Evaluating Trypanosomal DNA Repair Mechanisms

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

Isatou Drammeh

Supervisor: Dr Shane Wilkinson
School of Biological & Behavioural Sciences
Queen Mary University of London

Statement of originality

I, Isatou Drammeh 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.

Date: 16/10/2022

Details of publications:

1. Ambika Dattani, Isatou Drammeh, Aishah Mahmood, Mahbubur Rahman, Joanna Szular, Shane R Wilkinson. Unraveling the antitrypanosomal mechanism of benznidazole and related 2-nitroimidazoles: From prodrug activation to DNA damage.
   Molecular Microbiology 116, 674-689. https://doi.org/10.1111/mmi.14763 (2021).

Acknowledgements

My sincere gratitude to the Islamic Development Bank (IDB) and School of Biological and Behavioural Sciences, Queen Mary University of London for providing funding and facilities for my PhD.

Special thanks to all other research laboratories that have provided antibodies, cell lines compounds or research facilities for this project. I am particularly thankful to Professor John Kelly of the LSHTM and his group, the McCulloch laboratory at the University of Glasgow. Prof Dr Eva Gluenz and Dr Samuel Dean for providing the Cas9 cell lines and pPOT plasmids used in this project.

A massive thank you to my supervisor, Dr Shane Wilkinson, for steering me through this PhD and for always making sure work goes on even amid a difficult pandemic. I appreciate your wisdom and guidance throughout my PhD and will forever be grateful for your timely feedback and intervention. I admire your dedication and steadfastness to work and responsibilities. I hope to emulate that in everything I do. Without your constant support and wise guidance this work would not be the same.

Sincere appreciation to my Wilkinson lab teammates, Dr Ambika Dattani, Dr Monica Zavala and Katerina. And to all past undergraduate and masters' students from the Wilkinson group. Special thanks to Mr. Luke Tappouni, Mr. Mahbubur Rahman, Farhan Miah, for your contributions to this work. To my office mates, Dr Katherine Fenn, Nancy, Maria, and Layla, I thank you for all the special moment, especially our online tea sessions during the Covid-19 pandemic. Special thanks to my friends and colleagues at the Medical Research Council unit, The Gambia (now MRC at LSHTM). To my friends Dr Parveen Gul for your friendship and support. Fozia and Mamtuti for all the laughter and time together. I am also grateful to my Nusrat & UTG alumni mates for your continued friendship and support.

To my children I apologise for missing out on so many special events in your lives, I promise to make up! Thanks for your understanding and always bringing me joy. To my mum and siblings, I can't thank you enough for looking after my little children as I travelled across the world for my PhD. Thank you for your prayers, support, and love! To the men in my life my brother, dad and husband, I thank you from the bottom of my heart. I appreciate all the love and support you have given through the times. Special thanks to my aunties, uncles, and entire extended family for always checking on me to make sure I was okay. And finally, special appreciation to my sisters for picking your phones and just listening!

Abstract

The anti-trypanosomal agents nifurtimox, benznidazole and fexinidazole are nitroheterocyclic prodrugs that must undergo type I nitroreductase (NTR1)-mediated activation before promoting their anti-parasitic activities. Such reactions generate toxic products some of which are hypothesised to promote DNA damage in the parasite's genomes. Here, and using several complementary approaches, we assessed how components from *Trypanosoma brucei's* homologous recombination (HR), nucleotide excision repair (NER), base excision repair (BER) and DNA interstrand crosslink (ICL) repair pathways influence the mode of action of the above nitroheterocycles. CRISPR-Cas9 genome editing was utilised to engineer parasites expressing mNeonGreen-tagged DNA repair enzymes from the above pathways. These were phenotyped with regard to protein localisation, and how protein expression is affected by DNA damaging/nitroaromatic prodrug treatments. Next, the susceptibility of parasites null or heterozygous for certain DNA repair genes to various DNA damaging/nitroheterocyclic prodrug treatments was determined while the effect of RNA interference on expression and function of selected NER factors was followed. Together, our findings indicate that nifurtimox, benznidazole and fexinidazole (as well as most DNA damaging treatments) elicit a globalized alteration in expression of most DNA repair factors evaluated even when that factor supposedly plays no role in repairing any damage generated. Phenotypic and protein expression studies revealed that the principal mechanism by which benznidazole-induced damage is resolved involves the cooperation of components of the trypanosome's HR repair pathway (TbMRE11, γH2A, TbRAD51), indicating that the lesions produced by this agent are most likely double stand DNA breaks, while the sequence/recruitment kinetics of these factors parallels that in other eukaryotes HR systems. When extended to fexinidazole, our findings demonstrate this compound functions as a TbNTR1-activated prodrug with this interaction generating metabolites that promote forms of DNA damage which requires the cooperation of TbCSB and TbMRE11 from NER and HR, respectively, to resolve. Understanding the mode of activation of these compounds and how they interact with DNA repair factors will aid the development of novel therapies containing similar functional groups and facilitate the design of inhibitors of various DNA repair networks which when added to novel or current therapies makes them more potent.

Table of Contents

Statement of originality 2
Acknowledgements 3
Abstract. 4
Table of Contents 5
List of Figures 9
List of Tables 13
Summary of the localisation of the different proteins studied in this project 13
List of Symbols and Abbreviations 14
CHAPTER 1: Introduction 16
`1.1 Trypanosomatids. 16
1.2. *Trypanosoma brucei* and Human African Trypanosomiasis. 17
1.2.1 History of Human African Trypanosomiasis. 17
1.2.2. Historical Perspective of Human African Trypanosomiasis. 19
1.2.3. Distribution of Disease. 21
1.2.4. The Causative Agent. 24
1.2.5. Life Cycle of *Trypanosoma brucei*. 24
1.3. Disease Staging, Signs and Symptoms. 26
1.4. Diagnosis of Human African Trypanosomiasis. 28
1.5. Treatment of human African trypanosomiasis 30
1.6. *Trypanosoma brucei* molecular biology 34
1.6.1. The genome of *Trypanosoma brucei* 34
1.6.2. Gene expression in Trypanosomes 37
1.6.3. Genetic tools for studying *Trypanosoma brucei* biology 39
1.7. DNA repair pathways in *Trypanosoma brucei*. 41
1.8. Nitroheterocycles and DNA damage repair pathways in Trypanosomes 42
1.9. Research Aims 43

CHAPTER 2: Materials and Methods 45
2.1. Culturing of trypanosome cell lines and bacterial strains 45
2.1.1.Cell lines 45
2.1.2.Culturing and long-term storage of *Trypanosoma brucei* 46
2.1.3 Culturing and long-term storage bacterial strains 47
2.2. Trypanocidal and genotoxic agents 47
2.3 Plasmids and oligonucleotide primers 48
2.4. Bioinformatics 49
2.5 Nucleic Acid Purification. 50
2.6 DNA synthesis 51
2.6.1 cDNA synthesis 51
2.6.2 Standard PCR protocols 51
2.6.3 Generation of sgRNA amplicons for CRISPR 52
2.6.4 Generation of targeting fragment amplicons for CRISPR 52
2.6.5 Reverse transcription quantitative PCR 52
2.7 DNA manipulation 53
2.7.1 DNA electrophoresis 53

2.7.2 Restriction digestion 54
2.7.3 DNA ligation 54
2.7.4 *E. coli* DNA transformation 55
2.7.5 *T. brucei* DNA transformation 55
2.7.6 DNA sequencing 56
2.8 Construction and validation of recombinant parasite lines 56
2.8.1 CRISPR/Cas9 genome edited line 56
2.8.2 RNAi lines 57
2.8.3 Complementation lines 57
2.8.4 *T. brucei* TbCSB null lines that express elevated levels of TbNTR1. 58
2.9 Phenotyping recombinant parasite lines 58
2.9.1 Cumulative growth assays 58
2.9.2 *T. brucei* growth inhibition assays 58
2.9.3 Fluorescence microscopy and image analysis 59
2.9.4 Monitoring temporal changes in protein expression using fluorescence 60
2.10 Protein work 61
2.10.1 Preparation of trypanosomal cell extracts 61
2.10.2 SDS-PAGE electrophoresis 61
2.10.3 Immunoblot analysis 62

CHAPTER 3: Unravelling the Anti-Trypanosomal Mechanism of Nitroheterocycles: From Prodrug Activation to DNA Damage 63
3.1.Double strand DNA breaks and their repair. 63
3.2 Informatic analysis of TbMRE11 67
3.3 Informatic analysis of TbRAD51 71
3.4 Unravelling the mechanism of action of fexinidazole 74
3.5 Editing the Tb*mre11* and Tb*rad51* loci in *T. brucei*. 79
3.6 Localisation of TbMRE11 and TbRAD51 in *T. brucei*. 84
3.7 Nitroheterocycles and their effect on the homologous recombination pathway 86
3.8 Nitroheterocyclic prodrugs and their effects on temporal expression of the trypanosomal HR pathway. 89
3.9 Effect of UV on temporal expression of the trypanosomal HR pathway. 93
3.10 Effect of ICL damage on temporal expression of the trypanosomal HR pathway. 95
3.11. Chapter Summary 98

CHAPTER 4: Evaluation of Factors Specific to ICL DNA Repair 99
4.1. DNA interstrand cross links and their repair 99
4.2 Assessing the role of TbSNM1 in *T. brucei*. 110
4.2.1 Informatic analysis of TbSNM1 110
4.2.2 Editing the Tb*snm*1 loci in *T. brucei* 114
4.2.3 Localisation of TbSNM1 in *T. brucei*. 119
4.2.4 Effect of exogenous treatments on TbSNM1 expression. 120
4.2.5 Development of insect form *T. brucei* null for Tb*snm1*. 122
4.2.6 Phenotypic analysis of insect form *T. brucei* Tb*snm1Δ*. 124
4.3 Assessing the role of TbFAN1 in *T. brucei*. 127

4.3.1 Informatic analysis of TbFAN1. 127
4.3.2 Editing the Tb*fan1* loci in *T. brucei* 130
4.3.3 Localisation of TbFAN1 in *T. brucei*. 131
4.3.4 Effect of exogenous treatments on TbFAN1 expression. 133
4.4. Chapter Summary 136
CHAPTER 5: Unravelling the Role of the *T. brucei* BER Pathway. 137
5.1 Overview of base excision repair 137
5.2. Informatic analysis of components of the *T. brucei* base excision repair pathway 142
5.2.1. Informatic analysis of TbAPE1 142
5.2.2. Informatic analysis of TbMYH 144
5.3 Editing the Tb*ape1, Tbmyh* and *Tbnth1* loci in *T. brucei*. 150
5.4 Localisation of TbAPE1, TbMYH and TbNTH1 in *T. brucei*. 154
5.5 Effect of exogenous treatments on TbAPE1, TbMYH or TbNTH1 expression. 158
5.6 Disrupting Tb*ape1, Tbmyh* or *Tbnth1* in the genome of procyclic form *T. brucei*. 167
5.7 Phenotypic analysis of null lines 172
5.8. Chapter Summary 177

CHAPTER 6: Analysis of the *T. brucei* NER Pathway. 179
6.2 Informatic analysis of the *T. brucei* nucleotide excision repair pathway 184
6.2.1 Informatic analysis of TbCSB 184
6.2.2 Informatic analysis of TbDDB1 187
6.2.3 Informatic analysis of TbRAD23 190
6.2.4 Informatic analysis of TbXPC 192
6.2.5 Informatic analysis of TbERCC1 194
6.2.6 Informatic analysis of TbXPF 197
6.2.7 Informatic analysis of TbXPG 201
6.3 Editing the genes coding for NER factors. 205
6.4 Localisation of NER factors in *T. brucei*. 210
6.4.1 Localisation of the TC-NER factor, TbCSB. 212
6.4.2 Localisation of GG-NER factors. 214
6.4.3 Localisation of NER incision factors. 215
6.5 Effect of exogenous treatments on TbCSB, TbXPG or TbXPF expression. 219
6.6 Functional analysis of the trypanosomal NER pathway using RNA. interference. 225
6.7. Chapter Summary 230

CHAPTER 7: Discussion 231
7.1. Elucidating the trypanocidal activity of fexinidazole 232
7.2. Kinetics of *T. brucei* DNA repair and possible interactions 233
7.2.1 Recruitment of *T. brucei* DNA repairtoire 235
7.2.1.1. HR 235
7.2.1.2. ICL 237
7.2.1.3. BER 239
7.2.1.4. NER 240

7.3. RNAi of GG-NER **241**
7.4. Disruption of Key BER enzymes **242**
7.5. Disruption of TbSNM1(ICL) **243**
7.6. Informatic analysis of *T.brucei* DNA repairsome **244**

CHAPTER 8: Thesis Summary and Future Work **248**
8.1. Future work **249**
8.2. Thesis summary **249**
9. Appendix **251**
Appendix 1: Primer sequences for CRISPR/Cas9 based tagging **251**

Appendix 2: Primer sequences for validation of CRISPR/Cas9 based tagging **253**
Appendix 3: Primers for generation of RNAi vectors **254**
Appendix 4: Primers for construction of complementation vectors **254**

10. REFERENCES **271**

List of Figures

Figure number	Title of figure	Page number
1.1	Structure of a trypanosomatid cell.	16
1.2:	History of African trypanosomiasis.	18
1.3	Historical perspectives relating to human African trypanosomiasis	20
1.4	Geographical distribution of human African trypanosomiasis	21
1.5	Decline in human African trypanosomiasis between 2000-2018	23
1.6	Life cycle of Trypanosoma brucei	25
1.7	Disease staging, and sign and symptoms of human African trypanosomiasis	27
1.8	WHO guidelines for the detection of West African trypanosomiasis.	28
1.9	Detection of *T.b*. gambiense using the Card Agglutination Test for Trypanosomiasis.	29
1.10	Microscopic detection of *Trypanosoma brucei*	30
1.11	Structures of the drugs used to treat human African trypanosomiasis	31
1.12	Schematic showing generalised features found on a *T. brucei* chromosome	35
1.13	Schematic illustrates a generalized polycistronic transcription unit on *T. brucei* chromosome.	38
3.1	Double strand DNA repair pathways expressed in eukaryotic cells	64
3.2	Informatic analysis of TbMRE11.	68
3.3	3D prediction model of a TbMRE11 monomer	69
3.4	Manganese-binding residues and catalytic site	70
3.5	Protein: protein interaction regions of TbMRE11	70
3.6	Informatic analysis of TbRAD51	73
3.7	3D prediction model of TbRAD51	74
3.8	Location of motifs and loops in the TbRAD51 3D prediction model.	74
3.9	Susceptibility of *T. brucei* DNA repair mutants towards fexinidazole	75
3.10	Susceptibility of *T. brucei* Tb*mre11*Δ Tb*csb*Δ null line towards fexinidazole	76
3.11	Susceptibility of *T. brucei* expressing elevated levels of Tb*ntr1* towards fexinidazole	77
3.12	Linking fexinidazole activation to DNA damage	78
3.13	CRISPR/Cas9-mediated tagging of the endogenous loci	80
3.14	Validation of *T. brucei* lines expressing mNeonGreen tagged TbMRE11.	81
3.15	Validation of PCF parasites expressing mNeonGreen tagged TbMRE11	82

3.16	Validation of *T. brucei* lines expressing mNeonGreen tagged TbRAD51	83
3.17	Localisation of TbMRE11.	84
3.18	Localisation of TbRAD51	86
3.19	Effect of exogenous treatments on TbMRE11 expression	87
3.20	Effect of exogenous treatments on TbRAD51 expression	88
3.21	Effect of benznidazole on the temporal expression of the *T. brucei* homologous recombination DNA repair pathway	90
3.22	Effect of nifurtimox on the temporal expression of the *T. brucei* homologous recombination DNA repair pathway	91
3.23	Effect of fexinidazole on the temporal expression of the *T. brucei* homologous recombination DNA repair pathway	92
3.24	Effect of UV treatment on TbMRE11 and TbRAD51 expression	94
3.25	Effect of UV treatment on TbMRE11 and TbRAD51 expression	95
3.26	Effect of mechlorethamine on TbMRE11 & TbRAD51 expression	96
3.27	Effect of mechlorethamine on TbMRE11 & TbRAD51 expression	97

4.1	Types of ICL lesions. Distorting	99
4.2	Replication-independent ICL repair pathway	103
4.3	Replication-dependent ICL repair in higher eukaryotes.	106
4.4	Replication-dependent ICL repair in lower eukaryotes.	107
4.5	*T. brucei* ICL repair.	109
4.6	Informatic analysis of TbSNM1.	112
4.7	3D prediction models of TbSNM1.	113
4.8	Putative zinc-binding site of TbSNM1.	114
4.9	Validation of *T. brucei* lines expressing mNeonGreen tagged TbSNM1.	116
4.10	Validation of *T. brucei* Tb*snm1Δ* expressing mNeonGreen-TbSNM1.	117
4.11	Ectopic expression of TbSNM1 in *T. brucei*	118
4.12	Susceptibility of Tb*snm1Δ* cells ectopically expressing mNeonGreen-TbSNM1 to mechlorethamine	119
4.13	Localisation of TbSNM1 in *T. brucei*.	120
4.14	Effect of exogenous treatments on TbSNM1 expression	121
4.15	Effect of exogenous agents on the temporal expression of TbSNM1.	122
4.16	Gene deletion strategy	123
4.17	Effect of Tb*snmI* disruption on insect form *T. brucei*.	124
4.18	Growth of *T. brucei* Tb*snm1Δ* null mutant	125
4.19	Characterization of *T. brucei* Tb*snm1Δ* null mutants	125
4.20	Susceptibility of *T. brucei* Tb*snm1Δ* null mutants to anti-trypanosomal treatments	127
4.21	Informatic analysis of TbFAN1	129
4.22	3D prediction model of TbFAN1	130
4.23	Validation of *T. brucei* lines expressing mNeonGreen tagged TbFAN1	131

4.24	Localisation of TbFAN1 in *T. brucei* (part 1)	132
4.25	Localisation of TbFAN1 in *T. brucei* (part 2)	133
4.26	Effect of exogenous treatments on TbFAN1 expression	134
4.27	Effect of exogenous agents on the temporal expression of TbFAN1	135
5.1	Examples of various base lesions recognised by the BER pathway.	137
5.2	Mechanism of base excision repair (BER)	138
5.3	Informatic analysis of TbAPE1	143
5.4	3D prediction model of TbAPE1	144
5.5	Informatic analysis of TbMYH	146
5.6	3D prediction model of TbMYH	147
5.7	Informatic analysis of TbNTH1	149
5.8	3D prediction model of TbNTH1	150
5.9	Validation of *T. brucei* lines expressing mNeonGreen tagged TbAPE1	152
5.10	Validation of *T. brucei* lines expressing mNeonGreen tagged TbMYH	153
5.11	Validation of *T. brucei* lines expressing mNeonGreen tagged TbNTH1	154
5.12	Localisation of TbAPE1 in *T. brucei*	155
5.13	Localisation of TbMYH in *T. brucei*	156
5.14	TbNTH1 localisation in *T. brucei*	157
5.15	Effect of exogenous treatments on TbAPE1 expression	159
5.16	Effect of exogenous treatments on TbMYH expression	160
5.17	Effect of exogenous treatments on TbNTH1 expression	161
5.18	Effect of benznidazole on TbMYH and TbAPE1 temporal expression	162
5.19	Effect of nifurtimox on TbMYH and TbAPE1 temporal expression	163
5.20	Effect of fexinidazole on the temporal expression of the *T. brucei*	164
5.21	Effect of mechlorethamine on the temporal expression of TbMYH and TbAPE1	165
5.22	Effect of UV on temporal expression of TbMYH and TbAPE1	166
5.23	Effect of Tb*ape1* deletion on PCF *T. brucei*	169
5.24	Effect of Tb*myh* deletion in PCF *T. brucei*	171
5.25	Effect of Tb*nth1* deletion in PCF *T. brucei*	172
5.26	Cumulative growth of *T. brucei* mutant lines	173
5.27.	Characterization of *T. brucei* Tbape1 null mutants	174
5.28	Characterization of *T. brucei* Tb*myhΔ* “null” cells	174
5.29	Characterization of *T. brucei* Tb*nth1* “null” mutant.	175
5.30	Susceptibility of *T. brucei* mutants to trypanocidal nitroheterocyclic prodrugs	177
6.1	Mechanism of nucleotide excision repair (NER)	180
6.2	Informatic analysis of TbCSB	185
6.3	3D prediction model of TbCSB	186
6.4	Interaction of TbCSB with DNA	187

6.5	Informatic analysis of TbDDB1	188
6.6	3D prediction model of TbDDB1	189
6.7	Informatic analysis of TbRAD23	190
6.8	3D prediction model of TbRAD23.	191
6.9	Informatic analysis of TbXPC	193
6.10	3D prediction model of TbXPC.	194
6.11	Interaction of TbXPC/TbRAD23 with DNA.	194
6.12	Informatic analysis of TbERCC1	196
6.13	Informatic analysis of TbXPF	199
6.14	3D prediction model of TbXPF	200
6.15	3D prediction model showing TbXPF-TbERCC1-DNA interaction	201
6.16	Informatic analysis of TbXPG	202
6.17	3D prediction model of TbXPG	203
6.18	Interaction of TbXPG with DNA	204
6.19	Validation of *T. brucei* lines expressing mNeonGreen tagged TbCSB	208
6.20	Validation of *T. brucei* lines expressing mNeonGreen tagged GG-NER factors	209
6.21	Validating *T. brucei* lines expressing mNeonGreen tagged NER incision factors.	210
6.22	Localisation of TbCSB	213
6.23	Localisation of GG-NER factors	215
6.24	Localisation of incision factors of the *T. brucei* NER pathway	217
6.25	Effect of exogenous treatments on TbCSB, TbXPF or TbXPG expression	220
6.26	Effect of UV on the temporal expression of TbCSB and TbXPF	221
6.27	Effect of mechlorethamine on the temporal expression of TbCSB and TbXPF	222
6.28	Effect of fexinidazole on the temporal expression of TbCSB and TbXPF	223
6.29	Effect of benznidazole on the temporal expression of TbCSB and TbXPF	224
6.30	Characterization of RNAi cell lines targeting the *T. brucei* GG-NER pathway	228
6.31	Characterization of RNAi cell lines targeting incision factors of *T. brucei* NER	229

List of Tables

Table number	Title	Page number
1.1	Differences between the different forms of human African trypanosomiasis	24
1.2	Summary of drug treatments available to treat human African trypanosomiasis	32
1.3	Summary of the *T. brucei* genome	36
1.4	Summary of major cellular proteins encoded by the *T. brucei* genome	36
2.1	Organisms used in this study	46
2.2	Genes studied during this project	49
3.1	Examples of how double strand DNA breaks can arise	63
3.2	Primer combinations used in tagging Tb*mre11 & Tbrad51*	79
4.1	Primer combinations used in CRISPR/Cas9 tagging of Tb*snm1*	115
4.2	Primer combinations used in CRISPR/Cas9 gene deletion of Tb*snm1*.	123
4.3	Susceptibility of *T. brucei* Tb*snm1Δ* null mutants to DNA damaging agents.	126
4.4	Primer combinations used in CRISPR/Cas9 tagging of Tb*fan1*.	130
5.1	Components of the human and *T. brucei* base excision repair pathway	141
5.2	Primer combinations used in tagging Tb*ape1*, Tb*myh* & Tb*nth1*.	151
5.3	*T. brucei* lines expressing tagged base excision repair factors.	151
5.4	Primer combinations used in the CRISPR/Cas9 deletion of BER genes.	167
5.5	*T. brucei* lines potentially null/heterozygous for Tb*ape1*, Tb*myh* or Tb*nth1*	168
5.6	Susceptibility of *T. brucei* mutants to DNA damaging agents	175
6.1	Components of the human and *T. brucei* nucleotide excision repair pathway	
6.2	Primer combinations used in tagging `NER factors	187
6.3	*T. brucei* lines expressing tagged nucleotide excision repair factors	205
6.4	Localisation of NER factors based on predictive algorithms	205
6.5.	Localization of tagged NER factors in PCF *T. brucei*.	210
6.6	Summary of the localisation data.	211
		217
7.1	Summary of the localisation of the different proteins studied in this project	234
7.2	Key features of the different proteins studied in this project	247

List of Symbols and Abbreviations

BER	Base excision repair
BLA/*bla*	Blasticidin–S-deaminase protein/gene
BSF	Bloodstream form
CATT	Card agglutination test for trypanosomiasis
cDNA	Complementary DNA
CSF	Cerebral spinal fluid
DAPI	4',6-diamidino-2-phenylindole
DSB	Double strand break (dsDNA break)
dsDNA	Double stranded DNA
dsRNA	Double stranded RNA
E. coli	Escherichia coli
EC_{50}	Concentration of drug needed to kill 50% of cells
ESAG	Expression site associated gene
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/*hyg*	Hygromycin phosphotransferase protein/gene
ICL	Interstrand Crosslink
Kt (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
NEO/*neo*	neomycin phosphotransferase (*neo*)
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
NLS	Nuclear localisation signal
NTD	Neglected tropical disease
NTR 1	Type I nitroreductase 1
P2	Purine transporter 2

PAC/*pac*	Puromycin-N-acetyltransferase protein/gene
PCF	Procyclic form trypomastigotes
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PHLEO	Phleomycin
PSO2	Psoralen sensitive 2
RT-qPCR	Reverse transcription-quantitative polymerase chain reaction
SNM1	Sensitive to nitrogen mustard 1
SSB	Single strand break
SS-BSF	Short stumpy bloodstream form trypomastigotes
SIF	Stumpy form inducing factor
ssDNA	Single stranded DNA
sg RNA /RNA	Single guide DNA/ RNA
TC-NER	Transcription coupled-nucleotide excision repair
TLS	Translesion synthesis
TF	Tagging fragment
UV	Ultraviolet radiation
UTR	Untranslated region
VSG	Variant surface glycoprotein
WHO	World Health Organisation
WT/wt	Wild type

CHAPTER 1: Introduction

`1.1 Trypanosomatids.

Trypanosomatids are a collection of unicellular eukaryotic organisms classified in the Class Kinetoplastida; Order Trypanosomatida (Stuart et al., 2008). Most are obligate parasites of insects with several evolving heteroxenous life cycles involving a secondary invertebrate, vertebrate or plant host (Roberts et al 2000). The term Trypanosomatid is derived from the Greek *trypano-* (meaning boring) and *soma* (meaning body) and reflects that their cells, in at least one life cycle stage, have a long slender morphology and possess a single flagellum that together impart a corkscrew-like motion to that cell (Figure 1.1). As eukaryotes, trypanosomatid cells contain a nucleus and several other 'typical' organelles, including an endoplasmic reticulum system, Golgi bodies, peroxisomes, and mitochondrion, although the arrangement and function of these structures can be unusual (Figure 1.1). For example, they possess a single, large, and often ramified mitochondrion within which a bar-like structure called the kinetoplast is found (Vickerman 1965, Lukes et al., 2005; Cayla et al., 2019). This region contains the organelle's genome and is composed of a highly complex interconnected network of mini and maxi circle DNAs collectively known as kinetoplast DNA or ktDNAs (Shapiro, 1995; Ryan et al., 2003; Cayla et al., 2019). In the case of other organelles, the trypanosomatid peroxisomes, known as glycosomes, have a modified function with several enzymes of the glycolytic pathway being compartmentalized to this structure (Lopes, 2010; Parsons, 2004; Stuart et al., 2008) and the flagellar pocket, a small invagination of the plasma membrane where the flagellum emerges from the cell body the cytoplasm, serves as the site for endocytosis and exocytosis (Field and Carrington, 2009; Halliday et al., 2021).

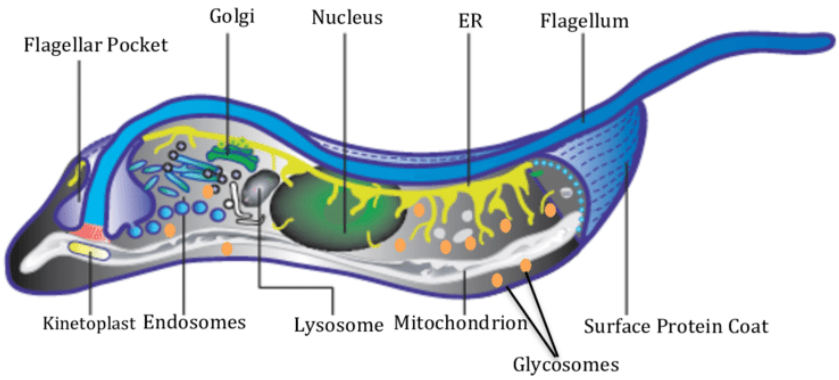


Fig. 1.1. Structure of a trypanosomatid cell. Modified from Overath and Engster 2004.

Trypanosomatids are responsible for several insect transmitted diseases of medical and veterinary importance. In humans, *T. brucei*, *T. cruzi* and 20+ species of *Leishmania* are the etiological agents of human African trypanosomiasis (HAT), Chagas disease, and Leishmaniasis, respectively, while animal African trypanosomiasis is caused by several species including *T. brucei*, *T. congolense* and *T. vivax*. There has been little interest from big pharmaceutical companies to develop specific therapies targeting these human infections resulting in them being classed as a neglected tropical disease (NTD) along with several other conditions that afflict low-income populations living in developing regions (Hotez *et al.,* 2020). In addition to their direct effect on human health, the burden associated with such infections has had a major impact on the socio-economic development of many poor rural communities in Latin America, Asia, and sub-Saharan Africa, effectively trapping sufferers, and their dependents in a disease/poverty cycle. For example, Chagas disease is estimated to cost the economies of Latin American ~US$7.2 billion per annum (Lee *et al.,* 2013), a figure that exceeds the total global burden of uterine, cervical & oral cancers (~US$6.7, 4.7 & 5.3 billion per annum, respectively (Lee et al*.,* 2013)). To help sufferers of NTDs, several non-government organizations such as Drugs for Neglected Tropical Diseases initiative and Médecins Sans Frontières with advice and support from the World Health Organization (WHO), Pan American Health Organization, and The Bill and Melinda Gates Foundation, have developed and implemented new preventative strategies, novel diagnostic procedures, and treatment schedules to combat such debilitating conditions. For trypanosomatid infections, this has resulted in a dramatic reduction in disease prevalence such that the number of new cases of HAT falling from an estimated peak of 325,000 in 1995 to fewer than 1,000 in 2019, while Chagas disease has been eliminated from Chile, Uruguay and several regions of Argentina and Brazil (Schofield et al., 2006, Franco et al., 2020).

1.2. *Trypanosoma brucei* and Human African Trypanosomiasis.

1.2.1 History of Human African Trypanosomiasis.

The association between *T. brucei* and African trypanosomiasis is linked to the trans-Atlantic slave trade, when slave traders incurred huge revenue losses due to sleeping sickness and asked ship surgeons to investigate the cause of this ‘eerie disease’. Thus, the first detailed description of the neurological symptoms of the late phase of the disease were made by the English Surgeon John Atkins in 1734. Another English physician Thomas Winterbottom in 1803 described the main

characteristic signs and symptoms of the disease such as localised lymphadenopathies (swollen lymph glands at the back of the neck), in the initial stage of the disease (Cox, 2004; Steverding, 2008), a characteristic long noted by Arab slave traders who avoided buying such slaves (Steverding, 2008). In 1895, the Scottish pathologist and microbiologist David Bruce discovered *T. brucei* as the cause of cattle trypanosomiasis (known as Nagana). In 1901 British Colonial surgeon, Robert Michael Forde first observed trypanosomes in the blood of a steamboat captain in The Gambia and physician Joseph Everett Dutton later identified these as trypanosomes and in 1902 suggested the name *Trypanosoma gambiense*, now renamed *Trypanosoma brucei gambiense* (*T. b. gambiense*). The discovery of Trypanosoma *brucei* in cerebrospinal fluid (CSF) was concurrently made by the Italian physician and pathologist Aldo Castellani when he examined CSF from patients with sleeping sickness and proposed that these microbes were the causative agents of the disease (Steverding, 2008). Figure (1.2) gives chronological overview of the history and discovery of African trypanosomes.

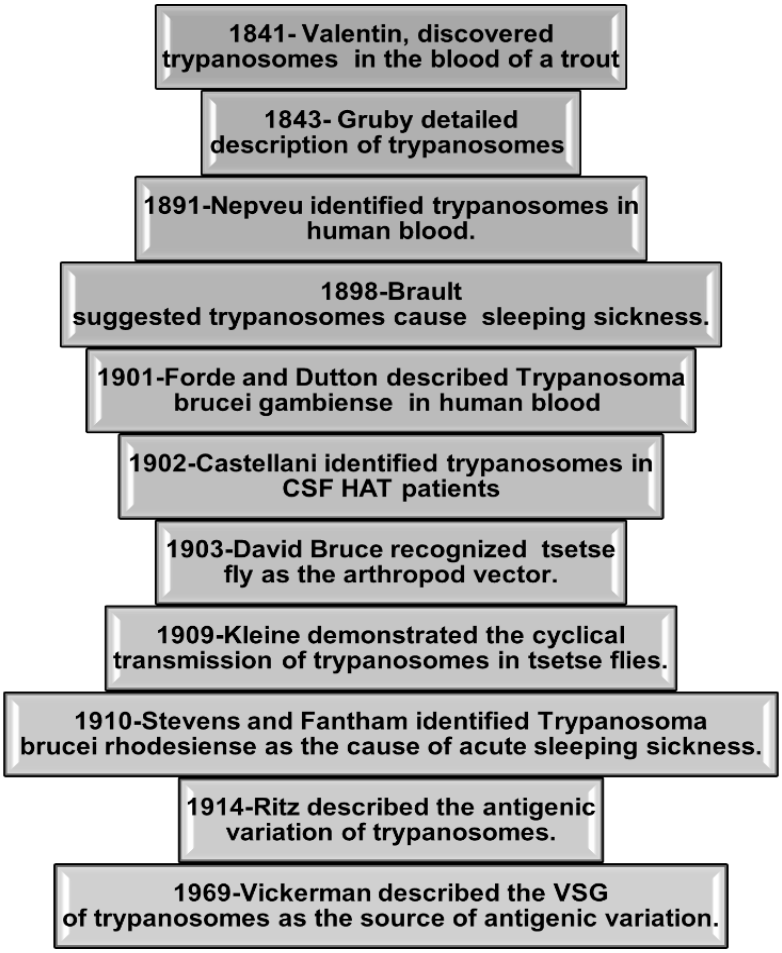


Fig. 1.2: History of African trypanosomiasis. Adapted from Cox, 2004; Steverding, 2008

1.2.2. Historical Perspective of Human African Trypanosomiasis.

African trypanosomiasis is a sporadic disease with the ability to re-emerge after a period of low prevalence or interepidemic periods that are characterised by low prevalence rates followed by increased infection rates that then escalate into an epidemic. Over the last 130 years three major epidemics have been recorded with earliest running from 1896 to 1906 (Figure 1.3). This primarily affected populations living in the Congo basin, and in the Busoga focus of Uganda and Kenya and is estimated to have killed up to 500,000 people (Steverding, 2008). Worried about this infection rates, colonial administration dispatched teams to investigate the cause of the disease and to develop a cure (Hide, 1999; De Raadt, 2005). This led to the identification of several compounds, some of which are still used in the fight against African trypanosomes. Sodium arsenite was shown by Charles Laveran and Félix Mesnil to be an effective treatment for infected laboratory animals while Harold Thomas and Anton Breinl reported that the arsenic-based drug atoxyl could cure experimentally infected animals and could be used to treat animal African trypanosomiasis (Figure 1.3) (Steverding, 2008). However, Robert Koch subsequently demonstrated that the latter was toxic to humans, noting 22 cases of atrophy of the optic nerve with complete blindness in up to 1622 atoxyl-treated patients. Encouraged by these findings, a series of other trypanocidal agents were identified notably Bayer 205 (Suramin) and the organo-arsenic drug tryparsamide that were effective against the early- or late-stage of HAT prevalent in east Africa, respectively (Figure 1.3). These compounds proved their worth in curbing the next major HAT outbreak that ran from 1920 through to the late 1940's across several sites in sub-Saharan Africa (Vickerman, 1965; Steverding, 2008). Here experiences learnt from the 1896-1906 epidemic led to the ruling powers introducing mobile teams to conduct systematic screening, treatment, and surveillance of populations. Continued drug development resulted in additional trypanocidal agents including pentamidine, a diamidine compound that can be used to treat patients in west Africa suffering from early-stage HAT, and melarsoprol, an arsenic compound that targets all forms of late-stage HAT. The combination of these new chemotherapies coupled with disease tracking by case detection and vector control measures by the start of the 1960's led to decline in incidence of sleeping sickness (Figure 1.3). However political instability and conflicts in the newly established independent African states combined with reduced funding for surveillance and disease control activities resulted in the outbreak of a third epidemic that began to emerge in the 1970s and lasted for three decades before finally coming under control in the late 1990s (Franco 2014, WHO Fact Sheet).

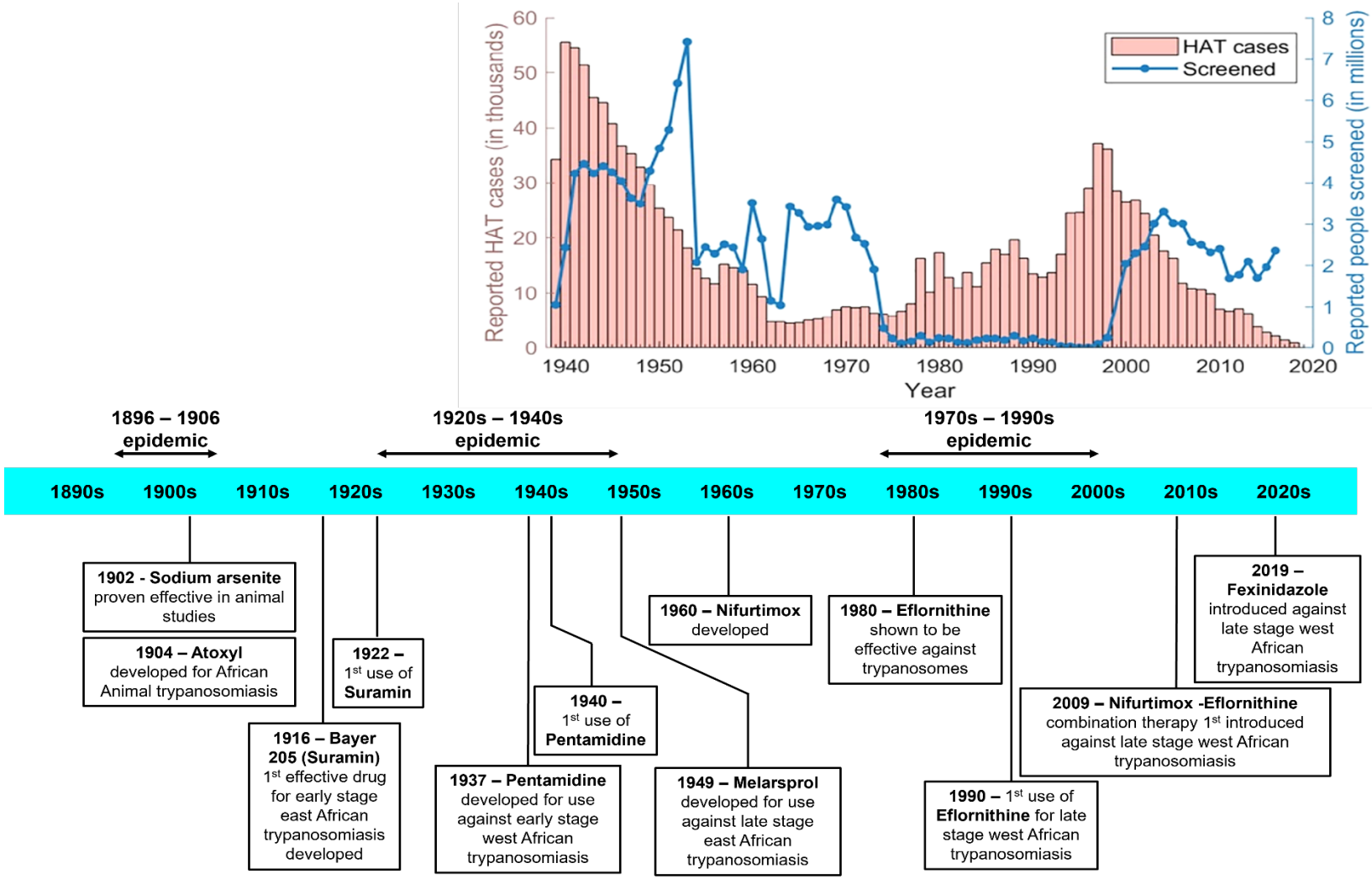


Fig. 1.3: Historical perspectives relating to human African trypanosomiasis.
Timeline of the three major recorded epidemics of human African trypanosomiasis is indicated. The upper graph shows the reported cases of (histogram), and the number of people screened for (line graph) HAT from 1940 to present day (taken from Davis et al., 2020). The lower section provides a timeline of drug development against human African trypanosomiasis (taken from Baker & Welburn, 2018).

1.2.3. Distribution of Disease.

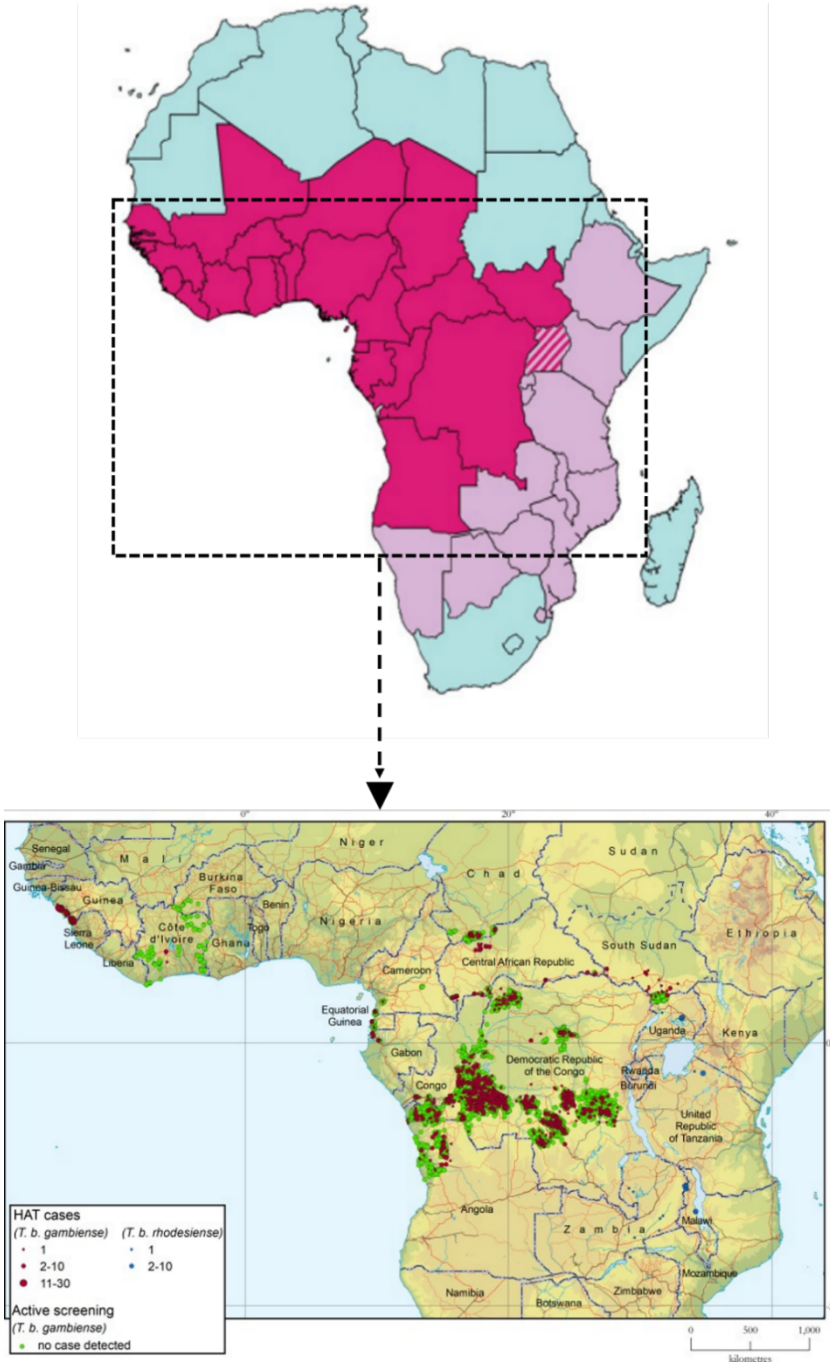


Fig. 1.4: Geographical distribution of human African trypanosomiasis.
In the upper map, the distribution of west (in cerise) and east (in pale pink) Africa for of human African trypanosomiasis is shown. Uganda, a country where both forms of the disease are prevalent, is highlighted (hatched). Regions that have low disease prevalence or absence of disease are highlighted in pale blue. In the lower map, an expanded view centred on the Democratic Republic of Congo and neighbouring countries shows where HAT cases have been recorded (Franco et al 2020).

Human African trypanosomiasis, also known as African sleeping sickness, is an insect transmitted infection recorded in 36 countries of sub-Saharan Africa (Figure 1.4) (Simarro et al., 2012). Of all African countries, the Democratic Republic of Congo (DRC) bears the highest disease burden with >70% new cases arising at this site: In 2012, for example, of the 7,106 new HAT cases reported in Africa, 5,968 (or 84%) were in DRC (Lumbala et al., 2015). Other sites where this disease is common include Angola, Cameroon, Central African Republic, Chad, Congo, Guinea, Malawi, South Sudan, and Zambia where moderate HAT prevalence ranging from 10-100 cases per annum are noted while Côte d'Ivoire, Equatorial Guinea, Gabon, Uganda, United Republic of Tanzania, and Zimbabwe have all recorded lower prevalence (1 and 10 cases) in 2019 (WHO, 2021). Additionally, and over the last decade, sporadic cases have been observed in Burkina Faso, Ghana, Kenya, and Nigeria. Although these cases are reported annually through the WHO, it is believed that the published figures do not reflect the true impact of HAT with estimates suggesting that the actual number of cases may be up to 3-fold higher than the reported number of cases (Odiit et al., 2005). Additionally, several countries in the HAT foci such as Gambia, Guinea-Bissau, Liberia, Senegal, Sierra Leone and Niger do not have any active HAT surveillance systems in place (Franco et al., 2020).

In the endemic countries, the areas where HAT develops can range from a single village to an entire region, with disease prevalence in the latter situation varying from village to village. Many of the affected individuals live in remote, rural communities where access to adequate health services is limited, and where disease surveillance/treatment is difficult to conduct. This situation is complicated by population displacements, political instability, and poverty. Current estimates indicate that approximately 53 million people live in "at risk" areas with 7.5 and 1 million individuals inhabiting moderate and high-risk regions, respectively. A major determinate in the distribution of HAT is the geographical range of the Tsetse fly insect vector. Several *Glossina* species including *G. palpalis*, *G. tachinoides*, *G. fuscipes*, *G. morsitans* and *G. swynnertoni,* can transmit the *T. brucei* subspecies responsible for HAT, with these distinguished from each other in terms of their habitats (woodland, riverine or savannah), behaviour and morphology. Unlike some other protozoal infections such as malaria and leishmaniasis, the heamatophagous feeding habits of both male and female Tsetse flies can spread the disease, with the parasite transferred between hosts in the insect's saliva.

To curb HAT transmission and aid in disease elimination, at the start of the 21st century public-private partnerships involving WHO, pharmaceutical companies Sanofi and Bayer, National Sleeping Sickness Control Programmes, non-governmental and philanthropic organizations, resulted in the implementation of novel diagnostic & vector control programmes as well as the introduction of the new therapies. This led to significant reduction in prevalence of HAT with the number of cases falling from record ~25,000 in 2000 to around 10,000 in 2007 (Figure 1.5) (Franco et al 2022). The introduction of the Nifurtimox Eflornithine combination therapy (NECT) in 2009 targeting late-stage *T. b. gambiense* HAT maintained the rate of reduction such that by 2018 a major milestone was reached with the number of reported HAT cases falling below 1,000 (977). With the arrival of Fexinidazole, the first orally administered treatment that is active against all stages of the HAT form present in West Africa, it is predicted that maintenance of this low disease level is now a realistic goal and makes the WHOs goals of eliminating HAT as a public health problem (<1 HAT reported case per 10,000 people per annum) and stopping disease transmission by 2030 achievable (Figure 1.5) (Franco et al., 2022).

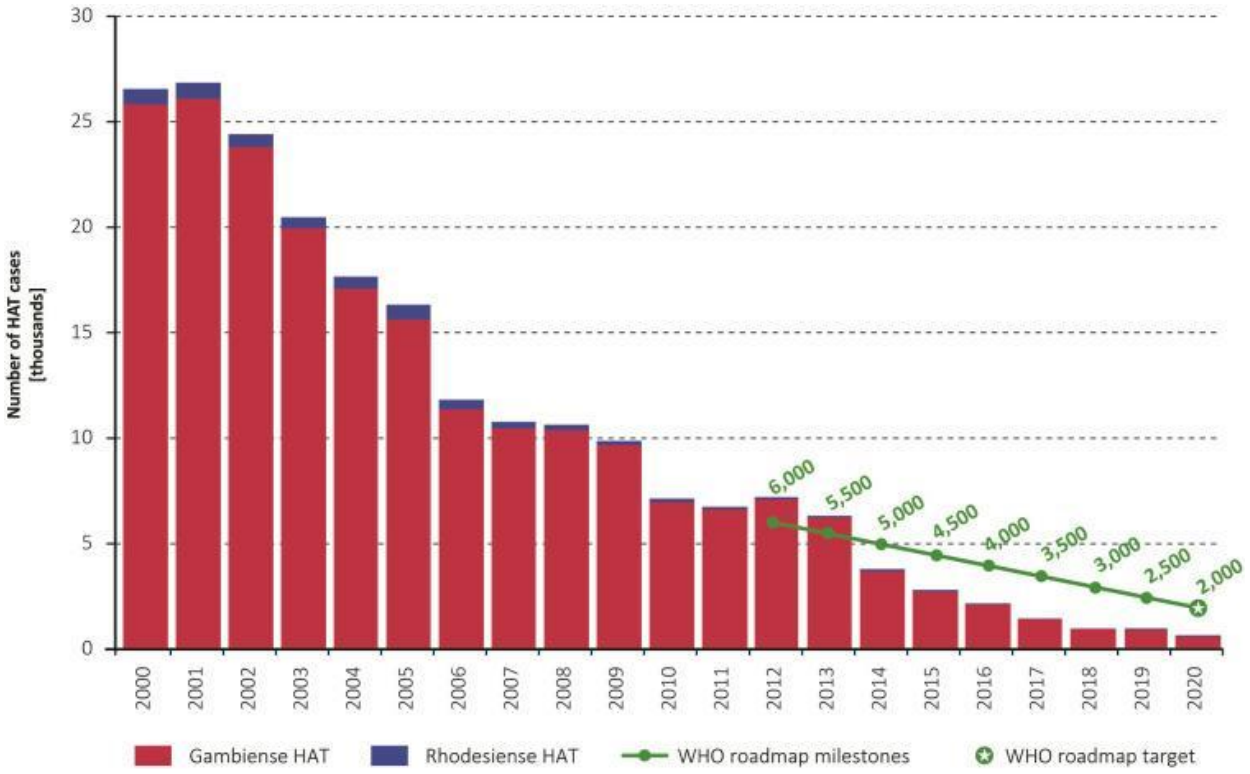


Fig. 1.5. Decline in human African trypanosomiasis between 2000-2020.
Image taken from Franco et al 2022.

1.2.4. The Causative Agent.

Based on geographic location, mechanism of transmission, epidemiology, and rate of disease progression, HAT can be split into two distinct variants, each caused by different sub-species of *T. brucei* (Table 1.1). In West and Central Africa, West African trypanosomiasis, caused by *T.b. gambiense*, predominates. This form of HAT accounts for more than 95 % of all HAT cases and is typically transmitted from human-to-human by riverine-type Tsetse flies (WHO, 2021). Parasitaemia in the mammalian host is generally low, with asymptomatic carriers being common. The disease is classed as a chronic infection as it can take months or even years for the disease to progress from the early to late stage. The second HAT variant, known as East African trypanosomiasis, is prevalent in East and Southern Africa. Caused by *T. b. rhodesiense*, this predominantly zoonotic infection only accounts for 5 % of all HAT cases and is spread via the blood-feeding habits of savannah-type Tsetse flies (WHO, 2021). In infected individuals, parasitaemia can reach high levels. This can promote an acute form of African trypanosomiasis rapidly progressing from the early to late stage in a little as 1 to 3 weeks.

	West African HAT	**East African HAT**
Causative agent	*T.b. gambiense*	*T.b. rhodesiense*
Vector	Riverine (*G. palpalis)*	Savannah (*G. morsitans)*
Ecology	Rainforest, rivers, and lakes	Woodland and dry bush
Location	West and Central Africa	East and Southern Africa
Transmission	Human to human	Zoonotic
Epidemiology	Endemic, some epidemics	Sporadic outbreaks
Prognosis	Slow progression to chronic disease	Rapid progression to death
Parasitaemia	Low	High
Asymptomatic carriers	Common	Rare
Non-human hosts	Rare	Wild and domestic animals (ungulates)

Table 1.1: Differences between the different forms of human African trypanosomiasis. Adapted from Wiser, 2011.

1.2.5. Life Cycle of *Trypanosoma brucei.*

To deal with the various environments *T. brucei* is exposed to during its life cycle, the parasite exhibits several distinct forms (Figure 1.6).

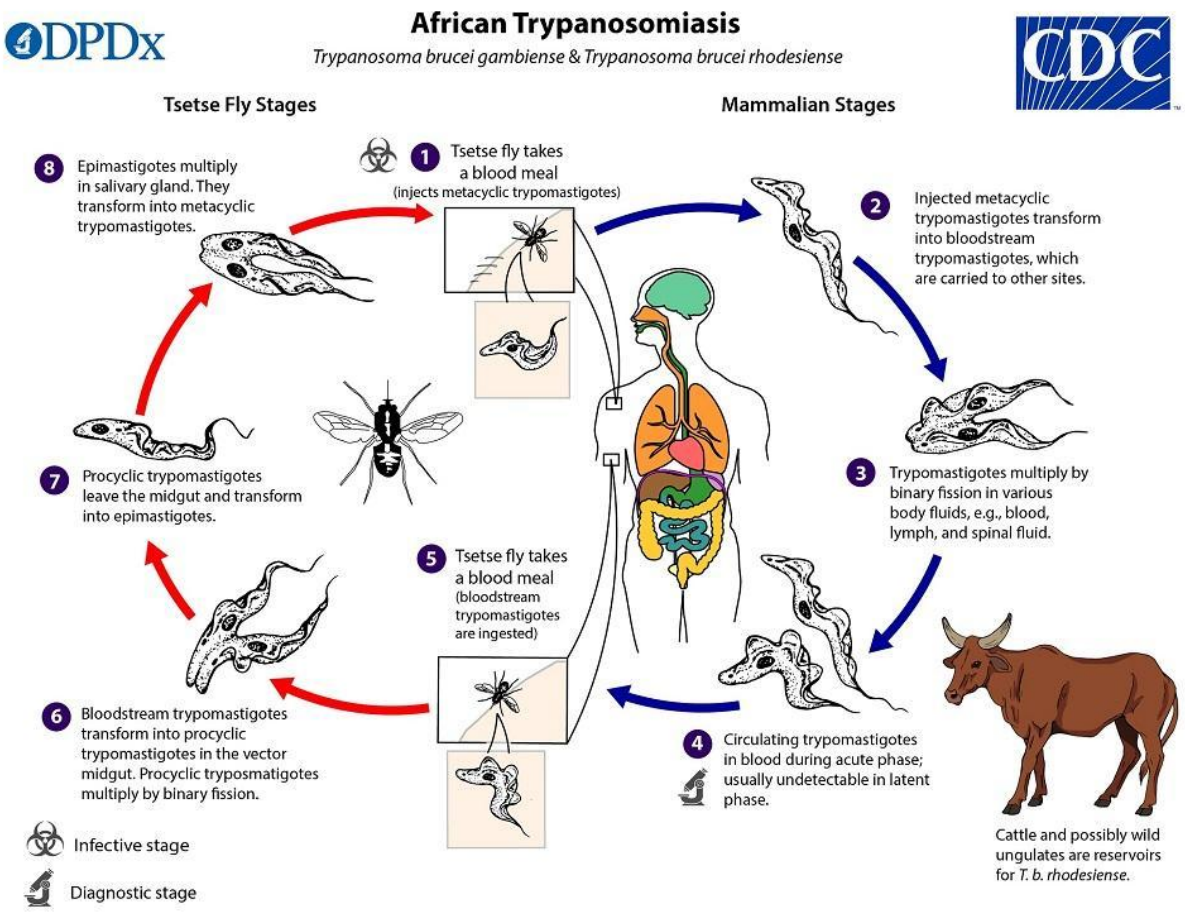


Fig. 1.6 Life cycle of *Trypanosoma brucei*.
Image sourced from (Center for Disease Control (CDC) 2020; http://www.cdc.gov/parasites/sleepingsickness/biology.html last accessed 9/11/2022).

Transmission of the parasite from the Tsetse fly to a mammalian host occurs when the vector takes a blood meal. Here, insect saliva containing highly motile, non-dividing metacyclic trypomastigotes is injected into the victim's blood and/or lymphatic systems. Possibly in response to an upshift in temperature (Kolev et al 2018), the parasite transforms into motile long slender bloodstream form trypomastigotes (LS-BSF) that divide by binary fission. Via an antigenic variation mechanism that involves the periodic alteration of the expressed dominant surface antigen, which in the case of LS-BSF *T. brucei* is the variant surface glycoprotein (VSG), the extracellular form parasites evade antibody- and complement-mediated immune destruction (Horn, 2014). They can go on to establish a systemic infection and if left untreated, can eventually cross the blood brain barrier to infect the CSF. Driven by the antigenic variation immune evasion process, the parasite numbers in the blood fluctuates to generate a characteristic "wave of

parasitaemia" profile. At the peak of a wave, a quorum sensing-like mechanism driven by the so-called "stumpy induction factor" operates, resulting in the differentiation of LS-BSF trypomastigotes to non-dividing, short stumpy form (SSF) trypomastigotes, a parasite stage that is pre-adapted for conditions found in the insect vector (Mony et al., 2014; Mony & Matthews, 2015; Rojas et al., 2019; Matthews, 2021).

When another tsetse fly takes a blood meal, the insect ingests blood or lymph containing LS-BSF and SSF parasites. Inside the insect's digestive tract, and surrounded by the peritrophic matrix, the LS-BSF die off while the SSF survive. Here, these cells undergo a further round of differentiation in response to the temperature drop and presence of metabolites (*e.g.,* citrate/*cis*-aconitate) to generate motile, procyclic form (PCF) trypomastigotes.

Next, the PCFs escape from the peritrophic matrix and become attached to the insect's midgut epithelium lining where they divide by binary fission and eventually differentiating into mesocyclic trypomastigotes. The mesocyclic trypomastigotes then transform into epimastigote (EP) with their nucleus located to the posterior side of the kinetoplast (sharma et al., 2008, Rotureau & Abbeele, 2013).

Next, the parasite is released into the midgut lumen and eventually finds its way to the salivary gland. Here, the EP binds *via* its flagellum to the epithelial lining of the salivary gland, divides by binary fission and undergoes genetic exchange, although this is not an obligatory part of the *T. brucei* life cycle (Sternberg and Tait, 1990; Gibson and Stevens 1999: Ralston et al. 2009). As the numbers of EP increase, some cells will transform into MT parasites, entering the lumen of the salivary gland, ready for transmission into the next mammalian host (Stevens et al., 2001; Mathews, 2005).

1.3. Disease Staging, Signs and Symptoms.

Human African trypanosomiasis manifests itself as two distinct phases; the early stage (also known as haemolymphatic or stage 1) and the late stages (also known as meningoencephalitic or stage 2) (Barrett et al., 2003; Stuart, et al., 2008). Disease progression from the former to the latter is dependent on the infecting sub-species, with these taking months/years in the case of the West Africa trypanosomiasis or a few weeks for East African trypanosomiasis (Figure 1.7) (Brun et al., 2010). During the early stage of the disease, parasites are restricted to the bloodstream and lymphatic system. Here, most patients are asymptomatic although approximately 50% of people

infected with *T.b. rhodesiense* develop a painful, localised skin inflammation known as trypanosomal chancre at the site of the insect bite (Figure 1.7) (Brun et al., 2010). When symptoms are displayed, they tend to be non-specific and are often mistaken for other infections including Malaria. These can range from an intermittent fever, headaches, and joint pains with further complications such as inflamed lymph glands at the back of the neck and spleen, local oedema, and cardiac abnormalities, also noted. If left unchecked, the parasite can eventually cross the blood brain barrier, gaining access to the CSF. It is at that point the disease enters the late stage. Here manifestations can include neurologically related conditions such as severe headaches, personality change, mental impairment, and weight loss with the patient often experiencing a reversal of their sleep-wake cycle, resulting in the alternative name of HAT, African Sleeping Sickness. If left untreated, the patient begins to experience convulsions, becomes comatose which typically results in death of that individual (Wiser, 2011).

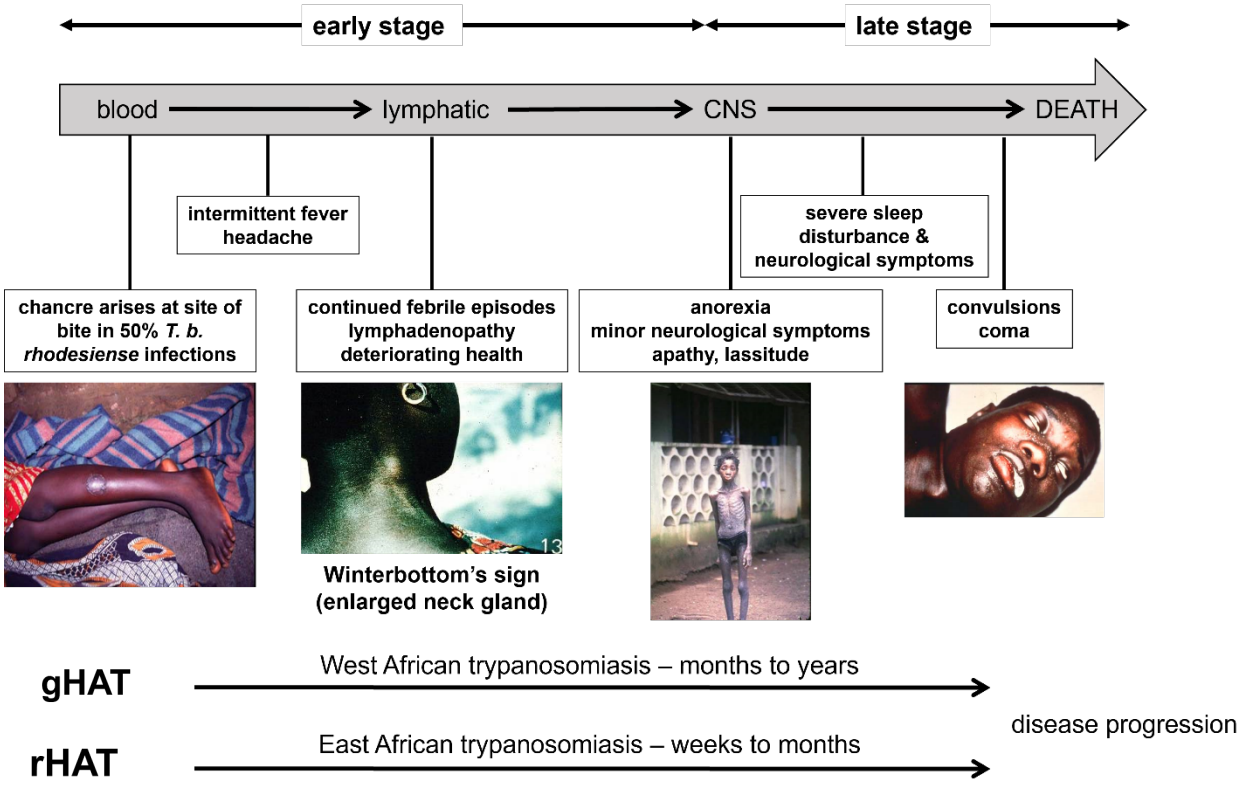


Fig. 1.7. Disease staging, and sign and symptoms of human African trypanosomiasis. Adapted from Wiser, 2011. gHAT and rHAT correspond to West African trypanosomiasis caused by *T.b. gambiense* and East African trypanosomiasis caused by *T.b. rhodesiense*, respectively while CNS is the central nervous system.

1.4. Diagnosis of Human African Trypanosomiasis.

The protocol used for HAT diagnosis is dependent on geographical location. In areas where west African trypanosomiasis is endemic, WHO recommends that the Card Agglutination Test for Trypanosomiasis (CATT) is used to identify infected individuals, with disease staging established by the microscopic detection of parasites in the patient's blood and CSF (Figure 1.8).

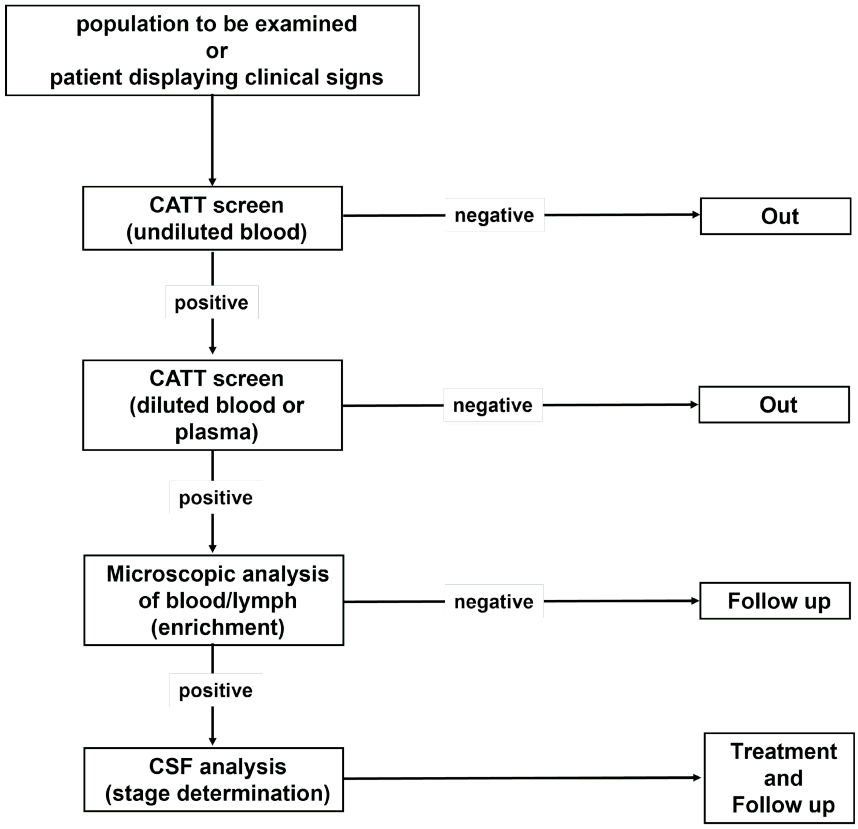


Fig. 1.8. WHO guidelines for the detection of West African trypanosomiasis.

In an initial triage, undiluted blood taken from patients displaying clinical signs of HAT or villagers living at an endemic region is screened using CATT (Figure 1.9). Developed during the 1970's, this serological assay relies on detection of the variable antigen type LiTat 1.3 that is expressed by *T.b. gambiense* and has a sensitivity and specificity of about 91 and 97 %, respectively (Bonnet et al., 2015): CATT cannot be used to *T.b. rhodesiense*. If the above results in a positive outcome, the CATT is repeated this time using a diluted blood sample. If a negative result is obtained for either assay, no further testing is required.

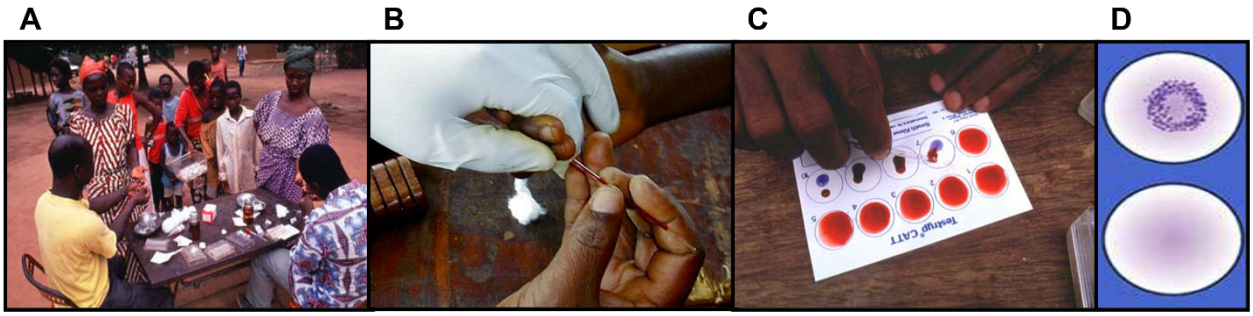


Fig. 1.9. Detection of *T.b. gambiense* using the Card Agglutination Test for Trypanosomiasis. HAT symptomatic patients or villagers living at an endemic site (A) present themselves to a surveillance team who then take a blood sample (B). The blood is drawn into a capillary tube and then mixed with fixed, stained parasites immobilized on a plastic card (C). If a patient's sera react with the fixed, stained parasites, a blue granular deposit is observed (D).

For individuals who are positive after the second CATT, microscopic screens are performed to directly detect the parasite firstly in blood samples and then in CSF (Figure 1.10). As parasitaemia is generally low for patients infected with *T.b. gambiense*, an enrichment step involving the use of DEAE cellulose minicolumns or microhaematocrit centrifugation is employed to increase the chance of observing the parasite under the microscope. If a patient is CATT positive but whose blood is microscopically negative for the parasite, further examination of this fluid is done to account for possible fluctuations in parasitaemia. For individuals whose blood is *T.b. gambiense* positive using both tests, microscopic analysis of the CSF obtained by lumbar puncture is performed with this done no more than 20 minutes after extraction. If the patient's CSF is deemed to be negative for the parasite, that individual is classified to be in the early stage of the disease, whereas if this sample does contain *T.b. gambiense*, that person is considered to be in the late stage of HAT. Based on this an appropriate course of chemotherapy can then be administered.

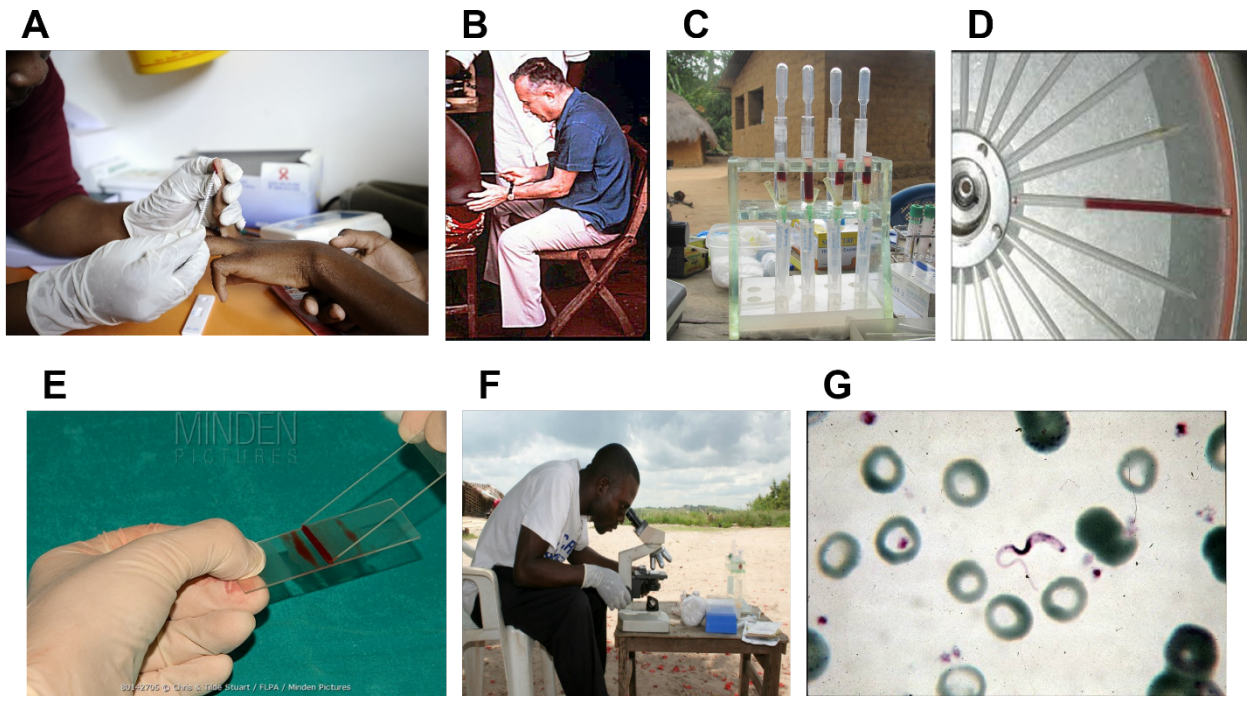


Fig. 1.10. Microscopic detection of *T. brucei*.
The blood or cerebrospinal fluid sample is taken from an individual as a pinprick (A) or lumbar puncture (B), respectively. The sample is smeared onto a slide (E), often treated with a cytologic stain such as Giemsa stain and then viewed microscopically (F) for the presence of parasites (G). For people living in areas where west African trypanosomiasis is endemic, any parasites in the blood sample may require concentrating using DEAE cellulose minicolumns (C) or microhematocrit centrifugation (D).

In areas where east African trypanosomiasis is prevalent, no tests equivalent to CATT is available and hence diagnosis is all based on microscopic detection of the parasite from blood and CSF samples. As the *T.b. rhodesiense* parasitaemia is generally high, no enrichment step is required when analysing blood with direct detection on a microscope slide sufficient to observe the parasite.

1.5. Treatment of human African trypanosomiasis

Currently, chemotherapy represents the only option available to treat HAT and involves a small number of licensed drugs which include: Suramin, Pentamidine, Melarsoprol, Nifurtimox-Eflornithine combination therapy (NECT) and Fexinidazole (Figure 1.11). Details relating to the mode of action of each drug, which disease stage/*T. brucei* subspecies they target, and problems associated with their use can be found in Table 1.2

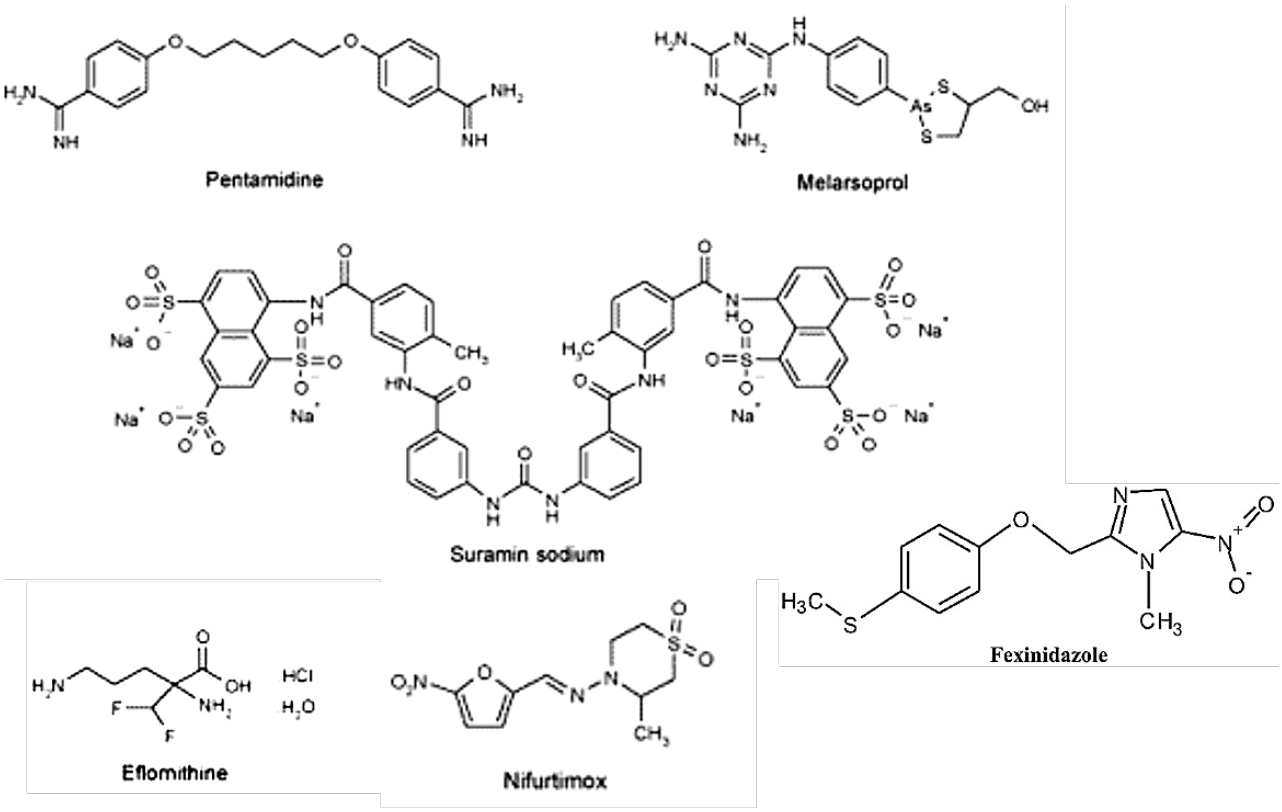


Fig. 1.11. Structures of the drugs used to treat human African trypanosomiasis. (See table 1.2 for details and references)

Drug (class)	Mode of action	Uses	Obstacles	References
Suramin (Polysulphonated naphthalene)	Glycoprotein ISG75 and four lysosomal proteins have been linked with efficacy and resistance. Function unclear 'promiscuous' inhibitor of an array of glycolytic and pentose phosphate pathway enzymes	First line of treatment for early stage of *T.b. rhodesiense* infections	Negatively charged cannot cross the blood-brain barrier. Administrated intravenously. Skin reactions, nerve damage, haematological toxicity.	Nok 2003; Legros et al., 2002; Bacchi 2009; Wilkinson & Kelly 2009; Alsford et al., 2012
Pentamidine (Aromatic diamidine)	P2 adenosine transporter 1 (AT1) and an aquaglyceroporin (AQP2/3) linked with uptake and resistance. Readily metabolized by human cytochrome P450 reductase. Accumulation of drug collapses trypanosome mitochondrial membrane potential. Nucleic acid binder to minor groove of minicircles.	First line of treatment for early stage of *T.b. gambiense* infections	Does not readily pass through the blood brain barrier. Generally, well tolerated	Shapiro & Englund 1990; Wilkinson & Kelly 2009; Alsford et al., 2012; Baker et al., 2013
Melarsoprol (Trivalent arsenical)	P2 adenosine transporter 1 (AT1) and an aquaglyceroporin (AQP2/3) linked with uptake and resistance. Prodrug with activation leading to formation of melarsen oxide (MelOx). Hinders adenine uptake. Inhibits a range of glycolytic enzymes and forms stable adducts with trypanothione thus block activity of trypanothione reductase.	Second line treatment for early stage of *T.b. rhodesiense* and *T.b. gambiense* The only treatment for late stage of *T.b. rhodesiense* infections.	Highly toxic causing 10% of patients to develop reactive encephalitis. Side effects include convulsions, loss of consciousness and vomiting.	Fairlamb et al. 1989; Stich et al., 2002; Bacchi 2009; Wilkinson & Kelly 2009; Schumann Burkard et al., 2011 Alsford et al., 2012; Baker et al., 2013
Nifurtimox (5-nitofuran)-Eflornithine (Fluorinated amino acid) combination therapy (NECT)	Uptake of Eflornithine *via* amino acid transporter (AAT6). Inhibitor of ornithine decarboxylase prevents production of putrescine. Nifurtimox is a prodrug and activated by a type I nitroreductase leading to cytotoxic products. Reduction of type I nitroreductase activity linked to resistance.	First line of treatment of late stage of *T.b. gambiense* infections.	Eflornithine – Poor transport across blood-brain barrier.; Costly, complicated, and difficult to administer (intravenous infusion); side effects include GI problems, convulsions, and fever. Nifurtimox – Orally administered; side effects include GI tract disturbances, abdominal pain, vomiting common. Convulsions, tremor & agitation, polyneuropathy have been reported.	Checchi et al., 2007; Steverding, 2008; Bacchi 2009; Wilkinson & Kelly 2009; Burri, 2010; Vincent et al., 2010; Hall et al., 2011; Wiser, 2011; Alsford et al., 2012; Franco et al., 2012; Alirol et al. 2013
Fexinidazole (nitroimidazole)	methylsulfanyl group metabolized by liver cytochrome P450 enzymes to sulfoxide & sulfone metabolites. Parent compound & metabolites function as type I nitroreductase activated prodrugs. Reduction of type I nitroreductase activity linked to resistance.	First oral treatment against early and late stage of *T.b. gambiense* infections.	Orally administered; side effects include headache, vomiting, insomnia, nausea, asthenia, tremor, decreased appetite, dizziness, hypocalcaemia, dyspepsia, back pain, upper abdominal pain, and hyperkalaemia	Torreele et al., 2010; Kaiser et al 2011; Wyllie et al., 2012; Wyllie et al 2016; Deeks, 2019; (European Medicines Agency, 2019)

Table 1.2: Summary of drug treatments available to treat human African trypanosomiasis

The use of anti-HAT drugs is problematic as they often have limited efficacy, may promote unwanted toxic side effects and can be costly to synthesize and/or administer, while non-responsive/resistant strains are emerging at endemic sites (Table 1.2). For example, drugs that have been used to treat HAT since 1922 and 1940, Suramin and Pentamidine, respectively (Figure 1.3), are both charged at physiological pH (Wilkinson and Kelly, 2009). As such, neither compound can cross the blood brain barrier to reach the therapeutic levels required to treat the parasite forms found in the CSF resulting in their use to be limited to patients in the early stage of HAT. Additionally, the use of these drugs is further restricted as Suramin is only used against east African trypanosomiasis while Pentamidine is used to target *T.b. gambiense* infections. Similarly, Melarsoprol, which has been in use since 1949, has long been the only effective treatment against the late stage of both HAT variants (Friedheim, 1949). However, the slow intravenous administration of this arsenic-based drug involves using the adjuvant propylene glycol that causes a very painful burning sensation due to scarring of veins and blood vessel collapse. Additionally, Melarsoprol itself promotes the potentially fatal condition reactive encephalitis in 5–10% of cases as well as other side effects including convulsions, pyrexia, loss of consciousness, rashes, bloody stools, nausea, and vomiting (Enanga et al 2002). Worryingly, parasite strains that are resistant towards Pentamidine and/or Melarsoprol are becoming more commonplace in the clinical setting with such lines generally having a reduced ability to take up these drugs (Bernhard et al 2007; Nerima et al 2007; Graf et al 2013).

Since the turn of this century, there have been concerted efforts to find alternative, cheaper, and safer anti-HAT therapies. Introduced in 2009, NECT currently represents the WHO Essential Medicine List of preferred treatment options against late stage west African trypanosomiasis. This superseded the previously favoured eflornithine monotherapy primarily based on cost and ease of administration: NECT requires fewer Eflornithine intravenous infusions relative to the monotherapy (14 instead of 56) and has a shorter treatment time (10 vs 14 days), features that make the combination more cost effective while requiring less hospitalization resulting in greater patient acceptance (Priotto et al., 2009; Yun et al., 2010). Additionally, NECT has a higher cure rate (95-98%) than Melarsoprol or Eflornithine monotherapies with few side effects. The use of this combination therapy alongside implementation of improved HAT surveillance programmes and political will, has resulted in the dramatic fall in the reported number of HAT cases seen today in sub-Saharan Africa.

Fexinidazole, a new, potentially ground-breaking anti-HAT therapy has been identified by a Drugs for Neglected Diseases initiative driven programme. In 2018 and after extensive data mining and appropriate safety trials, the 5-nitroimidazole compound Fexinidazole, was approved by the European Medicines Agency for use against both stages of west African trypanosomiasis (Deeks, 2019). This represents the first exclusively orally administrated therapy against HAT and eliminates the need for lumbar puncture when determining disease staging. Moreover, patients can be treated at home with this, for the first-time, making HAT therapy genuinely cost effective.

1.6. *Trypanosoma brucei* molecular biology

1.6.1. The genome of *Trypanosoma brucei*

T. brucei possess a nuclear and mitochondrial genome, with the latter known as the kinetoplast. The haploid nuclear genome is about 35 megabases (Mb) in size (El-Sayed et al., 2005) and divided across three linear chromosomal types, namely 11 pairs of megabase chromosomes, 1 to 5 intermediate chromosomes and 60 to 100 minichromosomes that are sized from 0.9-5.7 Mb, 300-900 kb and 50-100 kb, respectively (Table 1.3) (Berriman et al., 2005). The megabase chromosomes account for ~26 Mb of the genome and contains about 9000 pre-rRNA, tRNA, and snRNA, small nucleolar (sno) RNA, short interfering (si) RNA, spliced leader (SL) RNA, and mRNA coding genes (Figure 1.12; Table 1.4) (Berriman et al., 2005). Most coding sequences are intron-less with genes coding for proteins of unrelated function arranged in large polycistronic clusters and separated by polypyrimidine rich intergenic spaces (Mair et al., 2000). Many of the coding sequences in the trypanosome genome encode for a diverse repertoire of proteins involved in key cellular processes (Table 1.4). This includes those that code for cytoskeletal structural proteins (e.g. actin filaments and microtubules), metabolic enzymes involved in glycolysis, electron transport and oxidative phosphorylation, factors implicated in amino acid uptake and synthesis/biogenesis, various surface proteins that aid the parasite to evade detection and killing by host (e.g. the variant surface glycoprotein, GPEET and procyclin), as well as coding for a raft of intracellular proteins including GTPases for vesicle trafficking machinery, metabolism proteins, glycosylphosphatidylinositol anchor biosynthesis etc.

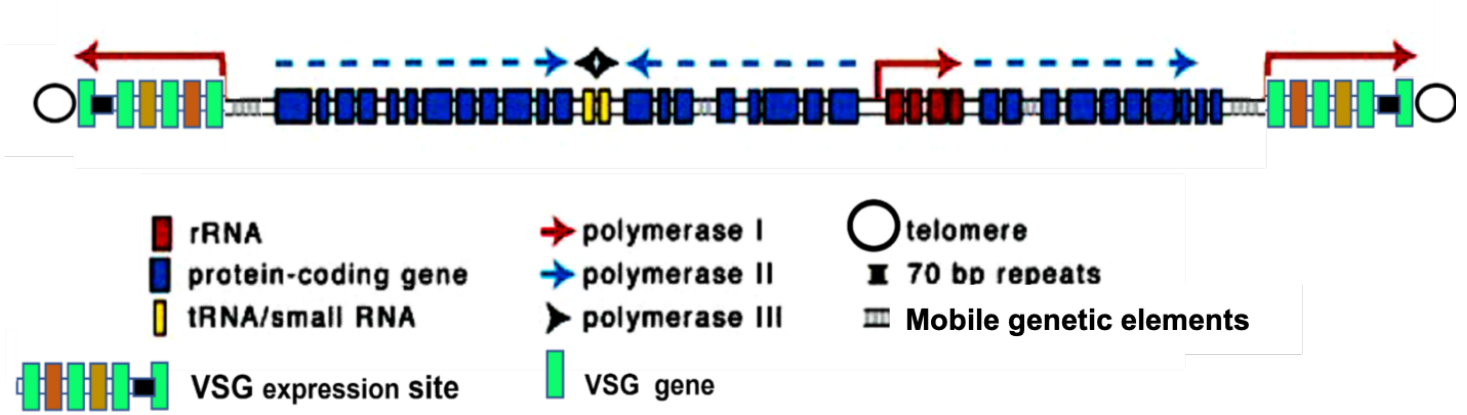


Fig. 1.12: Schematic showing generalised features found on a *T. brucei* chromosome
The diagram was taken from Clayton et al (2002), with modification and adaptation from Berriman et al, 2005. The genome of *T. brucei* is made up of protein coding genes, rRNA genes, tRNA and small RNAs (snRNA). Genes encoding for the variable surface glycoprotein (VSG) are located towards the telomeres of mega and intermediate chromosomes. The sub-telomeric regions contain special transcription units known as expression sites that house the expressed VSGs and additional genes known as the expression site associated genes (ESAGs) that code surface proteins that together with the VSGs contribute to parasite virulence.

Parameter		Number
Genome size (bp)		35000000
mega chromosomes	0.9-5.7 Mb	11
intermediate chromosomes	300-900 kb	5
Minichromosomes	50-100 kb	60-100
G+C content (%)		46.4
Percent coding sequence		50.5
Protein-coding genes		
Genes		9068
Pseudogenes		904
Mean CDS length (bp)		1592
G+C content (%)		50.9
Gene density (genes per Mb)		317
RNA genes		
transfer RNA		65
ribosomal RNA		56
slRNA		>28
small nuclear RNA		928 5
small nucleolar RNA		353

Table 1.3: Summary of the *T. brucei* genome.
Table taken from Berriman et al., 2005. Genome size and chromosome numbers are given for all 3-types of chromosomes.

Cellular process	Components
Antigenic Variation	VSG
Intracellular protein vesicle trafficking	endoplasmic reticulum & golgi network plus intermediary proteins, clathrin- dependent transport flagellar through pocket, Rab-GTPases, transport system
Cytoskeleton	microtubule and tubulin
Metabolism and transport	carbohydrate metabolism electron transport and oxidative phosphorylation
Anchor to major surface glycoproteins	glycosylphosphatidylinositol anchor
polyamine biosynthesis	trypanothione

Table 1.4: Summary of major cellular proteins encoded by the *T. brucei* genome.
Table taken from Berriman et al., 2005.

The mitochondrial genome is made of two classes of highly concatenated, circular DNAs known as maxi- and mini-circle DNAs. In each kinetoplast there are 20 to 50 maxi-circle DNAs

which are 23 to 40 kb in size. They code for ~16 proteins involved in mitochondrial oxidative phosphorylation (e.g., cytochrome b, subunits of cytochrome oxidase, and subunits of NADH dehydrogenase) as well as for ribosomal RNAs (Ramrath et al., 2018; Shapiro, 1995). In several cases, the 'coding' sequence represents a pseudogene with such transcripts needing post-transcriptional modification to form a functional mRNA, a process known as RNA editing (Hajduk & Ochsenreiter, 2010; Simpson et al., 2003). In addition, each kinetoplast also contains between (5000 to 10,000 mini-circle DNAs that are 0.5-0.5 kb in size. These exhibit a high level of sequence variability and code for specialised RNA molecules known as guide RNA (gRNA) (Cooper et al., 2019). The gRNAs function to direct a multi subunit protein complex called the editosome to sites in a pseudogene mRNA that require editing, thereby facilitating the insertion or deletion of nucleotides, generally uridine residues, at this region.

1.6.2. Gene expression in Trypanosomes

Although divergent, many aspects of trypanosomal gene expression are comparable to systems found in other eukaryotes. *T. brucei* expresses three RNA polymerases, Pol I, II and III. RNA Pol I transcribes the pre-rRNA (18S, 5.8S and 28S), the gene cluster that encode the bloodstream-form variant surface glycoproteins (VSGs) and the procyclic-form procyclins (Eps and GPEETs) (Günzl et al., 2003; Pays, 2005). In contrast, RNA Pol II transcribes mRNAs, the trypanosome-specific, cap 4 17-(m7Gpppm6,2AmpAmpCmpm3Um), and SL RNA genes (Bangs et al., 1992; Gilinger & Bellofatto, 2001) while RNA Pol III transcribes tRNAs, 5S RNA and the U-rich snRNAs. For genes where high protein levels are needed, genomic duplication events have occurred resulting in amplification of the open reading frames (ORFs). This often leads to multiple copies of the same gene arranged in tandem that generates higher levels of the transcript. In other cases, elevated transcription levels can be achieved by using RNA Pol I in place of RNA Pol II as noted in VSG expression.

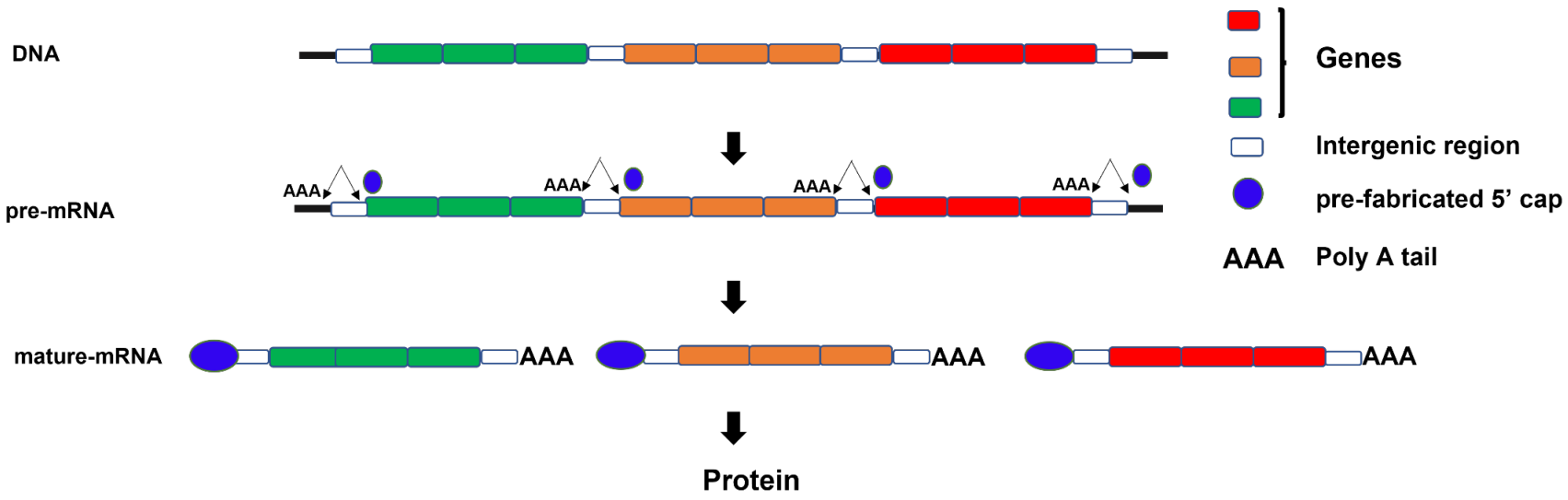


Fig. 1.13 Schematic illustrates a generalized polycistronic transcription unit on *T. brucei chromosome*.
Three different genes are transcribed into a polycistronic pre-mRNA, which is then post-transcriptionally modified by simultaneous capping with a 39-nucleotide sequence and polyadenylation. This results in the mature mRNA, which is then translated into the protein. Control of gene expression lies in the 3' UTR: this region controls mRNA abundance by rapid degradation of unwanted RNA before they are translated into proteins. Image adapted from Clayton et al (2002),

Most trypanosomal genes are constitutively expressed and not transcriptionally regulated, although there are several notable exceptions (e.g., in the expression of the highly regulated VSG, GPEET and procyclin genes) (Clayton, 2016; Clayton, 2019). Each gene cluster is transcribed into a single polycistronic pre-mRNA (Figure 1.13), which is then processed into mature monocistronic mRNAs. RNA maturation occurs via the simultaneous trans-splicing and polyadenylation of the 5' and 3' ends of the transcript. Here, a prefabricated 3' nucleotide splice leader RNA cap is added to the 5' end of each pre-mRNA with a polyadenylated sequence added to the pre-mRNA of the upstream sequence. These two events occur within 200 bp of each other in the intergenic space (LeBowitz et al., 1993; Ullu et al., 1993). Thus, a single round of Pol II transcription can produce several functional mRNAs. Polycistronic transcription saves the parasite from repeated transcription initiation and capping of the pre-mRNA because it does not require stalling of RNA Pol II to mobilise other enzymes needed to catalyse enzymatic capping processes to generate mature mRNA. Additionally, constitutive gene expression is advantageous to the parasite gene expression because it allows constant flow of transcripts to respond to rapid environmental stresses and changes in parasite's host environment (Palenchar & Bellofatto, 2006).

Control of gene expression in most cases occurs post-transcriptionally and is partly modulated by regulatory elements found within sequences in the 5' UTR and 3' UTR. Sequence elements in the 3' UTR-determine mRNA abundance and stability by modulating RNA degradation. Regulation of cell cycle switch also lies within the 5'UTR sequence (Clayton, 1999; Mahmood et al., 1999).

1.6.3. Genetic tools for studying *Trypanosoma brucei* biology

The development and application of molecular tools has had a major impact on our understanding of *T. brucei* biology. Initial work in this area centred on transient transfection protocols primarily aimed at studying trypanosomal gene expression and provided insights into transcription initiation, post-transcriptional regulation and protein targeting mechanisms in this parasite (Eid & Sollner-Webb, 1991; Ten Asbroek et al., 1990). In such experiments, circular plasmids were introduced into the parasite by electroporation with constructs invariably integrating into the *T. brucei* genome. The DNA vectors used in these experiments contained a reporter gene (e.g., chloramphenicol acetyltransferase, luciferase etc) flanked at its 5' end by a *T. brucei* trans-splicing signal required for mRNA processing, or strong promoter (e.g., procyclin Pol I or T7

polymerase) while at its 3' end by a *T. brucei* 3' untranslated region containing a polyadenylation signal (Beverley and Clayton, 1993). Further vector modifications allowed selection of stable, clonal recombinant lines thereby expanding the toolbox of approaches that could be used to study *T. brucei* cell and molecular biology (Ten Asbroek et al., 1990). Plasmids used in such studies contained a drug selectable marker and the genetic elements required for its expression, alongside sequences to non-transcribed regions (e.g., intergenic regions from the ribosomal or tubulin arrays) (Ten Asbroek et al., 1990) that facilitate integration of the construct into these sites within the parasite genome. The regions required for this homologous recombination (HR) mediated event were shown to be relatively small (20-50 bp) (Barnes & McCulloch, 2007; Dean et al., 2015; Beneke et al., 2017) with these observations facilitating construction of parasite lines heterozygous or null for a given gene/activity, and ectopic gene expression. In the latter case, elevated levels of a given transcript could be achieved using trypanosomal (e.g., from procyclin) or heterologous (e.g., bacteriophage T7) promoter sequences, with the latter requiring expression of an appropriate RNA polymerase (bacteriophage T7 polymerase) (Wirtz et al., 1994; McAndrew et al., 1998; Wirtz et al., 1998). Further adaptations resulted in constructs that facilitate the use of inducible gene expression, often using the tetracycline repressor/operator system (Ngô et al., 1998; E. Wirtz & Clayton, 1995; Elizabeth Wirtz et al., 1994).

The integration of genetic material into the *T. brucei* genome by homologous recombination can be enhanced by the application of CRISPR, an approach that can speed up construction of null lines (one round of DNA transformation can be used to delete a single copy, diploid gene), precision base editing, and facilitate the 5' or 3' tagging of a given gene at its endogenous loci (Beneke et al., 2017; Rico et al., 2018) . In this system, *T. brucei* lines expressing the *Streptococcus pyogenes* Cas9 nuclease and the T7 RNA polymerase are co-transformed with multiple amplicons. One of these amplicons is then transcribed by the bacteriophage enzyme to generate a short guide RNA (sgRNA) that represents a single stranded nucleic acid molecule that contains sequences which: 1. are homologous to the region of the parasite genome to be edited and 2. functions to bind the Cas9 nuclease. The sgRNA then directs the nuclease to the homologous region in the parasite genome with the bacterial enzyme cleaving the DNA template adjacent to the RNA/DNA interaction site to generate a double strand DNA break (DSB). The other transformed amplicons, known as healing or targeting fragments, function in the repair of the lesion and in doing so, are inserted into the parasite genome at the break site. Generally, the inserted DNAs contains

sequences that code for a drug selectable marker and the genetic elements required for its expression with some also possessing a gene coding for a suitable reporter (e.g., luciferase, fluorescent protein *etc*).

