction pathways, stimulating redox cycling, acting as enzyme inhibitors or alkylating other biological molecules (Sharma and Tomasz, 1994; Pagano et al., 2003; Siddik, 2003; Rabik and Dolan, 2007; Paz et al., 2012). One (or a combination) of these alternative modes of action (or possibly another unidentified mechanism) may account for cisplatin’s and mitomycin C’s trypanocidal properties, therefore negating the requirement for a TbSNM1-dependent ICL repair pathway.

Previous trypanocidal screening programmes have identified nitrobenzylphosphoramidate mustards and aziridi-
nito benzamides as having potent anti-parasitic activity (Bot et al., 2010; Hall et al., 2010; Hu et al., 2011). These agents function as prodrugs and must be activated before they can mediate their cytotoxic effects, a reaction catalysed by a NADH-dependent type I NTR. This reduction causes the conversion of a conserved electron withdraw-ing nitro-group present on the compound’s benzyl ring to an electron donating hydroxylamine derivative (Bot et al., 2010; Hall et al., 2010; Hu et al., 2011). This action effect-ively acts as an electronic switch that is believed to turn on the alkylating ability of the nitrogen mustard or aziridinyl moiety causing ICL-mediated DNA damage. Using Tbsnm1 null mutant parasites engineered to express elevated levels of TbNTR, we have now demonstrated a link between prodrug activation and ICL formation. Here we observed that wild-type parasites exhibited the highest IC50 values towards ICL formation and not the nucleus. 

family. The trypanosomal enzyme displays characteristics of its yeast counterpart and is able to repair the DNA damage caused by bi- and tri-functional alkylating agents. By exploiting parasites lacking this enzyme we were able to demonstrate that following TbNTR-mediated activation nitrobenzylphosphoramide mustard and aziridinyl nitrobenzamide agents, compounds previously shown to have potent trypanocidal properties with little/no cytotoxicity in mammalian cells, generate metabolites that promote ICL formation. Although not essential to survival of the medically relevant form of T. brucei, in the future TbSNM1 could be targeted through the use of inhibitors to improve the potency of other drugs that do cause parasite death through forma-tion of the extremely lethal ICL.

**Experimental procedures**

**Cell culturing**

Bloodstream form Trypanosoma brucei brucei [MITat 427 strain; clone 221a and a derivative (2T1) engineered to express elevated levels of TbNTR-myc] was maintained in HMI-9 (Invitrogen) medium supplemented with 3 g l−1 sodium bicarbonate, 0.014% (v/v) β-mercaptoethanol and 10% (v/v) foetal calf serum (Hirumi and Hirumi, 1989; Wilkinson et al., 2008) at 37°C under a 5% (v/v) CO2 atmosphere. The 2T1 cells were grown in the presence of 1 μg ml−1 phleomycin and 2 μg ml−1 puromycin. Transformed T. brucei cells were grown in the presence of 2.5 μg ml−1 hygromycin, 10 μg ml−1 blasti-cidin and/or 2 μg ml−1 puromycin.

Saccharomyces cerevisiae strains BY4742 (MATa his3-1 leu2-0 lys2-0 ura3-0) and a pso2 derivative obtained from the Open Biosystems (Thermo Scientific) knock-out collection were maintained in yeast extract-peptone broth containing 2% (w/v) glucose. Transformed cells were grown in Synthetic Complete Dropout medium lacking leucine (Sigma).

**Chemicals**

The DNA damaging agents were obtained from Drug Synthesis and Chemistry Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute except CB1954, H2Oz, MMS and hydroxyurea (all Sigma-Aldrich), NH1 (Helsby et al., 2004) and LH7, LH17, LH32-34 and LH37 (Hu et al., 2003; 2011; Li et al., 2003). Nifurtimox and benznidazole were obtained from Simon Croft (London School of Hygiene and Tropical Medicine) and DFMO from Mike Barrett (University of Glasgow).

**Plasmids**

The vectors used to delete Tbsnm1 from the T. brucei genome were generated as follows. Primers were designed to amplify 897 or 909 bp fragments from the 5’ Tbsnm1 untranslated region and 3’ region of the Tbsnm1 gene respectively. These
were cloned sequentially either side of a puromycin- (pac) or blasticidin- (bla) containing resistance cassette. The constructs were linearised (SacI/KpnI for the pac vector or SacII/ KpnI for the bla vector) then introduced into BSF *T. brucei* using the Human T-cell Nucleofector® kit and an Amaxa® Nucleofector™ (Lonzar AG) set to program X-001. Integration of the DNA constructs into the *T. brucei* genome results in deletion of 60% of the Tbsnm1 open reading frame (amino acids 1 to 425) including all of the non-canonical MB domain. As this region is essential for Pso2p/SNM1 function (Li and Moses, 2003), removal of the MBL encoding DNA sequence from the trypanosomal genome would generate parasites lacking TsSNM1 activity, effectively producing Tbsnm1 null mutant cells.

The Tbsnm1 trypanosomal expression vector was generated as follows: a 2166 bp DNA sequence corresponding to full-length Tbsnm1 was amplified from *T. brucei* genomic DNA using the primers cgccgctatttctgtagactcag and ggacgctcttatttctgtagactcag (lower-case italics correspond to restriction sites incorporated into the primers to facilitate cloning), digested with Sdi1/Sgs1 and ligated into the corresponding sites of vector pTubEX-LmSpSyn (Taylor et al., 2008), replacing LmSpSyn. The NotI/ XhoI digested construct was introduced into *T. brucei* wild type and Tbsnm1<sup>−/−</sup> cells using nucleofection and recombinant clones selected.

For the localisation construct, a 2166 bp DNA sequence corresponding to full-length Tbsnm1 was amplified from *T. brucei* genomic DNA using the primers tctagacgcaggtgcagtgcctgtagactcag and aagcttattttctgtagactcag (lower-case italics correspond to restriction sites incorporated into the primers to facilitate cloning), digested with XbaI/BgIII and ligated into the XbaI/BamHII sites of vector pRPα<sub>GFP</sub> (Alsford et al., 2005) to replace the Tdr1 coding sequence. The cloning was carried out such that the gene coding for the GFP was inserted in-frame at the 5’ end of the Tbsnm1-derived DNA fragment. The AscI/XhoI digested construct was introduced into *T. brucei* 2T1 parasites.

To construct the yeast complementation vector Tbsnm1 was amplified from the trypanosomal localisation plasmid using the primers tctagacgcaggtgcagtgcctgtagactcag and aagcttattttctgtagactcag (lower-case italics correspond to restriction sites incorporated into the primers to facilitate cloning). The resultant fragment was digested with XbaI/HindIII and ligated into the corresponding sites of a pYCYlac111 derivative containing a DNA sequence encoding for the FLAG-tag epitope (Novoselova et al., 2013). The plasmid was transformed into yeast strains BY4742 and pso2<sup>−</sup>. In this system recombinant TbsNM1 is tagged at its amino-terminus with a FLAG-tag epitope detectable with the anti-FLAG monoclonal antibody (Sigma).

**Localisation**

BSF trypanosomes expressing GFP-TbSNM1 were washed twice in phosphate-buffered saline (PBS), fixed in 2% (w/v) paraformaldehyde/PBS and washed again in PBS. Aliquots of the cell suspension (10<sup>5</sup> cells) were then air dried onto micro-scope slides. Parasite DNA was stained using Vectashield Mounting Medium containing 4',6-diamidino-2-phenylindole (DAPI) (Vectorshield Laboratories), and slides were viewed using a Leica SP5 confocal microscope (Leica Microsystems (UK) Ltd).

**Antiproliferative assays**

All assays were performed in a 96-well plate format. *T. brucei* BSF parasites were seeded at 1 x 10<sup>4</sup> ml<sup>−1</sup> in 200 μl growth medium containing different concentrations of compound. For UV irradiation, parasites were exposed to doses up to 900 J m<sup>−2</sup> using a Stratalinker® UV crosslinker (Stratagene). After incubation at 37°C for 3 days, 2.5 μg resazurin (20 μl of 0.125 μg ml<sup>−1</sup> stock in phosphate buffered saline) was added to each well, and the plates were incubated for a further 6–8 hours (Jones et al., 2010). Cell densities were determined by monitoring the fluorescence of each culture using a Gemini Fluorescent Plate Reader (Molecular Devices (UK) Ltd, Wokingham, UK) at an excitation wavelength of 530 nm, emission wavelength of 585 nm and a filter cutoff at 550 nm. The drug/treatment concentration that inhibits cell growth by 50% (IC<sub>50</sub>) was established using the non-linear regression tool on GraphPad Prism (GraphPad Software).

**Yeast complementation assay**

All assays were performed in a 96-well plate format. The cell density of overnight yeast cultures were equalised according to absorbance at 405 nm in medium containing different concentrations of mechlorethamine. The growth of each strain was then followed by monitoring the change in absorbance at 405 nm using an Absorbance Microplate Reader (BioTek Instruments Ltd). The % growth for each mechlorethamine-treated culture after 18 hours relative to untreated samples was determined.

**Acknowledgements**

Emma Louise Meredith is a recipient of a BBSRC Doctorial Training Studentship. We would like to thank Martin Taylor (London School of Hygiene and Tropical Medicine) for valuable discussions and comments on the manuscript.

**References**


© 2015 John Wiley & Sons Ltd, Molecular Microbiology


Distinct activation mechanisms trigger the trypanocidal activity of DNA damaging prodrugs

Emma Louise Meredith,1† Ambika Kumar,1† Aya Konno,1 Joanna Szular,1 Sam Alsford,2 Karin Seifert,2 David Horn 3 and Shane R. Wilkinson 1*
1 School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London, E1 4NS, UK.
2 Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London, UK.
3 The Wellcome Trust Centre for Anti-Infectives Research, School of Life Sciences, University of Dundee, Dundee, UK.

Summary

Quinone-based compounds have been exploited to treat infectious diseases and cancer, with such chemicals often functioning as inhibitors of key metabolic pathways or as prodrugs. Here, we screened an aziridinyl 1,4-benzoquinone (ABQ) library against the causative agents of trypanosomiasis, and cutaneous leishmaniasis, identifying several potent structures that exhibited EC50 values of <100 nM. However, these compounds also displayed significant toxicity towards mammalian cells indicating that they are not suitable therapies for systemic infections. Using anti-T. brucei ABQs as chemical probes, we demonstrated that these exhibit different trypanocidal modes of action. Many functioned as type I nitroreductase (TbNTR) or cytochrome P450 reductase (TbCPR) dependent prodrugs that, following activation, generate metabolites which promote DNA damage, specifically interstrand crosslinks (ICLs). Trypanosomes lacking TbSNM1, a nuclease that specifically repairs ICLs, are hypersensitive to most ABQ prodrugs, a phenotype exacerbated in cells also engineered to express elevated levels of TbNTR or TbCPR. In contrast, ABQs that contain substituent groups on the biologically active aziridine do not function as TbNTR or TbCPR-activated prodrugs and do not promote DNA damage. By unravelling how ABQs mediate their activities, features that facilitate the desired anti-parasitic growth inhibitory effects could be incorporated into new, safer compounds targeting these neglected tropical diseases.

Introduction

The protozoan parasites Trypanosoma brucei, T. cruzi and various Leishmania species are the etiological agents of human African trypanosomiasis (HAT), Chagas disease and Leishmaniasis, respectively. Spread by the hematophagous habits of insect vectors, these pathogens cause more than 55,000 deaths per year and are prevalent in many regions of the world least able to deal with the associated economic burden (http://www.dndi.org/diseases-projects/). Implementation of new surveillance and treatment programmes in conjunction with improved housing and vector control strategies has resulted in a dramatic reduction in disease prevalence. For example, the number of new cases of HAT has fallen from an estimated peak of 450,000 in 1997 to below 20,000 in 2014 while Chagas disease has been eliminated from Chile, Uruguay and several regions of Argentina and Brazil (Barrett, 2006; Schofield et al., 2006; WHO, 2014). The success of such strategies has led to the World Health Organization (WHO) aiming to eliminate HAT as a public health problem by 2020 (WHO, 2013).

Currently, drugs represent the only viable option to combat trypanosomal and leishmanial infections although their use is problematic. They often require supervision for administration, can be costly, have limited efficacy and may cause significant toxicity. Addition-ally, drug resistance is beginning to further limit the efficacy of the available chemotherapeutic arsenals, with antimonials no longer recommended as a first line treatment for leishmaniasis on the Indian sub-continent (Wilkinson and Kelly, 2009; Barrett and Croft, 2012; Alsford et al., 2013; Perry et al., 2015). In the case of HAT and Chagas disease, the front line treatments involve nitroheterocyclic-based prodrugs with nifurtimox

---

Accepted 4 August, 2017  *For correspondence. E-mail s.r.wilkinson@qmul.ac.uk; Tel. 144 (0)20 7882 3057; Fax 144 20 882 7732. †These authors contributed equally to this work.

Vc: 2017 The Authors Molecular Microbiology Published by John Wiley & Sons Ltd
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
or benznidazole monotherapies used to target T. cruzi while nifurtimox, in combination with difluoromethylornithine (DFMO), is used against a form of HAT prevalent throughout West and Central Africa (Priotto et al., 2009; Yun et al., 2010; Wilkinson et al., 2011). To mediate their trypanocidal effects, nifurtimox and benznidazole undergo an activation reaction catalysed by an FMN-containing, mitochondrial type I nitroreductase (NTR), generating cytotoxic reduction products (Wilkinson et al., 2008; Hall et al., 2011; Hall and Wilkinson, 2012). For nifurtimox, this results in the production of an unsatu-rated open chain nitrile while benznidazole is processed to glyoxal via a series of highly reactive intermediates. As type I NTRs are expressed by some unicellular eukar-yotes (including trypanosomes and Leishmania) but not by metazoan organisms, the bioreductive activity of this enzyme has been exploited to develop a series of novel antiparasitic nitroaromatic- and benzoquinone-based pro-drugs (Wilkinson et al., 2011; Patterson and Wylie, 2014). Several of these display significant potency against T. brucei in vitro while exhibiting little or no toxicity towards cultured mammalian cells.

Quinones represent a class of organic compounds that contain two carbonyl groups attached to a six mem-bered carbocyclic backbone. They are ubiquitous in nature, functioning in various oxidoreductase cascades, with some natural and synthetic variants being of phar-macological interest. In the latter case, many quinone-based agents often function as prodrugs and must undergo activation before they can mediate their cyto-toxic effects, reactions catalysed by quinone oxidoreduc-tases (McKeown et al., 2007; Siegel et al., 2012). Based on oxygen-sensitivity, such enzymes can be divided into two groups. Oxygen-sensitive quinone oxidoreductases, such as NADH cytochrome b5 reductase (CPR), and cytochrome P450 reductase (CPR), can mediate the 1e− reduction of one of the quinone’s carbonyl oxygens to form an unstable semiquinone radical that, under hypoxic conditions, can be further reduced to the hydro-quinone derivative (Powis, 1989; O’Brien, 1991). However, in the presence of oxygen, the semiquinone radical can undergo futile cycling, generating superoxide anions and regeneration of the parent compound. In contrast, oxygen-insensitive quinone oxidoreductases, such as NAD(P)H quinone oxidoreductase 1 (NQO1), catalyse the simultaneous 1e− reduction of both the quin-one’s carbonyl oxygens to form the hydroquinone directly (Ernster et al., 1962; Iyanagi and Yamazaki, 1970; Siegel et al., 2012). Dependent on structural context, this conversion can result in stable pharmacologically active products, as is the case for the benzoquinone ansamycin antibiotics, or unstable metabolites that undergo further rearrangement before exerting their toxic effects, as seen with the antitumor agents mitomycin A and b-lapachone (Siegel et al., 2012).

Due to their favorable redox and electrochemical properties, compounds containing a quinone pharma-cophore represent attractive scaffolds for antimicrobial drug development (Pinto and de Castro, 2009; Beena and Rawat, 2013). Screening against trypanosomes has resulted in the identification of various natural and synthetic lead structures postulated to function as prodrugs promoting oxidative stress within the parasite or through formation of toxic bioreductive products (Henderson et al., 1988; Kubata et al., 2002; Hoet et al., 2004; Pinto and de Castro, 2009; Ramos et al., 2009; Garavaglia et al., 2010; Hall et al., 2012). Using recombinant T. brucei that lack TbSNM1, a DNA repair enzyme that specifically fixes interstrand crosslinks (ICLs) (Sullivan et al., 2015), while expressing elevated levels of oxygen-sensitive (TbCPR) or –insensitive (TbNTR) quinone oxidoreductases, we demonstrate that many aziridinyl 1,4-benzoquinones (ABQs) possess significant potency towards trypanosomatid parasites, with different struct-ures undergoing distinct activation mechanisms to gen-erate metabolites that subsequently promote DNA damage within the parasite nucleus.

**Results**

Antiparasitic activity of aziridinyl benzoquinones

Compounds containing an ABQ core display potent anti-cancer properties particularly against cells where NADPH quinone oxidoreductase 1 (NQO1) expression is upregulated (Lin et al., 1972; Lusthof et al., 1989; Dehn et al., 2004). Recently, the antiparasitic activities of such compounds towards Plasmodium falciparum and T. brucei has been investigated with several poten-tial lead structures identified (Grellier et al., 2010; Hall et al., 2012). Here, we have expanded on these initial trypanosomal screens to evaluate the growth inhibitory properties of a larger ABQ library against T. brucei, T. cruzi and L. major (Table 1). Out of the 34 compounds tested, 20, 14 and 16 had no effect on T. brucei, T. cruzi and L. major growth, respectively, at concentrations of up to 10 mM. These were not analysed further. For the remaining compounds, RH1 exhibited potencies, expressed as the compound concentration that inhibits cell growth by 50% (EC₅₀), of <100 nM towards blood-stream form T. brucei, with a further five (DZQ, ABQ3, ABQ6, TZQ and ABQ22) exhibiting moderate potency towards this parasite (EC₅₀ values of 100–500 nM). Against T. cruzi epimastigotes, DZQ, TZQ and RH1 were highly active, yielding EC₅₀ values <10 nM, with others displaying high (EC₅₀ values between 10 and
Cytotoxicity against mammalian cells

To evaluate whether those ABQs that displayed activity against the three trypanosomatid parasites exhibited toxicity to mammalian cells, their inhibitory properties against differentiated THP-1 cells was determined (Table 1). For all structures tested, an in vitro toxicity was observed (EC50 values < 10 nM), with two compounds, including RH1, being extremely toxic and yielding EC50 values of < 100 nM. Comparison of the mammalian toxicity and antiparasitic EC50 data allowed a crude measure of each agent’s selectivity (the Selectivity Index or SI) toward the pathogen. In many cases, the ABQs displayed a higher potency against THP-1 cells relative to the parasite, resulting in a SI value < 1. Of those that did display preferential activity against the parasite, the observed selective toxicity of most was equivalent to that seen with nifurtimox, an agent whose use in humans is problematic. The toxicity of these compounds towards cultured mammalian cells precluded any attempt to establish the potency of these structures against the intracellular amastigote forms of T. cruzi and L. major.

As THP-1 cells are an immortal human monocyte line and ABQs have potent activity against cancerous cells (Tan et al., 1984; Lee et al., 1986; Moore et al., 1997), a series of screens using a selected quinone series on non-cancerous mouse peritoneal macrophages were performed. For DZQ, RH1, TZQ and ABQ22, substantial toxicity towards this line was observed with the compounds generating EC50 values of 255, < 100, < 100 and 120 nM, respectively. This clearly demonstrates that ABQs are toxic to primary cells, thus supporting our THP-1 findings. In light of the mammalian toxicity issues, our focus on ABQs switched from exploring them as potential treatments for infectious diseases to

100 nM; MeDZQ, ABQ22 and ABQ24) or moderate (EC50 values between 100 and 500 nM; ABQ4, ABQ6 and ABQ14) growth inhibitory effects. When tested against L. major promastigotes, TZQ exhibited an EC50 of < 10 nM, with three other compounds (DZQ, RH1 and ABQ25) exhibiting high potency and eight others yielding moderate potency. Of all the ABQs analysed, RH1 was the most potent agent tested against all three parasites, yielding EC50 values of 19.6 ± 1, 3.6 ± 1 and 68.6 ± 1 nM against T. brucei, T. cruzi and L. major, respectively.
using them as tools to dissect trypanosomal pro-drug activation pathways and the structure-activity relationships among this class of compound.

Different trypanosomal mechanisms activate aziridinyl benzoquinone prodrugs

The application of genome-scale RNAi screens has helped elucidate the trypanocidal mechanism of action of a number of agents and identify how the affected pathways may impinge on drug resistance (Alsford et al., 2012; 2013). As an initial step to understanding how ABQs mediate their antiparasitic activities, a loss-of-function screen was conducted on bloodstream form T. brucei, using RH1 as the selective agent (Fig. 1A). Treatment of the parasite RNAi library with this compound resulted in reduced T. brucei growth over the first week followed by the outgrowth of an RH1-resistant population. Genomic DNA was extracted from the RH1-selected cells and RNAi targets were amplified, from which a single 1.1 kbp fragment was identified as the major amplicon (Fig. 1B). Sequence analysis and comparison against the reference genome database revealed that this RH1-resistance-associated fragment mapped to a region encompassing the 5' untranslated region and 5' coding sequence of the open reading frame (Gene ID: Tb927.7.7230 on TriTrypDB - http://tritrypdb.org/tritrypdb/) for the type I nitroreductase, TbNTR (Fig. 1C). This was not unexpected given that this protein had been previously associated with the activation of trypanocidal nitroheterocyclic-based prodrugs, nifurtimox and benzni-dazole (Wilkinson et al., 2008; Baker et al., 2011; Hall et al., 2011; Hall and Wilkinson, 2012) and postulated to function as a NADH dependent quinone oxidoreductase (Wilkinson et al., 2008; Alsford et al., 2012; Hall et al., 2012).

To conclusively show that TbNTR is the key activator of ABQs, we evaluated the susceptibility of a T. brucei Tbntr heterozygote line to RH1. These recombinant cells displayed an EC₅₀ approximately 2.5-fold greater than that observed using wild type parasites (Fig. 2A) confirming that reduction of TbNTR expression through loss of one of the alleles encoding for this oxidoreductase is sufficient to generate resistance to RH1. As reduction of TbNTR activity leads to RH1 resistance, gain of function via over expression of the oxidoreductase should have the converse effect, generating cells that are more susceptible to the ABQ. To determine if
out of the additional parasites to a selected group of compounds, was tested (Table 2). Out of the additional

![Fig. 2. Susceptibility of T. brucei expressing altered levels of activator to RH1 or ABQ22.](image)

A. Dose response curve of wild type (solid line) and Tbntr heterozygous (dotted line) T. brucei to RH1. The extrapolated EC50 values (in nM) of wild type (white bar) and Tbntr heterozygous (grey bar) T. brucei to RH1 are shown.

B. Dose response curve of wild type (solid line) and Tbntr overexpressing (dotted line) T. brucei to RH1. The EC50 values (in nM) of wild type (white bar) and Tbntr overexpressing (grey bar) T. brucei to RH1 are shown.

C. Dose response curve of wild type (solid line) and Tbcpr2 overexpressing (dotted line) T. brucei to ABQ22. The extrapolated EC50 values (in nM) of wild type (white bar) and Tbcpr2 overexpressing (grey bar) T. brucei to RH1 are shown.

All data are means for experiments performed in quadruplicate 6 standard deviation. The asterisk indicates significant differences in susceptibility (P < 0.01) between wild type and genetically modified cells to RH1, as assessed by Student’s t test (GraphPad Software).

this was the case, the quinone-sensitivity of a T. brucei line engineered to express elevated levels of the enzyme was evaluated. In agreement with previous findings, parasites over expressing Tbntr were hyper-sensitive (approximately fourfold) to RH1 as compared to controls (Fig. 2B).

To determine whether Tbntr plays a role in activating other ABOs, the susceptibility of Tbntr over expressing parasites towards a selected group of compounds, was tested (Table 2). Out of the additional compounds screened, three (DZQ, MeDZQ and TZQ) showed the same hypersensitivity profile as RH1, with Tbntr over expressing T. brucei being up to 3.5-fold more susceptible to the aziridinyl agents than controls. In contrast, ABQ22–24 and 30 did not follow this pattern, with the Tbntr over expressers exhibiting an EC50 equivalent to that observed against control cells.

Dependent on backbone configuration, the carbonyl oxygen atoms in a quinone can undergo 1 or 2e reduction, with the former reaction carried out by a range of ubiquitous flavoproteins, including CPR, while the latter is catalysed by enzymes, such as NQO1 and type I NTRs. To determine whether the trypanocidal activity displayed by selected ABOs can undergo activation via a 1e reduction mechanism, the susceptibility of T. brucei lines engineered to express ectopic copies of the Tbcpr2 or 3 (Hall et al., 2011) to RH1, TZQ and ABQ22–24 was evaluated (Table 2). With RH1, TZQ, ABQ23 and ABQ24, the Tbcpr over expressing lines exhibited EC50 values similar to those observed using control cells, indicating that the 1e reduction pathway plays a significant role in their activation. When the T. brucei line expressing elevated levels of Tbcpr2 or 3 was tested against ABQ22, the recombinant lines were hypersensitive to this tetra-aziridinyl agent relative to controls (Fig. 2C).

Table 2. Growth inhibitory effect of ABOs towards T. brucei expressing elevated levels of potential activators.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Control</th>
<th>Tbntr11</th>
<th>Tbcpr211</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFMO</td>
<td>27500 6 108 30720 6 1330 (1.12)</td>
<td>33250 6 5123 (1.21)</td>
<td></td>
</tr>
<tr>
<td>Nifurtimox</td>
<td>2980 6 30 315 6 15 (0.11) *</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>DZQ</td>
<td>318 6 38 90 6 3 (0.28) *</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>MeDZQ</td>
<td>740 6 18 115 6 3 (0.16) *</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>RH1</td>
<td>19 6 1 3 6 1 (0.16) *</td>
<td>20 6 3 (1.05)</td>
<td></td>
</tr>
<tr>
<td>TZQ</td>
<td>225 6 26 60 6 9 (0.27) *</td>
<td>284 6 13 (1.26)</td>
<td></td>
</tr>
<tr>
<td>ABQ22</td>
<td>138 6 25 148 6 9 (1.07)</td>
<td>20 6 3 (0.14) **</td>
<td></td>
</tr>
<tr>
<td>ABQ23</td>
<td>2935 6 420 3315 6 540 (1.13)</td>
<td>3556 6 96 (1.21)</td>
<td></td>
</tr>
<tr>
<td>ABQ24</td>
<td>1470 6 120 2090 6 160 (1.42)</td>
<td>1698 6 23 (1.16)</td>
<td></td>
</tr>
<tr>
<td>ABQ30</td>
<td>7400 6 190 8670 6 1670 (1.12)</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

Data represent the EC50 values of parasites expressing wild type (control), elevated levels of Tbntr (Tbntr11) or Tbcpr2 (Tbcpr211). All values are means ± standard deviation of four independent experiments. The fold change in EC50 values of the over expressing lines relative to control is given in parentheses. Parasites over expressing Tbcpr3 behaved similarly to those expressing elevated levels of Tbcpr2.

a. Trypanocidal activity of DZQ, MeDZQ, RH1, ABQ6, TZQ and ABQ22 towards T. brucei cells over expressing Tbntr was previously reported (Hall et al., 2012).