The above innovations have greatly aided in deciphering how many biological pathways expressed in *T. brucei* operate, although they do have their limitations. For example, it is impossible to make null lines for genes whose products are essential for cell viability while it is impracticable to delete each coding sequence of a multiple copy gene from the parasite genome given the small number of selectable markers available. Against this backdrop, the identification that *T. brucei* possesses an active RNA interference pathway has partially helped overcome such obstacles (Ngo et al., 1998). Here, double stranded RNA is generated using DNA vectors that contain opposing promoters (Alibu et al 2005; LaCount et al., 2000; Wang et al., 2000) or single 'stem loop' (Alsford & Horn, 2008; Shi et al., 2000; Wang & Englund, 2001), often employing a bacteriophage T7 promoter/RNA polymerase whose expression is under the control of a tetracycline repressor/operator system. This approach has been extensively used on a "gene-by-gene" basis although with further adaptations and high throughput sequence analysis, has been employed at a genome wide level (Alsford et al., 2011; Glover & Horn, 2009; Horn, 2022). In the latter case, RNAi has proved a valuable tool in unravelling various aspects relating to *T. brucei* biology including in understanding the mechanisms of action and resistance to clinically used trypanocidal drugs (Alsford et al., 2013; S. Alsford et al., 2012; Baker et al., 2011), immune evasion (Alsford, 2014; Glover et al., 2016) and parasite differentiation (Mony et al., 2014).

1.7. DNA repair pathways in *Trypanosoma brucei*.

The genomes of all living organisms are constantly exposed to endogenous and exogenous DNA damage agents, including reactive oxygen or nitrogen species from cellular respiration, UV irradiation, DNA damaging drugs etc. These insults can promote alteration to the DNA sequence and result in genome instability. To resolve such damage a series of organism-specific, cell cycle dependent and independent strategies have evolved that function to recognize a lesion site facilitating the recruitment of factors to that site which then effect repair. Therefore, to combat the deleterious effects of a DNA damaging agent, an equilibrium between DNA damage and repair systems is required to prevent disease or cell death (Friedberg et al., 1995). Several 'classical' DNA repair systems such as the base excision repair (BER), homologous recombination (HR),

non-homologous end-joining (NHEJ), mismatch repair (MMR) and nucleotide excision repair (NER) pathways operate in model organisms (*E. coli*/*S. cerevisiae*/*H. sapiens*) with each mechanism designed to resolve specific types of DNA damage: NER tends to operate to resolve the bulky lesions induced by UV light, BER plays a major role in repair oxidatively damaged bases while both HR and NHEJ predominate in fixing DSBs. However, it is now recognized that there is a high degree of crosstalk between pathways. For example, the helicase CSB whose activity was thought to be restricted to transcription coupled NER (TC-NER) may also be recruited to lesion site where it can invoke transcription coupled HR or BER pathways (Stevnsner *et al.,* 2008; Menoni *et al.,* 2012; Batenburg *et al.,* 2015; Wei *et al.,* 2016; Menoni *et al.,* 2018). Informatic studies has shown that orthologues of many components from most of the 'classical' DNA repair pathways are expressed by *T. brucei* (Genois et al., 2014) with this, in several cases, supported by functional studies (Conway et al., 2002; Ulbert et al 2002; Bell et al 2004; Charret et al 2012; Machado et al 2014). Intriguingly, there are notable differences between this parasite's DNA repair mechanisms and those possessed by model organisms. For example, *T. brucei* reportedly lacks a functional NHEJ system (this trypanosome's genome does not contain genes that code for orthologues of DNA ligase IV, XRCC1 and XRCC4 although it does possess orthologues of Ku70 and Ku80) (Burton et al., 2007; Glover et al., 2008, 2011) whereas several components of the NER system (e.g., CSA, XPA) are apparently missing. The 'classical' DNA repair pathways and their role(s) in *T. brucei* will be discussed in this thesis in the introduction sections associated with appropriate Results chapter.

1.8. Nitroheterocycles and DNA damage repair pathways in Trypanosomes

Nitroheterocyclic compounds make up a significant component of the therapies available to treat trypanosomal infections. Nifurtimox, benznidazole and fexinidazole all function as prodrugs that must undergo activation to have cytotoxic effects. Functional genomic and biochemical approaches have demonstrated that trypanosomal expression of a type I nitroreductase (NTR) is key for their anti-parasitic selectivity (Wilkinson et al 2008; Baker et al 2011; Hall et al 2011; Hall & Wilkinson, 2012). Several of the NTR-generated, nitroheterocyclic derived metabolites including hydroxylamine and nitrenium ion derivatives, or glyoxal have potential to form adducts with various biological molecules and based on mutagenicity studies, DNA appears to represent a major target for these prodrugs (Ferreira & Ferreira, 1986; Toranzo et al., 1988; Hall et al 2011;

Hall & Wilkinson, 2012 Tweats et al 2012; Trochine et al., 2014; Campos *et al.,* 2017). What form this damage takes and how it arises is unknown although for the two nitroimidazole-based prodrugs this could be due to their ability to promote nucleotide oxidation causing mismatch base pairing leading to the formation of point mutation and/or DSBs (Gomes Passos Silva et al., 2018; Rajão et al., 2014). Support for this hypothesis comes from functional genomic studies. Here, *Trypanosoma cruzi* with elevated or reduced levels of certain MMR or HR enzymes generates trypanosomal lines that exhibit altered susceptibilities to benznidazole (Gomes Passos Silva et al., 2018; Rajão et al., 2014). Further, and in an NTR1-dependent process, benznidazole promotes DNA damage in the nuclear genome of *T. brucei* that requires components of the HR system to resolve, providing a direct link between prodrug activation and a downstream trypanocidal mechanism (Dattani et al., 2021). Intriguingly, no such link could be shown for nifurtimox indicating it affects DNA differently from benznidazole, if at all.

1.9. Research Aims

Trypanosomes are constantly exposed to DNA damaging agents and to maintain genome integrity, they possess functional DNA repair pathways that counter the effects of such agents.

Trypanocidal nitroheterocyclic compounds nifurtimox, benznidazole and fexinidazole all function as prodrugs that require nitroreduction by trypanosome specific nitroreductase, TbNTR1. TbNTR1 is important in drug selectivity as this enzyme is absent in humans (Wilkinson et al 2008; Baker et al 2011; Hall et al 2011; Hall & Wilkinson, 2012). Mutagenicity studies have also shown that one of the downstream effects of the reduced metabolites from benznidazole and nifurtimox involve interaction with DNA (Ferreira & Ferreira, 1986; Toranzo et al., 1988; Hall et al 2011; Hall & Wilkinson, 2012 Tweats et al 2012; Trochine et al., 2014; Campos *et al.,* 2017), forming adducts that in the case of benznidazole recruit factors from the HR to resolve the damage (Dattani et al., 2021). Mutagenicity studies have also shown that fexinidazole interacts with DNA. Whether this interaction results in recruiting various factors from *T. brucei* as seen above with benznidazole is yet to be demonstrated.

In this project I investigated the mode of activation of fexinidazole, and further elucidated the type of DNA repair pathways recruited to resolve such damage. Next, I assessed the effect of exogenous treatments *in T. brucei* to decipher the resultant DNA repair recruitment kinetics and how the different components involved in the various pathways interact in response to exogenous DNA damage. Understanding these

interactions will help in designing inhibitors for future exploitation in drug development. The specific aims of this project are to:

1. Study the mode of activation of fexinidazole
2. Understand the DNA repair pathways recruited to resolve fexinidazole induced DNA damage
3. Use CRISPR/Cas9 genome editing to engineer bloodstream and procyclic form parasite lines that express recombinant genes encoding DNA repair proteins endogenously tagged with sequences that encode the fluorescent marker mNeonGreen at the 5' or 3' ends of the endogenous gene.
4. Study the effect of exogenous agents including DNA damaging drugs (mechlorethamine, phleomycin), UV irradiation or Trypanocidal agents on recombinant parasites.
5. Evaluate the kinetics of DNA repair and how the various DNA repair proteins are triggered by these exogenous agents.
6. Understand how various components from these pathways interact

CHAPTER 2: Materials and Methods

2.1. Culturing of trypanosome cell lines and bacterial strains

2.1.1. Cell lines

The bacterial strains and trypanosomal cell lines used for this study are listed in Table 2.1.

Cell line	Source
Bloodstream form *T. brucei*	
Lister 427, antigenic type MITat 1.2, clone 221a (or '221').	Doyle et al 1980
Single Marker Bloodstream form (or 'SMB')	Wirtz et al 1999
221 *Tbsnm1Δ*	Sullivan et al 2015
221 *TbcsbΔ*	Dattani and Wilkinson 2019
221 *Tbmre11Δ*	Dattani and Wilkinson 2019
221 *TbcsbΔ Tbmre11Δ*	Dattani and Wilkinson 2019
2TAG1	Alsford et al. 2005
2TAG1 *Tbntr*$^{++}$	Wilkinson et al 2008
2TAG1 *Tbntr*$^{++}$ *Tbmre11Δ*	Dattani et al 2021
2TAG1 *Tbntr*$^{++}$ *TbcsbΔ*	Dattani et al 2021
SmOx B427 CAS9/pTB011	Beneke et al 2017
SmOx B427 CAS9/pTB011 mNG-TbMRE11	This study
SmOx B427 CAS9/pTB011 mNG-TbRAD51	This study
SmOx B427 CAS9/pTB011 mNG-TbSNM1	This study
SmOx B427 CAS9/pTB011 mNG-TbFAN1	This study
SmOx B427 CAS9/pTB011 mNG-TbAPE1	This study
SmOx B427 CAS9/pTB011 mNG-TbNTH1	This study
SmOx B427 CAS9/pTB011 mNG-TbCSB	This study
SmOx B427 CAS9/pTB011 mNG-TbERCC1	This study
SmOx B427 CAS9/pTB011 mNG-TbXPG	This study
SmOx B427 CAS9/pTB011 TbMRE11-mNG	This study
SmOx B427 CAS9/pTB011 TbRAD51-mNG	This study
SmOx B427 CAS9/pTB011 TbFAN1-mNG	This study
SmOx B427 CAS9/pTB011 TbAPE1-mNG	This study
SmOx B427 CAS9/pTB011 TbMYH-mNG	This study
SmOx B427 CAS9/pTB011 TbNTH1-mNG	This study
SmOx B427 CAS9/pTB011 TbCSB-mNG	This study
SmOx B427 CAS9/pTB011 TbDDB1-mNG	This study
SmOx B427 CAS9/pTB011 TbRAD23-mNG	This study
SmOx B427 CAS9/pTB011 TbXPC-mNG	This study
SmOx B427 CAS9/pTB011 TbERCC1-mNG	This study
SmOx B427 CAS9/pTB011 TbXPF-mNG	This study
SmOx B427 CAS9/pTB011 TbXPG-mNG	This study
221 *Tbsnm1Δ* pTubEX-nMG-TbSNM1	This study
SMB RNAi-Tb*ddb1*	This study
SMB RNAi-Tb*rad23*	This study
SMB RNAi-Tb*xpc*	This study
SMB RNAi-Tb*ercc1*	This study
SMB RNAi-Tb*rad23*	This study
SMB RNAi-Tb*xpg*	This study

Procyclic form *T. brucei*	
SmOx P927 CAS9/pTB011	Beneke et al 2017
SmOx P927 CAS9 Tb*ape1*$^{+/-}$*neo*	This study
SmOx P927 CAS9 Tb*ape1*$^{+/-}$*hyg*	This study
SmOx P927 CAS9 Tb*myh*$^{+/-}$*neo*	This study
SmOx P927 CAS9 Tb*myhΔ*	This study
SmOx P927 CAS9 Tb*nth1*$^{+/-}$*neo*	This study
SmOx P927 CAS9 Tb*nth1Δ*	This study
SmOx P927 CAS9/pTB011 mNG-TbMRE11	This study
SmOx P927 CAS9/pTB011 mNG-TbRAD51	This study
SmOx P927 CAS9/pTB011 mNG-TbSNM1	This study
SmOx P927 CAS9/pTB011 mNG-TbFAN1	This study
SmOx P927 CAS9/pTB011 mNG-TbAPE1	This study
SmOx P927 CAS9/pTB011 mNG-TbMYH	This study
SmOx P927 CAS9/pTB011 mNG-TbCSB	This study
SmOx P927 CAS9/pTB011 mNG-TbERCC1	This study
SmOx P927 CAS9/pTB011 TbMRE11-mNG	This study
SmOx P927 CAS9/pTB011 TbRAD51-mNG	This study
SmOx P927 CAS9/pTB011 TbFAN1-mNG	This study
SmOx P927 CAS9/pTB011 TbAPE1-mNG	This study
SmOx P927 CAS9/pTB011 TbMYH-mNG	This study
SmOx P927 CAS9/pTB011 TbNTH1-mNG	This study
SmOx P927 CAS9/pTB011 TbCSB-mNG	This study
SmOx P927 CAS9/pTB011 TbDDB1-mNG	This study
SmOx P927 CAS9/pTB011 TbRAD23-mNG	This study
SmOx P927 CAS9/pTB011 TbXPC-mNG	This study
SmOx P927 CAS9/pTB011 TbERCC1-mNG	This study
SmOx P927 CAS9/pTB011 TbXPF-mNG	This study
SmOx P927 CAS9/pTB011 TbXPG-mNG	This study
Escherichia coli XL-1 Blue	Laboratory stock

Table 2.1: Organisms used in this study

2.1.2. Culturing and long-term storage of *Trypanosoma brucei*

Trypanosoma brucei bloodstream form (BSF) trypomastigotes were grown in HMI-9 (Life Technologies) medium supplemented with 3 g l^{-1} sodium bicarbonate (Sigma Aldrich), 0.014 % (v/v) β-mercaptoethanol (Sigma Aldrich) and 10 % (v/v) heat-inactivated foetal bovine serum (Gibco) at 37 °C under a humidified 5 % (v/v) CO_2 atmosphere (Hirumi and Hirumi 1994). Recombinant parasites were cultured in this medium containing 5 µg ml^{-1} hygromycin, 10 µg ml^{-1} blasticidin, 2 µg ml^{-1} puromycin, 1 µg ml^{-1} phleomycin and/or 2 µg ml^{-1} G418. In some experiments, 1 µg ml^{-1} tetracycline was added to cultures to induce RNA interference.

Procyclic form trypomastigotes (PCF *Trypanosoma brucei*) were cultured at 27 °C in SDM-79 (Life Technologies) medium supplemented with 0.3 % (v/v) hemin (Sigma Aldrich) and 10 % (v/v) heat-inactivated foetal bovine serum (Gibco). Recombinant parasites were cultured in this medium containing 10 μg ml^{-1} blasticidin, 1 μg ml^{-1} puromycin and/or 25 μg ml^{-1} hygromycin.

For long term storage, parasite cell lines were stored in liquid nitrogen. To generate stabilities 1 ml of trypanosomal culture containing 10 % (v/v) sterile glycerol (Sigma-Aldrich) was transferred to labelled 1.2 ml cryogenic vials (Nunc). The tubes were frozen slowly (2-3 days) to -80 °C in a Mr. Frosty™ freezing container (Nalgene) containing isopropanol and subsequently transferred into liquid nitrogen. To recover stocks, the whole stabilate was rapidly thawed, transferred to fresh parasite growth medium (10 ml) and cultured using appropriate medium as described above.

2.1.3 Culturing and long-term storage bacterial strains

E. coli strains were cultured at 37 °C in a shaking incubator and in NZCYM broth (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} MgSO4) (Melford Laboratories Limited) supplemented with 100 μg mL^{-1} ampicillin (Sigma-Aldrich) where appropriate. In some experiments, this growth medium was solidified with 1.5 % (w/v) Agar Bacteriological (Agar No. 1) (Oxoid).

For long term storage, bacterial strains were stored at -80 °C. To generate a frozen stock, 1 ml of *E. coli* culture containing 10 % (v/v) sterile glycerol (Sigma-Aldrich) was transferred to a labelled 1.2 ml cryogenic vials (Nunc) and deposited directly into a -80 °C freezer. To revive, a small scraping (~10-50 μl) of the frozen bacterial line was placed onto NZCYM broth and cultured in a shaking incubator at 37°C.

2.2. Trypanocidal and genotoxic agents

Several antimicrobial or DNA damaging agents were used in this project. Mechlorethamine, an ICL inducing compound, was obtained from Cambridge Biosciences while the trypanocidal drugs benznidazole, fexinidazole and nifurtimox were supplied by Professor Simon Croft, and fexinidazole was kindly provided by Professor John Kelly (both London School of Hygiene & Tropical Medicine). The selective compounds blasticidin, puromycin, G418, hygromycin, phleomycin and ampicillin were sourced from Melford Laboratories Ltd while tetracycline was

purchased from Sigma-Aldrich. UV crosslinker (Stratagene) was used for UV irradiation experiments using multi-well petri dishes.

2.3 Plasmids and oligonucleotide primers

The precursor plasmids used to construct the trypanosomal expression and RNAi vectors or used to generate amplicons for use in CRISPR were pTubEX (Cross et al. 2002) p2T7^{TAblue} (Alibu et al. 2005), and pPOTv7-hygro-hygro-mNeonGreen (Dr Sam Dean, University of Oxford) (Dean et al., 2015).

The primers and primer combinations used in this study are listed in Appendix 1 to 4. All primers were synthesised by Sigma Aldrich or Integrated DNA Technologies (IDT) and supplied as a lyophilized pellet that required suspension in 1 ml sterile H_2O before use.

The primers used to amplify DNA fragments used in cloning generally contained, from 5' to 3', a nucleotide clamp to facilitate restriction enzyme binding, a restriction enzyme binding site and the region homologous to targeted sequence. Factors taken into consideration when designing primers include its length, T_m, GC content and ability to form secondary structures (hairpins, self-dimers etc).

The primers used to generate fragments for making RNA interference (RNAi) vectors were designed using RNAit website (last accessed 06-08-2022). In this system, the sequence under study is pasted into the appropriate box, and the primer selection options (melting temperature is usually 60 ºC)/PCR product size is between 400-600 bp) and the Blast parameters (stringency between 89-99%, subunit length of 20 bp and database used is Trypanosoma brucei TREU927) were inputted prior to primer design submission for.

The primers used to generate single guide RNA (sgRNA) amplicons specific for each targeted gene were designed as follows: The TriTrypDB (last accessed 06-08-2022) GeneID of the gene of interest was entered onto LeishGedit (last accessed 06-08-2022) database. Under the 'Chose your desired gene editing strategy' heading the 'N-terminal tagging' and 'C-terminal tagging' options were selected with the 'pPOT plasmids' choice made under in the 'Chose your desired modular plasmid system' question. After clicking on the 'Design primers' tab a CSV file was generated containing four primers for each tagging event. These primers were subsequently used to generate amplicons corresponding to: 1. The single guide RNA (sgRNA), and 2. The targeting fragment.

2.4. Bioinformatics

DNA sequences coding for a gene of interest (see Table 2.2) and its flanking regions were identified following PubMed (last accessed 07-08-2022) literature searches and/or orthology, textual or InterPro Domain searches of the TriTrypDB database (last accessed 07-08-2022). Orthologues and homologues of these were identified from this site or the National Center for Biotechnology Information and Saccharomyces genome databases (both last accessed 07-08-2022).

Gene	TriTryp GeneID	Chr	ORF (bp)	protein (aa)	TrypTag amino	TrypTag carboxy
Tb*ape1*	Tb927.8.5510	8	1191	396	nuc	nuc
Tb*csb*	Tb927.7.4080	7	3381	1126	cyt(r)	cyt/end
Tb*ddb*	Tb927.6.5110	6	3813	1270	cyt(r)	nuc(p)/ cyto(r)
Tb*ercc1*	Tb927.7.2060	7	801	266	no data	no data
Tb*fan1*	Tb927.10.7540	10	2949	982	nuc(l)/ cyt(r)	nuc(pt)/ cyt(p)
Tb*mre11*	Tb927.2.4390	2	2259	752	no data	no data
Tb*myh*	Tb927.11.11440	11	1533	510	no data	no data
Tb*nth1*	Tb927.11.12090	11	780	259	no data	no data
Tb*rad51*	Tb927.11.8190	11	1122	373	no data	no data
Tb*snm1*	Tb927.4.1480	4	2163	720	nuc(l)	nuc(l)/cyt
Tb*xpc*	Tb927.9.11930	9	2307	768	nuc(l)cyt(r)	nuc/cyt(p)
Tb*xpf*	Tb927.5.3670	5	3729	1242	no data	no data
Tb*xpg*	Tb927.9.11760	9	2241	746	nuc(l)/cyt(r)	nuc(l)

Table 2.2. Genes studied during this project.
TriTryp GeneID links last accessed on 07-08-2022. Chr is chromosome; ORF is open reading frame; aa is amino acids; TrypTag (last accessed on 07-08-2022) predicted localisations in PCF *T. brucei* based amino or carboxyl terminal tagging where cyt is cytoplasm; cyt(p) is cytoplasm (points); cyt(r) is cytoplasm(reticulated); end is endocytic; nuc is nucleus; nuc(l) is nuclear lumen; nuc(p) is nucleoplasm; nuc(pt) is nucleus (points).

Pairwise sequence analysis was done to understand the relatedness of *T. brucei* factors with orthologues from other organisms including humans, *T. cruzi, Leishmania spp*, yeast as well as bacteria. Percentage identities were generated by pairwise alignment of nucleotide sequences using NCBI blast suite or via EMBOSS Needle. Multiples sequence alignments analysis were performed using CLUSTAL Omega (https://www.ebi.ac.uk), MEGA11(Takamura et al., 2021) and

visualised using BoxShade (https://www.expasy.org/) or GeneDoc(Ref) to visualise conserved motifs.

Amino acid sequences for the various DNA repair proteins, were submitted on PFAM domain database, NCBI CDD, for HMMR analysis of structural domain architecture and function for each protein. Additionally, TriTrypDB, UniProt and PDB were used for in-silico functional analysis of the various factors using amino acid sequences and data from similar proteins.

In-silico prediction of subcellular protein localization was also done using localisation prediction algorithms (PSORTII and WoLFPSORT). A full-length 3D prediction model for each protein was constructed using Phyre2 (Kelley et al., 2015), 3-D prediction algorithm. and visualised using PyMOL. Predictions of disordered regions in the 3D-model were obtained by submitting amino acid sequences to generate a graphical representation with the residues that make up the disordered or structural domains of each protein (See different chapters).

2.5 Nucleic Acid Purification.

GeneJet™ Plasmid Miniprep kit (ThermoFisher Scientific) was used to extract plasmid DNA from overnight recombinant *E. coli* cultures following the manufacturer's online protocol (last accessed 07-08-2022). Once purified, plasmid DNA was stored at -20°C.

The 'Cultured Mammalian Cells Genomic DNA Purification Protocol' of the GeneJet™ Genomic DNA purification kit (ThermoFisher Scientific) was used to extract genomic DNA (gDNA) from approximately 5×10^7 Phosphate Buffered Saline (PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.76 mM KH_2PO_4, pH 7.6) (Sigma Aldrich)-washed *T. brucei*, following the manufacturer's online protocol (last accessed 07-08-2022). Once purified, plasmid DNA was stored short term at 4 °C, then at -20 °C for longer term.

The 'Mammalian Cultured Cells Total RNA Purification Protocol' of the GeneJet™ RNA Purification Kit (ThermoFisher Scientific) was used to extract total RNA from approximately 5×10^7 PBS-washed *T. brucei*, following the manufacturer's online protocol (last accessed 07-08-2022). Once purified, total RNA samples were stored at -20 °C.

DNA fragments (in an agarose gel slice or in liquid) were purified following the ThermoFisher Scientific online protocol (last accessed 07-08-2022). Once purified, DNA fragments were stored at -20 °C.

2.6 DNA synthesis

2.6.1 cDNA synthesis

Complementary DNA (cDNA) synthesis was carried out using qScript™ cDNA Synthesis Kit (Quantabio) following the manufacturer's online protocol (last accessed 07-08-2022). Purified *T. brucei* total RNA was synthesised by combining 5 µl RNA sample, 10 µl distilled water, RNAase free H_20, 4 µl 5X qScript Reaction Mix (containing an optimised buffer, magnesium, oligo(dT) and random primers, dNTPs) and 1 µl qScript reverse transcriptase in a 0.2 ml thin-walled micro-centrifuge tube to a total reaction volume of 20 µl. The tube was placed in a TC-412 thermal cycler (Techne) and subjected to one cycle at 5 minutes at 22°C, 60 minutes at 42°C, 5 minutes at 85°C and then held at 4°C. cDNA samples were stored at -20°C.

2.6.2 Standard PCR protocols

Oligonucleotide primers (Sigma Aldrich; IDT), dNTPs (ThermoFisher Scientific) and various DNA sources (parasite gDNA, cDNA, plasmid DNA) were used in PCRs to generate a variety of amplicons. In cases where an accurate copy of the template was required (generating DNA fragments used in DNA cloning, *T. brucei* cell line validation *etc*) Phusion® High Fidelity DNA polymerase (New England Biolabs) and its respective 5X GC buffer was used. In such reactions, template DNA (0.1 to 1 ng μl^{-1} *T. brucei* cDNA or gDNA), forward & reverse primers (0.5 µM each), dNTPs (200 µM each), DMSO (3 % (v/v)), 1X Phusion® buffer (containing 1.5 mM $MgCl_2$) and Phusion® High Fidelity DNA polymerase (0.02 U μl^{-1}) were combined in a 0.5 ml thin-walled microcentrifuge tube.

For amplifications where less accurate copies of the target sequence were needed (e.g., *T. brucei* cell line validation) Taq DNA Polymerase (New England Biolabs) and its 10X ThermoPol® buffer was employed. Here, a typical reaction was set up by combining in a 0.5 ml thin-walled microcentrifuge tube the template DNA (0.1 to 1 ng μl^{-1} *T. brucei* gDNA), forward & reverse primers (0.2 µM each), dNTPs (200 µM each), 1X ThermoPol® buffer (20 mM Tris-HCl, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 0.1% (v/v) Triton X-100, 1.5 mM $MgSO_4$) and *Taq* DNA Polymerase (0.025 U μl^{-1}).

Using a TC-412 thermal cycler (Techne), a PCR consisting of an initial denaturation step at 98 °C for 2 minutes, followed by 30 cycles of 98 °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/primer combination, the number of cycles and the temperature and time of the denaturation/annealing/extension conditions adjusted to optimise each amplification.

2.6.3 Generation of sgRNA amplicons for CRISPR

Amplicons corresponding to the sgRNA were made by combining sgRNA scaffold (G00) (2 µM) and the gene-specific primer (2 µM) with dNTPs (200 µM each), 1X Expand™ High Fidelity reaction buffer with 1.5 mM $MgCl_2$ (Roche) and Expand™ High Fidelity polymerase (0.05 Uμl^{-1}) (Roche) in a 0.5 ml thin-walled tube to give a final reaction volume of 20 µl. Using TC-412 thermal cycler (Techne), the reaction was subject to a single cycle at 98 °C for 120 seconds (initial denaturation), followed by 35 cycles at 98 °C for 10 seconds (denaturation), 60 ºC for 20 seconds (annealing), and 72 ºC for 15 seconds (elongation).

2.6.4 Generation of targeting fragment amplicons for CRISPR

Amplicons corresponding to the targeting fragment (TF) were made by combining pPOTv7-hygro-hygro-mNeonGreen (50-100 ng) with forward and reverse primers (0.1 mM each), dNTPs (200 µM each), DMSO (3 % (v/v)), $MgCl_2$(1.5 mM), 1X Expand™ High Fidelity reaction buffer with 1.5 mM $MgCl_2$ and Expand™ High Fidelity polymerase (0.025 U μl^{-1}) in a 0.5 ml thin-walled tube to give a final reaction volume of 40 µl. Using TC-412 thermal cycler (Techne), the reaction mixture was subjected to a single cycle at 98 °C for 120 seconds (initial denaturation), followed by 35 cycles at 98 °C for 10 seconds (denaturation), 60 ºC for 30 seconds (annealing), and 72 ºC for 90 seconds (elongation). At the end of the 35th cycle, a further extension step of 7 minutes at 72°C was added and the samples were then incubated at 4°C.

2.6.5 Reverse transcription quantitative PCR

Reverse transcription quantitative PCR (RT-qPCR) was performed on *T. brucei* cDNA using PowerUp™ SYBR™ Green PCR Master Mix, ThermoFisher Scientific, Cat # 4367659. A typical reaction contained 2X PowerUp™ SYBR™ Green Master Mix (SYBR™ Green Dye, Dual-

Lock™ DNA Polymerase, Uracil-DNA Glycosylase (UDG), ROX™ passive reference dye, dNTP mix with dUTP/dTTP and other optimized buffer components), cDNA (100 to 200 ng/ µl) and a pre-optimised concentration of forward/reverse primer (usually 1 pmol μl^{-1}), and distilled nuclease free water to make 10 µl reactions. Samples were loaded into the wells of a clear 96 well plate and an adhesive cover (Bio-Rad) used to seal the plate. PCR amplification was carried in a CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad using the following run conditions: single initial denaturation step at 95 °C for 2 minutes, which was followed by 30 cycles of 95 °C for 30 seconds (denaturation), 68 °C for 30 seconds (annealing) and 68 °C for 45 seconds (extension). QC to assess target specificity was done by melt curve analysis at the end of the amplification reaction. QPCR data generated was analysed using the CFX96 Touch Software (Bio-Rad) and the cycle threshold (CT values) were exported into Microsoft excel. Gene expression was evaluated using comparative CT analysis method (Schmittgen TD, 2008) and normalization was done using standardized housekeeping gene (Tb*tert*). Brenndörfer and Boshart, 2010), with normalisation control run simultaneously with samples.

2.7 DNA manipulation

2.7.1 DNA electrophoresis

Agarose gel electrophoresis was used to separate DNA fragments with the agarose concentration, voltage and run times altered to optimise separation over the desired range (100 bp to 10 kbp). An agarose gel was made by dissolving in a microwave an appropriate amount of Molecular Grade Agarose (Bioline) in 1X TAE buffer (40 mM Tris base, 40 mM acetic acid, 1 mM EDTA) (all Sigma-Aldrich) containing ethidium bromide (0.1 mg l^{-1}) (Sigma-Aldrich): Generally, 0.8 to 2 g agarose in 100 ml TAE buffer. The molten agarose was cooled to around 60°C then poured into a sealed gel tray containing a comb and allowed to solidify. The solidified gel was transferred to a gel electrophoresis tank containing 1X TAE supplemented with 0.1 mg l^{-1} ethidium bromide and the DNA samples loaded into the wells of the gel, with 1kb Plus GeneRuler (ThermoFisher Scientific) added to an additional well to act as a size marker. The DNA fragments were separated through the gel by applying a constant voltage (70-100 V) across the gel for run times of 1-2 hours: Some gels were run overnight (~18 hours) at 15V. Migration of the sample through the agarose matrix was followed by monitoring the dye fronts present in the Purple Gel

Loading Dye. Once the DNA had migrated the desired distance, the gel was visualized on a UV transilluminator and documented (Syngene).

2.7.2 Restriction digestion

Restriction enzymes and their 10X restriction buffers were sourced from New England Biolabs and/or ThermoFisher Scientific. To carry out a typical restriction digestion reaction DNA sample (1 to 5 µl or ~200 µg), 10X restriction buffer (2 µl) and restriction enzyme (0.5 µl or ~5 units) were combined in a 1.5 ml microcentrifuge tube and the total volume adjusted to 20 µl with an appropriate amount of sterile distilled water. The tube contents were consolidated by centrifugation and the reaction incubated at the appropriate incubation temperature, usually 37°C, for 30 to 60min. The reaction was halted by addition of 4 µl of 6X Purple Gel Loading Dye (2.5 % (w/v) Ficoll®-400, 10 mM EDTA, 3.3 mM Tris-HCl pH 8.0, 0.08 % (w/v) SD, 0.02 % (w/v) Dye 1, 0.0008 % (w/v) Dye 2; New England Biolabs) to the mixture.

For double restriction digestion reactions, the above was altered to reflect the addition of the second enzyme and use of a 10X restriction buffer compatible for both enzymes: the latter was done using ThermoFisher Scientific's DoubleDigest Calculator-Thermo Scientific (last accessed 08-08-2022) or New England Biolab's Double Digest Finder (last accessed 08-08-2022).

2.7.3 DNA ligation

DNA ligation reactions were performed using T4 DNA ligase and its 10X T4 DNA ligase buffer sourced from ThermoFisher Scientific. A typical ligation reaction was set up by combining purified digested vector (1-5 µl) and insert (5-10 µl) DNAs (depending on their concentration, determined using a NanoDrop) in a 1.5 ml microcentrifuge tube into which 10X T4 DNA ligase buffer (2 µl) (400 mM Tris-HCl, 100 mM MgCl2, 100 mM DTT, 5 mM ATP, pH 7.8 at 25°C) and T4 DNA ligase (0.5 µl or 2.5 Weiss units) were added and the total volume adjusted to 20 µl with appropriate amount of sterile distilled water. The tube contents were consolidated by centrifugation and incubated at room temperature for 3 to 4 hours or overnight. Reactions were topped by incubating at 65°C for 10 minutes. A control ligation was also set up replacing the insert DNA volume with sterile distilled H_2O.

2.7.4 *E. coli* DNA transformation

An aliquot (0.5 ml) from an *E. coli* stock culture was diluted 1:20 into NZCYM broth (10 ml) and incubated at 37 ºC for 2 hours with aeration. The cells were harvested by centrifugation (3,555 g for 5 minutes), the pellet suspended in an equal volume (generally 10 ml) ice cold 0.1 M $CaCl_2$ and held on ice for 10 minutes. The cells were re-harvested by centrifugation (3,555 g for 5 minutes) and the pellet suspended in a $1/10^{th}$ volume (generally 1 ml) of ice cold 0.1 M $CaCl_2$. An aliquot (0.1 ml) of the chemically competent cells was then incubated with DNA for 30 mins on ice prior to heat shocking at 42 °C for 2 minutes before returning to ice for a further 2 minutes. Transformed bacteria were plated onto ampicillin containing NZCYM agar plates and incubated overnight at 37 °C.

2.7.5 *T. brucei* DNA transformation

Bloodstream or procyclic form *Trypanosoma brucei* in the logarithmic phase of growth were harvested by centrifugation (1,600 g for 10 minutes) and the pellets washed in 10 ml Tb-BSF buffer. Cells were re-harvested by centrifugation (1,600 g for 10 minutes) and the pellet suspended at ~1×10^9 parasite per ml in Tb-BSF. An aliquot (200 µl) of cells (~5×10^7) was transferred into a 2 mm gap electroporation cuvette (Lonza) containing purified DNA and electroporated with a single 1.4 kV, 25 µF pulse using a Bio-Rad Gene Pulser II (without pulse controller or capacitance extender) or nucleofection using Amaxa® Nucleofector®(Lonza) (Burkard, Fragoso, and Roditi 2007).

Electroporated BSF parasites were immediately transferred into 55 ml pre-warmed HMI-11 growth medium incubated at 37 ºC for 4 to 6. A 5 ml aliquot of this culture was transferred into a 45 ml pre-warmed growth medium with this representing a 1:10 dilution. Selective agent was added to each culture and the cells transferred in 2 ml aliquots into the wells of a 24-well plate. Cultures were incubated at 37 ºC under a humid 5% CO_2 atmosphere with drug-resistant transformants apparent 5–7 days post transfection.

Electroporated PCF parasites were immediately transferred into 10 ml pre-warmed SDM-79 based growth medium incubated at 28 ºC overnight. The volume of this culture was then adjusted to 50 ml by addition of fresh SDM-79 medium into which a selective agent was added. The cells were then transferred in 2 ml aliquots into the wells of a 24-well plate and cultures incubated at 27

ºC under a humid 5% CO_2 atmosphere. Drug-resistant parasite populations apparent 2-3 weeks post transfection.

2.7.6 DNA sequencing

The sample's DNA concentration was determined using a NanoDrop™ UV spectrophotometer (ThermoFisher Scientific). The sample (~20-100 ng /µl) plus the appropriate primer (3.2 pmol/ µl) were sent for commercial sequencing by Source BioScience. The output electropherogram file was analysed using SnapGene viewer (last accessed 08-06-2020) and edited accordingly.

2.8 Construction and validation of recombinant parasite lines

2.8.1 CRISPR/Cas9 genome edited line

The PCR-based CRISPR/Cas9 technique described by Beneke *et al*. (2017), was used to endogenously tag at their 5' or 3' ends various genes encoding for DNA repair factors (TbAPE1, TbCSB, TbDDB1, TbERCC1, TbFAN1, TbMRE11, TbMYH, TbNTH1, TbRAD23, TbRAD51, TbSNM1, TbXPC, TbXPF and/or TbXPG) with sequencing coding for the fluorescent protein mNeonGreen. This procedure involved generating sgRNA and TF amplicons that recognise the targeted gene and the electroporation of these fragments into *T. brucei* SmOx P927 CAS9/pTB011or SmOx B427 CAS9/pTB011 cells. Engineered parasites were selected using hygromycin (the hygromycin phosphotransferase gene that encodes for resistance to this drug is present on the HF) with viable putative recombinant BSF clonal and PCF polyclonal lines visible 5 to 7 or 10 to 14 days post-transformation, respectively.

To demonstrate that the mNeonGreen-containing TF has integrated into the correct genetic loci, PCRs were performed on genomic DNA or cDNA templates generated from putative recombinant trypanosomes. To detect fusion events where the DNA repair gene had been tagged at its 5' end, amplifications were carried out using a mNeonGreen forward and a gene-specific reverse primer while in cases where the 3' end of the target gene had been tagged, a gene-specific forward and a mNeonGreen reverse primer were employed. In all situations, resulting amplicons were designed to contain sequences encompassing the fusion event. Fragments amplified from cDNA using Phusion® High Fidelity DNA polymerase were sequenced (Source BioScience) with the resultant

outputs used to confirm the in-frame fusion. For all DNA's analysed, additional reactions aimed at detecting Tb*tert* were performed to check the integrity of templates and as loading control. The primer sequences and combinations used in cell line validation are listed in Appendix 2.

2.8.2 RNAi lines

To construct *T. brucei* lines that can undergo RNAi targeting a transcript, gene specific fragments were amplified from *T. brucei* 221 gDNA with primers designed using the RNAit software. The resultant (~400bp) PCR products were BamHI/XhoI-digested and cloned into the corresponding sites of pRNAi-TbSpSyn (Taylor *et. al.* 2008), a BSF *T. brucei* RNAi plasmid derived from $p2T7^{TAblue}$, to replace the TbSpSyn sequence. Following validation by: 1. BamHI/XhoI restriction digestion and 2. DNA sequencing, one RNAi plasmid for each of the targeted transcripts was linearized with NotI and the gel purified fragments electroporated into BSF *T. brucei* SMB cells. Section for recombinant cells made using hygromycin with viable putative BSF clonal *T. brucei* visible 5 to 7 days post-transformation. To understand the effect of RNAi induction on the NER factors, recombinant lines were cultured in the presence of 1 µg/ml tetracycline alongside a control flask without tetracycline for each line. Daily counting of cells in the test and control flasks followed. Cell densities were plotted against each other to produce a +/- tetracycline growth curve. Next, to study the effect of RNAi induction on mRNA expression quantitative RT-PCR (qRT-PCR) was performed on cDNA from 48 hr cultures of induced and uninduced lines. Furthermore, the effect of RNAi induction on DNA damage was explored using recombinant lines subjected to mechlorethamine treatment following RNAi induction and without induction.

2.8.3 Complementation lines

To demonstrate that the *T. brucei* mNeonGreen recombinant proteins were functional, plasmids that facilitate the ectopic expression of *mNGTbSNM1, mNGTbMRE11, mNGTbCSB* and *mNGTbRAD51* were. To make these constructs, the complete open reading frame of the fusion gene was amplified from *T. brucei* genomic DNA extracted from the respective CRIPSR/CAS9-generated line. The resultant fragments were SbfI/AscI-digested, then cloned into the corresponding sites of pTub-EX-LmSpySyn Taylor et al 2008, a *T. brucei* constitutive expression

vector derived from pTub-EX, to replace the LmSpSyn sequence. Constructs were NotI/XhoI-digested with the purified DNA fragments then introduced into an appropriate *T. brucei* null mutant line by electroporation with recombinant parasite clones selected using phleomycin. Genetically modified *T. brucei* clones were observed after 5 to 7 days post-transformation. Cell lines were validated by PCR and as described for CRISPR Cas9 genome edited line, and fluorescence microscopy to confirm mNeonGreen expression.

2.8.4 *T. brucei* TbCSB null lines that express elevated levels of TbNTR1.

Construction of the DNA vectors used to delete the Tb*CSB* open reading frame from the *T. brucei* genome and containing the blasticidin-S-deaminase (*bla*) or puromycin N-acetyltransferase (*pac*) resistance cassette has been described elsewhere (Dattani & Wilkinson 2019). The constructs were linearized (SacI/KpnI or SacII/ KpnI for the *pac* or *bla* constructs, respectively) then sequentially transformed into BSF *T. brucei* (221 wild type or 2T1 engineered to express Tb*NTR1*) and clonal recombinant parasites selected using the appropriate drug selection. To demonstrate that integration of a gene deletion vector had occurred at the correct genetic loci and confirm that the *T. brucei* null mutant lines were no longer expressing Tb*CSB*, PCRs were performed on parasite genomic DNA or cDNA templates. In these reactions' primer combinations were used to generate amplicons specific for the intact Tb*CSB*, or the *pac*- and/or *bla*-disrupted allele. To check integrity of the template, and to act as a loading control, additional amplifications reactions aimed at detecting expression of the housekeeping gene Tb*tert* were performed.

2.9 Phenotyping recombinant parasite lines

2.9.1 Cumulative growth assays

T. brucei BSF parasites in the logarithmic phase of growth were seeded at 1×10^4 parasites ml^{-1} in 10 ml growth medium and incubated at 37 ºC 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 x 10^6 cells ml^{-1}, a new culture seeded at 1×10^4 parasites ml^{-1} was set up. This analysis was carried out over a 7–20-day period. Growth curves were generated using GraphPad Prism (GraphPad Software). All growth assays were performed in triplicate and each count at each time point expressed as a mean ± standard deviation.

2.9.2 *T. brucei* growth inhibition assays

BSF *T. brucei* in the logarithmic phase of growth were seeded at 1×10^4 cells ml^{-1} in 200 µl growth medium containing various concentrations of the compound under study and dispensed into the wells of a 96-well plate. Control assays including cell grown in the absence of drug (100 % growth) and a medium background control were set up in parallel. After incubation for 3 days at 37 °C under a humid 5 % (v/v) CO_2 atmosphere, resazurin (Sigma Aldrich) dissolved in PBS was added to each culture to a final concentration of 12.5 mg l^{-1} (or 2.5 µg resazurin per well). The plates were further incubated for 8 hours at 37 °C in a humid 5 % (v/v) CO_2 atmosphere with the fluorescence of each culture subsequently determined using a Gemini Fluorescent Plate Reader (Molecular Devices) set to a λ_{EX} =530 nm and λ_{EM} =585 nm with a filter cut off at 550 nm The change in fluorescence resulting from the reduction of resazurin is proportional to the number of live cells.

To calculate the fluorescence value at each compound concentration (fl_v), the following calculation was used:

$$fl_v = fl_{raw} - fl_{background} \qquad \text{equation 1}$$

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

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

$$\% \text{ growth at given [compound]} = fl_v/fl_{drug=0} \times 100 \qquad \text{equation 2}$$

Data expressed as % growth was used to construct dose response curves from which a compound's effective 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 t test calculator (GraphPad Software Inc).

2.9.3 Fluorescence microscopy and image analysis

T. brucei in the logarithmic phase of growth were harvested by centrifugation (800 g for 5 minutes), washed twice in PBS prior to fixing the cells by slowly suspending the pellet in 2 % (v/v) formaldehyde in PBS at a density of approximately 1 x 10^7 cells ml^{-1}. Samples were incubated at room temperature for 5 minutes, the cells harvested by centrifugation (800 g for 5 minutes) and the pellets suspended in PBS at an approximately density of 1 x 10^7 cells ml^{-1}. An aliquot (10 to

20 µl) of fixed parasites was dotted onto the well of a 10-well printed microscope slide (*Hendley Essex Ltd) and the sample* left to air dry. Slides were washed twice with PBS before being allowed to air dry again. Vectashield mounting medium containing 4′,6-diamidino-2-phenylindole DAPI (Vectashield Laboratories) was added to the surface of the slide and a coverslip applied. The edge of the slide was sealed with clear nail varnish and the varnish left to harden overnight in the dark and at room temperature. Slides were stored at 4 ºC until ready for analysis. Cells were visualized using a DMRA2 epifluorescent microscope (Leica) fitted with a C4742-95 digital camera (Hamamatsu Photonics) or a Deltavision Elite Deconvolution Microscope (GE Healthcare) fitted with an Olympus 60X NA 1.4 objective lens using Softworx software (GE Healthcare) to capture the images.

Images were processed using ImageJ (last accessed 08-08-2022). Cell images were converted into hyperstacks in the order of XYZCT, split into 2 channels and 31 frames. The 2 channels appeared red and green as staining DNA (mitochondrial and nuclear) with DAPI displayed as red while mNeonGreen tagged fluorescent proteins expressed as green. Z projects were selected, and stacks were converted to images to merge channels creating an image distinctly illustrating the 2-contrasting stained cellular features. With the brightfield image outlining parasites, separate channel images and merged images, stacks and montages were created and analysed for the location of expressed fluorescent proteins in individual cells.

2.9.4 Monitoring temporal changes in protein expression using fluorescence

T. brucei in the logarithmic phase of growth were seeded at a density of 1×10^6 cells ml^{-1} in medium containing a compound of interest. At time intervals (0, 1, 2, 3, 4 and 6 hours or 0, 0.5 1.5, 4 and 6 hours for UV irradiated cells) aliquots were removed from culture and the cells fixed, processed then imaged as previously described. Images were processed using ImageJ (last accessed 08-08-2022) and the corrected total cell *fluorescence* (CTCF) for a specified number of cells from each sample determined. Each CTFC data point was plotted as a scatter plot using GraphPad Prism (GraphPad Software Inc.) with the mean CTFC ± standard deviation calculated. The statistical significance of any differences in parasite susceptibilities was assessed using One-way analysis of variance (ANOVA). Kruskal–Wallis's nonparametric t-test ($p<0.0001$) (GraphPad Software Inc) was used to compare untreated samples against treated, Post-tests to

assess significance of each means by Dunn's comparative t-test was used to assess the significance of any differences between the means from each timepoint with 99% confidence.

2.10 Protein work

2.10.1 Preparation of trypanosomal cell extracts

Protein extracts derived from trypanosomal pellets were generated by lysis in a detergent–containing buffer or by freeze-thawing. To generate total extracts, parasites were harvested by centrifugation (1600 g for 5 minutes), the pellets washed twice with PBS before suspending the cells in 1x Laemmli buffer at a density of approximately 1×10^9 cellsml^{-1}. When soluble and insoluble fractions were required, the PBS washed cell pellets were suspended in PBS at a density of approximately 1×10^9 cellsml^{-1}. For mNeonGreen analysis, $5x10^7$ parasites were washed with PBS and suspended in 100ul 1X Laemmli buffer (50 ul and 50ul 1X PBS of 2X Laemmli). The cell suspension was then mixed by votexing and mechanically lysed using an ultrasonic water bath for 10 min or until lysate was no longer viscous, a consequence of DNA still bound to proteins. Alternatively, samples were placed in liquid nitrogen for 1 minute then defrosted at room temperature (approx. 5 minutes) with the freeze-thaw cycle repeated twice more. Extracts were clarified by centrifugation (20,238 g for 5 minutes) with the supernatant (now called the 'soluble fraction') transferred into a new 1.5 ml centrifuge tube to which an equal volume of 2x Laemmli buffer (4 % (w/v) sodium dodecyl sulphate, 20 % (v/v) glycerol, 0.004 % (w/v) bromophenol blue, 125 mM M TrisCl pH 6.8, 10 % (v/v) 2-mercaptoethanol) while the pellet was suspended in 100 µl 1x Laemmli buffer (referred to as the 'insoluble fraction').

2.10.2 SDS-PAGE electrophoresis

Prior to analysis, cell extracts in Laemmli buffer were heated to 96 °C for 5 minutes to denature the proteins. Samples alongside a Blue Prestained Protein Standard ladder (New England Biolabs) were loaded onto and separated through a 12 % (w/v) SDS-PAGE gel in a X SDS-PAGE running buffer (25 mM Tris-HCl, 200 mM glycine, 0.1% (w/v) SDS) using the Mini-PROTEAN® II cell system (Bio-Rad) run at a constant voltage of 120 V for 90 to 120 minutes at 4 °C. Gels were then: 1. stained for 1 h in 100 ml 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 ChemiDoc™ Touch Imaging System), or 2. transferred to 0.45 µm Protran nitrocellulose membrane (GE Healthcare) using the Trans-Blot Semi-Dry Transfer Cell (Bio-Rad) system running for 90 minutes at a constant voltage of 15 V and 1x Bjerrum Schafer-Nielsen transfer buffer (25 mM Tris, 192 mM glycine, 0.5 % (w/v) SDS, 20 % (v/v) methanol).

2.10.3 Immunoblot analysis

Nitrocellulose membranes containing cell extracts were washed at room temperature for a minimum of 1 h with blocking buffer (1X PBS, 0.1% (v/v) Tween 20, 10 % (w/v) Marvel powdered milk) then incubated overnight at 4 °C in the primary antibody diluted in blocking buffer. The following morning membranes were extensively washed (2 x 10 sec and 4 x 5 minutes washed) with a 1x PBS, 0.1% (v/v) Tween 20 buffer before being incubated for 1 h at room temperature with the secondary antibody diluted 1in blocking buffer. The membranes were extensively washed (2 x 10 sec and 4 x 5 minutes washed) again with a 1x PBS, 0.1% (v/v) Tween 20 buffer before finally being washed twice (5 minutes per wash at room temperature) in 1x PBS. Detection of the near infrared signal was monitored using an Odyssey® CLx infrared imaging system (LI-COR) or ChemiDoc™ ⸢(anti-mouse HRP) and images processed using the free software package ImageJ (last accessed 08/08/2022).

The primary antibodies used during this project were a mouse monoclonal antibody [32F6] that detects the *mNeonGreen* protein (Chromotek), a rabbit antiserum raised against *T. brucei* γH2AX (Dr R. McCulloch, University of Glasgow) and a rabbit antiserum raised against enolase from *T. brucei* (Prof Paul Michels, University of Edinburgh). These were used at a 1:1000, 1:1,000 dilution and 1:10,000 dilution, respectively. The secondary antibodies used were anti-mouse HRP, a goat anti-rabbit IRDye™ 800CW antibody (LI-COR), and a goat anti-mouse IRDye™ 680LT antibody (LI-COR), all diluted at 1:5,000.

CHAPTER 3: Unravelling the Anti-Trypanosomal Mechanism of Nitroheterocycles: From Prodrug Activation to DNA Damage

3.1. Double strand DNA breaks and their repair.

Caused by a variety of endogenous and exogenous factors (Table 3.1), double strand DNA breaks (DSBs) occur when both strands of DNA in a heteroduplex are severed within a few base pairs of each other resulting in the dissociation of the two strand ends (Jackson, 2002). If not correctly repaired, these can go on to promote mutations, loss of heterozygosity and/or genomic rearrangements events (e.g., deletions, translocations, fusions) that may have deleterious and potentially lethal effects (Kaye et al., 2004; Cannan and. Pederson, 2017). To prevent/minimise such events, eukaryotic cells have evolved multiple systems to resolve this damage including the error-free, homologous recombination (HR) pathway and the error-prone, non-homologous-end joining (NHEJ) pathway (Figure 3.1) (Chang et al., 2017). These mechanisms tend to predominate at different points of the cell cycle, with the former representing the main DSB repair system expressed in late S/G2 with the latter operating throughout although its activity is most pronounced during G1 (Dexheimer, 2013; Li & Heyer, 2008; Löbrich & Jeggo, 2017). In the absence of NHEJ, an error-prone microhomology end-joining (MMEJ) system can operate with this being most active during the S/G2 stages of the cell cycle (Truing et al 2013; Chang et al., 2017).

Damage	Reference
Endogenous	
Oxidised bases on both strands	Yang et al., 2004; Cannan et al., 2014
Replication fork accidents	Mirkin and Mirkin, 2007; García-Muse and Aguilera, 2016; da Silva et al., 2019
DSB generated spontaneously in the S-phase	Vilenchik and Knudson, 2003; Saleh-Gohari et al., 2005; Elango et al., 2017
Exogenous	
Ionising radiation - radiolysis of bases	Ward, 1994; Ma et al., 2012; Marin et al. 2018.
Anticancer chemotherapies: cross-linking agents (e.g., cisplatin) radiomimetic compounds (e.g., phleomycin)	Chen and Stubbe, 2005; Wyrobek et al., 2005; Jekimovs et al., 2014

Table 3.1. Examples of how double strand DNA breaks can arise.

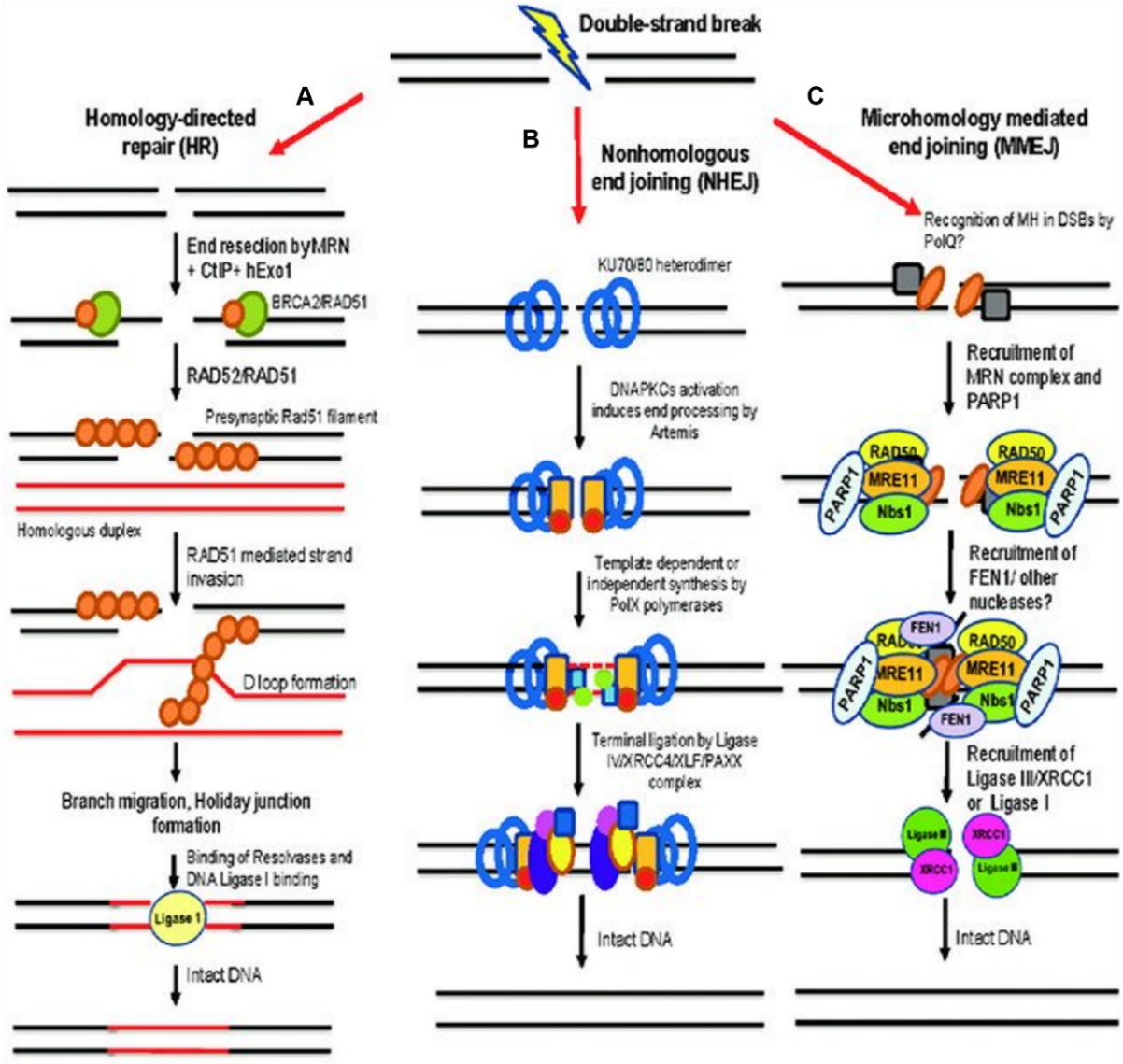


Fig. 3.1. Double strand DNA repair pathways expressed in eukaryotic cells.
A-C. Double strand DNA repair in eukaryotes can occur by error free-homology directed methods using sister chromatid via homologous recombination (**A**) or alternatively using error-prone non-homologous end-joining (**B**). In the absence of factors responsible for non-homologous end-joining, an alternative end joining mechanism called microhomology-mediated end-joining (**C**) can be used. Image adapted from Chang et al., 2017.

Homologous recombination is a high-fidelity DSB repair mechanism that uses a related but not always identical DNA sequence (usually a sister chromatid) as template to repair the lesion. This mechanism can be divided into three major steps: 1. DSB detection, 2. Initiation of DSB repair and 3. Repair of DSB by effector proteins. In the HR pathway (Figure 3.1A), a DSB is detected by the MRN (MRE11/RAD50/NBS1) complex (Gobbini et al., 2016). Here, MRE11 in conjunction with the ancillary protein CtIP initiates the exonucleolytic resection of the 5' ends to

generate 3' overhangs (Sartori et al., 2007; Gobbini et al., 2016) with the MRN complex recruiting and causing the autophosphorylation of ATM (Jones et al., 2014). Assembly of MRN/ATM complex at a damage site promotes phosphorylation of H2AX to γH2AX with this post-translational modification serving as an epigenetic marker of DSB (Bakkenist & Kastan, 2003; Houtgraaf et al., 2006; Kozlov et al., 2011). Formation of γH2AX modulates histone modifications that induces relaxing of chromatin structure (Ward &Chen 2001). This facilitates loading of downstream DNA repair factors such as the resection complex (EXO1, BLM and DNA2) that catalyse further nucleolytic processing of the 3' overhangs to generate longer single stranded DNA (ssDNA) overhangs that then serve as template onto which RPA binds with this interaction resulting in stabilisation of the resultant structure and facilitating effector proteins interactions (Tan et al., 2002; Mimitou & Symington, 2009; Nimonkar et al., 2011). Next, the RPA-ssDNA interaction recruits ATRIP and TOPBP1 that leads to activation of the ATR kinase causing cell cycle arrest (Lazzaro et al. 2009; Sancar et al. 2004; Warmerdam & Kanaar 2010). Activated ATR then facilitates exchange of DNA filaments from RPA to RAD51 which along with its paralogs (e.g., RAD51B, RAD51C, RAD51D, XRCC2 and XRCC3) and recombination mediators (e.g., BRCA1, BRCA2, PALB2, RAD52) that function to stabilise and protect the nucleoprotein filament as well as assisting identification of related sequences. Several mechanisms of how HR repairs the DSB have been proposed e.g., double strand DNA repair, synthesis dependent strand annealing and break induced repair, with these pathways extensively reviewed elsewhere (Ranjha et al., 2018).

In the NHEJ pathway of most eukaryotes (Figure 3.1B), the protein heterodimer Ku (composed of Ku70 and Ku80) recognise and bind the lesion leaving the DNA ends exposed (Gottlieb & Jackson, 1993; Ramsden & Geliert, 1998; Ranjha et al., 2018). This prevents enzymatic degradation of the DNA ends and allows mobilisation of other factors from the NHEJ pathway to this site. DNA-dependent protein kinase catalytic subunit (DNA-PKcs) tethers Ku bound exposed dsDNA ends after which the XRCC4-XLF complex along with DNA ligase IV are recruited to the lesion resulting in ligation of the two DSB ends. As formation of DSB may cause degradation of short regions (generally less than 10 bp long) flanking the break site, once consequence of the NHEJ pathway can be the deletion of sequence from the genome.

For DNA ends that cannot be ligated, for example, those containing blocking groups, gaps or overhangs, DNA-PKcs recruits Artemis to perform further nucleolytic processing of these ends.

(Ma et al. 2002; Chang et al 2017; Lobrich and Jeggo 2017). To generate DNA ends that can be ligated, DNA end processing may also involve phosphorylation of 5' hydroxyl or removal of 3' phosphate groups (Chappell et al. 2002). Another step in resolving DSB gap by NHEJ involve gap filling by DNA polymerases μ and λ (Bebenek et al. 2014; Moon et al. 2014). Once the ends are resolved, XRCC4-XLF and DNA ligase IV are recruited to ligate the two DNA ends and thereby restore the DNA duplex.

In the MMEJ pathway (Figure 3.1C) short overlapping homologies of 2-20 nucleotides between the two broken strands are used to join the two ends (Glover 2011; Chang et al, 2017). The initial stages of this system mirror that of the HR-pathway involving multiple repair proteins including the MRN complex, WRN, and CtlP (Audebert et al., 2004; Wang et al., 2003) that recognise and process the DSB to generate ssDNA. Initially, phosphorylated CtlP stimulates the 3'-5' endonuclease activity of MRE11 to extensively trim strand ends generating ssDNA overhangs that contain the microhomologies. DNA Polθ catalyses annealing reaction between the two homologous regions (Carvajal-Garcia., 2019). Next, overhanging bases are removed by flap endonuclease (FEN1) and any gaps filled by DNA polymerase activity. The DNA duplex is then restored through the concerted action of ligase III/XRCC1 (Audebert et al., 2004).

Orthologues and the accompanying activities of many components of the HR-based pathways as expressed by model eukaryotes have been identified in trypanosomes, reviewed by Genois et al. 2014; da Silva, 2021. These systems represent the major mechanisms by which DSBs are repaired in these parasites and parallels the pathway noted above, except the histone involved in the process is H2A, which undergoes phosphorylation to γH2A (Glover & Horn, 2009), while some ancillary proteins such as RAD52 are apparently missing (Passos-Silva et al 2010; Genois et al. 2014). In terms of its importance to trypanosomes, the HR pathway has been shown to play a major role in the antigenic variation process *T. brucei* employs to evade host immune responses where it facilitates gene switching mechanisms that allow the parasite to express new forms of the variant surface glycoproteins, the dominant surface antigen found on this organism, reviewed by Morrison et al 2009; Vink et al 2012; da Silva, 2021. Additionally, functional studies of components in this pathway (e.g., MRE11, RAD51) have helped to unravel the mechanism of action of nitroimidazole-based trypanocidal compounds such as benznidazole (Rajão et al., 2014; Gomes Passos Silva et al., 2018; Dattani et al 2021) and megazol (Enanga et al., 2003),

In this chapter, we explore the role *T. brucei* HR pathway plays in resolving DNA damage generated by trypanocidal nitroheterocyclic prodrugs. By employing existing parasite lines engineered to express altered levels of MRE11, CSB and/or NTR1, we show that a type I nitroreductase (TbNTR1) catalyses activation of the newly licensed prodrug fexinidazole. This generates metabolites that go on to promote formation of a type of DNA damage (presumably DSB) which are resolved in a pathway involving the concerted action of MRE11 and CSB. We hypothesise that CSB may play a role in transcription-coupled HR DNA repair mechanisms, an activity outside of its 'classical' parameter. Next, and using a PCR-based CRISPR-CAS9 genome editing strategy (Beneke et al., 2017), parasites expressing tagged versions of MRE11 and RAD51 were made. Initial experiments revealed the subcellular location of each enzyme in bloodstream and procyclic form parasites. Downstream experiments were performed to investigate the temporal expression of TbMRE11 and TbRAD51 in response to trypanocidal nitroheterocyclic prodrugs. The collected data confirm that both enzymes have a nuclear localisation throughout the cell cycle and in all life cycle stages examined with the expression levels of these enzymes showing temporal changes in response to external stimuli such as benznidazole, fexinidazole and nifurtimox. The findings reported here that relate to benznidazole form a significant component of the work covered in the publication Dattani et al 2021.

3.2 Informatic analysis of TbMRE11

Literature searches (Robinson et al., 2002; Tan et al., 2002) followed by BLAST analysis (Altschul et al., 1990) of *T. brucei* genome data held on TriTrypDB (Aslett et al., 2010) identified a single 2259 bp open reading frame (ORF) (Tb927.2.4390; designated as Tb*mre11*) from the *brucei* TREU927 strain. This sequence was located on chromosome 2 of the parasite's nuclear genome and had potential to encode for an 80.2 kDa enzyme (designated as TbMRE11). Pairwise analysis revealed the Tb*mre11* nucleotide sequence was 64 % identical to Tc*mre11* from *T. cruzi* CL Brener Esmeraldo-like haplotype (TcCLB.509099.70), 52 to 54% identical to the equivalent gene expressed by *Crithidia fasciculata* (CFAC1_210035200), *Leishmania donovani* (LdBPK.27.2.001790) and *Leishmania major* (LmjF.27.1890), and 47 to 49 % identical to its *S. cerevisiae* (YMR224C), *H. sapiens* (NM_005591) and *A. thaliana* (NM_124806) counterparts.

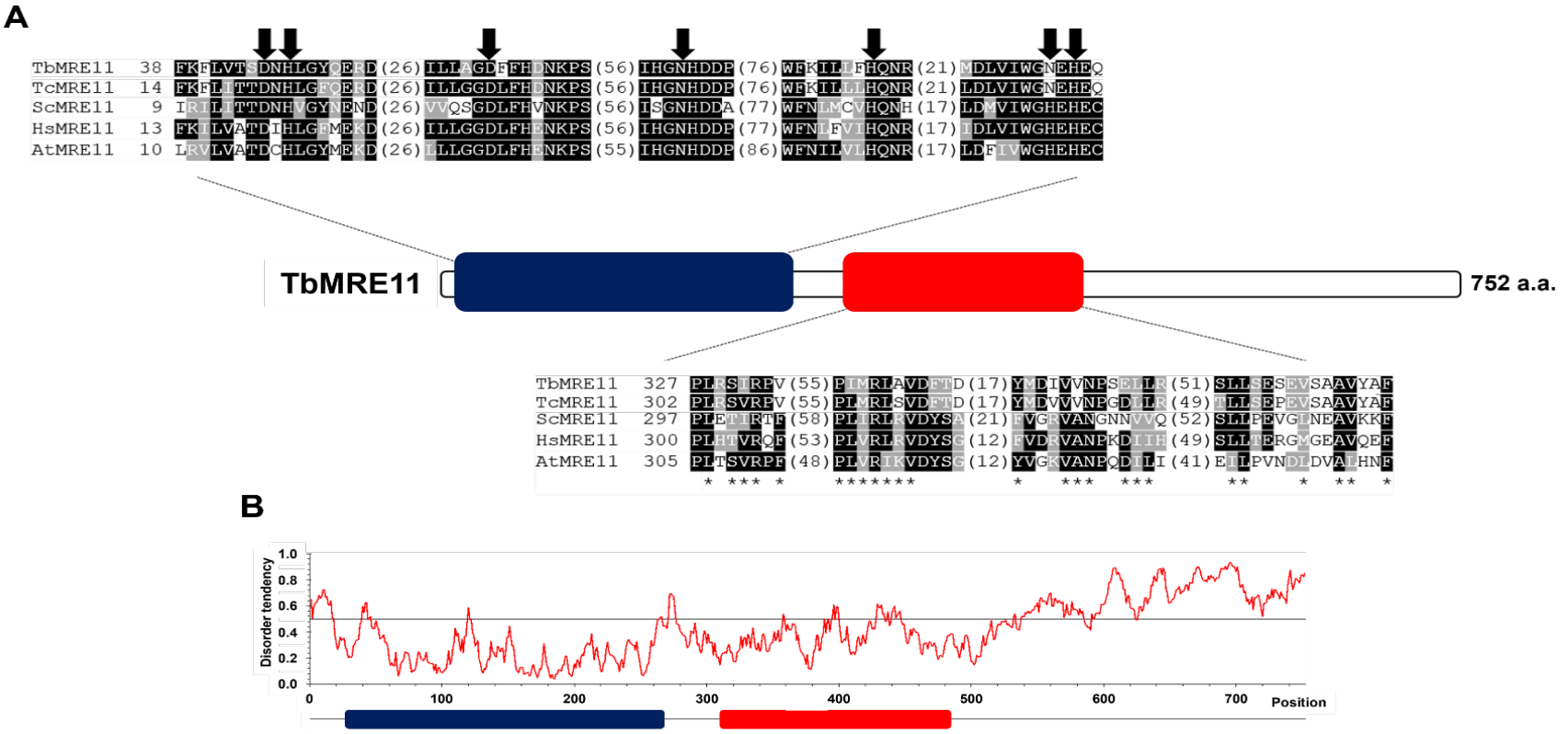


Fig. 3.2. Informatic analysis of TbMRE11.

A. The sequence corresponding to the metallophos calcineurin-like phosphoesterase (PF00149; residues 27 to 267; blue box) and Mre11 DNA-binding (PF04152; residues 310 to 484; red box) domains of TbMRE11 (XP_951632) were aligned with the equivalent regions from other members of the MRE11 family of nucleases: *T. cruzi* TcMRE11 (XP_816077), *Saccharomyces cerevisiae* ScMRE11 (P32829), *Homo sapiens* HsMRE11 (NP_005582) and *Arabidopsis thaliana* AtMRE11 (NP_200237). Residues that are identical in at least three of the five sequences and conserved substitutions are highlighted in black and grey, respectively. Amino acids in the calcineurin-like phosphoesterase domain postulated to coordinate metal (magnesium) co-factor binding are denoted with a down arrow while residues highlighted with an asterisk in the DNA-binding domain representing amino acids that constitute the HMM logo.

B. The metallophos calcineurin-like phosphoesterase and Mre11 DNA-binding domains are highlighted in blue and red, respectively. Disordered regions were obtained by submitting the TbMRE11 sequence to disorder prediction programs (IUPRED2; https://iupred2a.elte.hu/). A disorder value >0.5 indicates an unstructured protein region.

As with other members of this family of nucleases, TbMRE11 contains a calcineurin-like phosphoesterase domain (PF00149) towards its amino terminal (residues 27 to 267) and a centrally located MRE11 DNA binding domain (PF04152) (residues 310 to 484) (Figure 3.2A) with a 3D prediction of the monomer that covers residues 17-425 showing the structural organisation of these regions (Figure 3.3): Residues 1-28 and 426-752 could not be modelled as these regions lack homology to known structures and are predicted to be disordered (Figure 3.2B).

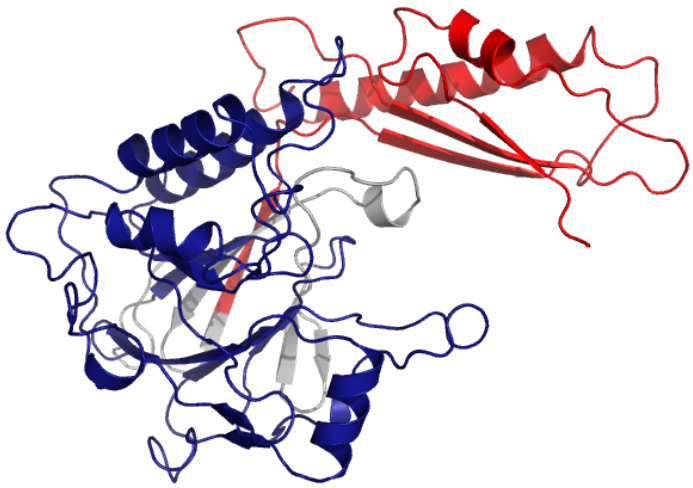


Fig. 3.3. 3D prediction model of a TbMRE11 monomer.
3D prediction of a TbMRE11 monomer (XP_951632) based on the *C. thermophilum* MRE11 (PDB: 4YKE; Seifert et al 2015) was constructed using Phyre2 (Kelley et al. 2015) and visualized using PyMOL: There was 100 % confidence over 52 % (394 residues) in the model to the template. The position of the calcineurin-like phosphoesterase (blue) and presumed Mre11 DNA-binding domains (red) are shown. Non-domain regions are in grey.

The calcineurin-like phosphoesterase domain (often referred to as the nuclease domain) of TbMRE11 consists of a double β-sheet sandwich that by comparison with the equivalent region from the *Chaetomium thermophilum* and *H. sapiens* orthologues contains several appropriately positioned residues that mediate metal ion binding (Figure 3.4) (Park et al., 2011; Seifert et al., 2016). In TbMRE11, D34, H36, D75, and H265 are postulated to be involved in binding one manganese ion with D75, N143, H231 and N263 potentially contributing to the binding of the second metal atom. In other MRE11s, several regions associated with this domain are involved in protein oligomerization and/or interactions with many of these discernible in the trypanosomal orthologue (Figure 3.5) (Park et al 2011). For example, TbMRE11 is predicted to dimerize through the interfaces between a loop (residues 141-152) found in monomer 1 with a loop found in monomer 2 (residues 166-174) and between a loop (residues 99-135) in monomer 1 with two α helices (residues 84-97 and 153-160) in monomer 2. Additionally, the loop formed by residues 99-135 is predicted to contribute to TbMRE11's interaction with NBS1. The Mre11 DNA binding domain (often called the capping domain) is predicted to be divided across four β-sheets, two α

helices plus the intervening loops (Figure 3.3). Three of these β-sheets have a parallel arrangement and are flanked to one side by the α helices while the remaining β-sheet is part of the double β-sheet sandwich of the calcineurin-like phosphoesterase domain.

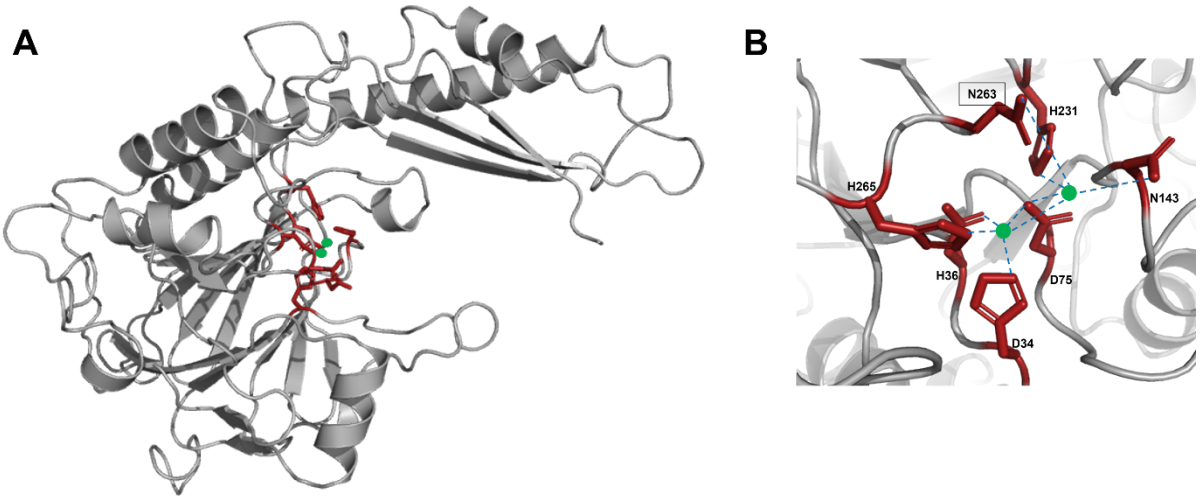


Fig. 3.4 Manganese-binding residues and catalytic site.
A & B. Wide (**A**) and zoomed (**B**) views showing residues from TbMRE11 postulated to be involved in manganese binding was made following comparison against MRE11 sequences from the *C. thermophilum* (PDB: 4YKE) and *H. sapiens* (PDB: 3T1I) (Park et al 2011; Seifert et al 2015). Binding of one manganese ion involves D34, H36, D75 and H265 while D75, N143, H231 and N263 potentially contribute to the binding of the second metal atom. The boxed residue (N250) represents a difference in the metal binding site noted between trypanosomal and fungal/human sequences.

Fig. 3.5. Protein: protein interaction regions of TbMRE11.
Comparison of the TbMRE11 approximation with the *C. thermophilum* (PDB: 4YKE) and *H. sapiens* (PDB: 3T1I) orthologues (Park et al 2011; Seifert et al 2015) identified regions that could mediate protein: protein interactions. TbMRE11 dimerization is predicted to involve residues 141-152 (purple) that form a loop in monomer 1 which interacts with a loop formed by residues 166-174 (green) in monomer 2. Further, residues 99-135 (red) in monomer 1 form a loop that interacts with two α helices (residues 84-97; 153-160; cyan) in monomer 2. Residues 99-135 (red) may contribute to NBS1 binding.

3.3 Informatic analysis of TbRAD51

Using the published literature (McCulloch & Barry, 1999) and data held at the NCBI resource, the *T. brucei* RAD51 orthologue was identified. Analysis of the TriTryp database revealed this to be a single copy 1297 bp gene (Tb927.11.8190; designated as Tb*rad51*) located on chromosome 11 of the *T. brucei* TREU927 genome with potential to code for a 40.4 kDa enzyme (designated as TbRAD51). Alignment of the Tb*rad51* nucleotide sequence with the orthologues expressed by other organisms revealed it to be between 67 to 70% identical to the *T. cruzi* (TcCLB.503801.30), *L. major* (LmjF.28.0550), *L. donovani* (LdBPK_280580.1) and *C. fasciculata* (CFAC1_300011500) orthologues while also sharing extensive identity (~52 to 55%) with its *S. cerevisiae* (P25454), *H. sapiens (*5H1B_A) and *A. thaliana* (NM_122092) counterparts. Analysis of its primary amino acid sequence revealed TbRAD51 has a domain structure typical of recombinases of a eukaryotic origin (Figure 3.6). Towards its amino terminal (residues 64-111) the trypanosomal enzyme contains a binding helix-hairpin-helix_5 (HHH_5; PF14520) domain, a dsDNA binding region that is absent from the prokaryotic RECA orthologues (Baumann & West, 1998), followed by a RAD51 (PF08423) domain (residues 116-370). Further analysis of the latter region revealed it contains several important motifs including a Walker A motif ($G(X)_4GKT$; residues 160 to 168) that functions in ATP and Mg^{2+} binding and a Walker B motif (LhhhD; h represents hydrophobic amino acids; residues 249 to 255) that catalyses ATP hydrolysis (Chi et al. 2006; Xu et al 2017). These sequences are key for enzyme activity with ATP-binding being required in the formation of the presynaptic filament complex and homologous strand exchange while ATP hydrolysis leads to the disassociation of RAD51-containing complexes (Bugreev & Mazin 2004; Chi et al. 2006; Modesti et al. 2007; Holthausen et al. 2010). Other sequences such as a short polymerization motif (residues 118-124) that tether individual RAD51 subunits together to form assemblies (Shin et al 2003) and two loop regions (residues 264-271 and 302-322) involved in ssDNA binding (Story & Steitz, 1992; Pellegrini et al., 2002) are present in TbRAD51 (Figure 3.6).

Construction of a 3D prediction of the trypanosomal enzyme that covers residues 56-373 showed the two-domain structure typical of the RAD51 family of recombinases (Figure 3.7). The amino terminal region of this model, often referred to as the NTD or N-terminal DNA binding domain, incorporates the entire HHH_5 domain and is predicted to fold into a compact region made up of four α helical bundles organised as two orthogonally packed α hairpins. The is followed

by an unstructured section in the RAD51 domain that contains the PM motif (Figure 3.8). The model predicts this to be on the surface of the protein outside interface site and as such is positioned to facilitate interactions that result in oligomerization of RAD51. The remainder of the RAD51 domain is predicted to contain nine α helices and six β sheets. Four of these β sheets have a parallel arrangement and are sandwiched between three α helices while the other two β sheets are proposed to have an anti-parallel arrangement with respect to each other. The Walker A and B motifs that mediate the enzyme's catalytic function form in a pocket with the former sequence found at the pocket entrance while the latter, which forms one of the parallel arranged β sheets, being buried deeper at this site. The two loops (loop 1 and 2) are predicted to be toward the surface of the protein where they can bind ssDNA with these being situated near to the Walker A- and B-containing pocket.

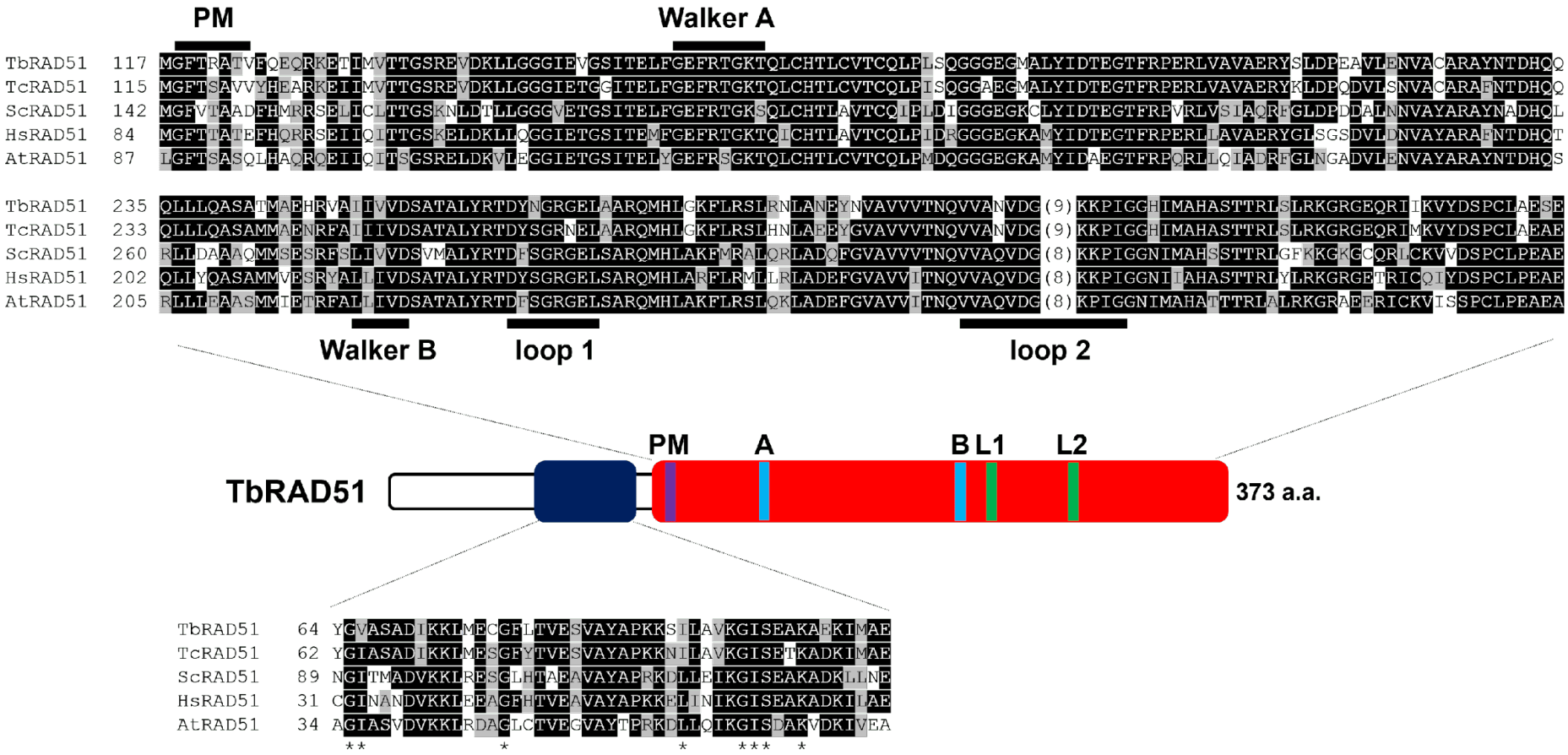


Fig. 3.6. Informatic analysis of TbRAD51.
The sequence corresponding to the helix-hairpin-helix_5 (HHH_5; PF14520; residues 64-111; blue box) and RAD51 (PF08423; residues ;116-370; red box) domains of TbRAD51 (XP_828893) were aligned with the equivalent regions from other RAD51 recombinases from: *T. cruzi* TcRAD51 (AAO72729), *S. cerevisiae* ScRAD51 (NP_011021), *H. sapiens* HsRAD51 (5H1B_A) and *A. thaliana* AtRAD51 (XP_020876169). Residues that are identical in at least three of the five sequences and conserved substitutions are highlighted in black and grey, respectively. The Walker A (A) and B (B) motifs, a polymerization motif (PM) and loop 1 (L1) & loop 2 (L2) are shown. Residues in the HHH_5 domain marked with an asterisk constitute parts of this regions HMM logo.

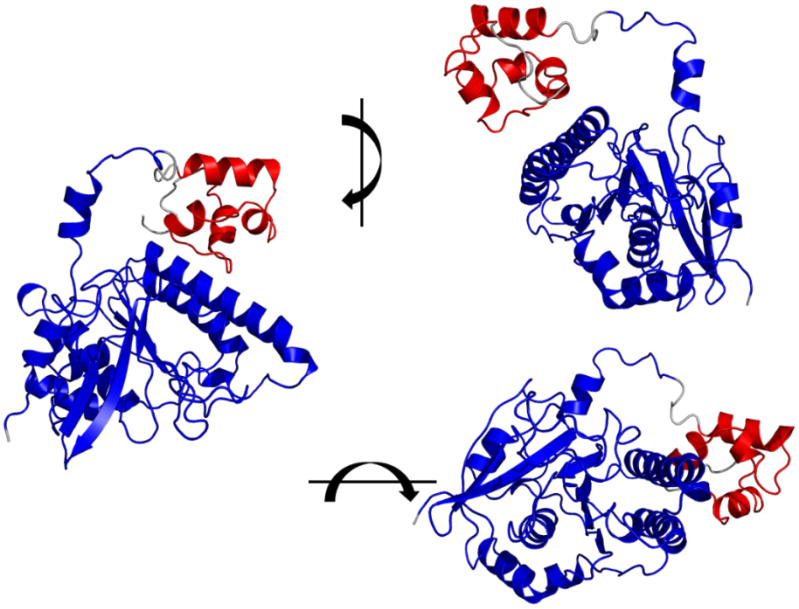


Fig. 3.7. 3D prediction model of TbRAD51.
3D prediction of a TbRAD51 (XP_828893) monomer based on *H. sapiens* RAD51 (5JZC; Short et al 2016) crystal structure was constructed using Phyre2 (Kelley et al. 2015) and visualized using PyMOL: There was 100% confidence over 83% (or 309 residues) of the model to the template. The position of the helix-hairpin-helix_5 (in red) and RAD51 (in blue) domains are shown. Non-domain regions are in grey. Several orientations of the approximation are shown.

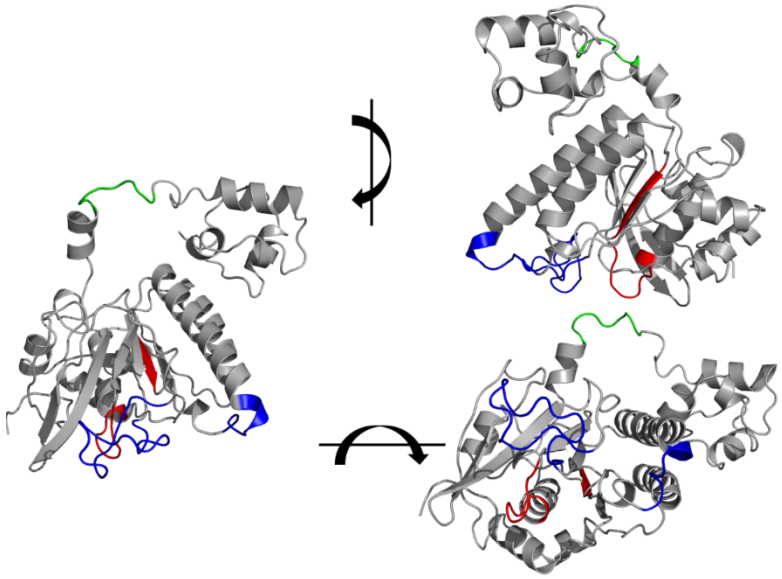


Fig. 3.8. Location of motifs and loops in the TbRAD51 3D prediction model.
The position of the Walkers A and B (in red), polymerization motif (green) and loops 1 and 2 (blue) in the TbRAD51 3D prediction model are shown. Several views of the motif/loop arrangements are shown.

3.4 Unravelling the mechanism of action of fexinidazole

Previous studies have shown that part of the trypanocidal activity displayed by the nitroheterocycles benznidazole and nifurtimox involves their ability to promote DNA damage (Goijman et al 1985; Ferreira & Ferreira, 1986; Campos *et al.,* 2017) with the type of lesions generated by the former requiring the HR pathway to resolve (Rajao *et al.,* 2014; Gomes Passos Silva *et al.,* 2018; Dattani et al 2021). To assess whether this DNA damaging effect extends to

fexinidazole, a newly licensed treatment that can be used against all stages of HAT (Deeks, 2019), the susceptibility of recombinant BSF *T. brucei* lines null for TbMRE11, TbCSB or TbSNM1 (Sullivan et al 2015; Dattani and Wilkinson, 2019), markers of the HR, TC-NER and ICL DNA repair pathways, respectively, was determined (Figure 3.9). Evaluation of the resultant dose response curves showed that BSF parasites lacking TbCSB or TbMRE11 were more susceptible (4-5-fold) to fexinidazole than wild type indicating that this nitroimidazole can promote DNA damage in the nuclear genome.

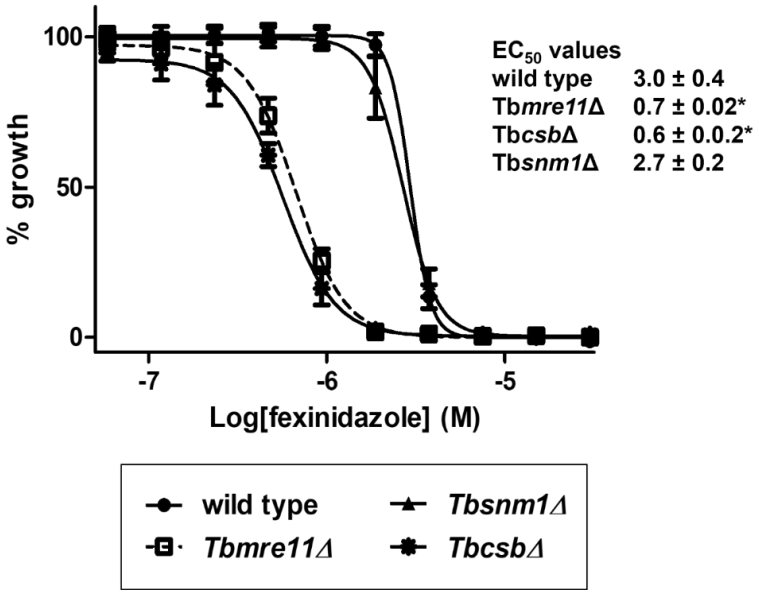


Fig. 3.9. Susceptibility of *T. brucei* DNA repair mutants towards fexinidazole.
The growth inhibitory effect of fexinidazole on BSF *T. brucei* wild type and various DNA repair null mutant lines was evaluated. Dose response curves were constructed from which EC50 values (in μM) were extrapolated. Data points are means ± standard deviation from experiments performed in quadruplicate. The asterisk (*) indicates significant differences in susceptibility ($P<0.0001$) between wild type and genetically modified cells, as assessed by Student's t test (GraphPad Software).

To investigate for any functional linkage between TbCSB and TbMRE11 in the repair of the fexinidazole-generated DNA damage, the susceptibility of parasites null for both activities was determined (Figure 3.10). This revealed that cells lacking TbCSB and TbMRE11 were equally as sensitive to this nitroimidazole as parasites lacking TbCSB or TbMRE11 alone, indicating that these repair factors function epistatically to resolve DNA damage arising from this treatment.

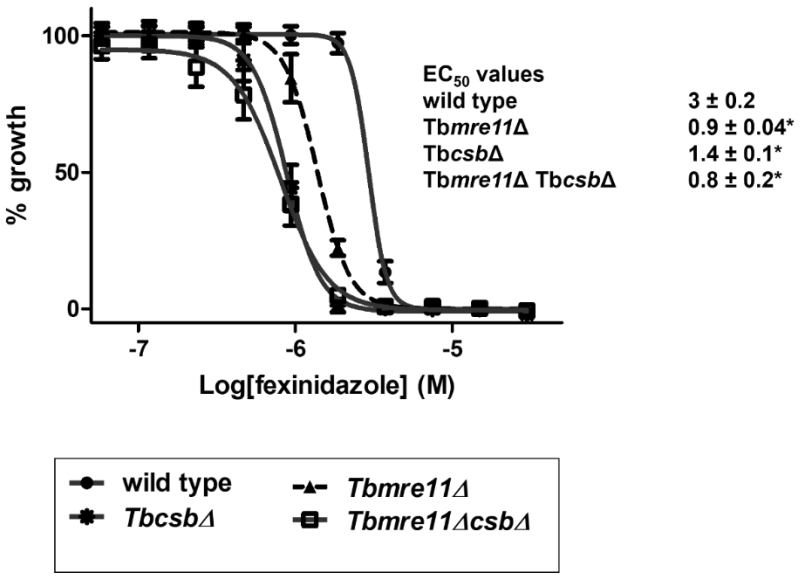


Fig. 3.10. Susceptibility of *T. brucei* Tb*mre11*Δ Tb*csb*Δ null line towards fexinidazole.
The growth inhibitory effect of fexinidazole on BSF *T. brucei* wild type, Tb*mre11∆*, Tb*csb∆* & Tb*mre11∆*Tc*csb∆* was assessed. Dose response curves were constructed from which EC_{50} values (in μM) extrapolated. Data points are means ± standard deviation from experiments performed in quadruplicate. The asterisk (*) Indicates significant differences in susceptibility (P<0.0001) between wild type and genetically modified cells, as assessed by Student's *t* test (GraphPad Software).

It is well documented that benznidazole and nifurtimox function as prodrugs with their activation involving a parasite specific type I nitroreductase (Wilkinson et al 2008; Baker et al 2011; Hall et al 2011; Hall & Wilkinson, 2012). To determine if this is the case for fexinidazole, the susceptibility of *T. brucei* that can be induced to express elevated levels of the major type I nitroreductase found in this parasite, TbNTR1 (Wilkinson et al 2008), was determined (Figure 3.11). Evaluation of the resultant dose response curves showed that BSF parasites containing higher levels of TbNTR1 were hypersensitive (11-fold) to this nitroimidazole than cells expressing wild type levels of the activator demonstrating that this enzyme plays a key role in mediating the trypanocidal activity of fexinidazole.

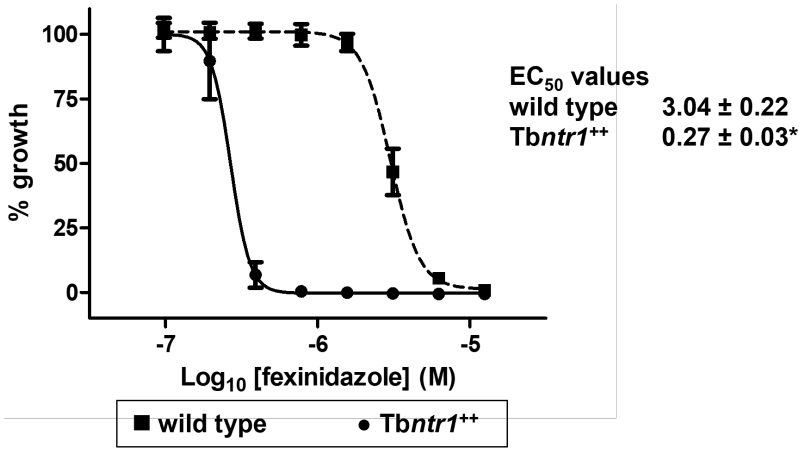


Fig. 3.11. Susceptibility of *T. brucei* expressing elevated levels of Tb*ntr1* towards fexinidazole. The growth inhibitory effect of fexinidazole on BSF *T. brucei* wild type and cells expressing higher levels of TbNTR1 (Tb*ntr1*$^{++}$) was assessed. Dose response curves were constructed from which EC_{50} values (in μM) were extrapolated. Data points are means ± standard deviation from experiments performed in quadruplicate. The asterisk (*) indicates significant differences in susceptibility (P<0.0001) between wild type and genetically modified cells, as assessed by Student's *t* test (GraphPad Software).

To determine if there is a link between activation of the fexinidazole and the TbMRE11-/TbCSB-mediated DNA repair pathways, the susceptibility of lines null for Tb*mre11* or Tb*csb* that also express ectopic copies of Tb*ntr1* was evaluated (Figure 3.12 A & B) (Dattani et al 2021). While assessing for a link between NTR1 activation and the HR pathways, the previously noted increase in sensitivity of the Tb*mre11* null line and *T. brucei* expressing elevated levels of Tb*ntr1* to the nitroheterocycle relative to controls was observed (Figure 3.12A). This susceptibility phenotype was enhanced in Tb*mre11*-deficient trypanosomes that also express an ectopic copy of Tb*ntr1*. For example, *T. brucei* lacking Tb*mre11* or over expressing Tb*ntr1* were 2-or 12-fold more susceptible to fexinidazole relative to wild type, respectively, with a 19-fold increase in potency noted to parasites deficient in TbMRE11 that also express elevated levels of TbNTR. Extending this investigation to look for any association between prodrug activation and the TbCSB repair pathway, treatment of the Tb*csb* null line or *T. brucei* expressing elevated levels of TbNTR1 with fexinidazole resulted in an increased susceptibility towards the nitroheterocycle with the Tb*csb* null line that also over expresses TbNTR1 displaying hypersensitivity (Figure 3.12B). This susceptibility data demonstrates a clear link between prodrug activation and DNA repair pathway(s) with the TbNTR1-mediated activation of fexinidazole promoting formation of lesions that are resolved by a pathway(s) in which TbMRE11 and/or TbCSB plays a key role.