*The difference in susceptibility of wild type and Tbntr11 lines to nifurtimox, DZQ, MeDZQ, RH1 and TZQ was statistically significant (P < 0.01), as assessed by Student’s t test.

**The difference in susceptibility of wild type and Tbcpr11 lines to ABQ22 was statistically significant (P < 0.01), as assessed by Student’s t test.
Fig. 3. RH1 resistance selection in T. brucei.
A. Selection of RH1 resistant T. brucei was performed by the stepwise increase of the ABQ (black line). The cumulative cell density of three independent cultures (grey lines) were monitored throughout the selection with only one of these (RH1\textsuperscript{R}) generating an outgrowth in medium containing >20 nM RH1. Clones from the RH1\textsuperscript{R} culture were then generated following limiting dilution.
B. Cumulative cell density of wild type T. brucei (black lines) and a clone (RH1\textsuperscript{RC1}) derived from RH1\textsuperscript{R} (grey lines) grown in medium lacking (solid lines) or containing 40 nM RH1 (dashed lines). Data points are averages ± standard deviation from experiments performed in quadruplicate. A second RH1\textsuperscript{R} derived clone (RH1\textsuperscript{RC2}) analysed in parallel displayed similar growth properties (data not shown).
C. Dose response curves of wild type T. brucei (black solid line), RH1\textsuperscript{RC1} (black dashed line) and RH1\textsuperscript{RC2} (grey solid line) to RH1. All data points are averages from experiments performed in quadruplicate ± standard deviation.
D. Susceptibility of T. brucei to RH1 as judged by the EC\textsubscript{50} values (in nM). The asterisk indicates significant differences in susceptibility (P < 0.001) between wild type and genetically modified cells to RH1, as assessed by Student’s t test (GraphPad Software).

Based on the above data, the trypanocidal activity of ABQs can be divided into compounds such as RH1 and TZQ that function as TbNTR-dependent (oxygen-insensitive quinone oxidoreductase) prodrugs, structures including ABQ22 that function as TbCPR-dependent (oxygen-sensitive quinone oxidoreductase) prodrugs and agents such as ABQ23 and 24 that are not activated by either of these two mechanisms: ABQ23 and 24 may still function as prodrugs undergoing activation by an alternative mechanism(s) to that explored here.

Trypanosomes selected for resistance toward RH1 have reduced NTR expression

A further approach to determine how ABQs mediate their trypanocidal activity involved the selection of T. brucei lines resistant to RH1. Wild type bloodstream form parasites were grown in the presence of RH1 over a 100 day period, starting with a compound concentration that inhibits T. brucei growth by 10% (5 nM) and progressively increasing this to twice the EC\textsubscript{50} (40 nM) (Fig. 3A). The cumulative cell density of selected cultures was followed throughout the experiment. Of the three replica cultures analysed, two stopped growing when the selective pressure was increased to 20 nM. In contrast, cells in the third culture remained viable in 40 nM RH1 and parasites in this population were cloned by limiting dilution (designated as RH1\textsuperscript{RC}). We then evaluated the growth properties of two RH1\textsuperscript{RC} lines (RH1\textsuperscript{RC1} and RH1\textsuperscript{RC2}) (Fig. 3B). In the absence of RH1, the doubling time of the two clones was comparable to the parental line (doubling time of 7.5 hours). When grown in medium containing RH1 (40 nM), the doubling time of the resistant clones increased 1.5-fold (doubling time of 10.5 hours); by comparison, no growth of wild type T. brucei was detected after 24 hrs. Next, we determined the susceptibility of the RH1\textsuperscript{RC} lines towards RH1 with this showing that both lines were 3.5-fold less sensitive to the ABQ than wild type (Fig. 3C and D); wild type, RH1\textsuperscript{RC1} and RH1\textsuperscript{RC2} exhibited EC\textsubscript{50} values of 25 6 6, 83 6 2 and 92 6 1 nM against RH1,
respectively. This resistance phenotype was not lost following 100 generations in medium lacking RH1. Indeed, the resistant lines were still 4.5-fold less sensitive to the ABQ than controls; wild type, RH1RC1 and RH1RC2 exhibited EC50 values of 19.6, 86.6 and 91.6 nM against RH1, respectively. When the above susceptibility studies were extended to TbNTR-activated prodrugs, cross-resistance towards nitrofurimox, benznidazole, megazol, CB1954 and LH17 was observed (Table 3). This phenotype was specific to this class of compound, as the RH1 selected cells were equally sensitive to the non TbNTR-activated compounds, including ABQ22 and other ICL inducing agents, as wild type T. brucei.

Studies aimed at deciphering the RH1 resistance mechanism demonstrated that although the Tbntr copy number was similar and Tbntr gene sequences were identical in the compound selected and wild type lines, the RH1RC1 cells had a lower (50%) Tbntr mRNA expression level (Fig. 4). Intriguingly, when extended to genes expressed in the same polycistron as Tbntr, all the tested ORFs exhibited the same trait in relation to copy number and transcript levels. The observed reduction in gene expression within the Tbntr-containing poly-cistron in the RH1 selected cells is specific to this region as additional ORFs present elsewhere in the T. brucei genome (for example Tb927.7.4940 and Tb927.7.6620, and the genes encoding for the DNA repair enzymes TbSNM1, TbMRE11, TbRAD51, TbXPG, TbEXO1, TbCSB and TbMSH3) are expressed at equivalent levels in RH1RC and control parasites.

Aziridinyl benzoquinones promote cell cycle arrest and DNA damage in trypanosomes

Aziridinyl benzoquinones mediate their toxicity against mammalian cells through formation of ICLs that promote cell cycle arrest and DNA damage (Kim et al., 2004; Dehn et al., 2005; Begleiter et al., 2007). To determine if this is the case in T. brucei, parasites at various stages in the cell cycle were identified in asynchronous cultures by staining their nuclear (N) and mitochondrial (known as the kinetoplast (K)) genomes with DAPI; a non-dividing trypanosome cell usually contains one mito-chondrion and hence one kinetoplast. The ratio of these two DNA-containing structures represents an excellent marker for the trypanosomal cell cycle, with T. brucei in the G1/S phase having a 1N1K arrangement, those in G2/M phase possessing a 1N2K ratio while cells dis-playing a 2N2K profile are in the post-M stage (Wood-ward and Gull, 1990; Siegel et al., 2008; Glover and Horn, 2012). For untreated T. brucei and cells exposed to phleomycin, most (approximately 80%) were in the G1/S stage of the cell cycle with a smaller percentage (12 to 20%) in G2/M (Fig. 5). In contrast, when trypanosomes were treated with RH1 or mechlorethamine (a well characterized ICL inducing agent), 58% and 45% of the cells displayed a configuration typical of the G2/M phase, respectively, with only 37% and 53% exhibiting a 1N1K DNA staining pattern, respectively. This indicates that treatment of T. brucei with RH1 (or mechloreth-amine) results in a G2/M cell cycle arrest phenotype, potentially as a result of damaged DNA and is consist-ent with observations made with mammalian cells (Kim et al., 2004; Dehn et al., 2005).

In most cells, ICLs are repaired through the concerted action of a number of complementary and overlapping pathways with enzymes belonging to the SNM1/PSO2 nuclease family playing a specific role. In T. brucei, the enzyme TbSNM1, a member of the SNM1/PSO2 nuclease family, has been shown to play a specific and key role in repairing such damage (Sullivan et al., 2015). To evaluate whether deletion of both copies of Tbsnm1 from the T. brucei genome altered the cells sensitivity to ABQs, the null mutant trypanosomes were grown in the presence of selected trypanocidal quinones and the EC50 values for each compound determined (Table 4; Fig. 6). For all the ABQ prodrugs tested, cells lacking the nucle-ase were up to 6.6-fold more susceptible to these
Fig. 4. Evaluating the mechanism underlying RH1 resistance in T. brucei.

A. Schematic of T. brucei chromosome 7 that contains Tbntr. This gene is part of a 200 kbp polycistronic region (see Expanded view) that encodes for 60 open reading frames (ORF) (light grey arrowed box) and is approx. 75 kbp downstream from a 'strandswitch' region, areas on the trypanosomal chromosome that are postulated to be regions where polycistronic transcription is initiated (Imboden et al., 1987; Muhich and Boothroyd, 1988). The position of selected ORFs from the Tbntr-containing polycistronic transcriptional unit plus the telomeric repeats (black arrow repeats) are shown. Other selected polycistrons on chromosome 7 are also highlighted (dark grey arrowed boxes).

B and C. The mRNA levels (panel B) and gene copy number (panel C) of selected ORFs located on chromosome 7 from wild type (grey bar) and RH1 RC1 (white bar) cells was evaluated by qPCR. This was compared against the expression level/gene copy number of a standardized control (Tbter t) and the average fold difference, as judged by 2–(ΔΔCt) from reactions performed in triplicate 6 standard deviation, plotted as a measure of the relative expression level. The asterisk and double asterisk indicates significant differences in relative mRNA expression levels (P <0.01 and <0.02, respectively) between the wild type and RH1 RC1, as assessed by Student’s t test (GraphPad Software).

Linking prodrug activation with DNA damage

To establish whether there is a link between activation of the trypanocidal ABQs and the TbSNM1-mediated DNA repair pathway, the susceptibility of Tbsnm1 null-mutants expressing ectopic copies of Tbntr or Tbcpr2 towards selected compounds was evaluated (Fig. 5). When these lines were treated with RH1 or TZQ, an
increased susceptibility was observed relative to wild type controls (Fig. 7A); Tbsnm1 null-mutants and T. brucei over expressing Tbtr were threefold and fivefold more susceptible to RH1, respectively, and 2.7- and 3.4-fold more susceptible to TZQ than wild type T. brucei. For Tbsnm1 null parasites expressing an ectopic copy of Tbtr, this increase in potency was magnified further, with these cells showing a 15-fold increase in susceptibility to both compounds as compared against wild type. Extending these studies to investigate the susceptibility of Tbsnm1 null-mutant parasites engineered to express an ectopic copy of Tbcpr2 towards ABQ22 yielded similar results (Fig. 7B). When Tbsnm1 null mutants or wild type cells engineered to express an ectopic copy of Tbcpr2 were treated with ABQ22 a sevenfold increase in susceptibility was observed by both lines compared against controls. The growth inhibitory effect of this compound was further enhanced in trypanosomes that lack Tbsnm1 activity while expressing elevated levels of Tbcpr2 (Fig. 7B), with these cells being 12.5-fold more sensitive to ABQ22 than wild type. The above alterations in susceptibilities were specific to the ABQs as all lines tested exhibited equivalent sensitivities to the non-DNA damaging trypanocidal agent DFMO.

### Table 4. Susceptibility of T. brucei Tbsnm1 null mutants to ABQ compounds.

<table>
<thead>
<tr>
<th>Compound</th>
<th>EC50 (nM) Control</th>
<th>DTbsnm1</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nifurtimox</td>
<td>2850 6 20</td>
<td>2250 6090</td>
<td>0.79</td>
</tr>
<tr>
<td>DFMO</td>
<td>27500 6 108</td>
<td>27100 8850</td>
<td>0.99</td>
</tr>
<tr>
<td>DZQ</td>
<td>157 6 23</td>
<td>43 62*</td>
<td>0.27</td>
</tr>
<tr>
<td>RH1</td>
<td>19 6 1</td>
<td>5.61*</td>
<td>0.26</td>
</tr>
<tr>
<td>TZQ</td>
<td>253 6 33</td>
<td>46.61*</td>
<td>0.18</td>
</tr>
<tr>
<td>ABQ22</td>
<td>138 6 25</td>
<td>21 65*</td>
<td>0.15</td>
</tr>
<tr>
<td>ABQ23</td>
<td>2300 6 60</td>
<td>2410 6160</td>
<td>1.01</td>
</tr>
<tr>
<td>ABQ24</td>
<td>1470 6 120</td>
<td>1130 6100</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Data represent the EC50 values of selected ABQ compounds towards wild type T. brucei (control) and T. brucei Tbsnm1 null mutants (DTbsnm1). All values are means ± standard deviation of four independent experiments. The ration of EC50 values between the two parasite lines are given. *Indicates significant differences in susceptibility (P < 0.01) between the wild type and DTbsnm1 lines, as assessed by Student’s t test (GraphPad Software).

Discussion

ICLs represent a highly toxic form of DNA damage that blocks processes dependent on DNA strand separation, such as replication and transcription. As rapidly dividing cells are particularly susceptible to this type of lesion, agents that are able to promote ICLs are of interest in the treatment of cancer and infectious diseases (Chen and Hu, 2009; Deans and West, 2011; Wilkinson et al., 2011). Here, we report that several ABQs, ICL inducing chemicals originally developed as anti-cancer agents, can also function as highly potent anti-trypanosomatid prodrugs with different structures able to undergo activation via different pathways. In these cases, activation promotes a type of DNA damage that is repaired by a pathway where SNM1, a nuclease that only functions in ICL repair, plays a key role.

Invariably, aziridine-containing compounds function as prodrugs with activation requiring protonation of the aziridinyl nitrogen. The ease with which this reaction can occur is governed by the compound’s chemical composition and how this influences the pKa of the aziridine. For certain structures (e.g., triethylenemelamine, thio-TEPA and ‘simple’ ABQs including ABQ11) protonation can readily take place in aqueous solution due to the relatively high pKa of the aziridine within such back-bones (Akhtar et al., 1975). However, for other agents, the pKa of the heterocycle is comparatively low.

Fig. 6. Susceptibility of T. brucei lines expressing altered levels of TbSNM1 to selected azirindinyl benzoquinones. Growth inhibitory effects, as judged by EC50 values, of RH1, TZQ and ABQ22 (all in nM) in the T. brucei wild type, Tbsnm1 null mutant (DTbsnm1) and Tbsnm1 null mutants expressing an ectopic copy of Tbsnm1 (DTbsnm1 Tbsnm11) lines (Sullivan et al., 2015). DFMO (in nM) was used as a non-DNA damaging agent control. Data are mean values ± standard deviations from experiments performed in quadruplicate. The asterisk indicates significant differences in susceptibility (P <0.001) between wild type and DTbsnm1 or between DTbsnm1 and DTbsnm1 Tbsnm11 cells to RH1, as assessed by Student’s t test (GraphPad Software).
Fig. 7. Linking aziridinyl benzoquinone prodrug activation to DNA damage.

A. Growth inhibitory effects, as judged by EC_{50} values, of RH1, TZQ, ABQ22 (all in nM) and DFMO (in uM) in the T. brucei wild type, T. brucei expressing an ectopic copy of Tbntr (Tbntr^{11}), Tbsnm1 null mutant (DTbsnm1) and Tbsnm1 null mutants expressing elevated levels of Tbntr (DTbsnm1 Tbntr^{11}) lines. The asterisk and double asterisk indicates significant differences in susceptibility (P < 0.001 and <0.03, respectively) between the DTbsnm1 Tbntr^{11} line relative to the Tbntr^{11} or DTbsnm1 parasites to RH1 and TZQ, as assessed by Student’s t test (GraphPad Software).

B. Growth inhibitory effects, as judged by EC_{50} values, of RH1, TZQ and ABQ22 (all in nM) in the T. brucei wild type, T. brucei expressing an ectopic copy of Tbcpr2 (Tbcpr2^{11}), Tbsnm1 null mutant (DTbsnm1) and Tbsnm1 null mutants expressing elevated levels of Tbcpr2 (DTbsnm1 Tbcpr2^{11}) lines: the latter lines was validated as previous described (Hall et al., 2011; Sullivan et al., 2015). DMFO (in uM) was used as a non-DNA damaging agent control. The asterisk and double asterisk indicates significant differences in susceptibility (P < 0.01 and P < 0.003, respectively) between the DTbsnm1 Tbcpr2^{11} line relative to the Tbcpr2^{11} or DTbsnm1 parasites to ABQ22, as assessed by Student’s t test (GraphPad Software).

Data in panels A and B are mean values 6 standard deviations from experiments performed in quadruplicate.

therefore preventing the above activation event. To circumnavigate this issue, the aziridine motif can be incorporated into a chemical backbone that also contains a bioreductive moiety, such as a quinone or nitroaromatic. Here, enzymatic reduction of carbonyl oxygen or nitro groups triggers a redistribution of electrons around the associated cyclical structure that can raise the pKa of the aziridinyl ring, and thus facilitating nitrogen protonation (Fig. 8A) (Hargreaves et al., 2000).

Using a combination of a whole genome loss-of-function screen, selection of resistant lines and/or functional genomic approaches, we have shown that T. brucei expresses at least two ABQ bioreductive sys-tems. For the eight compounds evaluated as potential prodrugs, half (DZQ, MeDZQ, RH1 and TZQ) were shown to be NTR activated, one (ABQ22) was shown to be dependent on a CPR activity, while the remainder (ABQ23, 24 and 30) were not reduced by any of the systems tested. For those that appeared not to function as prodrugs, this may be due to their structures, as these compounds contain methyl substituted aziridines (Fig. 8B). These alkyl groups may block protonation and thus activation of the aziridinyl nitrogen by lowering this motif’s pKa and/or hinder access of oxidoreductases to the carbonyl groups (Lusthof et al., 1989; Phillips et al., 1999). These possibilities may also account for why other compounds that contain methyl/dimethyl aziridine substitutions (e.g., ABQs 25–30 and 31–34) consistently displayed no or a lower activity against T. brucei than their non-substituted counterparts, mirroring observations made in high throughput oncological screens (Kim et al., 2016).

For those compounds that do function as trypanocidal prodrugs, the nature of the side chains found at the 3,6 positions on the benzoquinone ring appears to influence which activation mechanism predominates (Fig. 8B). For structures possessing small substituents at one or both sites, reduction occurs via a type I NTR activity with this mechanism unaffected by the presence of equivalent (e.g., hydrogen for DZQ, methyl for MeDZQ) or divergent (e.g., hydroxymethyl and methyl in RH1, aziridine and hydrogen in TZQ) groupings at the 3,6 positions. Such small side chains may aid type I NTR/compound interaction or help the prodrug (or its activated products)
Antiparasitic activities of aziridinyl benzoquinones

Fig. 8. Proposed mechanism of action and trypanocidal structure activity relationships of ICL-forming ABQs.

A. Quinones can be reduced to hydroquinones by two distinct pathways. In one pathway, enzymes such as cytochrome P450 reductase (CPR) mediate the 1e⁻ reduction of the quinone to a semiquinone radical. This then undergoes further reduction to the hydroquinone. In the presence of O₂, the semiquinone can undergo futile cycling to form O₂⁻² and the parental quinone. In the other system, reduction of the quinone to the hydroquinone occurs directly via a 2e⁻ reduction event as typified by a type I nitroreductase (NTR) activity. Formation of the hydroquinone then facilitates protonation of the nitrogen atoms in the aziridine rings with these electrophilic intermediates undergoing nucleophilic attack (e.g., by DNA), promoting opening of the aziridinyls, formation of aliphatic amine intermediates and alkylation of the target. As ABQs contain two (or more) aziridinyls they can covalently bind the two complementary DNA strands together resulting in an interstrand crosslink.

B. ABQs (e.g., DZQ, MeDZQ, RH1 and TZQ) possessing small (e.g., hydrogen, methyl, hydroxymethyl) substituent groups (small solid spheres) at the 3,6 positions on a benzoquinone core are more likely to be activated in T. brucei by the type I nitroreductase (NTR) but not by cytochrome P450 reductase (CPR) (i). ABQs that contain methyl substituted (small solid spheres) aziridines (e.g., ABQ23, 24 and 30) are not activated in T. brucei by NTR or CPR (ii). ABQ22, a compound that contains aziridinyl rings at the 3,6 positions on the benzoquinone core, is readily activated within the trypanosome by CPR but not by NTR (iii). ABQs (e.g., ABQ9–19) having bulky substituent groups (large solid spheres) at the 3,6 positions on the benzoquinone ring generally display no significant trypanocidal activity (iv).

Gain access to the site(s) of action. In contrast, the presence of bulkier groups (e.g., aziridines for ABQ22) at both sites results in CPR activation. Why this mechanism operates to reduce this particular compound is open to speculation but it may reflect a steric hindrance effect such that the four aziridine groups found in ABQ22 preclude TbNTR from interacting with and reducing its quinone carbonyl groups while facilitating associations with TbCPR isoforms. Alternatively, this difference could be due to subcellular localisation such that ABQ22 may be readily transported into the T. brucei endoplasmic reticulum, the organelle where the TbCPRs are believed to reside, but unable to gain access to the parasite’s mitochondrion, the site where TbNTR is found (Wilkinson et al., 2008).

Following protonation of the aziridinyl nitrogen, the resultant electrophilic intermediate undergoes nucleophilic attack leading to opening of the heterocyclic ring, formation of an aliphatic amine then alkylation of the target (Fig. 8A) (Hargreaves et al., 2000). If the targeting compound contains multiple aziridinyl groups, as is the case for the ABQs tested here, and the nucleophile is DNA, then two nucleobases within the nucleic acid double helix become covalently crosslinked. In many eukaryotes the SNM1 (also known as PSO2) nuclease plays a key role in resolving such DNA damage, with cells lacking this activity being specifically and highly susceptible to ICL inducing agents (Cattell et al., 2010). By phenotypically screening a selected set of trypanocidal ABQs against a T. brucei snm1 null mutant line, we demonstrated that trypanosomes lacking this DNA repair enzyme were more susceptible to NTR and CPR activated prodrugs relative to wild type. Moreover, in parasites where TbSNM1 was absent but the activation mechanism (either TbNTR or TbCPR) is elevated, this susceptibility phenotype was exacerbated, providing a direct link between the two oxidoreductase activities and DNA damage. Extending these assays to investigate the trypanocidal action of ABQs that do not undergo NTR or CPR activation, revealed that these do not promote a form of DNA damage that can be repaired through an SNM1-dependent pathway. It is plausible that these agents may still function as monofunctional or bifunctional DNA alkylating agents, leading to formation of a single site adduct or intrastrand crosslink or could act as modifiers of other biological molecules via their aziridinyl or quinone ring structures.

The continuous in vitro culturing of bloodstream form T. brucei in the presence of RH1 generated a parasite line that displayed cross-resistance to a range of NTR-activated prodrugs. Analysis of clones revealed that these cells had reduced TbNTR expression. Additionally, lower mRNA levels were also noted for other open reading frames across the 200 kb TbNtr-containing
polycistron; however, this effect did not extend to the surrounding transcriptional units. In contrast to trypanosomes selected for resistance towards other NTR activated prodrugs such as nifurtimox or fexinidazole (Wilkinson et al., 2008; Wyllie et al., 2016), the copy number of all tested open reading frames in the Tbntr-containing transcriptional unit were equivalent to wild type, indicating that the observed reduction in gene expression was not due to a large (>10 kb) DNA deletion event; the ‘strandswitch’ region, areas on the trypanosomal chromosome postulated to be regions where polycistronic transcription is initiated (Imboden et al., 1987; Muhich and Boothroyd, 1988), between the Tbntr-containing and adjacent transcriptional units is 6 kb. Although the exact molecular basis that leads to RH1 resistance has yet to be deciphered, it is plausible that the transcription initiation site of one of the two Tbntr-containing polycistrons (T. brucei has a diploid nuclear genome) may have been compromised resulting in the reduction in mRNA levels. We have demonstrated that ABQs display significant growth inhibitory activities against T. brucei, T. cruzi and L. major, with two compounds (RH1 and TZQ) exhibiting high potency (EC50 values < 100 nM) against all three pathogens. However, this is accompanied by toxicity against macrophage (like) mammalian cells, a trait previously noted in primary mouse splenocytes (Miliukiene et al., 2014). This unwanted activity may be because such immune cells have the ability to constitutively express or to upregulate quinone oxidoreductases including NQO1 and NQO1-independent activities that promote oxidative stress and/or DNA damage (Tudor et al., 2005; Potts-Kant et al., 2012; Miliukiene et al., 2014). Although these cytotoxicity issues may preclude the use of these ABQs as therapies for the treatment of systemic infections, understanding how they mediate their trypanocidal activities can inform drug development. By exploiting the chemical features that facilitate NTR or CPR specific activation such as those reported here, novel trypanocidal prodrugs could be generated. Following their activation, these compounds could be designed to release cytotoxic metabolites that affect a range of biological processes potentially targeting different biochemical pathways. Such agents would be best used in a polytherapy, as is the case with the anti-HAT nifurtimox-eflornithine combination therapy, in order to minimize potential resistance.

### Experimental procedures

#### Compounds

The compounds used in this study were obtained from the following sources: The aziridinyl benzoquinones (DZQ, MeDZQ, AZQ, TZQ and ABQs 4–6, 19, 22–34) (Table 5)

---

Table 5. Structure of aziridinyl 1,4-benzoquinones used in this study.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Structure</th>
<th>Compound</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABQ1 (DZQ)</td>
<td>R1=5R2=5X5YSYH</td>
<td>ABQ18</td>
<td>R1=5R2=5NHCO(CH3)2;X5YSYH</td>
</tr>
<tr>
<td>ABQ2 (MeDZQ)</td>
<td>R1=5R2=5SCH3;X5YSYH</td>
<td>ABQ19</td>
<td>R1=5R2=5NHCOPhenyl;X5YSYH</td>
</tr>
<tr>
<td>ABQ3</td>
<td>R1=5R2=5Br;X5YSYH</td>
<td>ABQ20 (AZQ)</td>
<td>R1=5R2=5NHCO(O)CH3;X5YSYH</td>
</tr>
<tr>
<td>ABQ4</td>
<td>R1=5R2=5S;X5YSYH</td>
<td>ABQ21 (TZQ)</td>
<td>R1=5azirindyl;R2=5X5YSYH</td>
</tr>
<tr>
<td>ABQ5 (RH1)</td>
<td>R1=5CH2OH;R2=5CH3;X5YSYH</td>
<td>ABQ22</td>
<td>R1=5R2=5azirindyl;X5YSYH</td>
</tr>
<tr>
<td>ABQ6</td>
<td>R1=5R2=5OCH2;X5YSYH</td>
<td>ABQ23</td>
<td>R1=5R2=5X5HS;Y5CH3 (r orientation)</td>
</tr>
<tr>
<td>ABQ7</td>
<td>R1=5R2=5OCH3;X5YSYH</td>
<td>ABQ24</td>
<td>R1=5R2=5X5HS;Y5CH3 (s orientation)</td>
</tr>
<tr>
<td>ABQ8</td>
<td>R1=5R2=5SO(CH2)2;X5YSYH</td>
<td>ABQ25</td>
<td>R1=5R2=5Br;X5HS;Y5CH3</td>
</tr>
<tr>
<td>ABQ9</td>
<td>R1=5R2=5Nh2;X5YSYH</td>
<td>ABQ26</td>
<td>R1=5R2=5Br;X5YSYH</td>
</tr>
<tr>
<td>ABQ10</td>
<td>R1=5R2=5NHCH3;X5YSYH</td>
<td>ABQ27</td>
<td>R1=5R2=5Cl;X5HS;Y5CH3</td>
</tr>
<tr>
<td>ABQ11</td>
<td>R1=5R2=5NHCH2CH3;X5YSYH</td>
<td>ABQ28</td>
<td>R1=5R2=5Cl;X5YSYH</td>
</tr>
<tr>
<td>ABQ12</td>
<td>R1=5R2=5N(H)(CH2)2;X5YSYH</td>
<td>ABQ29</td>
<td>R1=5R2=5F;X5HS;Y5CH3</td>
</tr>
<tr>
<td>ABQ13</td>
<td>R1=5R2=5N(H)CH3;X5YSYH</td>
<td>ABQ30</td>
<td>R1=5R2=5F;X5YSYH</td>
</tr>
<tr>
<td>ABQ14</td>
<td>R1=5morpholyl;R2=5F;X5YSYH</td>
<td>ABQ31</td>
<td>R1=5R2=5NHCOCH3;X5H;Y5CH3</td>
</tr>
<tr>
<td>ABQ15</td>
<td>R1=5R2=5NHCOCCH3;X5YSYH</td>
<td>ABQ32</td>
<td>R1=5R2=5NHCOCH3;X5YSYH</td>
</tr>
<tr>
<td>ABQ16</td>
<td>R1=5R2=5NHCOCCH2;X5YSYH</td>
<td>ABQ33</td>
<td>R1=5R2=5NHCOCH2;X5H;Y5CH3</td>
</tr>
<tr>
<td>ABQ17</td>
<td>R1=5R2=5NHCOCCH2;X5YSYH</td>
<td>ABQ34</td>
<td>R1=5R2=5NHCOOCH3;X5YSYH</td>
</tr>
</tbody>
</table>