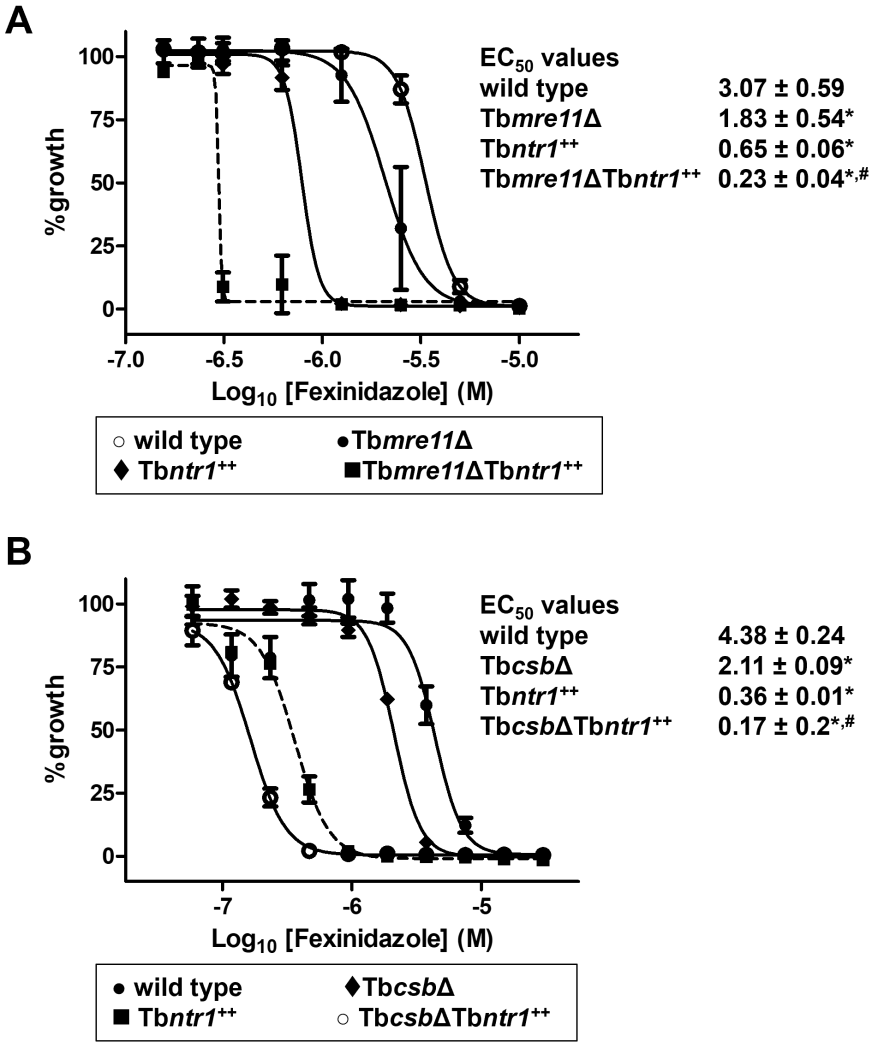


Fig. 3.12. Linking fexinidazole activation to DNA damage.

A. Growth inhibitory effect of fexinidazole on BSF *T. brucei* wild type, *T. brucei* expressing an ectopic copy of Tb*ntr1* (Tb*ntr1*$^{++}$), a Tb*mre11* null line (Tb*mre11Δ*) and a Tb*mre11* null line expressing elevated levels of Tb*ntr1* (Tb*mre11Δ* Tb*ntr1*$^{++}$).

B. Growth inhibitory effect of fexinidazole on BSF *T. brucei* wild type, *T. brucei* expressing an ectopic copy of Tb*ntr1* (Tb*ntr1*$^{++}$), a Tb*csb* null line (Tb*csbΔ*) and a Tb*csb* null line expressing elevated levels of Tb*ntr1* (Tb*csbΔ* Tb*ntr1*$^{++}$).

In **A** and **B**, dose response curves were constructed from which EC_{50} values (in μM) were extrapolated. Data points are means ± standard deviation from experiments performed in quadruplicate. The asterisk (*) indicates significant differences in susceptibility (P<0.0001) between the Tb*mre11Δ*, Tb*csbΔ*, Tb*ntr1*$^{++}$, Tb*mre11Δ* Tb*ntr1*$^{++}$or Tb*csbΔ* Tb*ntr1*$^{++}$ lines relative to wild type, while # indicates significant differences in susceptibility (P<0.0001) between Tb*mre11Δ* Tb*ntr1*$^{++}$cells and the Tb*mre11Δ* or Tb*ntr1*$^{++}$lines, and between the Tb*csbΔ* Tb*ntr1*$^{++}$ cells and the Tb*csbΔ* or Tb*ntr1*$^{++}$lines. Significance assessed by Student's *t* test (GraphPad Software).

3.5 Editing the Tb*mre11* and Tb*rad51* loci in *T. brucei*.

Based on the PSORTII and WoLFPSORT algorithms TbMRE11 was predicted to be in the parasite's nucleus, targeted to this organelle with three "four pattern" or one "seven pattern" sequences postulated to function as nuclear localisation signals (NLSs): The three proposed "four pattern" NLSs are RHKK (residues 533-536), RKPK (residues 698-701) and KPKR (residues (699-702) while the "seven pattern" is PARGRKP (residues 694-700. When these informatic screens were extended to TbRAD51, a cytoplasmic location was predicted contrary to previous findings (Glover and Horn 2012).

To confirm/assess the localisations of TbMRE11 and TbRAD51, a genome editing approach was employed and involved the CRISPR/Cas9-mediated tagging of the endogenous Tb*mre11* or Tb*rad51* loci at their 5' or 3' end with nucleotide sequences coding for the fluorescent protein mNEONGREEN (Figure 3.13) (Beneke et al., 2017). Here, two amplicons were electroporated into BSF or PCF *T. brucei* with recombinant, drug (hygromycin) resistant clones selected. The smaller of the two amplicons contained sequences that when expressed in the parasite generate a single guide RNA (sgRNA). These nucleic acids bind a resident Cas9 before recognising/binding to the 5' or 3' region of the target gene. The nuclease then introduces a DSB at this site in the genome. The larger DNA fragment corresponded to a targeting fragment that serves as template to repair the Cas9-generated DSB in the 5' or 3' region of target gene and in doing so introduces the *mNeonGreen* ORF along with a selectable marker (*hyg*) at the appropriate site. The primers combinations used to generate the different amplicons are listed below (Table 3.2)

	Primer combinations	
	sgRNA	**Targeting fragment**
Tb*mre11*		
5' tagging	G00/5'sgRNA Tbmre11	Tbmre11-UFP1/Tbmre11-DRF1
3' tagging	G00/3'sgRNA Tbmre11	Tbmre11-DFP1/Tbmre11-DRTAGP
Tb*rad51*		
5' tagging	G00/5'sgRNA Tbrad51	Tbrad51-UFP1/Tbrad51-DRF1
3' tagging	G00/3'sgRNA Tbrad51	Tbrad51-DFP1/Tbrad51-DRTAGP

Table 3.2. Primer combinations used in tagging Tb*mre11 & Tbrad51*.
The primer combinations used to generate amplicons corresponding to the single guide RNA (sgRNA) and targeting fragment used in tagging the 5' or 3' ends of the endogenous Tb*mre11* or Tb*rad51* loci with a sequences coding for the fluorescent protein mNEONGREEN. The sequence of each primer can be found in Appendix 1.

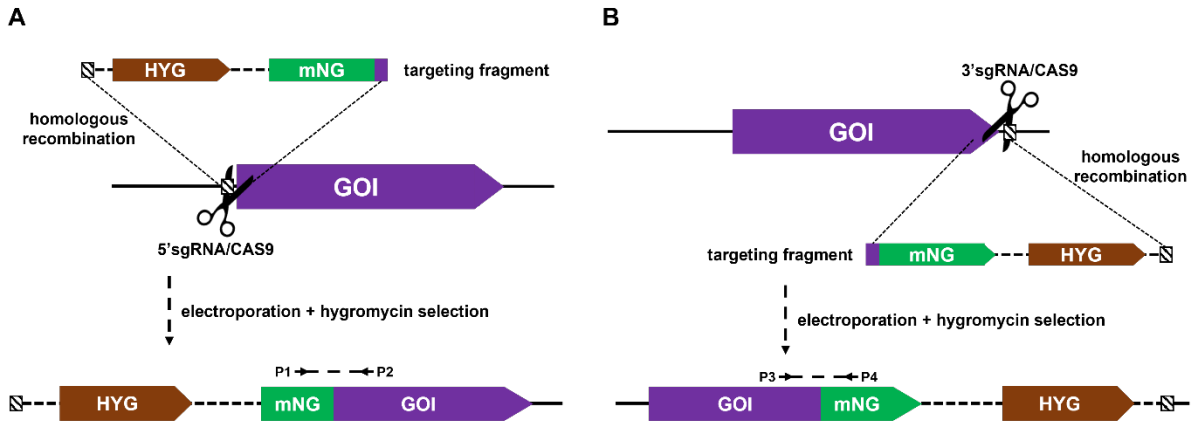


Fig. 3.13. CRISPR/Cas9-mediated tagging of the endogenous loci.
A & B. Schematic diagram outlining the CRISPR/Cas9-mediated tagging of the 5' (**A**) and 3' (**B**) regions of a gene of interest (GOI), in this case Tb*mre11* and Tb*rad51*. The sgRNA/Cas9 complex (indicated by scissors) introduces a DSB at the target site which can be repaired using a targeting fragment: This latter DNA contains sequences that encode for the resistance determinate hygromycin B phosphotransferase (HYG) and mNeonGreen (mNG), flanked by 30 bp sequences homologous to the targeted region (hatched and purple boxes). The lower image shows the edited alleles with P1-P4 corresponding to primers sequences used to validate the lines: The primer combinations and sequences used for each amplification are listed in Table 3.1 and Appendix 1, respectively.

To confirm that the introduced targeting fragments had integrated at the correct loci, DNA amplification reactions using gDNA or cDNA templates were performed with primer combinations that span the gene fusion event (Figure 3.14A & C). When using primer combinations intended to detect the *mNeonGreen*-Tb*mre11* fusion, PCR generated a 1.0 kb band, the expected size, from nucleic acid templates (cDNA) originating from BSF recombinant parasites with no band(s) observed in templates from wild type cells (Figure 3.14B). A similar outcome was noted for Tb*mre11-mNeonGreen* expressing BSF *T. brucei*: here a band of ~1.0 kb was noted in templates derived from genetically modified cells but not from wild type (Figure 3.14). For both tagging events, the cDNA-derived amplicons were sequenced, with analysis showing that the Tb*mre11* and *mNeonGreen* coding sequences were in-frame with each other (data not shown). When these validations were extended across to templates derived from procyclic form parasites, the same outcomes as seen in Figure 3.14 were obtained. A final validatory step was performed on PCF parasites expressing *mNeonGreen*-Tb*mre11* (Figure 3.15). Here, western blot analysis using an anti-mNeonGreen antibody detected a band of the expected 109 kDa size in extracts generated from recombinant parasites with no band(s) observed in lysates from wild type cells.

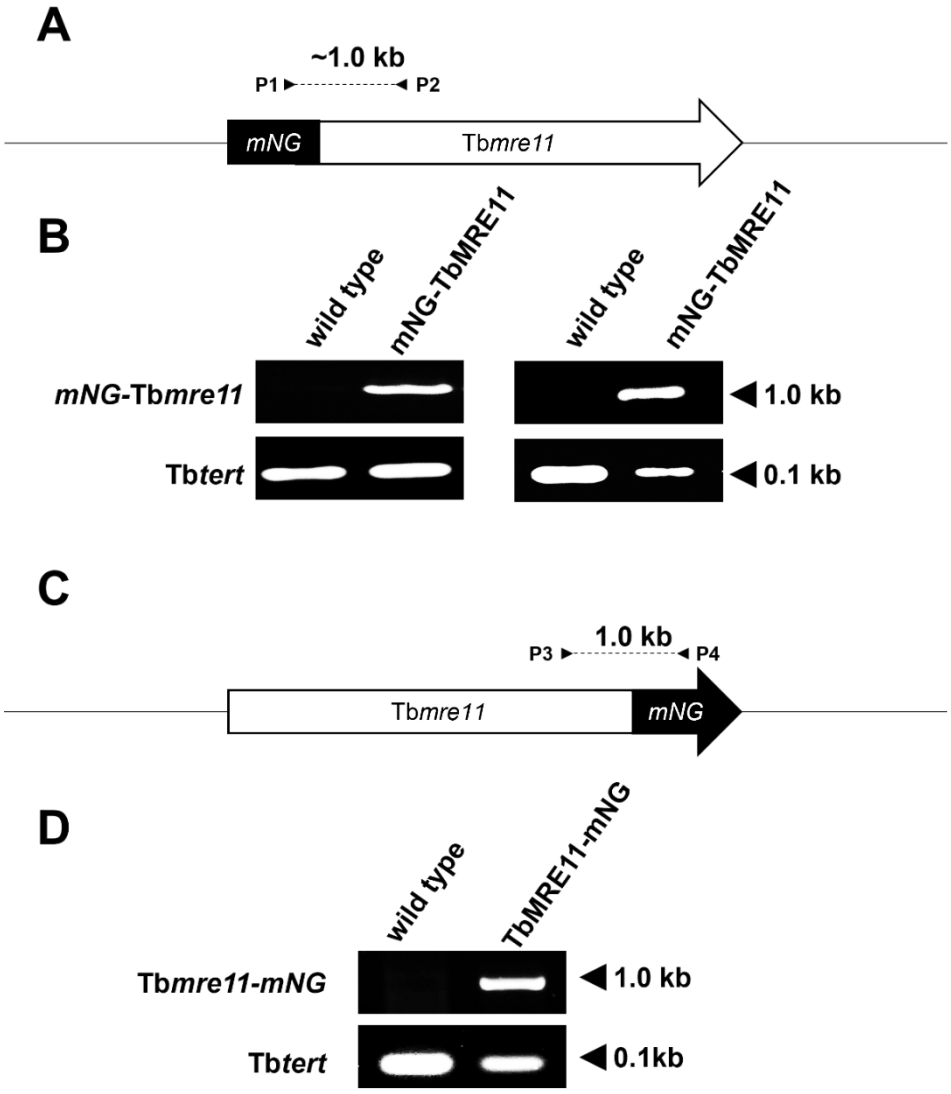


Fig. 3.14. Validation of *T. brucei* lines expressing mNeonGreen tagged BSF TbMRE11.
A & C. Schematic showing integration of *mNeonGreen* into the *T. brucei* genome at the 5' (**A**) or 3' (**C**) region of Tb*mre11*. P1 to P4 indicate the primers indicate primers used to generate the validation amplicons.
B & D. Amplicons (in kb) corresponding to the *mNeonGreen*-Tb*mre11* (**B**) or Tb*mre11-mNeonGreen* (**D**) were generated from template cDNAs derived from total RNA extracted from the *T. brucei* lines indicated. The integrity of DNA samples was evaluated by amplification of a 0.1 kb control fragment, Tb*tert* using primers Tbtert-F and Tbtert-R. Amplification using primer combinations mNeonGreen–1 (P1) & Tbmre11-KO-2 (P2) produced a band of ~1.0 kb, while Tbmre11-K03(P3) & mNeonGreen–2 (P4) produced a band of ~1.0 kb. Primer sequences and combinations used for each amplification are listed in Appendix 2,

Fig. 3.15. Validation of PCF parasites expressing mNeonGreen tagged TbMRE11 or TbRAD51.
Total cell extracts derived from 1 x 106 procyclic form *T. brucei* were size fractionated on a 10% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane. Blots were probed using an antibody to detect the mNeonGreen (mNG) tag. Bands of 109 and 77 kDa were detected and corresponded to mNG tagged versions of TbMRE11 and TbRAD51, respectively. No band was detected in extracts made from the wild type *T. brucei*. An extract from T. cruzi expressing mNG tagged luciferase (100 kDa), was used as positive control.

The same strategies were employed to validate *T. brucei* lines expressing tagged TbRAD51 (Figure 3.16). Here, PCRs using primer combinations intended to detect the *mNeonGreen*-Tb*rad51* or the Tb*rad51-mNeonGreen* fusions generated bands of the expected size (0.6 kb and 0.4 kb, respectively) from gDNA/cDNA templates originating from recombinant parasites with no band(s) observed in templates from wild type cells (Figure 3.16B & D). Again, and for both tagging events, the cDNA-derived amplicons were sequenced, with analysis showing that the Tb*rad51* and *mNeonGreen* coding sequences to be in-frame with each other (data not shown). When these validations were extended across to templates derived from procyclic form parasites, the same outcomes were obtained. A final validatory step was performed on PCF parasites expressing *mNeonGreen*-Tb*rad51* (Figure 3.15). Here, western blot analysis using an anti-mNeonGreen antibody detected a band of the expected 77 kDa size in extracts generated from recombinant parasites with no band(s) observed in lysates from wild type cells.

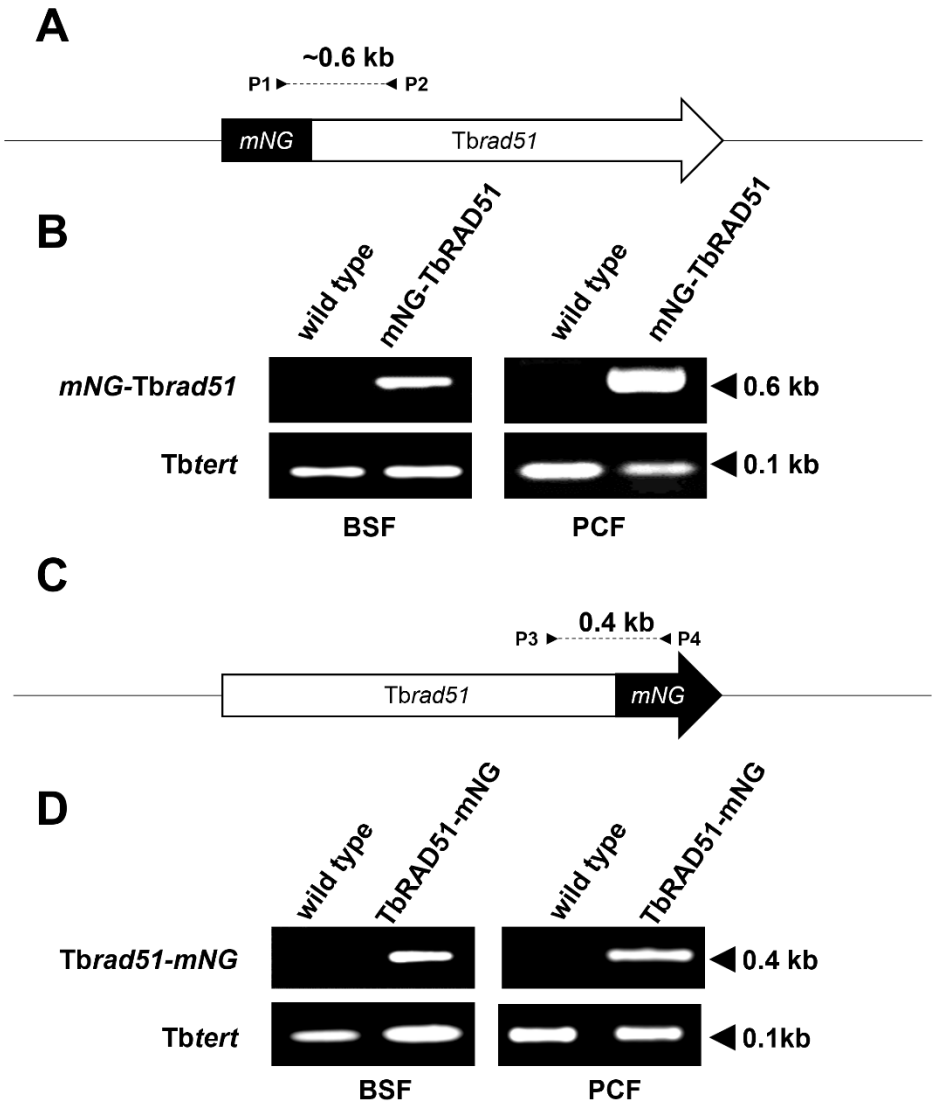


Fig. 3.16. Validation of *T. brucei* lines expressing mNeonGreen tagged TbRAD51.
A & C. Schematic showing integration of *mNeonGreen* into the *T. brucei* genome at the 5' (**A**) or 3' (**C**) region of Tb*mre11*. P1 to P4 indicate the primers indicate primers used to generate the validation amplicons. **B & D.** Amplicons (in kb) corresponding to the *mNeonGreen*-Tb*mre11* (**B**) or Tb*mre11-mNeonGreen* (**D**) were generated from template cDNAs derived from total RNA extracted from the *T. brucei* lines indicated. The integrity of DNA samples was evaluated by amplification of a 0.1 kb control fragment, Tb*tert* using primers Tbtert-F and Tbtert-R. Primers P1 and P4 represent mNeonGreen specific primers, mNeonGreen–1 &2, and P2 and P3 represent *Tbmre11* gene specific primers, Tbmre11-KO-2 & Tbmre11-K03 respectively. Primer sequences and combinations used for each amplification are listed in Appendix 2.

3.6 Localisation of TbMRE11 and TbRAD51 in *T. brucei*.

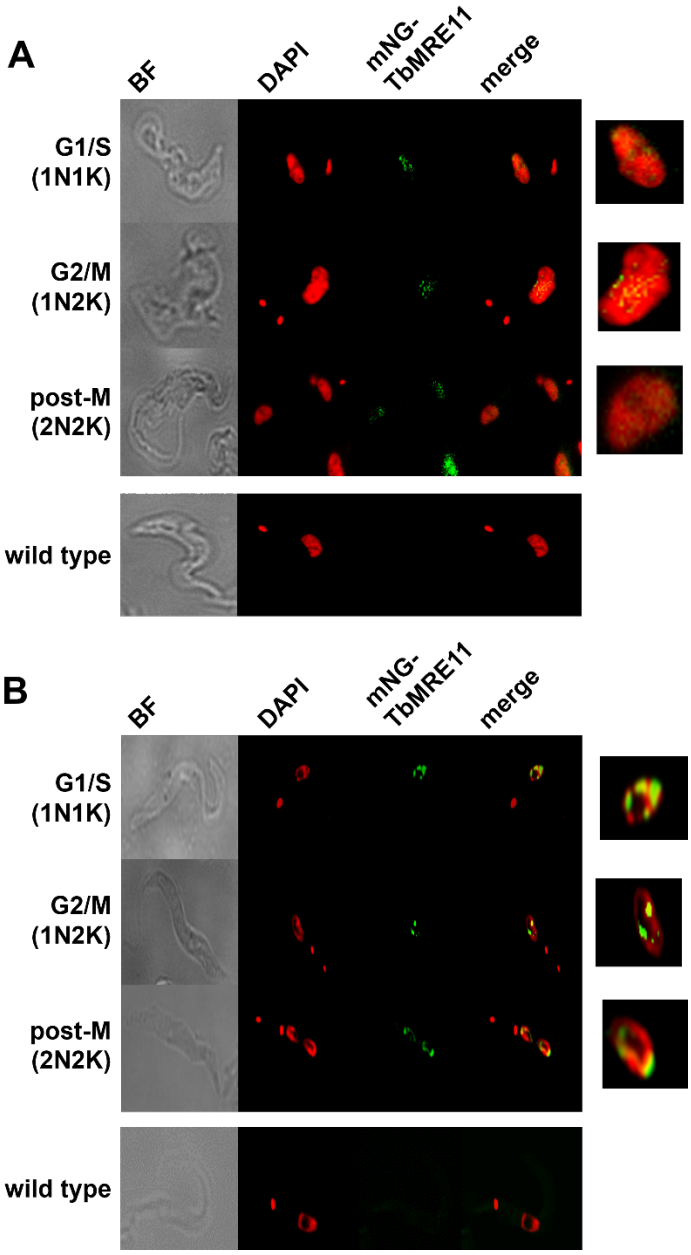


Fig. 3.17. Localisation of TbMRE11.

A & B. Bloodstream (**A**) and procyclic (**B**) form trypomastigotes *T. brucei* expressing TbMRE11 tagged at its amino terminal with mNeonGreen (mNG-TbMRE11) (green) from its endogenous loci were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow), including a close up image of the nucleus, in trypanosomes at different stages of the cell cycle are shown: The ratio of the nuclear (N) and mitochondrial (known as the kinetoplast; K) genomes in a single parasites represents a 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 (Glover & Horn, 2012; Siegel et al., 2008; Woodward & Gull, 1990). *T. brucei* expressing TbMRE11 tagged at its carboxyl terminal with mNeonGreen (TbMRE11-mNG) from the endogenous loci gave a similar pattern (data not shown).

To evaluate localisation, BSF or PCF *T. brucei* expressing TbMRE11 or TbRAD51 tagged at their amino or carboxyl termini with mNeonGreen were fixed, their DNA stained with DAPI then imaged by fluorescent microscopy (Figures 3.17 & 3.18). In both insect and mammalian form parasites, all wild type and about a third of *T. brucei* expressing a TbMRE11 variant no signal was detected. For the remaining recombinant trypanosomes, a weak signal throughout the DAPI-staining region of the nucleus (i.e., the nucleoplasm) was observed with some (10-20%) possessing

a few discrete foci: the non-DAPI staining region present in the trypanosome nucleus corresponds to the nucleolus (Schumann Burkard et al 2013; Bakari-Soale et al 2021). Localisation in the discrete punctate foci in the nucleoplasm were more in cells at the G1 phase of the cell cycle, whereas for those cells at the late S and G2/M phases, the localisation pattern shifts to a more coalesced with fewer dot foci or line foci in the nucleolus (Figure 3.17). Majority (90%) of untreated cells were in the G1/S phase of the cell cycle. Additional cytoplasmic signal was also observed around the kinetoplast in some cells. These cell cycle localisation patterns were observed in both life cycle forms examined.

For BSF and PCF trypanosomes engineered to express TbRAD51 tagged at its amino or carboxyl termini with mNeonGreen, most (~90%) cells presented a few (1 to 5) discrete foci in the nucleoplasm at all stages of the cell cycle with no discernible signal observed in the remaining (~10%) parasites. Cells undergoing nuclear duplication at the late S phase showed a more dispersed signal, with RAD51 signal seen dispersed across the enlarged nucleus. This may be due to involvement of TbRAD51 in DNA replication, and more compact signal seen in the S phase, dispersed signal as the nucleus divides, becoming more compact again in the G2/M phase.

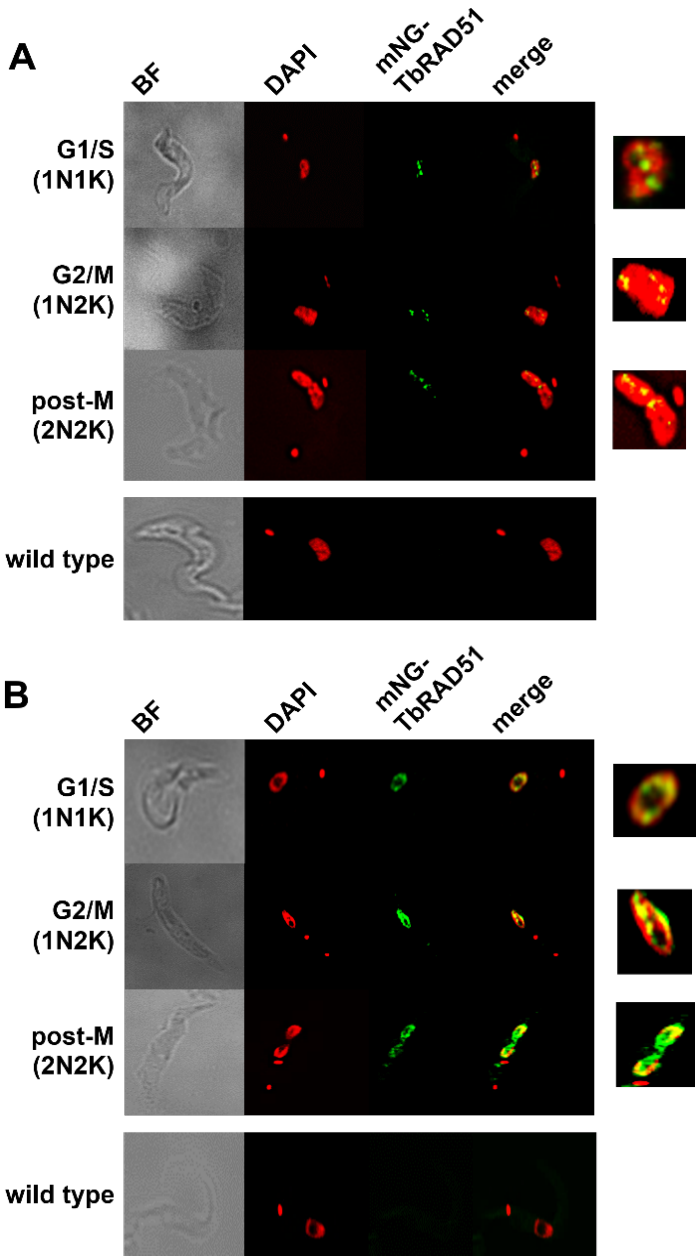


Fig. 3.18. Localisation of TbRAD51.
A & B. Bloodstream (**A**) and procyclic (**B**) form trypomastigotes *T. brucei* expressing TbRAD51 tagged at its amino terminal with mNeonGreen (mNG-TbRAD51) (green) from its endogenous loci were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow), including a close up image of the nucleus, in trypanosomes at different stages of the cell cycle are shown: The ratio of the nuclear (N) and mitochondrial (known as the kinetoplast; K) genomes in a single parasites represents a 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 (Glover & Horn, 2012; Siegel et al., 2008; Woodward & Gull, 1990). *T. brucei* expressing TbRAD51 tagged at its carboxyl terminal with mNeonGreen (TbRAD51-mNG) from the endogenous loci gave a similar pattern as that shown here.

3.7 Nitroheterocycles and their effect on the homologous recombination pathway.

In yeast and mammalian cells the localization and/or expression level of HR DNA repair factors changes in response to DNA damage (Lisby et al., 2004; Eckert-Boulet et al., 2011). To determine if benznidazole, nifurtimox and fexinidazole promote similar alterations in *T. brucei*, the fluorescent signal of BSF parasite lines engineered to express mNeonGreen variants of TbMRE11

or TbRAD51 grown in the presence of each nitroheterocycle was evaluated (Figure 3.19).

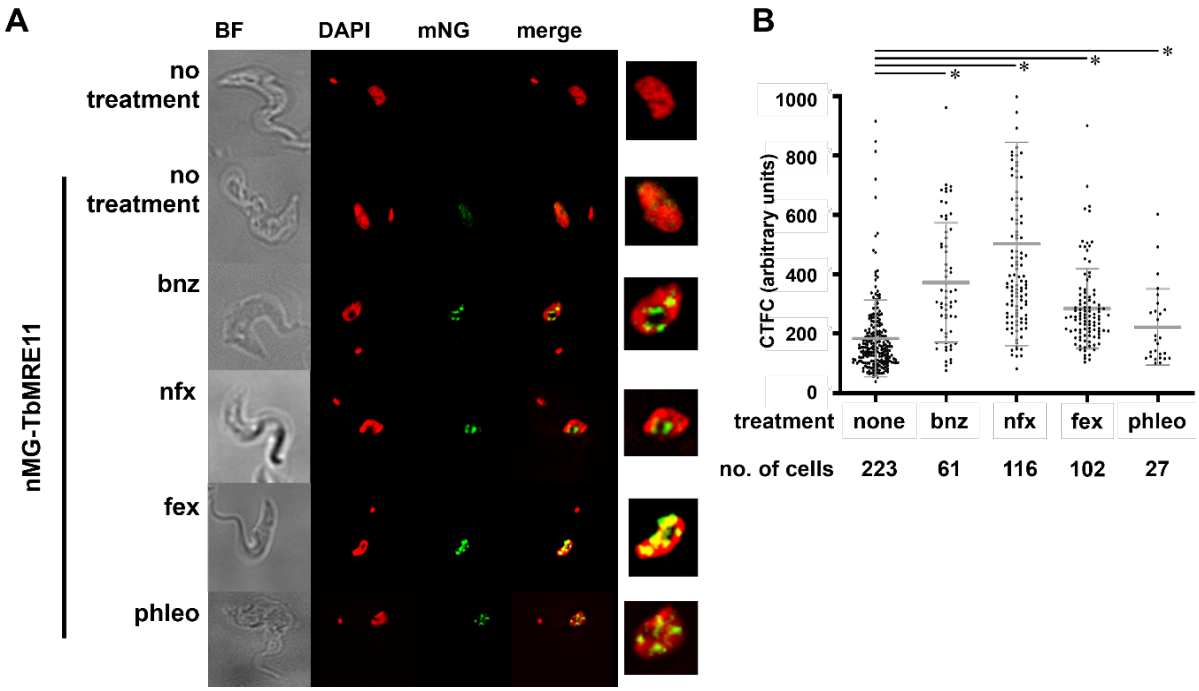


Fig. 3.19. Effect of exogenous treatments on TbMRE11 expression.
A. Bloodstream form trypomastigotes *T. brucei* expressing mNeonGreen-TbMRE11 (mNG; green) treated with benznidazole (bnz), nifurtimox (nfx), fexinidazole (fex) or phleomycin (phleo) were co-stained with the DNA stain, DAPI (red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merge; yellow), including a close-up image of the nucleus, in treated and untreated trypanosomes were compared.
B. The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen tagged TbMRE11 following benznidazole (bnz), nifurtimox (nfx), fexinidazole (fex) or phleomycin (phleo) treatment was determined and compared against untreated controls (none). Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).
A & B. Treatments used were: 30 µM benznidazole for 2 hours, 3 µM nifurtimox for 6 hours), 3 µM fexinidazole for 1 hour and 5 µg ml-1 phleomycin overnight.

Following treatment and at a single time point (30 µM benznidazole for 2 hours; 3 µM nifurtimox for 6 hours; 3 µM fexinidazole for 1 hour) the number of mNeonGreen-TbMRE11 expressing trypanosomes now possessing an intense punctate nuclear signal increased relative to untreated controls, a characteristic also noted in phleomycin-treated parasites (5 $\mu g/ml^{-1}$; overnight; Figure 3.19). These fluorescent centres were found in the nucleoplasm (Figure 3.19)

and present in cells at all stages of the cell cycle (data not shown). When the above was extended to mNeonGreen-TbRAD51 expressing BSF trypanosomes, in 60%–70% of cases, the number and intensity of the discrete foci increased relative to untreated controls for all agents tested. These fluorescent centres were found throughout the nucleoplasm (Figure 3.20) in cells at all stages of the cell cycle (data not shown). Intriguingly, some (5%–10%) drug-treated trypanosomes presented with a single, intense TbRAD51 spot possibly indicating that the multiple foci previously noted may have coalesced.

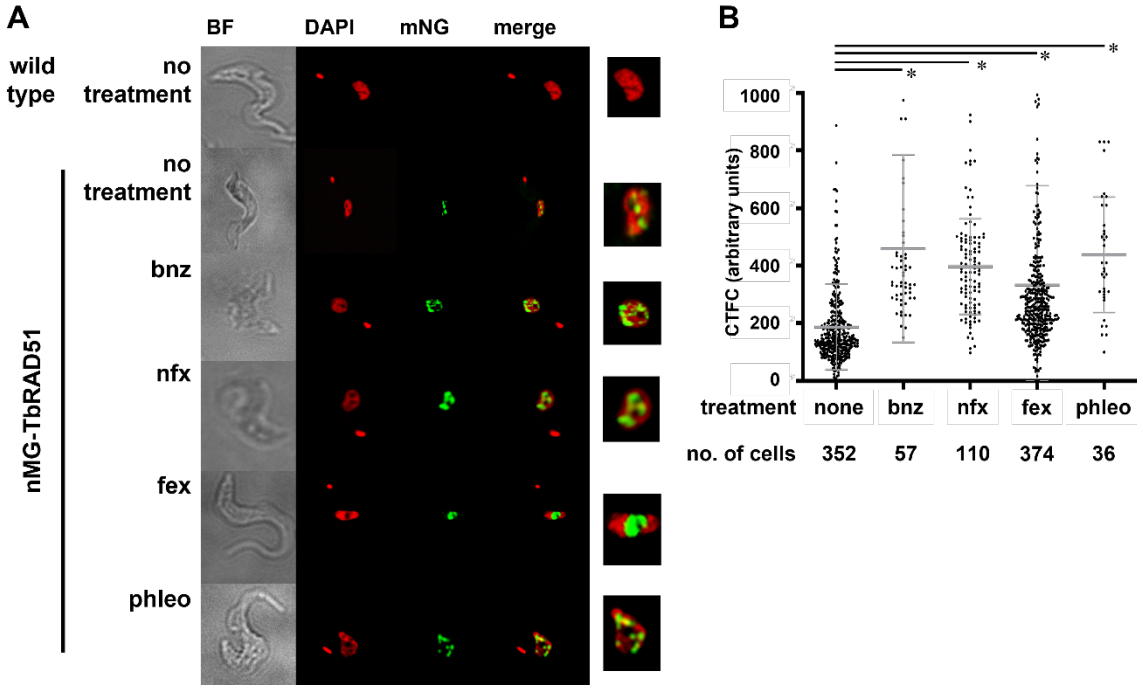


Fig. 3.20. Effect of exogenous treatments on TbRAD51 expression.
A. Bloodstream form trypomastigotes *T. brucei* expressing TbRAD51 tagged at its amino terminal with mNeonGreen (mNG; green) treated with benznidazole (bnz), nifurtimox (nfx), fexinidazole (fex) or phleomycin (phleo) were co-stained with the DNA stain, DAPI (red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merge; yellow), including a close-up image of the nucleus, in treated and untreated trypanosomes were compared. **B.** The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen tagged TbRAD51 following benznidazole (bnz), nifurtimox (nfx), fexinidazole (fex) or phleomycin (phleo) treatment was determined and compared against untreated controls (none). Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software). **A & B.** Treatments used were: 30 µM benznidazole for 2 hours, 3 µM nifurtimox for 6 hours), 3 µM fexinidazole for 1 hour and 5 µg ml^{-1} phleomycin overnight.

3.8 Nitroheterocyclic prodrugs and their effects on temporal expression of the trypanosomal HR pathway.

To assess the temporal response of the *T. brucei* HR pathway to benznidazole, the change in TbMRE11 and TbRAD51 expression, and in γH2A formation (an epigenetic DNA damage marker) was followed (Figure 3.21). BSF parasites expressing mNeonGreen-TbMRE11 or mNeonGreen-TbRAD51 were treated with a single bolus of benznidazole (30 μM) and periodically the fluorescent signals in individual cells determined (Figure 3.21A & B).

For mNeonGreen-TbMRE11 expressing trypanosomes, the previously noted increase in the numbers of cells possessing intense fluorescent centres in their nucleoplasm was shown to take place within the first hour of adding benznidazole to the culture medium (Figure 3.21A). This signal then remained at approximately the same level over the next 2 hours before declining 4-hours post treatment and reverting to untreated levels by 6 hours. Similarly, when BSF parasites expressing mNeonGreen-TbRAD51 were examined, the previously noted benznidazole-induced alterations to TbRAD51 expression were temporal, becoming apparent 4-hour posttreatment, a time point where the TbMRE11 signal had started to decline, and increasing at the final time point (6-hours post treatment) (Figure 3.21B).

In trypanosomes, the phosphorylation of H2A can be used as a biomarker for DNA damage including DSB formation (Glover & Horn, 2012). To determine if benznidazole can promote this posttranslational modification in *T. brucei*, protein extracts generated from wildtype parasites grown in the presence of the 2-nitroimidazole (30 μM) were probed for γH2A formation using an antiserum against this modified histone (Figure 3.21C): As a control, enolase was analysed in parallel. Comparison of the normalized γH2A and enolase band intensities revealed that the γH2A signal increased over the first 4 hr of treatment then declined over the next 4 hours indicating possible resolution of the DNA damage (Figure 3.21D). A summary of the combined temporal changes in γH2A formation and in TbMRE11 or TbRAD51 expression in response to benznidazole is presented in Figure 3.21E.

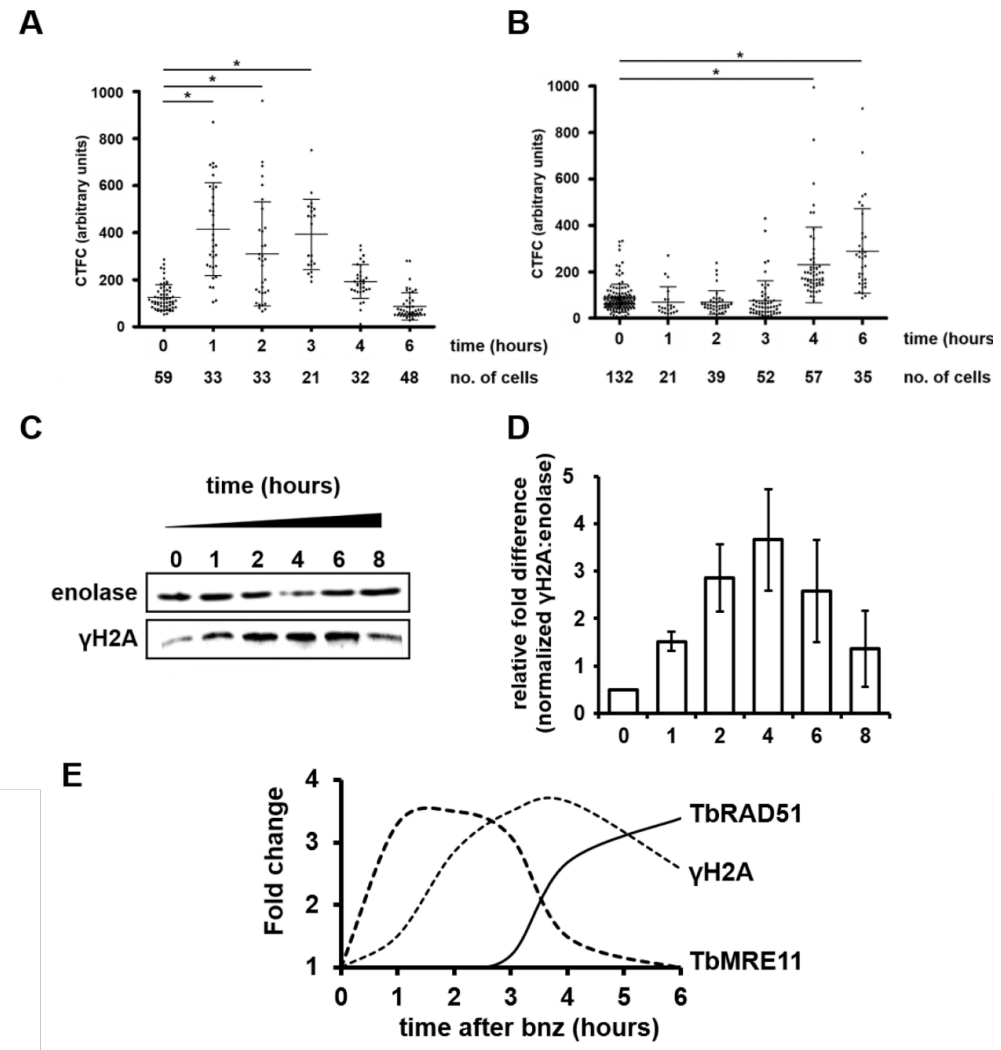


Fig. 3.21. Effect of benznidazole on the temporal expression of the *T. brucei* homologous recombination DNA repair pathway.

A & B. The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen tagged TbMRE11 (**A**) or TbRAD51 (**B**) was determined at time intervals following benznidazole (30 μM) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Student's *t* test (GraphPad Software).

C. *T. brucei* cells were treated with 30 μM benznidazole. Cell lysates were generated at time intervals (0, 1, 2, 4, 6, and 8 h) and analysed by western blot using anti-*T. brucei* enolase (loading control) and anti-*T. brucei* γH2A antiserum.

D. The γH2A and enolase signal intensities obtained from western blots containing extracts from three independent benznidazole treated cultures were determined using Image Studio™ Lite (Li−COR Biosciences) and normalised against untreated controls. The data derived from three independent blots is expressed as a mean ± standard deviation relative fold difference between the normalised γH2A and enolase signals.

E. Schematic showing the temporal changes in expression of TbMRE11 or TbRAD51, or γH2A formation following addition of benznidazole (bnz) to cultures.

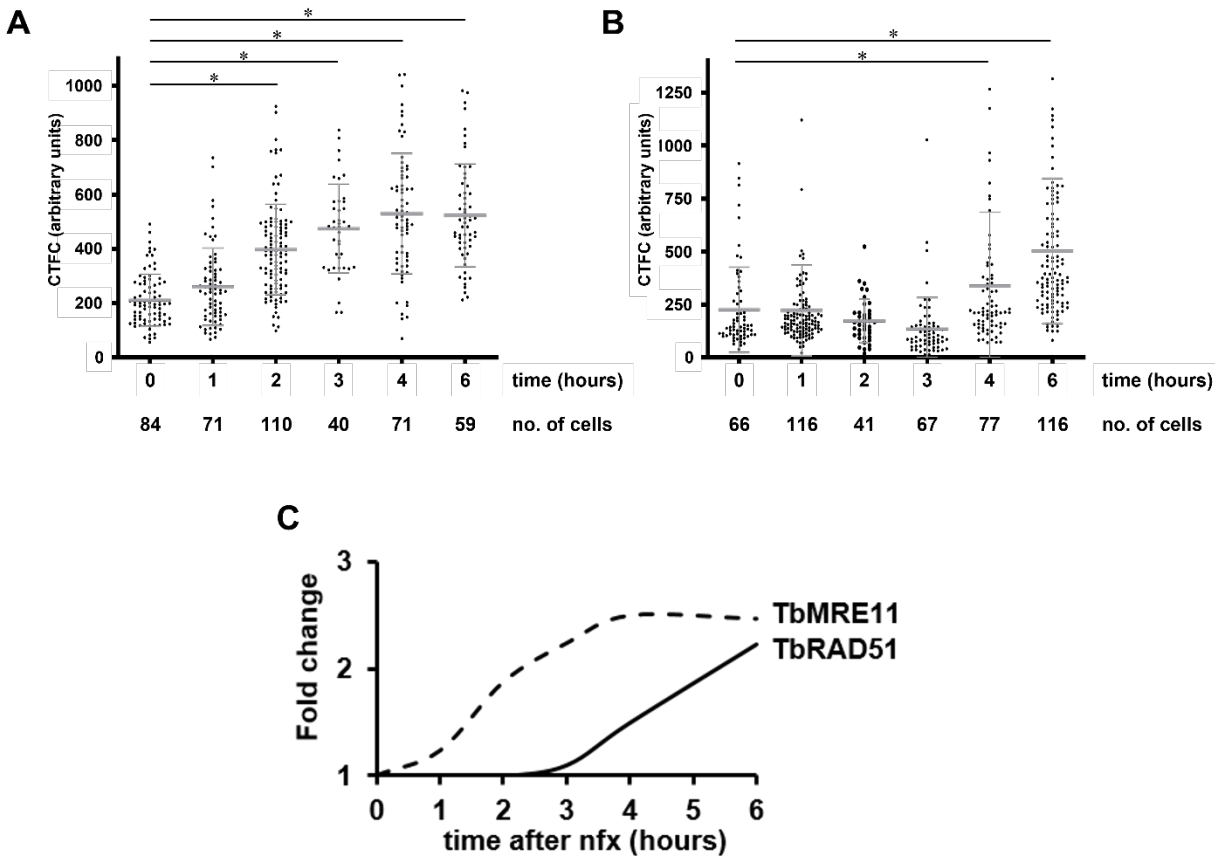


Fig. 3.22. Effect of nifurtimox on the temporal expression of the *T. brucei* homologous recombination DNA repair pathway.
A & B. The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen tagged TbMRE11 (**A**) or TbRAD51 (**B**) was determined at time intervals following nifurtimox (3 μM) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).
C. Schematic showing the temporal changes in expression of TbMRE11 or TbRAD51 following addition of nifurtimox (nfx) to cultures.

The above analysis was extended to evaluate if nifurtimox and/or fexinidazole promote similar changes in TbMRE11 and TbRAD51 expression as that noted for benznidazole (Figures 3.22 & 3.23): γH2A formation was not followed due to the lack of the appropriate antisera. When analysing whether nifurtimox caused any temporal changes to TbMRE11 expression, the numbers of cells possessing intense fluorescent centres in their nucleoplasm was shown to increase albeit at a slower rate than that noted when using benznidazole. Here, a slight but non-significant increase in TbMRE11 expression was noted over the 1st hour of treatment, with the fluorescent signal

gradually rising and reaching a maximum by 4-hours posttreatment, remaining at this level at the final time point (6-hours posttreatment) (Figure 3.22A). For *T. brucei* expressing mNeonGreen-TbRAD51, an increase in TbRAD51 expression became apparent 4-hours posttreatment and continued to rise to the final time point (6-hours posttreatment) (Figure 3.22B). A summary of the combined temporal changes in TbMRE11 or TbRAD51 expression in response to nifurtimox is presented in Figure 3.22C.

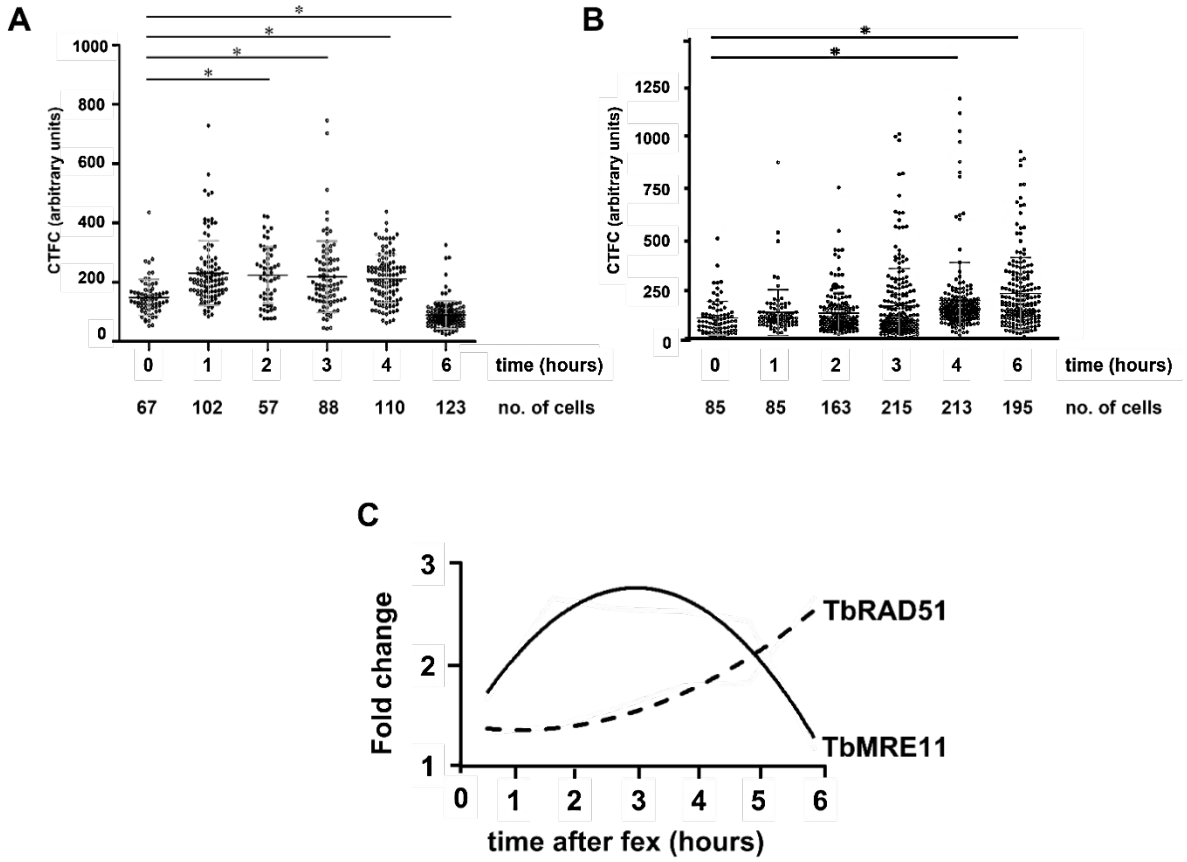


Fig. 3.23. Effect of fexinidazole on the temporal expression of the *T. brucei* homologous recombination DNA repair pathway.

A & B. The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen tagged TbMRE11 (**A**) or TbRAD51 (**B**) was determined at time intervals following fexinidazole (3 μM) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).

C. Schematic showing the temporal changes in expression of TbMRE11 or TbRAD51 following addition of fexinidazole (fex) to cultures.

When analysing whether fexinidazole caused any temporal changes to TbMRE11 and TbRAD51 expression, a pattern reminiscent of that observed using benznidazole was observed. When examining mNeonGreen-TbMRE11 expression, the numbers of cells possessing intense fluorescent centres in their nucleoplasm increased within the first hour of adding nitroimidazole to the culture medium (Figure 3.23A). This signal remained at approximately the same level over the next 2 hours before declining 4-hour posttreatment and reverting to untreated levels by 6 hours. When BSF parasites expressing mNeonGreen-TbRAD51 were examined, an increase in TbRAD51 expression become apparent 4-hour posttreatment, a time point where the TbMRE11 signal had started to decline, then increasing at the final time point (6-hours posttreatment) (Figure 3.23B). A summary of the combined temporal changes in TbMRE11 or TbRAD51 expression in response to fexinidazole is presented in Figure 3.23C.

3.9 Effect of UV on temporal expression of the trypanosomal HR pathway.

Exposure of eukaryotic cells to UV promotes formation of DSB with HR components playing a key role in resolving such lesions (Limoli et al 2002; Aboussekhra & Al-Sharif, 2005; Dattani & Wilkinson, 2019). To assess the temporal response of the *T. brucei* HR pathway to this mutagen, BSF parasites expressing mNeonGreen-TbMRE11 or mNeonGreen-TbRAD51 were treated with a single dose of UV (1500 J m^{-2}) and the cells allowed to recover with, periodically, the fluorescent signals in fixed individual cells determined (Figure 3.24).

When observing parasites expressing tagged TbMRE11, most cells 80-90%) presented with several small discrete foci throughout the nucleoplasm, the pattern noted in section 3.6. For the remaining trypanosomes (10-20%) alternative arrangements were seen including a subset of cells that harboured fewer but larger fluorescent centres in their nucleoplasm and others that contained a single, large spot (Figure 3.24A). In addition, a fluorescent signal outside of the nucleus was sometimes observed in cells displaying either of these two variants. Examination revealed that parasites possessing a single large TbMRE11 centre often displayed a rounded-up appearance, indicative that they were stressed/dying and thereby accounting for the non-nuclear signal. When these observations were extended to mNeonGreen-TbRAD51 expressing trypanosomes, UV treatment appears to promote re-distribution of this DNA repair enzyme. Here, most (90 %) untreated cells presented with several discrete fluorescent centres although a minority (10 %) possessed a single coalesced spot located in the nucleolus (Figure 3.24B). Exposure to UV

mutagen resulted in an increase in the number of cells upto16-30%) at all stages of the cell cycle containing the single TbRAD51 foci (Figure 3.24B).

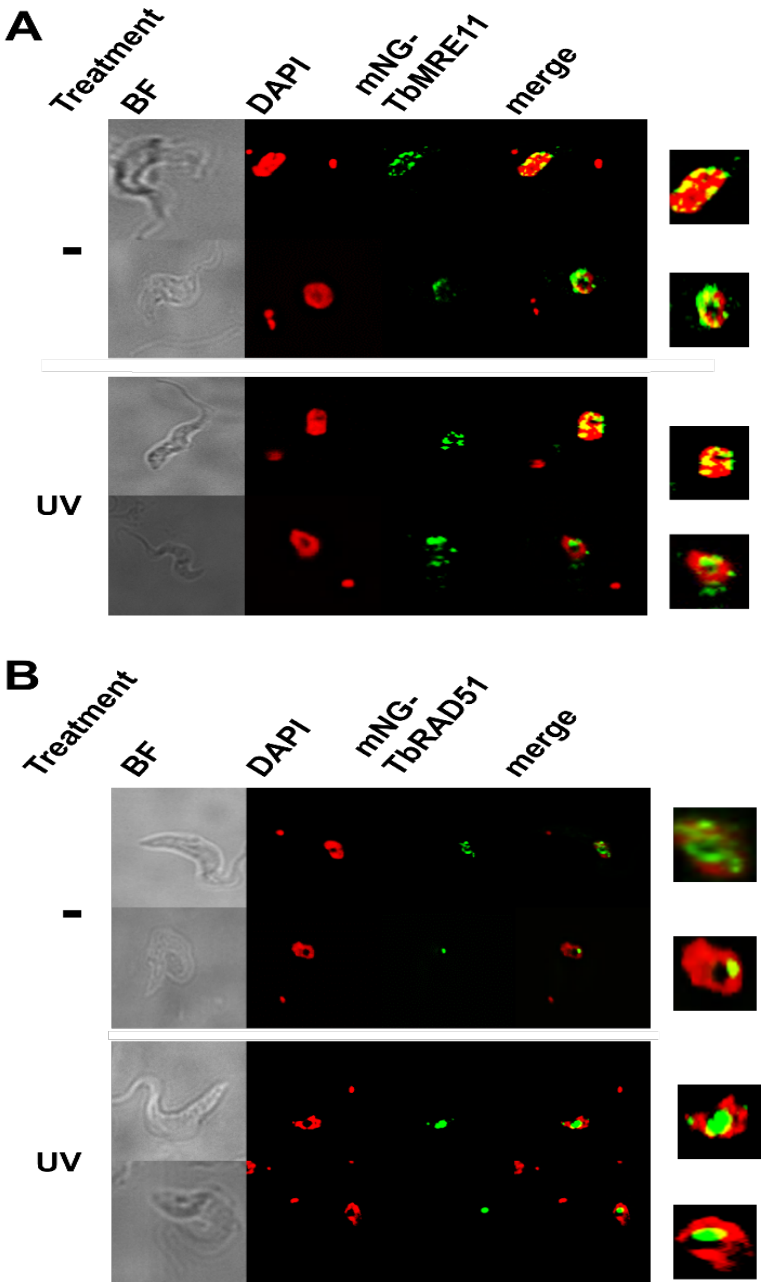


Fig. 3.24. Effect of UV treatment on TbMRE11 and TbRAD51 expression.
A & B. UV treated (UV) parasites expressing TbMRE11 (**A**) or TbRAD51 (**B**) tagged at their amino terminal with mNeonGreen (mNG-TbMRE11 & mNG-TbRAD51, respectively; green) were co-stained with DNA stain, DAPI (red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of co-localization (merge; yellow), including a close-up image of the nucleus are shown. Untreated (-) parasites were examined in parallel.

When analysing whether exposure to UV light caused any temporal changes to TbMRE11 or TbRAD51 expression, no significant difference in fluorescence was noted for both enzymes over the first 30 minutes of treatment. However, over the next few hours, the intensity of these signals

increased and reaching a peak by 4-hours post treatment before reverting to (TbMRE11) (Figure 3.25A) or declining towards (TbRAD51) (Figure 3.25A) untreated levels by 6 hours (Figure 3.25C).

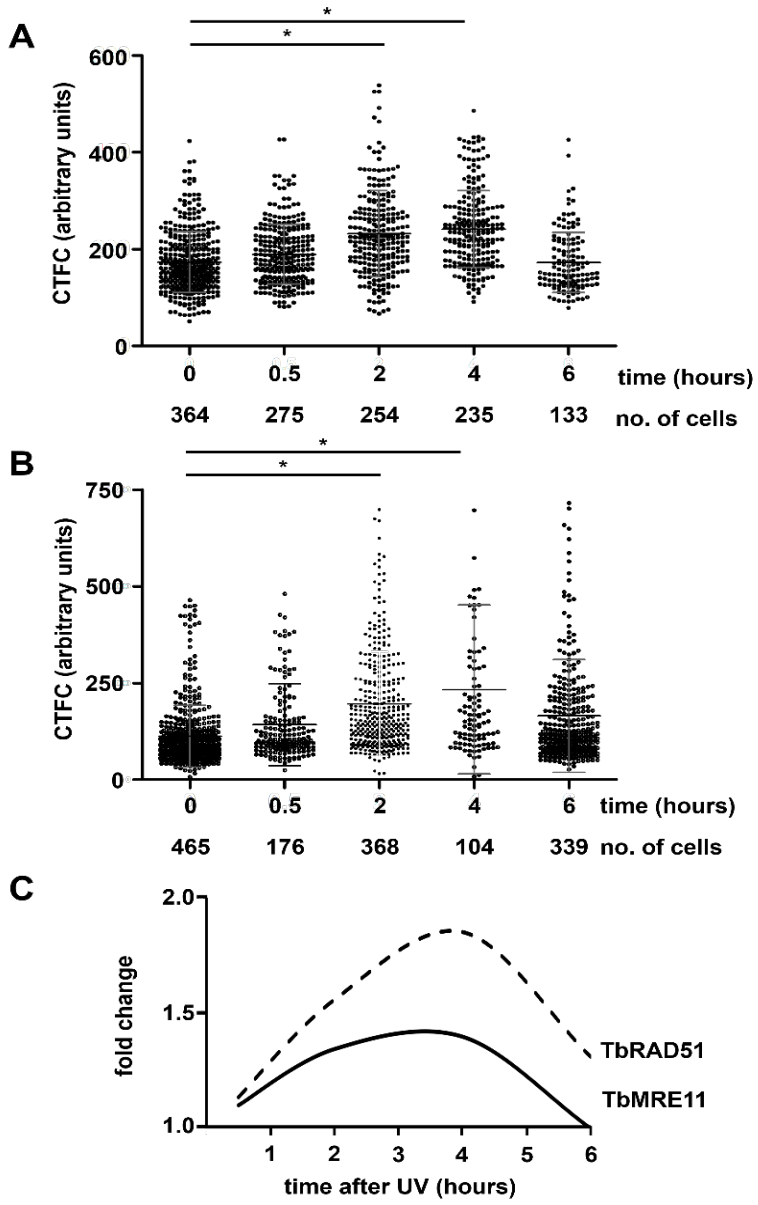


Fig. 3.25. Effect of UV on TbMRE11 & TbRAD51 expression.

A & B. The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen tagged TbMRE11 (**A**) or TbRAD51 (**B**) was determined at time intervals following exposure to UV (1500 J m^{-3}). Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).

C. Schematic showing the temporal changes in expression of TbMRE11 or TbRAD51 following exposure of cultures to UV light.

3.10 Effect of ICL damage on temporal expression of the trypanosomal HR pathway.

Interstrand crosslinks (ICLs) represent a highly toxic form of DNA lesion that blocks essential cellular processes such as DNA replication and transcription. To resolve this type of damage all cells, including trypanosomes, express several cell cycle dependent DNA repair mechanisms with

components of the HR pathway playing a key role in DNA replication-dependent pathways (further details given in Chapter 4) (Dattani & Wilkinson, 2019; Lehoczky et al., 2007; McVey, 2010;) Using BSF *T. brucei* lines engineered that constitutively express mNeonGreen tagged TbMRE11 and TbRAD51, the effect of mechlorethamine, the archetypal ICL inducing agent that displays trypanocidal activity (Dattani & Wilkinson, 2019; Sullivan et al., 2015) on the localisation and temporal expression of both HR factors was assessed.

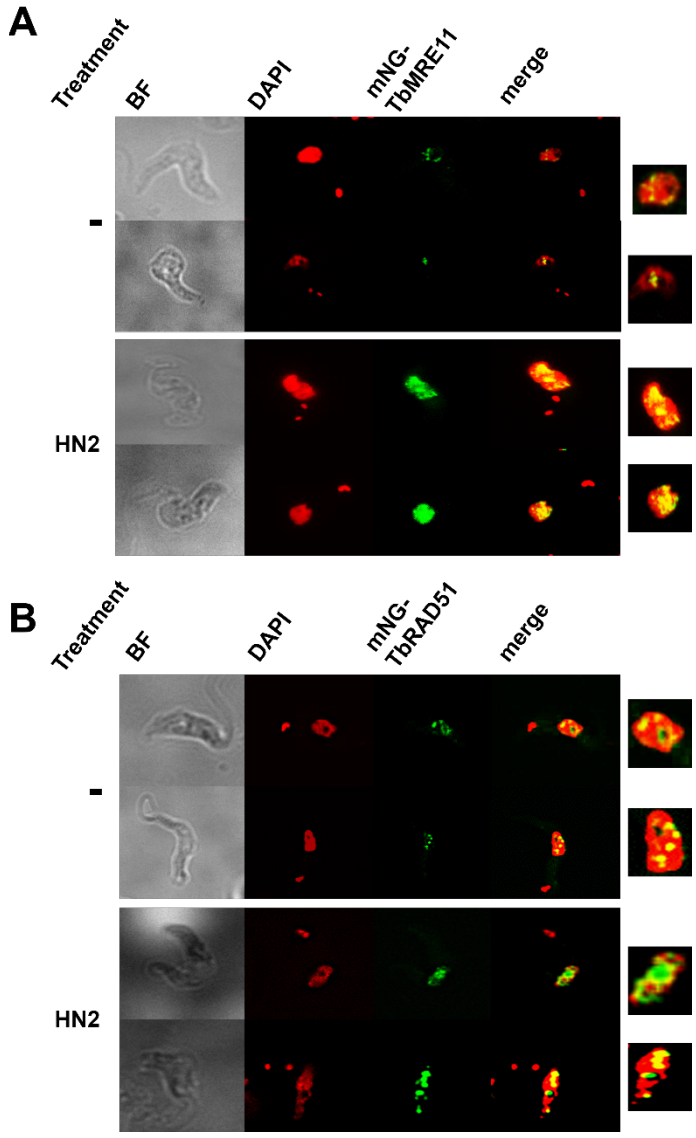


Fig. 3.26. Effect of mechlorethamine on TbMRE11 & TbRAD51 expression.
A & B. Mechlorethamine treated (HN2) parasites expressing TbMRE11 (**A**) or TbRAD51 (**B**) tagged at their amino terminal with mNeonGreen (mNG-TbMRE11 & mNG-TbRAD51, respectively; green) were co-stained with DNA stain, DAPI (red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of co-localization (merge; yellow), including a close-up image of the nucleus are shown. Untreated (-) parasites were examined in parallel.

When examining BSF mNeonGreen-TbMRE11 or mNeonGreen-TbRAD51 expressing *T. brucei*, the mechlorethamine treated parasites presented the same multiple foci pattern within their nucleoplasm as seen in controls (Figure 3.26A & B; section 3.6). In most cases, the fluorescence

centres observed in treated cells appeared slightly more intense in samples taken from the later time points (from 2- or 3- hours post treatment for TbMRE11 and TbRAD51, respectively).

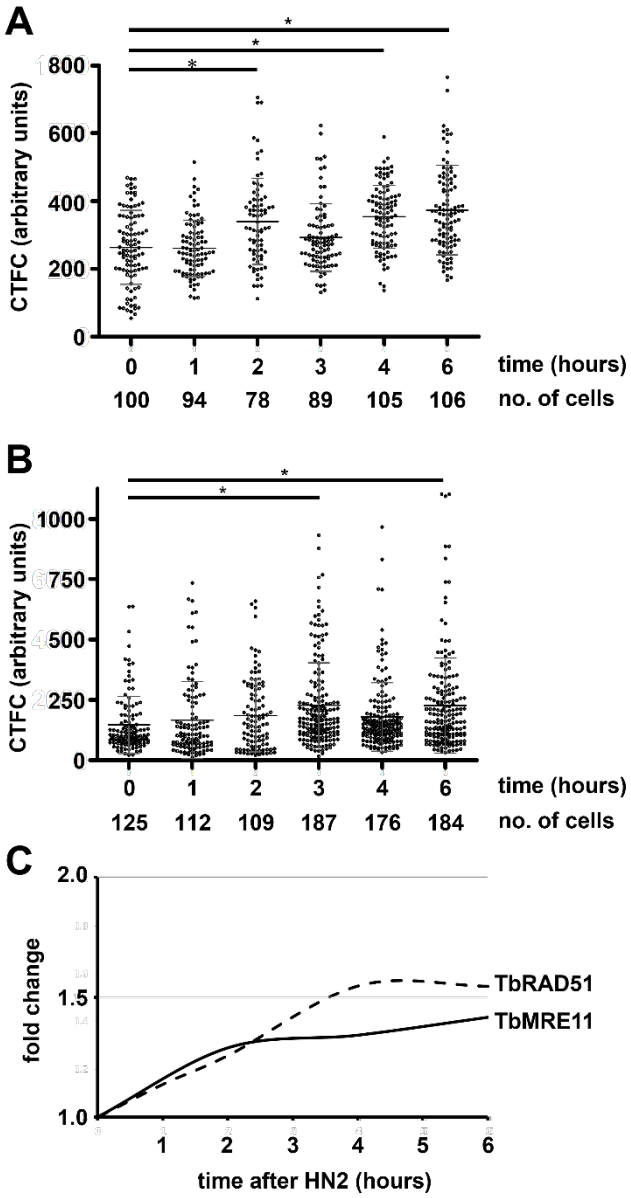


Fig. 3.27. Effect of mechlorethamine on TbMRE11 & TbRAD51 expression.
A & B. The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen tagged TbMRE11 (**A**) or TbRAD51 (**B**) was determined at time intervals following exposure to mechlorethamine (30 μM; HN2)). Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).
C. Schematic showing the temporal changes in expression of TbMRE11 or TbRAD51 following treatment with mechlorethamine.

When analysing whether mechlorethamine caused any temporal changes to TbMRE11 or TbRAD51 expression, a slight but non-significant increase in fluorescence was noted over the first 1 hour of treatment. For TbMRE11 the signal intensity gradually rose becoming significant at 2-hours posttreatment and remaining at this level at all other timepoints (Figure 3.26A & C). In the case of TbRAD51, a steady increase in fluorescence was observed that became significant at 3-hours posttreatment before reaching a maximum 4-hours posttreatment and remaining at this level

at 6-hours posttreatment (Figure 3.27B & C). The maximum increase in fluorescence signals induced by mechlorethamine was around 1.4- and 1.6-fold greater in treated mNeonGreen-TbMRE11 or mNeonGreen-TbRAD51 cells, respectively relative to controls. The fold difference noted for the ICL inducing agent is much lower than that determined for nitroimidazoles, nitrofurans or UV. This could reflect that a fraction (25 %) of trypanosomes within a population are replicating their nuclear genome at any given time (da Silva et al 2017; da Silva et al 2019; Vieira-da-Rocha et al., 2019), a point when the activity of the two HR factors is most pronounced in ICL repair.

3.11. Chapter Summary

In this chapter, we investigated the role *T. brucei* HR pathway plays in repairing DNA damage induced by trypanocidal nitroheterocyclic prodrugs. We demonstrated the mode of activation of fexinidazole by TbNTR1, using parasite lines engineered to express altered levels of MRE11, CSB and/or TbNTR1. From these results, we postulate that CSB may be involved in transcription-coupled HR. Using cell lines tagged to express mNeonGreen MRE11 and RAD51, we demonstrated the temporal expression of TbMRE11 and TbRAD51 in response to trypanocidal nitroheterocyclic prodrugs such as benznidazole, fexinidazole and nifurtimox.

Here we showed that:

1. Trypanocidal activity of fexinidazole occurs via activation by parasite specific TbNTR1
2. Fexinidazole damage mobilises TbCSB and TbMRE11 with both proteins involved in an epistatic relationship in response to fexinidazole damage.
3. Temporal recruitment of TbMRE11 and TbRAD51 occurs with benznidazole and fexinidazole repair. For benznidazole this involved γH2A mobilization and interaction of all 3 factors in response to possible DSB induced by this drug.
4. *T. brucei* HR pathway also detects DNA damage caused by Nifurtimox, UV and mechlorethamine but did not seem to be the main pathway involved in resolving these types of damage. Both TbMRE11 and RAD51 remained elevated throughout the exposure period.

CHAPTER 4: Evaluation of Factors Specific to ICL DNA Repair

4.1. DNA interstrand cross links and their repair

DNA interstrand crosslinks (ICLs) are cytotoxic lesions formed when two nucleotides on complementary strands in the DNA double helix become covalently linked (Lawley & Phillips, 1996). They are particularly dangerous form of damage as they can block essential cellular processes such as DNA replication and transcription where separation of nucleic acid strands is required and if not resolved, their accumulation in a genome can lead to chromosomal breakage, rearrangements, and eventually cell death (Dronkert and Kanaar, 2001; McHugh et al., 2001; Deans and West, 2011; Sengerova et al., 2011). The toxicity of these lesions is underlined as a single ICL can kill a repair-deficient microbe while as few as 40 can kill repair-deficient mammalian cells (Dronkert and Kanaar, 2001).

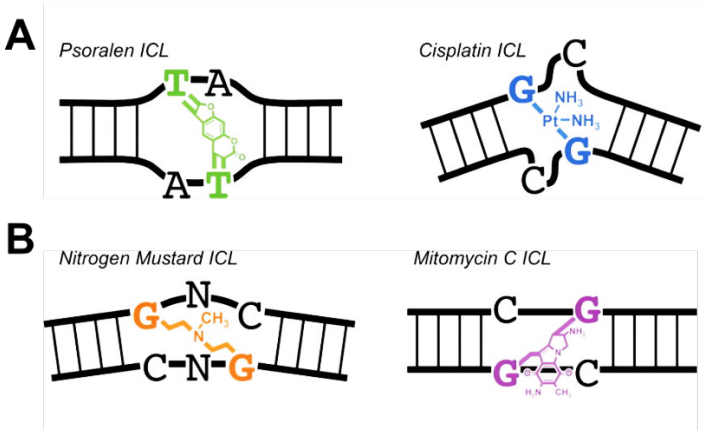


Fig 4.1. Types of ICL lesions. Distorting (A) and non-distorting (B) lesions caused by major DNA crosslinking agents. ICLs caused by cisplatin or psoralen result in major or moderate DNA distortion, respectively. In contrast, ICLs caused by nitrogen mustards (e.g., mechlorethamine) and mitomycin C lead to minor distortion to the DNA template. Image taken from (Rycenga & Long, 2018)

A variety of treatments can promote ICLs including environmental mutagens (e.g. UV light, ionizing radiation and furocoumarins present in plants and cosmetics), anti-cancer chemotherapies (e.g. mitomycin C, cisplatin, mechlorethamine, nitrosourea, and their derivatives) (Scott et al., 1976; Noll et al., 2006; Jones & Huang, 2012; Lopez-Martinez et al., 2016) and endogenous metabolic processes (e.g. nitrous acid, acetaldehyde and malondialdehyde) (Basu et al., 1984; Edfeldt et al., 2004; Kellum et al., 2021; Stone et al., 2008). These have different affinities towards nucleic acid templates with most functioning as bifunctional alkylating agents, containing two reactive groups that bind to different sites within different nucleotides to form a covalent bond (Figure 4.1) (Räschle et al., 2008). This interaction can then affect the sugar-phosphate backbone in diverse ways. As such, some ICL inducing agents such as cisplatin, psoralen, and UV can cause

significant distortions to the DNA double helix while others including mechlorethamine and mitomycin C have little or no effect (Figure 4.1) (Räschle et al., 2008). Below, certain features associated with selected ICL-inducing mutagens are listed below.

Psoralen and its derivatives are furanocoumarin-based natural compounds found in plants such as parsley and bergamot. They function as photoreactive prodrugs that following activation by irradiation with long wavelength UV light (e.g., UV-A) form products that preferentially bind thymidine bases at 5'-TA-3′ and 5′-AT-3′ sequences causing localised approximately 25° distortion to the DNA helix (Derheimer et al., 2008; Haran & Crothers, 1988; Wood, 2010). In medicine, they are used in PUVA (psoralen + UVA light) therapies that target a range of skin conditions such as psoriasis, eczema, and vitiligo (Vangipuram & Feldman, 2015).

Nitrogen mustard is a term given to a family of chemicals that contain at least two *N, N*-bis-(2-chlorethyl) amine functional groups with mechlorethamine being the prototype compound. In the presence of DNA, the chlorine atom in one of the reactive centres becomes displaced to form an aziridinium ion that then attacks a guanine nucleotide at its N7 position to form an N7-alkylated guanine derivative. This triggers the activation of a second reactive centre and at 5'-GNC-3' or 5'-CNG-3' sequences, mediates attachment to a second guanine nucleotide on the opposing DNA strand to form the ICL (Bauer & Povirk, 1997; Rink & Hopkins, 1995; Noll et al., 2006). For mechlorethamine this results in a slight (approximately 14°) distortion to the DNA helix (Rink and Hopkins, 1995). In medicine, nitrogen mustards are used in combinational therapies to treat a range of cancers. For example, mechlorethamine has been used to treat Hodgkin's disease, lymphosarcoma, chronic myelocytic leukaemia, polycythaemia vera, and bronchogenic carcinoma, melphalan is used against multiple myeloma, ovarian cancer, primary amyloidosis, and malignant melanoma while cyclophosphamide is used in the treatment of lymphomas, some forms of brain cancer, neuroblastoma, leukaemia, and some solid tumours (Stork & Schreffler, 2022)

The aziridine-based family of chemicals are classed together as they possess a three-membered heterocycle $(CH_2)_2NH$ in their structure. For many compounds, the ring can be found as a 'free' side chain (e.g., triethylenemelamine, ThioTEPA, CB1954, RH-1) with the heterocycle often requiring some form of activation to generate a reactive center. In many cases this activation event takes place at sites elsewhere in the structure such as through the reduction of carbonyl or nitro groupings that cause a retribution of electrons in the chemical backbone and as with nitrogen

mustards results in a reactive aziridinium ion (Bauer & Povirk, 1997; Rink & Hopkins, 1995; Noll et al., 2006).

This then targets a nucleotide, usually guanine at its N7 position to form an alkylated intermediate (Noll et al., 2006). In compounds that have two (or more) aziridinyl motifs, formation of the intermediate results in activation of the additional grouping(s). At appropriate sites within a genome, the second reactive aziridinium ion centre interacts with and covalently binds to a nucleotide, usually guanine, on the opposing DNA strand to form the ICL (Noll et al., 2006). In some aziridine-containing compounds such as mitomycin c, the heterocycle can be found 'trapped' within the chemical backbone. Again, enzymatic activation is required to promote functionality of the reactive centre which for mitomycin c involves the two-electron reduction of carbonyl motifs found on its quinone ring, a reaction catalysed by NADPH-cytochrome P-450 reductase and xanthine oxidase, leading to formation of a hydroquinone intermediate (Iyer & Szybalski, 1963; Pan et al., 1984; Noll et al., 2006). Tautomerization followed by reaction with the N2-amino group of guanine produces an alkylated monoadduct. Elimination of a carbamoyl group results in a highly reactive vinylogous hydroquinone methide intermediate that alkylates a guanine on the opposite strand of DNA at 5'-GC-3' or 5'-CG-3′ sequences to produce, after oxidation, the ICL (Rajski and Williams, 1998; Noll et al., 2006).

Generally, the distortion caused by an aziridine-induced ICL to the DNA helix is modest (Iyer and & Szybalski, 1963; Pan et al., 1984). In relation to their medical use, mitomycin c is employed to treat gastro-intestinal, anal and breast cancers while ThioTEPA is used in palliation of many neoplastic diseases including in the treatment of adenocarcinoma of the breast and ovary, and in papillary thyroid and bladder cancer.

Cisplatin (cis-diamminedichloroplatinum (II)) is a platinum-based, planar compound. Its activation results in displacement of one or both of its chloride groups, allowing it to interact with nucleic acids to form DNA or RNA monoadducts, intrastrand DNA crosslinks, ICLs, and DNA-protein crosslinks (Dasari & Tchounwou, 2014). In ICL formation, which accounts for 5 to 8% of the above interactions, the activated cisplatin targets the N7 position of guanine at 5'-GC-3' or 5'-CG-3' sequences before interacting with the opposing strand. The resultant short G-G interaction causes cytosine to flip out of the helix leading to severe distortion (Roberts & Friedlos, 1987; Jamieson & Lippard, 1999; Jamieson et al., 2006). Cisplatin is used to treat various cancers

including sarcomas, some carcinomas, lymphomas, bladder cancer, cervical cancer, and germ cell tumours, and is particularly effective against testicular cancer (Noll et al., 2006).

The type of ICL generated influences the repair pathway deployed by a cell to resolve the lesion (Beljanski et al., 2004; Smeaton et al., 2008). Distorting ICLs can be detected and repaired throughout the cell cycle while non-helix distorting lesions are often encountered during the S phase when the replisome stalls upon collision with the lesion (Smeaton et al., 2008; Zhang et al., 2015; Nepal et al., 2017). As such all cells possess DNA replication-independent or dependent ICL repair pathways with these mechanisms involving the cooperation of components from the 'classical' DNA repair pathways (e.g., from the nucleotide excision repair (NER), base excision repair (BER), mismatch repair (MMR), homologous recombination repair (HR) and translesion synthesis (TLS) pathways) alongside several dedicated enzymes (Muniandy et al., 2010). The situation is further complicated as evolutionary diverse organisms often utilise different components and/or systems at contrasting times in the cell cycle in order to resolve ICLs (Dong et al., 2015; McVey, 2010).

In most of the replication-independent mechanisms, NER factors play a key role in lesion detection. At transcriptionally silent regions of a genome and for helix distorting mutagens, recognition of the ICL-induced DNA confirmational change is mediated by components of the global genome nucleotide excision repair (GG-NER) pathway including XPC and HR23B (Figure 4.2) (Thoma et al., 2005; Zhao et al 2009). In contrast, at transcriptionally active sites, and usually for agents that form ICLs which do not significantly distort the DNA helix, the lesion causes stalling of an advancing RNA polymerase complex. This results in alteration to the chromatin structure, a signal that promotes deposition of CSA and CSB from the transcription-coupled nucleotide excision repair (TC-NER) pathway (Figure 4.2B) (Zheng et al., 2003; Iyama et al., 2015; Hashimoto et al., 2016; Iyama and Wilson, 2016). Next, both ICL detection mechanisms facilitate recruitment of other DNA repair factors to the lesion site. Initially, XPF-ERCC1 and XPG from the NER incision complex mediate the nucleolytic processing of sites 5' and 3' to the ICL resulting in 'unhooking' of the lesion from one strand of the double helix (Wood, 2010; Abdullah et al., 2017; Faridounnia et al., 2018) with DNA sequence associated with the 'unhooked' region then degraded by SNM1A up to and across where the covalent bond is found (Wang et al., 2011; Sengerova et al 2012). The resultant single strand DNA gap and the DNA backbone reformed via TLS polymerases (Pol ζ, Pol η, Pol ι, Pol Θ or Pol κ) and *DNA ligase*

activities *(*Sarkar et al 2006; Ho & Schärer, 2010; Ho et al., 2011; Sharma and Cannan, 2012*)*. The Translesion synthesis (TLS) machinery allow the repair of the DNA gap opposite the ICL by bypassing the ICL. It involves the synthesis of DNA opposite a bulky adduct using the damaged DNA strand containing the ICL as template. This type of repair helps cells avoid the chromosome damage potentially leading to double strand break or prolonged replication fork stalling (Vinh Ho., 2011). Complete removal of the ICL from the genome is achieved through a second round of NER incision, DNA polymerase and DNA ligase activities.

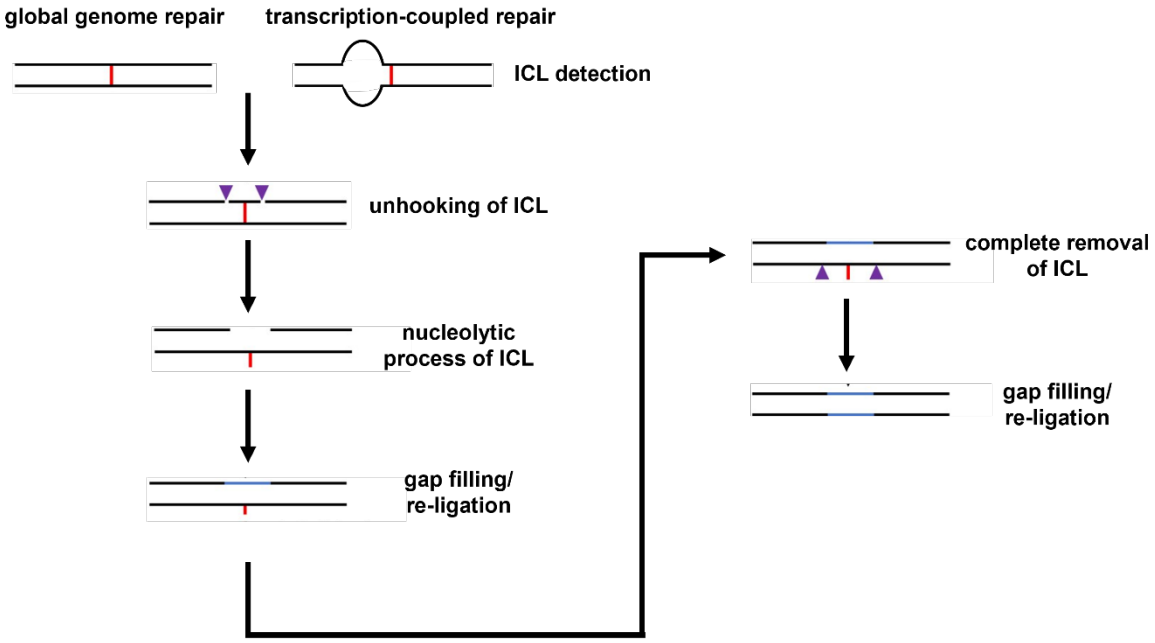


Fig 4.2. Replication-independent ICL repair pathway
Cells can detect ICLs (red line) throughout the cell cycle using components from the global genome (XPC; HHR23B) or transcription-coupled (CSA; CSB) NER pathways. Both detection mechanisms result in the nucleolytic processing (XPF-ERCC1; XPG) at sites flanking the ICL (purple triangles) causing the 'unhooking' of the crosslink from the DNA. The 'unhooked' sequences are then degraded (SNM1A/PSO2) with the gap on the other DNA strand filled and DNA backbone reformed by TLS polymerase (Pol ζ, Pol η, Pol ι, Pol Θ or Pol κ) and DNA ligase activities (newly synthesised DNA is blue line). The ICL is completely removed from the genome by a second round of nucleolytic processing (XPF-ERCC1; XPG)/ TLS polymerase (Pol ζ, Pol η, Pol ι, Pol Θ or Pol κ) and DNA ligase activities. Image adapted from (Huang and Li, 2013).

During the S phase of the cell cycle, most eukaryotic cells remove ICLs from the nuclear genome using the so-called Fanconi Anaemia (FA) pathway (Figure 4.3A). In cases involving stalling of a single replisome, changes in chromatin structure and in DNA configuration result in the recruitment of FANCM to the lesion site. In a process aided by ancillary factors such as FAAP24, MHF1 and MHF2, this guides assembly of FANCA, -B, -C, -E, -F, -G (XRCC9) and -

L for form the FA core complex (Moldovan and D'Andrea, 2009; Tan and Deans, 2017). Once bound to DNA, this complex acts as an E3 ubiquitin ligase, mediating the monoubiquitination of lysine residues in the FANCD2 component of the FANCD2-FANCI heterodimer (also known as the FA recruitment complex) resulting in its activation (Kim and D'Andrea, 2012; Liang et al., 2016). Next, the nuclease scaffold protein FANCP (SLX4) is recruited to ubiquitinated FANCD2 through its UBZ4 ubiquitin-binding zinc finger domain (Yamamoto et al., 2011) with this guiding structure-specific nucleases such as FAN1, XPF-ERCC1 (NER pathway), MUS81-EME1 and SLX1 to the lesion site to catalyse nucleolytic incision of the DNA backbone at sites 5' and 3' to the ICL resulting in 'unhooking' of the lesion from one strand of the double helix (Moldovan and D'Andrea, 2009; Zhang and Walter, 2014; Hashimoto et al., 2016; Siddiqui et al., 2017). The gap created in the remaining strand is then filled by REV1 or Pol ζ TLS polymerases with DNA ligases reforming the sugar-phosphate backbone (Räschle et al., 2008). A second round of nucleolytic processing/TLS polymerase/ligase activities then takes place completely removing the 'unhooked' ICL from the genome, resulting in formation of a repaired, intact double stranded DNA as well as introducing DSB in the newly synthesized DNA strands generated by the replisome. This newly formed secondary lesion is then repaired by the activity of components from the HR pathway (FANCD1 (BRCA2), FANCN (PALB2), FANCO (RAD51C), FANCR (RAD51), FANCS (BRCA1), FANCU (XRCC2)) that mediate recombination between the broken dsDNA molecule and the newly repaired DNA structure generated from the second round of nucleolytic processing/TLS polymerase/ligase activities. Upon repair, a Y-shaped forked structure is produced that can then function as a template for the DNA replication complex (Wang, 2007). In cases where an ICL causes the stalling of two converging replisomes (Chaudhury et al., 2014; Zhang and Walter, 2014), an X shaped structure can be generated (Figure 4.3B). Here, the FA mechanism removes the lesion from the genome but in doing so introduces two DSBs in the two newly synthesized DNA strands. These secondary lesions are resolved via HR activities using the newly repaired DNA strand as template resulting in formation of two complete dsDNA molecules and disassociation of the replisome.

Many single celled lower eukaryotes lack some of the orthologues that constitute the mammalian FA pathway (McHugh et al 2012; Daee and Myung, 2012; Dattani and Wilkinson, 2019). In such circumstances alternative replication dependent ICL repair systems operate. In *Saccharomyces cerevisiae* the ICL-induced stalling of a replisome promotes MPH1 (FANC),

MUTSα and MGM101 deposition to the lesion site resulting in regression of the replication fork (Figure 4.4) (Daee et al., 2012; Ward et al., 2012). Guided by the nuclease scaffold protein SLX4 (FANCP), NER nucleases mediate the 'unhooking' of the crosslink from one DNA strand with EXO1 removing nucleotides up to but not beyond the ICL (Muñoz et al., 2009; Ward et al., 2012). The gap formed from this process is then filled by damage-tolerant TLS polymerases and the DNA backbone reformed via DNA ligase activities. The 'unhooked' crosslink is completely removed by a second round of NER activity with the dsDNA restored by components of the single strand annealing HR pathway involving the MGM101 recruitment of RAD52 to the site of DNA with ancillary factors such as CHL1 (FANCJ) aiding this event (Daee et al., 2012; Ward et al., 2012).

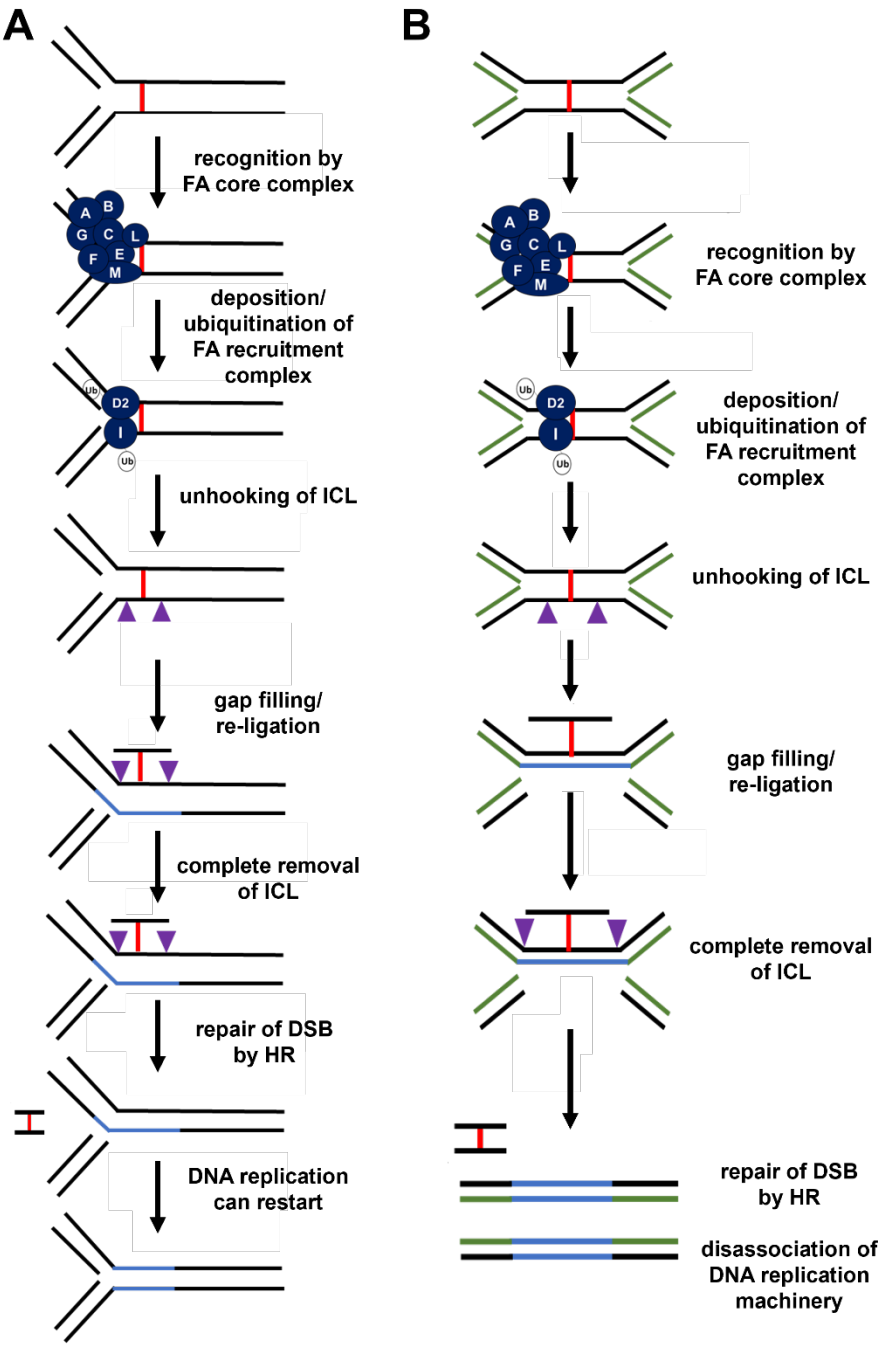


Fig 4.3 Replication-dependent ICL repair in higher eukaryotes.
A & B. During the S-phase of the cell cycle, FA core and recruitment complexes (blue spheres) recognise crosslinked DNA (red line) through stalling of a single replication fork (**A**) or convergence of two replication forks (**B**) resulting in the recruitment of FA effector proteins. The effectors (purple triangles) 'unhook' the ICL with TLS DNA polymerases/ligase filling in any gaps and reforming the DNA backbone: newly synthesised DNAs are shown as a blue line). This process generates DSB(s) that are subsequently repaired by the HR pathway. Image adapted from (Huang and Li, 2013; Krishnan et al., 2015).

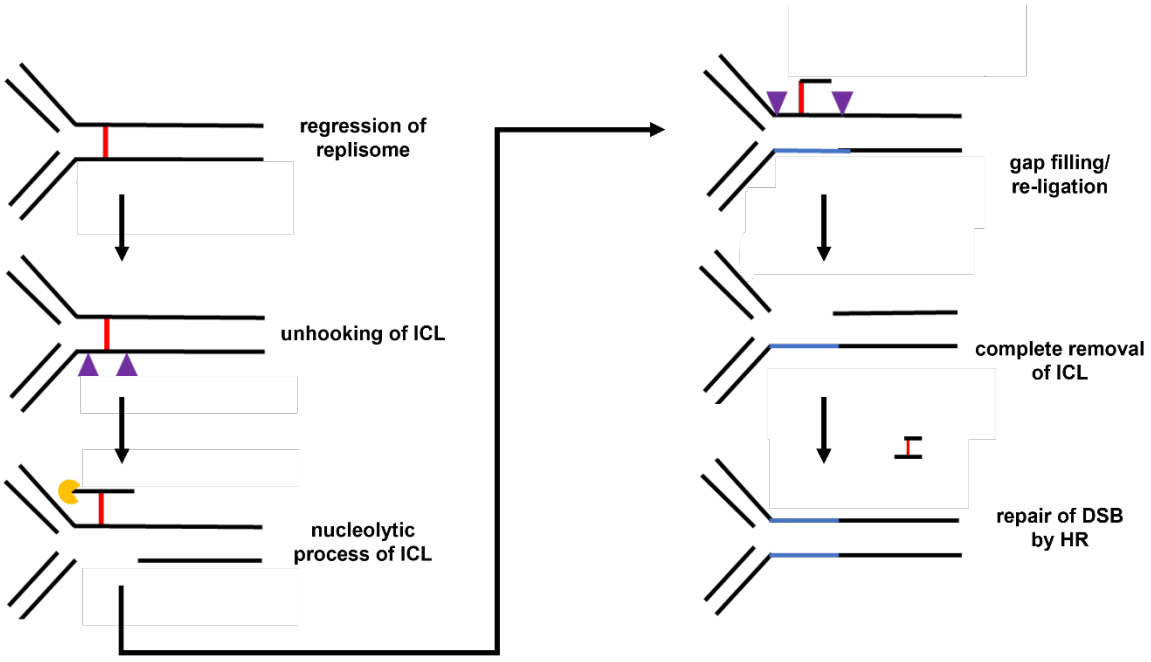


Fig 4.4 Replication-dependent ICL repair in lower eukaryotes.
In *Saccharomyces cerevisiae*, the ICL-induced stalling of a replisome results in recruitment of factors that cause regression of the DNA replication machinery. Next, NER factors (purple triangles) 'unhook' the lesion from one of the DNA strands with the released sequence degraded (orange sphere) up to but not beyond the lesion. The gap in the other DNA strand is then filled by TLS polymerases and the DNA sugar phosphate backbone reformed of DNA ligase. The ICL is completely removed from the genome by a second round of nucleolytic processing with the newly synthesised DNA (blue line) used as template by the HR pathway to repair DSBs. Image adapted from (Barber et al., 2005).

Bloodstream form *T. brucei* have been shown to be highly susceptible to ICL inducing agents with the most potent including a nitrogen mustard or aziridinyl grouping (Hall et al 2010; Bot et al 2010; Sullivan et al 2015). Analysis of the system(s) expressed in this organism to repair such lesion indicates this trypanosome form possesses several effector proteins from the FA pathway e.g., FANCD1 (BRCA2), FANCJ (CHL1), FANCO (RAD51C), FANCQ (XPF), FANCR (RAD51) and FANCV (MAD2) but lacks all components of the FA core and recruitment complexes (Genois et al., 2014; Dattani and Wilkinson, 2019). Instead, and primarily through the phenotypic characterisation of null mutant lines, it has been shown that bloodstream form *T. brucei* expresses at least three different ICL repair mechanisms reminiscent of the *S. cerevisiae* replication-dependent and the two replication-independent pathways (Figure 4.5) (Dattani and Wilkinson, 2019). In the two replication-independent systems (Figure 4.5A), ICL recognition is mediated by the global genome surveillance complex, involving TbXPC and TbDDB from the GG-NER pathway (Machado et al., 2014), or by the stalling of transcription complex, represented by TbCSB from TC-NER (Dattani and Wilkinson, 2019). These interactions are proposed to

trigger the NER-mediated excision of the DNA backbone, possibly requiring TbXPF-TbERCC1 and TbXPG, creating an "unhooked" crosslinked oligonucleotide. This 'unhooked sequence is then degraded, a process catalysed by TbSNM1 and/or TbEXO1 (Sullivan et al 2015; Kumar, 2018; Dattani and Wilkinson, 2019), leaving a single nucleotide linked to the complementary DNA strand by a covalent bond plus a ssDNA gap. TbRPA (Replication Protein A) is proposed to load to the ssDNA, leading to the initial recruitment of TbPCNA (Proliferating Cell Nuclear Antigen) and then TLS polymerases. The latter activities mediate nucleotide synthesis to fill the gap, with the DNA's sugar-phosphate backbone reformed by DNA ligase. The ICL is completely removed from the *T. brucei* genome by a second round of incision mediated by NER/ TLS DNA polymerase/DNA ligase activities.

In the replication-dependent mechanism (Figure 4.5B), the ICL causes the collapse of two converging replication forks. The resultant DNA structure is recognized by an unidentified mechanism that guide incision of the DNA backbone at sites flanking the lesion, possible by TbXPF-TbERCC1, TbMUS18 or TbFAN1 endonucleases (Kumar, 2018), leading to 'unhooking' of the ICL, formation of a ssDNA gap on the strand opposite the 'unhooked' sequence and generation of DSBs, as detected by the TbMRE11-mediated formation of γH2A (Dattani and Wilkinson, 2019). The DNA sequence associated with the 'unhooked' region may then be degraded, a process that occurs independently of TbSNM1 and possibly involves TbEXO1 (Kumar, 2018; Dattani and Wilkinson, 2019), with the ssDNA gap filled and DNA structure reformed through the concerted actions of TLS polymerases and DNA ligase. A second round of NER incision/TLS polymerase/DNA ligase activities result in complete removal of the ICL from the genome. The DNA molecules that contain the DSBs are then repaired by the HR pathway, where TbRAD51 and TbBRCA2 play roles, using the newly repaired dsDNA as template (Dattani and Wilkinson, 2019).