All compounds tested satisfy Lipinski’s Rule of 5 (http://pubchem.ncbi.nlm.nih.gov/) although all are predicted to have toxicity risks (http://www.organic-chemistry.org/prog/peo).
were supplied by the Drug Synthesis and Chemistry Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, United States, while RH1 was donated by Prof Frank Guziec Jr (Southwestern University, TX). The chemical, physical and potential toxicological properties of each compound can be accessed via the PubChem database (http://pubchem.ncbi.nlm.nih.gov/) or OSIRIS Property Explorer software (http://www.organic-chemistry.org/prog/). Nifurtimox was obtained from Prof Simon Croft (London School of Hygiene and Tropical Medicine, UK), DFMO from Prof Mike Barrett (University of Glasgow, UK), mechloroethamine HCl from Stratech Scientific, while puromycin, blastidicin, phleomycin, hygromycin b and tetracycline were purchased from Melford Laboratories.

Cell culturing

Bloodstream form T. brucei brucei (MiTat 427 strain; clone 221a) cells were cultured in HMI-9 media (Life Technologies) supplemented with 36 mM sodium bicarbonate, 0.014% (v/v) b-mercaptoethanol and 10% (v/v) heat-inactivated foetal calf serum (GE Healthcare) as described (Hirumi and Hirumi, 1989). A cell line (2T1) engineered to constitutively express the tetracycline repressor protein was grown in the modified HMI-9 medium, supplemented with 1 µg ml⁻¹ phleomycin (Alsford et al., 2005). T. brucei Tbntr heterozygote lines were cultured in the presence of 10 blasticidin against wild-type parasites expressing Tbntr, Tbcpr2 or Tbcpr3 were maintained in the modified HMI-9 medium supplemented with 1 lg ml⁻¹ phleomycin, 2.5 lg ml⁻¹ hygromycin b and, where appropriate, 1 lg ml⁻¹ tetracycline to induce protein expression (Hall et al., 2011). T. brucei lines expressing altered levels of the DNA repair enzyme TsSNM1 were grown in the presence of 1 lg ml⁻¹ phleomycin, 10 lg ml⁻¹ blastidicin and/or 2 lg ml⁻¹ puromycin, as appropriate (Sullivan et al., 2015).

The epimastigote form of T. cruzi was cultured in RPMI-1640 medium (Lonza) supplemented with 5 g l⁻¹ trypticine, 5: g l⁻¹ HEPES pH 8.0, 20 mg l⁻¹ haemin, 0.34 g l⁻¹ sodium glutamate, 0.22 g l⁻¹ sodium pyruvate, 2500 U l⁻¹ penicillin, 0.25 g l⁻¹ streptomycin (all Sigma-Aldrich) and 10% (v/v) heat-inactivated foetal calf serum (GE Healthcare) at 25°C (Kendall et al., 1990). The protomastigote form of L. major MHOM/IL/80/Friedlin was cultured at 25°C in modified M199 medium (Invitrogen) supplemented with 4 mM sodium bicarbonate, 40 mM HEPES pH 7.4, 0.1 mM adenine, 0.005% (w/v) haemin (all Sigma-Aldrich), 25000 U l⁻¹ penicillin, 25 mg l⁻¹ streptomycin (GE Healthcare) and 10% (v/v) heat-inactivated foetal bovine serum (GE Healthcare) (Kapler et al., 1990).

The human acute monocytic leukemia (THP-1) cell line was grown at 37°C under a 5% (v/v) CO₂ atmosphere in RPMI-1640 medium (Lonza) supplemented with 2 mM pyruvate, 2 mM sodium glutamate, 2.5 U ml⁻¹ penicillin and 2.5 lg ml⁻¹ streptomycin, 20 mM HEPES pH 7.4 and 10% (v/v) heat-inactivated foetal calf serum (GE Healthcare). Differentiation of THP-1 toward macrophage-like cells was carried out using 20 ng ml⁻¹ phorbol 12-myristate 13-acetate (Sigma-Aldrich) (Rovera et al., 1979).

Selective screening of the T. brucei RNAi library

Selection of the T. brucei RNAi library was performed as described (Glover et al., 2015). Parasites (5 3 10⁶ cells) were treated for 24 hours with 1 lg ml⁻¹ tetracycline to induce RNAi prior to addition of 30 nM RH1. Periodically, the cell density of the culture was determined and when appropriate, diluted in growth medium containing fresh RH1 (30 nM) and tetracycline (1 mg ml⁻¹). Once an outgrowth population had been selected, genomic DNA was prepared for RNAi target identification (Alsford et al., 2012).

Selection of laboratory generated RH1-resistant T. brucei

Wild type bloodstream form T. brucei seeded at 1 3 10⁴ cells ml⁻¹, were cultured in the presence of sub-lethal concentration of RH1 (5 nM). Over a 100 day period, the ABQ concentration in the culture media was increased in a step-wise manner reaching a final concentration of 40 nM. The resulting line (designated as RH1 R) was cloned by limiting dilution in the presence of RH1 and the phenotype of two clones (RH1RC1 and RH1RC2) analysed.

Parasite growth inhibition assays

All growth inhibition assays were carried out in a 96-well plate format (ThermoScientific). Bloodstream form T. brucei, L. major promastigotes and T. cruzi epimastigotes parasites in the logarithmic phase of growth were seeded at 1 3 10⁵, 5 3 10⁵ and 5 3 10⁶ cells ml⁻¹ respectively in 200 ll growth medium containing different concentrations of the compound under study. After incubation at 37°C for 3 days (T. brucei), 25°C for 6 days (L. major) or 25°C for 14 days (T. cruzi), resazurin (Sigma-Aldrich) was added to each well at a final concentration of 12.5 mg ml⁻¹ (or 2.5 mg per well). The plates were further incubated at 37°C for 8 hours (T. brucei), 25°C for 24 hours (L. major) or 25°C for 24–48 hours (T. cruzi) before measuring the fluorescence of each culture using a Gemini Fluorescent Plate reader (Molecular Devices) set at kEX 5 530 nm and kEM 5 585 nm with a fil-ter cut off at 550 nm. The change in fluorescence resulting from the reduction of resazurin is proportional to the num-ber of live cells. A compound’s EC₅₀ value was established using the non-linear regression tool on GraphPad Prism (GraphPad Software) and the statistical significance of any differences in parasite susceptibilities assessed using the Student’s t test calculator (GraphPad Software).

Mammalian inhibition assays

Differentiated THP-1 cells seeded at 2.5 3 10⁴ cells ml⁻¹ were incubated for 3 days at 37°C in a 5% (v/v) CO₂ atmosphere in 200 ll modified RPMI-1640 growth medium containing various concentrations of the compound under
study. Resazurin (2.5 mg per well) was added to each culture and the plates incubated for a further 8 hours before determining the fluorescence of each sample. Experiments involving animals were approved by the Animal Welfare and Ethics Review Board at LSHTM and con-ducted under license in accordance with the Animals (Scientific Procedures) Act 1986 (UK Home Office Project Licence PPL70/6997). Peritoneal macrophages were isolated from BALB/c mice by lavage 24 hours after intraperitoneal injection of 2% (w/v) soluble starch (Sigma). Macrophages (40,000) in RPMI-1640 medium (Lonza) sup-plemented with 10% (v/v) heat-inactivated foetal calf serum (GE Healthcare) were adhered overnight at 37°C in a 5% (v/v) CO₂ atmosphere onto a 16-well chamber slide. Cells were then treated with various concentrations of the compound under study and the cultures incubated at 37°C up to 5 days under a 5% (v/v) CO₂ atmosphere. Untreated controls were analysed in parallel on each slide. Macro-phages were fixed in 100% (v/v) methanol, stained with 10% (w/v) Giemsa, visualized with Leica DMRA2 microscope using a 3100 oil immersion objective and images captured using a Retiga EXi Fast 1394 digital cam-era. The number of cells per field of view was determined with a minimum of 10 fields examined for each drug treat-ment. The average number of cells per field of view was determined and expressed as a % relative to untreated controls. All treatments were performed in quadruplicate.

Quantitative PCR (qPCR)

The copy number and mRNA levels of various genes were analysed by qPCR. Genomic DNA or total RNA was prepared from wild type or RH1-resistant T. brucei using the DNeasy Blood and Tissue or RNeasy Extraction kits (Qiagen), respectively. cDNA was synthesized from total RNA using the Superscript VILO cDNA synthesis kit (Invitrogen). All qPCRs were performed in triplicate with the PerfeCTa qPCR FastMix kit (Quanta Biosciences) on a CFX96 Touch Real-Time PCR Detection System (BioRad). Fluorescence data was collected using the CFX Manager Software (BioRad), the products analysed by a melt curve after the final cycle and normalized against telomerase reverse transcrip-tase (Tbtert; Tb11.01.10190) using the comparative CT method (Schmittgen and Livak, 2008; Brenndorfer and Boshart, 2010).

Cell cycle arrest assay

Wild type T. brucei (5 3 10⁵ ml⁻¹) were cultured for 16 hours in the presence of mutagen (1 mg ml⁻¹ phenolycin; 35 nM meclothorane; 100 nM RH1). Cells were washed in phosphate buffered saline (PBS), fixed in 2% (w/v) para-formaldehyde in PBS and washed once more in PBS. Ali-quotsof the cell suspension (10⁵ cells) were air-dried onto microscope slides and the samples mounted in Vectashield containing 4, 6-diamidino-2-phenylindole (Vector Laboratories). Images were captured using Leica DMRA2 epi-fluorescent microscope fitted with a digital cam-era (Hamamatsu Photonics).

Acknowledgements

The Wellcome Trust provided funding to DH (Project grant 093010/B/10/Z, Investigator Award 100320/Z/12/Z) and SA (Institutional Strategic Support Fund fellowship). KS was sup-portated by a grant jointly funded by the Medical Research Council and the Department for International Development (grant MR/ J008702/1). ELM and AK were recipients of a BBSRC Doctorial Training Studentship and QMUL College studentship, respectively.

References


1. Introduction

The Trypanosomatida represent an order of parasitic protozoa be-longing to the Class Kinetoplastidae [1]. They are responsible for se-veral infections in humans with Trypanosoma brucei and Trypanosoma cruzi causing human African trypanosomiasis (HAT) and Chagas dis-ease, respectively. In addition, more than 20 Leishmania species can trigger a spectrum of pathologies collectively termed leishmaniasis. Spread by the hematopha-gous habits of in-sect vectors, these pathogens are endemic throughout tropical and sub-tropical regions of the world and cause more than 55,000 deaths per year (http://www.dndi.org/).

All organisms living in an aerobic environment are exposed to a range of reactive oxygen species (ROS) primarily generated as by-pro-ducts of respiration. These can readily react with various macro-molecules leading to formation of other toxic metabolites and/or da-mage to the target. One target that ROS have a considerable effect on are pro-teins, resulting in oxidation of certain residues such as cysteine, histidine, tyr-o-sine, and methionine (Met), which can reducing equiva-lents into change and, in some cases, modification of protein function [2,3]. Oxidation of Met produces methionine sulfoxide (Met(SO)) that exists as a mixture of two epimers, methionine-(S)-sulfoxide (Met(SO)) and methionine-(R)-sulfoxide (Met(R)O). To combat the potentially deleterious effect of these diastereomers, cells express several unrelated enzymes, known as methionine sulfoxide re-ductases (MSRs), that cat-alyse reduction of MetSO back to Met [4,5]. These distinct activities can be distinguished based on sequence and substrate specificity. Metabo-lism of both free and protein-bound Met(S)SO is performed by me-thionine sulfoxide reductase A (MSRA) [6,7] while reduction of free or protein-bound Met(R)O is mediated by free methionine-(R)-sulfoxide reductase (fRMSR) or methionine sulfoxide reductase B (MSRB), re-spectively [8–10]. The activity of these enzymes is generally driven by a thioredoxin/thioredoxin reduc-tase-depending redox cascade that facil-itate transfer of reducing equivalents from NADPH to the MSR, al-though other molecules such as metallothionein or glutaredoxin, may also act as source of reductant [8,11–14]. The importance of these pathways has been demonstrated as organisms lacking MSR(s) are more susceptible to oxidative stress, often have a shortened life span and, in the case of bacterial pathogens, reduced virulence [7,15–19].