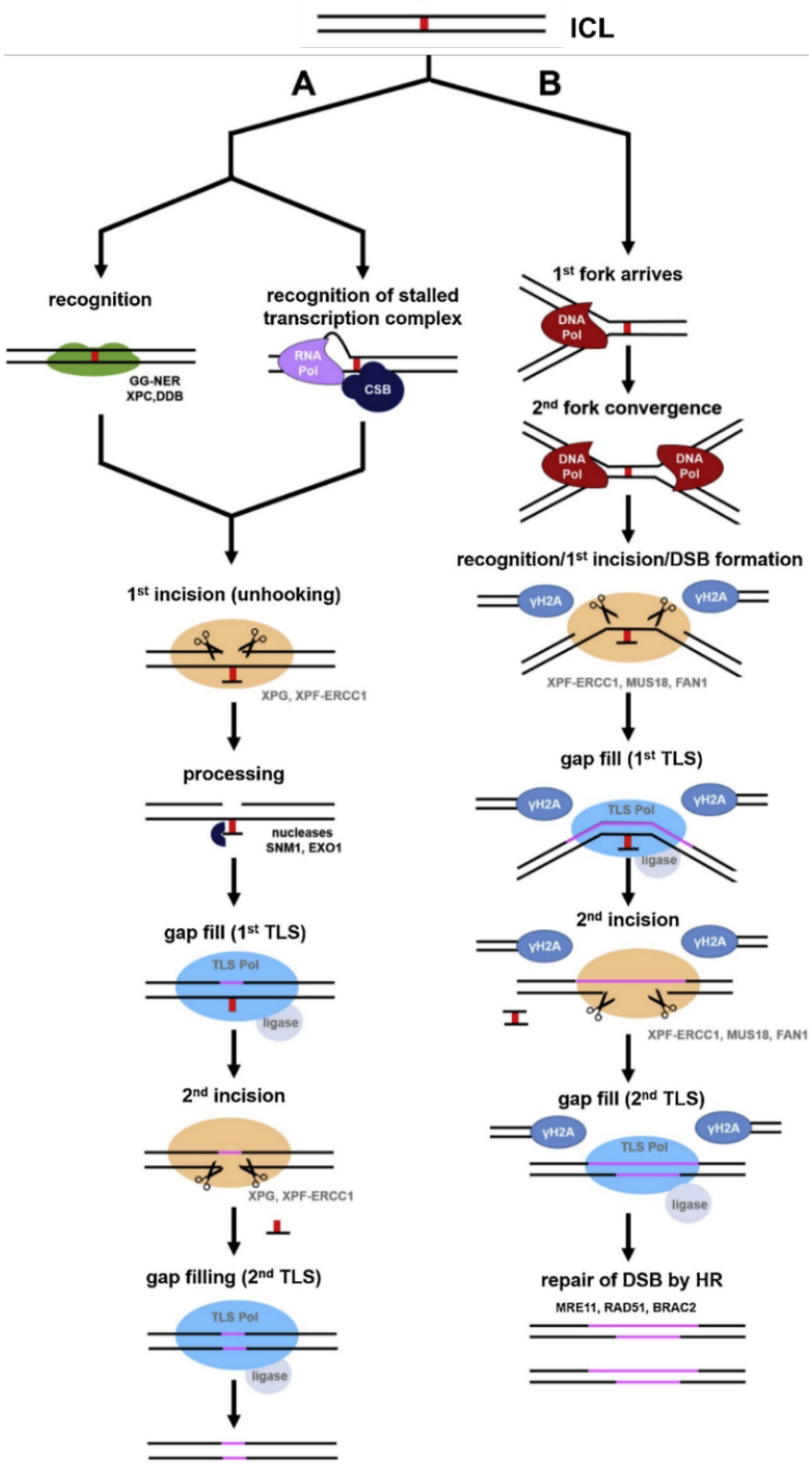


Fig 4.5. ***T. brucei*** **ICL repair.**
T. brucei utilizes two replication-independent **(A)** and replication-dependent **(B)** pathways to resolve DNA interstrand crosslink damage. Details of these proposed mechanisms is given in the text. Figure taken from Dattani & Wilkinson 2019.

In this chapter, we build on the above to further study the role of TbSNM1 and TbFAN1 in the *T. brucei* ICL repair system. Using a PCR-based CRISPR/Cas9 genome editing strategy (Beneke et al 2017), bloodstream and procyclic parasites expressing tagged versions of either enzyme, were made with this approach used to construct procyclic form *T. brucei* null for Tb*snm1*. The resultant lines were used to assess the subcellular location of each enzyme in both parasite stages, evaluate how protein expression is affected by ICL inducing agents (mechlorethamine) and trypanocidal nitroheterocyclic prodrugs and in the phenotyping (growth and susceptibility studies) of insect form parasites lacking TbSNM1.

4.2 Assessing the role of TbSNM1 in *T. brucei*.

4.2.1 Informatic analysis of TbSNM1

Literature searches (Sullivan et al., 2015) and BLAST analysis of *T. brucei* data held on TriTrypDB (Aslett et al., 2010) identified a single 2163 bp ORF (Tb927.4.1480; designated as Tb*snm1*) from the *brucei* TREU927 strain. This sequence is located on chromosome 4 of the parasite's nuclear genome and has potential to encode for a 79.5 kDa enzyme (designated as TbSNM1). Pairwise analysis revealed TbSNM1 was 43% identical to TcSNM1 from *T. cruzi* CL Brener Non-Esmeraldo-like haplotype (XP_816034), between 27 to 32% identical to orthologue expressed by *Leishmania* species (*L. major* SNM1, XP_001686430; *L. donovani* SNM1, XP_003864463) and 15 to 24% identity to its yeast (*S. cerevisiae* PSO2, NP_013857), mammalian (*H. sapiens* SNM1A, NP_001258745) and plant (*A. thaliana* SNM1, NP_189302) counterparts (Sullivan et al., 2015).

Searches of data embedded on TriTrypDB revealed that in keeping with other SNM1/PSO2 members, TbSNM1 contains a Lactamase_B_2 domain (PF12706) located towards its amino terminus (residues 37 to 234) and a DNA repair metallo-β-lactamase (DRMBL) domain (PF07522) found between residues 449 to 522 (Figure 4.6). Within the former domain a series of highly conserved histidine and aspartate residues, including a diagnostic HxHxDH motif, present in members of the SNM1/PSO2 family of nucleases are found in TbSNM1 with these functioning to co-ordinate metal (usually zinc) ion binding (Cattell et al., 2010).

Using the Phyre2 package, 3D predictions of the full-length protein based on multiple structures (Figure 4.7A) and a partial model based on the human SNM1A (4B87; Allerston et al., 2015) crystal structure (Figure 4.7D) were generated and compared (Figure 4.7E). This revealed that a

considerable part of the full-length sequence, including about 290 amino acids at the carboxyl terminal, had no homology to known structures (Figure 4.7B) with most of this region predicted to be disordered (Figure 4.7C). However, for both structures a common central core could be discerned. In keeping with other SNM1/PSO2 and related β-CASP proteins, this TbSNM1 core could be divided into two distinct structural regions (Mandel et al., 2006; Dorléans et al., 2011; Baddock et al., 2020; Allerston et al., 2015) with their relative positions mapped onto the partial 3D prediction (Figure 4.7E). The Lactamase_B_2 domain forms a characteristic MβL-fold (metallo-β-lactamase) composed of 11 β-sheets split across two regions and having a mixed arrangement. These two regions are sandwiched between two sets of α helices giving an αβ/βα organisation and representing a metallo-β-lactamase (MβL) -fold. The TbSNM1 DRMBL domain is inserted between the 10th and 11th β sheet of the MβL-fold.

When the residues involved in metal ion binding were incorporated onto the partial 3D prediction, these were shown to be the spatial positioned to facilitate binding of one or metal atoms. Here we postulate that residues H78, H80, H159 and D221 to be involved in binding one metal ion while D221 along with D82 and H83 may mediate binding to the second zinc atom

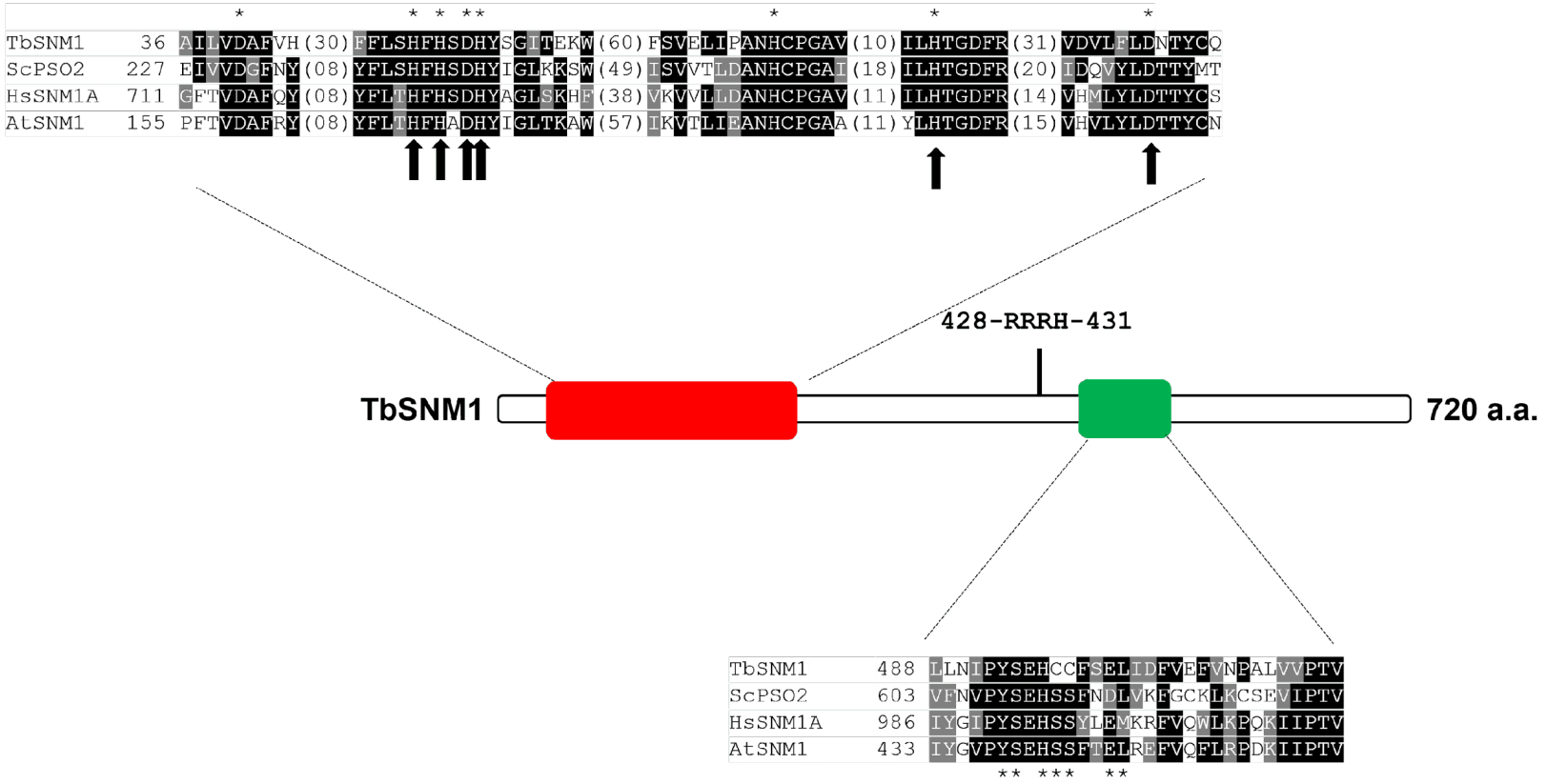


Fig 4.6. Informatic analysis of TbSNM1.
The sequence corresponding to the to β-lactamase superfamily (Lactamase_B_2; PF12706; residues 37-234; red box) and DNA repair metallo-β-lactamase (DRMBL; PF07522; residues 449-522; green box) of TbSNM1 (AAZ10739) was aligned with the equivalent regions from the orthologues expressed by Saccharomyces cerevisiae (ScPSO2; NP_013857), Homo sapiens (HsSNM1A; NP_001258745) and Arabidopsis thaliana (AtSNM1; NP_189302). Residues that are identical or are conserved substitutions are highlighted in black and grey, respectively. Amino acids marked with an asterisk correspond to residues that constitute the HMM logo of the corresponding domain. Amino acids marked with an arrow (H78, H80, D82, H83, H178, D221) correspond to residues that co-ordinates metal (zinc) cofactor binding. The location of potential 4 pattern nuclear localisation signal (RRRH) is shown.

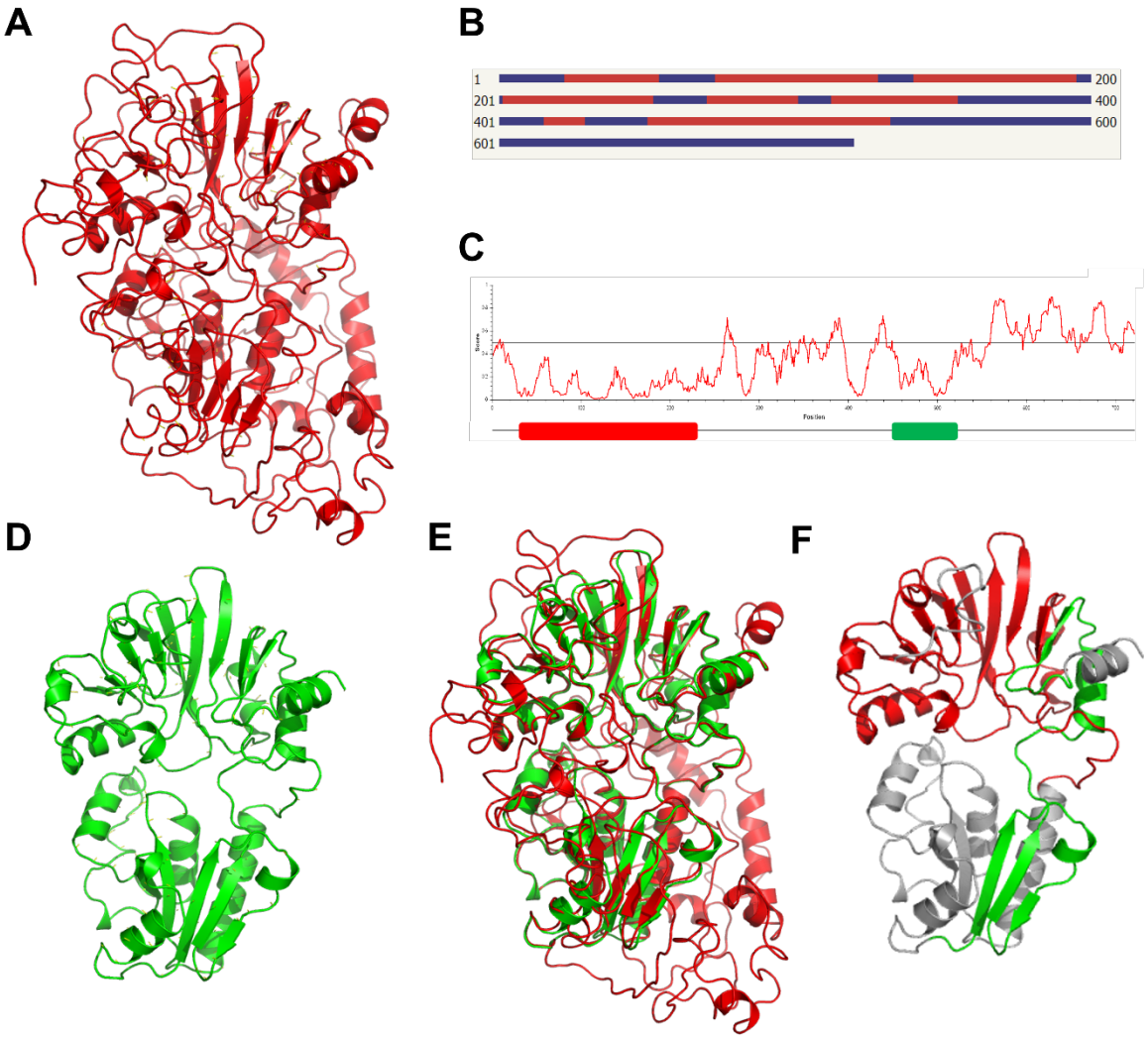


Fig 4.7. 3D prediction models of TbSNM1.

A & D. Full-length (**A**; red) or partial (**D**; green) 3D predictions of TbSNM1 (AAZ10739) were constructed using Phyre2 (Kelley et al. 2015) and visualized using PyMOL. The full-length approximation (**A**) was generated from two templates, 4B87, 3ZDK, with 360 residues (50%) modelled at >90% confidence. The model for 357 residues, including ~290 residues at the carboxyl terminal, are deemed highly unreliable. For the partial prediction (**D**), based on human SNM1A (4B87) crystal structure, there was 100% confidence over 33% of residues modelled.

B. Schematic showing the degree of confidence in the full-length model (**A**). Areas in red and blue correspond to regions where there is high or low confidence, respectively.

C. Regions of TbSNM1 identified as having an intrinsically disordered structure were obtained by submitting the protein sequence to the disorder prediction program (IUPRED2; https://iupred2a.elte.hu/). A disorder value >0.5 indicates an unstructured protein region. Based on this, the ~290 residues in the carboxyl terminal of TbSNM1 are predicted to form an intrinsically disordered region. The Lactamase_B_2 (PF12706; red box) and DRMBL (PF07522; green box) domains are highlighted.

E .Comparison of the full-length (red)and partial (green) 3D predictions of TbSNM1.

F. The position of the Lactamase_B_2 (PF12706; red) and DRMBL (PF07522; green) domains on the partial TbSNM1 model (**C**) are shown. Non-domain regions are noted in grey.

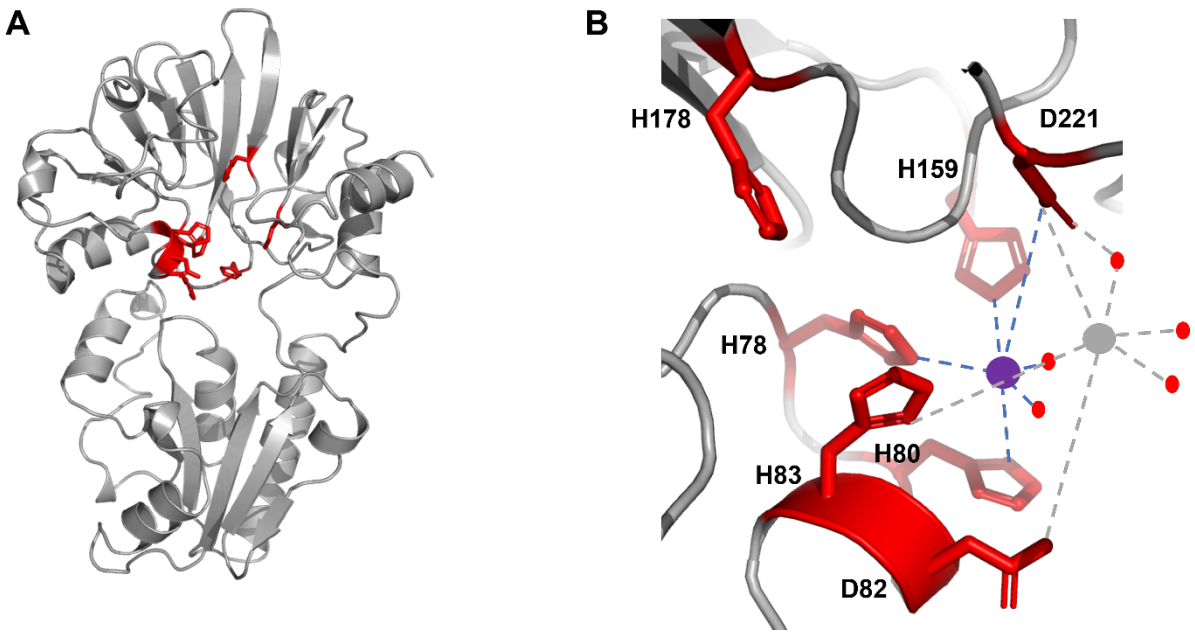


Fig 4.8. Putative zinc-binding site of TbSNM1.
A. Position of the amino acids (red) postulated to be involved in zinc-binding are shown on the or partial 3D predictions of TbSNM1. The residues highlighted correspond to those noted with an arrow in Figure 4.6.
B. A zoomed image of the amino acids (red) postulated to be involved in zinc-binding. The assignment of these residues involved in metal ion was made following comparison against the human SNM1A and SNM1B sequences (PDB: 5AHR and 5AHO, respectively (Allerston et al 2015). The residues H78, H80, H159 and D221 involved in binding (blue dotted line) to zinc ion 1 (purple sphere) and D82, H83 and D221 postulated to involved in binding (grey dotted line) to zinc ion 2 (grey sphere) are shown. The red spheres correspond to activated water/hydroxylate ions (Allerston et al 2015).

4.2.2 Editing the Tb*snm*1 loci in *T. brucei*

Based on localisation signal detection algorithms (PSORTII; WoLFPSORT), TbSNM1 is predicted to be found in the *T. brucei* nucleus with targeting is believed to involve a "4 pattern" RRRH nuclear localisation signal between residues 428 to 431 (Figure 4.7). Mutational studies performed on ectopically expressed, GFP tagged TbSNM1 confirmed that the two central residues (R429 & R430) play a key role in targeting this enzyme to the nucleus (Kumar, 2018). To confirm/assess TbSNM1 localisation throughout the cell cycle, the previously described CRISPR/Cas9 genome editing approach (see Figure 3.13) was employed to tag the endogenous loci at its 5' end with nucleotide sequences coding for the fluorescent protein mNEONGREEN (Beneke et al., 2017). Following the scheme set out in Chapter 3 and using appropriate primer combinations (Table 4.1), two amplicons that correspond to the sgRNA and targeting fragment

were electroporated into BSF or PCF *T. brucei* with recombinant, drug (hygromycin) resistant clones selected.

	Primer combinations	
	sgRNA	**Targeting fragment**
5' tagging	G00/5'sgRNA Tbsnm1	Tbsnm1-UFP1/Tbsnm1-DRF1

Table 4.1. Primer combinations used in CRISPR/Cas9 tagging of Tb*snm1*.
The primer combinations used to generate amplicons corresponding to the short guide RNA (sgRNA) and targeting fragment used in tagging the 5' end of the endogenous Tb*snm1* loci with a sequence coding for the fluorescent protein mNEONGREEN. The sequence of each primer can be found in Appendix 1.

To confirm expression of *mNeonGreen*-Tb*snm1*, DNA amplification reactions on cDNA templates derived from wild type and recombinant lines were performed using primer combinations that span the gene fusion event (Figure 4.9A). This generated a band of the expected size (0.6 kb) from templates originating from BSF recombinant parasites with no band(s) observed in templates produced from wild type cells (Figure 4.9B). Analysis of sequence from the cDNA-derived amplicon confirmed that the *mNeonGreen* and Tb*snm1* coding sequences were in-frame (data not shown). When this validatory approach was extended to cDNA templates derived from PCF form parasites, the same outcomes were obtained. As a further validation step, western blot analysis to detect mNeonGreen was performed using extracts generated from PCF parasites expressing *mNeonGreen*-Tb*snm1* (Figure 4.9C). Here, the anti-mNeonGreen antibody detected a band of the expected size (106 kDa) only in extracts generated from recombinant cells: no band(s) were observed in lysates from wild type cells.

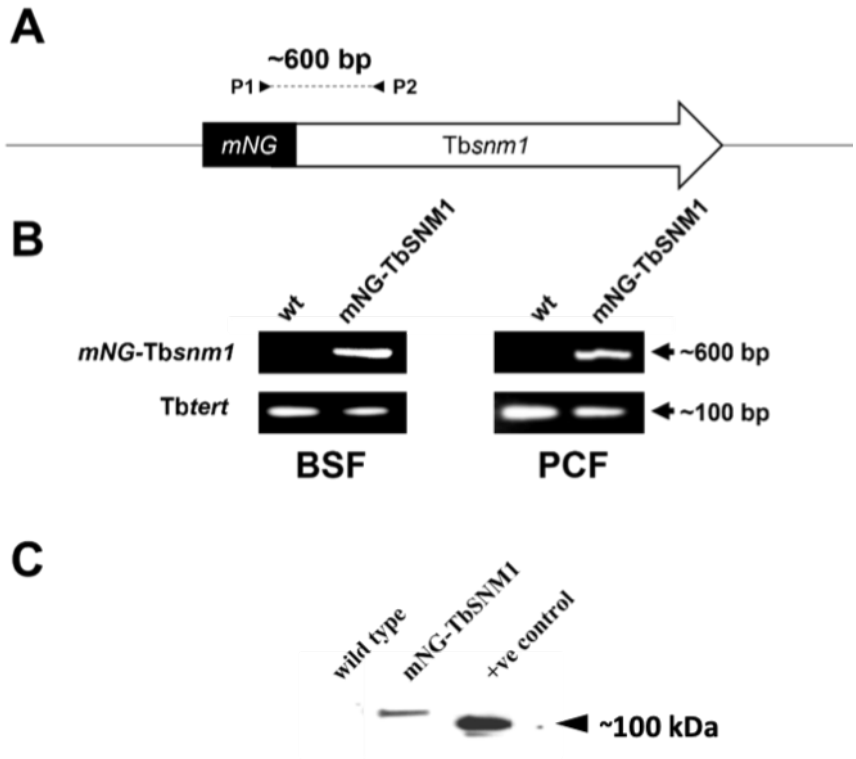


Fig 4.9. Validation of *T. brucei* lines expressing mNeonGreen tagged TbSNM1.
A. Schematic showing integration of *mNeonGreen* (*mNG*) into the *T. brucei* genome at the 5' region of Tb*snm1*. P1 (GGAGTTATACCACAGGAAACG) & P2 (CACCAGATCACCATTGCGGAG) indicate the primers used in the obtention of the validation amplicons.
B. Amplicons diagnostic of bloodstream form (BSF) or procyclic form (PCF) *T. brucei* expressing *mNeonGreen*-Tb*snm1* (*mNG*-Tb*snm1*) were generated from cDNA templates derived from total RNA extracted from wild type (wt) or recombinant parasites. The integrity of DNA samples was evaluated by amplification of a 0.1 kb control fragment, Tb*tert*.
C. Total cell extracts derived from 1 x 10^6 procyclic form *T. brucei* were size fractionated on a 10% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane. Blots were probed using an antibody to detect the mNeonGreen (mNG) tag. A band of 106 kDa was detected and corresponded to mNG tagged version of TbSNM1. No band was detected in extracts made from the wild type *T. brucei.* An extract from *T. cruzi* expressing mNG tagged luciferase (100 kDa), Costa et al., 2018, was used as positive control.

To assess functionality of mNeonGreen-TbSNM1, the entire fusion gene was amplified from genomic DNA derived from appropriate PCF *T. brucei* expressing *mNeonGreen*-Tb*snm1* and cloned into the trypanosomal expression vector, pTubEX-LmSpSyn (Taylor et al 2008) resulting in replacement of the existing *Lm*S*pSyn* ORF with *mNeonGreen*-Tb*snm1*. The ensuing pTubEX-mNGTbsnm1 plasmid was linearised, the purified DNA fragments transformed into BSF *T. brucei*

Tb*snm1Δ* (Sullivan et al., 2015) and drug (phleomycin) resistant, recombinant clonal lines selected. To confirm that the TbSNM1 null parasites were expressing *mNeonGreen*-Tb*snm1*, PCRs were performed on cDNA templates using primer combinations that span the gene fusion event. This generated an amplicon of the expected size (0.6 bp) from templates from wild type and Tb*snm1Δ* parasites both engineered to express *mNeonGreen*-Tb*snm1* with no band(s) observed in cDNAs produced from wild type or Tb*snm1Δ* cells (Figures 4.10).

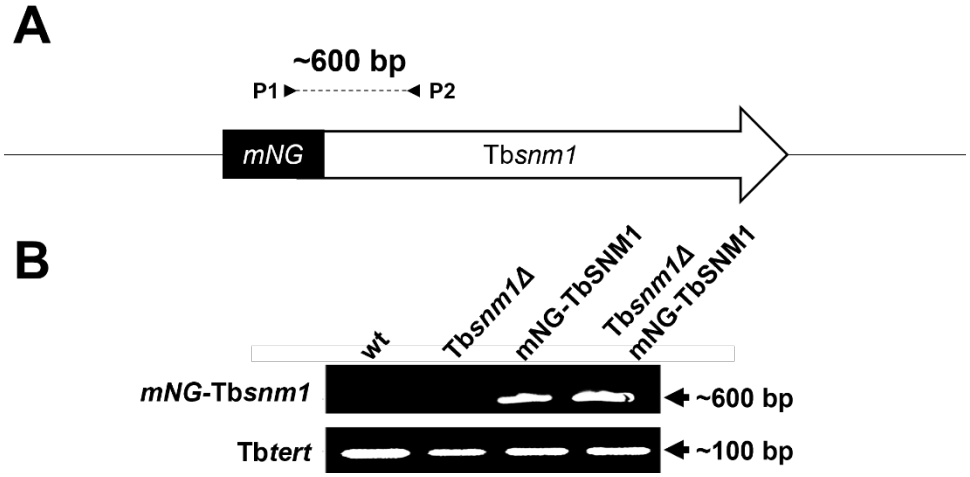


Fig 4.10. Validation of *T. brucei* Tb*snm1Δ* expressing mNeonGreen-TbSNM1.
A. Schematic showing integration of *mNeonGreen* (*mNG*) into the *T. brucei* genome at the 5' region of Tb*snm1*. P1 (GGAGTTATACCACAGGAAACG) & P2 (CACCAGATCACCATTGCGGAG) indicate the primers used in the obtention of the validation amplicons.
B. Amplicons diagnostic of bloodstream form *T. brucei* expressing *mNeonGreen*-Tb*snm1* (*mNG*-Tb*snm1*) were generated from cDNA templates derived from total RNA extracted from wild type cells (wt), parasites null for Tb*snm1* (Tb*snm1Δ*), trypanosomes expressing *mNeonGreen*-Tb*snm1* (mNG-TbSNM1) and *T. brucei* null for Tb*snm1* & expressing *mNeonGreen*-Tb*snm1* (Tb*snm1Δ* mNG-TbSNM1). The integrity of DNA samples was evaluated by amplification of a 0.1 kb control fragment, Tb*tert*. The amplified fragments are indicated (*mNG*-Tb*snmI*; Tb*tert*).

Next, the BSF Tb*snm1Δ* parasites expressing *mNeonGreen*-Tb*snm1* were fixed, their DNA stained with DAPI, and the cells imaged by fluorescent microscopy (Figures 4.11). Cells at each stage of the cell cycle presented with an intense signal throughout their nucleus, in keeping with observations made using GFP tagged TbSNM1 (Sullivan et al., 2015), with no appreciable fluorescence observed in wild type cells.

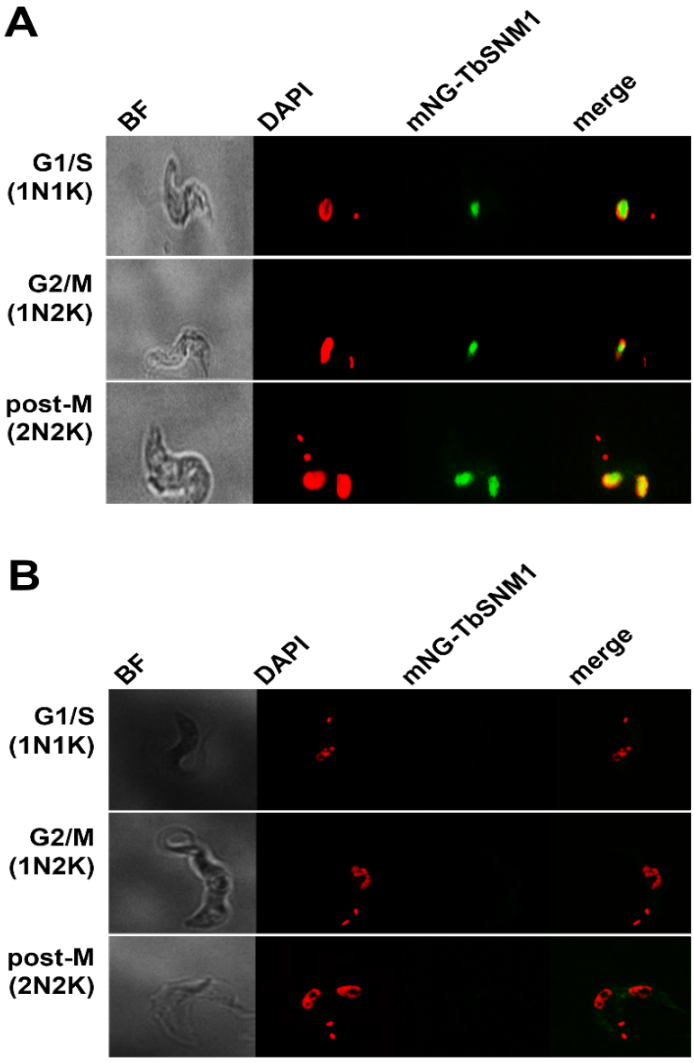


Fig 4.11. Ectopic expression of TbSNM1 in *T. brucei* *snm1Δ line*

A & B. Bloodstream form *T. brucei* ectopically expressing mNeonGreen-TbSNM1 (green; mNG-TbSNM1) from a tubulin array (**A**) were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow) at various stages of the cell cycle are shown. Wild type parasites (**B**) were analysed in parallel.

Finally, a screen was used to assess whether the previously reported Mechlorethamine susceptibility displayed by BSF Tb*snm1Δ* parasites could be complemented for by the ectopic expression of mNeonGreen-TbSNM1 in the null background (Figure 4.12) (Sullivan et al., 2015; Dattani & Wilkinson 2019). From the outputs, trypanosomes lacking TbSNM1 were shown to be approximately 4-fold more susceptible to the ICL inducing agent relative to wild type controls with expression of the tagged nuclease in the null line complementing this phenotype: wild type, Tb*snm1Δ* and Tb*snm1Δ* expressing *mNeonGreen*-Tb*snm1* parasites yielded EC_{50} values of 29.6, 7.4 and 52.6 μM towards mechlorethamine, respectively.

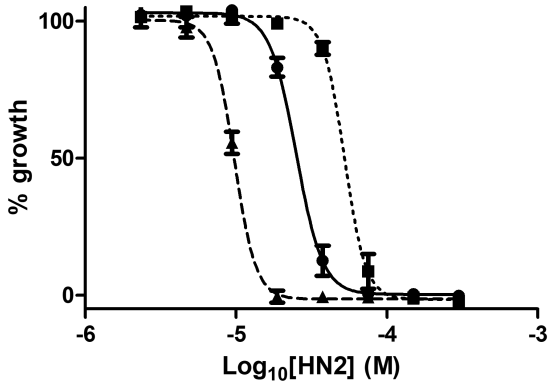


Fig 4.12. Susceptibility of Tb*snm1Δ* cells ectopically expressing mNeonGreen-TbSNM1 to mechlorethamine.
Susceptibility of bloodstream form *T. brucei* wild type (solid line), Tb*snm1Δ* (dashed line) and Tb*snm1Δ* expressing *mNeonGreen*-Tb*snm1* (dotted line) parasites to mechlorethamine (HN2). All data are mean values ± standard deviations from experiments performed in quadruplicate.

4.2.3 Localisation of TbSNM1 in *T. brucei*.

To evaluate whether expression of TbSNM1 tagged at its amino terminal with mNeonGreen from the endogenous Tb*snm1* loci influenced localisation, BSF or PCF *T. brucei* were fixed and their DNA stained with DAPI then imaged by fluorescent microscopy (Figures 4.13). Most (~90%) BSF and PCF recombinant parasites presented an mNeonGreen-TbSNM1 signal, several fluorescent foci were detected in the nucleoplasm (Figure 4.13). In some cells in the G1/S phase of the cell cycle, these fluorescent foci were associated with the non-DAPI stained nucleolus while during the G2/M phase, the SNM1-associated signal became more dispersed throughout the enlarging nucleus becoming more coalesced and becoming nucleolar again post-mitosis. This pattern is slightly different to that previously noted where an intense pattern throughout the nucleus was observed (Sullivan et al., 2015). This discrepancy most likely reflects any difference in expression between the two genomic sites where the tagged protein was inserted: In the earlier finding, TbSNM1 tagged at its amino terminal with GFP was ectopically expressed at a trypanosomal rRNA loci using a T7 rRNA promoter under the control of a TET operator while here expression was at the endogenous Tb*snm1* loci. In contrast, no signal corresponding to mNeonGreen fluorescence was observed in wild type cells or in about 10 % of recombinant insect of mammalian form parasites

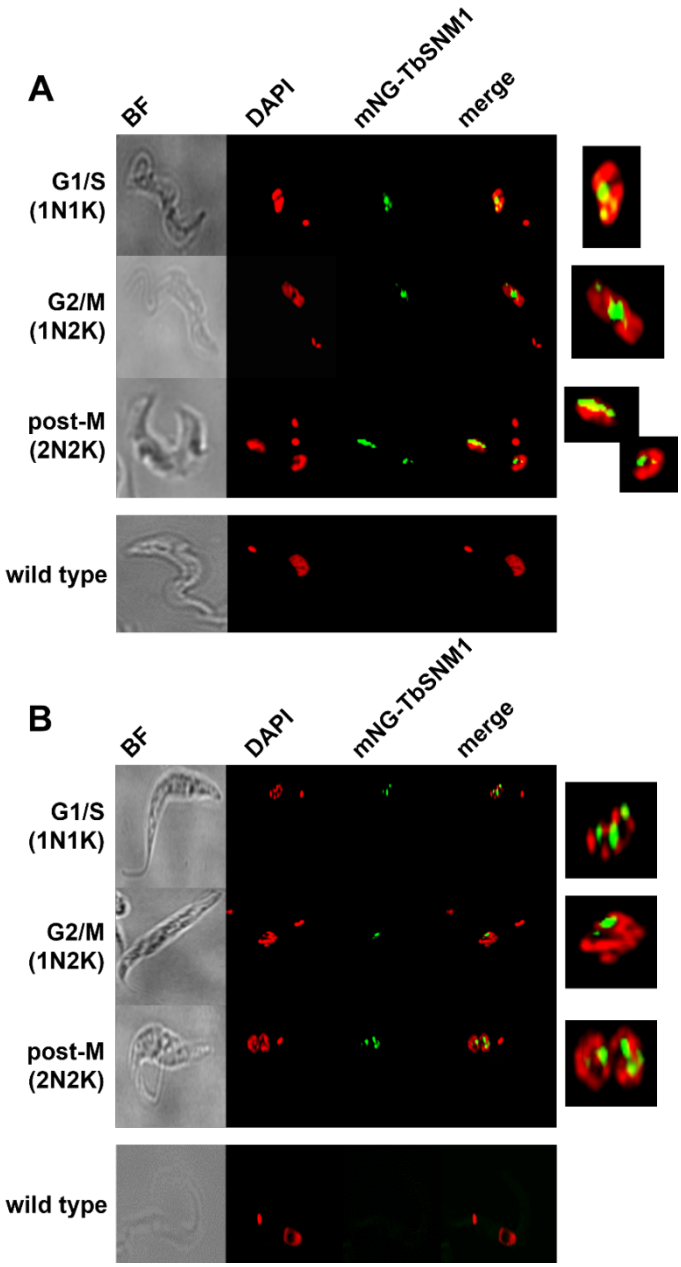


Fig 4.13. Localisation of TbSNM1 in *T. brucei*. **A & B**. Bloodstream (**A**) and procyclic (**B**) form trypomastigote *T. brucei* expressing TbSNM1 tagged with mNeonGreen (in green) at its amino (mNG-TbSNM1) terminal were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow), including a close-up image of the nucleus, in trypanosomes at various stages of the cell cycle are shown.

4.2.4 Effect of exogenous treatments on TbSNM1 expression.

As for the HR factors (Chapter 3), the effect of exogenous agents on the expression of mNeonGreen-TbSNM1 from its endogenous Tb*snm1* loci in bloodstream form parasites was assessed. Following treatment and at a single time point (30 µM mechlorethamine for 4 hours; 3 µM nifurtimox for 6 hours; 30 µM benznidazole for 2 hours), no change in the overall TbSNM1 localisation was observed (cells presenting multiple fluorescent foci in the nucleoplasm) irrespective of the agent used (Figure 4.14A) with an increase in intensity of these centres detected in a significant number of trypanosomes relative to untreated controls (Figure 4.14B).

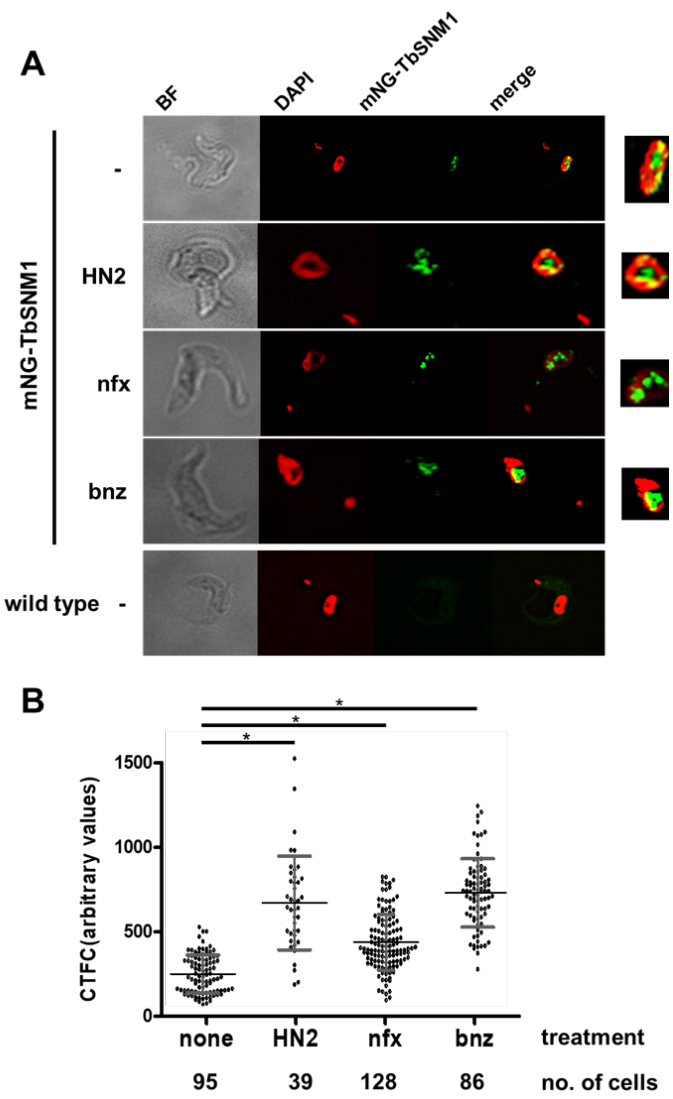


Fig 4.14. Effect of exogenous treatments on TbSNM1 expression

A. Bloodstream form *T. brucei* expressing mNeonGreen-TbSNM1 (mNG-TbSNM1; green) treated with mechlorethamine (HN2; 30 µM for 2 hours), nifurtimox (nfx; 3 µM for 3 hours) or benznidazole (bnz; 30 µM for 2 hours) were co-stained with DAPI (red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merge; yellow), including a close-up image of the nucleus, in treated and untreated (-) trypanosomes were compared.

B. The corrected total cellular fluorescence (CTFC) of individual *T. brucei* expressing mNeonGreen tagged TbSNM1 following mechlorethamine (HN2), nifurtimox (nfx) or benznidazole (bnz) treatment was determined and compared against untreated controls (none). Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).

The above analysis was extended to assess the temporal expression of TbSNM1 in response to the mechlorethamine (30 µM), nifurtimox (3 µM) or benznidazole (30 µM) (Figure 4.15). Following addition of exogenous agent mechlorethamine, a significant increase in the fluorescence intensity displayed by individual cells within the population was shown to take place in the first hour of treatment and remain at approximately the same level to the final time point (6-hours posttreatment) (Figure 4.15A). When analysing whether the two nitroheterocycles caused any temporal changes, a significant increase in TbSNM1 expression was noted over the 1st hour of treatment when using nifurtimox, and by 2-hour posttreatment for benznidazole. In both cases, the

fluorescent signal continued to rise reaching a peak at 3-hours posttreatment before declining at 4- and then 6-hours posttreatment (Figure 4.15B & C).

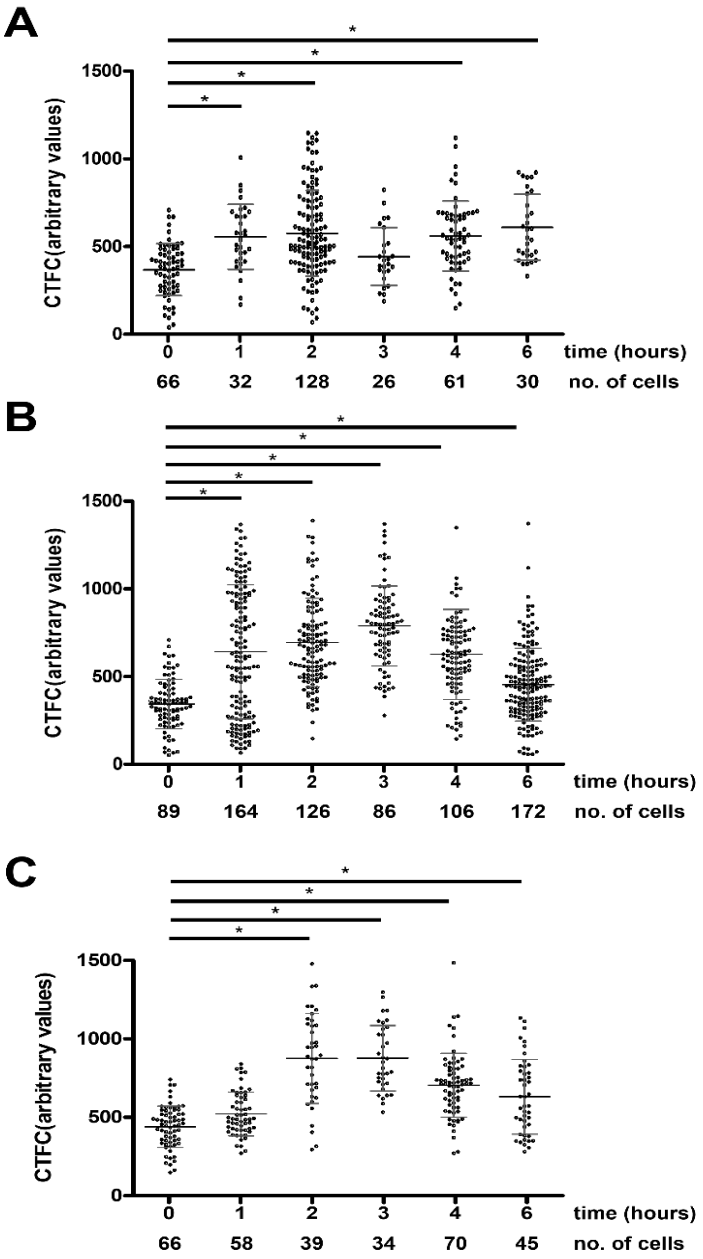


Fig 4.15. Effect of exogenous agents on the temporal expression of TbSNM1.
A-C. The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen TbSNM1 was determined at time intervals following mechlorethamine (30 µM) (**A**), nifurtimox (3 µM) (**B**) or benznidazole (30 µM) (**C**) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Student's t test (GraphPad Software).

4.2.5 Development of insect form *T. brucei* null for Tb*snm1*.

To assess the role of TbSNM1 in insect form *T. brucei*, a gene disruption approach was taken. Here, a CRISPR/Cas9 genome editing strategy was used to completely replace both copies of

Tb*snm1* in the parasite's genome with DNA sequences coding for the hygromycin and neomycin drug selectable markers (Figure 4.16) (Beneke et al 2017). To do this, two small amplicons that contain sequences that when expressed in the parasite generate sgRNA and two large amplicons that contain sequences used to delete the targeted gene were electroporated into PCF *T. brucei* with recombinant clones selected: A heterozygote line was constructed in parallel by omitting the hygromycin targeting amplicon/corresponding drug selection. Primer combinations used to generate the amplicon are given in Table 4.2.

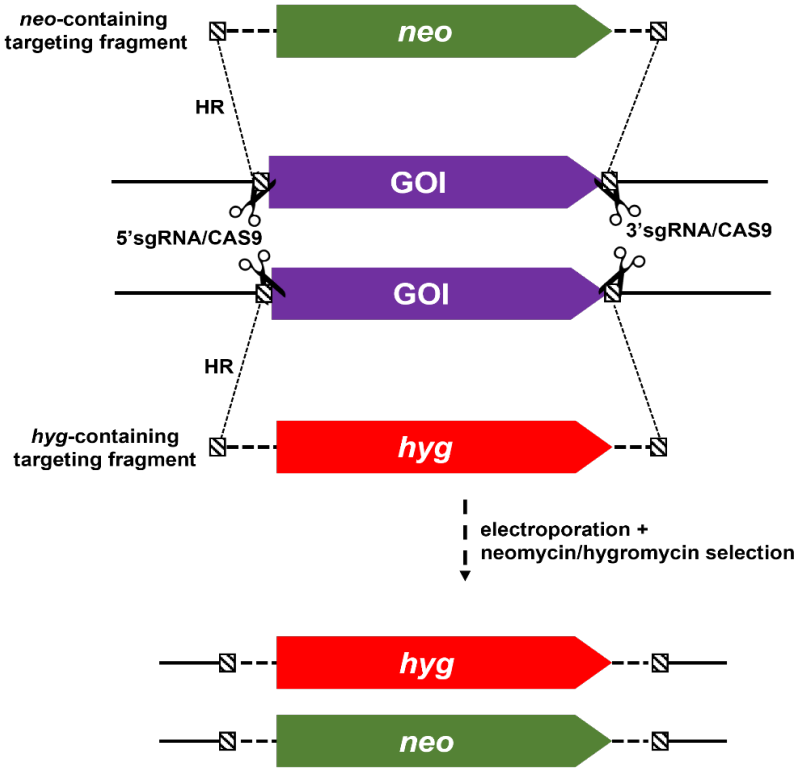


Fig 4.16. Gene deletion strategy.
Schematic diagram outlining the CRISPR/Cas9-mediated procedure used to delete a gene of interest (GOI) from the *T. brucei* genome and its replacement with resistance cassettes that contain sequences coding for the hygromycin B phosphotransferase (*hyg*) or neomycin phosphotransferase (*neo*) selectable markers plus the *T. brucei* untranslated regions required for mRNA processing. sgRNA/Cas9 complexes (indicated by scissors) introduce DSBs at sites 5' and 3' to the GOI with the lesion subsequently repaired via the homologous recombination pathway (HR) using NEO- and/or HYG-containing targeting fragments.

Amplicon	Primer combinations
5' sgRNA-Tbsnm1	G00/5'sgRNA-TbSNM1
3' sgRNA-Tbsnm1	G00/3'sgRNA-TbSNM1
HYG-TF-Tbsnm1	TbSNM1-UFP1/ TbSNM1-DRKOP1
NEO-TF-Tbsnm1	TbSNM1-UFP1/ TbSNM1-DRKOP1

Table 4.2. Primer combinations used in CRISPR/Cas9 gene deletion of Tb*snm1*.
The primer combinations used to generate small amplicons (designated as 5'sgRNA-Tbsnm1 & 3'sgRNA-Tbsnm1) that correspond to the short guide RNA (sgRNA) which target the 5' or 3' ends of Tb*snm1* and large amplicons (designated as HYG-TF-Tbsnm1 & NEO-TF-Tbsnm1) that correspond to the hygromycin or neomycin targeting fragments used to replace Tb*snm1*. The sequence of each primer can be found in Appendix 1.

To show that the resultant clonal line no longer expressed Tb*snm1*, DNA amplification reactions were performed on cDNA. When using a primer combination designed to detect intact Tb*snm1*, a single band of the expected size (670 bp) was detected in templates made from wild type parasites and the Tb*snm1*$^{+/-}$*neo* heterozygote line but not cDNA derived from the Tb*snm1Δ* line (Figure 4.17). To confirm that RNA had been extracted from all lines and that cDNA had been made, control reactions amplifying Tb*tert* were conducted in parallel. For all tested samples, an amplicon of the expected size (approximately 100 bp) was detected.

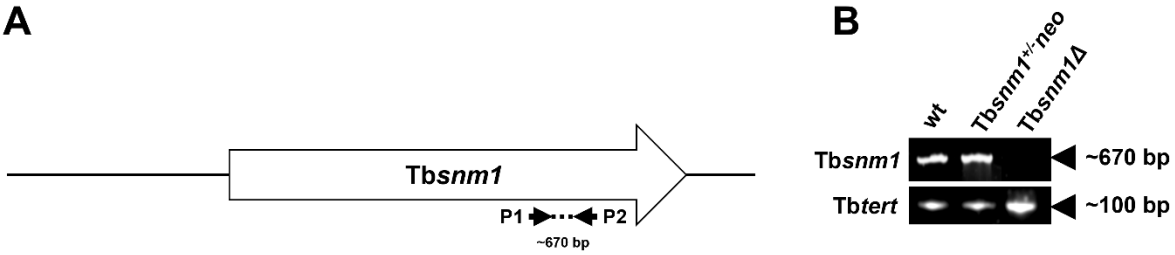


Fig 4.17. Effect of Tb*snm1* disruption on insect form *T. brucei*.
A. Schematic representation of the Tb*snm1* allele. P1 (aaaggatccCTCGTTAGCCCT GTTTCTCG) and P2 (gggctcgagTCTCCTCATCCTCCACATCC) correspond to the primers used to generate a Tb*snm1* specific amplicon from cDNA templates.
B. Amplicons corresponding to intact Tb*snm1* were generated from template cDNAs derived from total RNA extracted from procyclic form *T. brucei* wild type (wt), Tb*snm1* heterozygote (Tb*snm1*$^{+/-}$*neo*) and null (Tb*snm1Δ*) lines. The integrity of cDNAs (and hence RNAs) was evaluated by amplification of a 100 bp control fragment, Tb*tert*. The amplified fragment (Tb*snm1* or Tb*tert*) is indicated.

4.2.6 Phenotypic analysis of insect form *T. brucei* Tb*snm1Δ*.

The ability to make a PCF *T. brucei* line null for TbSNM1 clearly shows that this nuclease is not essential for growth of this parasite stage, consistent with the observations using BSF cells (Sullivan et al., 2015). To phenotype the insect line, its growth and susceptibility characteristics were assessed.

To perform the PCF *T. brucei* growth assays, independent cultures containing 1 x 10^4 parasites ml^{-1} were established and every few days over a period of 22 days, the number of cells in each culture was determined generating cumulative cell density growth curves that allowed us to calculate their doubling time (Figure 4.18). Over this period, cells were maintained in the exponential phase of growth: cultures that had a cell density between 1 and 5 x 10^7 parasites ml^{-1} were used to establish a new culture containing 1 x 10^4 cells ml^{-1}. This established that Tb*snm1*Δ

cells exhibited a growth difference as compared to the wild type with the mutant having a doubling time of around 23 hours and the control being around 13 hours.

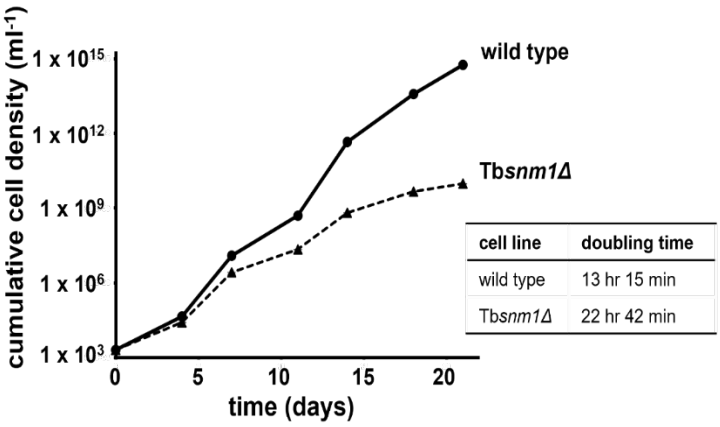


Fig 4.18. Growth of *T. brucei* Tb*snm1Δ* null mutant.
The cumulative cell density of procyclic form *T. brucei* Tb*snm1Δ* (▲) was followed for 22 days and compared against the growth of wild type (wt) *T. brucei* (●) cultures grown in parallel. Each data point represents the mean cell density ± standard error from two independent cultures. Insert table shows the doubling time of each cell line.

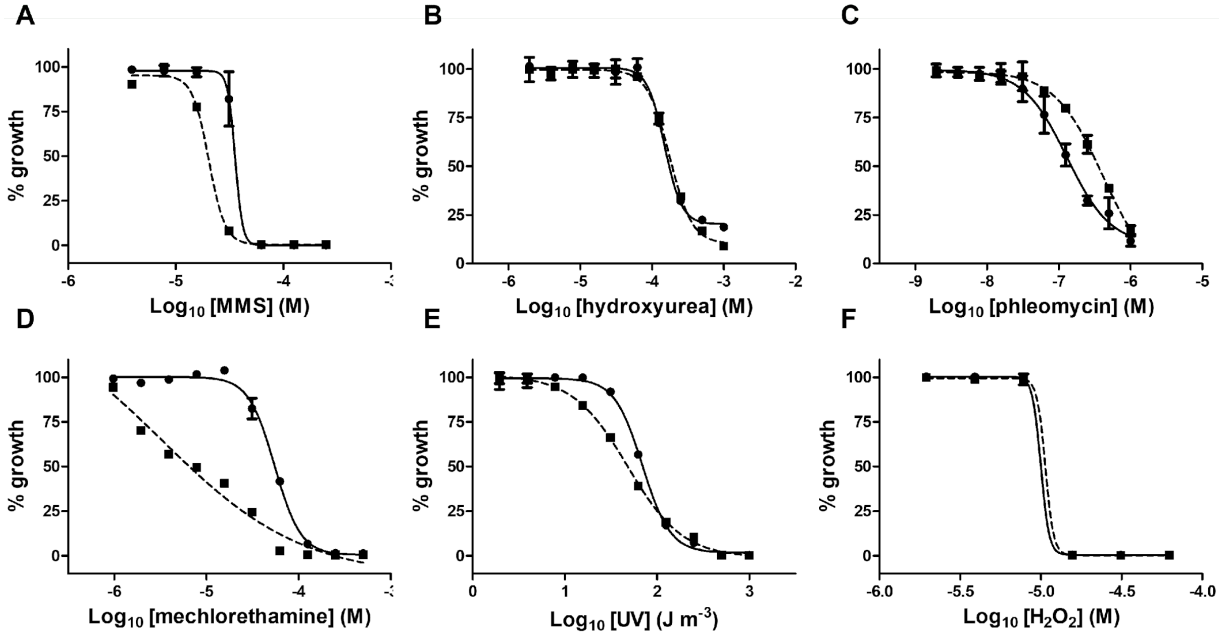


Fig 4.19. Characterization of *T. brucei* Tb*snm1Δ* null mutants. A-F. Susceptibility of procyclic form *T. brucei* wild type (solid line) and Tb*snm1Δ* (dashed line) to methyl methane sulfonate (MMS) (**A**), hydroxyurea (**B**), phleomycin (**C**), mechlorethamine (**D**), UV light (**E**) and H_2O_2 (**F**). All data are mean values ± standard deviations from experiments performed in quadruplicate.

To evaluate whether lack of TbSNM1 activity effects the susceptibility of PCF *T. brucei*'s to several types of DNA damaging agent, a series of growth inhibition assays were performed (Figure 4.19). The treatments used were hydroxyurea and H_2O_2 (which promote base oxidation),

phleomycin (causes double strand breaks), methyl methanesulfonate (MMS; alkylation leading to dsDNA breaks), ultraviolet radiation (UV; pyrimidine dimer formation) and mechlorethamine (ICL formation). In these assays, parasites were grown for 3 days in the presence of different concentrations of compound (or for UV, exposed to a single bolus prior to culturing for 3 days) 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 (Figure 4.19) from which the treatment EC_{50} value (Table 4.3) was calculated.

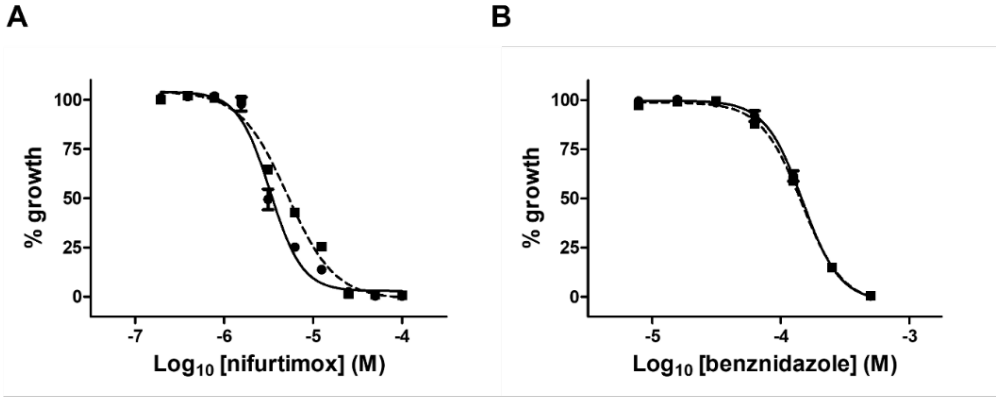
Treatment	**EC_{50}**		**Fold difference**
	wild type	**Tb*snm1Δ***	
MMS (μM)	36.8 ± 3.2	20.6 ± 0.1*	1.8
Hydroxyurea (μM)	150.4 ± 3.0	177.6 ± 7.1	0.8
Phleomycin (nM)	126.6 ± 16.0	466.4 ± 176.4	0.3
Mechlorethamine (μM)	54.6 ± 1.8	3.3 ± 1.2*	16.4
UV light (J m^{-3})	69.2 ± 0.8	49.8 ± 1.2	1.4
H_2O_2 (μM)	10.3 ± 0.7	9.5 ± 0.2	1.1

Table 4.3. Susceptibility of *T. brucei* Tb*snm1Δ* null mutants to DNA damaging agents.
The susceptibility of PCF *T. brucei* Tbsnm1Δ to various treatments, as judged by EC50 values extrapolated from dose response curves, were determined, and expressed as a fold difference relative to wild type. The asterisk (*) indicates significant difference in susceptibility (P<0.0001) between wild type and Tb*snm1Δ* cells, as assessed by the on-line Student's t test calculator (https://www.graphpad.com/quickcalcs/ttest1/?Format=SD).

Based on dose response curves and extrapolated EC_{50} values, trypanosomes lacking TbSNM1 are as equally sensitive to hydroxyurea, UV light and H_2O_2 as wild type indicating that this DNA repair factor does not play a front-line role in resolving the resultant lesions generated by these treatments. In contrast, PCF Tb*snm1Δ* cells were almost 2-fold and 16-fold more susceptible to MMS and mechlorethamine, respectively, indicating that in insect form *T. brucei*, TbSNM1 plays a role in protecting the parasite from alkylating agents, particularly those that generate ICLs. For MMS, the observed susceptibility alterations are at odds with the phenotypes obtained when analysing the BSF counterparts (Sullivan et al 2015): BSF Tb*snm1Δ* cells were as equally susceptible to this agent than wild type. This suggests that the systems used to repair the DNA

damage induced by MMS is differentially regulated, a trait previously noted (Vieira-da-Rocha et al 2019; Dattani et al 2021). Intriguingly, cells lacking TbSNM1 were about 3-fold more resistant to phleomycin. This unexpected finding may suggest that TbSNM1 may interact with and potentially interfere with trypanosomal DNA repair factors associated with repair of lesions generated by this agent.

When the above susceptibility studies were extended to nifurtimox and benznidazole, PCF Tb*snm1Δ* cells were as equally susceptible to these agents as wild type (Figure 4.20). This clearly shows that the trypanocidal nitroheterocyclic prodrugs do not generate DNA lesions that require TbSNM1 activity to resolve and is in keeping with previous findings on BSF trypanosomes (Sullivan et al 2015).



C

Treatment	EC_{50}		Fold difference
	wild type	Tb*snm1Δ*	
Nifurtimox (μM)	3.8 ± 0.1	5.1 ± 0.3	0.7
Benznidazole (μM)	133.7 ± 15.1	145.2 ± 3.7	1.1

Fig 4.20. Susceptibility of *T. brucei* Tb*snm1Δ* null mutants to anti-trypanosomal treatments.
A & B. The susceptibility of procyclic form *T. brucei* wild type (solid line) and TbsnmIΔ (dashed lines) to nifurtimox (A) and benznidazole (B) were determined.
C. The EC50 values of procyclic form *T. brucei* Tbsnm1Δ were extrapolated from the dose response curves shown (A & B) and expressed as a fold difference (in parenthesis) relative to wild type.
All data are mean values ± standard deviations from experiments performed in quadruplicate.

4.3 Assessing the role of TbFAN1 in *T. brucei*.

4.3.1 Informatic analysis of TbFAN1.

In mammalian cells, FAN1 is an ancillary protein of the FA pathway that has both 5'-3' exonuclease and 5' flap endonuclease activities (Kratz et al 2010; Liu et al 2010; MacKay et al., 2010; Smogorzewska et al. 2010; Zhao et al., 2014; Pizzolato et al 2015). It is recruited to lesion sites through its interaction with the ubiquitinated FAND2 component of the FA recruitment

complex and there participates in the ICL 'unhooking' process (Zhang & Walter, 2014). Searches of data held on TriTrypDB (Aslett et al., 2009) using human FAN1 (NP_055782) as bait identified a single 3091 bp ORF (Tb927.10.7540; designated as Tb*fan1*) from the *brucei* TREU927 strain. This sequence is located on chromosome 10 and has potential to code for protein of 79.5 kDa, designated here as TbFAN1. Pairwise analysis revealed TbFAN1 was approximately 40 % identical to TcFAN1 from *T. cruzi* CL Brener Non-Esmeraldo-like haplotype (XP_819887) and between 16 to 21% identical to orthologues expressed by *Leishmania* species (XP_001686839; CAC5434848) *Schizosaccharomyces japonicus* (XP_002175115), *H. sapiens* (NP_055782) and *A. thaliana* (Q5XVJ4).

Searches of data embedded on the TriTrypDB or following analysis using HMMER revealed that TbFAN1 contained three Tetratricopeptide Repeat_2 domains (TPR_2; PF07719) dispersed throughout the first half of the sequence (between residues 28 to 41, 340 to 360 and 501 to 526) and represent sequences believed to mediate protein: protein interactions (Lamb et al 1995) (Figure 4.21). Additionally, and in keeping with FAN1 from other organisms, TbFAN1 was shown to possess a virus-type replication-repair nuclease domain (VRR-NUC) (PF08774) towards its carboxyl terminus (residues 837 to 948) (Figure 4.21). This region contains a variant of PD-(D/E) XK, a motif characteristic of the VRR-NUC domain, although in the trypanosomal enzyme the lysine has been replaced by leucine, with this sequence proposed to be involved in divalent metal binding, a feature not required for FAN1's nuclease activity (Hadden et al 2007; Wang et al 2014). In comparison to counterparts expressed by other organisms, TbFAN1 lacks discernible domains present in other FAN1 members that mediate ubiquitin or DNA binding. In human FAN1, such interactions are mediated by the TRAF-type zinc finger (PF02176) and SAP (PF02037) domains, respectively (Smogorzewska et al., 2010; Zhao et al., 2014) while the latter role in plant enzymes is performed by a HIRAN (PF08797) domain (Iyer et al 2006). The absence of the ubiquitin binding domain in TbFAN1 is perhaps not a surprise given that in the human enzyme this region guides interaction between the nuclease and monoubiquitylated FANCD2 (*T. brucei* lacking a FANCD2 orthologue) while the TbFAN1 motif used to bind DNA has yet to be established.

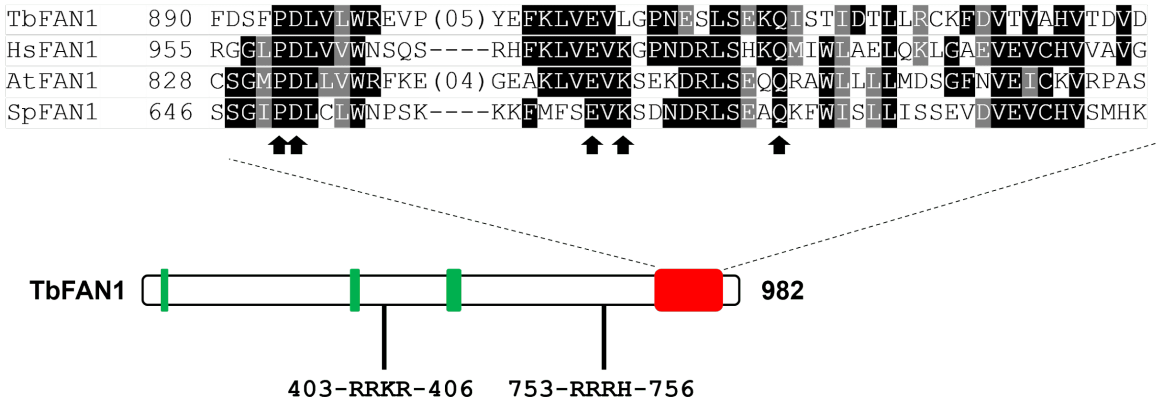


Fig 4.21. Informatic analysis of TbFAN1.
The sequence corresponding to the to VRR_NUC domain (PF08774; residues 837-948; red box) of TbFAN1 (EAN78160) was aligned with the equivalent regions from the orthologues expressed by *Homo sapiens* (HsFAN1; NP_055782), *Arabidopsis thaliana* (AtFAN1; Q5XVJ4) and *Schizosaccharomyces pombe* (SpFAN1; Q9Y804). Residues that are identical or are conserved substitutions are highlighted in black and grey, respectively while dashes have been inserted to optimise alignment. Arrows correspond to conserved residues that constitute the HMM logo of the VRR_NUC domain (Finn et al 2016 Nuc Acid Res 44, D279-D285). Three regions corresponding to Tetratricopeptide Repeat_2 domains (TPR_2; PF07719; residues 28-41, 340-360 & 501-526; green boxes) in TbFAN1 are highlighted. The location of potential 4 pattern nuclear localisation signals is shown.

Using Phyre2, a 3D prediction of full-length TbFAN1 monomer based on multiple structures was generated (Figure 4.22). Using the human FAN1 crystal structure (4RIB; Wang et al 2014), the TbFAN1 model could be divided into three regions. Here, we postulate that the globular amino terminal region may contain sequences that mediate DNA binding as is the case of all other FAN1 enzyme examined to date with the central section, which in TbFAN1 comprising a series of α helices including three that correspond to the TPR_2 domains, represent regions that facilitate protein: protein interactions. The final portion of TbFAN1 corresponds to the catalytic region and contains the VRR-NUC domain. The model for this latter region (Figure 4.22) is similar to proteins of the restriction endonuclease-like superfamily of which FAN1 belongs is characterised by mixed α/β fold with αββαβ topology, containing parallel β-sheet bordered by α-helices at one end (Iyer et al., 2006).

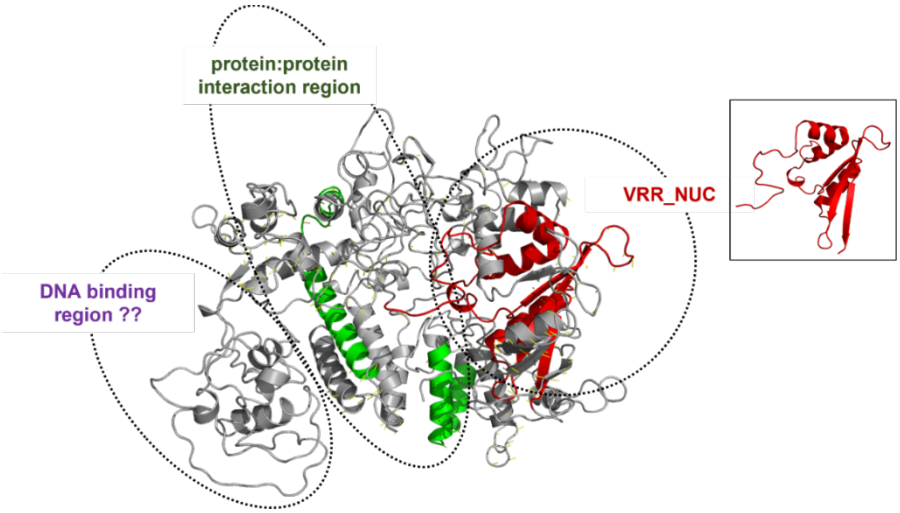


Fig 4.22. 3D prediction model of TbFAN1.
Full-length 3D prediction of TbFAN1 (EAN78160) was constructed using Phyre2 (Kelley et al. 2015) and visualized using PyMOL. The full-length approximation was generated from three templates (6WJE, 4R8A & 4RIB) with 677 residues (69%) modelled at >90% confidence. The position of the VRR_NUC (PF08774; red) and three TPR_2 (PF07719; green) domains are shown. A putative DNA binding region is indicated. Non-domain regions are in grey. Insert shows the TbFAN1's VRR-NUC domain in isolation.

4.3.2 Editing the Tb*fan1* loci in *T. brucei*

Based on localisation signal detection algorithms (PSORTII; WoLFPSORT), TbFAN1 is predicted to be found in the *T. brucei* nucleus possibly targeted to this organelle via a RRKR motif found at residues 403 to 406 or a RRRH sequence located at residues 753 to 756 (Figure 4.21). To assess localisation of this protein the previously described CRISPR/Cas9 genome editing approach (see Figure 3.13) was employed to tag the endogenous Tb*fan1* at its 5' or 3' end with nucleotide sequences coding for the fluorescent protein mNEONGREEN (Beneke et al., 2017). Following the scheme set out in Chapter 3 and using appropriate primer combinations (Table 4.4), two amplicons that correspond to the sgRNA and targeting fragment were electroporated into BSF or PCF *T. brucei* with recombinant, drug (hygromycin) resistant clones selected.

	Primer combinations	
	sgRNA	**Targeting fragment**
5' tagging	G00/5'sgRNA TbFAN1	TbFAN1-UFP1/TbFAN1-DRF1
3' tagging	G00/3'sgRNA TbFAN1	TbFAN1-DFP1/TbFAN1-DRTAGP

Table 4.4. Primer combinations used in CRISPR/Cas9 tagging of Tb*fan1*.
Primer combinations used to generate amplicons corresponding to the short guide RNA (sgRNA) &targeting fragment used in tagging the 5' or 3' ends of endogenous Tb*fan1* with a sequence coding for mNEONGREEN. The sequence of each primer can be found in Appendix 1.

A set of validatory DNA amplifications similar to those described for Tb*snm1* were performed on cDNA templates generated from parasites expressing *mNeonGreen*-TbFAN1 or TbFAN1-*mNeonGreen* (Figure 4.23). Using primer combinations to detect expression of these fusion gene, a single band of the expected size (1000 bp) was detected in cDNA derived from all recombinant BSF and PCF lines with no band(s) detected from wild type templates (Figure 4.23B and D).

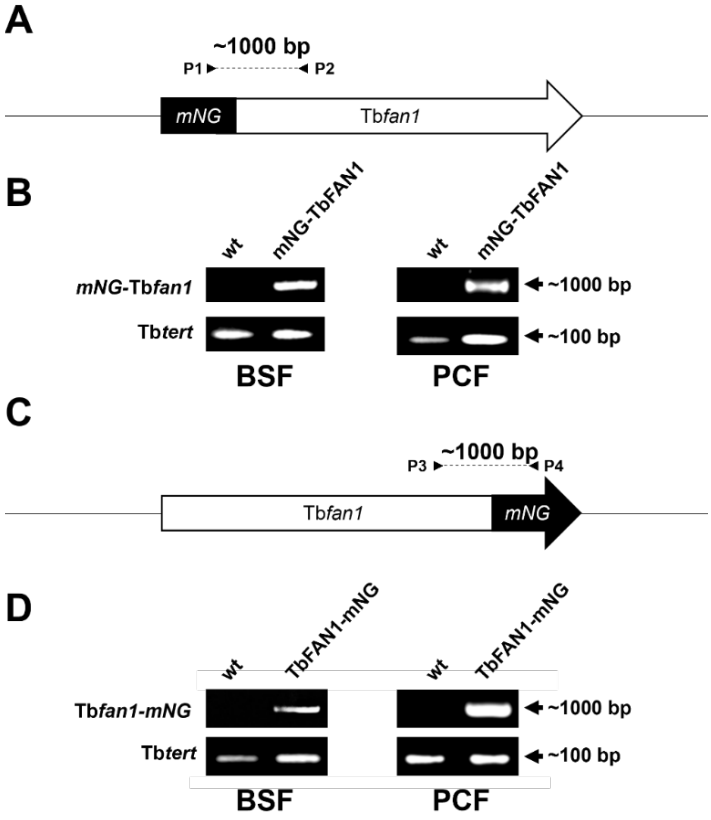


Fig 4.23. Validation of *T. brucei* lines expressing mNeonGreen tagged TbFAN1.
A & C. Schematic showing integration of *mNeonGreen* (*mNG*) into the *T. brucei* genome at the 5' (**A**) or 3' (**C**) region of Tb*fan1*. P1 (GGAGTTATACCACAGGAAACG), P2 (ATTCCAACACAATATCAGCGC), P3 (TTTTGGACCCTGAGGAGTTTCCT) and P4 (ACAAGAATCCACGGGGAAAAC) indicate the primers used in the obtention of the validation amplicons.
B & D. Amplicons diagnostic of bloodstream form (BSF) or procyclic form (PCF) *T. brucei* expressing *mNG*-Tb*fan1* (**B**) or Tb*fan1-mNG* (**D**) were generated from cDNA templates derived from total RNA extracted from wild type (wt) or recombinant parasites. The integrity of DNA samples was evaluated by amplification of a 0.1 kb control fragment, Tb*tert*. The amplified fragment is indicated (*mNG*-Tb*fanI*; Tb*fanI-mNG* or Tb*tert*). The primer sequences and combinations used for each amplification are listed in Appendix 2, respectively.

4.3.3 Localisation of TbFAN1 in *T. brucei*.

To evaluate localisation BSF or PCF *T. brucei* expressing TbFAN1 tagged at its amino or carboxyl termini with mNeonGreen were fixed, their DNA stained with DAPI then imaged by fluorescent microscopy (Figures 4.24 & 4.25). For most (80 to 90 %) insect and mammalian form parasites and at each cell cycle stage, several fluorescent foci were detected in the nucleoplasm of trypanosomes presenting a mNEONGREEN-TbFAN1- or TbFAN1-mNEONGREEN signal. In some cells, a faint punctate signal was noted in the cytoplasm but not along the flagellum (data not shown). These patterns are consistent with the nuclear/cytoplasmic (reticulated) localisation attributed to TbFAN1 as part of the TrypTag project (Dean et al., 2017). In all cases, an

mNeonGreen signal was not detected in 10-20 % of recombinant parasites with no fluorescence observed in wild type cells.

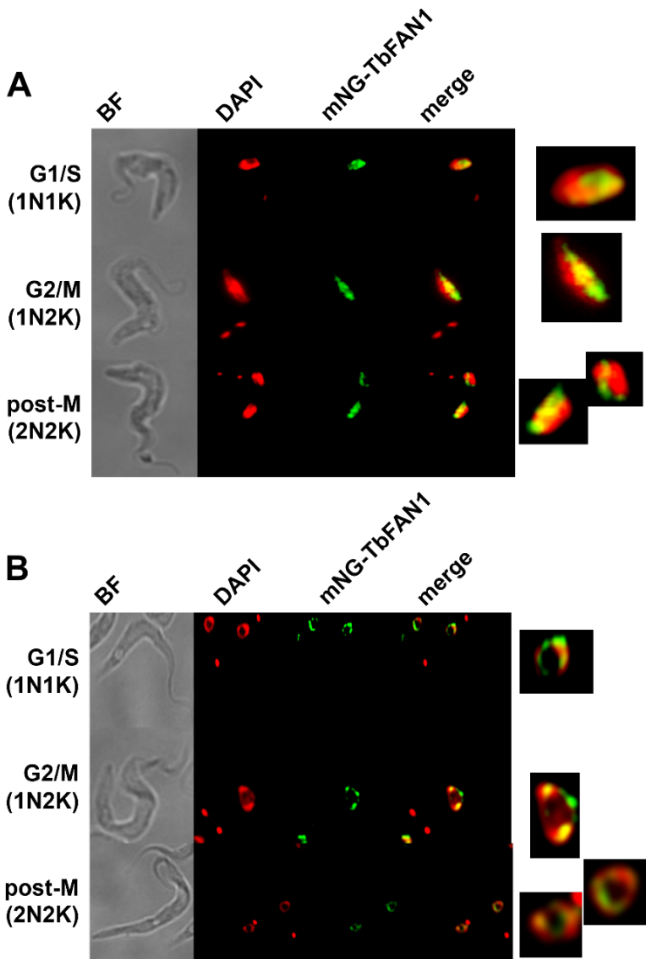


Fig 4.24. Localisation of TbFAN1 in *T. brucei* (part 1).
A & B. Bloodstream (A) and procyclic (B) form trypomastigotes *T. brucei* expressing TbFAN1 tagged with mNeonGreen (in green) at its amino (mNG-TbFAN1) were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow), including a close-up image of the nucleus, in trypanosomes at various stages of the cell cycle are shown.

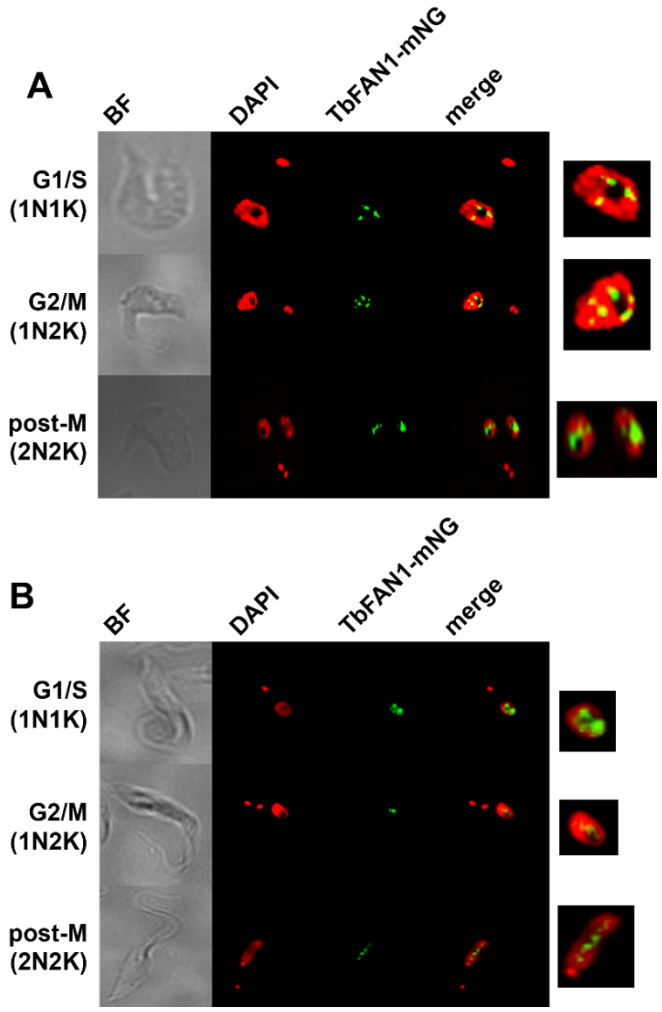


Fig 4.25. Localisation of TbFAN1 in *T. brucei* (part 2).
A & B. Bloodstream (**A**) and procyclic (**B**) form trypomastigotes *T. brucei* expressing TbFAN1 tagged with mNeonGreen (green) at its carboxyl (TbFAN1-mNG) termini were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow), including a close-up image of the nucleus, in trypanosomes at various stages of the cell cycle are shown.

4.3.4 Effect of exogenous treatments on TbFAN1 expression.

The effect of exogenous agents on the expression of mNeonGreen-TbFAN1 from the endogenous Tb*fan1* loci was assessed. Upon treatment with mechlorethamine (30 µM for 6 hours), nifurtimox (3 µM for 6 hours) or benznidazole (30 µM for 2 hours) the TbFAN1 signal was shown to switch from being found mainly as discrete foci within the nucleoplasm to predominantly forming multiple, discernible centres in the nucleolus (Figure 4.26A). The change in nuclear localisation was accompanied by a significant increase in the fluorescence intensity displayed in a substantial number of trypanosomes relative to untreated controls (Figure 4.26B).

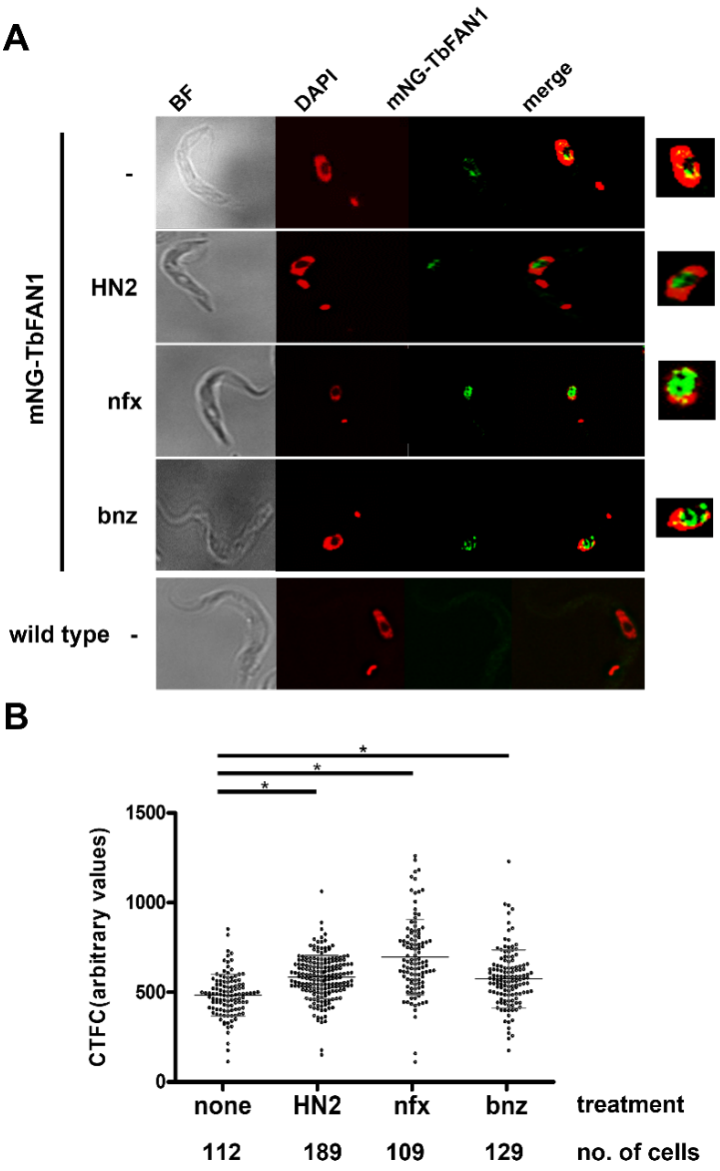


Fig 4.26. Effect of exogenous treatments on TbFAN1 expression.

A. Procyclic form *T. brucei* expressing mNeonGreen-TbFAN1 (mNG-TbFAN1; green) treated with mechlorethamine (HN2; 30 µM for 6 hours), nifurtimox (nfx; 3 µM for 2 hours) or benznidazole (bnz; 30 µM for 2 hours) were co-stained with the DNA stain, DAPI (red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merge; yellow), including a close-up image of the nucleus, in treated and untreated trypanosomes were compared. Untreated (-) recombinant and wild type parasites were analysed in parallel.

B. The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen tagged TbSNM1 following mechlorethamine (HN2), nifurtimox (nfx) or benznidazole (bnz) treatment was determined and compared against untreated controls (none). Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).

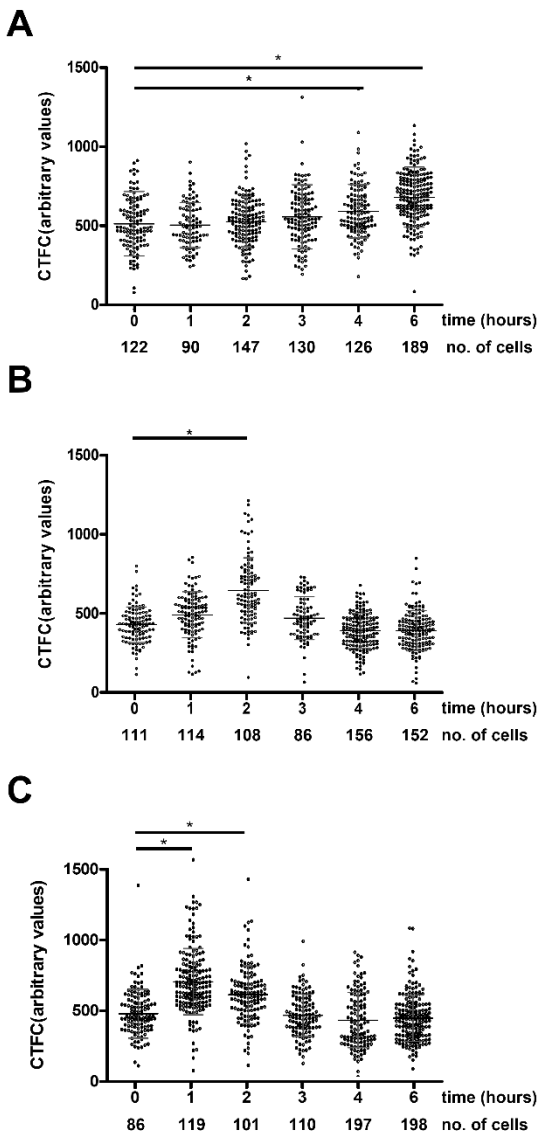


Fig 4.27. Effect of exogenous agents on the temporal expression of TbFAN1.

A-C. The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen TbFAN1 was determined at time intervals following mechlorethamine (30 µM) (**A**), nifurtimox (3 µM) (**B**) or benznidazole (30 µM) (**C**) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Student's t test (GraphPad Software).

To assess the temporal response of TbFAN1 expression to exogenous agents, PCF parasites expressing mNeonGreen-TbFAN1 were treated with a single bolus of mechlorethamine (30 µM), nifurtimox (3 µM) or benznidazole (30 µM) and periodically the fluorescent signals in individual cells determined (Figure 4.27). When mechlorethamine was added to cultures, a slight but non-significant increase in TbFAN1 expression was noted over the first 3 hours of treatment, with the fluorescent signal gradually rising becoming significant by 4-hours posttreatment and continuing to rise at the final time point, 6-hours posttreatment (Figure 4.27A). Treatment of the recombinant line with nifurtimox resulted a slight but non-significant increase in TbFAN1 expression was noted

over the 1st hour of treatment, becoming significant at 2-hours posttreatment before declining and returning to the levels noted in untreated samples by 4-hours posttreatment (Figure 4.27B). In the case of benznidazole, a pattern reminiscent of the profile obtained when analysing mNeonGreen-TbMRE11 was observed (Figure 4.27C). Here, an increase in mNeonGreen-TbFAN1 expression was shown to take place, reaching a maximum within the first hour of adding the nitroimidazole to the culture medium. The fluorescence signal remained at approximately this level at 2-hours posttreatment before reverting to untreated levels by 6-hours posttreatment.

4.4. Chapter Summary

In this chapter, we expanded on the role of TbSNM1 and TbFAN1 in the *T. brucei* ICL repair system. First, we utilised PCR-based CRISPR/Cas9 genome editing technique (Beneke et al 2017) to tag parasites to express tagged versions of both proteins in insect or bloodstream form parasite lines and additionally to produce procyclic form parasite lines that lacked Tb*snm1*. We also investigated the effect of ectopic expression on *T. brucei* SNM1 response to DNA damage. Recombinant parasites were phenotyped with respect to subcellular location of each enzyme in both parasite stages and protein expression in response to ICL inducing agents (mechlorethamine) and trypanocidal nitroheterocyclic prodrugs was assessed. Parasite lines null for TbSNM1 were phenotyped to assess growth profiles and susceptibility to DNA damage

Here we showed:

1. The sub-cellular localisation of mNeonGreen tagged TbFAN1 andTbSNM1 in bloodstream and procyclic form *T. brucei*.
2. *T. brucei* ICL factors respond in a temporal manner to ICL as well to trypanocidal nitroheterocyclic prodrugs, nifurtimox and benznidazole
3. *Tb*SNM1 is non-essential in *T. brucei* procyclic form parasites since it was possible to delete both copies. However, depletion of TbSNM1 caused growth defects in procyclic form parasites.
4. Reduction *Tb*SNM1 expression caused increased susceptibility to mechlorethamine, and MMS, while causing resistance to phleomycin. NO effect was seen with hydroxyurea or hydrogen peroxide. increased resistance to phleomycin
5. Ectopic expression of mNGTbSNM1 complements the growth defect of TbSNM1Δ in the presence of ICL (mechlorethamine).

CHAPTER 5: Unravelling the Role of the *T. brucei* BER Pathway.

5.1 Overview of base excision repair

The base excision repair (BER) pathway is an evolutionary conserved mechanism that functions to resolve small non-helix distorting base lesions from a genome (Robertson et al., 2009). This type of injury can be induced by endogenously generated metabolites or environmental toxins with these able to promote base oxidation, alkylation, or deamination (Lindahl, 1993; Kavli et al., 2007; Sedgwick et al., 2007). The role of the BER pathway is to detect/fix the myriad of modifications that can arise from the above reactions (Figure 5.1), damage that if unresolved can promote incorrect base-pairing leading to mutation or cause DNA breaks during replication.

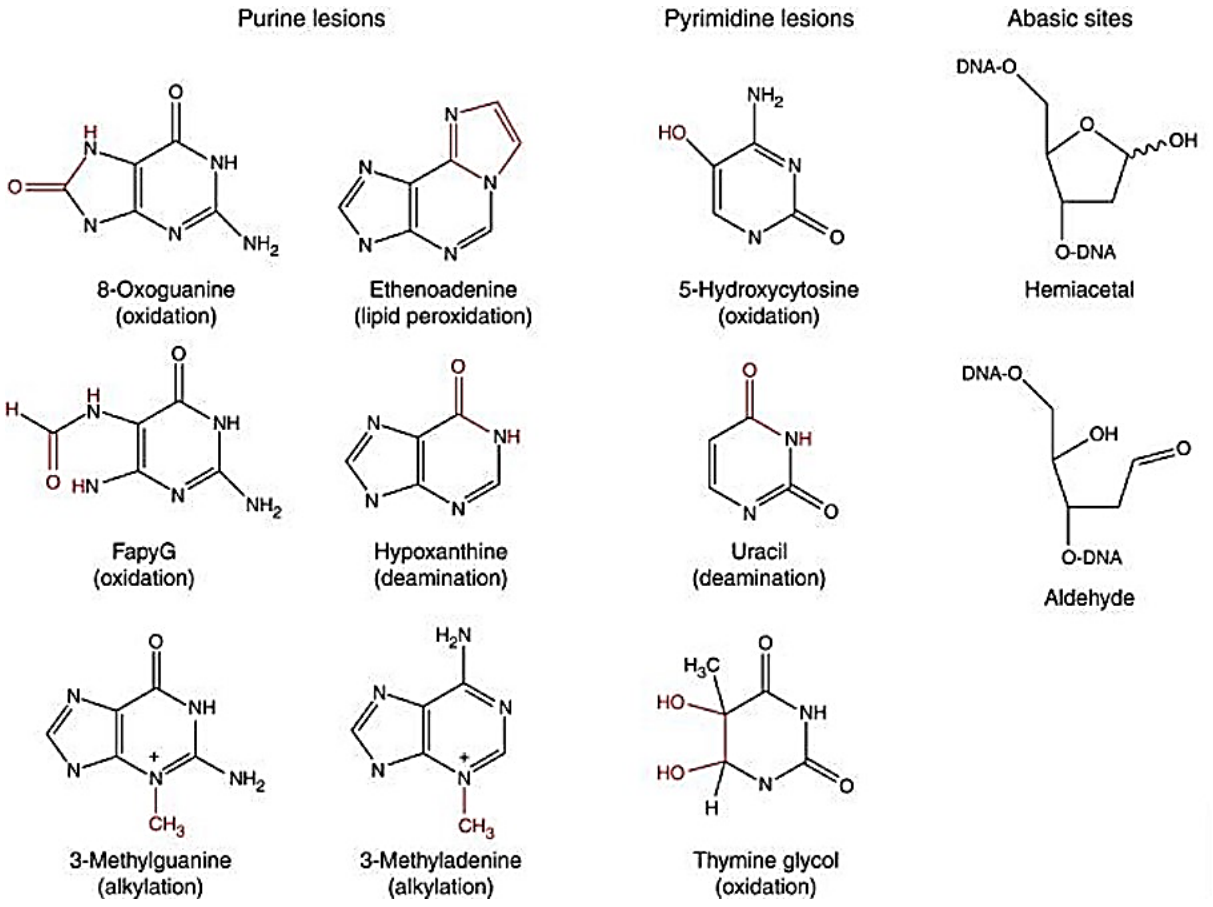


Fig 5.1. Examples of various base lesions recognised by the BER pathway.
Image taken from Krokan & Bjørås, 2013.

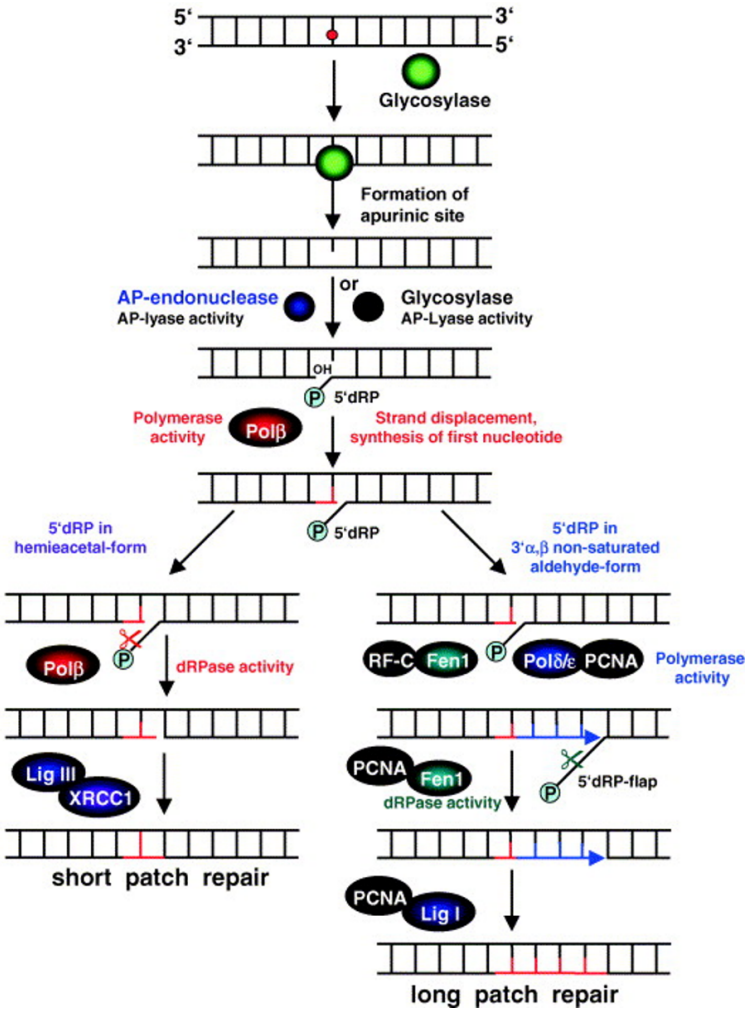


Fig 5.2. Mechanism of base excision repair (BER).

Recognition of the DNA lesion occurs by DNA glycosylase that remove the damaged base by hydrolysing the *N*-glycosidic bond. The resulting apurinic/apyrimidinic (AP) site is processed by AP endonucleases leading to a nick in the sugar-phosphate backbone of DNA. Repair of this gap is performed via the short- or long-patch BER pathways involving the concerted action of DNA polymerases and DNA ligases. FEN1, a flap endonuclease, plays a key role in the long-patch BER mechanism. Image taken from (Christmann et al., 2003).

The BER pathway (Figure 5.2) is initiated by DNA glycosylases (Robertson et al 2009). Each organism expresses variants of this enzyme to counteract the various modified bases that can be generated within a genome with some having an affinity for specific substrates while others possess overlapping activities (Genois et al., 2014). The DNA glycosylase function to recognize the lesion before catalyzing the hydrolytic cleavage of the *N*-glycosidic bond to remove the damaged base from the DNA strand, leaving an apurinic/apyrimidinic (AP) site (McCullough et al., 1999; Jacobs & Schär, 2012). Next, an AP endonuclease (e.g., APE1 or APE2) cleaves the DNA's sugar phosphate backbone immediately 5' to the AP site, generating a single-strand break with a 3'-OH group adjacent to a 5'-deoxyribose phosphate (Connor & Laval, 1989; Robson & Hickson, 1991). Two distinct classes of DNA glycosylase operate in BER (Krokan et al., 1997). Monofunctional DNA glycosylases such as MPG, UNG-1, UNG2, SMUG, MBD4, TDG, MYH, possess only glycosylase activity and participate in the initial excision step of the modified base from the DNA strand. In contrast, bifunctional DNA glycosylases including OGG1, NTH1, NEIL-

1, NEIL-2, also possess an AP lyase activity such that in addition to the base excision (glycosylase activity), they mediate cleavage of the phosphodiester bond of DNA to create a single-strand break without the need for an AP endonuclease.

In the final steps, DNA polymerases and DNA ligases are recruited to the damage site and mediate DNA gap filling and strand-sealing, respectively. Different mechanisms can operate at this point leading to two distinct BER sub-pathways referred to as the short-patch BER or long-patch BER (Frosina et al., 1996; Kubota et al., 1996). In the short-patch repair mechanism, the single-strand break is filled by the action of DNA polymerase β with DNA ligase III, accompanied by the scaffolding protein XRCC1, reforming the DNA's sugar-phosphate backbone (Caldecott, 2019; Caldecott et al., 1995; Singhal & Wilson, 1993). In the long-patch mechanism, which may operate when there is an ATP shortage (Petermann et al., 2003), replication factor C interacts with the single-strand break to recruit PCNA and DNA polymerases δ or β (Matsumoto et al., 1994). The polymerases then add several nucleotides (between 2 to 13) starting at the break site and in doing so displaces the existing original nucleotides resulting in an oligonucleotide overhang called the flap. The FLAP endonuclease FEN1 is recruited which then removes the overhang before DNA ligase I seals the nick (Kim & Wilson III, 2012; Klungland & Lindahl, 1997; Stucki et al., 1998).

Sequence analysis has shown *T. brucei* has potential to express many components of the mammalian BER pathway (Table 5.1) (Genois et al., 2014). Of these, functional studies have shown that TbUNG functions as a DNA glycosylase that aids in removal of uracil from a DNA template (Fárez-Vidal et al., 2001) while TbAPE1 plays a key role in AP site processing (Charret et al 2012)., Work on *T. cruzi* additionally demonstrates that TcNTH1 possesses an AP lyase activity (Ormeno et al., 2016). Overexpression of APE1 in *T. cruzi* confers cellular resistance to oxidative DNA challenge (Charret et al., 2012; Valenzuela et al., 2012). Whether TbNTH1 also possesses such activity is yet to be determined. The sub-cellular localization of TbAPE1, TbMYH and TbNTH1 have all been determined through the TrypTag project (Dean et al., 2017), although TbMYH and TbNTH1 proteins are yet to be characterized. The T. cruzi orthologs, TcNTH1 as described above lacks glycosylase activity (Ormeno et al., 2016), Whereas TcMYH catalyse the removal of adenine in 8-oxoG:A base lesions in the nuclear and mitochondrial genomes of these parasites and can additionally complement mutY gene in deficient bacteria (Kunrath-Lima et al., 2017). Polymerases involved in BER have also been identified and characterised in trypanosomes, for example BER polymerase beta (Polβ) have been described in *T. brucei* as well as in *T. cruzi*,

with both proteins also possessing 5'dRP lyase activity Additionally, Polβ localizes in the kinetoplasts of these parasites (Saxowsky et al., 2003; Lopes et al., 2008; Maldonado et al., 2021), in contrast to higher eukaryotes where it is found predominantly in the nucleus (Foster & Gurney, 1976). Interestingly, DNA Polβ in leishmania (*L. infantum*), localizes in the nucleus similar to their higher eukaryotic counterparts (Taladriz et al., 2003). Analysis of the Tritryps genomes however, revealed the absence of DNA polymerase lambda (Polλ), and mu (Polμ) (El-Sayed et al., 2005).

Here, we report a series of experiments aimed at further understanding the role of TbAPE1 within *T. brucei* as well as unravelling the function of the monofunctional DNA glycosylase TbMYH and the bifunctional DNA glycosylase TbNTH1. Using a PCR-based CRISPR/Cas9 genome editing strategy (Beneke et al 2017), bloodstream and procyclic form parasites expressing tagged versions of each enzyme were made with attempts made to use this approach to construct procyclic form *T. brucei* lines null for Tb*ape1*, Tb*myh* or Tb*nth1*. The resultant lines were employed to assess protein subcellular location in parasite stages, for any alteration of protein expression following treatment with DNA damaging agents of trypanocidal nitroheterocycles, and in the phenotyping (growth and susceptibility studies) of insect form parasites lacking Tb*ape1* or heterozygous for Tb*myh* or Tb*nth1*.

Enzyme	Function	*H. sapiens**	*T. brucei*#
UNG/UDG2	Uracil DNA N-glycosylase/ Uracil DNA glycosylase 2	M_003362/NM_080911	Tb927.10.13970
SMUG1	Single-strand-selective Monofunctional Uracil-DNA glycosylase 1	NM_014311	X
OGG1	8-Oxo-guanine glycosylase 1	NM_016821	Tb927.4.2480
TDG	Thymine DNA glycosylase	NM_003211	X
MBD4	Methyl-CpG-binding domain 4 DNA glycosylase	NM_003925	X
MYH	**Mut Y homolog DNA glycosylase**	**NM_012222**	**Tb927.11.11440**
NTH1	**Endonuclease three homolog 1 DNA glycosylase**	**NM_002528**	**Tb927.11.12090**
MPG	Methyl purine DNA glycosylase	NM_002434	X
NEIL1	Nei-like DNA glycosylase1	NM_024608	X
NEIL2	Nei-like DNA glycosylase 2	NM_145043	X
NEIL3	Nei-like DNA glycosylase 3	NM_018248	X
APE1/APE2	**AP endonuclease 1**/AP endonuclease 2	**M_001641**/NM_014481	**Tb927.8.5510**
LIGI/LIGIII	DNA ligase involved: I - mainly in long-patch BER; III - only short-patch BER	M_000234/NM_013975	Tb927.6.4780
Polβ	DNA polymerase involved in long and short-patch BER	NM_002690	Tb927.5.2780
Polε	DNA polymerase involved in long-patch BER	NM_006231	Tb09.211.1820
Polδ	DNA polymerase involved in long-patch BER	NM_002691	Tb927.2.1800
XRCC1	X-ray repair cross-complementing protein 1	NM_006297	X
PCNA	Proliferating cell nuclear antigen	NM_182649	Tb927.9.5190
RF-C	Replication factor C	NM_002913	Tb927.11.5650
FEN1	Flap endonuclease 1	NM_004111	Tb927.3.830

Table 5.1. Components of the human and *T. brucei* base excision repair pathway.
Evaluation of sequence data held at TriTrypDB revealed that *T. brucei* has potential to express orthologues of most components of the human base excision repair pathway. * and # represent gene sequence identifiers from the NCBI or TriTrypDB respectively. X represents BER sequences present in humans but not found in *T. brucei*. Enzymes highlighted in red correspond to those analysed in the remainder of this chapter. Note the *T. brucei* ID numbers for: 1. UNG & UDG2, 2. APE1 & APE2 and 3. LIGI & LIGIII are the same and reflect that unlike the human counterparts, this parasite has only one form of enzyme Modified from Genois et al., 2014.

5.2. Informatic analysis of components of the *T. brucei* base excision repair pathway

5.2.1. Informatic analysis of TbAPE1

A *T. brucei* gene coding for an apurinic/apyrimidinic endonuclease was identified from the published literature (Charret et al., 2012; Genois et al., 2014) and data held on the TriTryp database. This 1191 bp long ORF (Tb927.8.5510), designated as Tb*ape1*, is found on chromosome 8 of the *T. b. brucei* TREU927 genome and has potential to code for a 44 kDa protein, designated as TbAPE1. Pairwise alignment revealed Tb*ape1* was 50% identical to its *T. cruzi* CL Brener Esmeraldo-like haplotype (TcCLB.507083.30) orthologue, 56 to 58% identical to the equivalent genes from *C. fasciculata* (CFAC1_120014800), *L. donovani* (LdBPK.16.2.000680) and *L. major* (LmjF.16.0680), and between 39 to 47 % identical to its *S. cerevisiae* (YOL043C), *H. sapiens* (NM_001641) and *A. thaliana* (NM_129709) counterparts. Analysis of the primary amino acid sequence revealed TbAPE1 contained a single domain, namely the Exo_endo_phos domain (PF03372) that is sited between residues 97 to 394) (Figure 5.3). Alignment of this region with the human and yeast counterparts coupled with HMMR analysis identified several conserved motifs in the trypanosomal enzyme's Exo_endo_phos domain that correspond to this region's HMM logo. These motifs were: 1. (S/T)(W/Y/F)N(V/L/I), 2. (D/Y)(V/I/L)(L/V/T).(L/F/I/V)QE, 3. GD(L/V)N, 4. R(L/I/V)D(R/Y/H)(I/V/A)(L/F/I) and 5. SDH and were ordered as listed. In the human orthologue, several residues within these sites, including the asparagine residue (N212) in the GD(L/V)N motif, function in metal ion binding/enzyme activity (Beernink et al., 2001; Rothwell & Hickson, 1996). Mutational studies on the asparagine residue (N287) in TbAPE1's GD(L/V)N motif has shown it plays a crucial role in mediating AP endonuclease activity although whether involves co-factor binding has not been reported (Charret et al., 2012).

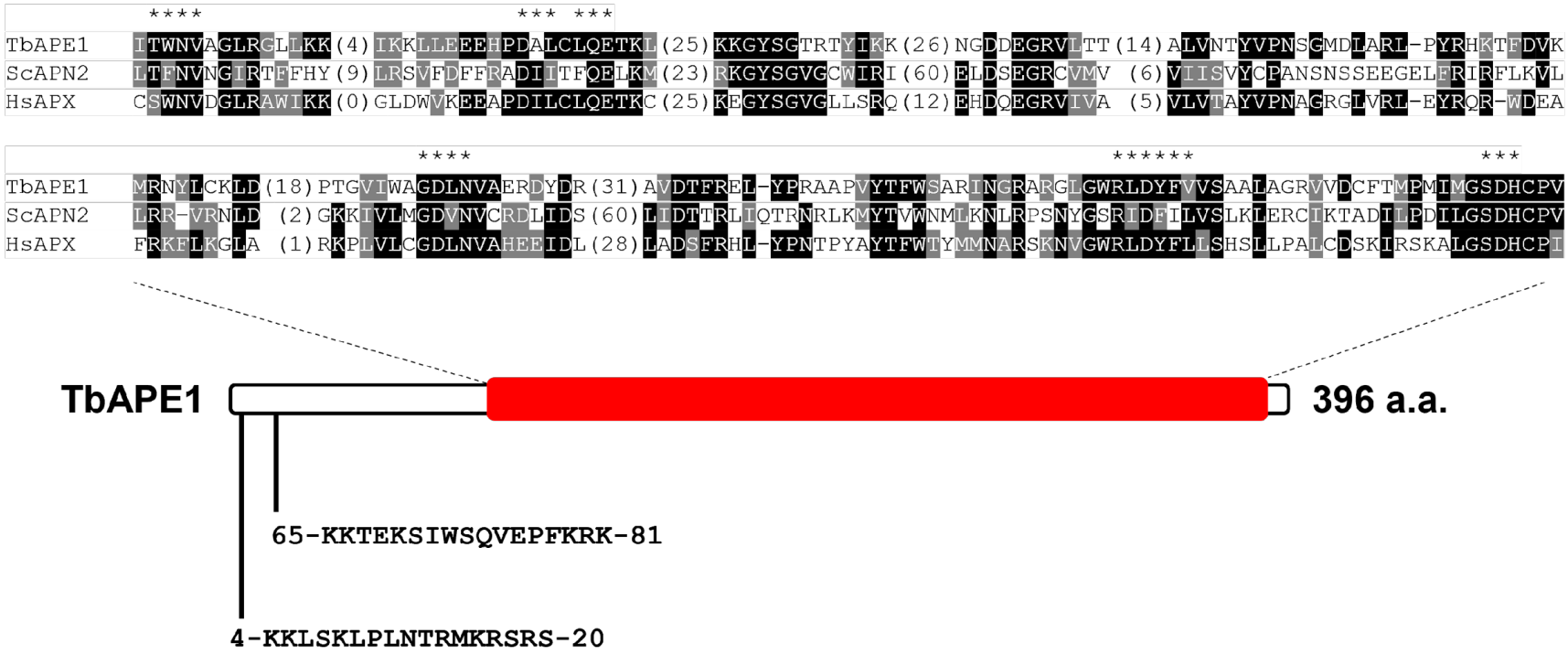


Fig 5.3. Informatic analysis of TbAPE1.
The sequence corresponding to the Exo_endo_phos (PF03372; residues 97-394; red box) domain of TbAPE1 (XP_847381) were aligned with the equivalent regions from other apurinic/apyrimidinic endonucleases expressed by S. cerevisiae (ScAPN2; AJP82740) and H. sapiens (HsAPX; BAA02633). Residues that are identical in two of the sequences and conserved substitutions are highlighted in black and grey, respectively. Residues in the Exo_endo_phos domain that are marked with an asterisk constitute parts of this regions HMM logo. Two potential nuclear localisation signals (Robbins & Dingwall consensus sequences) found towards the amino terminal of TbAPE1 are shown**.**

Using the Phyre2 package, a 3D prediction of the full-length TbAPE1 based on the *L. major* orthologue (2J63; Vidal et al., 2007) was generated (Figure 5.4). The amino terminal region (residues 1-62) in the trypanosomal enzyme had no homology to its *Leishmania* (or any other) counterpart with the model having high confidence in the sequence covering the Exo_endo_phos domain. The structure for this region is predicted to form a typical four-layered α/β-sandwich fold as noted for other AP endonucleases (Gorman et al., 1997; Mol et al., 1995, 2000; Vidal et al., 2007). This consisted of two curved but parallel β-sheets each made up of six β-strands surrounded by helices. By comparison to the human counterpart, an approximate location for any metal ion binding site could be discerned with this position in other AP endonucleases also being designated as the DNA binding site and catalytic centre.

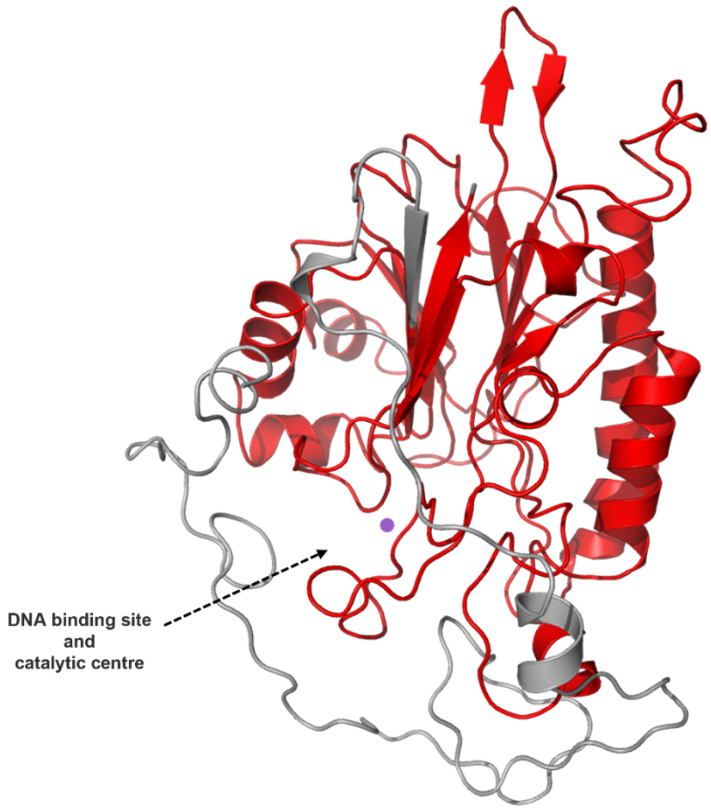


Fig 5.4. 3D prediction model of TbAPE1.
A full-length 3D prediction of TbAPE1 (XP_847381) was constructed using Phyre2 (Kelley et al. 2015) and visualized using PyMOL: There was >90% confidence over 84% of residues modelled. The position of the Exo_endo_phos (PF03372) (red) domain is shown. Non-domain regions are in grey. APE1s are metal binding enzymes although the identity (Mn2+; Mg2+) and number of metal ions bound plus the residues at the binding site(s) remains controversial (Oezguen et al 2007; Tsutakawa et al 2013; He et al 2014). Based on structures generated by Freudenthal et al 2015, the DNA binding site and catalytic centre along with the region of TbAPE1 where a metal ion co-factor(s) is believed to bind (purple spheres) are shown.

5.2.2. Informatic analysis of TbMYH

A *T. brucei* ORF coding for a putative MYH was identified following literature searches (Genois et al., 2014) and BLAST analysis of data held on the TriTryp database using the *T. cruzi* orthologue (Furtado et al., 2012) as bait. The 1,533 bp long *T. brucei* gene (Tb927.11.11440), designated *Tbmyh*, is found on chromosome 11 of the *T. b brucei* TREU927 genome and has potential to code for a 57 kDa protein (designated TbMYH). Pairwise alignment revealed Tb*myh* was 55% identical to the orthologue expressed by *T. cruzi* CL Brener Esmeraldo-like haplotype (TcCLB.511803.20), 50 to 52 % identical to the *myh* genes from *C. fasciculata* (CFAC1_300035500), *L. donovani* (LdBPK_282290.1) and *L. major* (LmjF.28.2140), and 47 % identical to its *H. sapiens* (NM_012222) counterpart: *S. cerevisiae* does not express a MYH orthologue. Domain searches revealed that TbMYH contains a HhH-GPD domain (PF00730) located between residues 112 to 236 and a nucleoside diphosphate linked to X (NUDIX; accession number PF14815) domain towards its carboxyl terminal between residues 342 to 424 (Figure 5.5A). 3D predictions of the full-length trypanosomal enzyme based on mouse (7EF8) and human (1RRQ) MUTY structures (Fromme & Verdine, 2004; Nakamura et al., 2021) showed that the two domains formed distinct structural units (Figure. 5.6). The sequence containing the HhH-GPD domain forms a 6-helix barrel structure while the NUDIX region consists of a four-stranded, mixed β-sheet sandwiched between two α-helices. Analysis of the region between the two domains identified four cysteine residues (C274; C282; C285; C291) that form a connector sequence which coordinates binding of a 4Fe-4S cluster (Figure 5.5A and B) (Fromme & Verdine, 2004; Luncsford et al., 2011; 2013). In the mammalian MUTY counterparts, this connector region (called the interdomain connector) also contains amino acids that facilitate bind Zn^{2+} ions (Engstrom et al., 2014) although alignment indicates the equivalent residues are missing from TbMYH. By comparison with mammalian structures (Fromme & Verdine, 2004; Nakamura et al., 2021), the HhH-GPD domain and 4Fe-4S binding region represent TbMYH's catalytic centre while an interface between the HhH-GPD and NUDIX domains may represent the enzyme's DNA binding site.

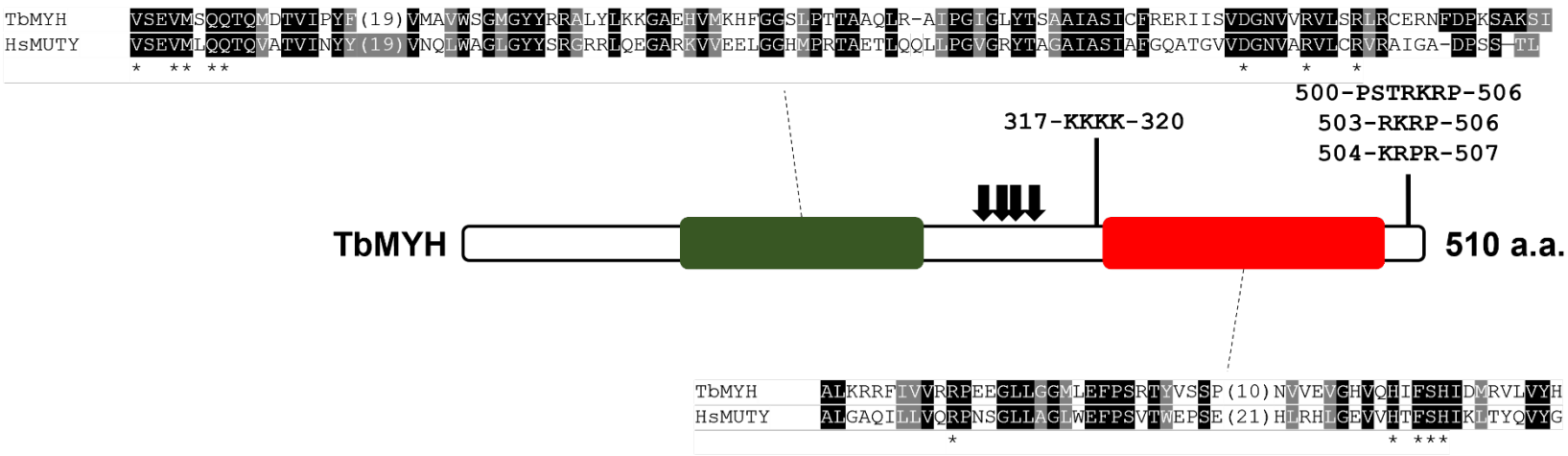


Fig 5.5 Informatic analysis of TbMYH.

Comparative sequence analysis of TbMYH and Hs MUTY. The sequence corresponding to the HhH-GPD (PF00730; residues 113-236; green box) and NUDIX_4 (nucleoside diphosphate linked to X; PF14815; residues 342-424; red box) domains of TbMYH (XP_829209) were aligned with the equivalent regions from MUTY, the orthologue expressed by *H. sapiens* (HsMUTY; NP_001041638). Residues that are identical or are conserved substitutions are highlighted in black and grey, respectively. The arrows correspond to 4 cysteine residues (C274; C282; C285; C291) that coordinate binding to Fe-S clusters (Luncsford et al 2011). Residues marked with an asterisk are characteristic of the HMM logo for the HhH-GPD or NUDIX_4 domains. The location of four potential nuclear localisation signals (three 4 pattern and one 7 pattern), three of which are towards the carboxyl terminal of TbMYH, are shown

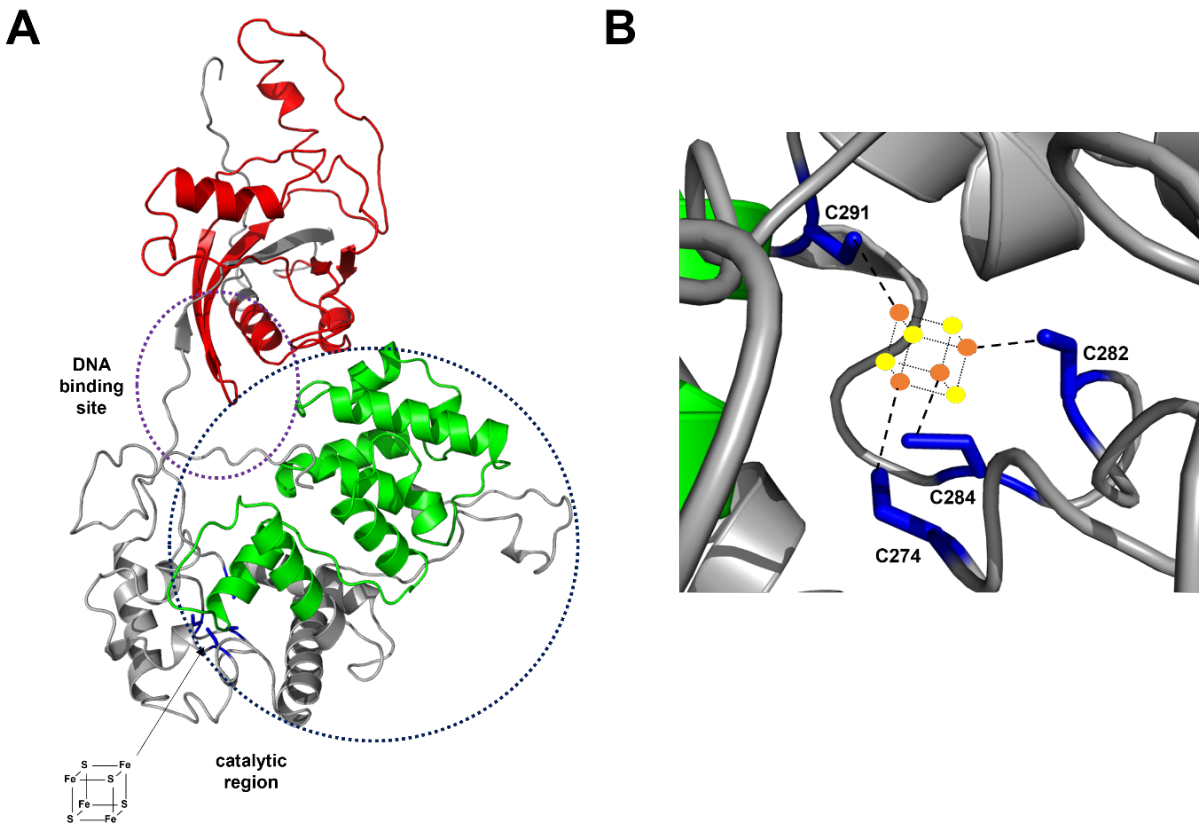


Fig 5.6. 3D prediction model of TbMYH.
A. A full-length 3D prediction of TbMYH (XP_829209) was constructed using Phyre2 (Kelley et al. 2015) and visualized using PyMOL: There was >90% confidence over 76 of residues modelled. The position of the HhH-GPD (PF00730; green box) and NUDIX_4 (PF14815; red box) domains are shown. Non-domain regions are in grey. The TbMYH orthologue is postulated to bind a 4Fe-4S cluster *via* four cysteine residues (C274; C282; C285; C291, in blue) present in a connector sequence found between the HhH-GPD and NUDIX_4 domains ((Luncsford et al., 2011)). The HhH-GPD and connector region represents this enzyme's catalytic centre. Based on mammalian MUTY orthologues (Fromme & Verdine, 2004; Nakamura et al., 2021), a potential DNA binding site is shown and located at the interface between HhH-GPD and NUDIX_4 domains.
B. A zoomed view showing four cysteines (C274; C282; C285; C291) found in the interdomain chain of TbMYH postulated to be involved in binding a 4Fe-4S cluster. Iron and sulphur atoms are shown in orange and yellow spheres, respectively.

5.2.3. Informatic analysis of TbNTH1

A *T. brucei* gene coding for endonuclease III was identified based on homology with its T. *cruzi* orthologue (Ormeño et al., 2016), literature searches (Genois et al., 2014) and following BLAST analysis of data held at the TriTryp database using the *H. sapiens* NTH1 (P78549) and *S. cerevisiae* NTG2 (EGA76950) as bait. The resultant 780 bp long ORF (Tb927.11.12090), designated as Tb*nth1*, is located on chromosome 11 of the *T. b. brucei* TREU927 genome and has potential to code for a 29 kDa protein, designated here as TbNTH1. Pairwise alignment revealed Tb*nth1* was 67% identical to the *T. cruzi* CL Brener Esmeraldo-like haplotype (TcCLB.504005.10) orthologue,

58 and 61 % identical to the *L. donovani* (LdBPK.09.2.000070) and *L. major* (LmjF.09.0050) *nth1* genes, 50 to 58% identical to the three *nth1* ORFs in the *C. fasciculata* (CFAC1_100006900; CFAC1_100007000; CFAC1_100007100) genome, and 39 to 43 % identical to its *S. cerevisiae* (YOL043C), *H. sapiens* (NM_002528) and *A. thaliana* (NM_128702) counterparts. Domain searches revealed TbNTH1 contained a single domain, namely a HhH-GPD domain (PF00730) located at the centre of the protein (between residues 53 to 189) within which a Helix-Hairpin-Helix motif (HhH; PF00633; residues 119-145) is found (Figure 5.8). Towards the carboxyl terminal four cysteine residues (C209; C216; C219; C225) are present that by analogy to the mammalian counterpart coordinate binding to an 4Fe-4S cluster (Carroll et al., 2021). 3D prediction of full length TbNTH1 based on the structure determined for the human enzyme (7RDT) (Carroll et al., 2021) revealed that the trypanosomal enzyme comprises two globular helical regions joined by two linkers. One globular region contains a six-helical bundle domain that includes the HhH DNA-binding motif while the second which is composed of part of the HhH-GPD domain, and the carboxyl terminal contains the 4Fe- 4S binding site (Figure 5.9). Intriguingly, the *T. cruzi* orthologue possess many sequence traits associated with a NTH1 bifunctional DNA glycosylase although it has been reported to only function as an AP endonuclease (Ormeño et al., 2016). Whether the *T. brucei* enzyme displays this trait has yet to be shown.

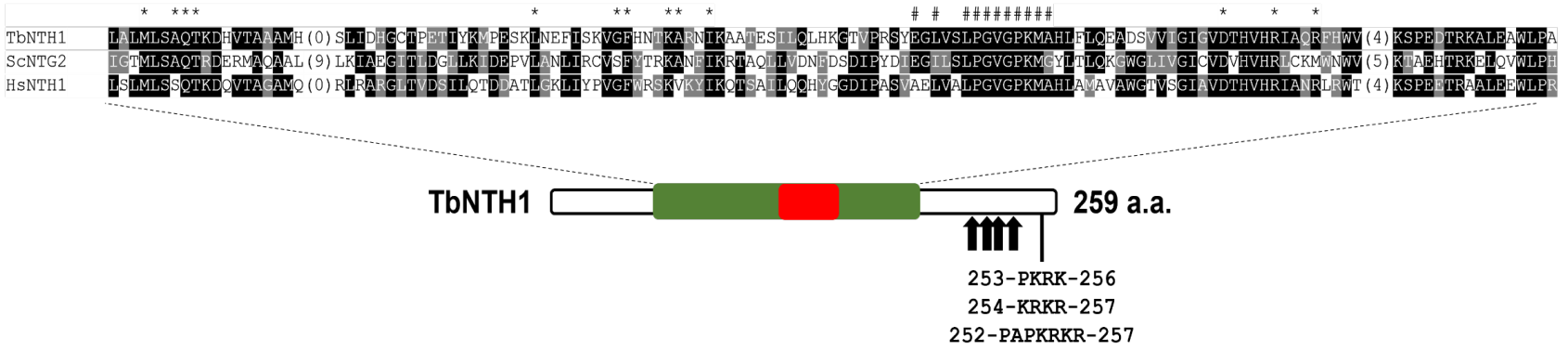


Fig 5.7. Informatic analysis of TbNTH1.

The sequence corresponding to the helix-hairpin-helix GPD (HhH-GPD; PF00730; residues 53-189; green box) domain that contains a Helix-hairpin-helix motif (HHH; PF00633; residues 119-145; red box) of TbNTH1 (XP_829271) were aligned with the equivalent regions from endonuclease IIIs expressed by *S. cerevisiae* (ScNTG2; EGA76950) and *H. sapiens* (HsNTH1; P78549) Residues that are identical or are conserved substitutions are highlighted in black and grey, respectively. The arrows correspond to four cysteine residues (C209; C216; C219; C225) that coordinate binding to Fe-S clusters The location of three potential nuclear localisation signals (two 4 pattern and one 7 pattern) found towards the carboxyl terminal of TbNTH1 are shown. Residues marked with * or # are characteristic of the HMM logo for the HhH-GPD domain or HHH motif, respectively.

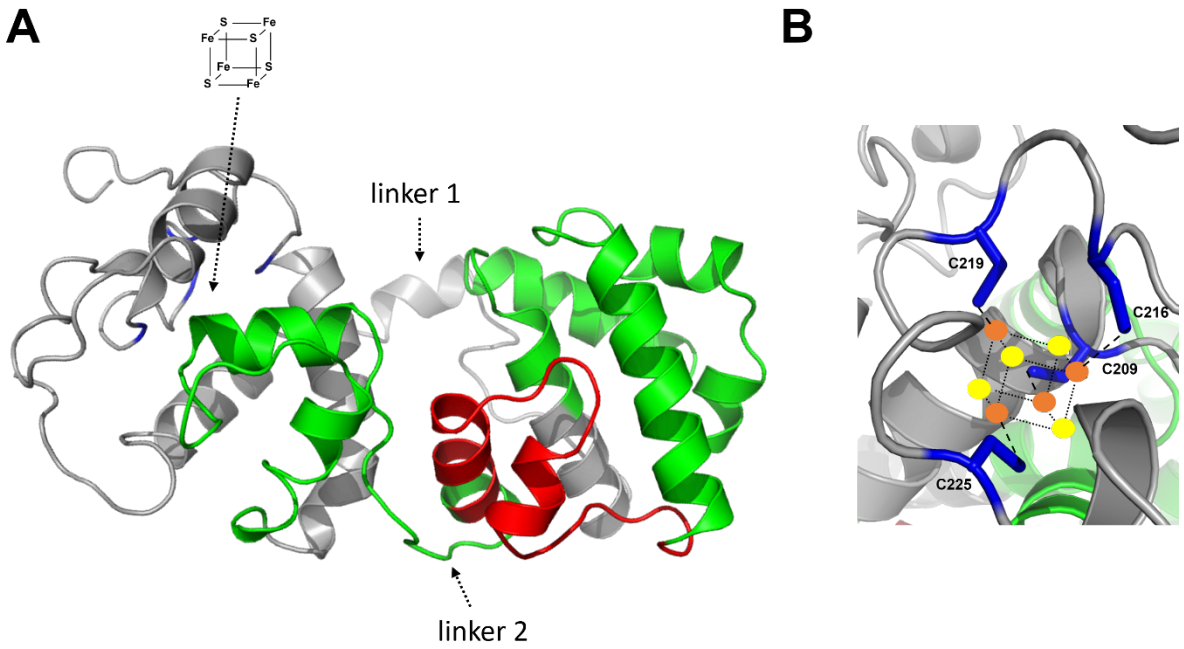


Fig 5.8. 3D prediction model of TbNTH1.
A. A full-length 3D prediction of TbNTH1 (XP_829271) was constructed using Phyre2 (Kelley et al. 2015) and visualized using PyMOL: There was >90% confidence over 86% of residues modelled. The position of the HhH-GPD domain (PF00730; green) and HHH motif (PF00633; red) are shown. Non-domain regions are in grey. TbNTH1 orthologues bind a 4Fe-4S cluster *via* four cysteine residues present towards the carboxyl terminal The position of these amino acids (C209; C216; C219; C225) (in blue) and the region where the 4Fe-4S cluster is postulated to be sited in the *T. brucei* enzyme are shown.
B. A zoomed view showing four cysteines (C209; C216; C219; C225) found towards the carboxyl terminal of TbNTH1 postulated to be involved in binding a 4Fe-4S cluster. Iron and sulphur atoms are shown in orange and yellow spheres, respectively.

5.3 Editing the Tb*ape1, Tbmyh* and *Tbnth1* loci in *T. brucei.*

Based on localisation prediction algorithms (PSORTII and WoLFPSORT), TbAPE1, TbMYH and TbNTH1 are predicted with 88, 70 and 60 % certainty to be nuclear proteins with in-silico analysis identifying nuclear localisation signals in each protein sequence (see Figures 5.4, 5.6 and 5.8): The number and type of each signal for each protein is given in the appropriate figure. For APE1, this nuclear localisation has been confirmed experimentally in mammalian and insect form parasites by immunofluorescence using an antibody against the trypanosomal enzyme (Charret et al 2012). In the case of TbNTH1, and as part of the TrypTag project (Dean et al 2017), expression of the glycosylase at its amino or carboxyl terminal with sequences coding for mNeonGreen in procyclic form *T. brucei* resulted in an intense signal throughout the nucleus. When this approach was applied to TbMYH, amino or carboxyl terminal tagged protein gave a cytoplasmic or mitochondrial location, respectively.

To confirm/assess the localisations of the three BER proteins in bloodstream form and/or insect form *T. brucei*, the CRISPR/Cas9-mediated tagging of the endogenous Tb*ape1,* Tb*myh,* or Tb*nth1* loci at their 5' or 3' end with nucleotide sequences coding for the fluorescent protein mNeonGreen was performed (see Figure 3.13) using appropriate primer combinations (Table 5.2) to generate the sgRNA and targeting fragments. The cells lines generated from these studies are listed in Table 5.3: note not all the desired lines were made despite multiple electroporations.

	Primer combinations	
	sgRNA	**Targeting fragment**
Tb*ape1*		
5' tagging	G00/5'sgRNA-Tbape1	Tbape-UFP1/Tbape-DRF1
3' tagging	G00/3'sgRNA-Tape11	Tbape1-DFP1/Tbape11-DRTAGP
Tb*myh*		
5' tagging	G00/5'sgRNA-Tbmyh	Tbmyh-UFP1/Tbmyh-DRF1
3' tagging	G00/3'sgRNA-Tbmyh	Tbrad51-DFP1/Tbmyh-DRTAGP
Tb*nth1*		
5' tagging	G00/5'sgRNA-Tbnth1	Tbmyh-UFP1/Tbmyh-Tbnth1
3' tagging	G00/3'sgRNA-Tbnth1	Tbrad51-DFP1/Tbmyh-Tbnth1

Table 5.2. Primer combinations used in tagging Tb*ape1*, Tb*myh* & Tb*nth1*.
The primer combinations used to generate amplicons corresponding to the short guide RNA (sgRNA) and targeting fragment used in tagging the 5' or 3' ends of the endogenous Tb*ape1*, Tb*myh* or Tb*nth1* loci with a sequences coding for the fluorescent protein mNEONGREEN. The sequence of each primer can be found in Appendix 1.

tagged gene	**Bloodstream form**		**Procyclic form**	
	amino	**carboxyl**	**amino**	**carboxyl**
Tb*ape1*	✓	✓	✓	✓
Tb*myh*	X	✓	✓	✓
Tb*nth1*	✓	✓	X	✓

Table 5.3. *T. brucei* lines expressing tagged base excision repair factors.
Attempts to tag the 5' or 3' region of Tb*ape1*, Tb*myh* or Tb*nth1* in bloodstream or insect form *T. brucei* were performed although not all were successful. A tick (✓) corresponds to recombinant cells generated from in this analysis while X represents situations where the desired line was not obtained.

To confirm that the desired fusion gene was being expressed, DNA amplification reactions using cDNA templates were performed with primer combinations that span the gene fusion event: the gene fusion could only be generated if the introduced targeting fragments had integrated at the correct loci. In all cases where recombinant parasites were analysed, the expectedly sized amplicons were observed. For example, when using primer combinations intended to detect the *mNeonGreen*-Tb*ape1* fusion, PCR generated a 600 bp band, the expected size, from nucleic acid

(cDNA) originating from BSF and PCF recombinant parasites with no band(s) observed in templates from wild type cells (Figure 5.9B). A similar outcome was noted for Tb*ape1-mNeonGreen* expressing BSF or PCF *T. brucei* although in this case, a band of ~550 bp was detected only in templates from the recombinant cells (Figure 5.9D).

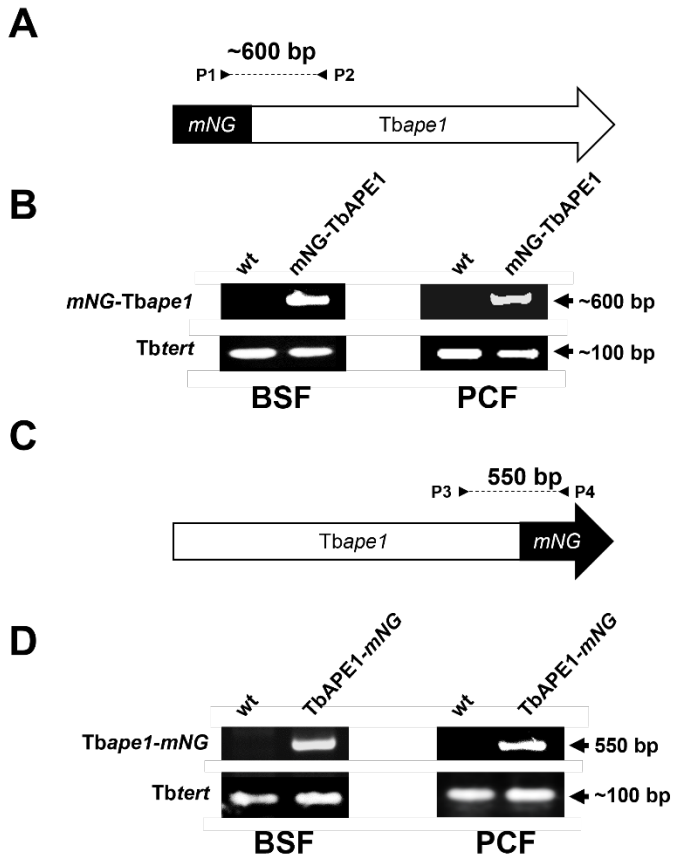


Fig 5.9. Validation of *T. brucei* lines expressing mNeonGreen tagged TbAPE1.

A & C. Schematic showing integration of *mNeonGreen* (*mNG*) into the *T. brucei* genome at the 5' (**A**) or 3' (**C**) region of Tb*ape1*. P1 to P4 indicate the primers used in the obtention of the validation amplicons.

B & D. Amplicons (in kb) diagnostic of bloodstream form (BSF) or procyclic form (PCF) *T. brucei* expressing *mNG*-Tb*ape1* (**B**) or Tb*ape1-mNG* (**D**) were generated from cDNA templates derived from total RNA extracted from wild type (wt) or recombinant parasites. The integrity of DNA samples was evaluated by amplification of a 0.1 kb control fragment, Tb*tert*. The amplified fragment Is indicated (*mNG*-Tb*APE1*; Tb*ape1-mNG* or Tb*tert*). The primer sequences and combinations used for each amplification are listed in Appendix 2.

For cells expressing tagged TbMYH (Figure 5.10), PCR using primer combinations intended to detect the *mNeonGreen*-Tb*myh* or the Tb*myh -mNeonGreen* fusions generated bands of the expected size of 600 and 650 bp, respectively, only from the genetically modified lines.

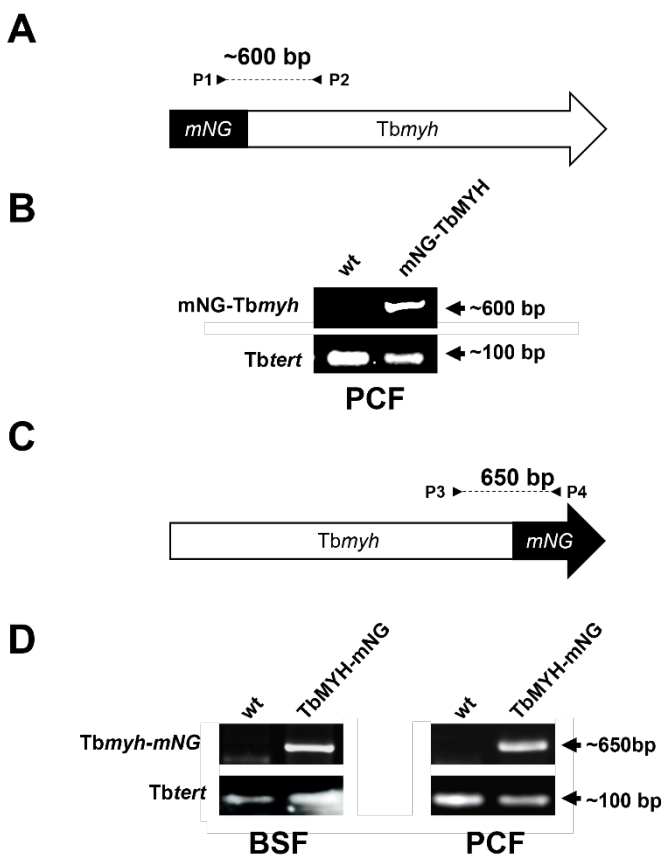


Fig 5.10 Validation of *T. brucei* lines expressing mNeonGreen tagged TbMYH.

A & C. Schematic showing integration of *mNeonGreen* (*mNG*) into the *T. brucei* genome at the 5' (**A**) or 3' (**C**) region of Tb*myh*. P1 to P4 indicate the primers used in the obtention of the validation amplicons.

B & D. Amplicons (in kb) diagnostic of bloodstream form (BSF) or procyclic form (PCF) *T. brucei* expressing *mNG*-Tb*myh* (**B**) or Tb*myh1-mNG* (**D**) were generated from cDNA templates derived from total RNA extracted from wild type (wt) or recombinant parasites. The integrity of DNA samples was evaluated by amplification of a 0.1 kb control fragment, Tb*tert*. The amplified fragment Is indicated (*mNG*-Tb*myh*; Tb*myh-mNG* or Tb*tert*). The primer sequences and combinations used for each amplification are listed in Appendix 2.

A similar outcome was noted when validating BSF and/or PCF parasites expressing *mNeonGreen*-Tb*nth1* or the Tb*nth1-mNeonGreen* (Figure 5.11). Templates derived from cells expressing the former gene fusion generated a band the expected 620 bp size while for in the latter case, a band of 650 bp was observed

For each gene fusion, the confirmatory cDNA-derived amplicons were sequenced (data not shown) and the Tb*ape1*, Tb*myh* or Tb*nth1* sequence shown to be in-frame with that coding for mNeonGreen. Also, and for all recombinant and wild type samples tested, control PCRs were performed to amplify a 100 bp fragment from the Tb*tert* transcript. In all case (Figures 5.9-5.11), a band of this size (100 bp) was observed confirming the integrity of all nucleic acid templates

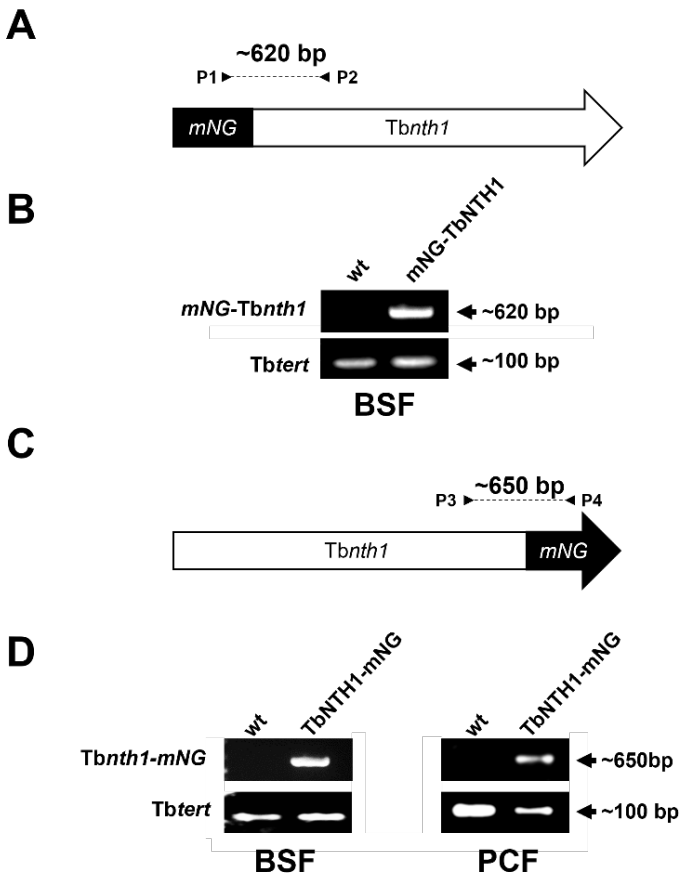


Fig 5.11. Validation of *T. brucei* lines expressing mNeonGreen tagged TbNTH1.

A & C. Schematic showing integration of *mNeonGreen* (*mNG*) into the *T. brucei* genome at the 5' (**A**) or 3' (**C**) region of Tb*nth1*. P1 to P4 indicate the primers used in the obtention of the validation amplicons.

B & D. Amplicons (in kb) diagnostic of bloodstream form (BSF) or procyclic form (PCF) *T. brucei* expressing *mNG*-Tb*nth1* (**B**) or Tb*nth1-mNG* (**D**) were generated from cDNA templates derived from total RNA extracted from wild type (wt) or recombinant parasites. The integrity of DNA samples was evaluated by amplification of a 0.1 kb control fragment, Tb*tert*. The amplified fragment Is indicated (*mNG*-Tb*nth1*; Tb*nth1-mNG* or Tb*tert*). The primer sequences and combinations used for each amplification are listed in Appendix 2.

5.4 Localisation of TbAPE1, TbMYH and TbNTH1 in *T. brucei*.

To evaluate localisation, BSF or PCF *T. brucei* expressing mNeonGreen tagged TbAPE1, TbMYH or TbNTH1 were fixed, their DNA stained with DAPI then imaged by fluorescent microscopy. In relation to the cell cycle, most (80-85%) trypanosomes from both life cycle forms exhibited a 1N1K profile showing they were in the G1/S phase, around 10 -15 % possessed a 1N2K ratio indicative of cells in G2/M phase while about 5 % had a 2N2K configuration and were therefore the post-M stage. When assessing TbAPE1 location, a fluorescence signal was noted in > 90% of recombinants. For BSF parasites from all life cycle stages that express *mNeonGreen*-Tb*ape1* or Tb*ape1-mNeonGreen*, punctate foci were observed throughout the nucleoplasm (Figure 5.12A & C) although for the latter fusion this was accompanied by an intense signal present in the non-DAPI-staining nucleolus (Figure 5.12C). For PCF *T. brucei*, an intense nucleolar signal was observed accompanied by a faint fluorescence elsewhere in the nucleus (Figure 5.12B and D). Because of the intense nature of the nucleolar foci, any potential punctate pattern throughout the nucleoplasm could not be discerned.

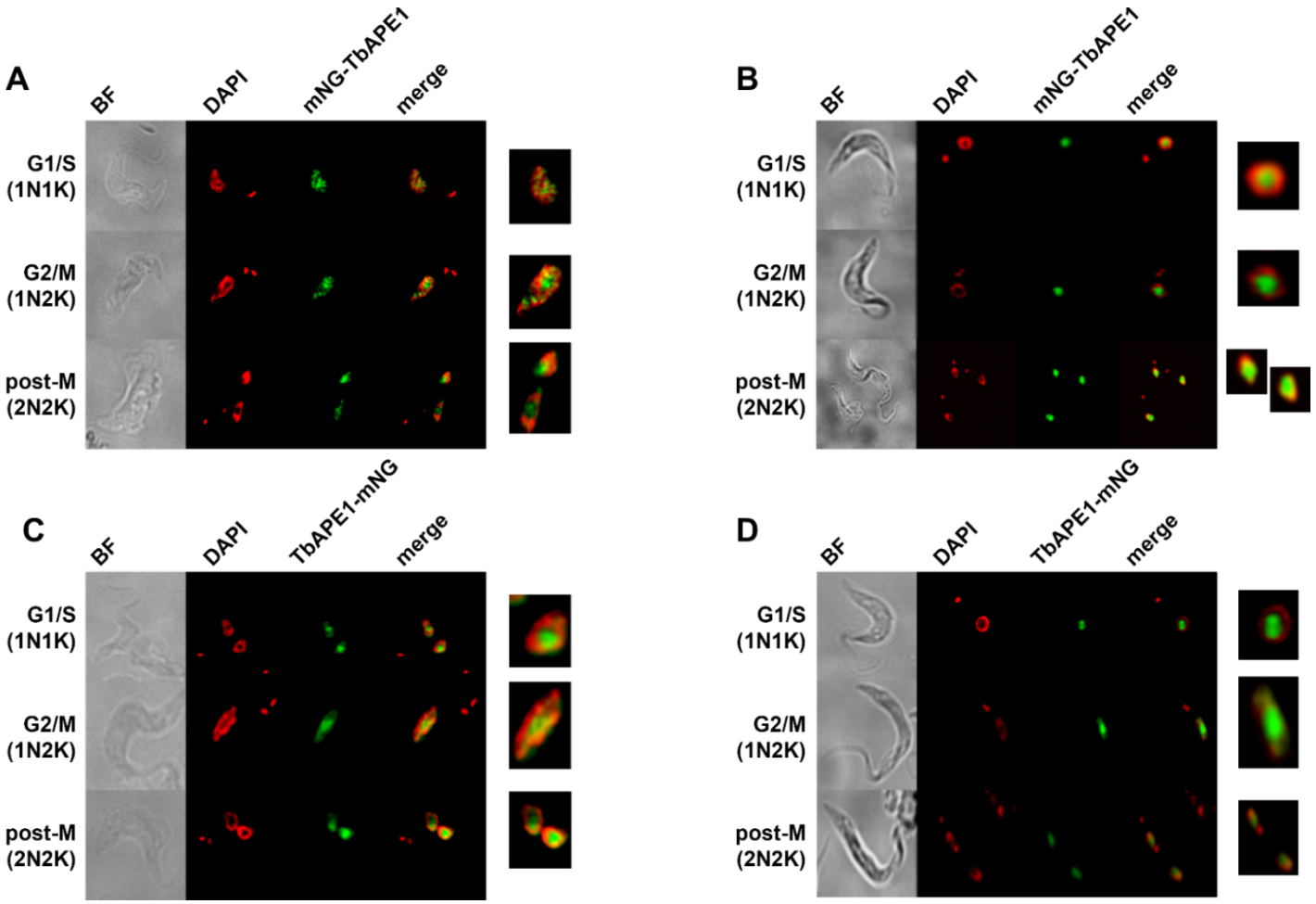


Fig 5.12. Localisation of TbAPE1 in *T. brucei*.
A-D. Bloodstream (**A & C**) and procyclic (**B & D**) form trypomastigotes *T. brucei* expressing TbAPE1 tagged with mNeonGreen (in green) at its amino (mNG-TbAPE1) (**A & B**) or carboxyl (TbAPE1-mNG) (**C & D**) termini were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow), including a close-up image of the nucleus, in trypanosomes at various stages of the cell cycle are shown.

When evaluating TbMYH localisation, a few (two or three) foci coincident with the non-DAPI-staining nucleolus were observed in BSF *T. brucei* expressing *mNeonGreen*-Tb*myh* (Figure 5.13A). In contrast, recombinant PCF cells, irrespective as to which terminal of TbMYH was tagged, a reticular fluorescence signal throughout the cytoplasm was detected (Figure 5.13B and C).

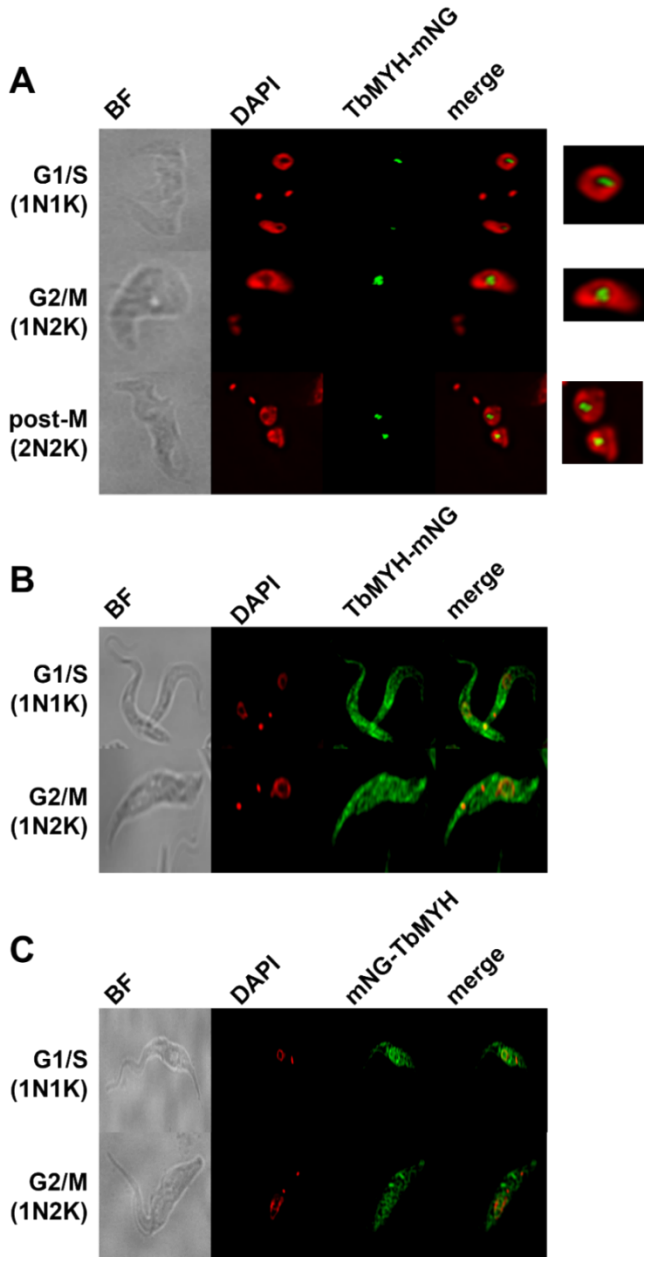


Fig 5.13. Localisation of TbMYH in *T. brucei*.
A-C. Bloodstream (**A**) and procyclic (**B & C**) form trypomastigotes *T. brucei* expressing TMYH tagged with mNeonGreen (in green) at its carboxyl (TbMYH-mNG) (**A & B**), or amino (mNG-TbMYH) (**C**) termini were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow), including a close-up image of the nucleus, in trypanosomes at various stages of the cell cycle are shown.

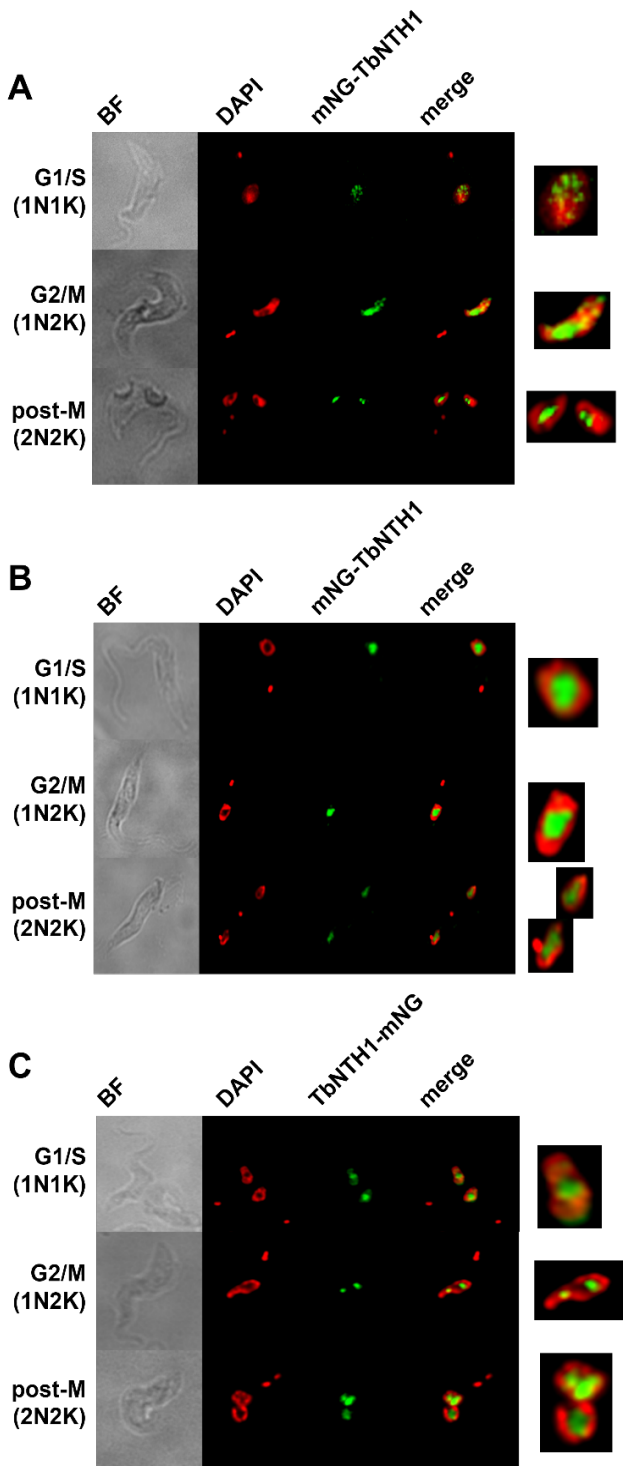


Fig 5.14. TbNTH1 localisation in *T. brucei*.
A-C. Bloodstream (**A & C**) and procyclic (**B**) form trypomastigotes *T. brucei* expressing TbNTH1 tagged with mNeonGreen (in green) at its amino (mNG-TbNTH1) (**A & B**) or carboxyl (TbNTH1-mNG) (**C**) termini were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow), including a close-up image of the nucleus, in trypanosomes at various stages of the cell cycle are shown.

In relation to TbNTH1 localisation, a pattern reminiscent of that noted for TbAPE1 was seen. For *mNeonGreen*-Tb*nth1* or Tb*nth1-mNeonGreen* expressing BSF parasites at most stages of the

cell cycle (G1/S/G2/M), a (sometimes) faint punctate signal throughout the nucleoplasm was noted although many cells also presented with a few intense foci in their nucleoli (Figure 5.14A and C). For trypanosomes in the post-M phase of the cell cycle, the nucleolar signal predominated. PCF parasites expressing *mNeonGreen*-Tb*nth1*, and irrespective of the cell cycle stage, only the intense nucleolar signal was apparent, possibly masking faint foci elsewhere in the nucleus.

5.5 Effect of exogenous treatments on TbAPE1, TbMYH or TbNTH1 expression.

The BER pathways is conserved across all domains of life and is particularly important in aerobic organisms since it neutralises the effect of oxidative stress induced by cellular metabolism and exogenous sources (Demple & Harrison, 1994; Imlay, 2003). To assess the effect of *heterochthonous* agents on the *T. brucei* BER pathway, selected parasite lines engineered to express mNeonGreen variants of TbAPE1 (BSF cells expressing TbAPE1-mNeonGreen), TbMYH (BSF cells expressing TbMYH-mNeonGreen) and TbNTH1 (PCF cells expressing mNeonGreen-TbNTH1) were exposed to single bolus of DNA damaging agents (mechlorethamine; UV light) or trypanocidal nitroheterocycle prodrug (nifurtimox; benznidazole; fexinidazole) and the change in fluorescence signal of individual cells measured at a single time point with the total corrected cell fluorescence values measured in individual parasites compared against data obtained using untreated controls.

Treatment with all the above agents resulted in a significant increase in fluorescence displayed by BSF parasites expressing TbAPE1-mNeonGreen or TbMYH-mNeonGreen (Figure 5.15 & 5.16). For TbAPE1, this was associated with a subtle change in protein localisation such that there was an apparent reduction in relative size and intensity of the nucleoplasm foci with a concomitant increase in the nucleolar signal (Figure 5.15A). In the case of TbMYH, treatment resulted in most of the detected signal being retained in the nucleolus although a few small, discrete foci were seen elsewhere in the nucleus: see the close-up image of the nucleus in mechlorethamine-, nifurtimox- and fexinidazole-treated parasites (Figure 5.16A). It is plausible that these faint extranucleolar foci were present in untreated cells but masked by the intensity of the signal observed in the nucleolus.

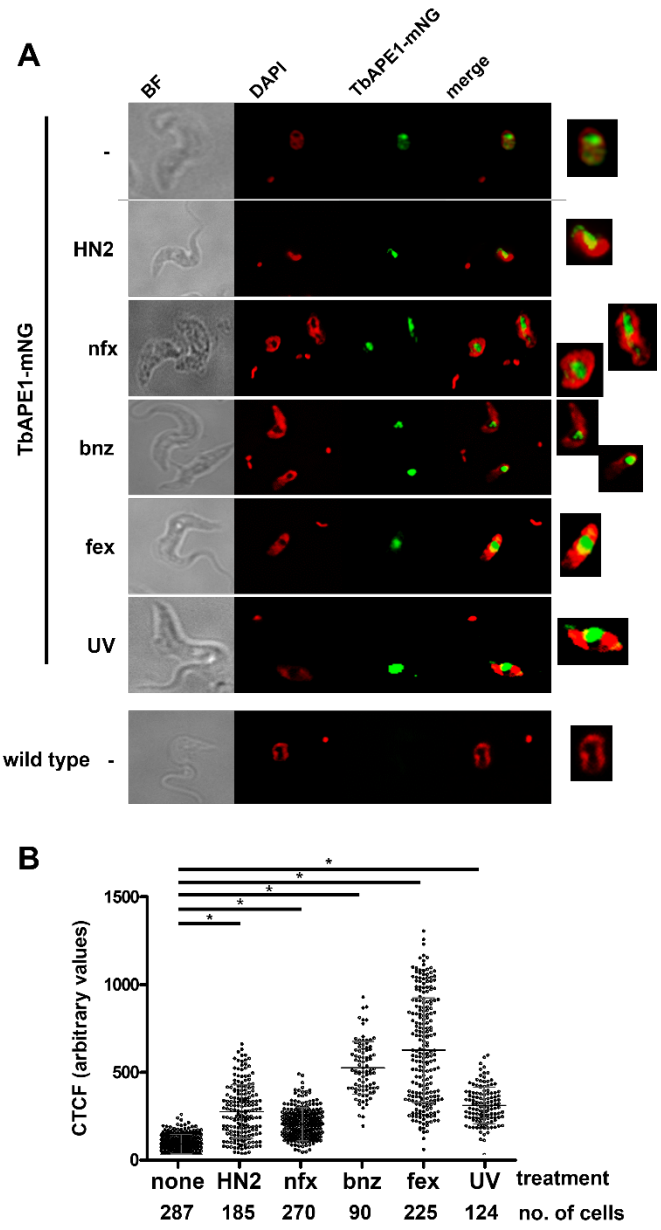


Fig 5.15. Effect of exogenous treatments on TbAPE1 expression

A. Bloodstream form trypomastigotes *T. brucei* expressing TbAPE1-mNeonGreen (TbAPE1-mNG; green) treated with mechlorethamine (HN2; 30 μM for 2 hours), nifurtimox (nfx; 3 μM for 6 hours), benznidazole (bnz; 30 μM for 4 hours), fexinidazole (fex; 3 μM for 3 hours) or UV light (1500 J m^{-2} for 90 minutes) were co-stained with the DNA stain, DAPI (red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. Untreated (-) recombinant and wild type parasites were analysed in parallel. The pattern of colocalization (merge; yellow), including a close-up image of the nucleus, in treated and untreated trypanosomes were compared.

B. The corrected total cell fluorescence (CTCF) of individual *T. brucei* cells expressing mNeonGreen tagged TbAPE1 following mechlorethamine (HN2), nifurtimox (nfx), benznidazole (bnz), fexinidazole (fex) or UV light treatment was determined and compared against untreated controls (none). Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).

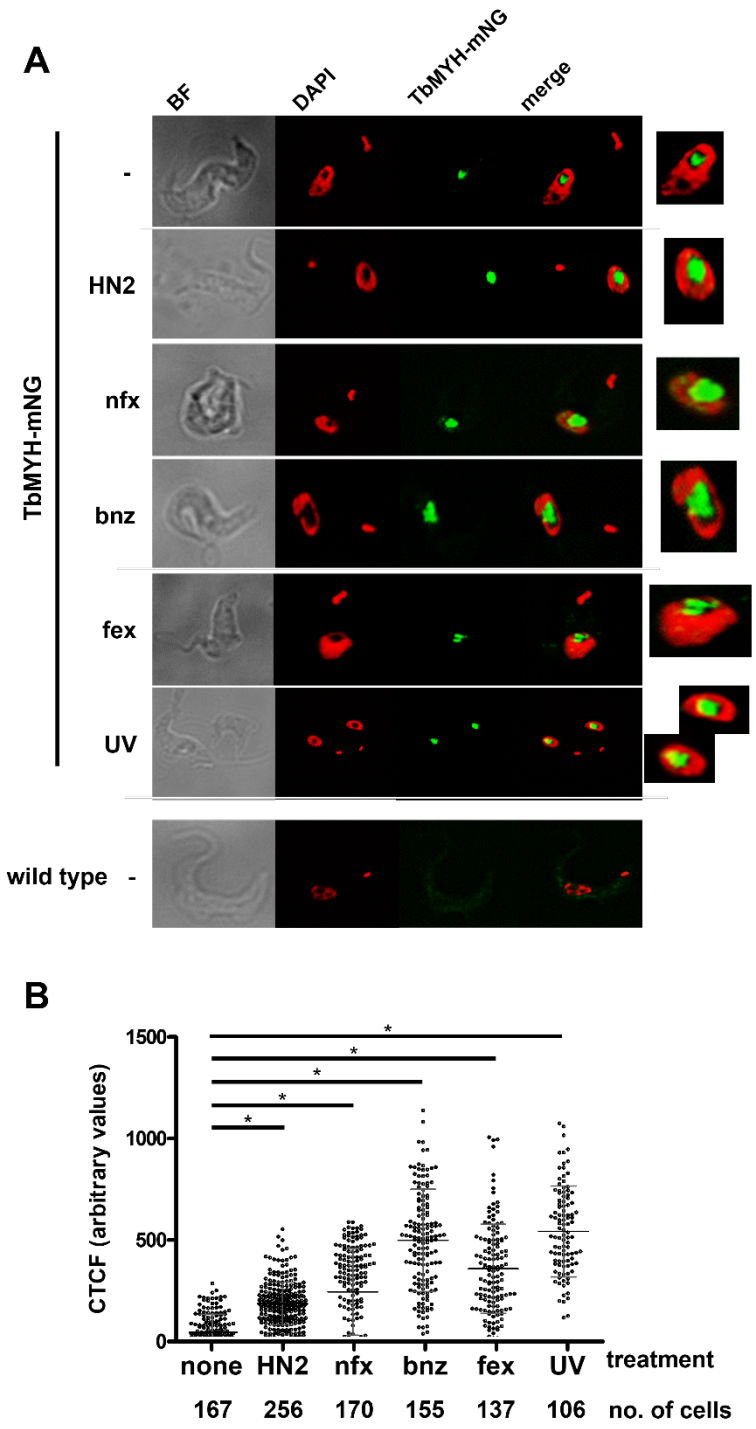


Fig 5.16. Effect of exogenous treatments on TbMYH expression.
A. Bloodstream form trypomastigotes *T. brucei* expressing TbMYH-mNeonGreen (TbMYH-mNG; green) treated with mechlorethamine (HN2; 30 µM for1 hour), nifurtimox (nfx; 3 µM for 2 hours), benznidazole (bnz; 30 µM for 2 hours), fexinidazole (fex; 3 µM for 3 hours) or UV light (1500 J m^{-2} for 90 minutes) were co-stained with the DNA stain, DAPI (red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured: Untreated (-) recombinant and wild type parasites were analysed in parallel. The pattern of colocalization (merge; yellow), including a close-up image of the nucleus, in treated and untreated trypanosomes were compared.
B. The corrected total cell fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen tagged TbMYH following mechlorethamine (HN2), nifurtimox (nfx), benznidazole (bnz), fexinidazole (fex) or UV light treatment was determined and compared against untreated controls (none). Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).

In the case of PCF cells expressing mNeonGreen-TbNTH1, a significant increase in total cellular fluorescence was noted following treatment with mechlorethamine, UV light or nifurtimox with the signal found exclusively in nucleolus (Figure 5.15): no extranucleolar signal was

observed. In contrast, no meaningful change in fluorescence was detected after treatment with the two trypanocidal nitroimidazole compounds, benznidazole and fexinidazole.

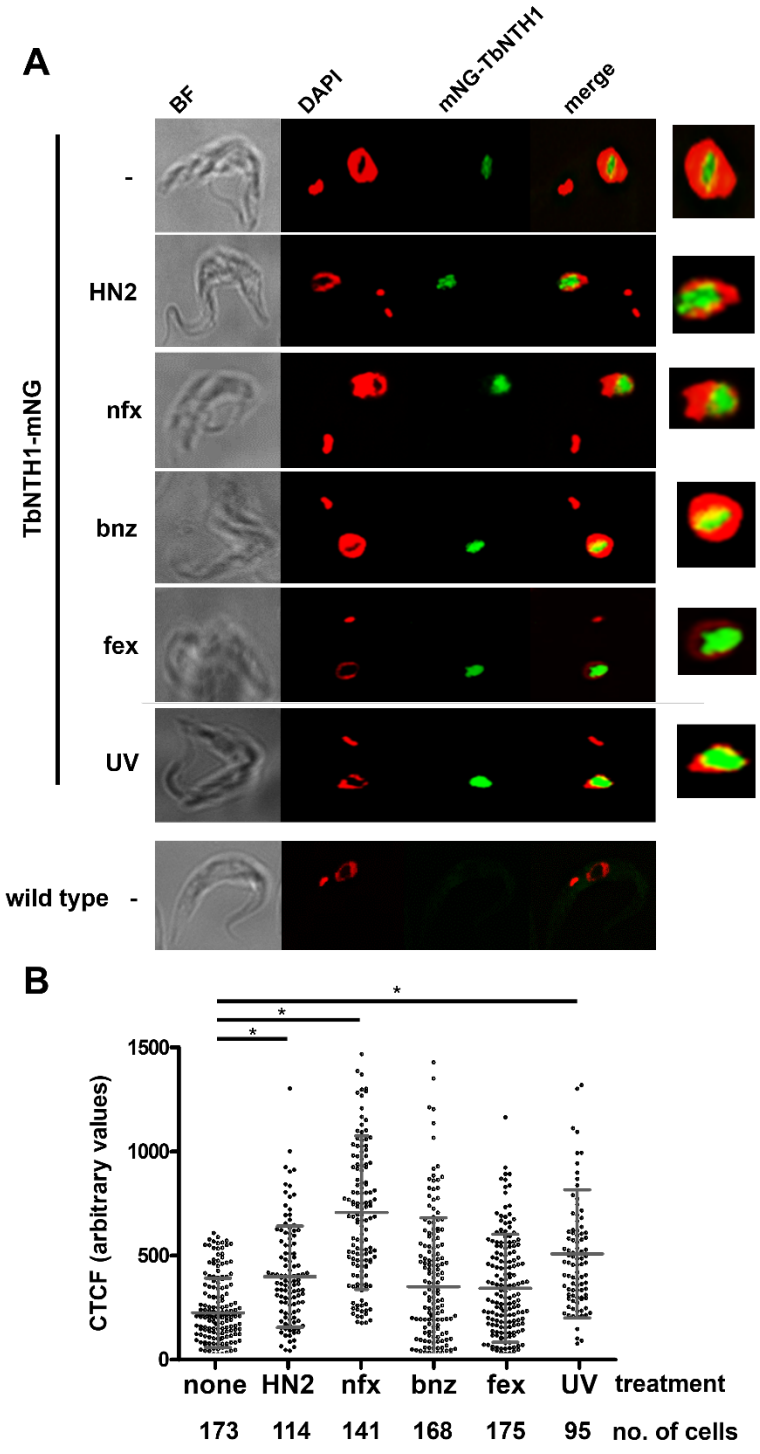


Fig 5.17. Effect of exogenous treatments on TbNTH1 expression. Procyclic form *T. brucei* expressing mNeonGreen-TbNTH1 (mNG-TbNTH1; green) treated with mechlorethamine (HN2; 30 µM for 4 hours), nifurtimox (nfx; 3 µM for 4 hours), benznidazole (bnz ; 30 µM for 4 hours), fexinidazole (fex; 3 µM for 3 hours) or UV light (1500 J m^{-2} for 4 hours) were co-stained with the DNA stain, DAPI (red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured: Untreated (-) recombinant and wild type parasites were analysed in parallel. The pattern of colocalization (merge; yellow), including a close-up image of the nucleus, in treated and untreated trypanosomes were compared. **B.** The corrected total cell fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen tagged TbNTH1 following mechlorethamine (HN2), nifurtimox (nfx), benznidazole (bnz), fexinidazole (fex) or UV light treatment was determined and compared against untreated controls (none). Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).

In human cells, MYH and APE1 have a temporal and synergistic relationship in the repair of damaged bases with the former enzyme able to bind to the latter via an interdomain connector and thereby enhancing its AP endonuclease activity (Luncsford et al., 2013). To assess for any temporal response of the *T. brucei* BER pathway, BSF *T. brucei* expressing TbMYH-mNeonGreen or TbAPE1-mNeonGreen were treated with trypanocidal nitroheterocyclic prodrugs (30 µM benznidazole; 3 µM nifurtimox; 3 µM fexinidazole) or with DNA damaging agents (30 µM mechlorethamine; 1500 J m^{-2} UV) and the corrected total cell fluorescence in individual parasites monitored over time.

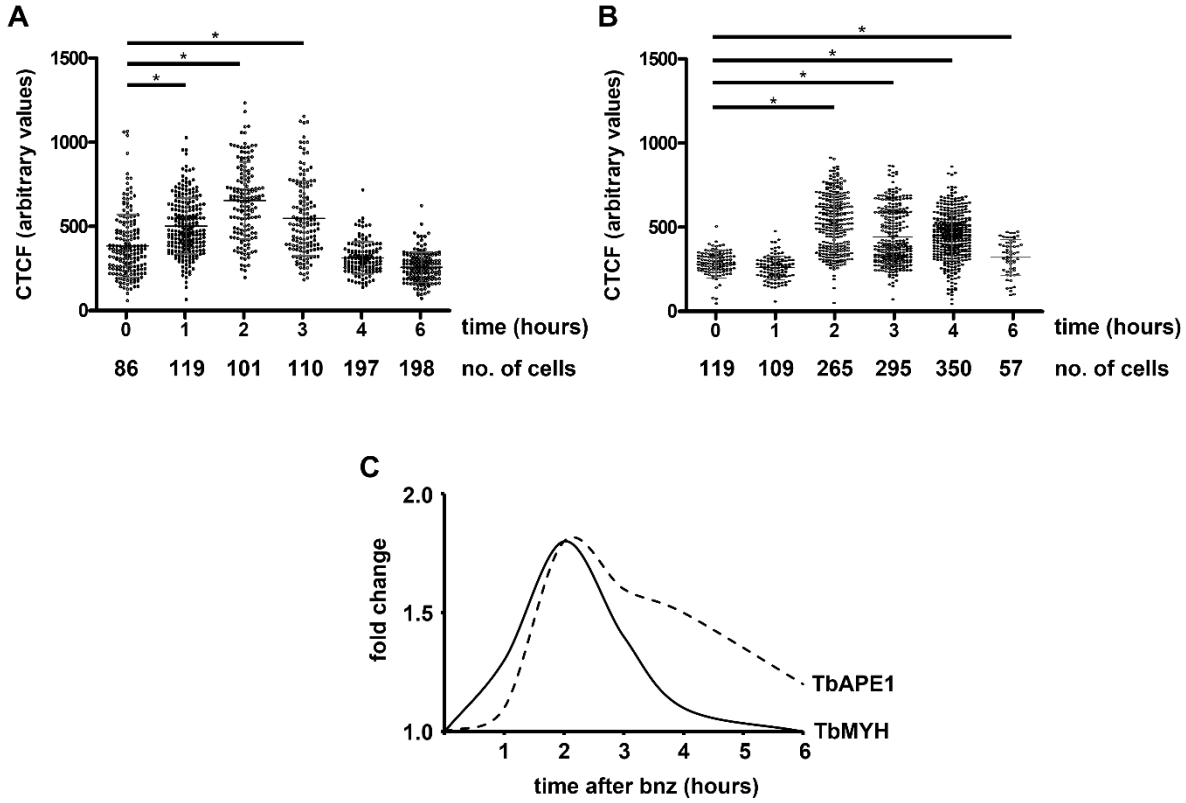


Fig 5.18. Effect of benznidazole on TbMYH and TbAPE1 temporal expression.
A & B. The corrected total cell fluorescence (CTCF) of individual *T. brucei* cells expressing mNeonGreen tagged TbMYH (**A**) or TbAPE1(**B**) was determined at time intervals following benznidazole (30 µM) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software)
C. Schematic showing the temporal changes in expression of TbMYH or TbAPE1 following addition of benznidazole (bnz) to cultures.

Following addition of benznidazole to *T. brucei* cultures, TbMYH-mNeonGreen expression was shown to significantly increase in the first hour of treatment, reaching maximum fluorescence by 2-hours posttreatment. The signal subsequently declined, reverting to untreated levels by 4 hours posttreatment (Figure 5.18A). For TbAPE1, the fluorescence signal remained at untreated levels over the first hour of treatment before rapidly rising and reaching a peak 2-hours posttreatment before slowly declining over the remaining time points (Figure 5.18B). A summary of the combined temporal changes in TbMYH and TbAPE1 expression in response to benznidazole is presented in Figure 5.18C.

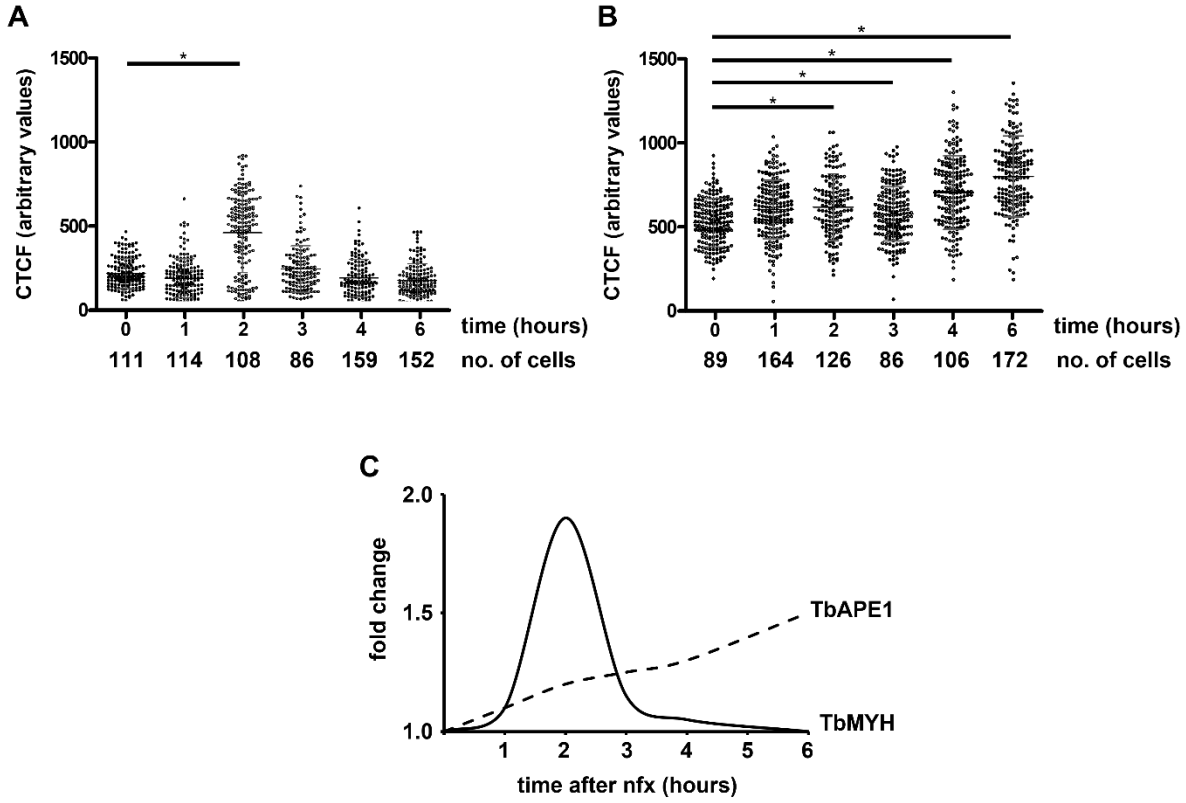


Fig 5.19. Effect of nifurtimox on TbMYH and TbAPE1 temporal expression.
and B. The corrected total cell fluorescence (CTCF) of individual *T. brucei* cells expressing mNeonGreen tagged TbMYH (**A**) or TbAPE1(**B**) was determined at time intervals following nifurtimox (30 μM) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).
C. Schematic showing the temporal changes in expression of TbMYH or TbAPE1 following addition of nifurtimox (nfx) to cultures.

When the effect of nifurtimox on the temporal expression of the *T. brucei* BER factors was assessed, fluorescence associated with TbMYH-mNeonGreen remained at untreated levels during the first hour of treatment, then rapidly rose and reach a peak 2-hours posttreatment before rapidly falling and reverting back to untreated levels by 3-hours posttreatment (Figure 5.19A). In contrast, for TbAPE1 a slow progressive increase in expression was seen over the course of the experiment, becoming significant 2-hours posttreatment (Figure 5.19B). A summary of the combined temporal changes in TbMYH and TbAPE1 expression in response to nifurtimox is presented in Figure 5.19C.

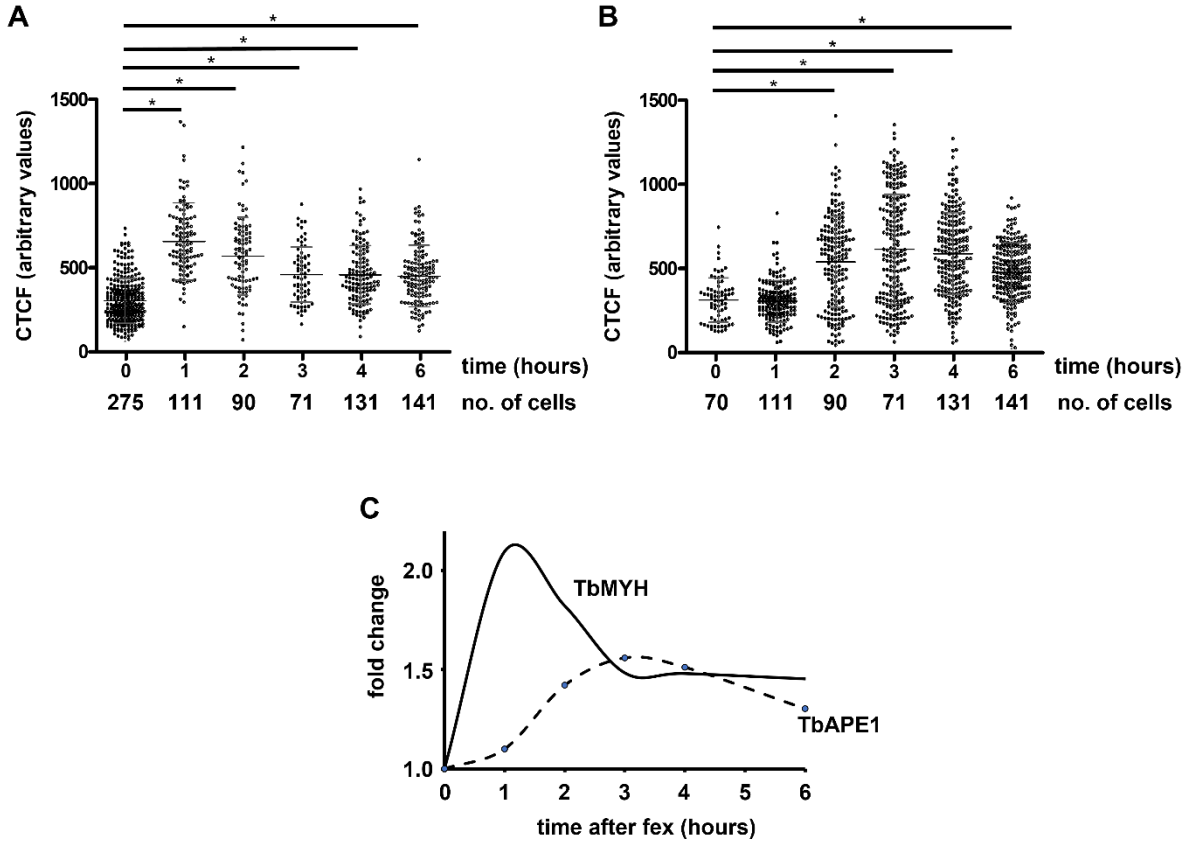


Fig 5.20. Effect of fexinidazole on the temporal expression of the *T. brucei* base excision repair pathway.

A & B. The corrected total cell fluorescence (CTCF) of individual *T. brucei* cells expressing mNeonGreen tagged TbMYH (**A**) or TbAPE1(**B**) was determined at time intervals following fexinidazole (30 μM) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software). **C.** Schematic showing the temporal changes in expression of TbMYH or TbAPE1 following addition of fexinidazole (fex) to cultures.

In fexinidazole-treated trypanosomes, TbMYH-mNeonGreen fluorescence was shown to reach a peak during the first hour of treatment with the signal falling slightly over the next two hours before plateauing and remaining at a level significantly above that noted for untreated cells (Figure 5.20A). For TbAPE1, fluorescence remained at untreated levels during the first hour of treatment, rose over the new 2 hours and reached a peak 3-hours posttreatment before starting to decline Figure 5.20B). A summary of the combined temporal changes in TbMYH and TbAPE1 expression in response to fexinidazole is presented in Figure 5.20C.

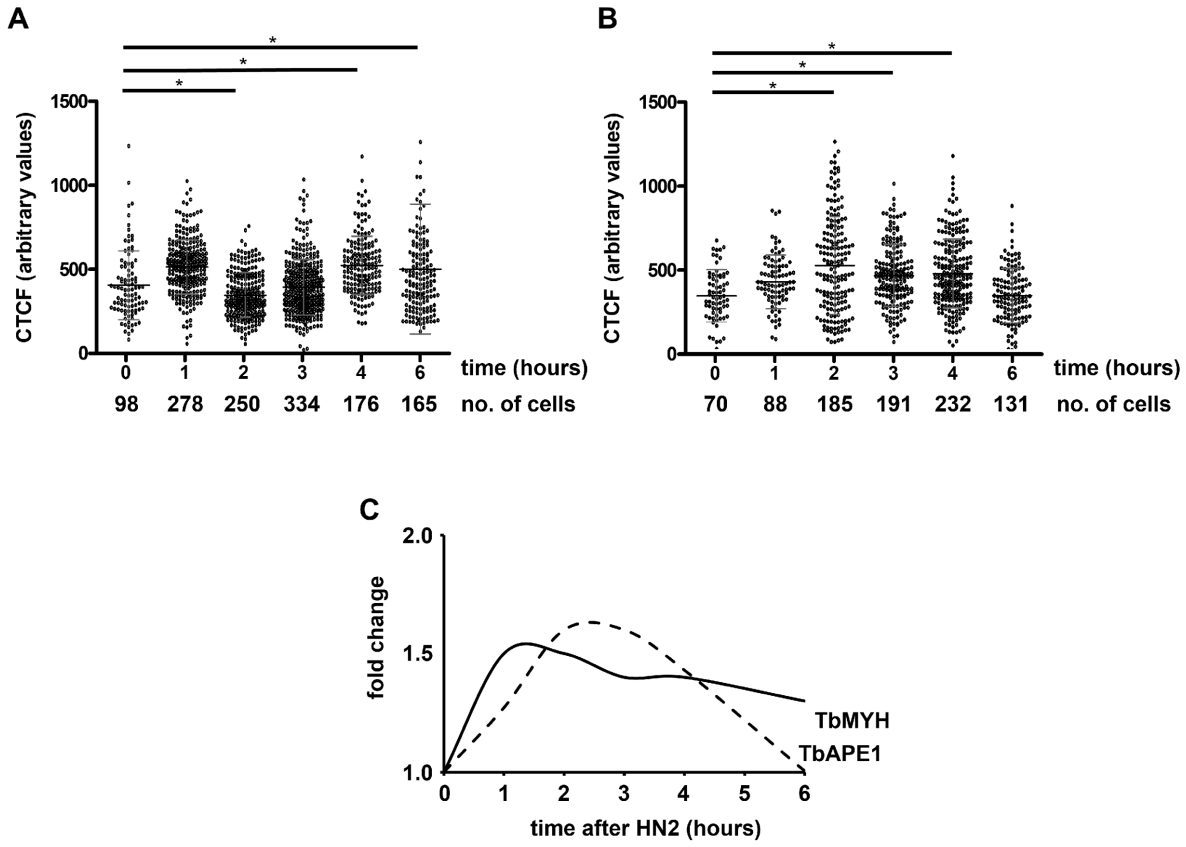


Fig 5.21. Effect of mechlorethamine on the temporal expression of TbMYH and TbAPE1.

A and B. The corrected total cell fluorescence (CTCF) of individual *T. brucei* cells expressing mNeonGreen tagged TbMYH (**A**) or TbAPE1 (**B**) was determined at time intervals following mechlorethamine (30 μM) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).
C. Schematic showing the temporal changes in expression of TbMYH or TbAPE1 following addition of mechlorethamine (HN2) to cultures.

When the above analysis was extended to the ICL inducing agent mechlorethamine, a slight but significant increase in the fluorescence signal associated with TbMYH-mNeonGreen or TbAPE1-mNeonGreen was noted. For the DNA glycosylase, expression peaked 1-hour posttreatment and remained at roughly this level over the course of the experiment (Figure 5.21A) while the endonuclease expression reached a maximum 3-hours posttreatment before slowly declining and reaching untreated levels 6-hours posttreatment Figure 5.21B). A summary of the temporal changes in TbMYH and TbAPE1 expression caused by mechlorethamine is presented in Figure 5.21C.

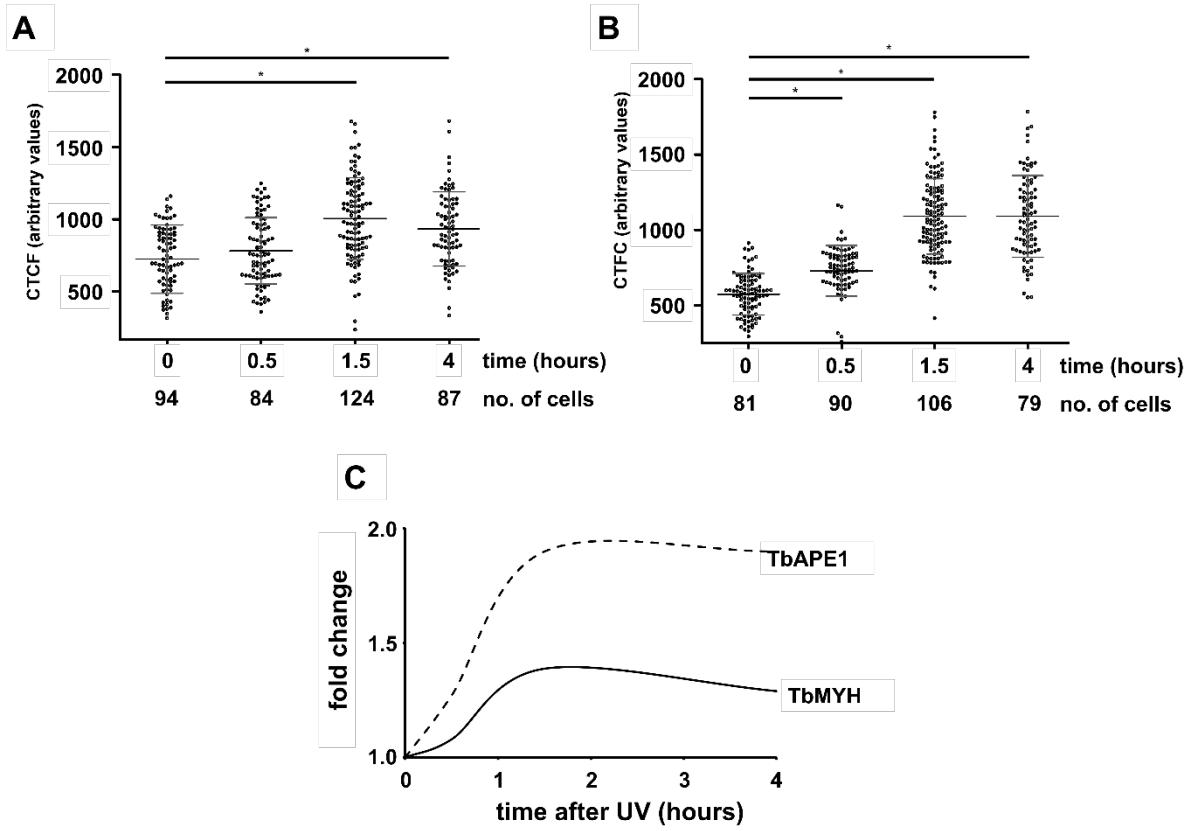


Fig.5.22. Effect of UV on temporal expression of TbMYH and TbAPE1.
A and B. The corrected total cell fluorescence (CTCF) of individual *T. brucei* cells expressing tagged TbMYH-mNeonGreen (**A**) and TbAPE1-mNeonGreen (**B**) was monitored following ultraviolet light (UV) irradiation (1500 J m^2) with the resultant data compared against untreated controls (none). Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).
C. Schematic showing the temporal changes in expression of TbMYH or TbAPE1 following UV irradiation.

UV irradiation of BSF *T. brucei* expressing TbMYH-mNeonGreen or TbAPE1-mNeonGreen resulted in a slight increase in fluorescence over the 30 minutes posttreatment (Figure 5.22A and B). For the DNA glycosylase this initial change was non-significant while for the AP endonuclease this alteration was deemed significant. For both enzymes, the signal became significant and reached a peak 90 minutes, posttreatment then remained at an elevated level at the remaining timepoint (4-hours posttreatment). A summary of the temporal changes in TbMYH and TbAPE1 expression caused by UV light is presented in Figure 5.22C.

5.6 Disrupting Tb*ape1, Tbmyh* or *Tbnth1* in the genome of procyclic form *T. brucei.*

To examine the function of TbAPE1, TbMYH and TbNTH1 in PCF *T. brucei*, the previously described CRISPR/Cas9-mediated gene deletion approach was employed (See Figure 4.16) to replace the Tb*ape1-, Tbmyh*- or *Tbnth1*-containing alleles with sequences coding for the hygromycin B phosphotransferase (*hyg*) or neomycin phosphotransferase (*neo*) selectable markers. The primer combinations and amplicons required to target and replace each allele are given in Table 5.4. For each gene, electroporations coupled with appropriate drug selection were performed to obtain clonal lines null for a targeted activity while in other situations heterozygote lines were generated. The parasite lines generated from this analysis are listed in Table 5.5.