In contrast to their mammalian hosts, trypanosomatids lack a raft of enzymes including Cu/Zn- and Mn-superoxide dismutase (SOD), cata-lase, selenium-dependent glutathione peroxidase, glutathione reductase and thioredoxin reductase, activities that help maintain the redox bal-ance in many other eukaryotic cells. Instead, they express alternative mechanisms that fulfil the above activities, with many of these
proposed as potential targets for chemotherapeutic intervention. For example, removal of superoxide anions by trypanosomes and leishmania parasites is exclusively catalysed by Fe-SOD, enzymes normally found in bacteria, lower eukaryotes and the chloroplasts of plants, while maintenance of the intracellular thiol redox homeostasis is centred upon the trypanosomatid-specific molecule trypanothione [20–22]. In a pathway analogous to the glutathione (GSH)/glutathione reductase system, trypanothione is maintained in its reduced, dithy-drotypanothione (T(SH)2) form by the activity of a NADPH-dependent flavoprotein trypanothione reductase (TR) [23,24]. T(SH)2 then drives a series of two component cascades, facilitating flux of reducing equivalents into tryparedoxins (TXNs), GSH or ascorbate which in turn reduce various peroxidases, reductases, glyoxalases and transferases [25,26].

One recipient of the electron flux via the T(SH)2/TXN cascade is MSRA, with analysis of the T. cruzi and Leishmania major homologues revealing its role as an anti-oxidant enzyme within the parasite [27,28]. Here, we report the dissection of two MetSO metabolising pathways expressed by T. brucei focusing on the biochemical properties, sub-cellular localisation, and functional importance of its MSRA and MSRB complement.

2. Materials and methods

2.1. Parasites

Bloodstream form (BSF) T. brucei SMB and 2T1 trypomastigotes that constitutively express the tetracycline repressor protein were grown at 37 °C under a 5% (v/v) CO2 atmosphere in modified Iscove’s medium containing 2.5 μg ml−1 G418 (SMB) or 1 μg ml−1 phleomycin (2T1) [29–31]. Transformed parasites were maintained in this growth medium supplemented with 2.5 μg ml−1 hygromycin. DNA and total RNA were extracted from parasites using the DNeasy mini kits (Qiagen), respectively. T. brucei transgenic clones were grown at 37 °C with aeration. Extracts were prepared as previously described [32].

2.2. Protein purification

DNA fragments containing the full length coding sequence of TbMSRA and a version of TbMSRB lacking its 5’ (130 bp) region were amplified from T. brucei genomic DNA using the primer combinations TbMSRA-1/TbMSRA-2 or TbMSRB-1/TbMSRB-2 (Table 1).

Table 1

Oligonucleotides used in this study. The sequences in lower case italics correspond to restriction sites incorporated into the primers to facilitate cloning.

<table>
<thead>
<tr>
<th>Function</th>
<th>gene</th>
<th>Primer name</th>
<th>sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein expression</td>
<td>TbmsrA</td>
<td>TbMSRA-1</td>
<td>aagggatctGAACCCAAATGCTGGCTA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TbMSRA-2</td>
<td>ggaagctttCATTCCACAGTAGAGACGTT</td>
</tr>
<tr>
<td></td>
<td>TbmsrB</td>
<td>TbMSRB-1</td>
<td>aagggatctGAACACCTGGCAGAATAGAGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TbMSRB-2</td>
<td>ggaagctttTACTTCTCGGATTGAAAACG</td>
</tr>
<tr>
<td>RNAi</td>
<td>TbmsrA</td>
<td>TbMSRA-3</td>
<td>aagggatctACTTCTCGGATTAAAAACG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TbMSRA-4</td>
<td>aacagtgaATACCCATTGGGGTTTTCC</td>
</tr>
<tr>
<td></td>
<td>TbmsrB</td>
<td>TbMSRB-3</td>
<td>aagggatctGCCGCTCTTATTTTCTTCG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TbMSRB-4</td>
<td>aacagtgaCCGCTGGGTAGGTTGGATT</td>
</tr>
<tr>
<td>Localisation</td>
<td>TbmsrA</td>
<td>TbMSRA-5</td>
<td>aacagtgaATATACCCACTGCTGCTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TbMSRA-6</td>
<td>aacagtgaATATACCCACTGCTGCTG</td>
</tr>
<tr>
<td></td>
<td>TbmsrB</td>
<td>TbMSRB-5</td>
<td>aacagtgaATATACCCACTGCTGCTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TbMSRB-6</td>
<td>aacagtgaATATACCCACTGCTGCTG</td>
</tr>
<tr>
<td>qPCR</td>
<td>TbmsrA</td>
<td>TbMSRA-7</td>
<td>TAGATGAGTCAATGGGTTAAC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TbMSRB-7</td>
<td>AGGAAATTTTATGCAATCGG</td>
</tr>
<tr>
<td></td>
<td>Thbet</td>
<td>ThBet-F</td>
<td>AGGAAATTTTATGCAATCGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ThBet-R</td>
<td>GAGGCTGTCGCCGCAAAGGG</td>
</tr>
</tbody>
</table>

respectively. The products were digested with BamHI/HindIII (TbmsrA) or BglIII/HindIII (TbmsrB) and cloned into the BamHI/HindIII sites of the expression vector pTrcHis-C (Invitrogen).

Protein expression and purification were conducted as previously described [27]. Overnight cultures of E. coli BL21 (DE3) transformed with the expression plasmid were diluted 1/100 in Terrific Broth (12 g l−1 peptone, 24 g l−1 yeast extract, 4 ml l−1 glycerol, 2.3 g l−1 KH2PO4, 12.5 g l−1 K2HPO4, pH 7.0) supplemented with 100 μg ml−1 ampicillin and grown to exponential phase at 37 °C with aeration. Pre-extraction of the HIS-tagged recombinant protein was induced with 0.5 mM IPTG, followed by incubation at 25 °C. After 4 h, cells were harvested and bacterial pellets stored at −20 °C. Purification of recombiant protein was performed using a Ni2+-HiTrap column (GE Healthcare). Briefly, the bacterial pellet was resuspended in binding buffer (20 mM Tris.HCl, pH 7.5, 10 mM imidazole and 400 mM NaCl) and disrupted by sonication. The lysate was centrifuged (10,000 g, 30 min) to remove cell debris. The resultant crude extract was loaded onto a Ni2+-HiTrap column (1 ml) previously equilibrated with binding buffer. After washing with 15 bed volumes of binding buffer plus 30 mM imidazole, the recombinant protein was eluted with elution buffer (20 mM Tris.HCl, pH 7.5, 400 mM NaCl, 300 mM imidazole). Purified enzyme fractions were pooled, concentrated by ultrafiltration, and stored at −80 °C in 20 mM Tris.HCl pH 7.5; 100 mM NaCl and 10% (v/v) glycerol.

2.3. Enzyme activity

MSR activity was measured by monitoring NADPH oxidation at 340 nm by means of a coupled assay that guaranteed the regeneration of TXNI to its reduced form [27]. All enzyme assays were performed at 30 °C using a Multiskan Ascent one-channel vertical light-path filter photometer (Thermo Electron Co.). The reaction mixture (final volume of 50 μl) contained (unless otherwise specified) 100 mM Tris.HCl, pH 7.5, 2 mM EDTA, 300 μM NADPH, 2 U ml−1 TcTR, 100 μM TcTXNII (Bachem), 10 μM TcTXNI, and the respective MSR included in a specific range of concentrations (0.5–3.5 μM TbMSRA or TbMSRB). TcTXNI was used in these assays as it was readily available in our laboratory and was assumed to function in an equivalent manner to its T. brucei counterpart (TcTXNI and TbTXNI share 62% identity [33]). Reactions were started by the addition of 5 mM MetSO substrate (racemic N-acetyl MetSO (N-AcMetSO) (Bachem), racemic L-MetSO (Sigma-Aldrich), L-Met(S)SO, L-Met(R)SO or N-Acetyl Met(R)SO). The enantiomers were prepared as previously described [34]. Addition of EDTA into the reaction did not affect TbMSRB activity and as such was included in assays to minimise heavy metal mediated thiol oxidation.

For TbMSRB steady-state kinetic analysis, the assay was performed using 20–2500 μM N-AcMet(R)SO and 0.5–20 μM TcTXNI. Kinetic data were plotted as initial velocity (μM min−1) versus substrate concentration (μM). The kinetic constants were acquired by fitting the data with a nonlinear least-squares formula and the Michaelis–Menten equation using the program Origin 7.0. Kinetic constants were the means of at least three independent sets of data, and they were re-producible within ± 10%.

2.4. Yeast complementation

In vivo MSR activity of T. brucei enzymes was checked using the triple msr mutant GY202 (msrA mbsrB Rmsrs) [18]. The Saccharomyces cerevisiae strain GY202 strain was transformed with parental plasmid p425GPD, p425GPD–TbmsrA or p425GPD–TbmsrB and selected for leucine prototrophy on Yeast Nitrogen Based (YNB)-agar medium supplemented with L-Met [35]. To perform the complement assay, each recombinant clone was cultivated on YNB-agar medium supplemented with L-Met, a L-MetSO racemic mix, L-Met(S)SO or L-Met(R)SO (all 100 μM) at 30 °C until growth was visualized.
The full length coding sequences of TbmsrA and TbmsrB were amplified from genomic DNA using the primer combinations TbMSRA-5/TbMSRA-6 or TbMSRB-5/TbMSRB-6 (Table 1), respectively. The products were digested with HindIII/XbaI and cloned into the corresponding sites of the vectors pRP6-C-GFP or BSF-9E10 [31,40]. The cloning was carried out such that the sequences coding for the green fluorescence protein (GFP) or 9E10 epitope from the human c-myc protein were inserted in-frame at the 3′ end of the Tbmsr-derived DNA fragment. The Ascl digested constructs were introduced into T. brucei 2T1 and transformants selected using hygromycin. Expression of the recombinant protein in the parasite was initiated by adding 1 μg ml⁻¹ tetracycline to the culture.

2.7. Localisation

BSF trypanosomes expressing TbMSRB-GFP were suspended at 5 × 10⁶ cells ml⁻¹ in medium containing 100 nM MitoTracker Red (Molecular Probes) and incubated at 37 °C for 5 min. Cells were washed twice in phosphate buffered saline (PBS), fixed in 2% (w/v) paraformaldehyde/PBS then washed again in PBS. Aliquots of the cell sus-pension (10⁵ cells) were then air dried onto microscope slides. Parasite DNA was stained with Vectorshied containing 200 μM 4’,6-diamidino-2-phenylindole (DAPI) (Vector Laboratories Ltd) and slides were viewed using a Leica SP5 confocal microscope. BSF trypanosomes expressing TbMSRA-GFP were treated similarly except the MitoTracker Red step was omitted.

2.5. RNA interference

Fragments corresponding to internal sequences of TbmsrA (491 bp) and TbmsrB (426 bp) were amplified from T. brucei genomic DNA using the primers TbMSRA-3/TbMSRA-4 or TbMSRB-3/TbMSRB-4 (Table 1), respectively. The products were digested with BamHI/XhoI (TbmsrA) or BglII/XhoI (TbmsrB) and cloned into the BamHI/XhoI sites of the vector p2T7 [36]. In this vector, the inserted DNA is flanked by two opposing T7 promoters with each promoter under the control of a tetracycline operator. Constructs were linearised with NotI, electroporated into T. brucei SMB and normalized against standardized control TbTERT and TbMSRB-Amploid curves, the cycle threshold (CT) value was determined in parallel on two independently generated cDNA samples. From the resultant sigmoidal curves, the cycle threshold (CT) value was determined and normalized against standardized control (TbTert; primer combination TbTERT-R and TbTERT-F) amplified in parallel [38] using the com-parative CT method [39].
Fig. 2. T. brucei expresses functional methionine sulfoxide reductases. (A). Coomassie-stained SDS-PAGE gel (15%) containing size standards (in kDa; lane M) and purified, recombinant TbMSRA (lane 1) and TbMSRB (Lane 2). (B) MSR activity of TbMSRA or TbMSRB was followed by monitoring NADPH oxidation at 340 nm using a coupled assay that maintained tryparedoxin (10 µM) to its reduced form (see panel E). Reactions were carried out in 100 mM Tris–HCl pH 7.5, 2 mM EDTA buffer containing NADPH (300 µM), trypanothione reductase (2 U ml⁻¹) and T(SH)₂ (100 µM), to generate reduced tryparedoxin, and different concentrations of TbMSRA (0.05–3.5 µM) or TbMSRB (4–270 µM). Assays were initiated by addition of N-AcMetSO (5 mM) to reactions. (C). The MetSO isomer specificity of TbMSRA and TbMSRB (20 µM) was performed using L-Met(SO) or N-AcMet(R)SO (both 5 mM) as substrate (see Section 2). The data are presented as mean % activity values (± standard deviations) derived from three independent experiments relative to assays performed using racemic L-MetSO (5 mM) as substrate. (D). Proposed scheme for the metabolism of Met(R)SO via a tryparedoxin-mediated TBMSRB pathway. Tryparedoxin (TXN1) acts as a redox shuttle between trypanothione and TBMSRB. Trypanothione disulfide (TSS₂) is converted to dihydrotrypanothione (T[SH]₂) at the expense of NADPH by the activity of tryparedoxin reductase (TR) with "red" and "oxi" representing the reduced and oxidized form of protein/substrates, respectively. The interactions of TBMSRB with TXN1 (reaction I) and Met(R)SO (reaction II) are indicated. Kinetic parameters of recombinant TbMSRs is shown in the associated Table. To study Reaction I, the interaction between TXN1 with TBMSRB, TBMSR activity was assayed as described in panel B using different concentrations of TXN1 (0.3–20 µM) and a fixed concentration of N-AcMet(R)SO (2.5 mM). For Reaction II, the interaction between TBMSRB with N-AcMet(R)SO, TbMSR activity was assayed as described in panel B using a fixed concentration of TXN1 (10 µM) and different concentrations of N-AcMet(R)SO (40–2500 µM). *Data relating to TbMSRA activity taken from Arias et al. [27].