Targeted gene	Amplicon	Primer combinations
Tb*ape1*	5' sgRNA-Tbape1	G00/5'sgRNA-Tbnth1
	3' sgRNA-Tbape1	G00/3'sgRNA-Tbnth1
	HYG-TF-Tbape1	Tbnth1-UFP1/ Tbnth1-DRKOP1
	NEO-TF-Tbape1	Tbnth1-UFP1/ Tbnth1-DRKOP1
Tb*myh*	5' sgRNA-Tbmyh	G00/5'sgRNA-Tbnth1
	3' sgRNA-Tbmyh	G00/3'sgRNA-Tbnth1
	HYG-TF-Tbmyh	Tbnth1-UFP1/ Tbnth1-DRKOP1
	NEO-TF-Tbmyh	Tbnth1-UFP1/ Tbnth1-DRKOP1
Tb*nth1*	5' sgRNA-Tbnth1	G00/5'sgRNA-Tbnth1
	3' sgRNA-Tbnth1	G00/3'sgRNA-Tbnth1
	HYG-TF-Tbnth1	Tbnth1-UFP1/ Tbnth1-DRKOP1
	NEO-TF-Tbnth1	Tbnth1-UFP1/ Tbnth1-DRKOP1

Table 5.4. Primer combinations used in the CRISPR/Cas9 deletion of BER genes.
The primer combinations used to generate small amplicons (designated as 5'sgRNA-TbXXX & 3'sgRNA-TbXXX) that correspond to the short guide RNA (sgRNA) which target the 5' or 3' ends of the targeted gene and large amplicons (designated as HYG-TF-TbXXX & NEO-TF-TbXXX) that correspond to the hygromycin or neomycin targeting fragments used to replace the targeted gene. The sequence of each primer can be found in Appendix 1.

Targeted gene	Name of line	comments
Tba*pe1*	Tb*ape1Δ*	Putative null line; neomycin & hygromycin resistant
	Tb*ape1*$^{+/-}$neo	Putative heterozygote; neomycin resistant
	Tb*ape1*$^{+/-}$hyg	Putative heterozygote; hygromycin resistant
Tb*myh*	Tb*myhΔ*	Putative null line; neomycin & hygromycin resistant
	Tb*ape1*$^{+/-}$neo	Putative heterozygote; neomycin resistant
Tb*nth1*	Tb*nth1Δ*	Putative null line; neomycin & hygromycin resistant
	Tb*ape1*$^{+/-}$neo	Putative heterozygote; neomycin resistant

Table 5.5. *T. brucei* lines potentially null or heterozygous for Tb*ape1, Tbmyh* or *Tbnth1*

For Tb*ape1*, to confirm that the introduced DNA amplicons had integrated into the correct genetic loci, PCRs were performed on gDNA extracted from wild type, heterozygote (Tb*ape1*$^{+/-}$neo; Tb*ape1*$^{+/-}$hyg) and null (Tb*ape1Δ*) cells (Figure 5.23). Such reactions used primer combinations to generate biomarker fragments specific for the intact Tb*ape1* gene or the *hyg* and *neo* disrupted alleles (*Tbape1::hyg* and *Tbape1::neo*, respectively). When using a primer combination designed to detect intact Tb*ape1*, amplification generated a band of the expected size, ~620 bp, from templates derived from the wild type and heterozygote (Tb*ape1*$^{+/-}$neo; Tb*ape1*$^{+/-}$hyg) lines with no band observed from null (Tb*ape1Δ*) cells (Figure 5.23B). In contrast, primer combinations that detect the Tb*ape1::neo* or Tb*ape1::hyg* alleles generated amplicons of the predicted size (~1100 bp and ~1300 bp, respectively) only from the gDNA purified from *T. brucei* Tb*apeΔ* with no bands observed when using *T. brucei* wild type gDNA as template. Note the expected disrupted Tb*ape1::neo* and Tb*ape1::hyg* alleles were amplified from Tb*ape1*$^{+/-}$neo and Tb*ape1*$^{+/-}$hyg, respectively.

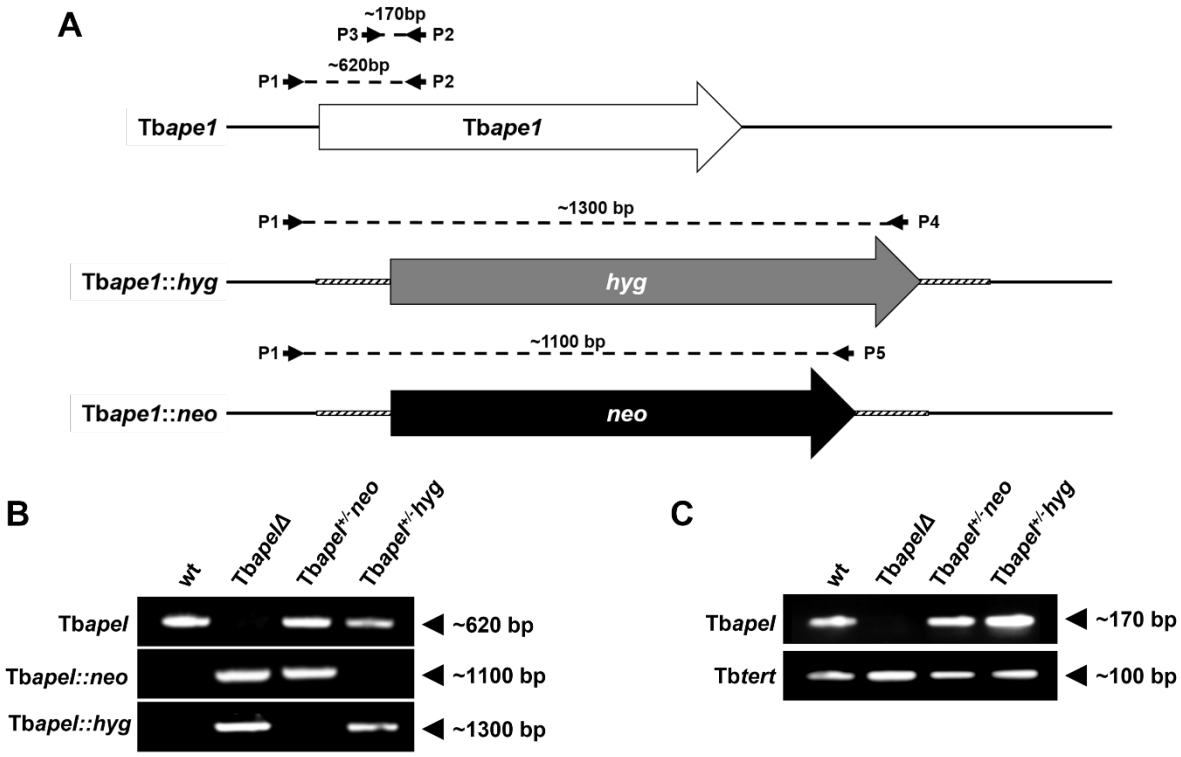


Fig. 5.23 Effect of *Tbape1* deletion on PCF *T. brucei*.

A. Schematic representation of the Tb*ape1* allele and the effect of its disruption with DNA fragments containing sequences encoding for neomycin phosphotransferase (*neo*) or hygromycin B phosphotransferase (*hyg*), plus *T. brucei* intergenic elements required for processing their mRNAs (hatched boxes). P1+P2 and P1+P3 correspond to the primers used to generate a Tb*ape1* specific amplicons from cDNA or genomic DNA templates respectively. P1+P4 and P1+P5 correspond to the primers used to generate specific amplicons used to detect Tb*ape1* alleles interrupted with *hyg* or *neo*, respectively. The size of the predicted amplicons, denoted by a dashed line, is given. The primers sequences used are: P1 – TGCTTGCACCAGGAGGAGAGG (TbAPE1-1), P2 – AGCTTCATTACCCTTGTCCCC (TbAPE1-seq1), P3 – GACTTTAAGCCTCGGGAGATG (TbAPE1-2), P4 – ggg*atcgat*CTATTCCTTTGCCCTCGGACG (HYG-2) and P5 – ggg*atcgat*TCAGAAGAACTCGTCAAGAAG (NEO-2).

B. Amplicons corresponding to intact Tb*ape1,* and the *neo*- or *hyg*-interrupted counterparts were generated from template genomic DNAs extracted from *T. brucei* wild type (wt), Tb*ape1* heterozygotes (Tb*ape1*$^{+/-}$*neo*; Tb*ape1*$^{+/-}$*hyg*) and Tb*ape1* null mutant (Tb*ape1Δ*) lines**.** The amplified fragment is indicated (Tb*ape1*; Tb*ape1::neo* or Tb*ape1::hyg*).

C. Amplicons corresponding to intact Tb*ape1* were generated from template cDNAs derived from total RNA extracted from *T. brucei* wild type (wt) or Tb*ape1* heterozygotes (Tb*ape1*$^{+/-}$*neo*; Tb*ape1*$^{+/-}$*hyg*) and Tb*ape1* null mutant (Tb*ape1*).

To confirm that the Tb*ape1Δ* was no longer expressing the interrupted gene, PCRs were performed using a primer combination (TbAPE1-seq1/ TbAPE1-2) designed to generate a Tb*ape1* specific amplicon against cDNA generated from total RNA (Figure 5.23C). When using wild type or heterozygote cDNA as template, a single band of ~170 bp, the expected size, was observed with

no band(s) detected in material derived from *T. brucei* Tb*ape1*Δ. To confirm that RNA had been extracted from both cell lines and that cDNA had indeed been made, control reactions amplifying Tb*tert* 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.

The above validation strategy was then extended to Tb*myh* with PCRs performed on gDNA extracted from wild type, heterozygote (Tb*myh*$^{+/-}$neo) and null (Tb*myhΔ*) cells (Figure 5.24). Here, primers designed to detect the *hyg* and *neo* disrupted alleles (*Tbmyh:hyg* and *Tbmyh::neo*, respectively) generated amplicons of the predicted size (~1025 bp and ~1255 bp, respectively) only from the gDNA purified from Tb*myhΔ* with no bands observed when using *T. brucei* wild type gDNA as template (Figure 5.24B): the *Tbmyh::neo* was also detected from Tb*myh*$^{+/-}$neo-derived templates. When amplifications to detect intact Tb*myh* were performed, all sample generate a band of the expected size, ~550 bp. This indicates that the putative Tb*myh* null line, although possessing both gene disruption markers at the correct loci has still retained a copy of the targeted gene.

To assess if the "null" line does express the Tb*myh* transcript, DNA amplification were performed on cDNA derived from wild type, Tb*myh*$^{+/-}$neo and Tb*myhΔ* cells. For all samples, a ~140 bp sized band diagnostic of an intact Tb*myh* was obtained thereby confirming that Tb*myhΔ* cells have retained this activity (Figure 5.24C).

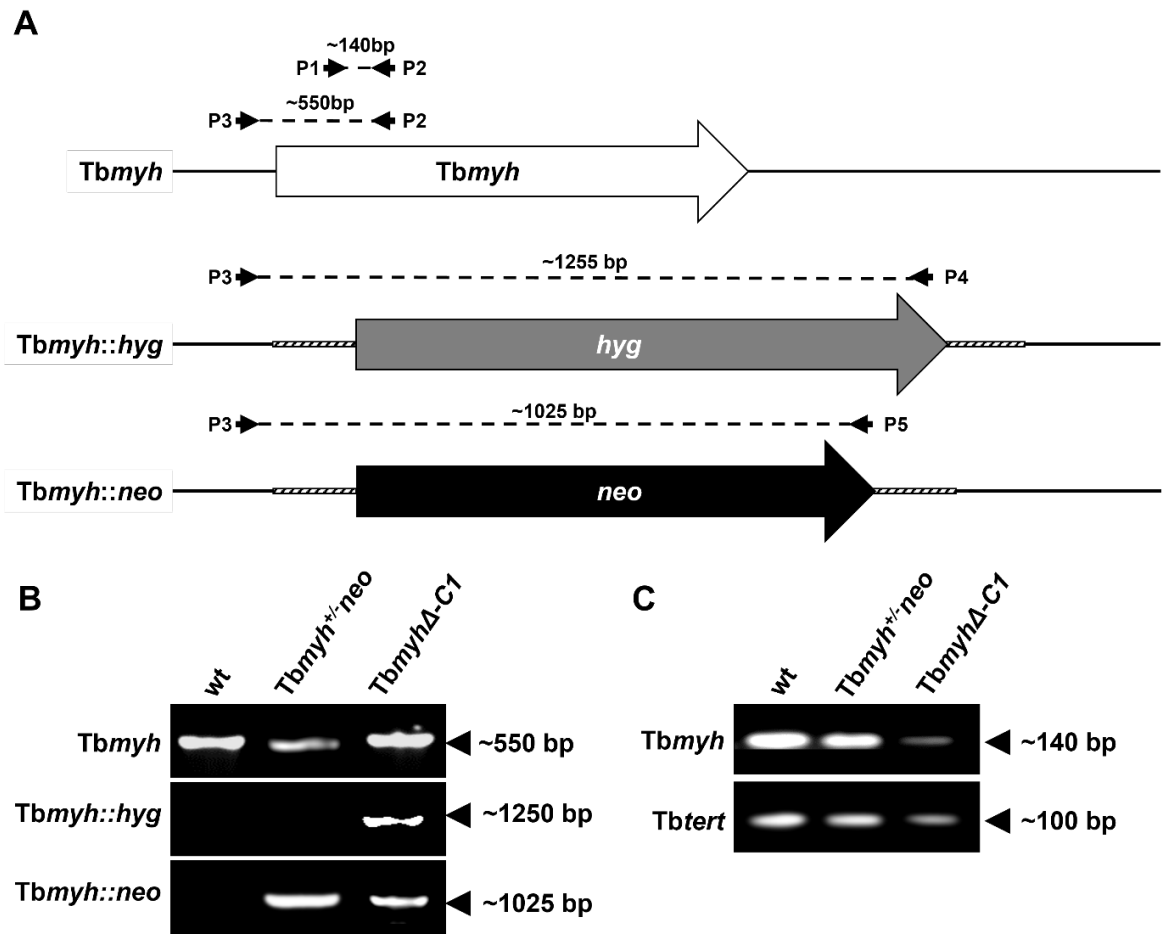


Fig. 5.24 Effect of Tb*myh* deletion in PCF *T. brucei*.

A. Schematic representation of the Tbmyh allele and the effect of its disruption with a DNA fragment containing sequences encoding for neomycin phosphotransferase (*neo*) or hygromycin B phosphotransferase (*hyg*), plus *T. brucei* intergenic elements required for processing their mRNAs (hatched boxes). P1+P2 and P2+P3 correspond to the primers used to generate Tb*myh* specific amplicons from cDNA or genomic DNA templates respectively. P3+P4 and P3+P5 correspond to the primers used to generate specific amplicons used to detect Tb*myh* alleles interrupted with *hyg* or *neo*, respectively. The size of the predicted amplicons, denoted by a dashed line, is given. The primers sequences used are: P1 – CTCGATGGTGTGGCAACTGCG (TbMYH-1), P2 – GGAAAGACCGTATCCATCGC (TbMYH-seq1), P3 – GACACAGATGCGTGTCATCCG (TbMYH-2), P4 – ggg*atcgat*CTATTCCTTTGCCCTCGGACG (HYG-2) and P5 – ggg*atcgat*TCAGAAGAACTCGTCAAGAAG (NEO-2).

B. Amplicons corresponding to intact Tb*myh,* and the *neo*- or *hyg*-interrupted counterparts were generated from template genomic DNAs extracted from *T. brucei* wild type (wt), Tb*myh* heterozygote (Tb*myh*$^{+/-}$*neo*) and "null" mutant (Tb*myhIΔ-C1*) lines. The amplified fragment is indicated (Tb*myh*; Tb*myh: neo* or Tb*myh: hyg*).

C. Amplicons corresponding to intact Tb*myh* were generated from template cDNAs derived from total RNA extracted from *T. brucei* wild type (wt), Tb*myh* heterozygote (Tb*myh*$^{+/-}$*neo*) and "null" mutant (Tb*myhIΔ-C1*) lines**.** The integrity of cDNAs (and hence RNAs) was evaluated by amplification of a 100 bp control fragment, Tb*tert*. The amplified fragment (Tb*myh* or Tb*tert*) Is indicated. The amplicon profile for one Tb*myh* "null" clonal line is shown: another putative "null" line gave the same pattern.

The above PCR approach was then used to evaluate if Tb*nth1* had been successfully disrupted in PCF *T. brucei* cells. Using cDNA derived from wild type, Tb*nth1*$^{+/-}$neo and *Tbnth1Δ* cells and primers designed to generate a Tb*nth1* specific amplicon, a band of expected size, ~410 bp, was detected from all tested templates (Figure 5.25). As with Tb*myh*, this indicates that the line "null" for Tb*nth1* has still retained at least one copy of the targeted gene.

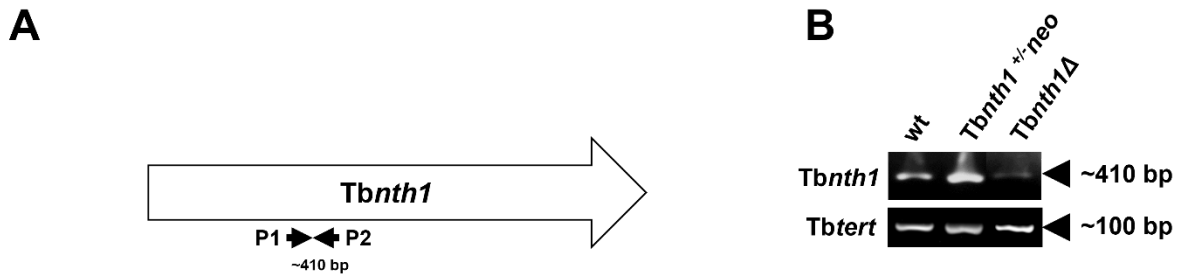


Fig. 5.25. Effect of Tb*nth1* deletion in PCF *T. brucei*.
A. Schematic representation of the Tb*nth1* allele. P1 (gggGGCAAGCTATTTAAACCACCA) and P2 (TTGGGCCCAACACCTGGAAGG) correspond to the primers used to generate a Tb*nth1* specific amplicon (predicted size noted in bp) from cDNA templates.
B. Amplicons corresponding to intact Tb*nth1* were generated from template cDNAs derived from total RNA extracted from *T. brucei* wild type (wt), Tb*nth1* heterozygote (Tb*nth1*$^{+/-}$*neo*) and "null" (Tb*nth1Δ*) lines. The integrity of cDNAs (and hence RNAs) was evaluated by amplification of a 100 bp control fragment, Tb*tert*. The amplified fragment (Tb*nth1* or Tb*tert*) is indicated. The amplicon profile for one Tb*nth1* "null" clonal line is shown: two other putative "null" lines gave the same pattern.

Taken together, the above DNA amplifications confirm that both copies of Tb*ape1* had indeed been deleted from the PCF *T. brucei* genome thereby validating the Tb*ape1Δ* line. This demonstrates that, as in BSF parasites (Charret et al 2012), this AP endonuclease is not essential for viability of insect form parasites when grown in culture. Intriguingly, lines null for TbMYH or TbNTH1 activity could not be generated with cells still retaining at least one copy of the encoding gene. This suggests that both DNA glycosylases may be important, if not essential, to the survival and growth of PCF *T. brucei*.

5.7 Phenotypic analysis of null lines

The base excision repair pathway plays a key role in the removal of small non-helix distorting base damage. Despite its importance in maintaining genome integrity, absence of certain key DNA repair enzymes (e.g., DNA glycosylases) generally does not lead to adverse growth defects in the absence of genotoxic challenge (Parsons & Elder, 2003; Friedberg, 2006; Jacobs & Schär, 2012)

although an accrual of base lesions in the genome may be detected (Russo et al., 2004; Castillo-Acosta et al., 2012). In contrast, deletion of genes that code other BER enzymes such as Polβ (Sobol et al., 1996), APE1 (Xanthoudakis et al., 1996), FEN1 (Kucherlapati et al., 2002) and XRCC1 (Tebbs et al., 1999) is often lethal with inhibition of these activities leading to predisposition to cancer (Xanthoudakis et al., 1996; Ludwig et al., 1998; Kucherlapati et al., 2002). Here, the growth and/or susceptibility characteristics of PCF *T. brucei* null for TbAPE1 (Tb*ape1Δ*) and cells with (presumably) lower TbMYH (Tb*myhΔ*) or TbNTH1 (Tb*nth1Δ*) activities was assessed.

Using the previously described assays (see sections 2.9.1 & 4.2.6), the cumulative growth of Tb*ape1Δ* and Tb*myhΔ* PCF *T. brucei* was followed over 22 days. For both recombinant lines, profiles were reminiscent of that obtained with wild type were seen from which comparable doubling times were extrapolated (Figure 5.26): the *Tbape1Δ* and *TbmyhΔ* cells had a doubling time of around 14 and 12 ¼ hours respectively with the control being 13¼ hours. As such, absence of TbAPE1 or reduced levels of TbMYH in PCF *T. brucei* has no significant effect on parasite growth.

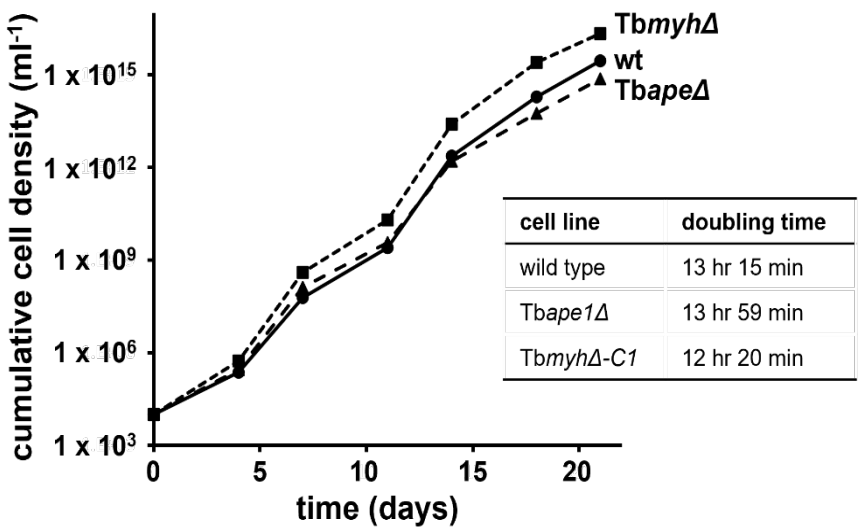


Fig. 5.26. Cumulative growth of *T. brucei* mutant lines. Cumulative cell density of Tb*ape1Δ* (▲) and Tb*myhΔ* (■) were followed for 22 days and compared against the growth of PCF wild type (wt) *T. brucei* (●) cultures grown in parallel. Each data point represents the mean cell density ± standard deviation from two independent cultures. Insert shows the doubling time of each cell line.

A series of growth inhibition assays were then performed on each line to evaluate whether lack of APE1 or reduced levels of TbMYH or TbNTH1 effects that line's susceptibility to hydroxyurea, H_2O_2, phleomycin, MMS, UV light and mechlorethamine. The resultant data was then plotted to generate dose response curves (Figures 5.27 to 5.29) from which EC_{50} values (Table 5.5) were derived.

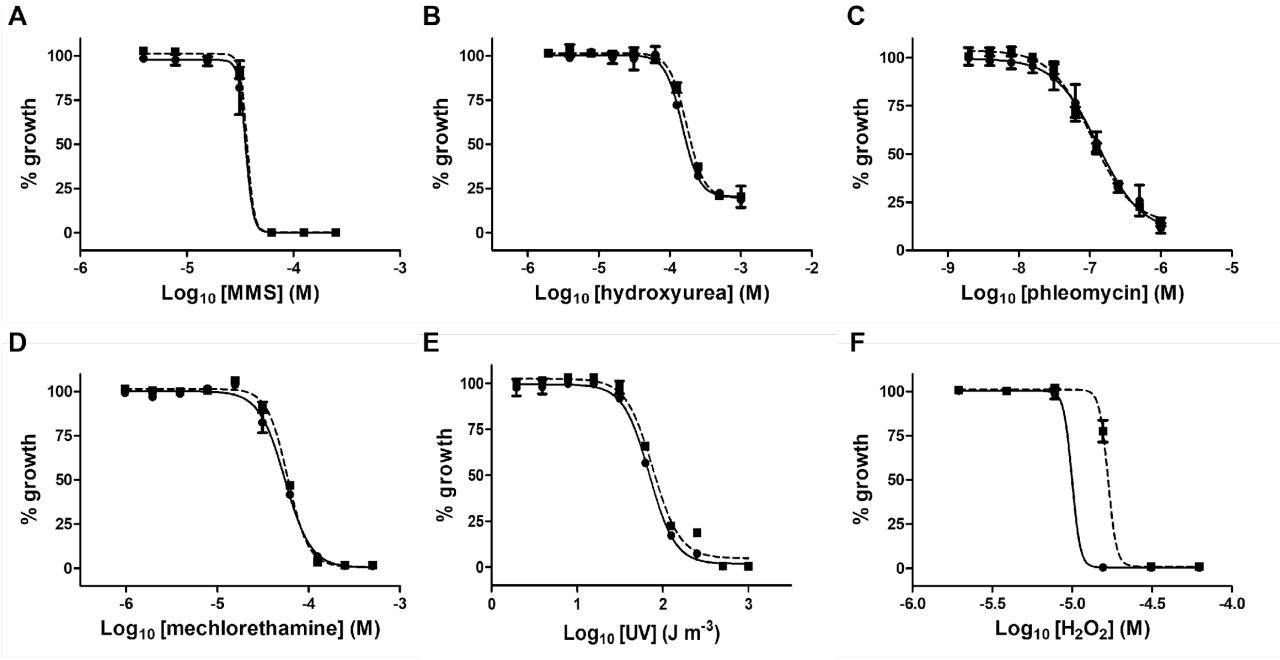


Fig. 5.27. Characterization of *T. brucei* Tbape1 null mutants.
Susceptibility of *T. brucei* wild type (solid line) and Tb*ape1Δ* (dashed line) to methyl methanesulfonate (MMS) (A), hydroxyurea (B), phleomycin (C), mechlorethamine (D), UV light (E) and H_2O_2 (F). All data are mean values ± standard deviations from experiments performed in quadruplicate.

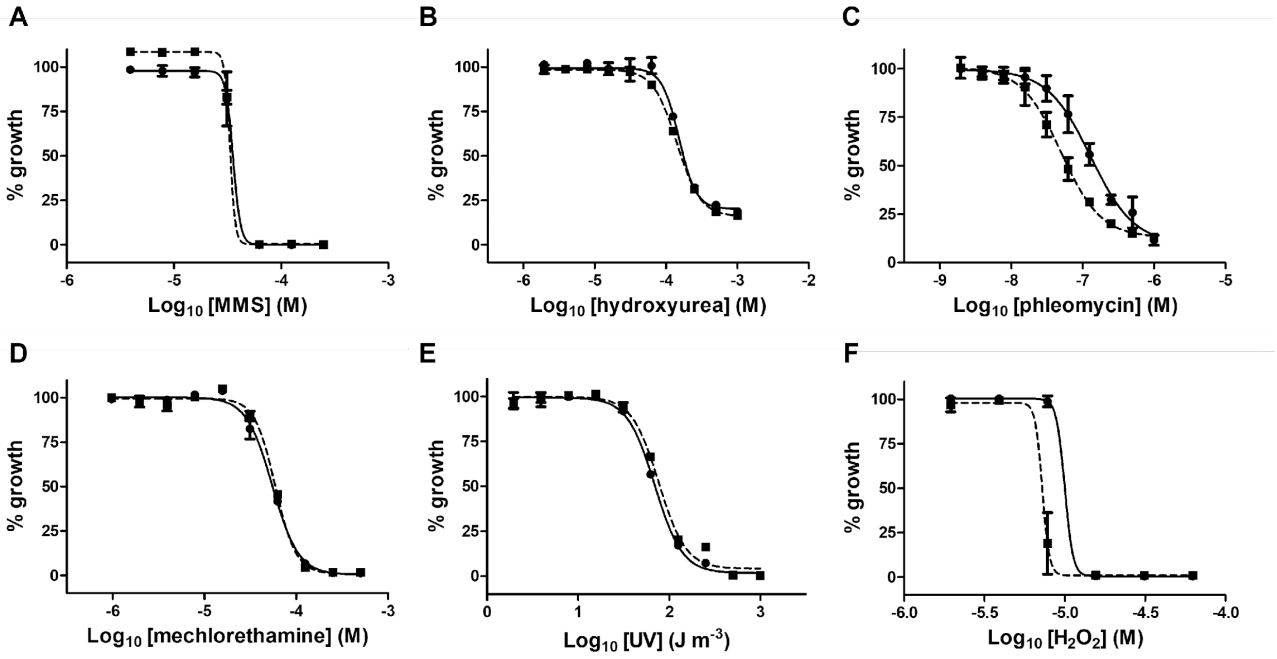


Fig. 5.28. Characterization of *T. brucei* Tb*myhΔ* "null" cells
Susceptibility of *T. brucei* wild type (solid line) and Tb*myhΔ* (dashed line) to methyl methanesulfonate (MMS) (A), hydroxyurea (B), phleomycin (C), mechlorethamine (D), UV light (E) and H_2O_2 (F). All data are mean values ± standard deviations from experiments performed in quadruplicate.

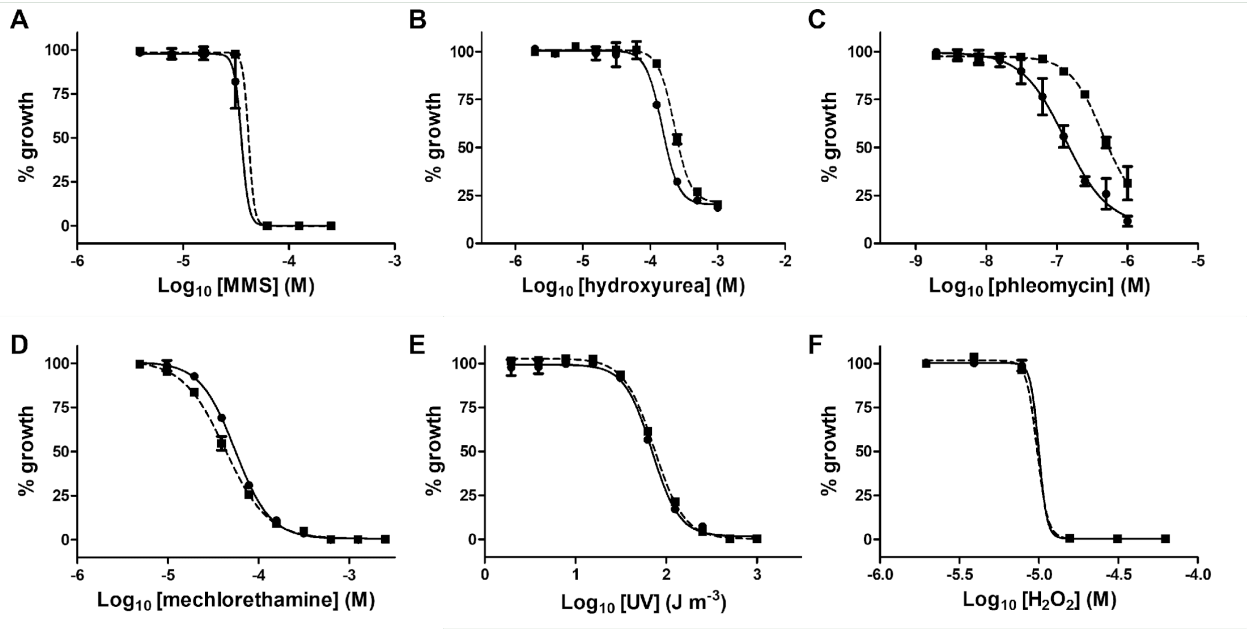


Fig. 5.29 Characterization of *T. brucei* Tb*nth1* "null" mutant.
Susceptibility of *T. brucei* wild type (solid line) and Tb*nth1Δ* (dashed line) to methyl methanesulfonate (MMS) (A), hydroxyurea (B), phleomycin (C), mechlorethamine (D), UV light (E) and H_2O_2 (F). All data are mean values ± standard deviations from experiments performed in quadruplicate.

Treatment	EC_{50}			
	wild type	Tb*ape1Δ*	Tb*myhΔ*	Tb*nth1Δ*
MMS (μM)	36.8 ± 3.2	36.4 ± 0.2 (1.0)	33.6 ± 1.0 (1.0)	44.5 ± 0.6 (+1.2)
Hydroxyurea (μM)	150.4 ± 3.0	173.8 ± 11.8 (+1.1)	141.8 ± 11.8 (-1.1)	229.9 ± 7.0* (+1.5)
Phleomycin (nM)	126.6 ± 16.0	105.9 ± 7.6 (-1.3)	49.8 ± 6.7$ (-2.5)	650.8 ± 95.3* (+5.1)
Mechlorethamine (μM)	54.6 ± 1.8	59.0 ± 1.9 (+1.1)	59.0 ± 1.1 (+1.1)	42.8 ± 2.8 (-1.3)
UV light (J m^{-3})	69.2 ± 0.8	76.8 ± 1.2 (+1.1)	77.2 ± 0.9 (+1.1)	73.5 ± 1.4 (+1.1)
H_2O_2 (μM)	10.5 ± 1.1	16.8 ± 0.5* (+1.6)	7.2 ± 0.4# (-1.4)	9.5 ± 0.2 (-1.1)

Table 5.6 Susceptibility of *T. brucei* mutants to DNA damaging agents.
The susceptibility of PCF *T. brucei* Tb*ape1Δ*, Tb*myhΔ* and Tb*nth1Δ* lines to various treatments, as judge by EC_{50} values extrapolated from dose response curves, were determined, and expressed as a fold difference (in parenthesis) relative to wild type (- and + indicates increased sensitivity or resistance, respectively). *, $ and # indicates significant difference in susceptibility (P<0.0001, P=0.0001 or P=0.0002, respectively) between wild type and mutant cells, as assessed by the on-line Student's *t* test calculator (https://www.graphpad.com/quickcalcs/ttest1/?Format=SD).

Based on the dose response curves/extrapolated EC_{50} values, trypanosomes lacking TbAPE1 are as equally sensitive to MMS, hydroxyurea, phleomycin, mechlorethamine and UV light as wild

type indicating that this DNA repair factor does not play a front-line role in resolving the resultant lesions generated by these treatments. In contrast, PCF Tb*ape1Δ* cells were slightly (1.5-fold) more resistant to H_2O_2. This phenotype could indicate that TbAPE1 could inhibit the AP lyase activity displayed by other enzymes including bifunctional DNA glycosylase (e.g., TbOGG1). When TbAPE1 is absent, any alternative activities are effectively freed up and able to participate more readily in mechanisms that operate to repair H_2O_2 induced base damage, analogous to the AP endonuclease independent BER systems noted in human cells (Wiederhold et al., 2004).

When the susceptibility of PCF cells with reduced levels of TbMYH was evaluated, the Tb*myhΔ* line behaved similarly to wild type with regards MMS, hydroxyurea, mechlorethamine and UV light treatment but were 1.4- and 2.5-fold more sensitive to H_2O_2 and phleomycin than controls, respectively. As the MYH family of DNA glycosylase is primarily involved in repairing oxidatively damaged based, it is not surprising that a reduction in the level of TbMYH in PCF parasites resulted in the phenotype observed towards the oxidant. However, the effect of this group of enzymes in repair of such lesions is usually masked by other DNA glycosylases such as OGG1 that have overlapping activities (Russo et al., 2004). Similarly, the phenotype displayed by the putative Tb*myhΔ* "null" toward phleomycin may reflect that this compound promotes H_2O_2 formation and base damage during its cytotoxic mechanism of action (Sleigh, 1976).

PCF cells that have disrupted *Tbnth1* were as equally sensitive to MMS, mechlorethamine, UV light and H_2O_2 as wild type while displaying 1.5-fold and 5.1-fold resistance to hydroxyurea or phleomycin, respectively. How TbNTH1 contributes to these resistance phenotypes is unclear but could suggest it may function in pathways outside of its role in BER, possibly involving its AP lyase activity. Intriguingly, the *T. cruzi* orthologue reportedly does not function as a DNA glycosylase with it possessing only as an AP lyase activity (Ormeño et al., 2016).

When the above studies were performed on the Tb*ape1Δ* and Tb*nth1Δ* lines and using benznidazole and nifurtimox, growth inhibition profiles reminiscent of wild type controls were obtained (Figure 5.30). This similarity in susceptibility plots extended across to *TbmyhΔ* cells treated with the nitroimidazole but not to the nitrofuran. Parasites with disrupted *Tbmyh* gene were > 4-fold more resistant to nifurtimox than controls. This may indicate that the reactive oxygen species generated following activation of nifurtimox by type II nitroreductases promotes formation of oxidatively damaged bases, that with disruption of *Tbmyh* are preferentially repaired by other components of the BER pathway, for example by the overexpression of OGG1.

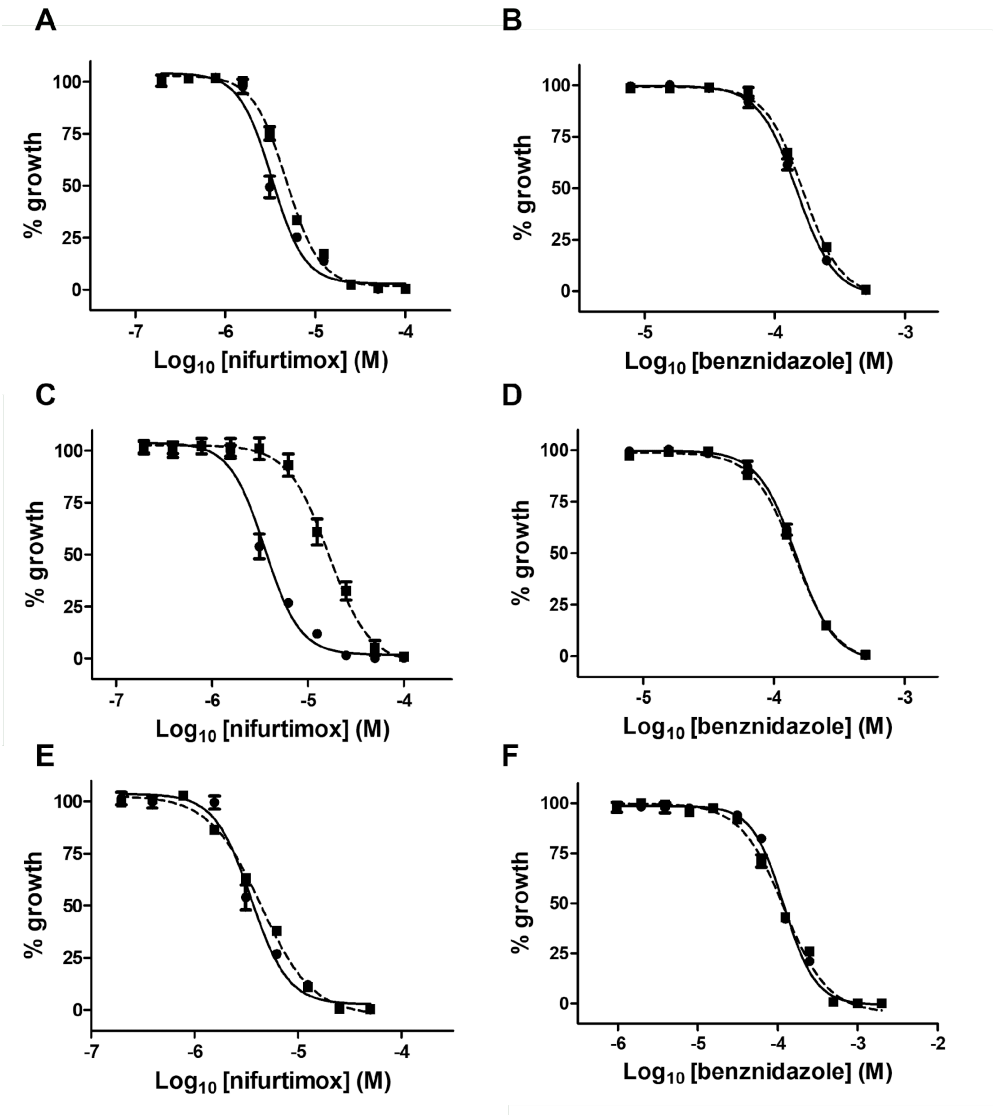


Fig. 5.30 Susceptibility of *T. brucei* mutants to trypanocidal nitroheterocyclic prodrugs.
A-F. The susceptibility of *T. brucei* wild type (solid line), Tb*ape1Δ* (**A & B**), Tb*myhΔ* (**C & D**) or Tb*nth1Δ* (**E & F**) (all dashed lines) to nifurtimox (**A, C & E**) and benznidazole (**B, D & F**) treatments were determined. All data points are mean values ± standard deviations from experiments performed in quadruplicate

5.8. Chapter Summary

In this chapter, we investigated the role *T. brucei* BER pathway in resolving DNA damage generated by trypanocidal nitroheterocyclic prodrugs, UV irradiation and ICL inducing agent(mechlorethamine). We also generated recombinant parasite lines that lacked single or both

copies of *Tbape1, Tbmyh* or *Tbnth1* and investigated the effect of DNA damage on survival of these lines.

Here we showed:

1. The sub-cellular localisation of mNeonGreen tagged TbAPE1, TbMYH and TbNTH1 in bloodstream and procyclic form *T. brucei.*
2. *T. brucei* BER factors respond in a temporal manner to UV irradiation and ICL as well to trypanocidal nitroheterocyclic prodrugs.
3. *Tbape1* is non-essential in *T. brucei* procyclic form parasites since it was possible to delete both copies. On the contrary single or double mutants of the genes encoding *Tbmyh* or *Tbnth1* still retained a copy of each gene in spite of the presence antibiotic resistance markers.
4. Downstream analysis of knockout *Tbape1Δ*, *Tbmyh or Tbnth1* lines revealed that the absence or reduction of these proteins does not affect parasite survival. In the presence of DNA damage different phenotypes were expressed. For example:
    - PCF Tb*ape1Δ* cells were 1.5-fold more resistant to $H_2O_{2.,}$ possibly because Tb*ape1Δ* lines mobilise alternative repair enzymes such as TbOG1 to restore these lesions.
    - *Tbmyh* lines were 1.4- and 2.5-fold more sensitive to H_2O_2 and phleomycin than wild type.
    - Reduction *Tbnth1* expression did not affect parasite survival, but instead caused increased resistance to hydroxyurea or phleomycin

CHAPTER 6: Analysis of the *T. brucei* NER Pathway.

6.1 Overview of nucleotide excision repair

The nucleotide excision repair (NER) pathway is responsible for the template-directed repair of bulky and helix distorting adducts that can be generated following exposure to certain chemical (e.g., cisplatin and psoralen) or environmental (e.g., UV light radiation) mutagens (Schärer, 2013; Marteijn et al., 2014; Petruseva et al., 2014; Kusakabe et al., 2019). The proteins that constitute this system are evolutionary conserved (Rouillon & White, 2011) and function in a stepwise process to remove the lesion. The first part of the pathway involves damage detection followed by excision of a short ssDNA segment containing the mutation with DNA polymerases synthesising a new DNA strand using the remaining undamaged homologous DNA as a template. Finally, the integrity of the DNA's sugar-phosphate backbone is restored through the action of DNA ligases (Cleaver et al., 2009; Rouillon & White, 2011). Deficiency in components of the NER pathway can result in build-up of DNA mutations leading to genome instability with this in higher eukaryotes associated with carcinogenesis, developmental abnormalities, and premature aging (Lazzaro et al., 2009; Oksenych & Coin, 2010; Rahimian et al., 2020). Such defects in humans can result in Cockayne syndrome, xeroderma pigmentosa and trichothiodystrophy (Vermeulen *et al*.,2001; Lehmann, 2003; Cameroni *et al*., 2010; Emmert *et al*.,2011).

Eukaryotes express two distinct NER sub-pathways (Figure 6.1). One is global genomic NER (GG-NER) that functions as a global surveillance system to remove lesions from any site within a genome irrespective of whether that sequence is being transcribed or not. The second mechanism is transcription-coupled NER (TC-NER) that operates to resolve damage encountered in transcribed DNA sequences. Each sub-pathway employs distinct sets of proteins to detect the DNA lesion which go on to recruit a common downstream cascade that mediate the actual repair. In mammalian GG-NER, XPC in conjunction with RAD23/CETN2 functions to recognise helix distorting mutations (Mu et al.,1995; Araujo *et al*., 2000; Araki et al., 2001; Nishi et al., 2005) while for bulkier adducts (e.g. UV-induced pyrimidine dimers) XPC recruitment is directed by the activity of several chromatin remodelling and modifying enzymes in a process ultimately guided by the DDB1/DDB2-containing CUL4A ubiquitin E3 ligase complex, DDB1-CUL4A^{DDB2} (Felderg and Grossman, 1976; Chu and Chang, 1988; Apelt et al., 2021). The GG-NER lesion detection mechanisms then promote unwinding of the DNA helix which facilitates loading of the

multi-subunit transcription factor IIH (TFIIH) to the damage site. In contrast, lesion recognition in TC-NER is initiated by stalling of a transcription complex at the damaged site (Hanawalt, 2003; Hanawalt & Spivak, 2008; Rouillon & White, 2011). The factors CSA and CSB then interact with the stalled RNA polymerase causing the complex to translocate away from the lesion to facilitate recruitment of other DNA repair factors (e.g., TFIIH). As such, this allows DNA repair to take place without disengaging the RNA polymerase (Sarker et al., 2005).

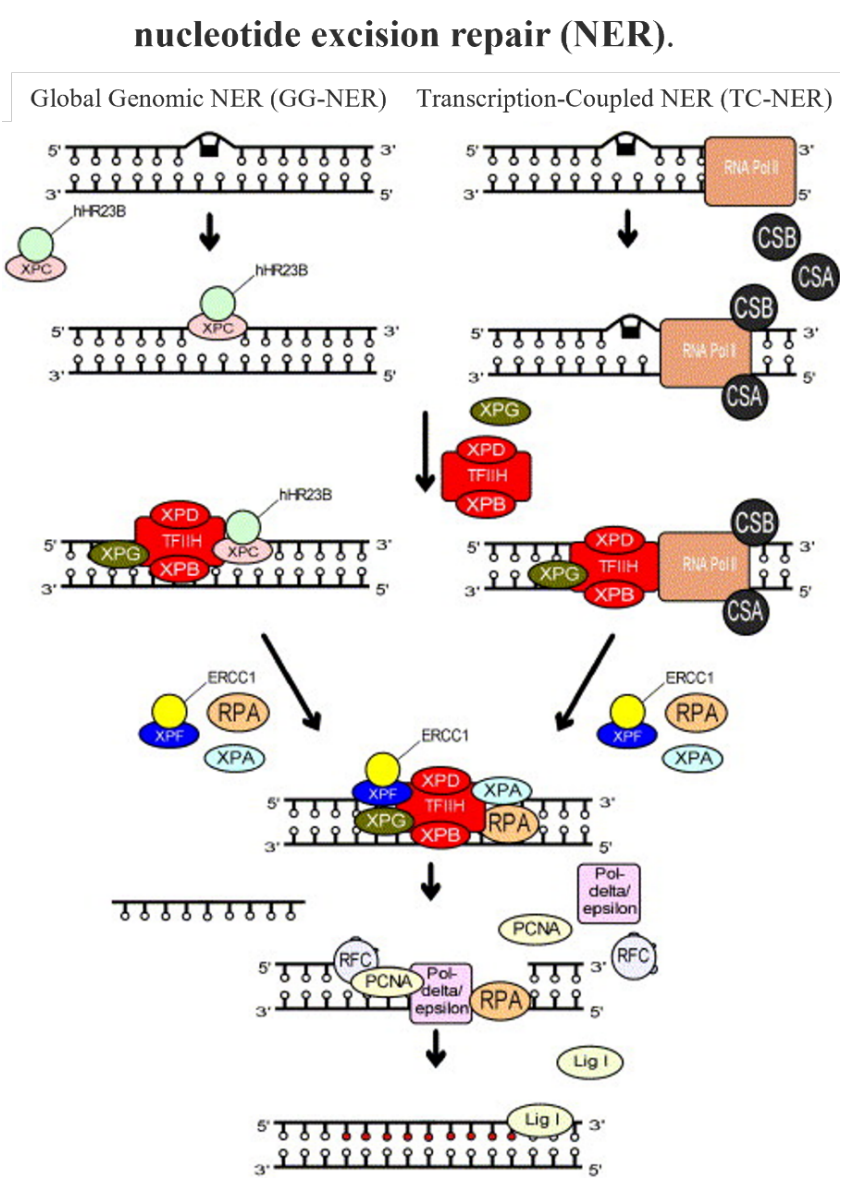


Fig. 6.1. Mechanism of nucleotide excision repair (NER). Nucleotide excision repair occurs through two sub-pathways, the global genomic NER (GG-NER) and transcription-coupled NER (TC-NER). The NER mechanism consists of the excision of a short oligonucleotide containing the DNA damage and replacement of the excised fragment by template directed synthesis using the non-damaged DNA strand. GG-NER detects DNA lesions in silent/non-transcribed regions of the genome or non-transcribed regions of actively transcribed genes. TC-NER removes lesions from the transcribed strand of active genes XPA–G, xeroderma pigmentosum complementation group A–G; hHR23B, human homologue of yeast RAD23B; RNA Pol II, RNA polymerase II; CSA and CSB, Cockayne syndrome factors A and B; TFIIH, general transcription factor IIH; ERCC1, excision repair cross complementing group 1 protein; RPA, replication protein A; PCNA, proliferating cell nuclear antigen; RFC, replication factor C; Polδ/ε, DNA polymerase delta/epsilon; Lig1, DNA ligase 1 (Olsen et al., 2005).

Following lesion detection, both sub-pathways then proceed via a common set of effector proteins beginning with recruitment of TFIIH complex and its co-factors. The TFIIH helicase sub-

units XPB and XPD extend the ssDNA region to form a bubble around the DNA lesion (Evans et al., 1997) that can accommodate other NER factors at the damage site to form the pre-incision complex (Min & Pavletich, 2007). Here, RPA and XPA load onto and protect the ssDNA in the bubble with this leading to release of the TFIIH complex and concomitant recruitment of the incision complex components, TbXPF/ERCC1 and XPG. These endonucleases then catalyse cleavage at sites 5′ and 3′ to the DNA lesion generating a 25–30 nucleotide gap (Mu et al., 1996; Tapias et al., 2004). Next, RFC and PCNA localise on the damage site and oversee the loading of the DNA synthesis machinery. DNA polymerases (Polδ in conjunction with Polκ or Polε alone) then regenerate the missing sequence using the remaining, undamaged DNA strand as template (Ogi et al., 2010) with the newly synthesised nucleic acid joined to the existing one via DNA ligase I or XRCC1/ DNA ligase III activities that function in dividing or non-dividing cells, respectively ((Moser et al 2007; Fagbemi et al., 2011).

Trypanosomes possess functional NER pathway that resembles the systems employed by higher eukaryotes although with some key differences (Table 6.1) (Machado et al., 2014). For example, informatic studies on *T. brucei* sequence data held on various databases indicates this parasite lacks orthologues of DDB2 from GG-NER, CSA from TC-NER and XPA from the pre-incision complex (Genois et al 2014; Machado et al., 2014). Additionally, the *T. brucei* TFIIH appears to be distinct (Lee et al., 2009). Here, several components of the transcription factor including the factors that form CAK complex, are apparently absent with the parasite expressing trypanosome-specific proteins (e.g., TSP1 and TSP2) or alternative paralogs (e.g., XPB and XPBz) instead (Lecordier et al., 2007; Lee et al., 2009; Badjatia et al., 2013). Functional studies have shown that TC-NER is the main NER mechanism employed by *T. brucei*, perhaps reflecting the highly transcribed nature of this parasite's genome (Machado et al., 2014). However, only certain factors (CSB, XPBz and XPG) appear to contribute to this process resulting in a system that can efficiently resolve bulky adducts generated by chemicals such as cisplatin but functions poorly in the repair of UV induced damage (Machado et al., 2014). As such, this suggests that other trypanosomal pathways may operate to fix the form(s) of lesion caused by UV. In the case of trypanosomal GG-NER mechanism, this does not significantly contribute to the repair of UV or cisplatin induced damage and appears to function primarily in the resolution of ICLs suggesting that this sub-pathway acts predominantly in genome maintenance, although how has yet to be established (Machado et al., 2014).

In this chapter, we build on the above to study the role of NER factors, focusing on DDB1, RAD23 and XPC from GG-NER, CSB from TC-NER and the endonucleases ERCC1, XPF and XPG (Table 6.1). We employed the PCR-based CRISPR/Cas9 genome editing strategy (Beneke et al., 2017)) to generate parasites expressing tagged versions of these enzyme while RNA interference (RNAi) was used to construct *T. brucei* lines expressing lowered levels of a targeted transcript. The cells were used to assess protein subcellular location, evaluate how protein expression is affected by exogenous DNA

Pathway	Enzyme	Function	Human*	*T. brucei*#
GG-NER	CETN2	Damage recognition; forms complex with XPC	NP_004335	Tb927.8.1080
	DDB1	**Damage recognition; forms complex with DDB2**	**NP_001914**	Tb927.6.5110
	DDB2	Damage recognition; recruits XPC	NP_000098	X
	RAD23A	**Damage recognition; forms complex with XPC**	**NP_005044**	**Tb927.6.4650**
	RAD23B	Damage recognition; forms complex with XPC	NP_002865	X
	XPC	**Damage recognition**	**XP_047304820**	**Tb927.9.11930**
TC-NER	CSA	Ubiquitin ligase complex; interacts with CSB & p44 of TFIIH	AAA82605	X
	CSB	**Transcription elongation factor; involved in transcription coupling/chromatin remodelling**	**NP_000115**	**Tb927.7.4080**
	XAB2	Damage recognition; interacts with XPA, CSA, & CSB	NP_064581	Tb927.5.1340
	UVSSA	Stabilization of CSB	NP_065945	X
TFIIH	XPD	ATPase and helicase activity; TFIIH subunit	NP_000391	Tb927.8.5980
	XPB	ATPase and helicase activity; TFIIH subunit	XP_011509096	Tb927.3.5100
	XPBz	ATPase and helicase activity; TFIIH subunit	+	Tb927.11.16270
	GTF2H1	TFIIH subunit	NP_005307	Tb927.11.9430
	GTF2H2	TFIIH subunit	NP_001506	Tb927.8.6540
	GTF2H3	TFIIH subunit	NP_001507	Tb927.11.16070
	GTF2H4	TFIIH subunit	NP_001508	Tb927.10.5210
	GTF2H5	TFIIH subunit	NP_997001	X
	TSP1	TFIIH subunit	+	Tb927.1.1080
	TSP2	TFIIH subunit	+	Tb927.11.14110
	CCNH	CAK subunit	NP_001230	X
	MNAT1	Stabilizes CAK complex	NP_002422	X
	CDK7	CAK subunit	NP_001790	X
	MMS19	Interacts with XPD & XPB	NP_071757	Tb927.8.3500/Tb927.8.3920
Pre-incision complex	XPA	Damage recognition	NP_000371	X
	RPA1	Subunit of RFA complex	NP_002936	Tb927.11.9130
	RPA2	Subunit of RFA complex	NP_002937	Tb927.5.1700
	RPA3	Subunit of RFA complex	NP_002938	Tb927.9.11940
incision	**ERCC1**	**Involved in incision on 3' side of damage: forms complex with XPF**	**NP_001974**	**Tb927.7.2060**
	XPF	**Involved in incision on 3' side of damage: structure specific endonuclease**	**NP_005227**	**Tb927.5.3670**
	XPG	**Involved in incision on 5' side of damage; stabilizes TFIIH; structure specific endonuclease**	**NP_000114**	**Tb927.9.11760**
Ligation	LIG1	Final ligation in dividing cells	NP_000225	Tb927.6.4780
	LIG3	Final ligation in non-dividing cells	NP_039269	X

Table 6.1. Components of the human and *T. brucei* nucleotide excision repair pathway.
T. brucei has potential to express orthologues of most NER components. * and # represent gene sequence identifiers from the NCBI or TriTrypDB respectively. X represents sequences present in humans but not in *T. brucei*. + represents sequences present in *T. brucei* but not in humans. Enzymes highlighted in red correspond to those analysed in the remainder of this chapter. Modified from Genois et al., 2014.

6.2 Informatic analysis of the *T. brucei* nucleotide excision repair pathway

6.2.1 Informatic analysis of TbCSB

Literature searches (Genois et al 2014; Machado et al., 2014; Dattani & Wilkinson, 2019), coupled with analysis of data held on the TriTryp database identified a single 3381 bp long ORF (Tb927.7.4080) as coding for the Tc-NER factor CSB. This gene, designated as Tb*csb*, is located on chromosome 7 of the parasite's nuclear genome and has potential to encode for a 126 kDa protein, designated as TbCSB, containing 1126 amino acids. Pairwise analysis revealed that Tb*csb* is 60% identical to Tc*csb* from the *T. cruzi* CL Brener Esmeraldo-like haplotype (TcCLB.508675.20), 49 to 51% identical to the equivalent gene from *C. fasciculata* (CFAC1_110017400), *L. donovani* (LdBPK.14.2.000900) and *L. major* (LmjF.14.0840), and 39 to 44% identical to *S. cerevisiae rad26* (X81635), *H. sapiens* csb/ercc6 (NM_000124) and *A. thaliana chr8* (NM_001335605) counterparts. As with other members of the CSB/ERCC6/RAD26 family, TbCSB possess a SNF2-related domain (PF00176) towards its centre (residues 420-742) with this region containing the characteristic DExD/H box, and a helicase conserved C-terminal domain (PF00271) located towards its carboxyl terminal (residues 798-911) (Figure. 6.2). These domains contain a series of motif-forming residues (designated motifs I, IA, II, III IV, V and VI) that facilitate ATP and DNA binding. Initial attempts to generate a 3D prediction using full-length TbCSB generated a model that was deemed unreliable over a sizeable portion (25%) of the protein (Figure 6.3A and B). This was most evident in disordered amino and carboxyl termini regions that flanked two central globular regions. Instead, a partial approximation covering residues 345 to 939 and based on *S. cerevisiae* RAD26 (5VVR; Xu et al 2017) was generated (Figure 6.3C). Comparison of the two TbCSB models confirmed a similar structure within the two central globular regions (Figure 6.3D and E) onto which, in the case of the partial model, the SNF2 and helicase domains could be mapped (Figure 6.3F). Both globular regions are predicted to be formed of five parallel β sheets surrounded by α helices with one region principally made up of the SNF2 domain while the other contains the carboxyl portion of the SNF2 domain, and the entire helicase domain.

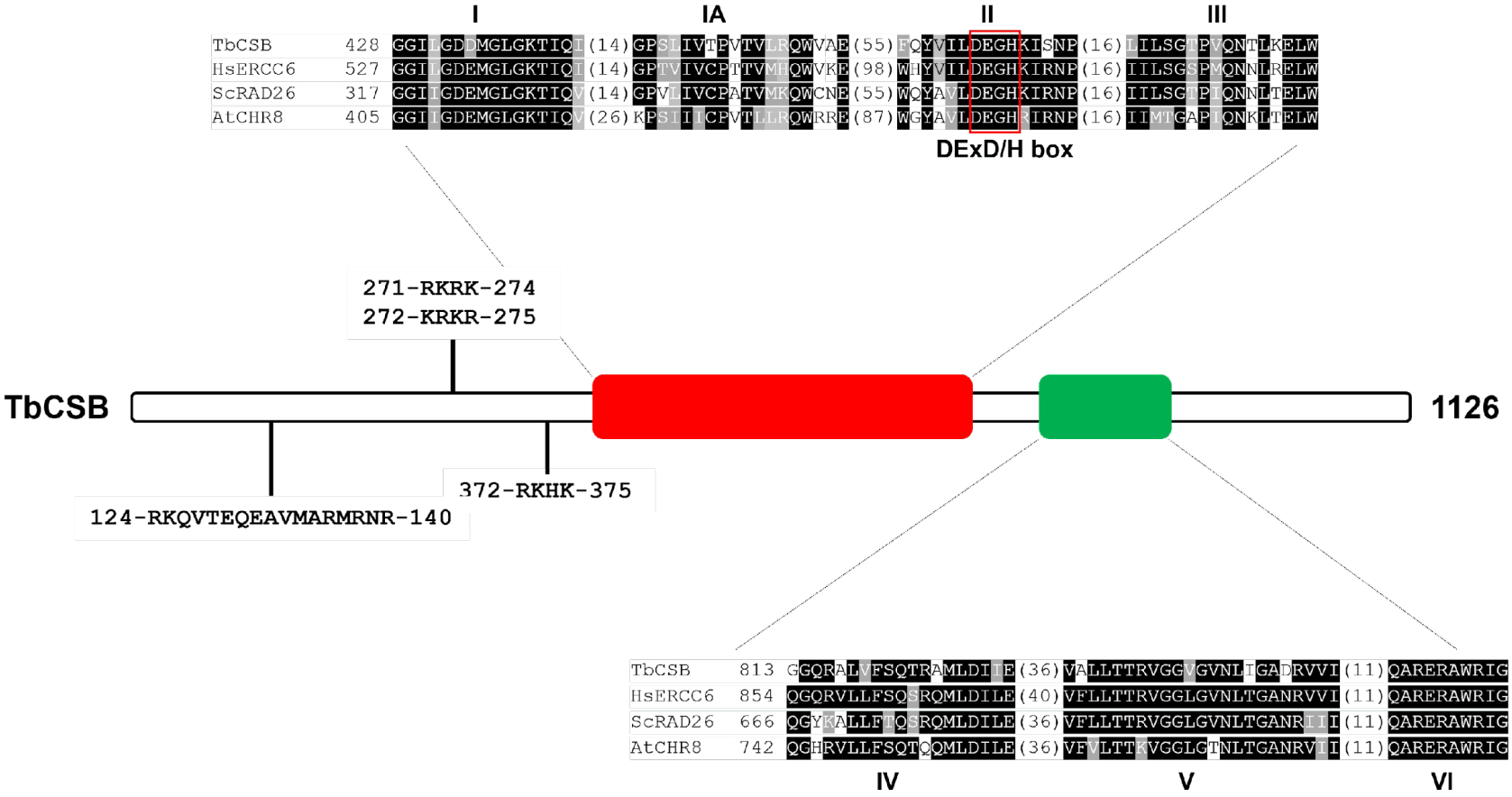


Fig. 6.2. Informatic analysis of TbCSB

The sequence corresponding to the SNF2-related (PF00176; residues 420-742; red box) and Helicase conserved C-terminal (PF00271; residues 798-911; green box) domains of TbCSB (XP_846031) were aligned with the equivalent regions from other members of the CSB family of helicase: *H. sapiens* HsERCC6 (NP_000115), *S. cerevisiae* ScRAD26 (P40352), and *A. thaliana* AtCHRB (NP_001318246). Residues that are identical in at least three of the sequences and conserved substitutions are highlighted in black and grey, respectively. The motifs I, IA, II (containing the DEXH sequence (orange box) and III in the SNF2-related domain and motifs IV, V and VI in the Helicase conserved C-terminal domain that mediate ATP and dsDNA binding plus the position of putative nuclear targeting signals (three 4 pattern and one Robbins & Dingwall consensus sequence) are shown.

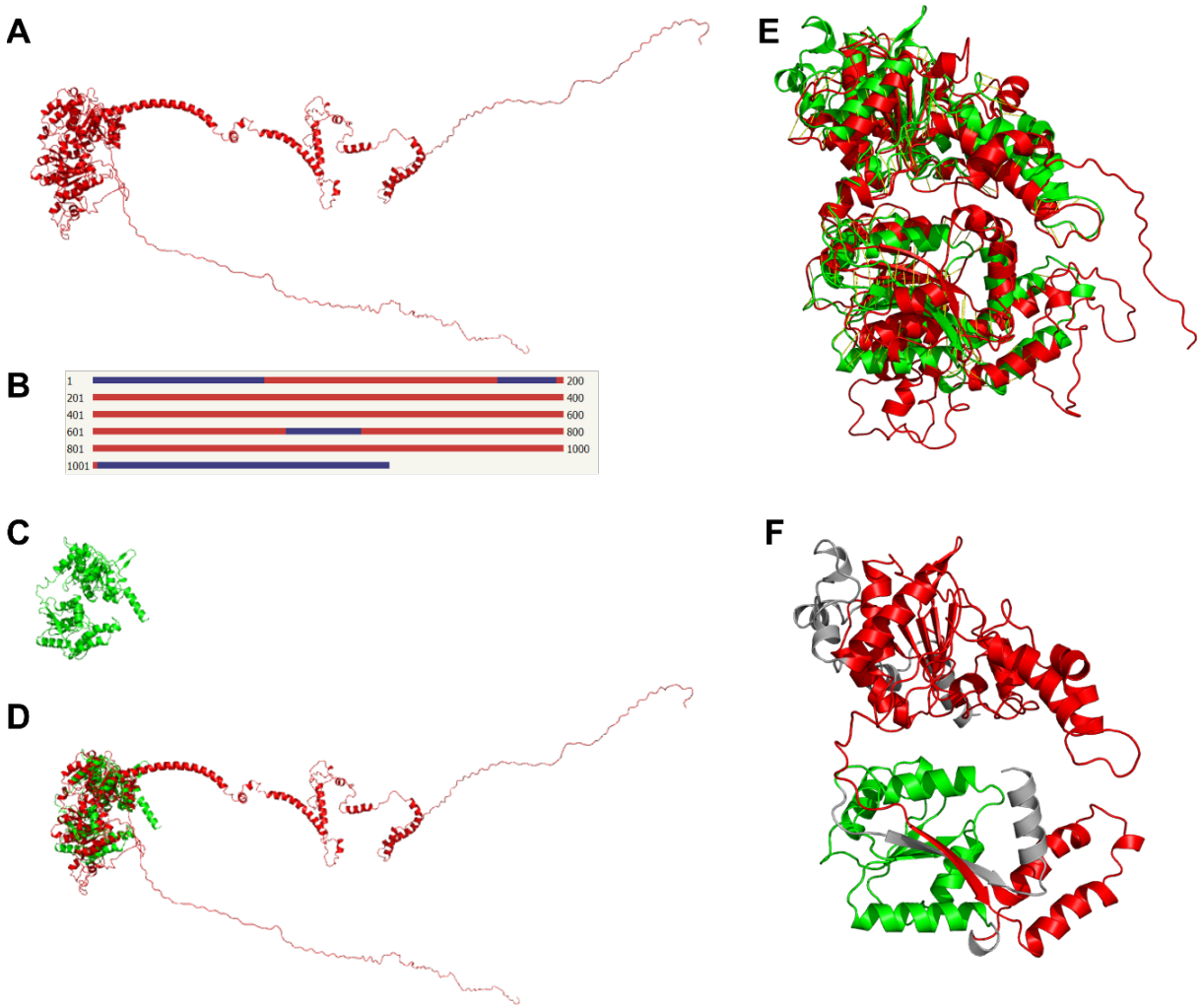


Fig 6.3. 3D prediction model of TbCSB.
A & C. Full-length (**A**) or partial (**C**) TbCSB (XP_846031) models were constructed using Phyre2 (Kelley et al., 2015) and visualized using PyMOL. The full-length approximation (**A**) was generated from six templates (6TDA, 7OOP, 6GEJ, 3MWY, 5VVR & 5O9G), with 867 residues (77%) modelled at >90% confidence. The model of 254 residues was deemed highly unreliable with these being primarily found at the protein's amino and carboxyl termini (see **B**). Such residues are predicted to have a disordered arrangement. For the partial approximation (**C**), based on *S. cerevisiae* RAD26 (5VVR) crystal structure, there was 100% confidence over 45% of residues modelled.
B. Schematic showing the degree of confidence in the TbCSB full-length model (A). Areas in red and blue correspond to regions where there is high or low confidence in the model, respectively.
D & E. Comparison of the full-length (red) and partial (green) 3D predictions of TbCSB. **D** represents a comparison between the entire models while **E** is of the domain-containing regions. The yellow lines indicate differences between the two models at the same point: these are more readily observed in **E**.
F. The position of the SNF2-related (PF00176; red box) and Helicase conserved C-terminal (PF00271; green box) domains in the partial TbCSB model (B). Non-domain regions are noted in grey.

To assess how TbCSB may interact with DNA, the partial model of the parasite enzyme was taken and aligned with the *S. cerevisiae* RAD26 component found in the cryo-EM structure of the ternary complex of RNA Pol II, transcription scaffold and Rad26 while bound to DNA (Xu et al

2017). Following alignment, the positions of the various yeast proteins were removed leaving TbCSB and nucleic acid in place (Figure 6.4). This indicates that DNA sits at a cleft between the two globular regions and tends to associate with a fold containing part of the SNF2 and the entire helicase domains.

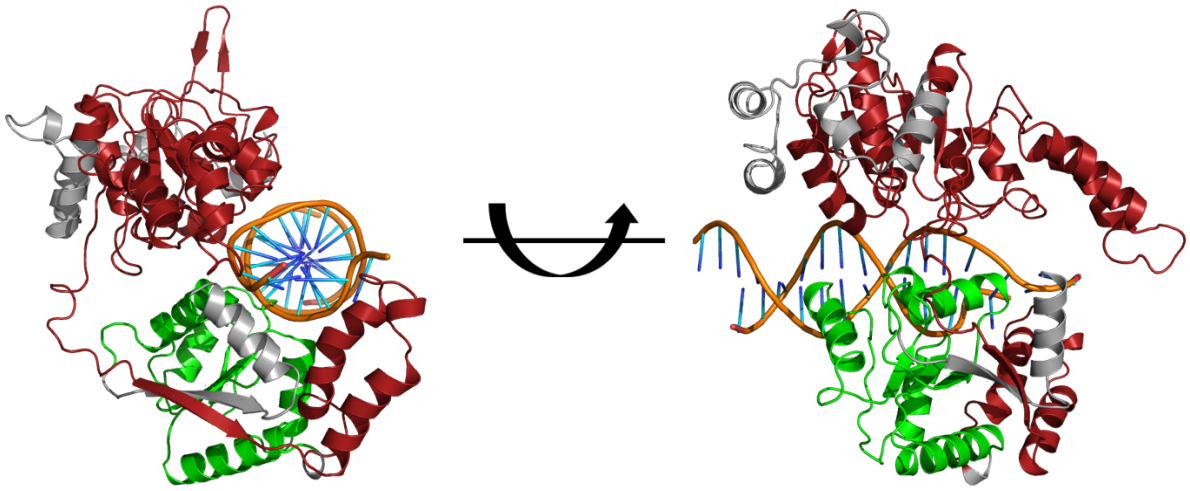


Fig 6.4. Interaction of TbCSB with DNA.
To assess how TbCSB interacts with DNA, the partial model (red) was aligned with the *S. cerevisiae* RAD26 component in the cryo-EM structure of the ternary complex of RNA Pol II, transcription scaffold and Rad26 (5VVR; Xu et al 2017). The yeast proteins were removed from the alignment, but DNA retained. The positions of the SNF2-related (red) and Helicase conserved C-terminal (green) domains are noted.

6.2.2 Informatic analysis of TbDDB1

Literature searches (Machado et al., 2014) and analysis of *T. brucei* data held at TriTrypDB identified a single 3813 bp ORF (Tb927.6.5110) from the *brucei* TREU927 strain coding for DDB1. This sequence, designated as Tb*ddb1*, is found on chromosome 6 of the parasite's nuclear genome and can encode for a 138 kDa protein, designated as TbDDB1, of 1270 amino acids. Pairwise analysis revealed Tb*ddb1* is between 49 to 50% identical to the equivalent genes from other Trypanosomatids (e.g., *C. fasciculata* (CFAC1_260062800), *L. donovani* (LdBPK.30.2.003770) and *L. major* (LmjF.30.3710) and is 44 and 47% identical to its *S. pombe* (NM_001019012) and *H. sapiens* (NM_001923) counterparts: *T. cruzi* CL Brener Esmeraldo-like haplotype (and other strains) apparently lack a complete *ddb1* ORF. Analysis of the primary sequence demonstrated that, as with other members of this family, TbDDB1 contains two domains, possessing an MMS1_N domain (PF10433) between residues 202 to 603 followed by a CPSF_A domain (PF03178) between residues 803 to 1183 (Figure 6.5).

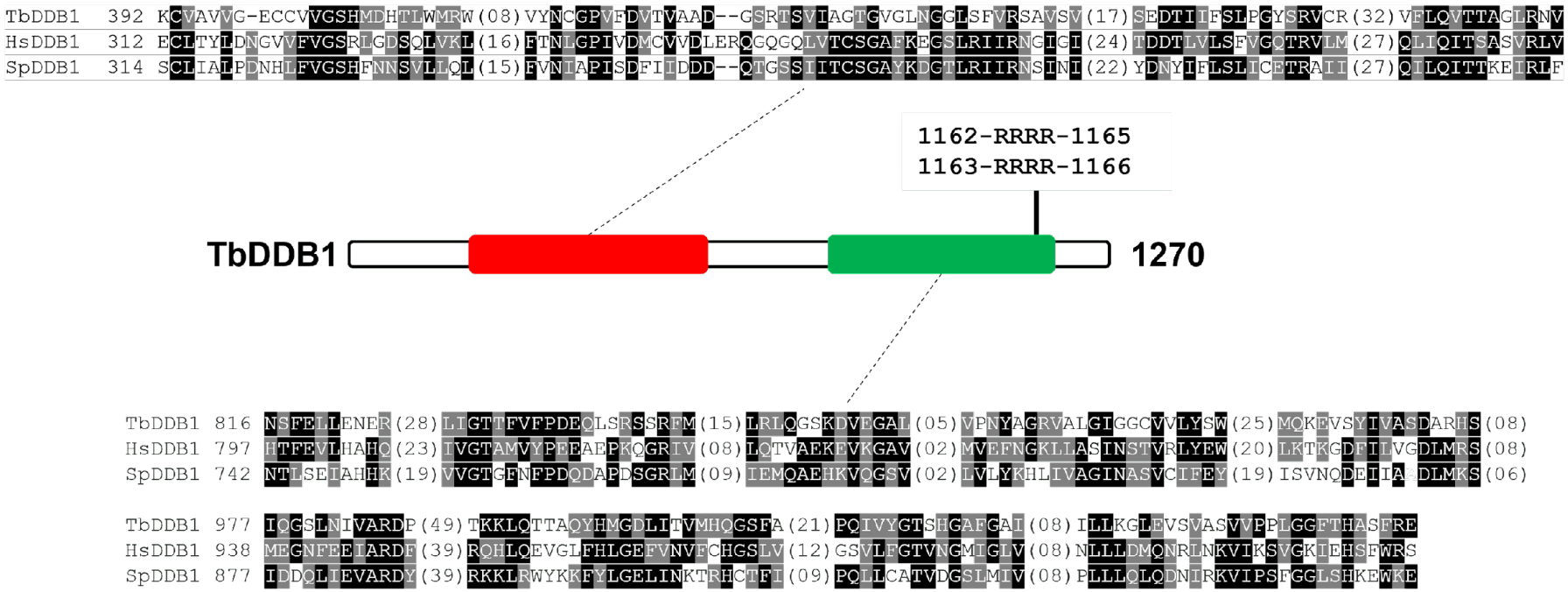


Fig 6.5 Informatic analysis of TbDDB1

The sequence corresponding to the mono-functional DNA-alkylating methyl methanesulfonate N-term (MMS1_N; PF10433; residues 202-603; red box) and cleavage/polyadenylation specificity factor A (CPSF_A; PF03178; residues 803-1183; green box) domains of TbDDB1 (XP_845628) were aligned with the equivalent regions from *H. sapiens* DDB1 (HsDDB1; NP_001914) and *Schizosaccharomyces pombe* DDB1 (NP_593580). Residues that are identical or conserved are highlighted in black and grey, respectively. The position of two 4 pattern putative nuclear targeting signals are shown.

3D predictions of full-length TbDDB1 revealed the trypanosomal protein can be split into four distinct structural regions (Figures 6.6A and B). Its carboxyl terminal, which consists principally of α-helices, is predicted to be located at the centre of the model where it forms the so-called carboxyl terminal domain (CTD). By analogy to the human orthologue, this is believed to form a rigid anchor that helps secure the other structural units of the protein (Li et al., 2006). The remaining regions of TbDDB1 are predicted to be structurally related and radiate out from CTD. Each section is predicted to form β-propeller structural unit and are designated as BPA, BPB, and BPC (Figures 6.6B and C). Analysis of each β-propeller revealed they are entirely formed of β-sheets with four sheets having an anti-parallel, zigzag arrangement to form a blade and each β-propeller then containing seven blades (Figures 6.6C). Mapping of the domains onto the model demonstrated that the MMS1-N was partitioned across all three β-propeller regions while CPSF_A was restricted BPC (Figures 6.6A and B). The predicted TbDDB1 structure allows it to function as a scaffold with the blades of each β-propeller facilitating protein-protein interactions. For example, the β-propellers in the human DDB1 can readily associate with DDB2 (Scrima et al., 2008) or CUL4 (Angers et al., 2006; Li et al., 2006).

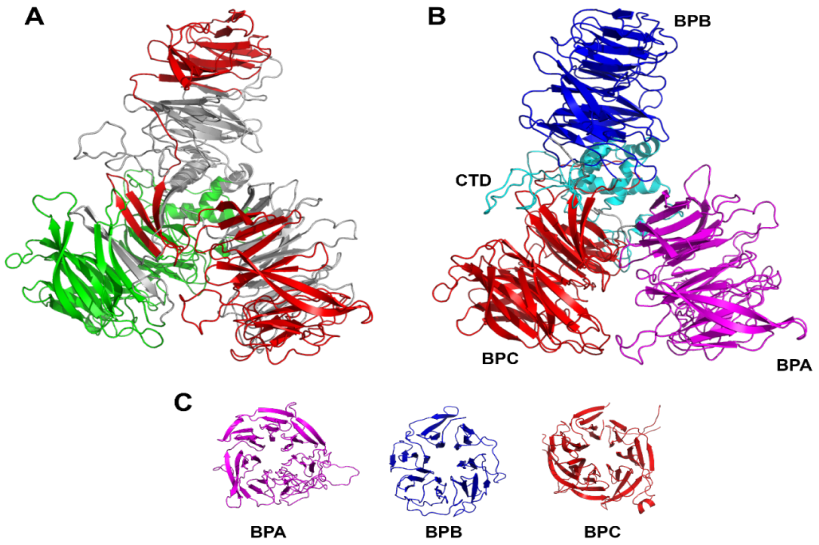


Fig 6.6. 3D prediction model of TbDDB1.
A 3D prediction of full-length TbDDB (XP_845628) was generated from six templates (6BM0, 5LGW, 3EI1, 5IFE, 5FDG, 5HY7), with 1181 residues (93%) modelled at >90% confidence. The position of the mono-functional DNA-alkylating methyl methanesulfonate N-term (MMS1_N; PF10433; red) and CPSF_A subunit (PF03178; green) domains in this model are highlighted. Non-domain regions are noted in grey.
B. The MMS1_N and CPSF_A domains form three β-propellers, designated as BPA (magenta), BPB (blue) and BPC (red). The position of these and the carboxyl terminal domain (CTD; cyan) in the TbDDB 3D prediction are highlighted.
C. Zoomed images showing the seven bladed structure of each β-propeller region.

6.2.3 Informatic analysis of TbRAD23

Analysis of published literature (Machado et al 2014; Genois et al 2014), and confirmed by BLAST analysis, identified a single 1071 bp long gene (Tb927.6.4650), designated as Tb*rad23*, located on chromosome 6 of the *T.b. brucei* TREU927 strain genome with potential to code for a 38 kDa RAD23 protein (designated as TbRAD23). Analysis revealed that Tb*rad23* was 55% identical to Tc*rad23* from the *T. cruzi* CL Brener Esmeraldo-like haplotype (Tb927.6.4650), between 48 to 50% identical to the equivalent genes from *C. fasciculata* (CFAC1_260058100), *L. donovani* (LdBPK.30.2.003350) and *L. major* (LmjF.30.3300), and between 44 to 48 % identical to its *S. cerevisiae* (L25428), *H. sapiens* (NM_002874.5) and *A. thaliana* (NM_101486) counterparts. Analysis using the HMMER algorithm revealed that TbRAD23 had a similar protein domain structure noted in orthologues expressed by other organisms (Figure 6.7). Based on its primary sequence TbRAD23 is predicted to contain a ubiquitin domain (UBL; PF00240) at its amino terminal (residues 3 to 72), two ubiquitin-associated domains (UBA-1 & UBA-2; PF00627) running from residues 142 to 179 and from 308 to 347 and a XPC binding domain (PF09280) positioned between residues 226 to 282.

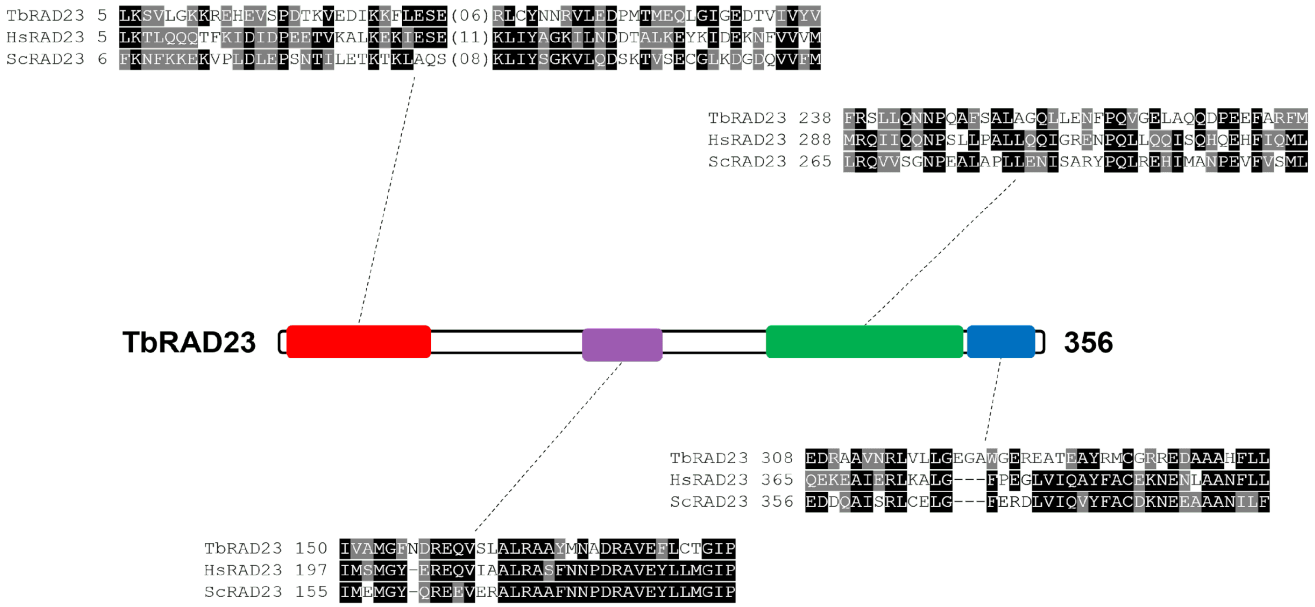


Fig 6.7 Informatic analysis of TbRAD23

The sequences corresponding to the ubiquitin-like (UBL; PF00240; residues 3-72; red box), two UBA protein (PF00627) (UBA-1: residues 142-179; purple box & UBA-2; residues 308-347; blue box) and XPC-binding (PF09280; residues 226-282; green box) domains of TbRAD23 (XP_845582) were aligned with the equivalent regions from the *H. sapiens* (HsRAD23; NP_002865) and *S. cerevisiae* (ScRAD23; AJU50350) orthologues. Residues that are identical in at least three of the sequences and conserved substitutions are highlighted in black and grey, respectively.

3D predictions of the full-length trypanosomal protein based on the human orthologue (Walters et al., 2003) showed that the four domains formed distinct structural units connected by flexible, unstructured linker regions (Figure. 6.8A). The two UBA and the XPC-binding domains have entirely α-helical arrangement possessing three or four helices, respectively, whereas the UBL domain contains four β-strands and an α-helix (Figure. 6.8B). For this latter domain, the first and last β-strands are oriented parallel to each other and anti-parallel to the central sheets. By analogy to the human RAD23 orthologue, the UBL domain can interact with the carboxy terminal α-helix of UBA-1 or UBA-2 leading to a 'closed' conformation (Walters et al., 2003). The same region of UBL can also interact with other proteins such as the proteasomal subunit S5a. In this situation, it is believed that ubiquitin can bind to the UBA's in a manner like that of UBL allowing RAD23 to adopt a more 'open' conformation'. In this state, RAD23 along with bound XPC could then be targeted for degradation thereby providing a mechanism in regulating protein levels.

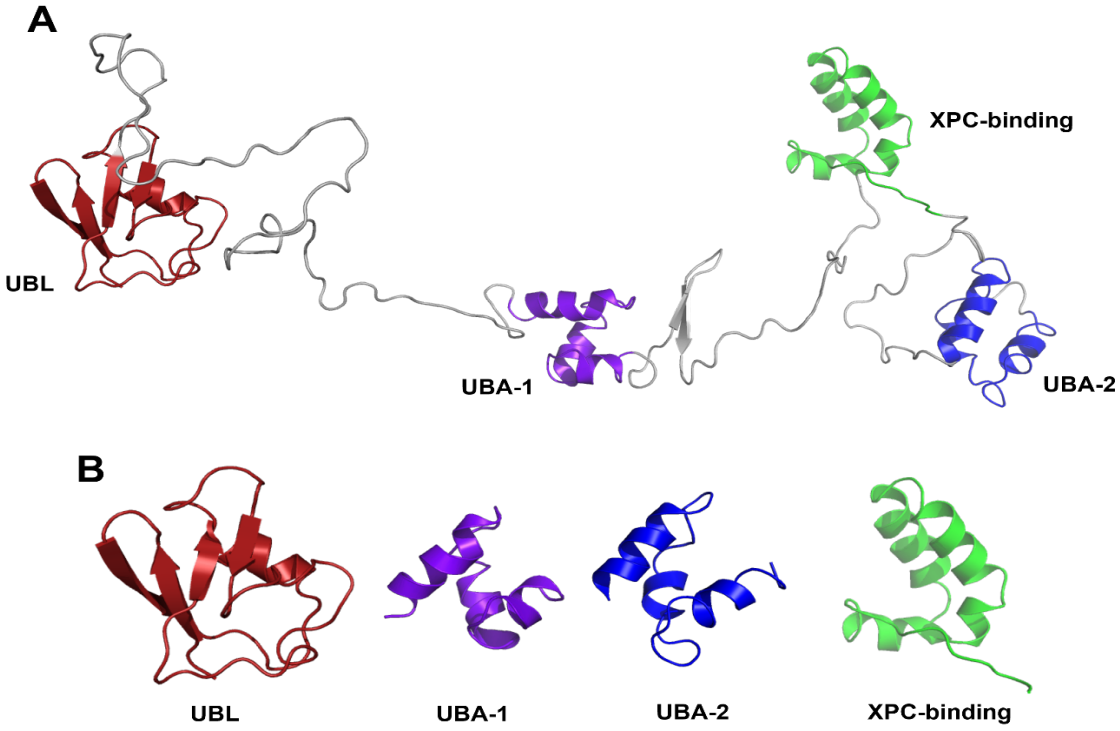


Fig 6.8. 3D prediction model of TbRAD23.
A. Full-length 3D prediction of TbRAD23 (XP_845582) was constructed using Phyre2 (Kelley et al. 2015) and visualized using PyMOL. The full-length approximation was generated from a single template (1OQY; Walters et al 2003), with 352 residues (99%) modelled at >90% confidence. More than half of the modelled residues (51%) are predicted to be in flexible, unstructured regions. The position of the UBL *(*PF00240; red), two UBAs (PF00627; UBA-1 in purple & UBA-2 in blue) and XPC-binding (PF09280; green) domains are shown. The four structural domains are connected by flexible, unstructured linker regions (grey).
B. Images showing the 4 structural domains in isolation.

6.2.4 Informatic analysis of TbXPC

Literature searches (Machado et al 2014; Genois et al 2014) coupled with analysis of data held on TriTrypDB identified an ORF (Tb927.9.11930) that could encode for an XPC orthologue. The 2307 bp long gene, designated as Tb*xpc*, is located on chromosome 9 of the *T.b. brucei* TREU927 strain genome and has potential to code for an 86 kDa protein, designated as TbXPC, of 768 amino acids. Pairwise analysis revealed that Tb*xpc* is 54% identical to Tc*xpg* from the *T. cruzi* CL Brener Esmeraldo-like haplotype (TcCLB.507011.140), 48 to 53% identical to the equivalent genes from *C. fasciculata* (CFAC1_300084900), *L. donovani* (LdBPK.35.2.003500) and *L. major* (LmjF.35.3450), and 45 to 50 % identical to its *S. cerevisiae* (M24928), *H. sapiens* (NM_004628) and *A. thaliana* (NM_121669) XPC/RAD4 counterparts. Domain searches revealed that TbXPC contains a Rad4 transglutaminase-like domain (TGD) (PF03835) located between 392 to 479 followed by three Rad4 beta-hairpin domains designated as BHD1 (PF10403), which is found between residues 525 to 576, BHD2 (PF10404) running from residues 580 to 628, and BHD3 (PF10405) located between residues 636 to 715 (Figure 6.9A). Structural studies on the full-length protein predicted that the amino terminal region of TbXPC (residues 1 to 190) was disordered (Figure 6.9B) and as such this region was excluded from the input sequence used in the construction of a 3D prediction. The truncated TbXPC model covering residues 190-768 and based on two *S. cerevisiae* Rad4 structures (Min & Pavletich 2007; van Eeuwen et al 2021) revealed that TGD forms a large lobe with the three BHDs extending away from it as a curved protrusion (Figure 6.10). To assess how TbXPC, TbRAD23 and DNA may interact, the models of truncated TbXPC and TbRAD23's XPC-binding domain were aligned with the *S. cerevisiae* RAD4-RAD23-DNA structure (Min & Pavletich, 2007a). The yeast proteins were then deleted from the alignment leaving the two trypanosomal sequences and nucleic acid in place (Figure 6.11). This revealed that the XPC-binding domain from RAD23 interacts with TbXPC's TGD while DNA appears to bind to the three BHDs. In yeast, RAD4 can bind to DNA in two ways: TGD and BDH1 both bind to undamaged dsDNA at a site (1-11 bp) adjacent to the mutation with BHD2 and BHD3 both binding to a 4 bp DNA segment that contains the lesion. The model generated here indicates that the three BDHs may be positioned to fulfil a similar role in TbXPC although whether its TGD is suitably situated to participate in this interaction is unclear. This may reflect an issue with the model or indicate a difference in the mechanism of action of the trypanosomal enzyme. Compared to its yeast counterpart

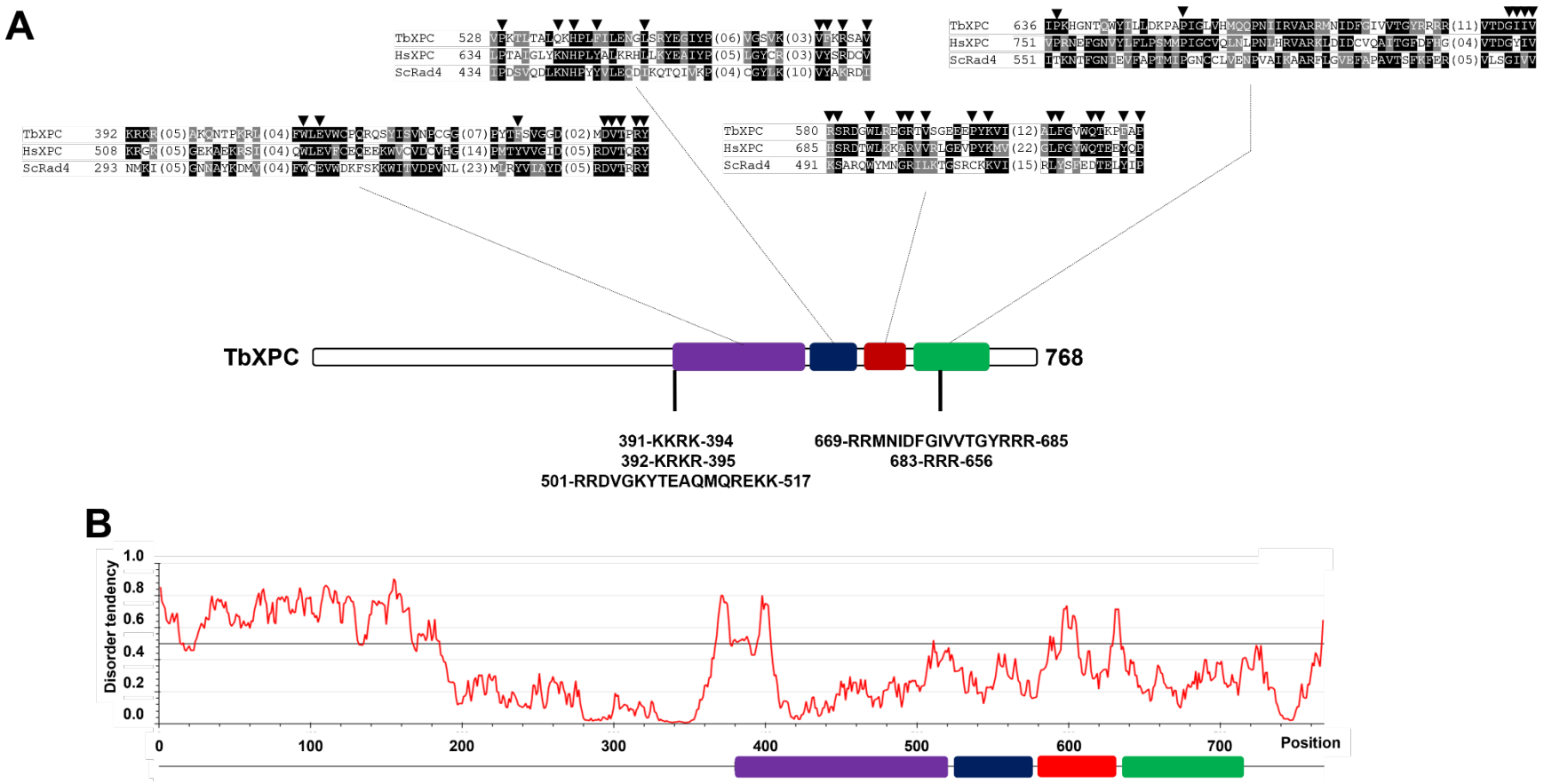


Fig. 6.9. Informatic analysis of TbXPC

A. The sequence corresponding to the Rad4 transglutaminase-like (TGD; PF03835; residues 381-520; purple box), Rad4 beta-hairpin domain 1 (BHD_1; PF10403; residues 525-576; blue box), Rad4 beta-hairpin domain 2 (BHD_2; PF10404; residues 580-628; red box) and Rad4 beta-hairpin domain 3 (BHD_3; PF10405; residues 636-715; green box) domains of TbXPC (XP_827449) were aligned with the equivalent regions from *H. sapiens* XPC (HsXPC; XP_047304820) and *S. cerevisiae* RAD4 (ScRAD4; AJV35225). Residues that are identical or are conserved substitutions are highlighted in black and grey, respectively. Amino acids marked with an arrow constitute key components in a domain's HMM logo. The location of five potential nuclear localisation signals (three 4 pattern and 2 Robbins & Dingwall consensus sequences) are shown.

B. Disordered regions were obtained by submitting the TbXPC sequence to disorder prediction program IUPRED2 (https://iupred2a.elte.hu/). A disorder value >0.5 indicates an unstructured protein region.

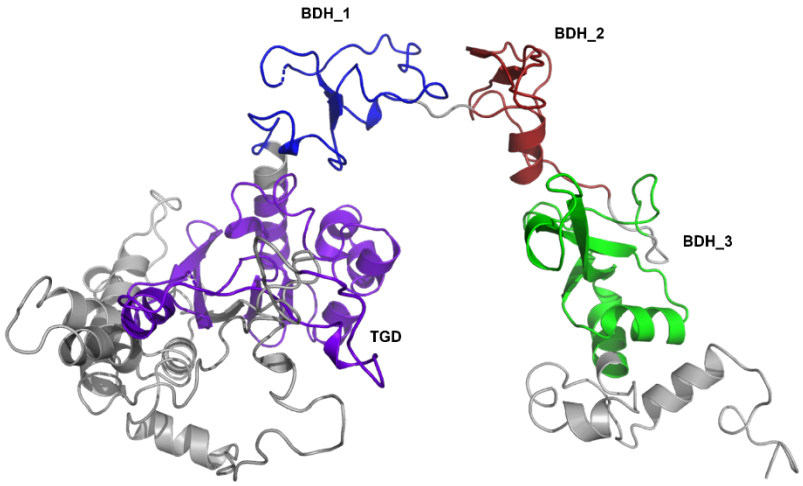


Fig 6.10. 3D prediction model of TbXPC.
A 3D prediction of a truncated version (residues 190-768) of TbXPC (XP_827449) was constructed using Phyre2 (Kelley et al. 2015) and visualized using PyMOL. The full-length approximation was generated from two templates (7K04, 2QSH), with 506 residues (66%) modelled at >90% confidence. This model lacks 190 residues present at the amino terminus that correspond to the previously described disorder region. The position of the TGD (purple) and three Rad4 beta-hairpin (BHD_1, BHD_2 & BHD_3 in blue, red, and green, respectively) domains are shown. Non-domain regions are noted in grey.

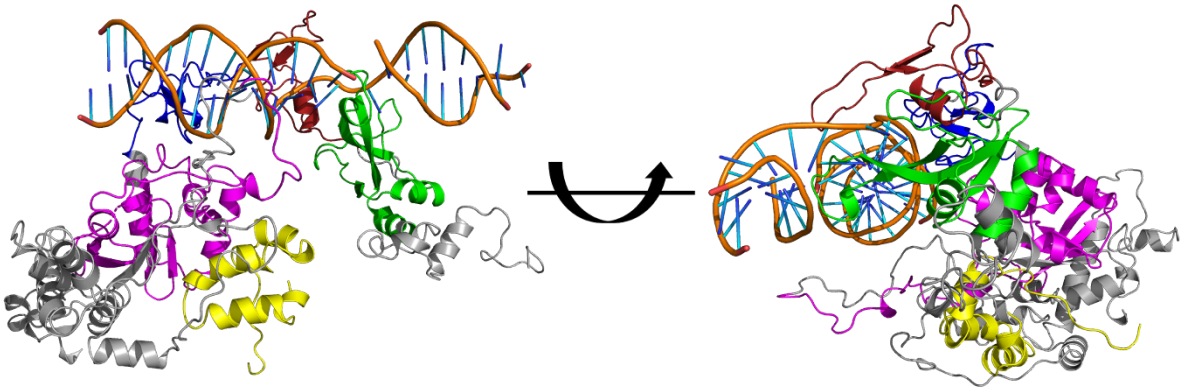


Fig 6.11. Interaction of TbXPC/TbRAD23 with DNA.
To assess how TbXPC/TbRAD23 interacts with DNA, the truncated TbXPC model and the XPC-binding domain from TbRAD23 were aligned with the structure of *S. cerevisiae* Rad4-Rad23 bound to a mismatch DNA (2QSH; Min & Pavletich 2007). The yeast sequences were removed from the alignment, but the DNA retained. The positions of the TGD (magenta), BHD_1 (blue), BHD_2 (red) & BHD_3 (green) domains plus XPC-binding domain from TbRAD23 (yellow) are noted.