2.8. Peroxide sensitivity experiments

T. brucei BSF parasites were seeded at 1 x 10⁴ ml⁻¹ in 200 µl growth medium containing different concentrations of H₂O₂ and tetracycline (1 µg ml⁻¹), where appropriate. After incubation at 37 °C for 3 days 2.5 µg resazurin (Sigma Aldrich) was added to each culture and the plates incubated for a further 8 h. Cell densities were determined by monitoring the fluorescence of each culture using a Gemini Fluorescent Plate Reader (Molecular Devices (UK) Ltd, Wokingham, UK) at an ex-citation λ = 530 nm, emission λ = 585 nm and a filter cut off at 550 nm, and EC₅₀ values established using the non-linear regression tool on GraphPad Prism (GraphPad Software Inc.).

3. Results

3.1. Trypanosoma brucei express functional methionine sulfoxide reductases

Reduction of the MetSO stereoisomers to Met is mediated by distinct MSRs. Analysis of the T. brucei genome database [32] identified two hypothetical genes that encode for such activities. One 543 bp open reading frame (TB927.8.550; designated as TbmSRA) located on chro-mosome 8 has potential to be translated into a 20 kDa protein related to peptide MSRs, a family of enzymes that mediate metabolism of Met(SO) [6,7]. The second 423 bp open reading frame (TB927.11.11930; designated as TbmSRB) located on chromosome 11, has potential to encode for a 16 kDa protein that has homology to SelR enzymes which catalyse reduction of Met(R)SO [8]. The two T. brucei MSRs share no homology but do have approximately 60% identity to their trypanosomal and leishmanial orthologues (Fig. 1). When compared to other sequences, TbMSRs exhibits similar % identities (35–40%) to counterparts from plants, humans, fungi and bacteria while TbMSRB has higher identity to SelR proteins of plant (~48%) and fungal (~38%) origin relative to bacterial and human (both ~25%) enzymes. Based on sequence, TbMSRB is composed of a single peptide methionine sulfoxide reductase (PF01625) domain that contains a characteristic G-CFWG motif [5]. In MSRAs from other organisms, cysteine (Cys13 in TbMSRA) at this site plays a key role to catalyse reduction of MetSO, with a second conserved cysteine (Cys171 in TbMSRB) found towards the carboxyl terminal helping maintain the catalytic cysteine in its reduced, active form (Fig. 1A). Further searches using localisation pre-diction algorithms (e.g. PSORT II) indicates that TbMSRA lacks any classical sub-cellular localisation signals, suggesting that this enzyme is mostly likely found in the parasite’s cytoplasm. Similarly, TbMSRB is composed of a single SelR (PF01641) domain containing several con-served cysteines [5]. MSRBs from other organisms also contains several conserved cysteine (or selenocysteine) residues. These are involved in oxidoreductase activity (Cys117 and Cys171 in TbMSRA), with the latter cysteine (or selenocysteine) catalysing substrate reduction and the former functioning to maintain the catalytic cysteine in its reduced, active form, or zinc co-factor binding (Cys99, Cys102, Cys145 and Cys148 in TbMSRB) (Fig. 1B). Here, localisation prediction algorithms (e.g. PSORT II, iPSORT and TargetP) suggest that the amino terminal of this enzyme, characterized by the presence of hydrophobic and basic amino acids and a lack of acidic residues, may function as a mi-tochondrial targeting signal.

To investigate whether the trypanosomal proteins can function as MSRs, the DNA sequences encoding for their catalytic domains were
cloned into pTrcHis-C (Invitrogen) and expressed in E. coli. In this system, the recombinant enzymes were tagged at their amino terminus with a histidine-rich sequence and an epitope detectable with the anti-Xpress monoclonal antibody (Invitrogen). For TbMSRA, expression of the full length gene generated soluble recombinant protein. In contrast, the only construct that gave functional ThMSRB was a deletion deri-vative in which the recombinant protein lacked the amino terminal, mitochondrial targeting extension (Fig. 1B). After induction with iso-propyl β-D-thiogalactoside, these constructs generated ~24 and ~21 kDa proteins corresponding to HIS-tagged TbMSRA and TbMSRB respectively, proteins that could be readily purified after one round of affinity chromatography on a nickel- HiTrAP column (Fig. 2A).

Previous studies on trypanosomal MSR revealed that their activity is driven by a TXN-dependent cascade, with NADPH being the source of reducing equivalents [27]. Using a coupled assay that guaranteed re-generation of TXN, the activity of TbMSRA and TbMSRB was monitored by following NADPH oxidation at 340 nm (Fig. 2B). Under the conditions employed, both enzymes were able to reduce N-AcMetSO con-firming that they exhibited MSR activities, with the rate of substrate reduction being dependent upon ThMSR concentration (Fig. 2B). When any of the constituents (TR, T(SH)2, TXN or TbMSRA/ThMSRB) of the pathway were missing, no activity was observed. TbMSRA was able to metabolise N-AcMetSO more readily than TbMSRB with TbMSRA ex-hibiting an apparent kcat > 110-fold higher than that of TbMSRB for this substrate: TbMSRA and TbMSRB display apparent kcat values of 12.3 and 0.11 min−1, respectively, towards N-AcMetSO.

3.2. TbMSRs display different MetSO isomer specificity

To determine TbMSRA or TbMSRB specificity, their activity was followed using different MetSO stereoisomers as substrate and com-pared to reactions using a MetSO racemic mix (Fig. 2C). In agreement with MSRs from other organisms [5], TbMSRA could only metabolise L-Met(S)SO while TbMSRB was specific for towards α-Met(R)SO.

To investigate substrate specificity further and provide additional evidence that the two T. brucei enzymes do function as MSRs, the TbmsrA and TbmsrB DNA sequences in pTrcHis-C were transferred into the yeast expression vector p425-GPD. The resultant plasmids were transformed into a S. cerevisiae strain lacking three MSR enzymes and growth of all fungal lines was supported. When using di- or B variants tagged at their carboxyl terminal with the 9E10 epitope from the human c-myc protein (see below). TbMSRA or TbMSRB could grow on medium where L-MetSO racemate was the sole source of l-Met. When using α-Met(S)SO or α-Met(R)SO supplemented YNB, only S. cerevisiae transformed with the TbMSRA could grow on the former medium whereas only yeast expressing ThMSRB displayed strong growth on the latter.

Previous work revealed that TbMSRA activity could be readily sa-turated by TXN and l-Met(S)SO [27]. Here, we showed that free l-Met (R)SO was not efficiently metabolised by ThMSRB indicating that this was not the physiological substrate for this enzyme (Table 1). Instead TbMSRB displayed Michaelis-Menten type kinetics towards TXN and N-AcMet(R)SO although metabolism of this particular substrate via this pathway was extremely low (kcat of 0.085 min−1) (Fig. 2D). Comparison of Km values suggests that the rate limiting step within this pathway may be the interaction of TbMSRB with N-AcMet(R)SO (TbMSRB has a Km of 1.8 ± 0.3 µM towards TXN and a Km of 429 ± 58 µM towards N-AcMet(R)SO). These low kinetic values are typical for those reported for bacterial, plant and mammalian MSRs [41–43] while the ability of the parasite enzyme to metabolise N-AcMet(R)SO indicates that it can ef-fectively reduce protein bound Met(R)SO: this form of MetSO is often used as substrate to test for peptide bond MSR activity [44].

3.3. TbMSRs are targeted to different cellular localisations

The subcellular location of the trypanosomal MSRs was examined by expressing GFP-tagged versions of each enzyme in BSF parasites. The DNA sequences encoding for TbMSRA and TbMSRB were amplified then cloned in-frame and upstream of the GFP gene in a trypanosomal vector that facilitated tetracycline inducible gene expression. The co-localisation constructs were electroporated into T. brucei and re-combinant parasites selected.

To induce expression of the tagged protein, cells were incubated in the presence of tetracycline for 48 h. The parasites were then examined by western blotting using an antibody against GFP (Fig. 4A), with each extract containing a single band of the expected size (bands of ~48 and ~45 kDa were observed in TbMSRA-GFP or TbMSRB-GFP containing lanes), or were fixed and examined by confocal microscopy (Fig. 4C and D). For T. brucei induced to express tagged TbMSRA, a fluorescence signal was observed throughout the main body of the cell but absent from the nucleus indicating that this protein is found in the parasite's cytoplasm (Fig. 4C). In contrast, for parasites expressing TbMSRB-GFP, a lattice-like structure spread throughout the cell and reminiscent of the pattern reported for proteins that localise to the parasite's large, single mitochondrion, was observed (Fig. 4D). To confirm this, cells were co-stained with the mitochondrial dye, MitoTracker. When the images were superimposed, a pattern of co-localisation (yellow staining) was noted indicating that TbMSRB was located in the same compartment as MitoTracker. When cells expressing untagged GFP were analysed, fluorescence was detected throughout the whole cell (Fig. 4B). The above localisation patterns were observed by immunofluorescence studies using parasites expressing TbMSRA or B variants tagged at their carboxyl terminal with the 9E10 epitope from the human c-myc protein (see below).

3.4. Functional analysis of TbMSRs in T. brucei

To assess the importance of MSR activity to BSF T. brucei, an RNAi-based approach was employed. DNA fragments corresponding to the central regions of TbmsrA and TbmsrB were amplified and cloned into the vector p2T711. The RNAi constructs were transformed into T. brucei and recombiant parasites selected.

The effect of inducing RNAi was examined by following the cell density of tetracycline-treated cultures over a 4 day period and com-pared against untreated controls (Fig. 5A). In the absence of RNAi, all recombiant cell lines were found to grow at roughly the same rate as wild type control cells. For cells undergoing RNAi targeting the TbmsrA transcript, no significant difference in rate of growth was observed over the initial 24 h period (Fig. 5B). However, over the following 24 h, a dramatic and reproducible reduction in the cell density was observed in all replicates and clones tested, with this correlating to ~75% fall in the Fig. 3. Yeast complementation assay demonstrates substrate preference of T. brucei MSRs. The growth of S. cerevisiae GY202 triple mutant (msrA msrB fmsr) transformed with plasmids that facilitate expression of TbmsrA (msrA) or TbmsrB (msrB) on YNB agar the vector plus dextrose agar supplemented with Met or different MetSOs was compared against control strains transformed with empty vector (p425 GPD).
The yellow pattern in panel 4 revealing the co-fractionation (panel 3) and phase image (panel 4) shown. (D). T. brucei cells expressing TmMSR-GFP (panel 2) were co-stained with DAPI (panel 1) with the merged signals (panel 3) and phase image (panel 4) shown. (D). T. brucei cells expressing TmMSR-GFP (panel 2) were co-stained with DAPI (panel 1) and Mitotracker (TM; panel 3). The merged GFP/DAPI/Mitotracker signals (panel 4) and phase image (panel 5) are shown, with the yellow pattern in panel 4 revealing the co-localisation on TmMSRB-GFP and Mitotracker. Scale bar in B, C and D = 5 µm (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Fig. 4. Localisation of TmMSRs in bloodstream form T. brucei. (A). Expression of GFP tagged TmMSRs was examined by probing a blot containing cell lysates from T. brucei wild type (lane 1) and TmMSRA-GFP or TmMSRB-GFP expressing cells (lanes 2 and 3 respectively) using an anti-GFP antibody (upper panel). Protein from 1 × 10^7 cells was loaded in each track and equal loading verified by Coomassie staining (lower panel). (B). T. brucei expressing un-tagged GFP. The blue spots correspond to the nuclear (N) and mitochondrial (K) genomes of a trypanosome expressing GFP alone. (C). T. brucei cells expressing TmMSRA-GFP (panel 2) were co-stained with DAPI (panel 1) and Mitotracker (TM; panel 3). The merged GFP/DAPI/Mitotracker signals (panel 4) and phase image (panel 5) are shown. (D). T. brucei cells expressing TmMSR-GFP (panel 2) were co-stained with DAPI (panel 1) and Mitotracker (TM; panel 3). The merged GFP/DAPI/Mitotracker signals (panel 4) and phase image (panel 5) are shown, with the yellow pattern in panel 4 revealing the co-localisation on TmMSRB-GFP and Mitotracker. Scale bar in B, C and D = 5 µm (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

The expression of TmMSR transcript, as judged by qPCR. Further, the presence of cell debris was noted in these cultures, suggestive that a proportion of the parasite population had undergone lysis, while any remaining viable cells exhibited reduced movement. From 48 h onwards, an outgrowth of viable parasites were observed in all replicates, a type of reversion previously observed when targeting transcripts important to the growth of BSF T. brucei [36,45,46].

In contrast, cells undergoing RNAi targeting the TmMSRB transcript continued to grow throughout the entire period albeit with a reduced rate: over the first 24 h, the growth rate of tetracycline-treated cells was roughly half that of untreated cultures with this falling further to approximately 20% in the next 24 h (Fig. 5B). Over the subsequent 48–72 h, an outgrowth of viable parasites were observed.