6.2.5 Informatic analysis of TbERCC1

An 801 bp ORF (Tb927.7.2060) located on chromosome 7 of the *T. b. brucei* TREU927 genome) and with potential to encode for 29 kDa ERCC1 was identified following literature searches (Machado et al., 2014; Genois et al., 2014). Pairwise analysis revealed the gene, designated as Tb*ercc1*, was 51 to 57% identical to the orthologues expressed by *T. cruzi* CL Brener

Esmeraldo-like haplotype (TcCLB.510165.20), *C. fasciculata* (CFAC1_240006800) and *L. major* (LmjF.22.0070), and between 39 to 45% identical to its *S. cerevisiae* (NM_001182454), *H. sapiens* (NM_001369409) and *A. thaliana* (NM_111394) *ercc1*/*rad10* counterparts. As with other members of the ERCC1/SW10 family, TbERCC1 contains an ERCC1/RAD10 DNA binding (PF03834) nuclease-like domain (NLD) towards its amino terminal (residues 8-132) and Helix-hairpin-helix motif (PF14520) towards its carboxy terminal (residues 166 to 212) (Figure. 6.12A). In the human protein, the NLD in enzymatically inactive but along with HHH does contribute to heterodimerization with XPF, and in DNA binding. The former interaction is mediated through a series of salt bridge-forming amino acid and hydrophobic interactions (Jones et al., 2020; Tsodikov et al., 2007). Comparison of the *T. brucei* and human ERCC1 sequences revealed that most residues involved in heterodimerization are missing from the trypanosomal protein. For example, at least 4 salt bridge-forming amino acids (K187, K193, D193, R207) and five hydrophobic residues (I190, T195, I197, Y208, Y212) in the NLD of human ERCC1 contribute to heterodimerization with XPF while the HHH domain participates in XPF binding *via* hydrophobic interactions in which F294 plays an essential role (Tsodikov et al 2005). Out of the ten residues noted above, TbERCC1 contains only three and lacks an equivalent residue to F294. In terms of DNA binding, again both domains contribute with several (at least 15) residues proposed to be implicated in this process ((Jones et al., 2020; Tsodikov et al., 2005). In this case, TbERCC1 contains half (7) of the residues (or their equivalent) implicated in binding the human protein to DNA with most of these (5 out of 6) found within the HHH domain.

Analysis of full-length TbERCC1 predicted that the carboxyl terminal region (residues 223 to 266) was disordered (Figure 6.12B) and as such this was excluded from the input sequence used in the construction of a 3D prediction. The resultant truncated model that covers residues 1-222, revealed that the NLD and HHH domains form distinct globular domains within the protein separated by a short linker region. Comparison of the TbERCC1 approximation with the human ERCC1 structure identified potential XPF interacting and DNA binding surfaces in NLD and HHH domains of the trypanosomal enzyme (Jones et al., 2020; Tsodikov et al., 2005).

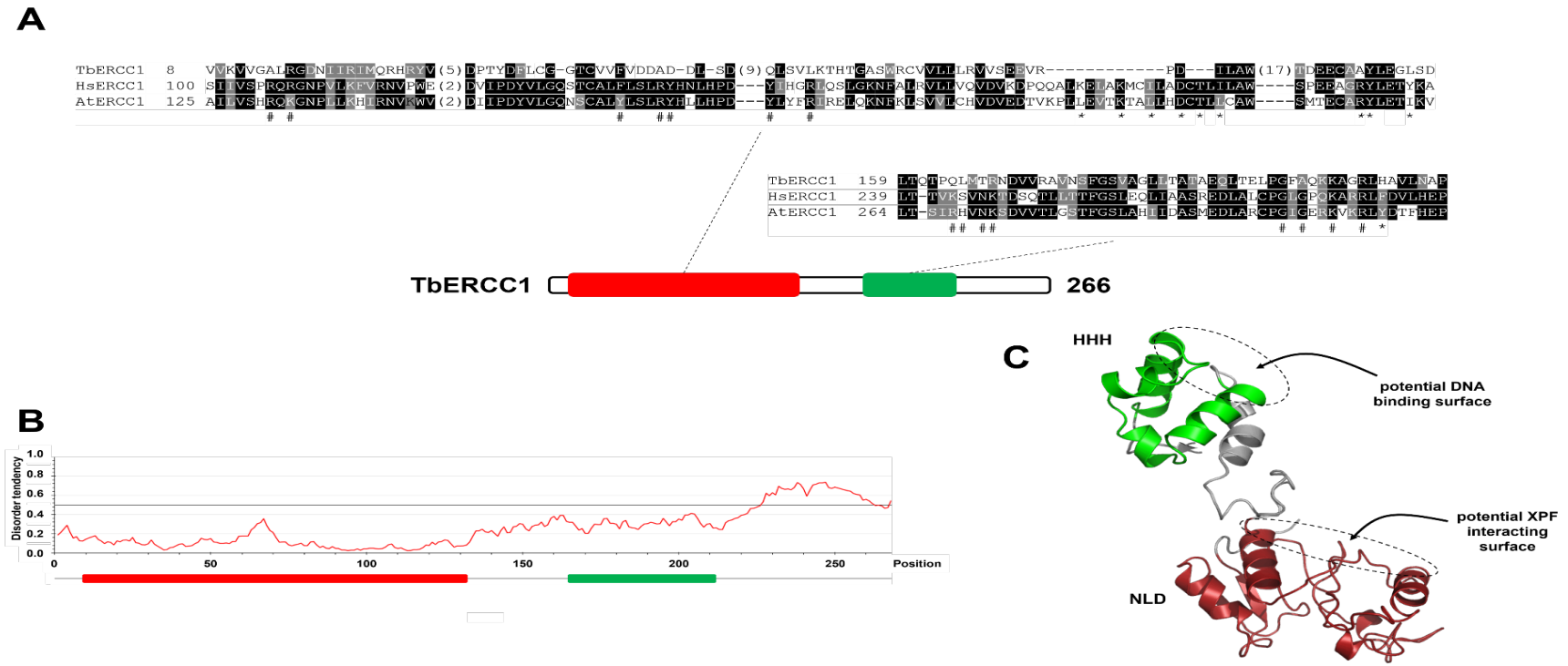


Fig 6.12. Informatic analysis of TbERCC1

A. The sequence corresponding to a nuclease-like domain (NLD; PF03834; residues 8-132; red) & Helix-hairpin-helix motif (HHH; PF14520; residues 166-212; green) domains of TbERCC1 (XP_845829) were aligned with the equivalent regions from *H. sapiens* HsERCC1 (NP_001974) & A. *thaliana* AtERCC1 (Q9MA98). Residues that are identical or are conserved substitutions are highlighted in black & grey, respectively. In relation to the human orthologue, residues highlighted with # or * may mediate interactions with DNA or XPF, respectively (Tsodikov et al 2005; Jones et al 2020).

B. Disordered regions were obtained by submitting the TbERCC1 sequence to disorder prediction program IUPRED2 (https://iupred2a.elte.hu/). Residues 223 to 266 have a disorder value >0.5 & indicates this is an unstructured region of the protein.

C. 3D prediction of a truncated version of TbERCC1 (XP_845829) was constructed. The approximation lacks 43 residues that correspond to the disorder region present at the carboxyl terminal & was generated from three templates (6SXB, 1V9P, 4BXO), with 2266 residues (85%) modelled at >90% confidence. The position of the NLD (red) and HHH (green) domains are shown. Non-domain regions are noted in grey. Comparison of HsERCC1 and TbERCC1 structures identified potential XPF interacting and DNA binding surfaces in the trypanosomal enzyme (Tsodikov et al 2005; Jones et al 2020).

6.2.6 Informatic analysis of TbXPF

Literature searches (Machado et al 2014; Genois et al 201473) followed by BLAST analysis of *T. brucei* TREU927 genome data held on TriTrypDB identified a single 3729 bp long ORF (Tb927.5.3670) as coding for an XPF orthologue. This gene, designated as Tb*xpf*, is found on chromosome 5 of the parasite's nuclear genome and has potential to encode for a 136 kDa enzyme, designated as TbXPF, composed of 1242 amino acids. Sequence comparisons revealed that Tb*xpf* is 54% identical to Tc*xpf* from *T.cruzi* CL Brener Esmeraldo-like haplotype (TcCLB.509779.10), 49 to 51% identical to the equivalent genes from *C. fasciculata* (CFAC1_200007300), *L. donovani* (LdBPK.08.2.000150) and *L. major* (LmjF.08.0140), and 43 to 45 % identical to its *S. cerevisiae rad1* (M15435), *H. sapiens ercc1* (XM_011522424) and *A. thaliana xpf* (NM_124608) counterparts. Primary sequence analysis and structural comparison against human ERCC1 (Jones et al 2020) revealed TbXPF can be divided into two 'modules' with these made up of several distinct sections (Figure 6.13). Starting at the amino terminal and forming ~75% of the protein, the so-called helicase-like module (HLM) could be discerned. This region contained two divergent RecA-like domains, designated as RECA1 and RECA2, that flank an all α-helical domain. Within the remaining sequence, a region known as the catalytic module (CM) was detected. This is composed of a nuclease domain (NUC; PF02732; residues 1012-1144) and includes a characteristic GDX_nERKX_3D motif, followed by a helix-hairpin-helix motif (HHH; residues 1177-1229). By comparison against the human enzyme, TbERCCI NUC contains several conserved residues that are important/essential for catalytic activity (D1014; R1016; E1017; D1042; E1052; R1053; K1054; D1058) with some (D1014; E1017; D1042; E1052) implicated in metal ion (Mg^{2+} or Mn^{2+}) binding.

3D predictions on full-length TbXPF based on the *H. sapiens* XPF component of the XPF-ERCC1-DNA structure (6SXB; Jones et al 2020) revealed that the five protein domains each form structurally distinct globular regions within the protein (Figure 6.14A). In terms of organization, the two divergent RecA-like domains contain 2 (RECA1) or 5 (RECA2) β-sheets arranged in parallel surrounded by α-helices, the NUC consists of 3 parallel arranged β-sheets flanked by 5 α-helices while HHH is made up of by 5 α-helices. Although HLM in the human enzyme is required for full XPF activity (Bowles et al., 2012; Klein Douwel et al., 2017), its precise role is unclear especially given it lacks residues necessary to bind and hydrolyse ATP (Gaillard et al 2000; Sgouros et al 1999). In contrast, the NUC and HHH domains which are structurally related to NLD

and HHH from TbERCC1, are important in heterodimerization with ERCC1 through predominantly hydrophobic interactions, in DNA binding and in nuclease activity via its metal co-factor (Figure 6.14B) (Jones et al 2020; McNeil and Melton, 2012).

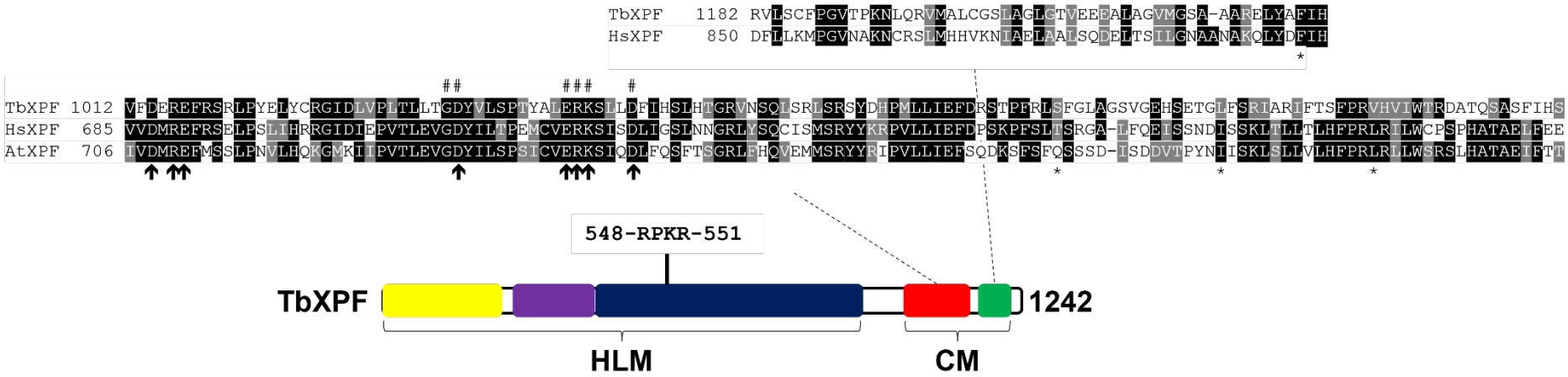


Fig 6.13. Informatic analysis of TbXPF

The sequences that correspond to the nuclease domain (NUC; PF02732; residues 1012-1144; red box) or Helix-hairpin-helix motif (HHH; PF14520; residues 1171-1229; green box) of TbXPF (XP_845037) were aligned with the equivalent regions from *H. sapiens* HsXPF (NP_005227) and/or *A. thaliana,* AtXPF (ABQ52659). Residues that are identical or are conserved substitutions are highlighted in black and grey, respectively. This region forms the "catalytic module" (CM) of this enzyme. Alignment of the human XPF structure (Jones et al 2020) with the TbXPF 3D prediction model identified two divergent RecA-like domains, designated RECA1 (yellow box) and RECA2 (blue box) that flank an all α-helical domain (purple). This region forms the helicase-like module (HLM) of this enzyme. In comparison to the human orthologue, * represents residues that may mediate interaction with ERCC1 (Jones et al 2020) while # or corresponds to amino acids that form the characteristic GDX_nERKX_3D motif or are important in catalytic activity, respectively (McNeil and Melton, 2012). The position of a 4-pattern putative nuclear targeting signal is shown.

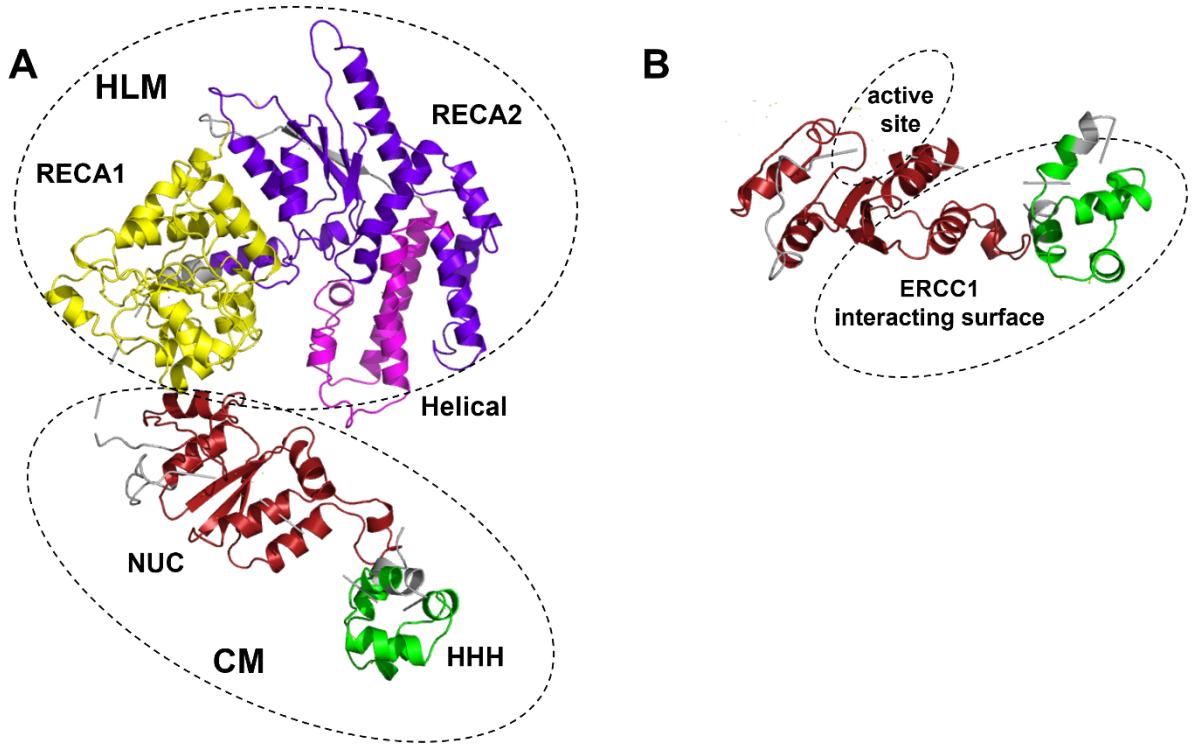


Fig 6.14. 3D prediction model of TbXPF.

A. A 3D prediction of TbXPF (XP_845037) was constructed using Phyre2 (Kelley et al. 2015) and visualized using PyMOL. The approximation was generated based on the *H. sapiens* XPF component of the XPF-ERCC1-DNA structure (6SXB; Jones et al 2020). There is 100% confidence over 25% of residues modelled. The RECA1 (yellow), helical (purple) and RECA2 (blue) regions that make up the HLM and the NUC domain (red) and HHH motif (green) that comprise the CM are shown. The former three regions were identified following alignment of the TbXPF 3D prediction with the *H. sapiens* XPF structure.
B. A zoomed image showing the CM of TbXPF. The NUC domain and HHH motif are given in red and green, respectively. The position of the active site (based on the location of the GDX_nERKX_3D active site motif in NUC) and a potential ERCC1 interaction surface are highlighted.

To assess how TbXPF, TbERCCI and DNA may interact, the truncated TbERCC1 and full-length TbXPF models were each aligned with the human XPF-ERCC1-DNA structure (Jones et al 2020). Next, the human proteins were removed from the alignment to leave the two trypanosomal approximations and nucleic acid in place (Figure 6.15). This revealed that both ERCC1 domains could potentially heterodimerise with their equivalent domains in TbXPF's CM, forming discrete NUC–NLD and 2×HHH functional units. Additionally, the 2×HHH functional unit is spatially positioned such that this region could mediate DNA binding.

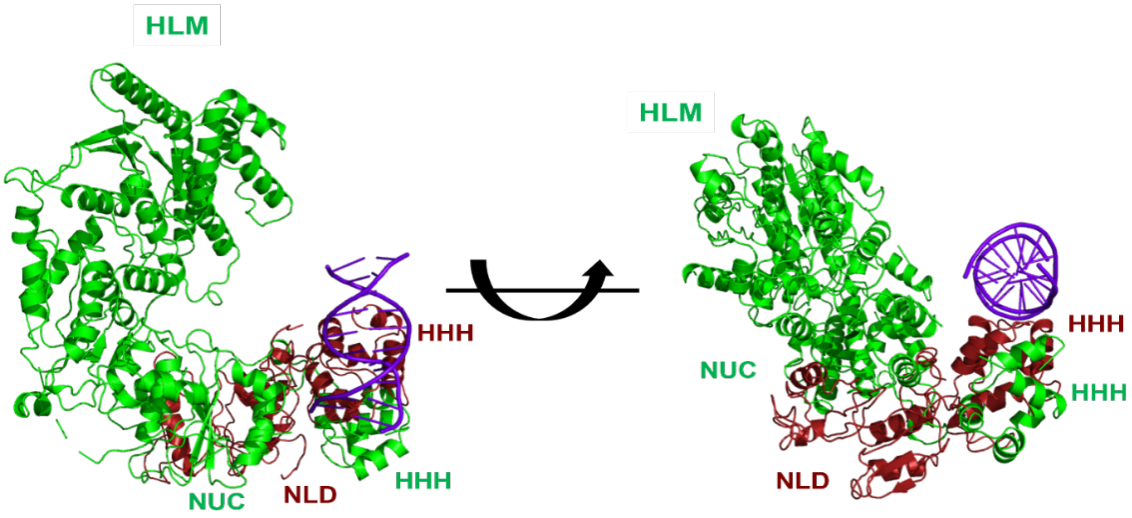


Fig 6.15. 3D prediction model showing TbXPF-TbERCC1-DNA interaction.
To illustrate how TbXPF, TbERCC1 and DNA interact, the truncated TbERCC1 (red) and full-length TbXPF (green) models were each aligned with the human XPF-ERCC1-DNA structure (6SXB; Jones et al 2020) using PyMOL. The human XPF-ERCC1 structures were removed from the alignment but DNA (purple) was retained. The HLM, nuclease domain (NUC) and helix-hairpin-helix motif (HHH) of TbXPF are labelled in green while the nuclease-like domain (NLD) and helix-hairpin-helix motif (HHH) of TbERCC1 are labelled in red.

6.2.7 Informatic analysis of TbXPG

Literature searches (Machado et al., 2014; Genois et al., 2014) followed by BLAST analysis of *T.b. brucei* TREU927 data held at NCBI identified a 2241bp long ORF (XM_822339) as coding for a XPG orthologue. This gene, designated as Tb*xpg*, is found on chromosome 9 of the parasite's nuclear genome and has potential to encode for an 82.4 kDa TbXPG enzyme, composed of 746 amino acids. Sequence comparisons revealed that Tb*xpg* is 57% identical to Tc*xpg* from *T. cruzi* CL Brener Esmeraldo-like haplotype (TcCLB.507009.120), 46 to 49% identical to the equivalent gene from *C.fasciculata* (CFAC1_300086200), *L.donovani* (LdBPK.35.2.003640) and *L. major* (LmjF.35.3590), and 39 to 45.0 % identical to *S.cerevisiae rad2* (NM_001181387), *H. sapiens xpg*/*ercc5* (NM_000123) and *A.thaliana xpg* (NM_100069). Analysis of TbXPG revealed it possessed an XPG N-terminal domain (XPG_N; PF00752) at its amino terminal (residues 1-96) and a XPG I-region/containing domain (XPG_I; PF00867 & PF12813) towards its carboxyl terminal (residues 461-701) (Figure 6.16). Several conserved sites were identified in these domains including a set of carboxylate amino acids (e.g., D30, D75, E485, D487, D507, D509, D557) that may mediate divalent metal ion (Mg^{2+} or Mn^{2+}) coordination, and three sequences (β-hairpin motif, helix-2turn-helix domain motif & hydrophobic wedge) that may facilitate in DNA binding (González-Corrochano et al., 2020).

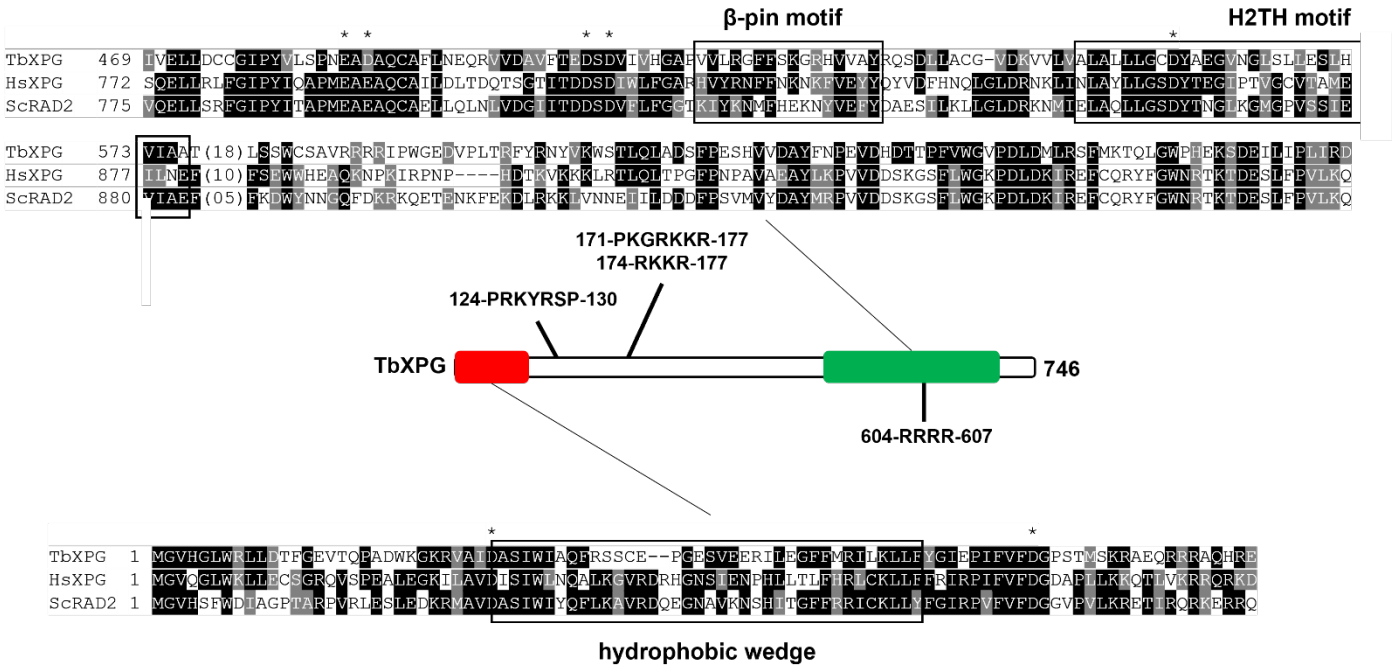


Fig 6.16. Informatic analysis of TbXPG

The sequences that correspond to the XPG N-terminal domain (XPG_N; PF00752; residues 1-96; red box) or XPG I-region/containing domain (XPG_I; PF00867 & PF12813; residues 461-701; green box) of TbXPG (XP_827432) were aligned with the equivalent regions from *H. sapiens* (HsXPF; NP_000114) and S. cerevisiae (ScRAD2; NP_011774). Residues that are identical or are conserved substitutions are highlighted in black and grey, respectively. Asterisk (*) highlights residues possibly involved in metal ion coordination while the β-pin motif, H2TH motif & hydrophobic wedge (boxed) are all regions in the human protein that contribute to DNA binding (González-Corrochano et al 2020). The position of two 4 pattern and two 7 pattern putative nuclear targeting signals are shown.

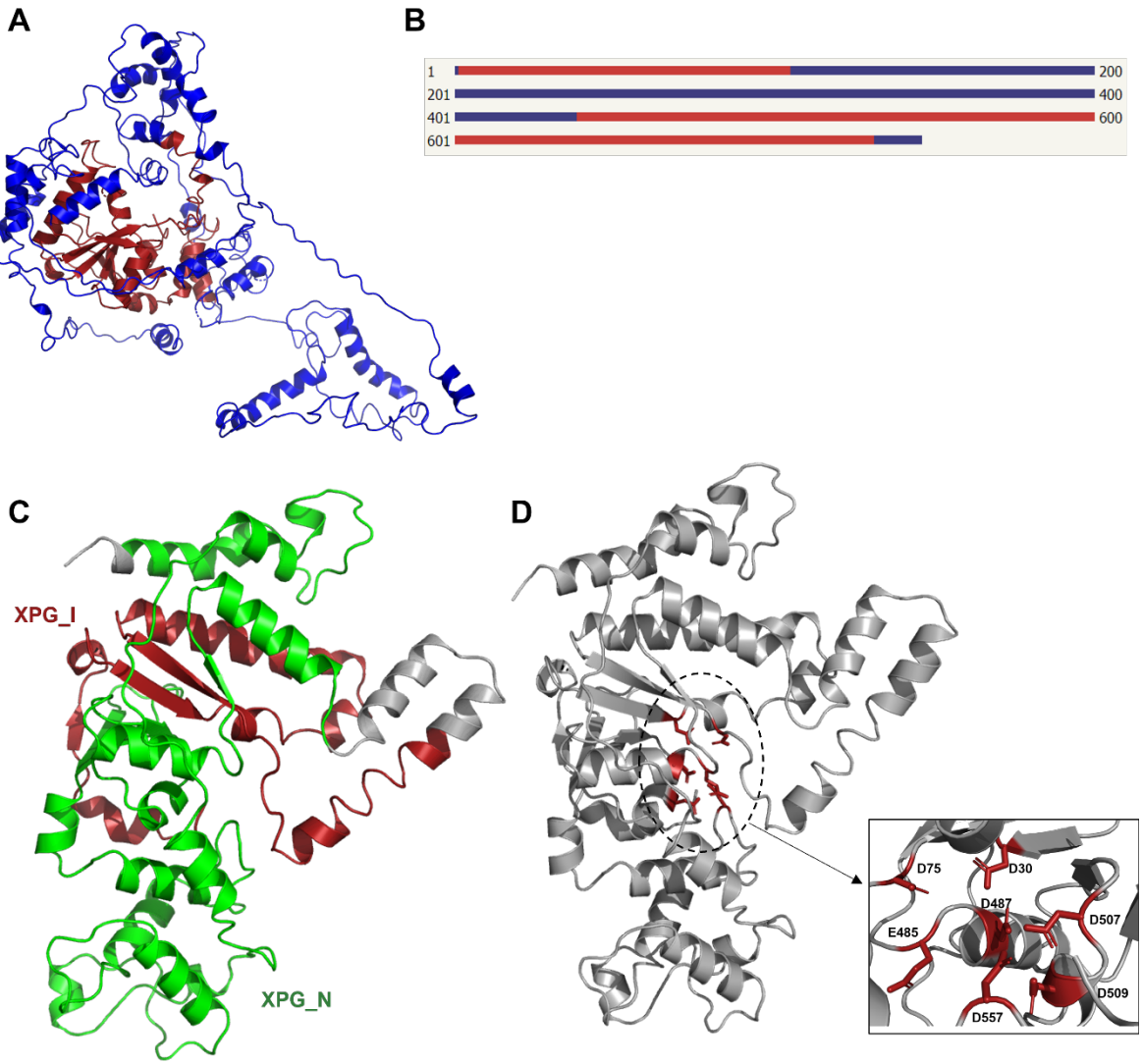


Fig 6.17. 3D prediction model of TbXPG.
A. 3D predictions of full-length TbXPG (XP_827432) based on six templates (5CNG; 4G0R; 1UL1; 6VBH; 3Q8L; 1B43) wase constructed using Phyre2 (Kelley et al. 2015) and visualized using PyMOL. Only 53% of residues were modelled at >90% confidence (red). Residues 97-476 and 578-746 could not be reliability modelled as these regions lack homology to known structures (blue).
B. Schematic showing the degree of confidence in the TbXPG full-length model (**A**). Areas in red and blue correspond to regions where there is high or low confidence in the model, respectively. The areas
C. A 3D prediction of a TbXPG monomer (XP_827432) was constructed using Phyre2 (Kelley et al. 2015) and visualized using PyMOL. The approximation was generated based on the catalytic domain of *H. sapiens* XPG (6VBH; Tsutakawa et al 2020). There is 100% confidence over 38% of residues modelled. The XPG_N (green) and XPG_I (green) domains are shown. Non-domain regions are noted in grey.
D. Putative catalytic site of TbXPG is in a groove lying roughly at the centre of the approximation. The insert shows a zoomed image of this region and highlights the proximity of the metal binding residues, D30, D75, E485, D487, D507, D509 & D557.

Initial attempts at generating 3D predictions using full-length protein produced an unreliable model across about 50% of the structure (Figure 6.17A and B). To overcome this, the 'low

confidence' regions (residues 97-476 and 578-746) were excluded from the input sequence and an approximation of a truncated TbXPG monomer made (Figure 6.17C). The resultant model predicted that the trypanosomal protein folded into a centrally located seven stranded β-sheet surrounded by 21 α-helices. When mapped onto the structure, the two domains combine to form one globular region resulting in an overall bean-like shape structure, reminiscent of other enzymes of the flap endonuclease family (González-Corrochano et al., 2020; Orans et al., 2011; Tsutakawa et al., 2011), while the residues that potentially form the active site are found in centrally located groove (Figure 6.17C & D).

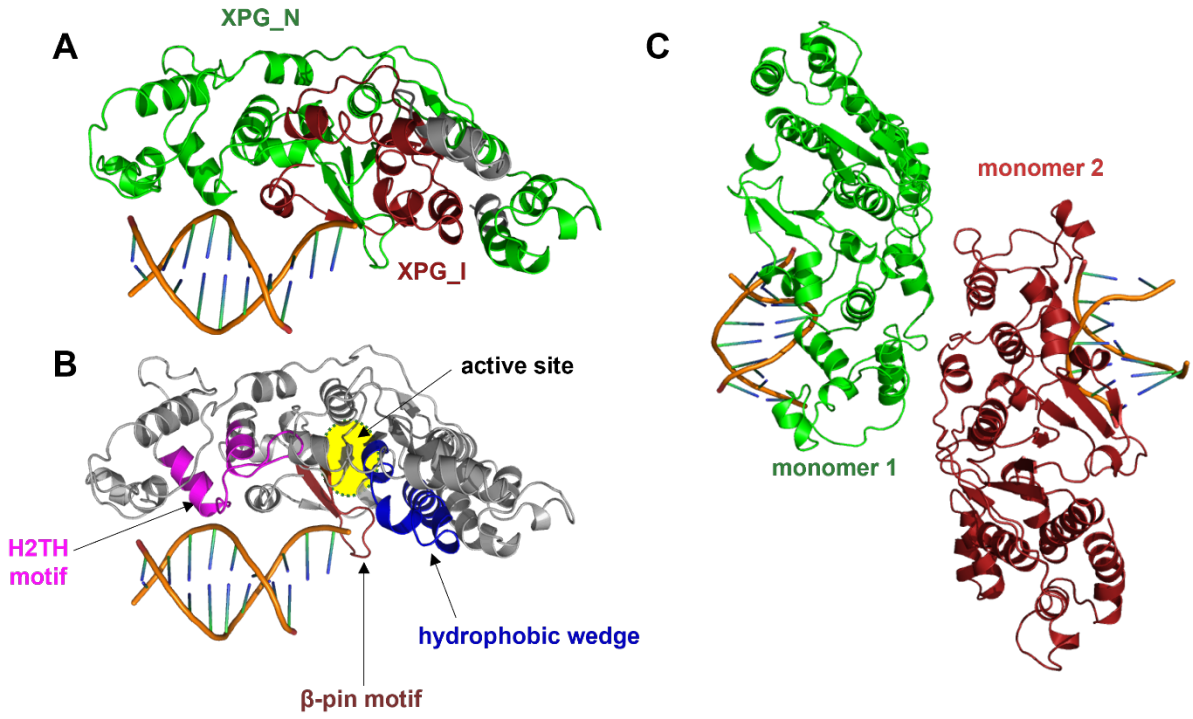


Fig 6.18. Interaction of TbXPG with DNA.

A & B. The TbXPG monomer model was aligned with the human XPG bound to DNA (6TUW; González-Corrochano et al 2021). The human structure was removed from the alignment, but DNA retained. In **A**, the XPG_N and XPG_I) domains are highlighted in green or red, respectively. In **B**, based on human enzyme, the position of the putative active site and regions that may contribute to DNA binding (β-pin motif, H2TH motif & hydrophobic wedge) in the trypanosomal enzyme shown

C. Two models of TbXPG monomers (green or red) were aligned with dimeric human XPG bound to DNA (6TUX; González-Corrochano et al 2021). The human structures were removed from the alignment, but DNAs retained.

To assess how an TbXPG monomer may interact with DNA, the truncated TbXPG model was aligned with the human XPG bound to DNA structure (6TUW; González-Corrochano et al 2021), with the human protein then removed from the alignment to leave the trypanosomal approximation and nucleic acid in place (Figure 6.18A and B). This revealed that both domains in TbXPG maybe

suitably located to interact with DNA through the β-pin and helix-2turn-helix domain motifs as well as the hydrophobic wedge such that they could position the nucleic acid substrate near to the active site. Further alignment modelling based on other arrangements of the human protein (6TUX; González-Corrochano et al 2021) generated an approximation indicating how TbXPG may dimerise (Figure 6.18C). By analogue to the human enzyme, the α-helical forming sequence 590-EQVRDMLSW-598 from one TbXPG monomer is predicted to be located to interact the α-helical forming sequence 638-SHVVDAYF-645 in the second TbXPG monomer, with the reciprocal interactions also occurring.

6.3 Editing the genes coding for NER factors.

Using the CRISPR/Cas9 genome editing procedure reported by (Beneke et al 2017) a library of BSF and/or PCF *T. brucei* lines were made where the endogenous genes coding for NER factors (Tb*csb*, *Tbddb1*, *Tbercc1* Tb*rad23*, *Tbxpc, Tbxpf* or *Tbxpg*) were tagged at their 5' or 3' region with nucleotide sequences coding for the fluorescent protein mNeonGreen (see Figure 3.13). This was achieved by using appropriate primer combinations (Table 6.2) to generate the sgRNA and targeting fragment amplicons that were subsequently electroporated into the parasite line with the suitable drug pressure applied to select for the desired integration. The cell lines generated by this approach are listed in Table 6.3. Despite multiple transformations (3 or more), some lines could not be made.

targeted gene	primer combinations	
	sgRNA	targeting fragment
Tb*csb*		
5' tagging	G00/5'sgRNA-Tbcsb	Tbcsb-UFP1/Tbcsb-URP1
3' tagging	G00/3'sgRNA-Tbcsb	Tbcsb-DFP1/Tbcsb-DRTAGP
Tb*ddb1*		
3' tagging	G00/3'sgRNA-Tbddb1	Tbddb1-DFP1/TbTbddb1-DRTAGP
Tb*ercc1*		
5' tagging	G00/5'sgRNA-Tbercc1	Tbercc1-UFP1/Tbercc1-DRF1
3' tagging	G00/3'sgRNA-Tbercc1	Tbercc1-DFP1/Tbercc1-DRTAGP
Tb*rad23*		
3' tagging	G00/3'sgRNA-Tbrad23	Tbrrad23-DFP1/Tbrad23-DRTAGP1
Tb*xpc*		
3' tagging	G00/3'sgRNA-Tbxpc	Tbxpc-DFP1/Tbxpc-DRTAGP1
Tb*xpf*		
5' tagging	G00/5'sgRNA-Tbxpf	Tbxpf-UFP1/Tbxpf-URP1
3' tagging	G00/3'sgRNA-Txpf	Tbxpf-DFP1/Tbxpf-DRTAGP
Tb*xpg*		
5' tagging	G00/5'sgRNA-Tbxpg	Tbxpg-UFP1/Tbxpg-URP1
3' tagging	G00/3'sgRNA-Tbxpg	Tbxpg-DFP1/Tbxpg-DRTAGP1

Table 6.2. Primer combinations used for tagging *T. brucei* NER factors.
The primer combinations used to generate amplicons corresponding to the short guide RNA (sgRNA) and targeting fragment used in tagging the 5' or 3' ends of the endogenous Tb*csb*, *Tbddb1*, *Tbercc1* Tb*rad23*, *Tbxpc, Tbxpf* or *Tbxpg* loci with a sequences coding for the fluorescent protein mNeonGreen. The sequence of each primer can be found in Appendix 1.

gene	bloodstream form		procyclic form	
	amino	carboxyl	amino	carboxyl
Tb*csb*	✓	✓	✓	✓
Tb*ddb1*	X	✓	X	✓
Tb*rad23*	X	✓	X	✓
Tb*xpc*	X	✓	X	✓
Tb*ercc1*	✓	✓	✓	✓
Tb*xpf*	✓	✓	✓	✓
Tb*xpg*	X	✓	X	✓

Table 6.3. *T. brucei* lines expressing tagged nucleotide excision repair factors.
Attempts to tag the 5' or 3' region of genes coding for NER factors in bloodstream or procyclic form *T. brucei* were performed although not always successfully. ✓ corresponds to recombinant cells generated by the CRISPR/Cas9 genome editing procedure while X represents situations where the desired line was not obtained.

To confirm that the desired fusion gene was being expressed, DNA amplification reactions using cDNA templates were performed using a primer combination that spans the gene fusion event. As such an amplicon could only be generated if the introduced targeting fragments had integrated at the correct loci resulting in the desired recombinant gene. In all cases where recombinant parasites were analysed, the expectedly sized amplicons were observed. For example, when using primer combinations intended to detect the *mNeonGreen*-Tb*csb* fusion, PCR generated a 1200 bp band, the expected size, from nucleic acid (cDNA) originating from recombinant BSF and PCF parasites with no band(s) observed in templates from wild type cells (Figure 6.19B). Likewise, a similar outcome was noted when analysing Tb*csb-mNeonGreen* expressing BSF or PCF *T. brucei* although in this case, a band of ~400 bp was detected in templates from the recombinant cells (Figure 6.19D).

When analysing cells expressing a tagged gene coding for a GG-NER factor (Figure 6.20), PCR using primer combinations designed to amplify across the fusion event only generated a band of the expected size from the genetically modified BSF and PCF lines with no amplicon obtained when using wild type-derived nucleic acids as template: For parasites expressing Tb*ddb1-mNG,* Tb*rad23-mNG* or Tb*xpc-mNG* the resultant amplicons were of 1000, 580 or 590 bp, respectively.

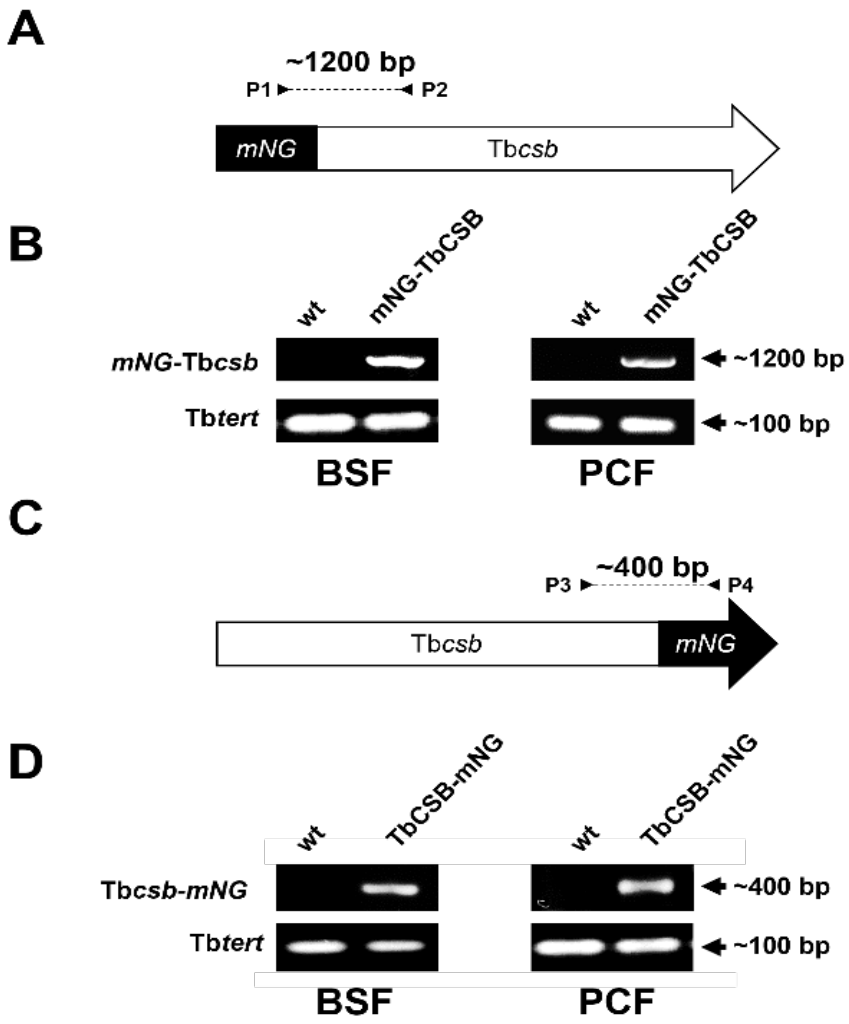


Fig 6.19. Validation of *T. brucei* lines expressing mNeonGreen tagged TbCSB.

A & C. Schematic showing integration of *mNeonGreen* (*mNG*) into the *T. brucei* genome at the 5' (**A**) or 3' (**C**) region of Tb*csb*. P1 to P4 indicate the primers used in the obtention of the validation amplicons.

B & D. Amplicons (in kb) diagnostic of bloodstream form (BSF) or procyclic form (PCF) *T. brucei* expressing *mNG*-Tb*csb* **(B)** or Tb*csb*-*mNG* **(D)** were generated from cDNA templates derived from total RNA extracted from wild type (wt) or recombinant parasites. The integrity of DNA samples was evaluated by amplification of a 0.1 kb control fragment, Tb*tert*. The amplified fragment Is indicated (*mNG*-Tb*csb*; Tb*csb*-*mNG* or Tb*tert*). The primer sequences and combinations used for each amplification are listed in Appendix 2/

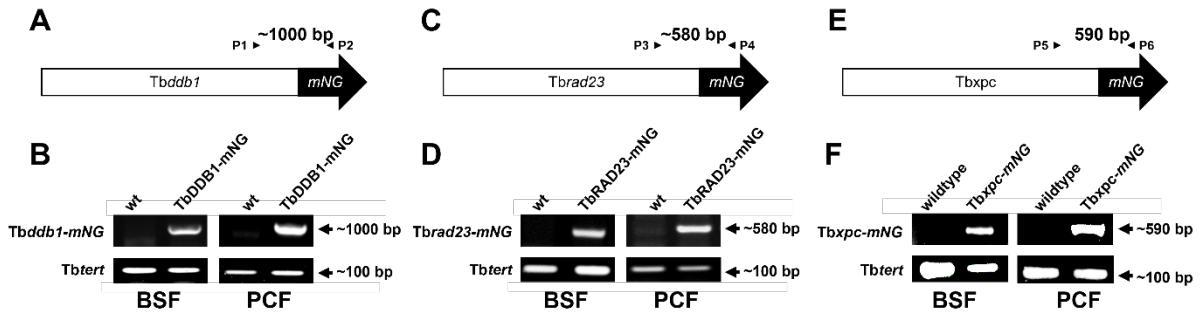


Fig 6.20. Validation of *T. brucei* lines expressing mNeonGreen tagged GG-NER factors.
A, C & E. Schematic showing integration of *mNeonGreen* (*mNG*) into the *T. brucei* genome at the 3' region of Tb*ddb1* (**A**), Tb*rad23* (**C**) or Tb*xpc* (**E**). P1 to P6 indicate the position of the primers used in the obtention of the validation amplicons.
B, D & F. Amplicons diagnostic of bloodstream form (BSF) or procyclic form (PCF) *T. brucei* expressing Tb*ddb1-mNG* (**B**), Tb*rad23-mNG* (**D**) or Tb*xpc-mNG* (**F**) were generated from cDNA templates derived from total RNA extracted from wild type (wt) or recombinant parasites. The integrity of DNA samples was evaluated by amplification of a 0.1 kb control fragment, Tb*tert*. The amplified fragment Is indicated (Tb*ddb1-mNG*, Tb*rad23-mNG,* Tb*xpc-mNG* or Tb*tert*). The primer sequences and combinations used for each amplification are listed in Appendix 2, respectively.

When the above validation process was extended to parasites expressing mNeonGreen tagged genes coding for NER incision factors (TbERCC1, TbXPF or TbXPG), again an amplicon of the expected size was detected only from cDNAs derived from recombinant BSF or PCF parasites and not from wild type (Figure 6.21). PCR from templates stemming from *T. brucei* expressing *mNG*-Tb*ercc1*, *mNG*-Tb*xpf* or *mNG*-Tb*xpg* resulted in amplicons of 600, 630 or 630 bp, respectively, while bands of 630, 600 and 600 bp were obtained from nucleic acids originating from Tb*ercc1-mNG*, Tb*xpf-mNG* or Tb*xpg-mNG* expressing trypanosomes, respectively.

For each gene fusion, the confirmatory cDNA-derived amplicons were sequenced (data not shown) and the *Tbcsb, Tbddb1, Tbercc1 Tbrad23, Tbxpc, Tbxpf* or *Tbxpg* sequence shown to be in-frame with that coding for mNeonGreen. Also, for all recombinant and wild type samples tested, control PCRs were performed to amplify a 100 bp fragment from the Tb*tert* transcript. In all cases (Figures 6.19-6.21), a band of this size (100 bp) was observed confirming the integrity of all nucleic acid templates.

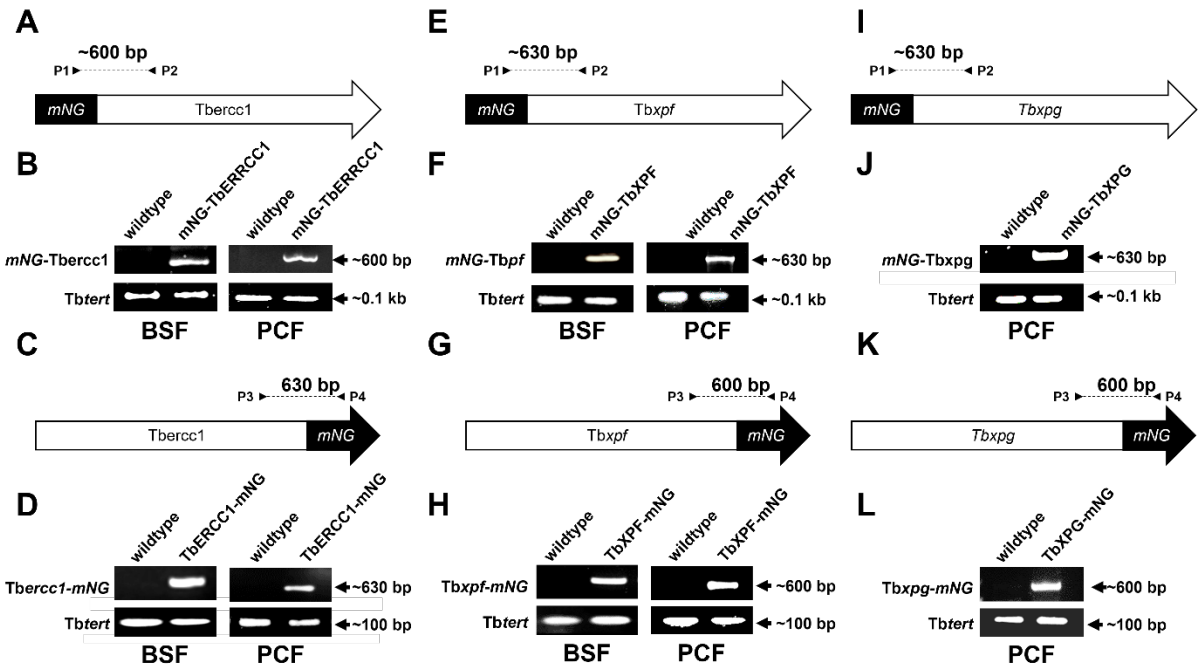


Fig 6.21. Validating T.brucei lines expressing mNeonGreen tagged NER incision factors.

A, C, E, H, I & K. Schematic showing integration of *mNeonGreen* (*mNG*) into the *T. brucei* genome at the 5' (**A**, **E** & **I**) or 3' (**C, G** & **K**) region of Tb*ercc1* (**A** & **C**), Tb*xpf* (**A** & **C**) or Tb*xpg* (**A** & **C**). P1-P12 indicate the primers used in the obtention of the validation amplicons.
B, D, F, H, J & K. Amplicons diagnostic of bloodstream form (BSF) or procyclic form (PCF) *T. brucei* expressing *mNG*-Tb*ercc1* (**B**), Tb*ercc1-mNG* (**D**), *mNG*-Tb*xpf* (**F**), Tb*xpf-mNG* (**H**), *mNG*-Tb*xpg* (**J**) or Tb*xpg-mNG* (**K**) were generated from cDNA templates derived from total RNA extracted from wild type or recombinant parasites. The integrity of DNA samples was evaluated by amplification of a 0.1 kb control fragment, Tb*tert*. The amplified fragment Is indicated (*mNG*-Tb*ercc1*; Tb*ercc1-mNG, mNG*-Tb*xpf*, Tb*xpf-mNG*, *mNG*-Tb*xpg*, Tb*xpg-mNG* or Tb*tert*). The primer sequences and combinations used for each amplification are listed in Appendix 2.

6.4 Localisation of NER factors in *T. brucei.*

Using the PSORT and PSORTII algorithms, potential subcellular localization and, where applicable, nuclear localisation signals (NLS) of the NER factors discussed in section 6.2 were assessed (Table 6.4). Based on at least one of these packages, TbCSB, TbRAD23, TbXPC and TbXPF were proposed to be found in the parasite's nucleus with organellar targeting of TbCSB, TbXPC and TbXPF possibly involving a predicted NLS. In the case of TbDDB1, TbERCC1 and TbXPG, predictions indicate they may be found at other sites in the parasite, although for TbDDB1 and TbXPG potential NLSs were detected from their amino acid sequences.

Protein	Localisation
TbCSB	PSORT - **nuclear** (0.96), microbody (peroxisome) (0.30). PSORTII – cytoplasmic (47.8%), **nuclear** (39.1%). NLS: 3 x '4 pat' & 1 x Robbins & Dingwall consensus sequence.
TbDDB1	PSORT - plasma membrane (0.60), Golgi body (0.4). PSORTII - endoplasmic reticulum (44.4%), plasma membrane (33.3%). NLS: 2 x '4 pat'.
TbRAD23	PSORT - cytoplasmic (0.65), lysosome (lumen) (0.16). PSORTII – **nuclear** (52.2%), mitochondrial (21.7%), cytoplasmic (17.4%). NLS: none detected.
TbXPC	PSORT - **nuclear** (0.91), ER (membrane) (0.85), plasma membrane (0.44). PSORTII – **nuclear** (43.5%); cytoplasmic (21.7%); mitochondrial (17.4%). NLS: 3 x 4 pattern and 2 x bipartite.
TbEERC1	PSORT – cytoplasmic (0.65), lysosome (lumen) (0.16). PSORTII – cytoplasmic (60.9%); mitochondrial (26.1%). NLS: none detected
TbXPF	PSORT – microbody (peroxisome) (0.30), **nucleus** (0.30). PSORTII – cytoplasmic (47.8%), nuclear (26.1%), mitochondrial (217.4%). NLS: 1 x '4 pat'
TbXPG	PSORT – microbody (peroxisome) (0.30), Golgi body (0.4), ER membrane (0.30). PSORTII – ER (47.4%), plasma membrane (22.2%). NLS: 2 x '4 pat' & 2 x '7 pat.

Table 6.4. Localisation of NER factors based on predictive algorithms.
Analysis using PSORT/PSORTII resulted in a prediction in given protein's subcellular localisation. For PSORT, the value in parentheses corresponds to the certainty value. The higher this value, the more likely the prediction. For PSORTII, the value in parentheses corresponds to the % that the prediction is correct. Again, the higher the value, the more likely the prediction. Cases where a nuclear localisation is predicted are highlighted in red. Sequences were analysed for four types of nuclear localisation signal (NLS) (4 pattern, 7 pattern, bipartite and Robbins & Dingwall consensus sequence) with the locations of these shown in Figures 6.2, 6.5, 6.13 and 6.15.

As part of the TrypTag project (Dean et al., 2017), the localization of several mNeonGreen tagged NER factors in PCF parasites has been confirmed experimentally (Table 6.5). From this study it was shown that fluorescence microscopic evaluation of insect form *T. brucei* expressing tagged TbCSB, TbDDB1, TbXPC, TbRAD23 or TbXPG yielded a nuclear-associated signal although this was often accompanied by foci elsewhere within the cell. For TbXPC and TbXPG, the nuclear fluorescence pattern was noted irrespective of which protein terminal was tagged. In contrast, nuclear localisation was noted for TbCSB only when tagged at their carboxyl terminal or for TbRAD23 when it was tagged at its amino terminal. This suggests that in these three cases, mNeonGreen was somehow interfering with nuclear targeting or potentially promoting the redirecting of the tagged protein to other sites.

Protein/Tritryps ID*	procyclic form	
	amino tag	carboxyl tag
TbCSB (Tb927.7.4080)	cytoplasm (reticulated)	**nucleus** endocytic
TbDDB1 (Tb927.6.5110)	cytoplasm (reticulated)	**nucleus (50%)** cytoplasm (reticulated)
TbRAD23 (Tb927.6.4650)	cytoplasm **nucleoplasm (points, strong)** flagellar cytoplasm nuclear lumen	cytoplasm flagellar cytoplasm
TbXPC (Tb927.9.11930)	**nucleoplasm** **nucleolus (peripheral)** cytoplasm (reticulated, weak)	**nuclear lumen** cytoplasm (weak, points)
TbERCC1(Tb927.7.2060)	X	X
TbXPF (Tb927.5.3670)	X	X
TbXPG (Tb927.9.11760)	**nucleus** cytoplasm (reticulated)	**nuclear** endocytic

Table 6.5. Localization of tagged NER factors in PCF *T. brucei*.
Data taken from the TrypTag project (TrypTag.org). In this study attempts have been made to tag every protein expressing gene in the PCF *T. brucei* genome with sequences coding for mNeonGreen (Dean et al 2017). *The TriTryp ID for each protein is given in parenthesis. X represents situations where the localization has not yet been determined. Cases where a nuclear localisation has been observed are highlighted in red.

Here, and using the mNeonGreen-expressing parasite lines reported in section 6.3, we evaluated the localisation of TbCSB1, TbDDB1, TbRAD23, TbXPC, TbERCC1, TbXPF or TbXPG in BSF or PCF *T. brucei*. Cells from exponentially growing cultures were fixed, their DNA stained with DAPI then imaged by fluorescent microscopy. In all cases, most (80 to 85%) trypanosomes from both life cycle forms exhibited a nucleus: kinetoplast genome ratio of 1N1K indicating they were in the G1/S phase of the cell cycle with 10 to 15% of parasites possessing ratio of 1N2K indicative of cells in G2/M phase while a 2N2K ration was noted in around 5% of *T. brucei* signifying they were in the post-M stage.

6.4.1 Localisation of the TC-NER factor, TbCSB.

Localisation of TbCSB depends on where the epitope is sited on the parasite protein (Figure 6.22). Most BSF or PCF *T. brucei* engineered to express mNG-TbCSB (i.e., TbCSB tagged at amino terminal with mNeonGreen) exhibited a reticular fluorescence signal throughout their cytoplasm with this, in some cases, extending along the flagellum cytoplasm, a pattern consistent with that previous noted (Figure 6.22A & C; Table 6.3). In relation to the cell cycle, this network-

like profile was observed in all PCF cells (G1/S, G2/M or post-M) and in G1/S BSF *T. brucei*. In contrast, discrete cytoplasmic foci (possibly from within the endocytic system) were noted in G2/M or post-M BSF trypanosomes. This apparent difference in localisation may reflect variations in mNG-TbCSB expression between the cell cycle stages.

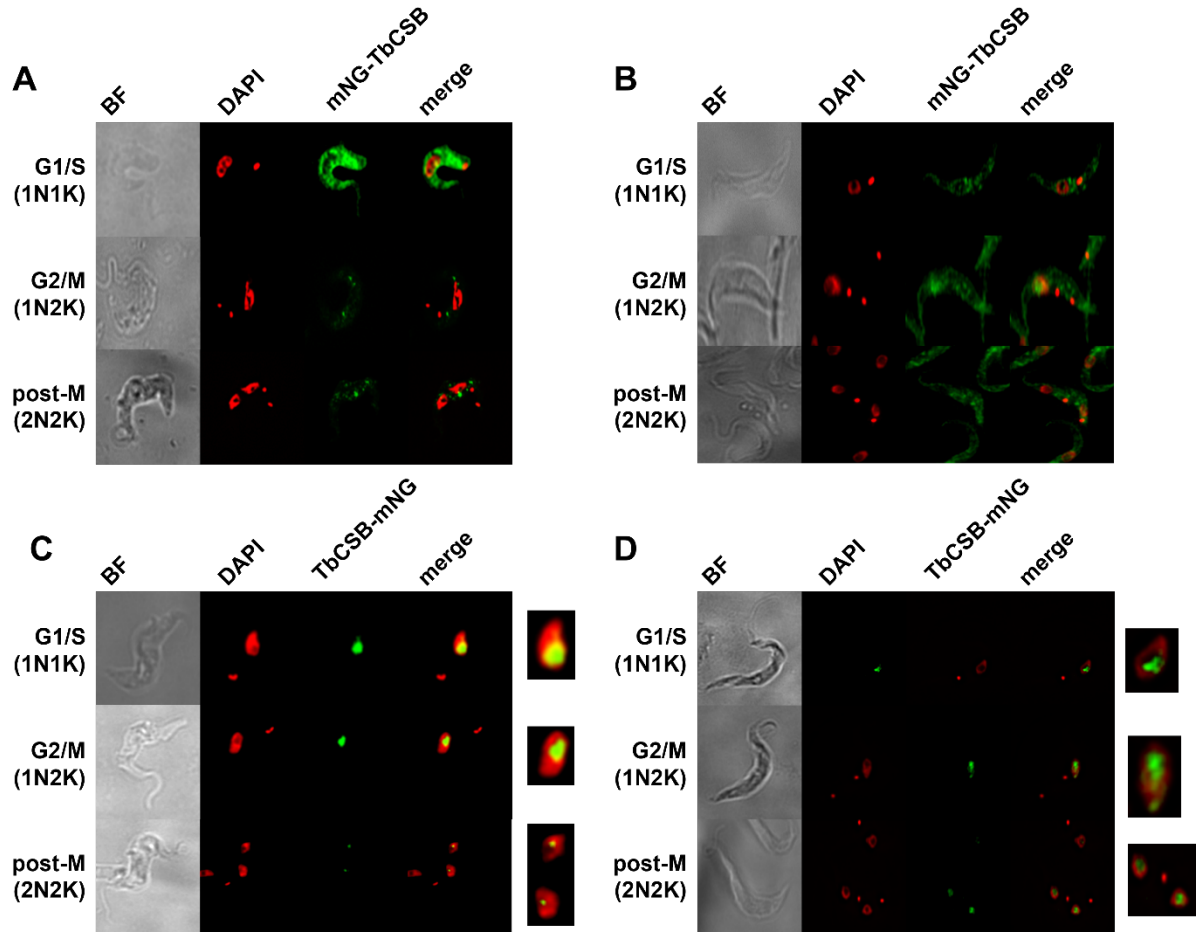


Fig 6.22. Localisation of TbCSB.
A-D. Bloodstream (**A & C**) and procyclic (**B & D**) form trypomastigotes *T. brucei* expressing TbCSB tagged with mNeonGreen (in green) at its amino (mNG-TbCSB) (**A & B**) or carboxyl (TbCSB-mNG) (**C & D**) termini were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow), including a close-up image of the nucleus, in trypanosomes at various stages of the cell cycle are shown.

When the localisation studies were extended to examine trypanosomes expressing TbCSB tagged at its carboxyl terminal with mNeonGreen (i.e., TbCSB-mNG), *T. brucei* from all cell cycle stages in both life cycle forms displayed a common pattern possessing several discrete fluorescent centres primarily within the non-DAPI staining nucleolus (Figure 6.22A & C): some cells also presented with the occasional fluorescent centre in the nucleoplasm or cytoplasm. As the nucleolus

is the major transcription site within the parasite nucleus, it is perhaps not surprising to find TbCSB, a TC-NER factor, here.

6.4.2 Localisation of GG-NER factors.

When evaluating the localisation of TbDDB1 or TbRAD23 in BSF or PCF *T. brucei* parasites engineered to express TbDDB-mNG or TbRAD23-mNG exhibited an intense reticular fluorescence signal throughout their cytoplasm, a pattern observed in cells from all stages of the cell cycle (Figure 6.23A-D). For TbRAD23, this localisation profile is consistent with that previous noted although for TbDDB1, there are differences (Table 6.5): based on TrypTag project data, TbDDB1 was shown to be within the nucleus and cytoplasm of PCF *T. brucei*. Why there is this discrepancy between our and the TrypTag project data is unclear.

When examining TbXPC-mNG expressing PCF *T. brucei*, discrete fluorescent centres were seen within the nucleus (Figure 6.23E). This punctate signal was throughout the nucleoplasm and was also present in the non-DAPI stained nucleolus This type of localization signal was present in cells in the G1/S and G2 stages of the cell cycle. For trypanosomes in the post-M phase of the cell cycle, a more coalesced nucleolar signal predominated.

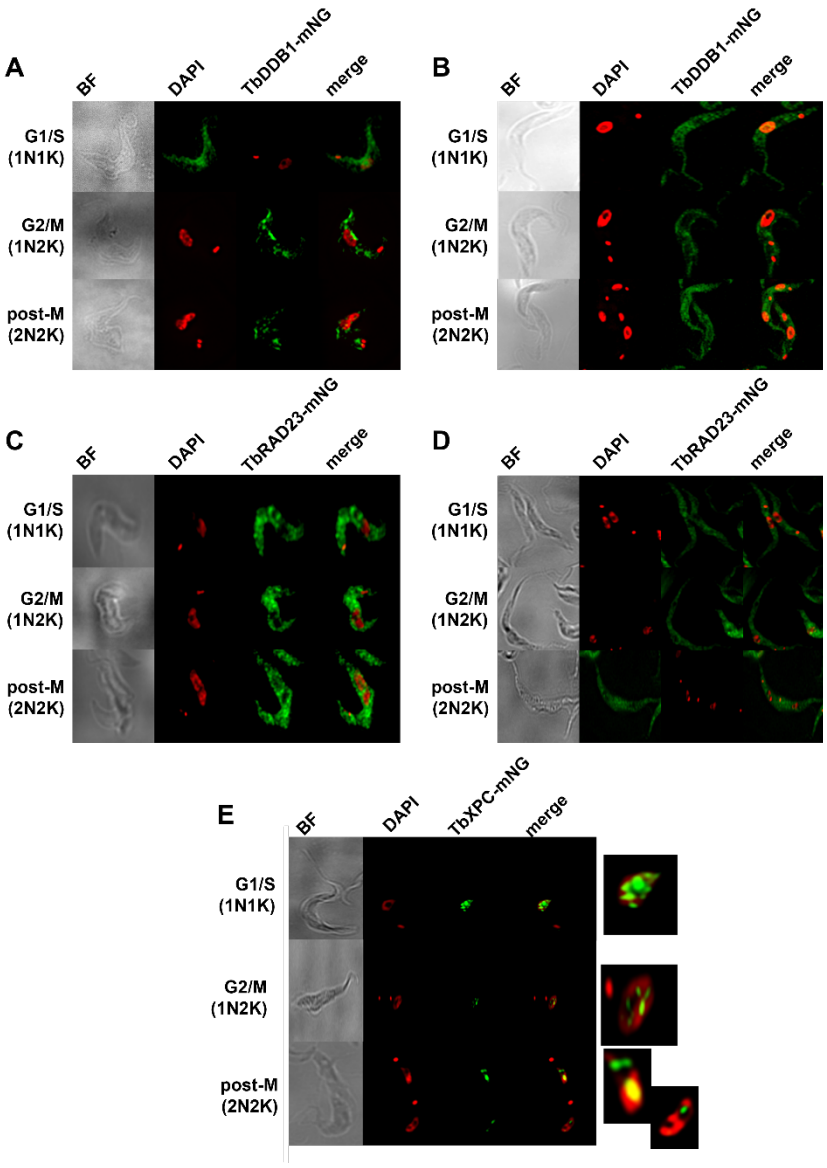


Fig 6.23. Localisation of GG-NER factors.
A-E. Bloodstream (**A & C**) or procyclic (**B, C & E**) form *T. brucei* expressing TbDDB1 (**A & B**), TbRAD23 (**C & D**) or TbXPC (**E**) tagged at their carboxyl terminal with mNeonGreen (TbDDB1-mNG; TbRAD23-mNG; TbXPC-mNG; in green) were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow), including for TbXPC a close-up image of the nucleus, in trypanosomes at various stages of the cell cycle are shown.

6.4.3 Localisation of NER incision factors.

When evaluating for TbEERC1 localisation using fixed/DAPI-stained TbERCCI-mNG or mNG-TbERCC1 expressing *T. brucei*, most cells displayed a fluorescent signal throughout their cytoplasm (Figure 6.24A-D). For BSF parasites from all cell cycle stages, and for G1/S phase PCF trypanosomes, this signal was relatively weak with the fluorescence pattern primarily presenting as discrete centres, suggesting the tagged protein was at a vesicular (possibly endocytic) location. In contrast, the fluorescence signal expressed by G2/M and post-M phase PCF *T. brucei* was intense and had a more reticular-like appearance. This difference could be due to higher expression

of the fusion protein in PCF *T. brucei* during the G2/M and post-M stages of the cell cycle relative to G1/S.

Extending the study to examine TbXPF (Figure 6.24E-H) and TbXPG (Figure 6.24I-J) localisations revealed both proteins had a nuclear localisation. In BSF and PCF G1/S stage *T. brucei* engineered to express mNG-TbXPF, discrete fluorescent centres were noted throughout the nucleoplasm while in the equivalent TbXPF-mNG parasites, the punctate pattern appeared more dispersed and present in the nucleolus. For mNG-TbXPF or TbXPF-mNG expressing BSF or PCF in the G2/M and post-M stages of the cell cycle, the punctate nucleoplasm/nucleolus pattern was also observed although in most cases, the coalesced nucleolar signal appeared to predominate. Similarly, PCF cells engineered to express mNG-TbXPG or TbXPF-mNG also presented with discrete fluorescent foci throughout the nucleoplasm/nucleolus pattern.

A summary of the localisation data obtained here is reported in Table 6.6. From this study, TbCSB, TbXPC, TbXPF and TbXPG were shown to be nuclear proteins generally forming discrete foci in the nucleoplasm and nucleolus.

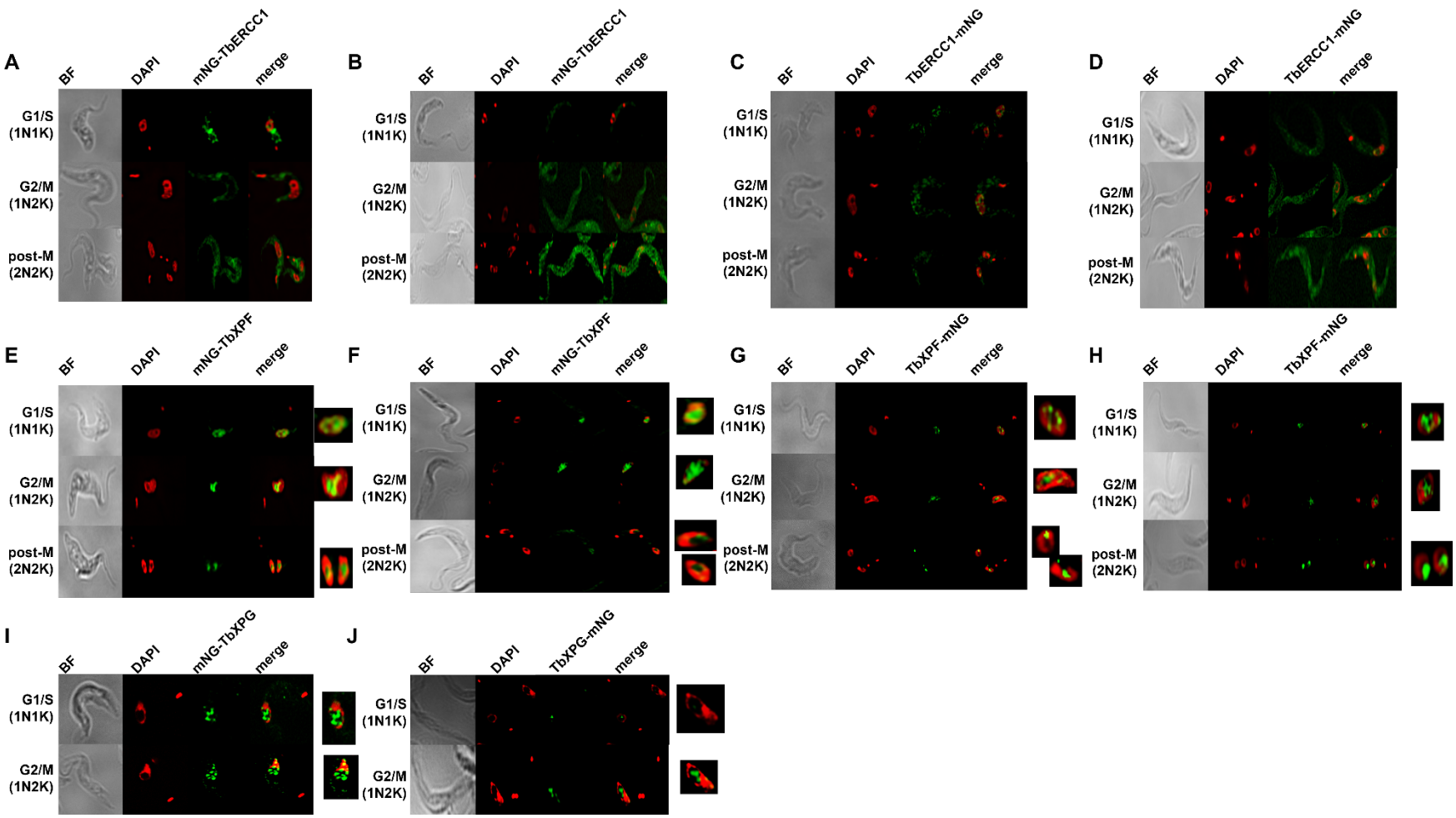


Fig 6.24. Localisation of incision factors of the *T. brucei* NER pathway.

A-J. Bloodstream (**A**, **C**, **E** & **G**) and procyclic (**B**, **D**, **F**, **H**, **I** & **J**) form *T. brucei* expressing TbERCC1 (**A-D**), TbXPF (**E-H**) or TbXPG (**I** & **J**) tagged with mNeonGreen (green) at their amino (mNG-TbERCC1 (**A & B**), mNG-TbXPF (**E** & **F**), mNG-TbXPG (**I**) or carboxyl (TbERCC1-mNG (**C & D**); TbXPF-mNG (**G & H**); TbXPG-mNG (**J**) were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow), including for TbXPF and TcXPG close-up images of the nucleus, in trypanosomes at various stages of the cell cycle are shown.

Protein	Bloodstream form		Procyclic form	
	mNG-TbXXX	TbXXX-mNG	mNG-TbXXX	TbXXX-mNG
TbCSB	cytoplasmic (reticulate) - G1/S vesicular - G2/M & post M	**nuclear (nucleolus; nucleoplasm) - G1/S,G2/M & post M**	cytoplasm (reticulated); flagellar cytoplasm - G1/S,G2/M & post M	**nuclear (nucleoplasm; nucleolus) - G1/S,G2/M & post M**
TbDDB	X	cytoplasm (reticulated) - G1/S,G2/M & post M	X	cytoplasm (reticulated) - G1/S,G2/M & post M
TbRAD23	X	cytoplasm (reticulated) - G1/S,G2/M & post M	X	cytoplasm (reticulated) - G1/S,G2/M & post M
TbXPC	X	X	X	**nuclear (nucleoplasm; nucleolus) - G1/S,G2/M & post M**
TbERCC1	X	cytoplasm (reticulated) - G1/S,G2/M & post M	X	cytoplasm (reticulated) - G1/S,G2/M & post M
TbXPF	**nuclear (nucleolus; nucleoplasm) - G1/S, G2/M & post M**	**nuclear (nucleolus; nucleoplasm) - G1/S, G2/M & post M**	**nuclear (nucleolus; nucleoplasm) - G1/S, G2/M & post M**	**nuclear (nucleolus; nucleoplasm) - G1/S, G2/M & post M**
TbXPG	X	X	**nuclear (nucleolus; nucleoplasm) - G1/S, G2/M & post M**	**nuclear (nucleolus; nucleoplasm) - G1/S, G2/M & post M**

Table 6.6. Summary of the localisation data.
Localisation data for each amino (mNG-TbXXX) or carboxyl (TbXXX-mNG) tagged NER protein in bloodstream or procyclic form *T. brucei* is presented. The parasites examined were at various stages (G1/S, G2/M, or post M) of the cell cycle. Cases where a nuclear localisation has been observed are highlighted in red. X indicates where protein localisation was not determined.

6.5 Effect of exogenous treatments on TbCSB, TbXPG or TbXPF expression.

The NER pathway is a versatile DNA repair mechanism that can respond to resolve a range of structurally unrelated bulky and/or helix distorting lesions (Schärer, 2013; Marteijn et al., 2014; Petruseva et al., 2014; Kusakabe et al., 2019). To assess the effect of exogenous treatments on expression of *T. brucei* NER components, PCF parasites engineered to express TbCSB-mNG, mNG-TbXPG or mNG-TbXPF were subjected to single bolus of DNA damaging (mechlorethamine, UV), or trypanocidal nitroheterocycle prodrug (nifurtimox; benznidazole; fexinidazole) agent and the fluorescence signal of individual cells within a population measured at a single time point post exposure, with the resultant data compared against untreated controls (Figure 6.25).

As previously noted, untreated trypanosomes expressing TbCSB-mNG presented with a punctate nuclear localization pattern with these foci seen in both the nucleolus and nucleoplasm: the fluorescence noted in the former structure were generally more intense than those observed in the latter. Treatment of the recombinant line with any of the above exogenous insults resulted in a significant increase in the total cellular fluorescence intensity displayed by most cells within a population (Figure 6.25A and B). Such trypanosomes still displayed the same overall nucleolus/nucleoplasm localisation pattern as observed in untreated controls although here the signal associated with the nucleolus appeared elevated.

When evaluating the effect of the above treatments on mNG-TbXPF expression, all resulted in an apparent subtle change in the overall localisation pattern (Figure 6.25C): in untreated controls, the fluorescence was primarily associated with nucleoplasm foci with exposure to any of the exogenous insults tested causing this signal to coalesce within the nucleolus. Intriguingly, the total fluorescence associated with cells exposed to nifurtimox, benznidazole or fexinidazole was equivalent to that noted in the untreated lines while in contrast, a significant change in fluorescence was noted in trypanosomes treated with the two DNA damaging agents mechlorethamine or UV (Figure 6.25D).

Extending the analysis to TbXPG, all treatments tested here led to an apparent but subtle change in the overall localisation pattern, akin to that observed for TbXPF (Figure 6.25E). Here, most of the fluorescent foci associated with the nucleoplasm in untreated cells appear to have now coalesced to the nucleolus posttreatment. This apparent change was not accompanied by increased fluorescence change within the population (Figure 6.25F).

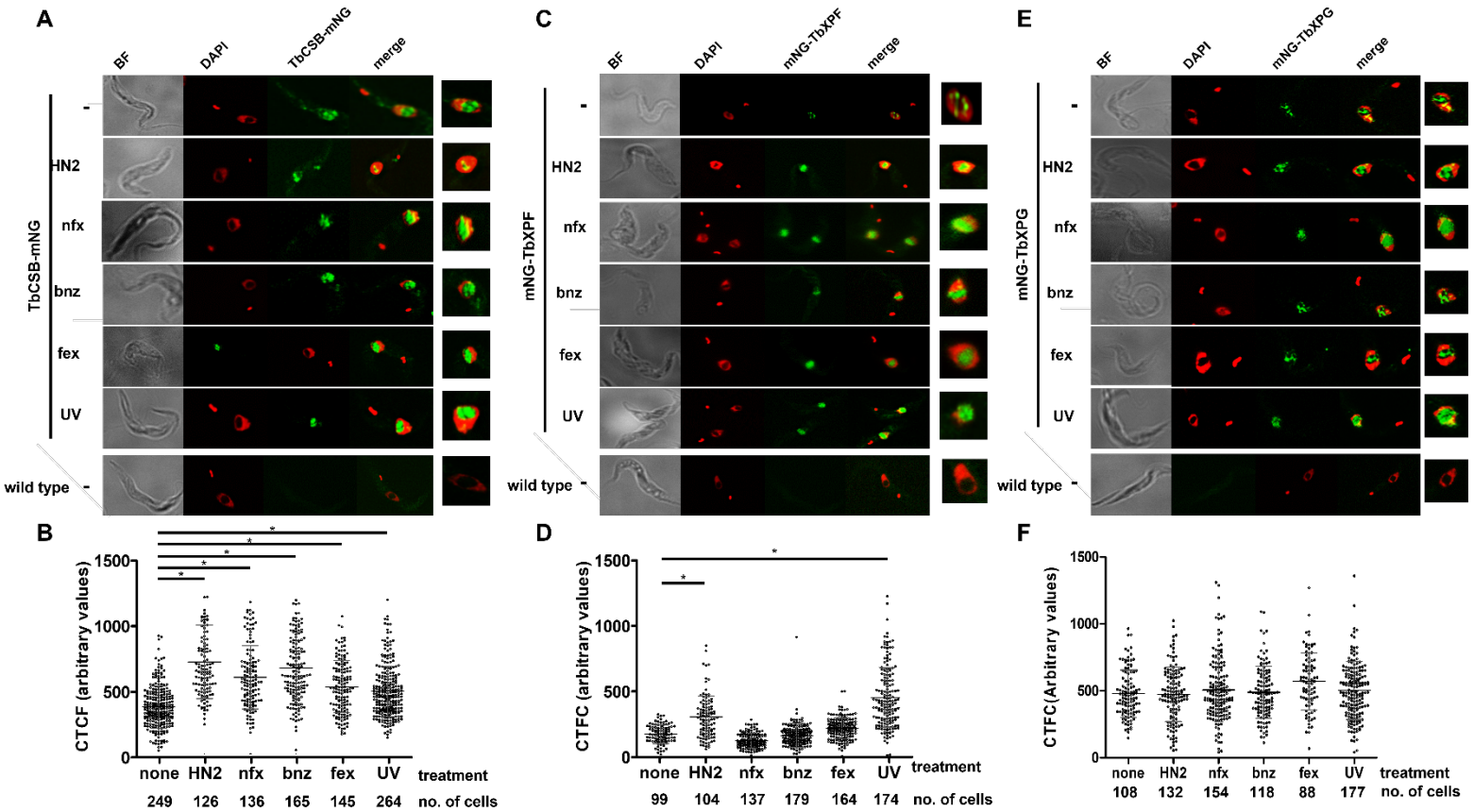


Fig 6.25. Effect of exogenous treatments on TbCSB, TbXPF or TbXPG expression.

A, C & E. PCF *T. brucei* expressing TbCSB-mNG (**A**), mNG-TbXPF (**B**) or mNG-TbXPG (**C**) (green) were treated with mechlorethamine (HN2; 30 µM for 6 hours), nifurtimox (nfx; 3 µM for 3 hours), benznidazole (bnz; 30 µM for 3 hours), fexinidazole (fex; 3 µM for 4 hours) or UV (1500 J m^{-2} for 6 hours) were co-stained with DAPI (red). The cells were examined by fluorescence microscopy & the brightfield (BF) image captured: Untreated (-) recombinant & wild type parasites were analysed in parallel. The pattern of colocalization (merge; yellow), including a close-up image of the nucleus, are shown.

B., D & F. The corrected total cell fluorescence (CTCF) of individual *T. brucei* cells expressing TbCSB-mNG (**B**), mNG-TbXPF (**D**) or mNG-TbXPG (**F**) following mechlorethamine (HN2), nifurtimox (nfx), benznidazole (bnz), fexinidazole (fex) or UV treatment was determined & compared against untreated controls (none). Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the horizontal lines. The number of cells analysed for each treatment is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).

The above analysis was then extended to assess the effect treatment with a single bolus of UV (1500 J m^{-2}), mechlorethamine (30 µM), fexinidazole (3 µM) or benznidazole (30 µM) has on the temporal expressing of TbCSB and TbXPF. Following exposure of *T. brucei* cultures to UV, TbCSB-mNG expression was shown to steadily increase, becoming significant 2-hours posttreatment, before reaching maximum fluorescence by 4-hours posttreatment. The signal subsequently declined over time point but still remained above values noted for the untreated control (Figure 4.26A). For TbXPF, the fluorescence signal rose rapidly over the first hour of treatment then continued to rise albeit at a slower rate, over the remainder of the time points analysed (Figure 4.26B). A summary of the combined temporal changes in expression in response to UV is presented in Figure 6.26C.

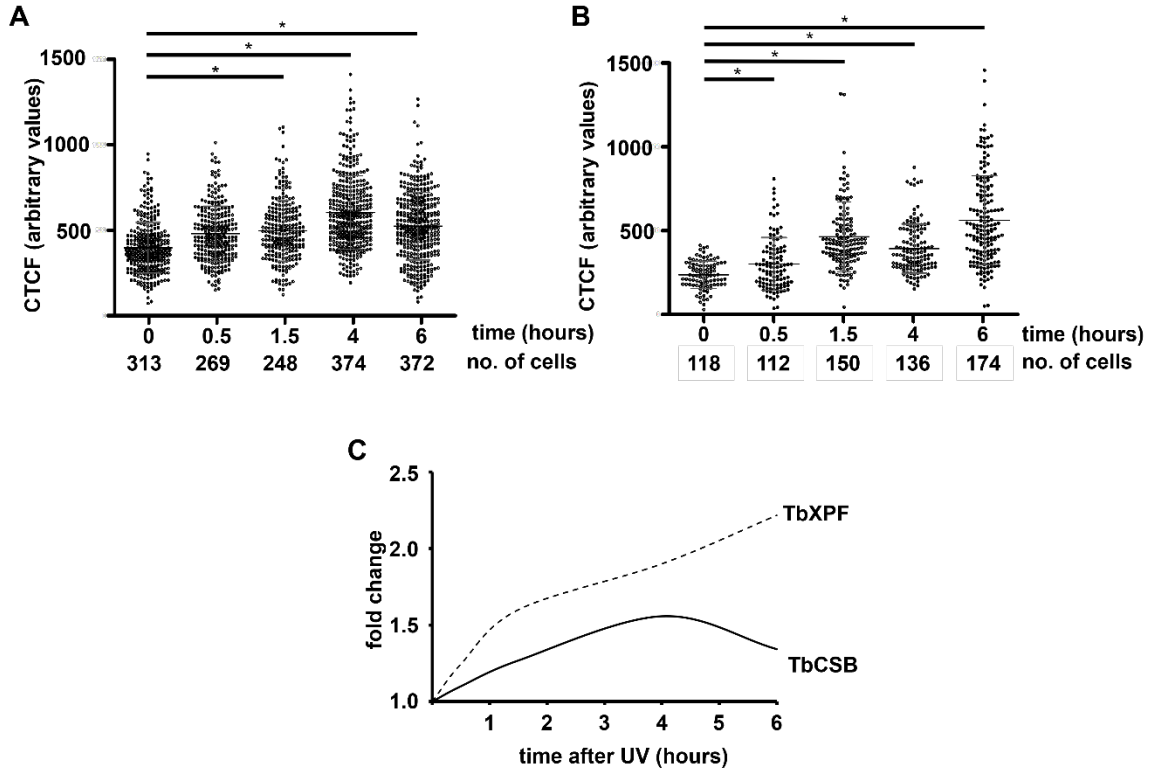


Fig 6.26. Effect of UV on the temporal expression of TbCSB and TbXPF.
A & B. The corrected total cellular fluorescence (CTCF) of individual *T. brucei* cells expressing tagged TbCSB-mNeonGreen (**A**) or mNeonGreen-TbXPF (**B**) was determined at time intervals following UV irradiation (1500 J m^{-2}). Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).
C. Schematic showing the temporal changes in expression of TbCSB or TbXPF following addition of UV irradiation.

When examining the effect of mechlorethamine, TbCSB and TbXPF were both shown to progressively rise, reaching a peak 3-hours posttreatment before plateauing and remaining at a level significantly above that noted for untreated cells (Figure 6.27A and B). A summary of these temporal changes in protein expression in response to mechlorethamine is presented in Figure 6.27C.

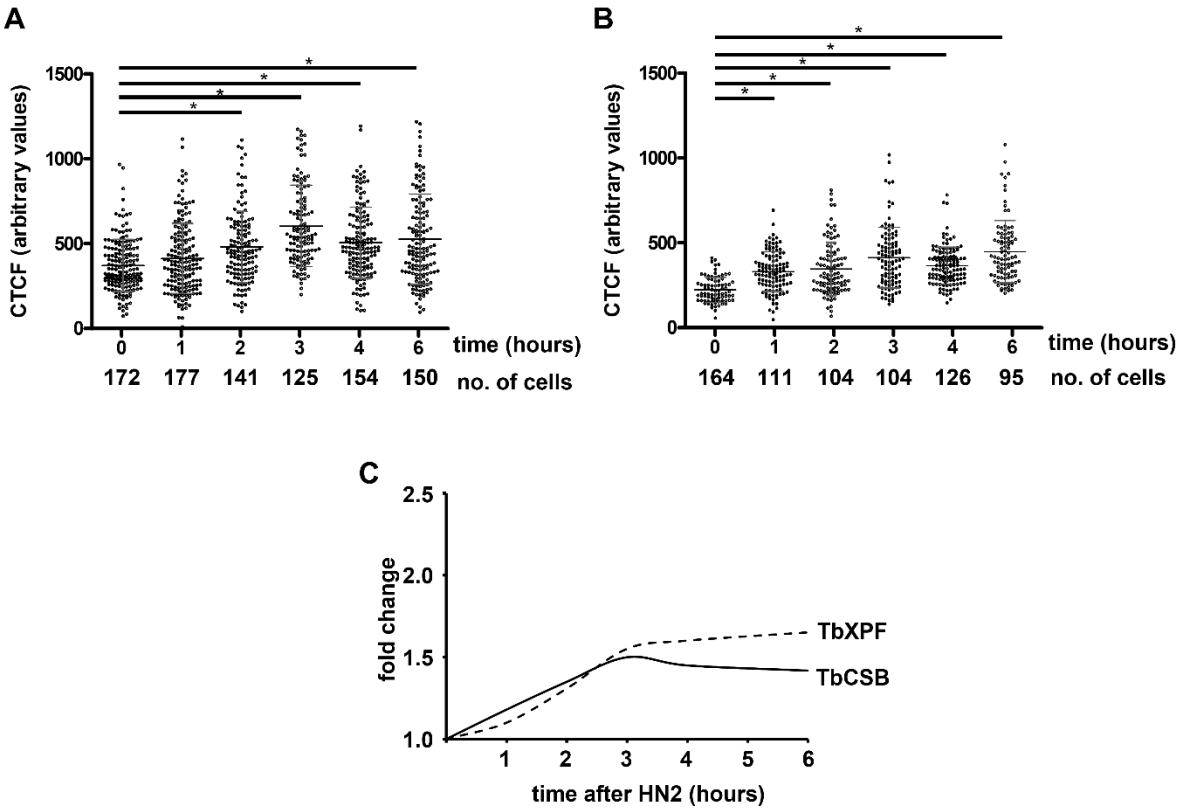


Fig 6.27. Effect of mechlorethamine on the temporal expression of TbCSB and TbXPF.
A & B. The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing tagged TbCSB-mNeonGreen (**A**) or mNeonGreen-TbXPF(**B**) was determined at time intervals following mechlorethamine (HN2, 30 μM) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).
C. Schematic showing the temporal changes in expression of TbCSB or TbXPF following addition of mechlorethamine (HN2) to cultures.

Exposure of TbCSB-mNG expressing PCF *T. brucei* to fexinidazole resulted in a slight increase in total cellular fluorescence, becoming significant 2-hours posttreatment (Figure 6.28A). The signal then remained roughly at the same level over the next four hours. For TbXPF, fexinidazole

treatment promoted alterations in protein expression, contrary to previous observations (Figure 6.28B). Here, the total cell fluorescence associated with mNG-TbXPF remained at untreated levels during the first two hours of treatment before rising and reaching a peak 4-hours posttreatment then falling back towards untreated levels by 6-hours posttreatment. The apparent discrepancy between the data sets displayed in Figures 6.25C and 6.28B may simply be a timing issue as to when samples were taken relative to each other and/or may reflect the transient nature of TbXPF expression in response to fexinidazole. A summary of the temporal changes in protein expression in response to fexinidazole is presented in Figure 6.28C.

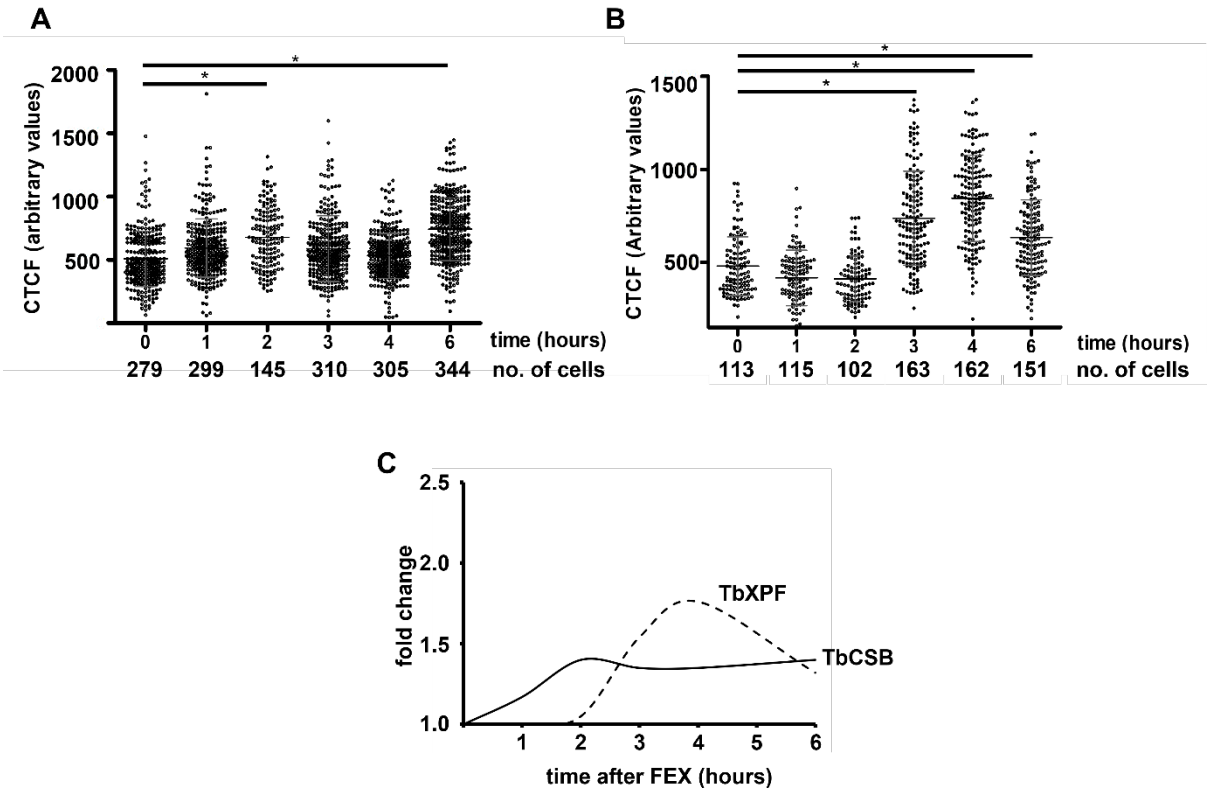


Fig 6.28. Effect of fexinidazole on the temporal expression of TbCSB and TbXPF.
A & B. The corrected total cellular fluorescence (CTCF) of individual *T. brucei* cells expressing tagged TbCSB-mNeonGreen (**A**) or mNeonGreen-TbXPF (**B**) was determined at time intervals following fexinidazole (FEX, 3 μM) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).
C. Schematic showing the temporal changes in expression of TbCSB or TbXPF following addition of fexinidazole (FEX) to cultures.

Following exposure to benznidazole, both TbCSB and TbXPF expression levels were shown to transiently rise. When TbCSB-mNG expressing parasites were exposed to the 2-nitroimidazole, the fluorescence associated with mNeonGreen remained at untreated levels during the first hour of treatment, then rapidly rising to reach a peak 2-hours posttreatment before slowly declining towards untreated levels over the next 4 hours (Figure 6.29A). In relation to TbXPF, expression of the tagged protein remained at untreated levels over the first two hours of treatment before transiently increasing 3-hours posttreatment and then subsequently declining back to untreated levels 4-hours posttreatment (Figure 6.29B). Again, the TbXPF pattern is contrary to previous observations (Figure 6.25C). This difference may reflect the transient nature of TbXPF expression in response to benznidazole or possibly be an operational error. A summary of the temporal changes in protein expression in response to fexinidazole is presented in Figure 6.29C.

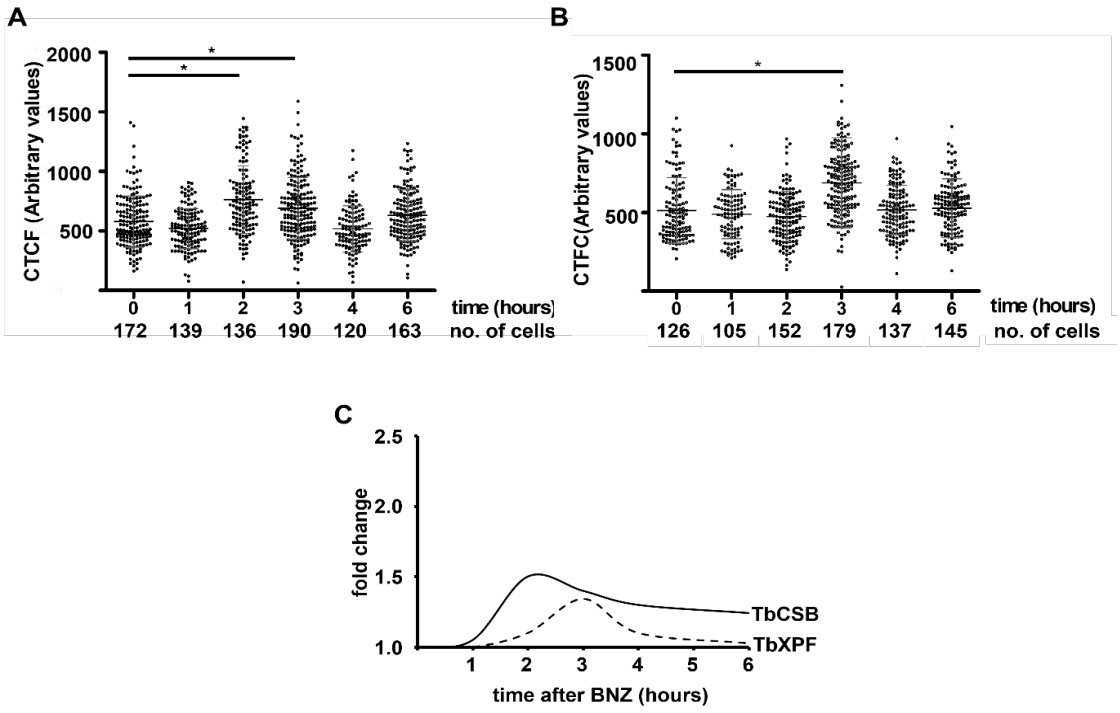


Fig 6.29. Effect of benznidazole on the temporal expression of TbCSB and TbXPF.
A & B. The corrected total cellular fluorescence (CTCF) of individual *T. brucei* cells expressing tagged TbCSB-mNeonGreen (**A**) or mNeonGreen-TbXPF (**B**) was determined at time intervals following benznidazole (BNZ; 30 μM) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the grey horizontal lines. The number of cells analysed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($P < 0.001$) between untreated and treated cells, as assessed by Kruskal-Wallis's test (GraphPad Software).
C. Schematic showing the temporal changes in expression of TbCSB or TbXPF following addition of benznidazole (BNZ) to cultures.

6.6 Functional analysis of the trypanosomal NER pathway using RNA interference.

RNA interference (RNAi) is a powerful technique for down regulating gene expression. This approach was first noted in *T. brucei* by Ngo and colleagues (Ngô et al., 1998) and for about 15 years was one of the mainstay approaches used to assess trypanosomal gene function (Horn, 2022). The whole RNAi process is driven by double stranded RNA (dsRNA) with two genetic-based systems widely used to promote formation of such molecules within *T. brucei* with one involving the expression of a stem-loop RNA structure from a single promoter and the other requiring expression of a short DNA fragment from two 'head-to-head', opposing promotes (Shi et al., 2000; Wang et al., 2000; LaCount & Donelson, 2001). Here, using the p2T7^{Ti} RNAi vector system developed by Alsford and Horn (2008) that relies on the second of these systems, internal DNA fragments from genes coding for GG-NER (Tb*ddb1,* Tb*xpc,* Tb*rad23*) or NER incision (Tb*ercc1,* Tb*xpf*, Tb*xpg*) factors were amplified from *T. brucei* gDNA (primers designed using RNAit (Redmond et al., 2003) then cloned into the pRNAiTbSpSyn vector backbone (Taylor et al., 2008): in this cloning the existing Tb*SpSyn*-derived DNA fragment was replaced with the amplicon containing sequence from a NER-coding gene. The resultant plasmids were linearized and transformed into the BSF *T. brucei* single marker bloodstream form (SMB) line (Wirtz et al., 1999), cells that constitutively express T7 RNA polymerase and the tetracycline repressor protein thus facilitating the tetracycline-inducible ectopic expression of dsRNA.