Alteration of MSR expression levels in other organisms has shown that they play an important role in protecting cells from exogenous oxidants [7,15,19,27,28]. Here, we investigated whether down-regulation of the trypanosomal MSR transcripts altered parasite susceptibility to H2O2. Tetracycline induced and non-induced cells harbouring the TmmsrA or TmmsrB RNAi constructs were grown in the presence of oxidant and the effective compound concentration that inhibits parasite

![Fig. 5. Phenotypic analysis of RNAi cell lines. (A). The TmmsrA or TmmsrB transcript levels from non-induced cells (dark grey) and trypanosomes induced for 48 h to undergo RNAi (light grey) was assessed by qPCR and compared against the expression level of a standardized control (TbH2). The relative fold difference, as judged by 2(-ΔΔCt) from reactions performed in triplicate ± standard deviation, was plotted as a measure of the relative expression level. The difference in relative expression levels between the non-induced and induced lines was judged to be statistically significant (P < 0.01), as assessed by the Student’s t-test. The TmmsrA or TmmsrB mRNAs were both ~3.5 fold lower in the corresponding RNAi line relative to controls. (B). Growth of cells induced to undergo RNAi (dashed lines) targeting the TbmsrA or TbmsrB transcript was compared against non-induced cultures (solid line). The growth patterns shown relate to a single clone. Two other clones analysed in parallel exhibited the same profiles. (C). The BSF RNAi-TbmsrA and RNAi-TbmsrB lines were grown for 24 h in the presence of tetracycline (1 µg ml^-1) (square, dotted line), seeded at 1 × 10^4 ml^-1 and then exposed to various concentrations of H2O2 (12.5–1000 µM): where applicable, RNAi induction was maintained in peroxide treated cultures by addition of fresh tetracycline to the growth medium. After 3 days at 37 °C, resazurin (2.5 µg) was added to each culture and used to determine cell density (Experimental Procedures). Untreated (circles, solid line) and wild type parasites were analysed in parallel. From the resultant dose response curves, the EC50 of each line towards H2O2 was calculated. All data points are means for experiments performed in quadruplicate ± standard deviation. The difference in susceptibility to H2O2 displayed by the tetracycline-treated and untreated RNAi-TbmsrA cells as judged by EC50 values was statistically significant (P < 0.0001), as assessed by Student’s t-test.](image-url)
Fig. 6. Overexpression of TbMSRs in bloodstream form T. brucei. (A). Expression of -myc (9e10) tagged TbMSRs was examined by probing a blot containing cell lysates from T. brucei wild type (lane 1) and TbMSRA-9e10 or TbMSRB-9e10 expressing cells (lanes 2 and 3 respectively) using c-myc (9e10) anti-serum (upper panel). Protein from 1 × 10^5 cells was loaded in each track and equal loading verified by Coomassie staining (lower panel). (B). The T. brucei TbmsraA-9e10 and Tbmsrb-9e10 lines were grown for 24 h in the presence of tetracycline (1 µg ml^-1) (diamond, dotted line), seeded at 1 × 10^5 ml^-1 and then exposed to various concentrations of H_2O_2 (50–500 µM). After 3 days at 37°C, resazurin (2.5 µg) was added to each culture and used to determine cell density (Experimental Procedures). Untreated (circles, solid line) and wild type parasites were analysed in parallel. From the resultant dose response curves, the EC_{50} of each line towards H_2O_2 was calculated. All data points are means for experiments performed in quadruplicate ± standard deviation. The difference in susceptibility to H_2O_2 displayed by the tetracycline-treated and untreated Tbmsra-9e10 cells as judge by EC_{50} values was statistically significant (P < 0.0001), as assessed by Student’s t-test.

growth by 50% (EC_{50}) determined (Fig. 5C). For cells expressing re-duced levels of TbMSRB, no significant difference in parasite sensitivity was observed: the EC_{50} values ranged from 210 to 225 µM. In contrast, recombinant T. brucei induced to undergo RNAi targeting the Tbmsra transcript were approximately 2-fold more susceptible to H_2O_2 than controls: the EC_{50} of non-induced RNAi parasites was 204.7 ± 9.4 µM in comparison to tetracycline treated cells that exhibited an EC_{50} of 111.4 ± 13.5 µM.

As parasites with reduced levels of Tbmsra are more susceptible to oxidative stress, we tested whether MSR overexpression leads to re-sistance. Tbmsra and Tbmsrb, minus their STOP codon, were amplified then cloned in-frame and upstream of the DNA sequence encoding for the 9e10 epitope from the human c-myc protein in a trypanosomal vector that facilitated tetracycline inducible gene expression. The constructs were introduced into T. brucei and recombinant parasites selected. To verify expression of the tagged TbMSRs, cells extracted generated from parasites cultured in the presence of tetracycline for 48 h were examined by western blotting using an antibody against c-myc epitope (Fig. 6A). A single band was observed in lanes where TbMSRA-9e10 (~21 kDa) or TbMSRB-9e10 (~20 kDa) expression had been induced. Tetracycline treated and untreated T. brucei were then grown in the presence of H_2O_2 and EC_{50} determined (Fig. 6B). For cells expressing TbMSRB-9e10, no significant difference in parasite sensi-tivity was observed. When these studies were extended to TbMSRA, parasites expressing the c-myc-tagged enzyme were 1.5-fold more re-sistant to H_2O_2 than controls: the EC_{50} of non-induced parasites was 212.9 ± 7.7 µM in comparison to tetracycline induced cells that ex-hibited an EC_{50} of 320.4 ± 2.9 µM.

4. Discussion

Throughout its life cycle, T. brucei is continually exposed to ROS. To combat their detrimental effects, this parasite expresses a series of novel protection strategies that collectively contribute to this pathogen’s oxi-dative defence system. As these mechanisms are distinct from those of its mammalian host [21,22,47–49], with several components being essential for growth of the medically relevant parasite stage, they are viewed as potential targets for development of new HAT che-motherapies [21,36,49,50]. Here, we report the characterisation of two additional components of the trypanosomal oxidative defence system, showing that T. brucei expresses two structurally unrelated MSRs (TbMSRA and TbMSRB) that can be distinguished on the basis of their sequence, substrate specificity and sub-cellular localisation. Although functionally related, TbMSRA and TbMSRB share no sig-nificant sequence homology with this diversity conferring difference in substrate specificities. In the case of TbMSRA, the presence of a GCFWG motif in its sequence coupled with it being only able to metabolise the S epimer of MetSO establishes this as a member of the peptide methionine sulfoxide reductase, or MSRA, family of antioxidant proteins (Figs. 1A, 2 and 3). In contrast, TbMSRB contains a domain structure that places it in the SelR, or MRSB, group of methionine sulfoxide reductases, with enzymatic and phenotypic screening showing that this parasite oxi-doreductase can only medi-ate conversion of peptide-bound Met(R)SO to Met(S)SO (Fig. 1B, 2 and 3).

Based on sequence and substrate specificity, the two parasite en-zymes are typical members of the MSRA or MRSB clades although they are different in relation to the pathways that facilitate their reduction. In many instances, the concerted action of thioredoxin reductase, thioredoxin and MSR act as intermediaries to shuttle reducing equiva-lents from NADPH to MetSO [11]. As trypanosomes lack thioredoxin reductase this pathway cannot operate in T. brucei. In the case of TbMSRA, its in vitro activity is dependent upon TR and trypanothione. Using NADPH as electron donor, these parasite specific factors fulfil the role of ‘thioredoxin reductase’, functioning to maintain TXN in its re-duced state [51], with reducing equivalents then transferred onto Met (S)ISO to form Met via the trypanosomal MSR (Fig. 2D) [27]. Steady state kinetics indicate that the rate limiting step of this pathway appears to be the TbMSRA mediated reduction of free Met(S)ISO to Met, a trait noted for MSRs from other organisms [27,52–54]. As all components of this cascade are widely distributed throughout the parasite cytosol.
Our data shows that TbMSRA and TbMSRB are located in the T. brucei cytoplasm and mitochondrion respectively, a distribution also observed with the S. cerevisiae counterparts (Fig. 4C) [47,55] and that TbMSRA exhibits reasonable kinetics towards both TXN and free Met(S)SO, it is hypothesized that the above pathway functions within T. brucei, acting to repair damaged Met, specifically free Met(S)SO, at this particular subcellular site. Intriguingly, our yeast complementation studies indicate that TbMSRA (and TbMSRB) can utilize non-trypanothione redox cascades as source of reducing equivalents suggesting that the activity of these enzymes within the trypanosome may be driven by an alternative system to that described above. This is analogous to the situation seen with other trypanoperoxidin-dependent enzymes, including the trypanoperoxidin peroxidases, which can exploit different thioredoxin/thioredoxin-like molecules as electron donor [48,56]. We postulate that the trypanosomal MSRs are able to interact with the yeast thioredoxin system and complement for the appropriate fungal mutation to produce the observed growth pheno-type.

In contrast, the pathway that operates to maintain TbMSRB in its reduced state is unclear. Biochemical studies on recombinant enzyme has shown that the TR/T[SH]2/TXN system can support TbMSRB ac-tivity resulting in slow turnover of N-AcMet(R)SO (Fig. 2). The kinetic values exhibited by the parasite protein to both TXN and MetSO are on par with those noted for other MSRBs, with the rate of N-AcMet(R)SO reduction limited by the enzyme/substrate interaction [8,57]. Such observations, coupled with the MSRBs preference to bind and meta-polise protein-bound Met(R)SO, has led to the idea that free MetSO may not be the major physiological substrate for these reductases [8] and may account for the fact that some organisms have evolved other free Met(R)SO metabolising activities (e.g. rMRSRs), with turnover of this particular substrate taking place at an appreciably faster rate than MSRBs [9,10]. Further, the components that support the trypanothione-dependent, TbMSRB reduction cascade are not present at the same cellular site: TR and TbMSRB are restricted to the cytosol and mi-tochondrion, respectively, while TXN1 is apparently found across both sites (Fig. 4B) [47,55]. By implication, either TbMSRB activity is sup-port ed by an unidentified pathway found entirely within the parasite mitochondrion or the TR/T[SH]2/TXN redox cascade is split across different cellular compartments, with comparative analysis indicating that the thiol constituent being the transferable factor. In other eu-karyotic cells, GSH is maintained at high (mM) levels in the mi-tochondrial lumen even though this organelle lacks the biosynthetic machinery to make this tripeptide. Instead, cytosolic pools of GSH are transferred into mitochondria via several transporters which may in-clude dicarboxylate and 2-oxoglutarate carriers [58–60]. If an equiva-lent T[SH]2 translocation mechanism(s) does function in trypanosomes then given the unique properties of the parasite specific thiol, any transport system would be mechanistically distinct from that which operates in the transfer of GSH and as such would be of particular in-terest as a target(s) for chemotherapy. Once in the mitochondrion, T[SH]2 can then facilitate transfer of reducing equivalents via TXN and TbMSRB to Met(R)SO. Intriguingly, data released as part of the TrypTag project [61] indicates that trypanosomes express a classical thioredoxin (Tb927.9.3370 on TrypDB [32]) which can be readily reduced by T[SH]2 [62,63], present throughout the parasite mitochondrion. If correct, this may also function as an intermediary in shuttling reducing equivalents from T[SH]2 to TbMSRB.

In summary, we have demonstrated that T. brucei expresses cyto-solic and mitochondrial methionine sulfoxide reducing pathways which together constitute a new arm of this parasite's oxidative reparative defence system. As the redox cascades associated with these activities rely upon trypanosome-specific factors and that the activity of the terminal reductase is important for pathogen growth, these mechanisms may have potential as chemotherapeutic targets.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Acknowledgments

We thank Guy Hanke (QMUL) for their critical review of this manuscript. We acknowledge the members of the T. brucei genome (http://tritrypdb.org/tritrypdb/) and TrypTag (http://trypact.org) projects for sequence and localisation data, respectively. A component of this work was supported by grants from ANPCyT (PICT-2015-1149; PICT-2014-2103), SAG, and DGA are investigator career members from CONICET. AK was a recipient of a Queen Mary University of London PhD studentship.

References

Our data shows that TbMSRA and TbMSRB are located in the T. brucei cytoplasm and mitochondrion respectively, a distribution also observed with the S. cerevisiae counterparts (Fig. 4) [64]. Why trypa-nosomas and yeast target these two enzymes to only these sites is un-clear given that mammalian cells possess multiple isoforms of each enzyme type at different sub-cellular sites (mammalian cells express cytosolic, mitochondrial and nuclear versions of MSRA and MSRB while an isoform of the latter is also present in the endoplasmic reticulum [65–67]). Taking into consideration their substrate specificity, the specific localisation displayed by TbMSRA and TbMSRB raises a number of interesting questions such as how is free or protein bound Met(S)SO metabolised in parasite organelles and how is free or protein bound Met(R)SO detoxified at non-mitochondrial sites. Part of this may be attributed to uncharacterised activities (e.g. T. brucei appears to have potential to encode for a cytoplasmic rMRS (Tb927.5.1250 on TrypDB)) or could reflect the importance of various MetSO forms in different compartments of the cell (e.g. it may be favourable to export free MetSO epimers out of the mitochondrion into the cytoplasm to facilitate conversion back to Met, rather than carrying out this repair within the organelle).

MSRs represent key components in an organism's oxidative defence armoury, functioning in processes such as bacterial and protozoal virulence [7,28,68–70] and ageing [16,17,71]. To evaluate the im-portance of trypanosomal enzymes to T. brucei, functional genomic approaches were used to generate parasite lines expressing altered le-vels of each reductase. Using RNAi, each enzyme was shown to be important but not essential for the growth of bloodstream form para-sites (Fig. 5). For trypanosomes engineered to express reduced levels of the TmmsrA transcript, the stalling of parasite growth was accompanied by an increased susceptibility to exogenous H2O2 with the reciprocal phenotype observed in cells engineered to over express this enzyme. In contrast, the reduction in cell growth observed when targeting the TbmssrB or the elevated expression of this reductase in T. brucei did not affect parasite sensitivity to H2O2. The observed susceptibility pheno-types suggest that the amount of exogenous oxidant added to cultures is sufficient to promote MetSO formation in targets found in the parasite's cytosol, damage that can be readily repaired by TbMSRA. However, these peroxide levels are not sufficient to cause significant Met oxida-tion in the trypanosomal mitochondrion, presumably because this oxi-dant is detoxified by the various trypanoperoxidin peroxidase systems ex-pressed by T. brucei [47,49].
S. Guillermo et al.


References


Cole, R. S. (1971) ‘Inactivation of Escherichia coli, F’ episomes at transfer, and


240


its full in vitro and in vivo functions.