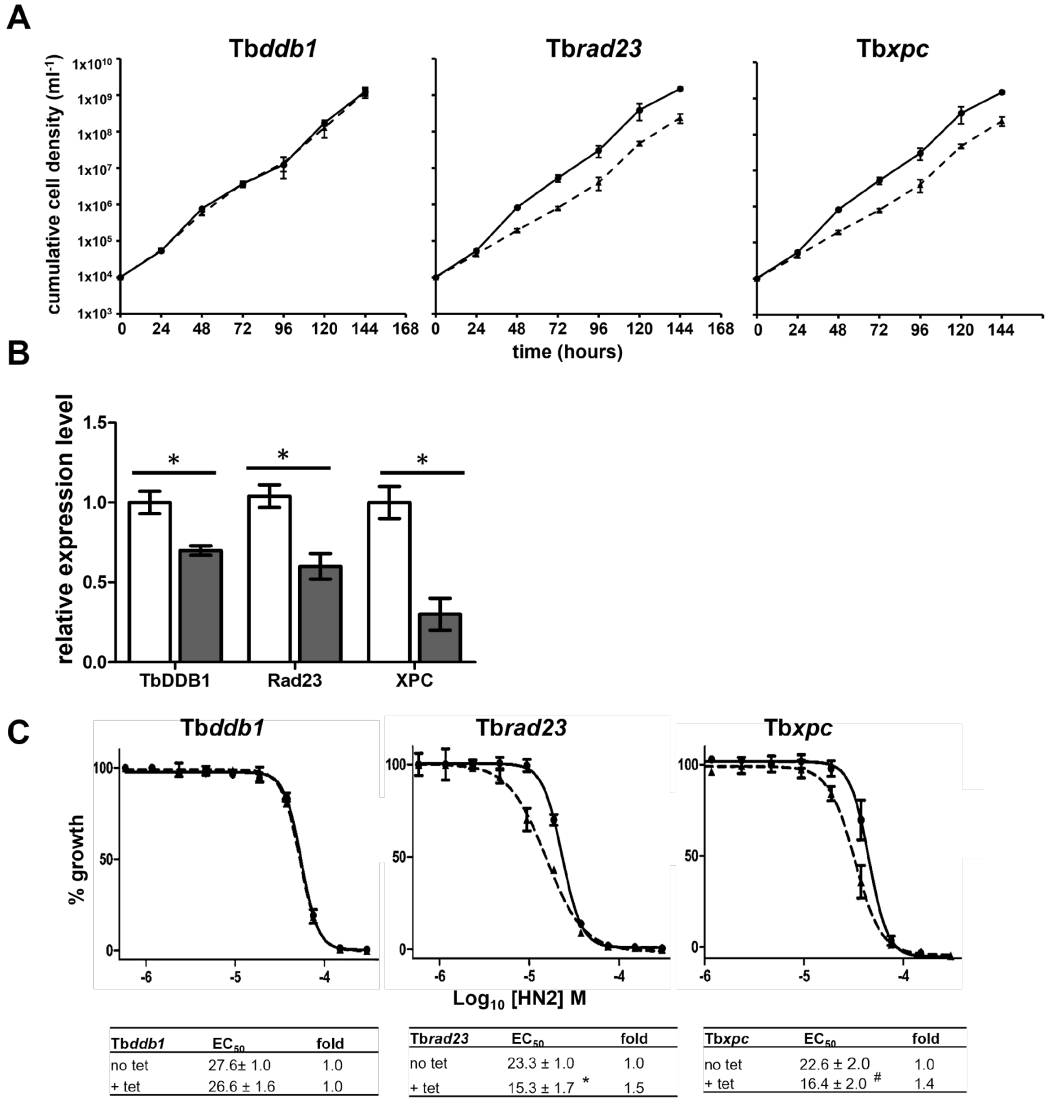
Following selection and establishment of drug (hygromycin) resistant recombinant lines, the effect of inducing RNAi on *T. brucei* growth was examined by determining the cumulative cell density of tetracycline-treated parasites relative to untreated controls (Figures 6.30A and 6.31A). In all cases, cells induced to undergo RNAi were shown to grow at an equivalent rate as to parasites cultured in the absence of the antibiotic irrespective of which mRNA was being targeted. This suggesting that none of the encoded proteins (TbDDB1, TbRAD23, TbXPC, TbERCC1, TbXPF or TbXPG) are essential for growth of BSF *T. brucei*. To confirm that down regulation of the targeted transcript had indeed taken place, RT-PCRs were performed on cDNA samples generated from total RNAs extracted from tetracycline-treated (48-hours) *T. brucei* cultures with outputs compared to untreated controls. For Tb*ddb1,* Tb*xpc*, Tb*rad23* and Tb*xpg*, a 30, 70, 40 or 50% reduction in the corresponding transcript was noted in cells induced to undergo RNAi, respectively (Figure 6.30B and 6.31B). In contrast, this analysis revealed that the Tb*xpf* mRNA level was unchanged between the two compared cultures while curiously, the Tb*ercc1* transcript was shown to be

elevated (approximately 1.5 higher) in the tetracycline-treated culture relative to untreated control (Figure 6.31B). Together, this data indicates that differing levels of down regulation towards Tb*ddb1,* Tb*xpc*, Tb*rad23* and Tb*xpg* can be promoted by the RNAi system used here without affecting cell growth. With regards to Tb*ddb1* and Tb*xpc*, the results generated here are contrary to the published literature (Machado et al 2014). In this earlier study, both GG-NER factors were shown to be essential for growth of BSF *T. brucei*. These discrepancies may be due to the different RNAi systems employed in the two studies: here the 'head-to-head' promoter system was employed to generate dsRNA while Machado and colleagues (2014) used a stem-loop system. For Tb*ddb1*, the 'head-to-head' promoter RNAi system may not have reduced the corresponding transcript sufficiently to generate the growth defect phenotype noted by Machado and colleagues (2014). In the case of Tb*xpc* (and Tb*ddb1*), the kinetics of RNAi targeting this mRNA the 'head-to-head' promoter RNAi system may have been slow such that, even though down regulation of the desired transcript was taking place, the parasites had a window of opportunity to adapt and alter their expression of other genes to compensate for the lowered levels of the TbXPC (or TbDDB1).

The susceptibility of all RNAi lines towards mechlorethamine was assessed to determine whether any of the selected GG-NER or incision NER factors are involved in ICL repair (Figure 6.30C and 6.31C). This revealed that cells undergoing RNAi targeting the Tb*rad23* or Tb*xpc* mRNA were up to 1.5-fold more susceptible to the agent tested than controls: no difference in mechlorethamine sensitivity was noted for Tb*ddb1*, Tb*xpf* or Tb*xpg*. Intriguingly, cells potentially expressing elevated levels of the Tb*ercc1* mRNA (see Figure 6.31B) are also more susceptible (about 2-fold) to mechlorethamine. In this case, elevated levels of TbERCC1 may promote off target effects that could impinge on TbERCC1/TbXPF heterodimer function that in human and yeast cells often leads to chromosomal aberrations (Hwang & Mendell, 2006; Jacobsen et al., 2017), a situation that could be magnified by an ICL inducing agent. Together, these findings support the idea proposed by Machado and colleagues (2014) that GG-NER factors play a key role in ICL repair and suggest this part of the NER pathway plays a critical process in trypanosomatid genome maintenance.

ERCC1 and XPF have a unique role in response to ICL damage, where they function as scissors tasked with unhooking of the ICL. XPF-ERCC1 and XPG unhook ICLs by nucleolytic processing of sequences 5' and 3' to the ICL on one strand of the double helix (Wood, 2010; Abdullah et al.,

2017; Faridounnia et al., 2018) to allow other enzymes such as SNM1A to be recruited to the damage (Wang et al., 2011; Sengerova et al.,2012). Depletion of XPF-ERCC1 eukaryotes causes increased susceptibility to ICL induced by cisplatin, psoralen and Mitomycin C (Wood, 2010). In our system, however, contrasting results were obtained in relation to ICL induction of RNAi TbXPF or ERCC1. Whereas RNAi TbERCC1 resulted in increased susceptibility to ICL, depletion of TbXPF by RNAi did not affect survival in ICL inducing agents.



Tb*ddb1*	EC_{50}	fold
no tet	27.6± 1.0	1.0
+ tet	26.6 ± 1.6	1.0

Tb*rad23*	EC_{50}	fold
no tet	23.3 ± 1.0	1.0
+ tet	15.3 ± 1.7 *	1.5

Tb*xpc*	EC_{50}	fold
no tet	22.6 ± 2.0	1.0
+ tet	16.4 ± 2.0 #	1.4

Fig 6.30. Characterization of RNAi cell lines targeting the *T. brucei* GG-NERA pathway.

A. Growth of cells induced to undergo RNAi (dashed lines) targeting the Tb*ddb1*, Tb*rad23* or Tb*xpc* transcript was compared against non-induced cultures (solid line). All data points represent the mean cell density ± standard deviation from two independent cultures, each performed in triplicate.

B. The Tb*ddb1*, Tb*rad23* or Tb*xpc* transcript levels from non-induced cells (white bars) and trypanosomes induced for 48 hours to undergo RNAi (grey bars) was assessed by qPCR and compared against the expression level of a standardized control (Tb*tert*). The relative fold difference, as judged by 2-(ΔΔCT) from reactions performed in quadruplicate ± 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 Tb*ddb1*, Tb*rad23* or Tb*xpc* mRNAs were approximately 1.4-, 1.7- and 3.0-fold lower in the corresponding RNAi line relative to controls.

C. The BSF RNAi lines (RNAi-Tb*ddb1*, RNAi-Tb*rad23* and RNAi-Tb*xpc*) were grown for 24 hours in the presence of tetracycline (1 μg ml^{-1}) (dashed line), seeded at 1 × ml–1 and then exposed to various concentrations of mechlorethamine (HN2). After 3 days at 37 °C, resazurin (2.5 μg) was added to each culture and used to determine cell density. Untreated (solid line) parasites were analysed in parallel. From the resultant dose response curves, the EC_{50} of each line towards HN2 was calculated (Table below graph). All data points are means for experiments performed in quadruplicate ± standard deviation. * and # indicates significant differences in susceptibility toward HN2 of $P = 0.0004$ and $P=0.0046$ ($P<0.001$), respectively, in tetracycline treated and untreated cultures, as assessed by Student's *t* test (GraphPad Software).

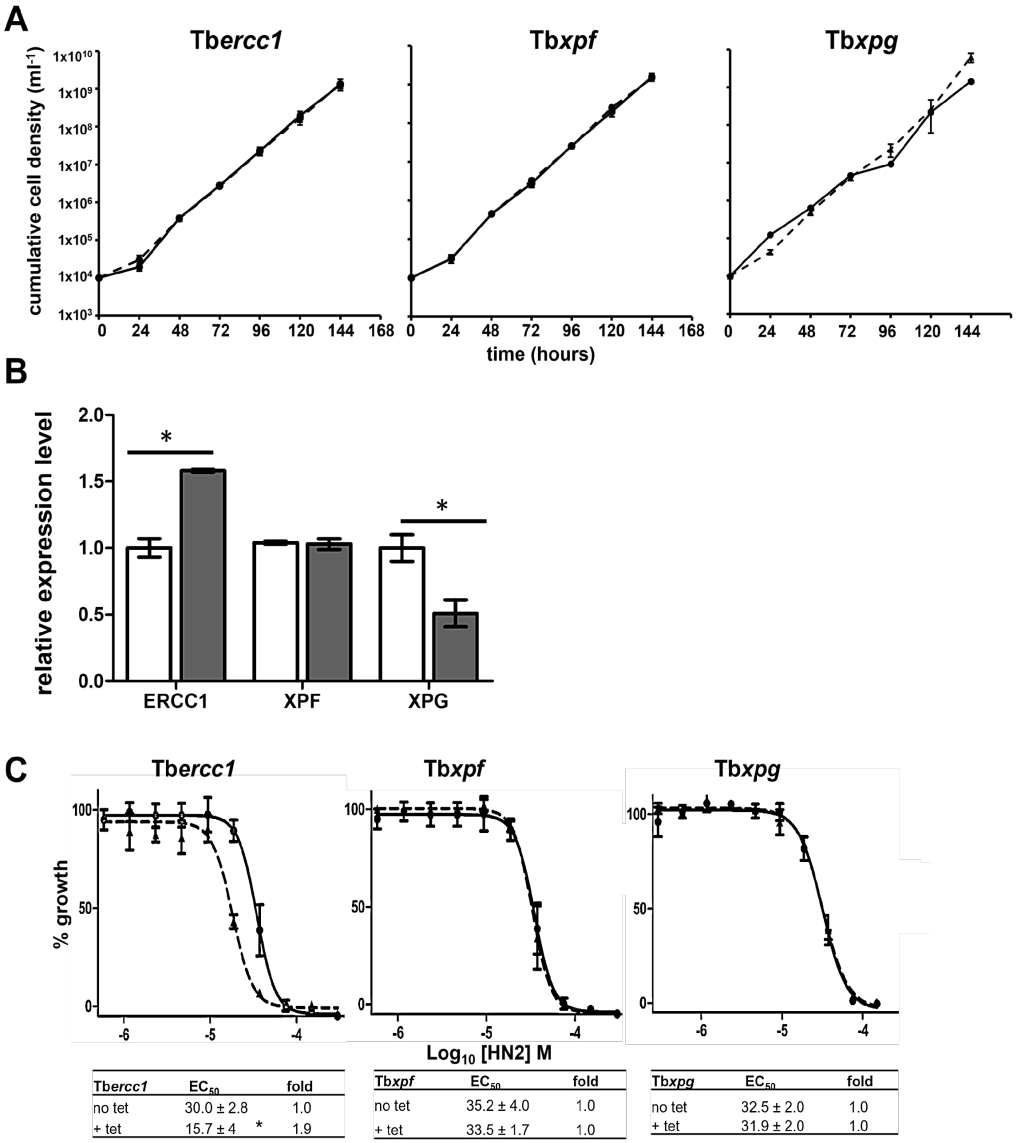


Fig 6.31. Characterization of RNAi cell lines targeting incision factors of *T. brucei* NER.

A. Growth of cells induced to undergo RNAi (dashed lines) targeting the Tb*ercc1*, Tb*xpf* or Tb*xpg* transcript was compared against non-induced cultures (solid line). All data points represent the mean cell density ± standard deviation from two independent cultures, each performed in triplicate.

B. The Tb*ercc1*, Tb*xpf* or Tb*xpg* transcript levels from non-induced cells (white) and trypanosomes induced for 48 hours to undergo RNAi (grey) was assessed by qPCR and compared against the expression level of a standardized control (Tb*tert*). The relative fold difference, as judged by 2-(ΔΔCT) from reactions performed in quadruplicates ± 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 Tb*ercc1* mRNA was deemed to be approximately 1.5 higher, the Tb*xpf* transcript level was deemed to be unchanged and the Tb*xpg* message was approximately 2.0-fold lower in the corresponding RNAi line relative to controls.

C. The BSF RNAi lines (RNAi-Tb*ercc1*, RNAi-Tb*xpf* and RNAi-Tb*xpg*) lines were grown for 24 hours in the presence of tetracycline (1 μg ml^{-1}) (dashed line), seeded at 1 × ml^{-1} then exposed to various concentrations of mechlorethamine (HN2). After 3 days at 37 °C, resazurin (2.5 μg) was added to each culture and used to determine cell density. Untreated (solid line) parasites were analysed in parallel. From the resultant dose response curves, the EC_{50} of each line towards HN2 was calculated. All data points are means for experiments performed in quadruplicate ± standard deviation. Asterisk (*) indicates significant differences in susceptibility toward HN2 of P = 0.0011 in tetracycline treated and untreated cultures, as assessed by Student's *t* test (GraphPad Software).

6.7. Chapter Summary

In this chapter, we explored the role of *T. brucei* NER factors, DDB1, RAD23 and XPC, CSB and the endonucleases ERCC1, XPF and XPG. PCR-based CRISPR/Cas9 genome editing strategy (Beneke et al., 2017)) was employed to engineer parasite lines that contain mNeonGreen Tag of these proteins. Additionally, the subcellular localisation of these proteins was determined using fluorescence microscopy. The kinetics of *T. brucei* NER pathway in repairing DNA damage generated by trypanocidal nitroheterocyclic prodrugs, UV irradiation and ICL inducing agent(mechlorethamine) was also determined. *T. brucei* lines expressing lower levels of target transcript (DDB1, RAD23, XPC, ERCC1, XPF and XPG) were obtained using RNA interference (RNAi). The effect of RNAi induction on expression of these proteins was evaluated with and without DNA damage.

Here we showed:

1. The sub-cellular localisation of mNeonGreen tagged *T. brucei* DDB1, RAD23 and XPC the GG-NER, CSB from TC-NER and the endonucleases ERCC1, XPF and XPG in bloodstream and/ or procyclic form parasites.
2. *T. brucei* NER factors CSB, XPG and XPF show increased expression in the presence of DNA damage. Like other NER systems, *T. brucei* CSB and XPF respond in a temporal manner to UV irradiation and ICL as well to trypanocidal nitroheterocyclic prodrugs.
3. Initiation RNAi of *T. brucei* GG-NER caused reduction in transcript levels for Tb*ddb1,* Tb*xpc*, Tb*rad23* and Tb*xpg*, respectively. Increased expression was seen with TbERCC1, whereas no effect was seen with TbXPF.

CHAPTER 7: Discussion

Nitroheterocyclic prodrugs are important in treatment of trypanosomal infections and have for several years constituted a significant proportion of therapies used against HAT and chagas disease (Deeks, 2019; Patterson & Fairlamb, 2019; Priotto et al., 2009; Wilkinson & Kelly, 2009). In recent years, rediscovery of nitroimidazole-based compounds in drug mining has resulted in the approval of pretomanid, delamanid, and fexinidazole for treatment of tuberculosis or HAT (Deeks, 2019; Liu et al., 2018; Patterson & Fairlamb, 2019; Thakare et al., 2020). Trypanocidal activity of benznidazole, fexinidazole and involve nitro-reduction by parasite specific NTR1 to generate cytotoxic metabolites that target macromolecules including DNA but are unharmful to humans (Hall et al., 2011; Hall & Wilkinson, 2012; Tweats et al., 2012; Trochine et al., 2014;). The type of DNA damage caused by these drugs, or the kinetics of DNA damage and repair is still emerging. Evidence from functional genomics studies link damage induced by benznidazole and nifurtimox treatment of *T. brucei* or *cruzi* to mutations in the parasite's genome, resulting in changes in gene copy number and the development of resistance to benznidazole and other nitroheterocyclic*s* (Campos et al., 2017; Díaz de Toranzo et al., 1988; Nozaki et al., 1996; Trochine et al., 2014). *T. brucei* or *T. cruzi* lines expressing elevated or reduced levels of DNA repair enzymes from MMR or HR also show altered susceptibilities to benznidazole (Rajao et al., 2014). Recent work from the Wilkinson group revealed that TbNTR1 activation of benznidazole promotes DNA damage in the nuclear genome of *T. brucei* that mobilises components of the HR system to resolve this damage, thus giving a direct link between prodrug activation and a downstream trypanocidal mechanism (Dattani et al 2021). In Chapter 3, we investigated whether fexinidazole also possesses a related activity, by exploring the relationship between its activation and downstream effect on TC-NER and HR. We also explored whether an epistatic interaction exists between TbMRE11 and CSB in resolving any damage generated by fexinidazole.

DNA repair is conserved across all domains of life and involve well-choreographed steps with interaction between different DNA repair factors and pathways to resolve the damage and generate repaired (intact) DNA heteroduplex (Beljanski et al., 2004). The genome of *Trypanosoma brucei* encode several canonical DNA repair genes but how these are expressed are in many ways divergent from other higher eukaryotes (Glover et al., 2013; Ubeda et al., 2014; Mehnert et al.,2021). In this project, we evaluated how trypanosomal DNA repairtoire respond to exogenous DNA damage. CRISPR/Cas9 based genome editing was used to engineer cell lines that expressed

recombinant mNeonGreen tagged enzymes from the BER, HR, ICL as well as NER pathways. Using these engineered lines, we investigated the role the above *T. brucei* DNA repair pathways play in resolving DNA damage induced by exogenous agents including Trypanocidal nitroheterocyclic prodrugs, ICL inducing agents, UV irradiation or double strand inducing agent(phleomycin). This enabled us to understand the temporal expression of *T. brucei* BER, HR, ICL and NER DNA repair upon exogenous DNA damage and elucidated how factors from these pathways are recruited upon damage induction (Chapter 3-6). RNA interference of key *T. brucei* GG-NER factors was used to shed light on the contribution of these enzymes to DNA repair (Chapter 6). Reverse genetics-based approach using CRISPR/Cas9 genome editing was also utilised to delete or disrupt genes that encode factors from the ICL and BER pathways to investigate their contribution to *T. brucei* DNA repair (Chapter 4 &5).

One of the initial aims of this project was to use the CRISPR/Cas9 generated mNeonGreen tagged cell lines to perform pulldown experiments in order to study protein: protein interactions between factors from the different DNA repair pathways. Unfortunately, we were unable to perform these experiments due to the interruptions by the COVID-19 pandemic. The cell lines generated in this project, however, can be utilised to expand knowledge on interaction of the different DNA repair pathways studied here.

7.1. Elucidating the trypanocidal activity of fexinidazole

Fexinidazole is a prodrug requiring activation by parasite specific NTR1(Deeks, 2019). Using cell lines that that overexpress TbNTR1(*Tbntr*$^{+/+}$) (Wilkinson et al., 2008; Baker et al 2011; Hall et al 2011; Hall & Wilkinson, 2012), I confirmed that activation of this compound is linked to nitro reduction by TbNTR1. Lines overexpressing TbNTR1(*Tbntr*$^{+/+}$) were ~11 times more susceptible to fexinidazole than wild type, Since activation fexinidazole generates cytotoxic metabolites, that cause DNA damage(Tweats et al.,2012), I investigated whether the downstream effect of fexinidazole activation involves the recruitment of *T.brucei* DNA repair factors from the HR, TC-NER and ICL DNA repair pathway using recombinant BSF *T. brucei* lines null for TbMRE11, TbCSB or TbSNM1 respectively (Sullivan et al 2015; Dattani and Wilkinson, 2019). Fexinidazole activation resulted in the mobilisation of TbCSB and TbMRE11 from the TC-NER and HR pathways respectively but not TbSNM1. TbCSB and TbMRE11 were 4-5-fold more susceptible

to fexinidazole than wild type, indicating that this nitroimidazole may promote DNA damage in the nuclear genome.

Next, I investigated whether the epistatic or non-epistatic relationship previously seen between CSB and MRE11 when treated with benznidazole and other nitroimidazoles (Dattani et al., 2021), is also involved in resolving fexinidazole induced DNA damage. Here. using lines deficient in TbCSB or TbMRE11 and null for both enzymes, I showed that just like benznidazole (Dattani et al., 2021), fexinidazole activation resulted in mobilisation and epistatic interaction of TbCSB or TbMRE11. In addition to transcription coupled, NER, CSB also contributes to repair of TC-HR, and BER (Batenburg et al., 2015; Wei et al., 2016; Menoni et al., 2012, 2018; Stevnsner et al., 2008). TbCSB and TbMRE11 are multi-functional proteins that cooperate in the repair several types of lesions in eukaryotes as well as *in T. brucei* and epistatic interactions between these proteins have previously been shown, where both proteins cooperate in the repair of transcription coupled HR. In this type of repair, CSB and MRE11 are mobilised to stalled replication forks. Here, the chromatin remodelling domain or ATPase activity of CSB enhances MRE11 recruitment and role in fork reversal by degradation (Batenburg, 2021). TbMRE11 and TbCSB contribute epistatically in ICL repair (Dattani & Wilkinson, 2019) as well as in the repair of DSB caused by benznidazole (Dattani et al., 2021).

Given the above findings, I evaluated whether mobilisation of TbCSB or TbMRE11 was as result of downstream effect of TbNTR1 activation of fexinidazole. I assessed the susceptibility of lines null for Tb*mre11 or* Tb*csb,* but express ectopic copies of Tb*ntr1* (Dattani et al 2021), alongside lines that overexpress only TbNTR. Here lines null for Tb*mre11* or Tb*csb but* expressed ectopic copies of Tb*ntr1* were hypersensitive to fexinidazole compared to controls lacking Tb*mre11* or Tb*csb* alone. Tb*mre11* or Tb*csb lines* containing ectopic copies of Tb*ntr1* were 13 and 25-fold more sensitive to fexinidazole than the wild type or Tb*ntr1*$^{+/+}$alone which was only 5-12-fold more sensitive to fexinidazole than the wild type. These results show a clear correlation between TbNTR1-mediated prodrug activation of fexinidazole generating cytotoxic lesions that are repaired by mobilisation of both TbCSB and TbMRE11.

7.2. Kinetics of *T. brucei* DNA repair and possible interactions

PCR based CRISPR/Cas9 genome tagging (Beneke et al., 2017) of the endogenous 5' or 3' region of *Tbmre11* and *Tbrad*51(HR), *Tbfan1* and *Tbsnm1*(ICL), *Tbape1, Tbmyh* and *Tbnth1*

(BER) as well as *Tbcsb*, *Tbddb1*, *Tbercc1* Tb*rad23*, *Tbxpc, Tbxpf* or *Tbxpg* from the NER pathway with nucleotide sequences coding for the fluorescent protein mNeonGreen (Shaner et al., 2013). All tagged lines were validated by PCR and fluorescence microscopy. Table 8.1 summarises the outcome of the tagging of all tagged lines generated. The results shown here suggests that tagging of the endogenous gene loci did not interfere with protein recruitment and localisation.

	BSF		**PCF**	
	mNG-TbXXX localisation	**TbXXX-mNG localisation**	**mNG-TbXXX localisation**	**TbXXX-mNG localisation**
TbAPE1	Nuclear - in nucleolus	Nuclear - in nucleolus	Nuclear - in nucleolus	Nuclear - in nucleolus
TbMYH	X	Nuclear - in nucleolus	cytoplasmic	Cytoplasmic
TbNTH1	Nuclear - in nucleolus	Nuclear - in nucleolus	X	Nuclear - in nucleolus
TbCSB	cytoplasmic	Nuclear - in nucleolus	cytoplasmic	Nuclear - in nucleolus
TbXPC	X	X	X	Nuclear - punctate pattern in nucleoplasm and in nucleolus
TbDDB1	X	Punctate pattern throughout cytoplasm	X	Punctate pattern throughout cytoplasm
TbRAD23	X	Punctate pattern throughout cytoplasm	X	cytoplasmic
TbERCC1	Punctate pattern throughout cytoplasm	Punctate pattern throughout cytoplasm	Punctate pattern throughout cytoplasm	Punctate pattern throughout cytoplasm
TbXPF	Nuclear - punctate pattern in nucleoplasm	Nuclear - punctate pattern in nucleoplasm	Nuclear - punctate pattern in nucleoplasm	Nuclear - punctate pattern in nucleoplasm
TbXPG	X	Nuclear - punctate pattern in nucleoplasm	X	Nuclear - punctate pattern in nucleolus
TbMRE11	Nuclear - punctate pattern in nucleoplasm	Nuclear - punctate pattern in nucleoplasm	Nuclear - punctate pattern in nucleoplasm	Nuclear - punctate pattern in nucleoplasm and in nucleolus

TbRAD51	Nuclear - punctate pattern in nucleoplasm nucleolar	Nuclear - punctate pattern in nucleoplasm and in nucleolus	Nuclear - punctate pattern in nucleoplasm and in nucleolus	Nuclear - punctate pattern in nucleoplasm and in nucleolus
TbFAN1	Nuclear - punctate pattern in nucleoplasm	Nuclear - punctate pattern in nucleoplasm	Nuclear - punctate pattern in nucleoplasm	Nuclear - punctate pattern in nucleoplasm and in nucleolus
TbSNM1	Nuclear - punctate pattern in nucleoplasm	X	Nuclear - punctate pattern in nucleoplasm	X

Table 7.1: Summary of the localisation of the different proteins studied in this project. 'X' denotes where no cell line was generated and TbXXX, represents the gene of interest.

To gain insight into how various DNA repair pathways are mobilised in response to DNA damage, selected lines were treated with exogenous agents and the sequential mobilisation of DNA repair evaluated by quantifying the fluorescence intensity of tagged proteins.

7.2.1 Recruitment of *T. brucei* DNA repairtoire

7.2.1.1. HR

Double strand breaks are induced endogenously or exogenously by 'genotoxic stressors' including replication break accidents, radiomimetic compounds, anticancer therapy as well as ionising radiation. Unrepaired DSB result in chromosome break leading to genome instability(da Silva, 2021). Trypanosomes also utilise double strand breaks to their advantage when exposed to host antibody pressure to maintain genome variability including antigenic variation as seen in *T. brucei* and maintenance of chromosome copy numbers in Leishmania. Several studies have shown that one of the mechanisms of action of trypanocidal drugs is via DNA damage induction. For example, damage induced by benznidazole and nifurtimox treatment of *T. brucei* or *cruzi* have been linked to mutations in the parasite's genome (Díaz de Toranzo et al., 1988; Nozaki et al., 1996; Trochine et al., 2014; Campos et al., 2017). Alterations in DNA repair enzymes from MMR or HR in *T. brucei* or *T. cruzi* lines also cause variable susceptibilities to benznidazole (Rajao et al., 2014). Recent work from the Wilkinson group (part of this thesis) demonstrated that the downstream effect of benznidazole activation involves DNA damage in the nuclear genome of *T. brucei* resulting in mobilisation of the HR system (Dattani et al 2021).

Exposure of parasites to exogenous DSB inducing agents often results in temporal mobilisation of HR pathway which is the main mechanism of DSB repair in trypanosomes. The first line of HR repair involves mobilisation of the MRN complex accompanied by an inverse relationship in

expression of TbMRE11 and TbRAD51, where MRE11 expression declines with phosphorylation of H2A and mobilisation and increase in expression of RAD51(Marin et al., 2018). In this project, treatment trypanocidal nitroheterocycles resulted in increased fluorescence intensities of TbMRE11 and TbRAD51. An accompanied temporal recruitment pattern was seen, with MRE11 and RAD51, just like in other eukaryotes (Eckert-Boulet et al., 2011; Lisby et al., 2004; Lisby& Rothstein, 2009). MRE11, as part of the tripartite M/R/N DNA damage recognition complex (Gomes Passos Silva et al., 2018) that detects and process the damaged DNA ends was mobilised after the first hour of drug treatment, similar to what is seen in mammalian MRE11 lines (Haince et al., 2008). Similarly, as MRE11 expression declined, γH2A was recruited and expression remained stable until TbRAD51 expression peaked as γH2A expression declined. This sequential mobilisation pattern of the *T. brucei* HR pathway was also noted in Fexinidazole treated mNG-TbMRE11 and RAD51 lines, although without γH2A. Unlike the benznidazole and fexinidazole treatment, exposure to nifurtimox caused a different type of mobilisation pattern expression of both proteins remaining elevated throughout the study period. Results from the nifurtimox treated lines was surprising because it contrasts phenotyping assays that showed that nifurtimox activation does not recruit factors from the HR pathway (Dattani et al., 2021). In another study, nifurtimox treatment affected the amount of protein in the mitochondrion but was not linked to DNA damage(Thomas et al., 2018). Alternatively, nifurtimox being a multi-target compound, may possibly trigger a type of damage that is recognised by both *Tb*MRE11 and *Tb*RAD51 but is repaired in a pathway independent of the HR system. Like nifurtimox, UV irradiation resulted in mobilisation of *Tb*MRE11 and *Tb*RAD51, however, the HR pathway did not seem to play a key role in repairing UV induced lesions. The results presented here suggest that TbRAD51 may be able to detect UV damage by being mobilised to the nucleolus and possibly is involved in recognising this lesion but does not seem to contribute significantly to resolving UV damage. Additionally, UV damage is repaired by factors from the TC-NER and since these lesions take a longer time to resolve than 6 hrs. It is possible that the duration of the experiment may not have been sufficient to study the complete repair of this type of lesion (Machado et al., 2014).

In addition to the above exogenous stimuli, ICL induction of HR factors using mechlorethamine resulted in the mobilisation of HR factors with a change in localisation pattern of both mNG-TbMRE11 and mNG-TbRAD51 seen. Temporary stalling of cells in the S phase of the cell cycle was also observed with ICL treatment. HR is one of the effectors of replication dependent ICL

repair systems. Here, temporal change in *Tb*MRE11 and *Tb*RAD51 expression did not involve interaction of both proteins. Both TbMRE11 and RAD51 expression remain elevated throughout the experiment. This is possibly because mechlorethamine causes spontaneous DNA damage which requires prolonged activation of DNA repair factors to resolve the damage. Additionally, since the cells were in solution containing mechlorethamine, several rounds of spontaneous DNA damage and repair may have been triggered.

Double strand break repair is important in maintaining genome stability and parasite survival through continuous antigenic variation to escape detection by host antibodies. Parasites deficient in dsDNA repair factors (TcRAD51) are more sensitive to damage, while those that overexpress these factors become more resistant (Vispé et al., 1998; Vieira-da-Rocha et al., 2019), HsRAD51 also plays a role in radioresistance in human, (Vispé et al., 1998). The results presented here show that the HR pathway is also important in modulating response from DNA damaging agents in *Trypanosoma brucei*. These results are important because understanding how DNA repair proteins are recruited, interact with DNA and with each other is crucial in designing inhibitors for these interactions ((Pearl et al., 2015; Weilbeer et al., 2022),

7.2.1.2. ICL

ICL repair involves cell cycle stage dependent cooperation of NER, FA, HR or TLS pathways to successfully restore DNA double helix. SNM1 and FAN1 are both ICL repair nucleases that operate independent or outside the FA pathway. FAN1 deficiency causes sensitivity to DNA crosslinking agents such as mitomycin C (MMC) (Kratz et al., 2010; Thongthip et al., 2016). Its role in ICL resolution includes structure specific unhooking of crosslinked DNA lesion in the replication dependent ICL repair. Here, FAN1 together with FANCP (SLX4), XPF-ERCC1 (NER pathway), MUS81-EME1 and SLX1 catalyse nucleolytic incision of the DNA backbone at sites 5' and 3' to the ICL resulting in 'unhooking' of the lesion from one strand of the double helix (Hashimoto et al., 2016). SNM1 is 5'-3' exonuclease, involved in the nucleolytic degradation of ICLS following unhooking by XPF-ERCC1 Moldovan and D'Andrea, 2009; Zhang and Walter, 2014; Hashimoto et al., 2016; Siddiqui et al., 2017). TbSNM1 is a structural homolog of Sc PSO2 and rescues the sensitivity of PSO2 mutants to DNA crosslinkers (Sullivan et al., 2015).

TbFAN1 is a standalone VRR-NUC domain like its bacterial counterpart. Its absence in *T. brucei* leads to increased sensitivity to ICL damage (Kumar 2019). TbFAN1 interacts with XPF-ERCC1 and MUS18 to unhook ICLs generating a DSB which in turn mobilise MRE11 to trim the

DNA ends (Dattani & Wilkinson, 2019). Results presented here show that TbFAN1 also responds to exogenous agents with temporal changes in expression seen after these treatments. In the presence of ICL damage (mechlorethamine), a rise in fluorescence intensity of treated cells took place in the first hour of treatment of mNeonGreen tagged TbFAN1 becoming significant at 4-hours posttreatment and continuing to rise until the final time point, 6-hours posttreatment (Figure 4.26A). Exposure to benznidazole and nifurtimox resulted in rapid increase in expression with a peak seen within the first 1-2 hours post treatment. In nifurtimox treated lines, a significant increase was noted 2-hours posttreatment declining to levels noted in untreated lines by 4-hours posttreatment (Figure 4.26 B). These results mirror what was seen in in cell free xenopus egg extracts where the kinetics of ICL repair of the eukaryotic interaction partner of FAN1, FANCD2-ID was studied and shown to resolve ICLs with catalysis of most incisions taking place within 2 hrs and peaking between 50-90 minutes after ICL induction (Klein Douwel et al., 2014; Knipscheer et al., 2009). The pattern seen in benznidazole treated lines is similar to that obtained with benznidazole treated mNeonGreen-TbMRE11 (Figure 3.27C). These results suggest possible interaction of TbFAN1 and TbMRE11. Recent work from my group pointed to possible epistatic interaction between TbFAN1 and TbMRE11. FAN1 mediated ICL repair leads to DSBs, the repair of which require downstream processing by the HR mediated recombinational repair requiring MRE11 processing and RAD51 recombination. (MacKay, 2010, Ducy 2019, Ronato 2020). A similar comparison using TbSNM1, indicated non-epistatic relationship was seen (Kumar, 2019). These observations are not unique to trypanosomes but also exist in other eukaryotes, where FAN1 and SNM1 contribute to ICL repair in a non-epistatic manner in mice and yeasts (Fontebasso, 2013).

TbSNM1 is a canonical ICL repair enzyme, involved in resolution of ICLs induced by DNA crosslinking agents (Dattani & Wilkinson, 2019; Sullivan et al., 2015). In line with these studies, we assessed the temporal recruitment of TbSNM1 upon exposure to various exogenous agents (mechlorethamine, nifurtimox or benznidazole). Addition of these agents resulted in temporal expression of TbSNM1 in response to the treatment Figure 4.15). Upon ICL induction by mechlorethamine, significant increase in the fluorescence intensity occurred in the first hour of treatment with no further change seen until the final time point (6-hours post treatment) (Figure 4.15A). In the case of nitroheterocycles benznidazole and nifurtimox significant increase in TbSNM1 expression was noted by 2-hour and1st hour of treatment respectively with a peak in

expression seen at 3-hours post treatment and declining at 4- and 6-hours post-treatment (Figure 4.15B & C). Benznidazole and nifurtimox cause genome wide mutations that is seen here to mobilize various DNA pathways including ICL repair factors, HR and NER pathways.

7.2.1.3. BER

Base lesions are highly mutagenic and require rapid removal by BER (David et al., 2007). Base excision repair catalyses the repair of base lesions resulting from base alkylation, deamination, hydrolysis, oxidation induced by cellular metabolism or exogenous treatments (Caldecott, 2020; Demple & Harrison, 1994; Imlay, 2003; Kim & Wilson III, 2012). The mobilisation of BER enzymes has been reconstituted in various organisms including *E. coli* (Dianov et al., 1992; Dianov & Lindahl, 1994), yeast (Wang et al., 1993), humans (Dianov et al., 1992) using purified and crude cell extracts. DNA glycosylases constantly scan the genome for base lesions and rapidly excise them. The initial steps of BER involve coordinated removal of modified bases via hydrolytic cleavage of the N-Glycosidic bond by damage specific DNA glycosylases and eventual removal of the AP site by AP endonucleases (Luncsford et al., 2013). To successfully carry out their roles in BER, DNA glycosylases (e.g. MYH) and AP endonucleases (e.g. APE1) physically interact with each other via inter-domain regions (Luncsford et al., 2013), This interaction causes the recruitment of other enzymes at this point for repair synthesis of the missing base(s) by DNA repair specific polymerases followed by ligation of the repaired strand by DNA ligase I or III, resulting in complete damage repair (Caldecott, 2020).

In addition to maintaining genome integrity, base excision repair is important in responding to DNA damage induced by exogenous treatment including cancer therapies. Enzymes from this pathway are important in modulating cellular response anticancer agents and their expression determine cancer prognosis (Lam et al., 2006a; Lu et al., 2021).

In this project we demonstrated sequential recruitment of TbMYH and TbAPE1(Dianov & Lindahl, 1994; Korolev, 2005; Hegde et al., 2008; Caldecott, 2020; Koczor et al., 2021) to exogenous DNA damage induced by mechlorethamine and UV light as well as by trypanocidal nitroheterocycle prodrugs (nifurtimox; benznidazole; fexinidazole). In all cases, rapid mobilisation of MYH occurred in the first hour of treatment followed by APE1 with slightly delayed recruitment in response to DNA damage, in line with what has been reported in the sequential mobilisation of BER proteins. The repair speed reported when studying crude or purified cell extracts and naked DNA above, was faster than those observed in this study. This is probably due to the complications

of dealing with chromatin and repair in live cells compared working directly with purified DNA (Kennedy et al., 2018).

7.2.1.4. NER

XPF as part of the XPF-ERCC1 heterodimer are involved in NER, HR and ICL repair. They are recruited to ICL by ubiquitinated FANCD2-ID complex and together with SLX4/FANCP to unhook the ICL by 3' -5' incision of the lesion (Klein Douwel et al., 2014), In NER, XPF-ERCC1 along with XPG are involved by 3' -5' incision (Mu et al.,1996, Tapias et al., 2004). CSB, recognises the stalled RNA Pol and mobilises repair factors TFIIH (Sarker et al., 2005).and modulates several transcription-coupled DNA repair networks including TC-NER, HR, BER and ICL) repair (Batenburg et al., 2015; Wei et al., 2016; Menoni et al., 2012, 2018; Stevnsner et al., 2008). NER repair involves sequential interaction of CSB and XPF with DNA in response to damage, with CSB being mobilised before XPF. This temporal expression was monitored with respect to exogenous DNA damage on mobilisation of TbCSB and TbXPF respectively and in all cases, temporal expression of both NER factors involve TbCSB mobilised before TbXPF(Tiwari et al., 2021; Volker et al., 2001).

ICL induction in actively transcribed genes causes stalling of RNA Pol, leading to the recruitment of TbCSB and other NER factors(Sarker et al., 2005, Dattani & Wilkinson 2019).Treatment with exogenous agents resulted in mobilisation of TbCSB to the DNA damage followed by TbXPF, in line with what has been previously seen in actively transcribed euchromatin, where CSB is rapidly recruited to ICLS(IIyama et al., 2015), before other factors such as XPF or XPG (Abdullah et al., 2017; Klein Douwel et al., 2014; le Page et al., 2000; Miura et al., 1996; Mocquet et al., 2008; Sarker et al., 2005). ICL induction by mechlorethamine, resulted in increased expression of TbCSB and TbXPF, with both at elevated levels, reaching a peak 3-hours post treatment before plateauing and remaining at a above the level of untreated cells.

In addition to the role of CSB in resolving other types of DNA damage, several studies have noted its contribution in repairing oxidative stress induced by UV or IR by facilitating transcription restart (Bradsher et al., 2002; Dattani & Wilkinson, 2019; Machado et al., 2014; Nakazawa et al., 2020; Ogi et al., 2010; Rajão et al., 2014; Sarker et al., 2005). In *T. brucei*, predominant pathway for treating UV induced DNA damage is via TC-NER (Machado et al., 2014). and this requires the contribution of TbCSB. XPF expression occurs later and is involved in 3'-5'incision of the DNA lesion (Bowles et al., 2012; Klein Douwel et al., 2017; Miura et al., 1996). Similar to what

was previously noted for ICLS, UV irradiation mNeonGreen tagged TbCSB resulted in rapid increase in TbCSB expression(Herv´ et al., 2018; Menoni et al., 2012) reaching maximum expression after 4hrs and thereafter declining. For TbXPF rapid increase was seen in the first hour remaining elevated throughout the period studied. Delayed expression of XPF occurs after DNA recruitment. Here we saw that the TC-NER pathway is mobilised in response to UV damage, with similar temporal changes also occurring in benznidazole and fexinidazole treated TbCSB and TbXPF. However, for TbXPF, discrepancies with the results means these need to be interpreted with caution.

7.3. RNAi of GG-NER

GG-NER is important in the maintenance of *T. brucei* genome integrity. GG-NER is the global sensor for helix distorting lesions in silent or non-transcribed regions of the genome. XPC-RAD23 dimer and DDB recognise varied substrates for NER repair. Localisation of XPC lead to recruitment of other effectors such as TFIIH, XPA, RPA, XPG, and ERCC1-XPF, resulting in dual incision of the DNA lesion and eventual DNA synthesis. (Aboussekhra et al. 1995; Mu et al. 1995; Araujo et al. 2000, Sugasawa et al. 2001). Here we studied the effect of RNAi induction on the contribution of key GG-NER enzymes on DNA repair.

RNAi of key NER factors have been previously studied (Machado et al 2014), initiation of RNAi of GG-NER initiation factors XPC and TbDDB1 were lethal. Contrary to what was reported in the above study, RNAi of TbDDB1 in our system did not affect growth, whereas depletion of XPC and RAD23 showed slight growth impairments. These differences in induction levels may have been caused by the two expression systems employed in the two studies, stem-loop versus head-to-head promoter used here. Our system may not have produced substantial induction to notice the effects seen in the stem-loop system. Similar to what is reported here, null mutant RAD4(XPC), RAD23, RAD1(XPF) and RAD10(ERCC1 are viable in yeasts. Additionally, in higher organisms such as A thaliana, deletion of XPF/ERCC1 does not affect survival (Giaeveret al., 2002). Deletion of XPC / RAD23 causes maternal sterility in C. elegans but does not affect overall survival (Kamathet al., 2003).

Among the key functions of GG-NER in *T. brucei* is sensing and initiating ICL repair (Machado et al 2014). ICL induction by mechlorethamine results in minor distortions that recruits GG-NER factors XPC/RAD23 to initiate repair, and thereafter mobilising other NER repairsomes to load to the DNA (Dattani and Wilkinson 2019). RNAi of Tb*rad23* or Tb*xpc* were 1.5-fold more

susceptible to mechlorethamine. For Tbddb1, Tb*xpf* or Tb*xpg* RNAi induction did not result in increased sensitivity to ICLs. Contrary to what was seen with Tb*xpf*, RNAi of its repair partner, ERCC1 induced elevated levels of the Tb*ercc1* mRNA with (about 2-fold) more susceptibility to mechlorethamine. The results presented here indicate GG-NER factors play a key role in ICL repair (Machado et al 2014).

7.4. Disruption of Key BER enzymes

The genome of eukaryotes encodes several DNA glycosylases whose roles overlap in recognising and excision of base lesions. Consequently, mutation of genes encoding DNA glycosylases in many cases do not affect cell viability (Parsons & Elder, 2003; Friedberg, 2006; Jacobs & Schär, 2012) in the absence of genotoxic challenge but may lead to genome wide accumulation of base lesions (Castillo-Acosta et al., 2012; Russo et al., 2004). However double mutation of overlapping enzymes leads to hypersensitivity to such DNA damage. In mice, for example, deletion of Mutyh alone, does not lead to tumorigenesis, but those expressing double knockout of *mutyh* and *ogg1* genes were 66 % more likely to develop tumours than wild type (Xie et al., 2004). Additionally, point mutations in humans *myh* gene changing Tyr165 residue to cysteine(Oishi et al., 2002; Sieber et al., 2003) has also been linked to the development of colorectal cancer. Similarly, deletion of genes that code other key BER enzymes such as Polβ (Sobol et al., 1996), APE1 (Ludwig et al., 1998; Xanthoudakis et al., 1996), FEN1 (Kucherlapati et al., 2002) and XRCC1 (Tebbs et al., 1999) are often lethal with inhibition of these activities increasing to carcinogenesis (Kucherlapati et al., 2002; Ludwig get al., 1998; Xanthoudakis et al., 1996).

Our work and that of others(Charret et al., 2012) demonstrate that AP endonuclease activity is non-essential in *T. brucei.* In contrast we could not delete *Tbmyh or Tbnth1.* All putative null mutant lines of *Tbmyh or Tbnth1* maintained a copy of the endogenous gene, possibly because both DNA glycosylases may be essential for survival and growth in PCF *T. brucei.*

Contrary to previous reports in *T. brucei*(Charret et al., 2012) or humans (Walker et al., 1994). Depletion of *Tb*APE1 did not affect parasite survival in double strand break and 3′-PG termini inducing agent, phleomycin. Instead APE1 deficiency induced increased resistance to this agent. In agreement with Charret and colleagues, depletion of TbAPE1 did not affect survival in MMS. This is probably because alternative repair pathways may exist to resolve such lesions, for example

repair using *Tb*snm1/*Tb*mre11 dependent ICL repair pathways to remove such lesions (Dattani & Wilkinson, 2019).

Whereas Tb*myhΔ* lines behaved like wild type with regards MMS, hydroxyurea, mechlorethamine and UV light treatment. Disruption of *Tbmyh* rendered parasite lines to be more sensitive to oxidative damage and strand break induced by H_2O_2 and phleomycin than controls. Tb*myhΔ* lines were slightly more sensitive to H_2O_2 and phleomycin than wild type. The low sensitivity to DNA damage may be due to damage tolerance caused by overlapping repair by DNA glycosylases such as OGG1 (Russo et al., 2004). Similarly, the phenotype displayed by the Tb*myhΔ* "null" toward phleomycin may imply H_2O_2 formation and resultant base damage are phleomycin cytotoxic mechanism of action (Sleigh, 1976).

TbNTH1 were as equally sensitive to MMS, mechlorethamine, UV light and H_2O_2 as wild type displaying 1.5-fold and 5.1-fold resistance to hydroxyurea or phleomycin, respectively. Similar to what is reported here, depletion of nth or Endonuclease VIII alone in *E. coli* does not automatically cause increased sensitivity to H2O2 or gamma-radiation (Cunningham & Weiss, 1985) but mutants lacking both enzymes become hypersensitive to hydrogen peroxide and ionizing radiation (Jiang et al., 1997; Saito et al., 1997). The resistance phenotype of *Tbnth1* is suggestive of possible *Tb*nth1 function in pathways outside of its role in BER or possibly involving its AP lyase activity like seen in *T. cruzi* orthologue, which lacks DNA glycosylase activity but functions only as an AP lyase (Ormeño et al., 2016).

In all cases, disruption of BER factors here did not have any effect in on parasite survival or growth in response to UV irradiation. This is probably because repair of UV induced lesions are predominantly carried out by factors from the *T. brucei* TC-NER.(Machado et al., 2014).

7.5. Disruption of TbSNM1(ICL)

Disruption of SNM1 in yeasts and *C elegans* lead to increased sensitivity to DNA cross-linking agents such as Mitomycin C, nitrogen mustards etc. In mice null mutants of *snm1* gene express phenotypes such as infertility, increased tumorigenesis, and sensitivity to mitomycin C. (Dronkert et al., 2000; Hemphill et al., 2007; Baddock et al., 2020). Loss of *T. brucei* SNM 1 mirrors that of yeast and *C. elegans*, above. TbSNM1 depletion results in increased sensitivity to DNA damage including nitrogen mustards but does not result in growth defects or cell cycle stalling (Dattani and Wilkinson 2019; Sullivan et al., 2015), Similarly, in this project, loss of TbSNM1 did not cause growth defects. The lack of growth defects with respect to SNM1 is probably because SNM 1's

role in ICL repair overlaps with nucleases such as FAN1 and can partly complement this protein as seen in mice (Thongthip et al., 2016)

Depletion of SNM1 resulted in increased sensitivity to ICL inducing compound mechlorethamine but not UV (Dattani and Wilkinson 2019; Sullivan et al., 2015). Contrary to what was previously reported by Sullivan and colleagues (2015) with MMS treatment of BSF Tb*snm1Δ*, treatment of PCF Tb*snm1Δ* resulted in 1.8-fold increased susceptibility with respect to wild type. This variability in DNA repair may possibly be because DNA damage induced by MMS may be regulated differentially in PCF *T. brucei* (Vieira-da-Rocha et al 2019; Dattani et al 2021). Surprisingly, phleomycin treatment caused 3-fold more resistance, possibly suggesting that TbSNM1 may be important in regulating trypanosomal DNA repair factors associated with phleomycin repair. Like BSF Tb*snm1Δ* (Sullivan et al 2015), loss of SNM in PCF parasite did not result in increased susceptibility to nifurtimox and benznidazole. However, time course experiments contrast these results as treatment with both benznidazole and nifurtimox caused temporal mobilisation of TbSNM1. This type of contradiction was also noted in mutant HsSNM1, which did not show growth defects in the presence of IR damage but is mobilised by IR to form repair foci (Riche et al.,2002) and colocalised with ATM in the presence of IR (Akhter & Legerski, 2008).

7.6. Informatic analysis of *T.brucei* DNA repairsome

Computational prediction of the 3-dimensional structures and function were evaluated and described in chapters 3-6, for all proteins studied in this project. For each protein analysed, we identified important residues predicted to be involved in metal ion binding, Protein::Protein interaction including dimerization, polymerization, oligomerization, or enzyme catalysis.

Mutation of key residues often affect enzyme activity, for example mutation of the asparagine residue (N287) in TbAPE1's GD(L/V) N motif has affects AP endonuclease activity (Charret et al., 2012).

The structural information generated is instrumental in understanding how DNA repair enzymes interact with DNA as well as DNA damaging drugs or drug products. These structures could also be harnessed to perform drug-enzyme docking to understand the interaction between studied DNA repair enzymes and DNA damaging agents. Enzyme-ligand docking simulations can also be used to identify lead compounds, understand the effect of loss of key residues on enzyme activity, drug interaction with

enzymes or drug-enzyme inhibition prior to performing wet lab experiments (Sliwoski, et al., 2014; Durrant et al., 2010)

Using PSORTII and WoLFPSORT algorithms (Nakai, K., & Kanehisa 1992), we also predicted the localisation of each protein using the "four pattern" or "seven pattern" sequences postulated to function as nuclear localisation signals (NLSs) (Robbin et al., 1991). In most cases the predicted localisation matched the results generated. However, in some instances the computationally predicted localisation signal differed from what we obtained or what has been previously reported. This was the case for TbRAD51, a nuclear protein in *T. brucei* and other eukaryotes was predicted to localize in the cytoplasm. However, our results confirmed what others have previously reported, i.e., RAD51 is a nuclear protein (Glover and Horn 2012). Similarly, tagging both ends of TbERCC1 or C-terminal of TbRAD23 and TbDDB1 did not result in nuclear localisation.

Additionally, tagging the N-terminal of TbCSB affected its localisation in the nucleus whereas tagging the C-terminal had no effect on protein localisation. Tagging also affected the localisation of *T. brucei* BER protein TbMYH. Despite prediction of nuclear localisation, tagged TbMYH localized in the cytoplasm in procyclic form parasites, possibly because MYH is regulated differently in the two lifecycle stages or alternatively the tagged protein may have been targeted for proteasomal degradation in the cytoplasm. Table 7.2 gives an overview of key features from the different proteins studied here.

.

Protein	Key features	Ligand	Tagging affects localisation
TbMRE11	Nuclease and DNA binding domains Metal binding residues Loop residues for Protein: Protein interaction	Mn^{2+}	No
TbRAD51	Walker A & B Polymerisation motif SS-DNA binding loop region	Mg^{2+}	No
TbSNM1	N-terminal Lactamase_B_2 domain DNA repair metallo-β-lactamase domain (DRMBL, PF07522) HxHxDH motif that co-ordinate metal (zinc) ion binding 4 pattern" RRRH" nuclear localisation signal	Zn^{2+}	No
TbFAN1	tetratricopeptide Repeat_2 domains (TPR_2; PF07719) that mediate protein: protein interactions VRR-NUC domain with characteristic PD-(D/E) XK, motif, involved in divalent metal binding	Ca^{2+}	No
TbAPE1	GD(L/V) N motif, function in metal ion binding/AP endonuclease activity	Mg^{2+}, Mn^{2+}	No
TbMYH	Catalytic HhH-GPD domain and 4Fe-4S binding region Interdomain putative DNA binding site at interface between the HhH-GPD and NUDIX domains	4Fe-4S	Yes,
TbNTH1	HhH-GPD domain HhH DNA-binding motif C-terminal 4Fe-4S binding site (C209; C216; C219; C225	4Fe-4S	No
TbCSB	SNF2-related domain (PF00176,) with characteristic DExD/H box Conserved helicase C-terminal domain (PF00271) ATP and DNA binding motif- (I, IA, II, III IV, V and VI) nuclear localisation signal	Mg^{2+}, Zn^{2+}	Yes
TbDDB1	3 β-propeller domains C-terminal domain	Zn^{2+}, Cu^{2+}	Yes
TbERCC1	N-terminal ERCC1/RAD10 DNA binding (PF03834) nuclease-like domain (NLD) C-terminal Helix-hairpin-helix motif (PF14520)		Yes

TbRAD23	N-terminal ubiquitin-like domain (UBL; PF00240) ubiquitin-associated domains (UBA-1 & UBA-2; PF00627) XPC binding domain (PF09280)		Yes
TbXPC	Rad4 transglutaminase-like domain (TGD) (PF03835) Rad4 beta-hairpin domains, BHD1-3 involved in DNA binding TbXPC's TGD domain interacts with RAD23 XPC binding domain		No
TbXPF	-helicase-like module (HLM) containing -RecA-like domains, RECA1 and RECA2 catalytic module (CM) divided into: -nuclease domain (NUC; PF02732; with GDXnERKX3D motif, -helix-hairpin-helix motif (HHH).	Mg^{2+}, Mn^{2+}	No
TbXPG	XPG N-terminal domain (XPG_N; PF00752) XPG I-region/containing domain (XPG_I; PF00867 & PF12813) carboxyl terminal β-hairpin motif helix-2turn-helix domain motif hydrophobic wedge facilitate putative DNA binding	Mg^{2+}, Mn^{2+}	No

Table 7.2: Key features of the different proteins studied in this project. A summary of key structural domains or motifs found on each protein and catalytic metal co-factors or ligands that bind to the active site for various enzymes. Information on the effect of tagging on protein localisation is also given.

CHAPTER 8: Thesis Summary and Future Work

Several studies have demonstrated that benznidazole and nifurtimox function as prodrugs with their activation involving a parasite specific type I nitroreductase (Wilkinson et al., 2008; Wilkinson & Kelly, 2009, Baker *et al.,* 2011; Hall et al., 2011; Hall & Wilkinson, 2012; Thomas et al., 2018). Studies fexinidazole's mode of the action also suggest that one mechanism of fexinidazole activation involves NTR (Kaiser et al., 2011; Tweats et al., 2012). In this project we confirmed that fexinidazole, activation is dependent on parasite specific NTR1. All three drugs once activated target nuclear genome as confirmed by comet and salmonella mutagenesis assays (Ferreira & Ferreira, 1986, Buschini et al., 2009). These mutations then trigger DNA repair cascades from the various DNA repair pathway that are activated to resolve the damage. Here we showed sequential mobilisation of *T. brucei* DNA repair pathways to exogenous DNA damaging agents including that caused by current therapies. In the future, it will be worthwhile to study the protein: protein interactions that drive these responses. Additionally, it will be interesting to decipher the cell cycle stage(s) associated with each treatment and pathway studied using flow cytometry (Marin 2019).

Finally, DNA repair pathways are important in maintaining genome integrity and cell survival(Rahimian et al., 2020). Several chemotherapies target these DNA repair pathways. However, with constant exposure to these drugs, parasites as well as tumours counteract the effects of drug pressure by over-expressing DNA repair genes leading to resistance to such chemotherapies as seen in cancer and other microbial therapies(de Almeida et al., 2021; Heinen et al., 2002). To overcome the rising resistance to DNA damaging drugs, development of molecular inhibitors targeting these repairsomes has been on the rise[(Arora et al., 2016; Kelm et al., 2022; Sadoughi et al., 2021)], these include those targeting SNM1 (Buzon et al., 2021; Yosaatmadja et al., 2021)., APE1(Laev et al., 2017), XPF-ERCC1(Arora et al., 2016) etc. Molecular inhibitors are used as adjuvants to current chemotherapies to suppress the activation of DNA repair pathways thus potentially making these drugs more effective. For example, cephalosporins, cefotaxime, ceftriaxone and hydroxamic acids target the nuclease activity of hSNM1A. However, no significant inhibition was seen using these compounds with TbSNM1 (Kumar 2019).

Since DNA repair pathways are conserved across all domains of life, development of molecular inhibitors of DNA repair in *T. brucei* will have larger implications including utilisation in cancer and other microbial therapies that effect cytotoxicity by inhibiting DNA repair pathways.

8.1. Future work

To follow on the work presented in this thesis it will interesting to unravel the following in *T. brucei* BSF or PCF.

1. The cell cycle stage associated with each exogenous treatment and activated pathway.
2. The Protein: Protein interaction that drive these DNA repair responses associated with each treatment using pull down experiment.
3. Develop small molecular inhibitors based on Protein::Protein interaction targeting DNA repair pathways or key proteins to improve therapeutic activity in *T. brucei.*
4. Functional studies using null mutant lines to understand which of these factors are essential for each exogenous treatment studied.

8.2. Thesis summary

8.2.1 A globalized DNA repair reaction in response to trypanocidal nitroheterocyclic prodrugs, UV irradiation and ICL:

a. *T. brucei* DNA repair pathway respond to DNA damage by initially upregulating all factors. This is manifested by temporal mobilisation of DNA repair proteins from various pathways. Here in response to DNA damage sequential mobilisation to DNA damage was seen with:
    i. TbMRE11 and TbRAD51
    ii. TbMYH and TbAPE1
    iii. TbCSB and TbXPF
b. Temporal mobilisation of DNA repair factors also revealed interacting partners for example with benznidazole treatment of lines expressing mNG-TbMRE11 and mNG-TbRAD51 showed that repair of DNA lesions induced by benznidazole involved mobilization and interaction of TbMRE11 TbRAD51 and γH2A response to possible DSB induced. These interactions can be further validated by western blot and immunoprecipitation.

8.2.2 Using gene deletion, it is possible to differentiate which DNA repair enzymes are important for the repair pathway studied. For example:

c. TbSNM1 does not seem to hydroxyurea or hydrogen peroxide but is important in resolving mechlorethamine, and MMS induced lesions.

d. *T. brucei* lacking BER protein, TbMYH were more sensitive to H_2O_2 and phleomycin than wild type, indicating possible involvement in repairing the lesion generated by this drug.

8.2.3 Trypanocidal activity of fexinidazole occurs via activation by parasite specific TbNTR1.

e. Fexinidazole damage mobilises *Tb*CSB and *Tb*MRE11 with both proteins involved in an epistatic interaction to resolve this damage.

9. Appendix

Appendix 1: Primer sequences for CRISPR/Cas9 based tagging

Gene	Primer combination	Sequence
mNG-Tbcsb	*Tbcsb-UFP1* *TbcsbDRF1* *5sgRNA-1 Tbcsb*	*TATTATTATTTATTATTTTGCGCTAGTCCAgtataatgcagacctgctgc* *CACACCCAACTG GGAGAGTTCATCTGCCATactacccgatcctgatccag* *gaaattaatacgactcactataggCAAAAAAGCGGAAGCAAACTgttttagagctagaaatagc*
Tbcsb-mNG	*Tbcsb-DFP1* *Tbcsb-DRTAGP* *3sgRNA Tbcsb*	*TTTGCGAGGCTTCCCAGGGTGGAGGGATCGggttctggtagtggttccgg* *CGAAACGCACGCACGGATGCGAGCAGGGCAccaatttgagagacctgtgc* *gaaattaatacgactcactataggAACTTCCTCTTTAAAAAAAGgttttagagctagaaatagc*
mNG-bmre11	*Tbmre11-UFP1* *Tbmre11DRF1* *5sgRNA-1 Tbmre11*	*TTAACGGCTTGCTTCTACTAATTTTTTCCCgtataatgcagacctgctgc* *ATG CGA CTG TGTGGATGCCCTCTCGGCCATactacccgatcctgatccag* *gaaattaatacgactcactataggCGACTCCCCATTTCAGAAGggttttagagctagaaatagc*
Tbmre11-mNG	*Tbmre11-DFP1* *Tbmre11-DRTAGP* *3sgRNA Tbmre11*	*ACAAGTAGTTCTGGCATGCCAGATAACTATggttctggtagtggttccgg* *TGTAACTGTCGAGAGGTACGTGTCTGTCCAccaatttgagagacctgtgc* *gaaattaatacgactcac tataggTTTTTTTAATTTGTACTTTGgttttagagctagaaatagc*
mNG-Tbrad51	*Tbrad51-UFP1* *Tbrad51DRF1* *5sgRNA-1 Tbrad51*	*ACAACACCGAAAGGGCTTCTCCAGATTCCCgtataatgcagacctgctgc* *GCGTTTCTTATTTTTGGTGCGAGTGTTCATactacccgatcctgatccag* *gaaattaatacgactcactataggCGCGTTAAGAAACCTTCACGgttttagagctagaaatagc*
Tbrad51-mNG	*Tbrad51-DFP1* *Tbrad51-DRTAGP* *3sgRNA Tbrad51*	*TATGAGAACGGTGTGGGAGACGTTAGGGACggttctggtagtggttccgg* *TCCGCTTTCGGAGTGAATGTTTTTCTACAAccaatttgagagacctgtgc* *gaaattaatacgactcactataggATGACGAGGAAAAAAAAGATgttttagagctagaaatagc*
mNG-Tbsnm1	*Tbsnm1-UFP1* *Tbsnm1DRF1* *5sgRNA-1 Tbsnm1*	*TGCATTTGTTTCAAAGGATAGGAGGGGCCGgtataatgcagacctgctgc* *AGTCACCTTACCTGCAGCTCCACCTGCCATactacccgatcctgatccag* *gaaattaatacgactcactataggCGAGGTTGATGACACCTTCTgttttagagctagaaatagc*
mNG-Tbape1	*Tbape1-UFP1* *Tbape1-URP1* *5'sgRNA-Tbape1*	*TGTCTATATCGCTTTCATTCCGATGTTCCTgtataatgcagacctgctgc* *TGGTAGCTTGGATAGTTTTTTCGGTGGCATactacccgatcctgatccag* *gaaattaatacgactcactataggCTTTGGATTCTACAGAGGCAgttttagagctagaaatagc*
Tbape1-mNG	*TbAPE1-DFP1* *TbAPE1-DFP1* *3'sgRNA-Tbape1*	*CACTGCCCTGTGCAGATGTGGCTCCGTAAGggttctggtagtggttccgg* *TGTGCACACCTTTCCTCAAAATTTTCACCTccaatttgagagacctgtgc* *gaaattaatacgactcactataggGCGCAGCTGCGGGTTAAGCAgttttagagctagaaatc*
Tbddb1-mNG	*Tbddb1-DFP1* *Tbddb1-DRTAGP1* *3' sgRNA-Tbddb1*	*GCGTTTATACAAACCATCCAAAGAATTCACggttctggtagtggttccgg* *CAC CTATCCAACCGCTGTGACCCACACCCGccaatttgagagacctgtgc* *gaaattaatcgactcactataggTACTCGTGTGCCGATGTGTGgttttagagctagaaatagc*
mNG-Tbnth1	*Tbnth1-UFP1* *Tbnth1-URP1* *5'sgRNATbnth1*	*TGCATATTTATATGTTTCCTTTCTCTTATCgtataatgcagacctgctgc* *GTTGGATGGTGGTTTAAATAGCTTGCCCATactacccgatcctgatccag* *gaaattaatacgactcactataggCGTTTACTTTAGCTGGGCAGgttttagagctagaaatagc*
Tbnth1-mNG	*Tbnth1-DFP1*	*GTGGTTCCAGCCCCAAAACGGAAGCGTATAggttctggtagtggttccgg* *CAACAACAAAGAAAATGACATAATAGTACAccaatttgagagacctgtgc*

	Tbnth1-DRTAGP1	*gaaattaatacgactcactataggAGTGGTAAAGAGATACAGAGgttttagagctagaaatagc*
	3'sgRNA-Tbnth1	
mNG-Tbmyh	*Tbmyh-UFP1*	*AGTCTTGTTTCACTTCCCCCCGCGCGGCCTgtataatgcagacctgctgc*
	Tbmyh-URP1	*GTTCAGTGAAAATTGAAGGGGAAAGTTCATactacccgatcctgatccag*
	5'sgRNA-Tbmyh	*gaaattaatacgactcactataggTTGTGGTCAATGACAGAGCGgttttagagctagaaatagc*
Tbmyh-mNG	*Tbmyh-DFP1*	*TCAACGAGGAAACGACCTAGATTGGCCGGAggttctggtagtggttccgg*
	Tbmyh-DRTAGP1	*ACCGCCGCTCCGACATATCCTCCCACAGAAccaatttgagagacctgtgc*
	3'sgRNA-Tbmyh1	*gaaattaatacgactcactataggGCCGGTTCGGCCAAATTGATgttttagagctagaaatagc*
mNG-Tbxpf	*Tbxpf-UFP1*	*CACGTACTGTGAGTAGTCTCATTACTACCAgtataatgcagacctgctgc*
	Tbxpf-URP1	*TGCAATCCGGCTGCTAACAGTCTGTGGCATactacccgatcctgatccag*
	5'sgRNA-Tbxpf	*gaaattaatacgactcactataggTAAGGAATGGGGAGAAAGTGgttttagagctagaaatagc*
mNG-Tbercc1	*Tbercc1-UFP1*	*ATAACAAAAACACTTATAAATGCTAACGTgtataatgcagacctgctgc*
	Tbercc1-URP1	*TTCACCACGCCCCGTGGGAGGGGAGGCATactacccgatcctgatccag*
	5'sgRNA-Tbercc1	*gaaattaatacgactcactataggTCATTTACAGTGCGTgttttgttttagagctagaaatagc*
mNG-Tb xpg	*Tbxpg-UFP1*	*ATAACAAAAACACTTATAAATGCTAACgtgtataatgcagacctgctgc*
	Tbxpg-URP1	*TTCACCACGCCCCGTGGGAGGGGAGGCATactacccgatcctgatccag*
	5'sgRNA-Tbxpg	*gaaattaatacgactcactataggTCATTTACAGTGCGTGTTTTgttttagagctagaaatagc*
Tbercc1-mNG	*Tbercc1-DFP1*	*GAAGATGAAGATGTGCAGGAAGGTGGCCAggttctggtagtggttccgg*
	Tbercc1-DRTAGP1	*ACTGTATTTAGACTCGAATACCACCACCCAccaatttgagagacctgtgc*
	3sgRNA-Tbercc1	*gaaattaatacgactcactataggGAGGCAAGCTCAGGTATAGAgttttagagctagaaatagc*
Tbxpc-mNG	*Tbxpc-DFP1*	*GGAGCTACACATGGACATCTTTCCTCTCACggttctggtagtggttccgg*
	Tbxpc-DRTAGP1	*GTTGATGAAGATATGTGGAACTGTGCCCCTccaatttgagagacctgtgc*
	3' sgRNA-Tbxpc	*gaaattaatacgactcactataggTCCCTTAGCGAATAAAAGTGgttttagagctagaaatagc*
Tbrad23-mNG	*Tbrad23-DFP1*	*CTTCTTTTCAACTTTCTCGGGATTACCGACggttctggtagtggttccgg*
	Tbrad23-DRTAGP1	*GCC AAG CCT CGA GCC GCG GAG CAC TTT CACccaatttgagagacctgtgc*
	3sgRNA-Tbrad23	*gaaattaatacgactcactataggAACAACCAACAAAATGGACGgttttagagctagaaatagc*
Tbxpf-mNG	*Tbxpf-DFP1*	*CATGGGGAATCTGTGCCACCTACTCAGACA ggttctggtagtggttccgg*
	Tbxpf-DRTAGP1	*CCAGGACACAGCCAGGACGCTTCGGTTCCT ccaatttgagagacctgtgc*
	3' sgRNA-Tbxpf	*gaaattaatacgactcactataggCAATATGTAGCAGTTGCTGGgttttagagctagaaatagc*
Tbxpg-mNG	*Tbxpg-DFP1*	*TCCTATTTGCGTGCAGCCAGGGGTGATCCGggttctggtagtggttccgg*
	Tbxpg-DRTAGP1	*CTTTATATGAGATAGTTTCTCGCATAACCG ccaatttgagagacctgtgc*
	3' sgRNA-Tbxpg	*gaaattaatacgactcactataggTTTTTCTGCGCTTGTGGTGGgttttagagctagaaatagc*

Appendix 2: Primer sequences for validation of CRISPR/Cas9 based tagging

Gene	Primer combination	Sequence
Tbtert	*Tbtert-F* *Tbtert-R*	*AGGAACTGTCACGGAGTTTGC* *GAGCGTGTGACTTCCGAAGG*
Tbcsb-mNG	*Tbcsb-seq1* *mNeonGreen2*	*TGC CAC CAC AAC GTC ACG AAC* *ACA AGA ATC CAC GGG GAA AAC*
mNG-Tbcsb	*Tbcsb-q2* *mNeonGreen1*	*CGCTACCCTCTGTCTCTTCACT* *GGA GTT ATA CCA CAG GAA ACG*
Tb-mre11-mNG	*Tbmre11-K03* *mNeonGreen2*	*AAA GGGCCC TGCGAACTGTTGAGTAAATGC* *ACA AGA ATC CAC GGG GAA AAC*
mNG-Tbmre11	*Tbmre11-K02* *mNeonGreen1*	*GGG TCTAGA TTGTTTCGTAAATGCCACACT* *GGA GTT ATA CCA CAG GAA ACG*
Tbrad51-mNG	*TbRAD51-seq1* *mNeonGreen-2*	*GGT TGT TGC CAA TGT GGA TGG* *ACA AGA ATC CAC GGG GAA AAC*
mNG-Tbrad51	*TbRAD51q2* *mNeonGreen1*	*TGG AAA ACC GTA GCG CGC GTG* *GGA GTT ATA CCA CAG GAA ACG*
mNG-Tbsnm1	*TbSNM1-seq1* *mNeonGreen1*	*CACCAGATCACCATTGCGGAG* *GGA GTT ATA CCA CAG GAA ACG*
mNG-Tbape1	*TbAPE1-seq1* *mNeonGreen1*	*AGC TTC ATT ACC CTT GTC CCC* *GGA GTT ATA CCA CAG GAA ACG*
Tbape1-mNG	*TbAPE1-seq2* *mNeonGreen-2*	*TTA TGA CCG ATA TTT CGC AGG* *ACA AGA ATC CAC GGG GAA AAC*
mNG-Tbnth1	*TbNTH1-seq1* *mNeonGreen1*	*TTG GGC CCA ACA CCT GGA AGG* *GGA GTT ATA CCA CAG GAA ACG*
Tbnth1 mNG	*TbNTH1-seq2* *mNeonGreen-2*	*CAC AAA GGA ACA GTG CCG CGG* *ACA AGA ATC CAC GGG GAA AAC*
mNG-Tbmyh	*TbMYH-seq1* *mNeonGreen-1*	*GGA AAG ACC GCT ATC CAT CGC* *GGA GTT ATA CCA CAG GAA ACG*
Tbmyh-mNG	*TbMYH-seq2* *mNeonGreen-2*	*GGT ACT GAC ATA AAA GGG AGC* *ACA AGA ATC CAC GGG GAA AAC*
Tbdbb1-mNG	*RNAi-Tb DDB1-1* *mNeonGreen-2*	*AAA GGA TCC TGA ATT GCG CGG TGT TAT GGA* *ACA AGA ATC CAC GGG GAA AAC*
Tbrad23-mNG	*RNAi-TbRad23-1* *mNeonGreen-2*	*AAA GGA TCC GAG GGC GCT GTC GGC GAT TCC* *ACA AGA ATC CAC GGG GAA AAC*
Tbxpc-mNG	*RNAi-TbXPC-1* *mNeonGreen-2*	*AAA GGA TCC GCT CCG GAG CCG TTA GGT GAG* *ACA AGA ATC CAC GGG GAA AAC*
Tbxpg-mNG	*TbXPG-KO3* *mNeonGreen-2*	*aaagggcccTGCCATTAACGCGGTTCTACC* *ACA AGA ATC CAC GGG GAA AAC*
mNG-Tbxpg	*TbXPG-KO2* *mNeonGreen-1*	*aaatctagaTTACCACTTCCGGGAGATCGA* *GGA GTT ATA CCA CAG GAA ACG*
mNG-Tbxpf	*TbXPF-seq2* *mNeonGreen-1*	*CAC GCG CAA CGC GGC GAT GCA* *GGA GTT ATA CCA CAG GAA ACG*
Tbxpf-mNG	*TbXPF-seq1* *mNeonGreen-2*	*TCT TCA CTT CCT TCC CGC GGG* *ACA AGA ATC CAC GGG GAA AAC*
mNG-Tbercc1	*TbERCC1-seq2* *mNeonGreen-1*	*AAA TAG GCA GCA CAT TCC TCA* *GGA GTT ATA CCA CAG GAA ACG*
Tbercc1-mNG	*TbERCC1-seq1*	*ATA GTA ACG TTG CTA CCG CGG* *ACA AGA ATC CAC GGG GAA AAC*

	mNeonGreen-2	
mNG-Tbfan1	*TbFAN1-KO2 mNeonGreen1*	*aaatctagaGCCGAAAGATCGCCTTACCAT GGA GTT ATA CCA CAG GAA ACG*
Tbfan1-mNG	*TbFAN1-KO3 mNeonGreen 2*	*TTTTGGACCCTGAGGAGTTTCCT ACA AGA ATC CAC GGG GAA AAC*

Appendix 3: Primers for generation of RNAi vectors

Gene	Primer combination	Sequence
RNAiTbddb1	*TbDDB-RNAi1* *TbDDB-RNAi2*	*aaa gga tcc TGA ATT GCG CGG TGT TAT GGA* *ggg ctc gag TCC TTT CAG AAG GAT AAA TGT*
RNAiTbrad23	*TbRAD23-RNAi1* *TbRAD23-RNAi2*	*aaa gga tcc GAG GGC GCT GTC GGC GAT TCC* *ggg ctc gag TGT GCT GCC GCA TCT TCC CTC*
RNAiTbxpc	*TbXPC-RNAi1* *TbXPC-RNAi2*	*AAA GGA TCC GCT CCG GAG CCG TTA GGT GAG* *AAA CTC GAG CCA AAT ACT GCT GAC GAA TAC*
RNAiTbxpf	*pRNAi-TbXPF-1* *pRNAi-TbXPF-2*	*GGG GGA TCC GAC GAT ACC GAG AAG CAG GG* *ggg ctc gag CGT CGT CAT CGG TGC TAT CA*
RNAiTbxpg	*pRNAi-TbXPG-1* *pRNAi-TbXPG-2*	*aaa gga tcc ACA TCT GTG GTA TCC GGG GA* *ggg ctc gag GCG CCA TGT ACA ATG ACG TC*
RNAiTbercc1	*pRNAi-TbERCC1-1* *pRNAi-TbERCC1-2*	*aaa gga tcc TCC TGT GCG GTG GAA CTT GT* *ggg ctc gag ACG TCA GCT ACT AGG CAG CG*

Appendix 4: Primers for construction of complementation vectors

Gene	*Primer combination*	*Sequence*
Tbsnm1	*TbSNM1-exp5* *mNeonGreen3*	*gggggcgcgccTTATTCTGAGTCACTACTCAG* *AAA CCT GCA GGA TGG GTT CTG GTA GTG GTT CC*
Tbcsb	*TbCSB-EXP1* *mNeonGreen3*	*AAA GGC GCG CCT TAC GAT CCC TCC ACC CTG GG* *AAA CCT GCA GGA TGG GTT CTG GTA GTG GTT CC*
Tbmre11	*TbMRE11-EXP1* *mNeonGreen3*	*AAA GGC GCG CCC TAA TAG TTA TCT GGC ATG CC* *AAA CCT GCA GGA TGG GTT CTG GTA GTG GTT CC*
TbRad51	*TbRAD51-EXP1* *mNeonGreen3*	*AAA GGC GCG CCC TAG TCC CTA ACG TCT CCC AC* *AAA CCT GCA GGA TGG GTT CTG GTA GTG GTT CC*

Appendix 6: Publication

Received: 3 February 2021 | Revised: 17 May 2021 | Accepted: 29 May 2021

DOI: 10.1111/mmi.14763

RESEARCH ARTICLE

WILEY

Unraveling the antitrypanosomal mechanism of benznidazole and related 2-nitroimidazoles: From prodrug activation to DNA damage

Ambika Dattani | Isatou Drammeh | Aishah Mahmood | Mahbubur Rahman | Joanna Szular | Shane R. Wilkinson

School of Biological & Chemical Sciences, Queen Mary University of London, London, UK

Correspondence
Shane R. Wilkinson, School of Biological & Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK.
Email: s.r.wilkinson@qmul.ac.uk

Present address
Ambika Dattani, School of Life Sciences, Queen's Medical Centre, University of Nottingham, Nottingham, UK
Joanna Szular, Małopolska Centre of Biotechnology, Jagiellonian University, Krakow, Poland

Funding information
Queen Mary University of London; Islamic Development Bank Studentships

Abstract

Nitroheterocycles represent an important class of compound used to treat trypanosomiasis. They often function as prodrugs and can undergo type I nitroreductase (NTR1)-mediated activation before promoting their antiparasitic activities although the nature of these downstream effects has yet to be determined. Here, we show that in an NTR1-dependent process, benznidazole promotes DNA damage in the nuclear genome of *Trypanosoma brucei*, providing the first direct link between activation of this prodrug and a downstream trypanocidal mechanism. Phenotypic and protein expression studies revealed that components of the trypanosome's homologous recombination (HR) repair pathway (TbMRE11, γH2A, TbRAD51) cooperate to resolve the benznidazole-induced damage, indicating that the prodrug-induced lesions are most likely double stand DNA breaks, while the sequence/recruitment kinetics of these factors parallels that in other eukaryotes HR systems. When extended to other NTR1-activated 2-nitroimidazoles, some were shown to promote DNA damage. Intriguingly, the lesions induced by these required TbMRE11 and TbCSB activities to fix leading us to postulate that TbCSB may operate in systems other than the transcription-coupled nucleotide excision repair pathway. Understanding how existing trypanosomal drugs work will aid future drug design and help unlock novel reactions/pathways that could be exploited as targets for therapeutic intervention.

KEYWORDS

benznidazole, CRISPR gene editing, DNA repair, fluorescence profiling, gene deletion, prodrug activation

1 | INTRODUCTION

Protozoan parasites belonging to the genus *Trypanosoma* are responsible for a series of medically and veterinary important infections. In humans, *Trypanosoma cruzi* and *Trypanosoma brucei* are the etiological agents of Chagas disease and human African trypanosomiasis (HAT), respectively, while animal African trypanosomiasis is caused by several species including *T. brucei*, *T. congolense*, and *T. vivax*. The burden associated with these insect transmitted infections has had a major impact on the public health and socio-economic development of many poor rural communities in Latin America and sub-Saharan Africa, effectively trapping sufferers and their families in a disease/

Ambika Dattani and Isatou Drammeh contributed equally to this work.

poverty cycle: Estimates indicate that Chagas disease costs the Latin American economy ~US$7.2 billion per annum (Lee et al., 2013), a figure that exceeds the total global burden of uterine, cervical, and oral cancers (~US$6.7, 4.7, and 5.3 billion per annum, respectively) (Lee et al., 2013), while the economic impact of African trypanosomiasis runs into the billions of US$ each year (Shaw et al., 2014). Encouragingly, the implementation of new vector control and disease surveillance strategies alongside use of novel drug formulations has resulted in the decline in prevalence of the human forms of the disease at endemic sites (Franco et al., 2020; WHO, 2015). However, due to congenital and sexual transmission, blood transfusion, and organ transplantation coupled with population migration and leisure/recreational activities, we have started to see the emergence of trypanosomal infections as a public health problem across the globe (Bennett et al., 2018; Migchelsen et al., 2011; Perez-Molina & Molina, 2018; Requena-Mendez et al., 2016; Simarro et al., 2012). With no immediate prospect of a vaccine, drugs represent the only viable option for treating trypanosomal infections.

Nitroheterocycles make up a significant part of the therapies used to combat trypanosomal infections (Patterson & Fairlamb, 2019; Patterson & Wyllie, 2014; Wilkinson et al., 2011; Wilkinson & Kelly, 2009). The 2-nitroimidazole, benznidazole, and the 5-nitrofuran, nifurtimox have been employed for more than 50 years to treat the acute stage of Chagas disease with the latter used as part of the nifurtimox-eflornithine combinational therapy employed against the meningo-encephalic stage of HAT. Use of these compounds is controversial as they are toxic, reportedly carcinogenic, and have limited efficacy (Babokhov et al., 2013; Wilkinson & Kelly, 2009). This results in considerable scope for the emergence of resistance with strains refractory to both heterocycles encountered in natural and laboratory settings. Despite these issues, a range of other nitroaromatic compounds have been evaluated as potential trypanocidal agents with such screens identifying a variety of structures that display significant potency against the parasites while exhibiting little/no toxicity towards mammalian cells (Patterson & Fairlamb, 2019; Patterson & Wyllie, 2014; Wilkinson et al., 2011). Of note is the Drugs for Neglected Diseases initiative sponsored program that has "rediscovered" fexinidazole as a potent anti-*T. brucei* agent (Mesu et al., 2018). The success of clinical trials assessing this 5-nitroimidazole has resulted in it being recently recommended by the European Medicines Agency for use against all stages of West African form of HAT with this approval paving the wave for its registration and distribution in endemic countries (Deeks, 2019).

Nifurtimox, benznidazole, and fexinidazole are prodrugs that must undergo activation to have cytotoxic effects (Patterson & Fairlamb, 2019; Patterson & Wyllie, 2014; Wilkinson et al., 2011; Wilkinson & Kelly, 2009). Functional genomic and biochemical approaches have clearly demonstrated that the antitrypanosomal selectivity of these compounds lies in the parasite's ability to express a type I nitroreductase (NTR1), an enzyme absent in higher eukaryotes (Alsford et al., 2012; Wilkinson et al., 2008). These NADH-dependent, FMN-containing oxidoreductases catalyze two sequential 2-electron reductions of the substrate's nitro group to generate a nitroso and then hydroxylamine forms (Hall et al., 2011; Hall & Wilkinson, 2012). The latter derivatives are unstable and metabolized via adduct-forming intermediaries to highly reactive products that can potentially interact with biological molecules causing cellular damage (Diaz de Toranzo et al., 1998; Hall et al., 2011; Hall & Wilkinson, 2012; Trochine et al., 2014). Extending the above studies to include laboratory selected nifurtimox or benznidazole resistant lines revealed that reduction in the level of type I NTR activity is invariably associated with resistance (Alsford et al., 2012; Campos et al., 2014; Mejia et al., 2012; Mejia-Jaramillo et al., 2011; Wilkinson et al., 2009; Wyllie et al., 2016).

Besides the NTR1 activation mechanism, other flavin-containing trypanosomal oxidoreductases can also metabolize benznidazole and/or nifurtimox. In vitro, enzymes such as dihydrolipoamide dehydrogenase, cytochrome P450 reductase, and trypanothione reductase can mediate an "oxygen sensitive" type II nitroreductase activity, catalyzing the 1-electron reduction of conserved nitro group to form a nitro anion (Blumenstiel et al., 1999; Schoneck et al., 1997; Viode et al., 1999). In an oxygenic environment this radical can be rapidly re-oxidized back to the parental nitro-compound with the concomitant production of superoxide anions, a reactive oxygen species (ROS) that could potentially promote oxidative stress within the parasite (Docampo & Stoppani, 1979; Moreno et al., 1982; Viode et al., 1999). Intriguingly, addition of benznidazole to trypanosomal extracts does not trigger the production of significant levels of ROS (Moreno et al., 1982) although functional studies on SODB1, a superoxide dismutase isoform located in the parasites' cytosol and glycosome, and the formation of 8-oxoguanine in prodrug-treated cells indicate that this does occur in vivo (Prathalingham et al., 2007; Rajao et al., 2014; Temperton et al., 1998). These apparently conflicting findings could indicate that activities which mediate the 1-electron nitroreduction of benznidazole are unstable and maybe readily inactivated during cell lysis or reflect that this 2-nitroimidazole is a poor substrate for type II nitroreductases with superoxide anion formation occurring slowly. In the latter case, components of the unusual oxidative defense network present in trypanosomal extracts could then efficiently remove superoxide anions from samples thereby rapidly reducing this ROS to low, undetectable levels (Dufernez et al., 2006; Mateo et al., 2008; Wilkinson et al., 2006). An additional *T. cruzi* nitroreduction activity that displays features characteristic of both type I and II nitroreductases has been reported and involves the "Old Yellow Enzyme" (TcOYE), a homolog of prostaglandin F2α synthase. In vitro this enzyme can mediate the 2-electron reduction of nifurtimox under anaerobic conditions but not benznidazole (Kubata et al., 2002) although recent functional studies have shown that in vivo this enzyme can in fact metabolize both prodrugs (Diaz-Viraque et al., 2018; Garcia-Huertas et al., 2017).

Based on the mutagenic properties benznidazole displays towards bacterial and trypanosomal cells, DNA appears to constitute a major target for this prodrug (Campos et al., 2017; Ferreira & Ferreira, 1986). This damage may be caused by this nitroimidazole promoting oxidation of nucleotides found in the nucleotide pool that

are subsequently incorporated into the parasite's genomes (Rajao et al., 2014). This can promote a multitude of effects potentially causing base mismatch pairing leading to the formation of point mutations or generation of double-strand DNA breaks (DSBs) (Rajao et al., 2014). Cells across all domains of life possess dedicated DNA repair pathways which combat such mutagenic and structural adducts. Multiple repair pathways fix different forms of damage with errors in base pairing repaired by the mismatch repair (MMR) pathway and breakage events on both strands of the DNA double helix fixed by the homologous recombination (HR) or nonhomologous end joining (NHEJ) pathways. Overexpression or reduction in the levels of MMR or HR enzymes in the trypanosome generates lines that exhibit altered susceptibilities to benznidazole (Gomes Passos Silva et al., 2018; Rajao et al., 2014).

Here, we evaluated the growth inhibitory properties of a series of 2-nitroimidazoles, compounds structurally related to benznidazole, against the bloodstream form (BSF) of *T. brucei*. Using recombinant lines deficient in TbCSB, a DNA helicase that functions in transcription-coupled nucleotide excision repair (TC-NER) (Machado et al., 2014) and/or TbMRE11, a key component of the HR pathway (Robinson et al., 2002; Tan et al., 2002), we show that the trypanocidal activity of several nitroheterocycles, including benznidazole, occurs by promoting DNA lesions with this effect being dependent upon NTR1 activation of the prodrug.

2 | RESULTS

2.1 | Growth inhibitory effects of 2-nitroimidazoles

The susceptibility of BSF *T. brucei* towards a library of 2-nitroimidazole derivatives was assessed (Table 1). An initial screen revealed that 16 of the 39 compounds tested did not affect the growth of trypanosomes at concentrations up to 100 μM. These were not analyzed further. For the remaining compounds, further inhibition studies revealed that 13 agents (including benznidazole) yielded EC_{50} (a concentration of compound that inhibited parasite growth by 50%) values between 30 and 100 μM, 7 generated values between 10 and 30 μM, and 3 displayed EC_{50} values of 1 to 10 μM. Previous work has shown that the anti-*T. brucei* properties exhibited by many nitroheterocycles are dependent on the bacterial-like type I nitroreductase, TbNTR1. To demonstrate whether this extends across to the most potent trypanocidal 2-nitroimidazoles (compounds that yielded EC_{50} values <30 μM and benznidazole), the susceptibilities of BSF trypanosomes with altered levels of the oxidoreductase were investigated (Table 1). In all cases tested here, parasites induced to express elevated levels of TbNTR were more sensitive (between 3- and 18-fold) to the tested compound relative to controls.

To assess whether the most potent trypanocidal nitroimidazoles (EC_{50} value of <30 μM and benznidazole) displayed toxicity to mammalian cells, their growth inhibitory properties towards undifferentiated THP-1 cells were determined (Table 1). Out of the 11 compounds screened, 4, including benznidazole, had no effect on cell growth at concentrations up to 300 μM, 3 yielded EC_{50} values between 100 and 300 μM, and 2 generated EC_{50} values between 30 and 100 μM with significant in vitro toxicity (EC_{50} values <10 μM) observed for the remaining structures. Comparison of the mammalian toxicity and antiparasitic EC_{50} data allowed a crude measure of each agent's selectivity (the Selectivity Index or SI) toward the pathogen (Table 1). For all the 2-nitroimidazoles tested, a preference towards the trypanosome relative to THP-1 cells was observed. For one compound (**7**) a selective toxicity superior to that displayed by the trypanocidal prodrug nifurtimox was observed (agents **7** and nifurtimox generated SI values of ~85 and ~65, respectively) although SI values could not be calculated for **5**, **11**, benznidazole, or **38** as they displayed no mammalian cell toxicity.

2.2 | Benznidazole and related 2-nitroimidazoles promote DNA damage

Nitroheterocycles are well known for promoting DNA damage (Edwards, 1993). To assess whether this is the case for the most potent trypanocidal 2-nitroimidazoles (EC_{50} value of <30 μM and benznidazole) identified here, cells null for the DNA repair enzymes TbCSB, TbMRE11, TbEXO1 or TbSNM1, markers of the TC-NER, HR, MMR, or interstrand crosslink repair systems, respectively, were grown in the presence of the selected compound and the resultant data plotted as dose response curves from which EC_{50} values for each treatment were determined (Figures 1 and S1). Cells lacking TbCSB or TbMRE11 were shown to be more susceptible (2.5- to 10-fold) to agents **6**, **10,** or **11** than wild type (Figure 1a–c) with the TbMRE11 deficient line also displaying an increased sensitivity (2.5-fold) to benznidazole (Figure 1d). This indicates that these four nitroimdazoles can all promote DNA damage in the nuclear genome of BSF *T. brucei* with **6**-, **10**-, or **11**-generated lesions that may potentially cause stalling of DNA transcription. In contrast, TbCSB- and TbMRE11-deficient trypanosomes displayed wild-type-like susceptibilities towards the nitroimidazoles **5**, **7**, **15**, **34**, **35**, **36,** and **38** and the trypanocidal nitrofuran prodrug nifurtimox (Figure 1e) while parasites lacking TbEXO1 or TbSNM1 yielded EC_{50} values equivalent to that exhibited by controls for all compounds tested (Figures 1 and S1).

To investigate whether there was a functional linkage between TbCSB and TbMRE11 in the repair of the **6**-, **10**-, or **11**-generated DNA damage, the susceptibility of parasites null for both activities was determined (Figure 1a–c). This revealed that cells lacking TbCSB and TbMRE11 were hypersensitive to **10** and **11** (~58- and ~11-fold, respectively) relative to wild type and the two single null lines showing that these two enzymes function in a nonepistatic fashion and therefore do not operate in the same pathway to repair any DNA damage arising from treatment with these two nitroimidazoles. In contrast, the susceptibility of the double null line towards **6** was similar to that noted for lines lacking TbCSB or TbMRE11 indicating that these repair factors function epistatically to resolve DNA damage arising from treatment with this nitroimidazole.

TABLE 1 Susceptibility of *Trypanosoma brucei* and mammalian cells to 2-nitroimidazoles

	$EC_{50} \pm SD$ (in μM)[b,c]			
		T. brucei TbNTR^{++}[d,e]		
Compound[a]	*T. brucei* wild type	−tet	+ tet	THP-1 (SI[f])
3, 13, 14, 16, 18, 21, 22, 24, 26, 27, 28, 29, 32, 33, 37 & 39	>100	–	–	–
1	87.9 ± 2.5	–	–	–
2	56.0 ± 5.7	–	–	–
4	60.9 ± 5.7	–	–	–
5	15.3 ± 0.5	11.6 ± 0.5	1.6 ± 0.2 (10)	>300 (>20)
6	1.4 ± 0.1	1.5 ± 0.1	0.1 ± 0.0 (15)	9 (6)
7	20.3 ± 1.7	22.2 ± 1.7	2.0 ± 0.1 (11)	172 (86)
8	52.7 ± 6.6	–	–	–
9	30.7 ± 0.7	–	–	–
10	8.5 ± 0.2	11.1 ± 1.3	0.9 ± 0.0 (12)	152 (18)
11	24.9 ± 3.2	25.2 ± 0.6	2.0 ± 0.1 (13)	>300 (>12)
12	58.9 ± 2.2	–	–	–
15	10.0 ± 0.1	13.7 ± 1.0	1.2 ± 0.1 (11)	48 (9)
17	35.2 ± 2.6	–	–	–
19	48.9 ± 2.2	–	–	–
20	74.0 ± 7.4	–	–	–
23	43.0 ± 3.1	–	–	–
25	73.0 ± 2.0	–	–	–
30 (benznidazole)	66.2 ± 0.8	65.5 ± 2.9	10.2 ± 0.7 (6)	>300 (>5)
31	75.7 ± 5.0	–	–	–
34	11.3 ± 0.3	13.2 ± 0.7	3.31 ± 0.2 (4)	127 (11)
35	23.8 ± 7.9	22.5 ± 1.9	6.97 ± 0.2 (3)	–
36	7.9 ± 0.2	7.6 ± 0.1	0.95 ± 0.1 (8)	23 (30)
38	18.5 ± 2.2	21.8 ± 2.8	6.96 ± 1.4 (3)	>300 (19)
Nifurtimox	3.6 ± 0.2	3.3 ± 0.2	0.5 ± 0.0 (7)	65 (18)

[a]Compound structures are given in Table S1.

[b]Data are means from four experiments ± standard deviation.

[c]Compounds not tested against the *T. brucei* TbNTR^{++} or TPH-1 cells are depicted as "–".

[d]*T. brucei* TbNTR^{++} represents the parasite line that when cultured in medium containing tetracycline (+tet) overexpresses TbNTR. When grown in medium lacking tet (−tet), these cells display susceptibility phenotypes similar to wild type.

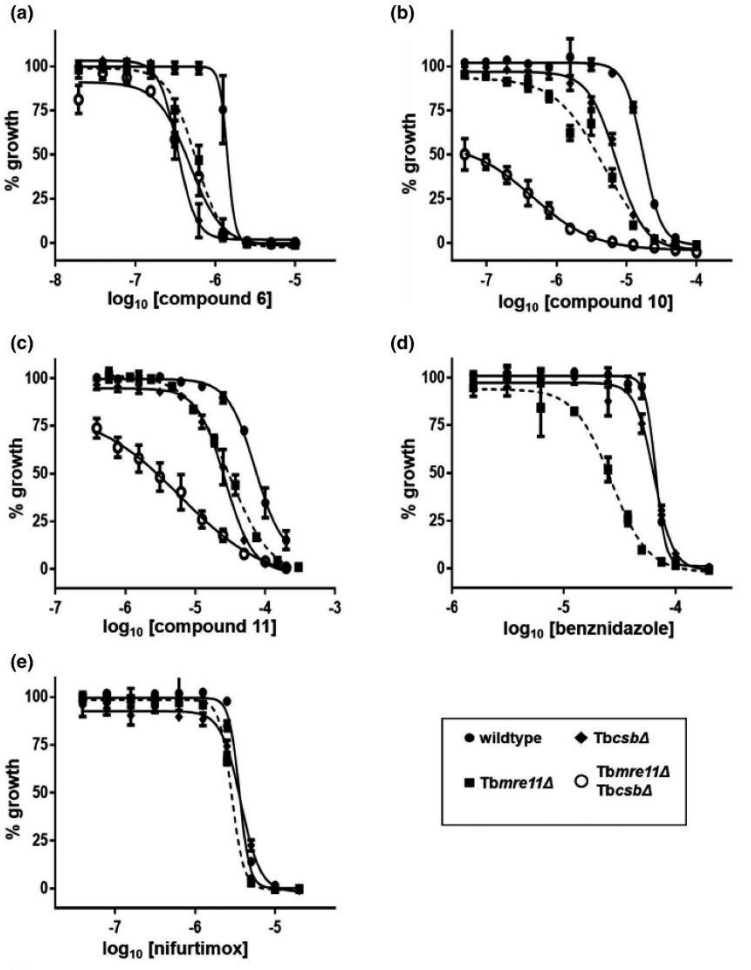
[e]The fold difference in susceptibility between +tet versus −tet treated *T. brucei* TbNTR^{++} is in parenthesis and is all judged to be statistically significant ($p < .01$), as assessed by Student's *t* test.

[f]Selective index (SI; in parenthesis) of a compound was calculated as a ratio of the EC_{50} against mammalian (THP-1) cells to the EC_{50} against wild type parasites.

2.3 | Linking prodrug activation to DNA damage

To establish whether there is a link between the activation of the trypanocidal nitroimidazoles and the TbMRE11-/TbCSB-mediated DNA repair pathways, the susceptibility of Tb*mre11* or Tb*csb* null mutants expressing ectopic copies of Tb*ntr* towards selected compounds was evaluated. To construct the cell lines, gene deletion (for Tb*csb*) or disruption (for Tb*mre11*) vectors based around blasticidin-S-deaminase (*bla*)- or puromycin N-acetyltransferase (*pac*) resistance cassettes were electroporated into *T. brucei* previously engineered to express elevated levels of Tb*ntr* (Wilkinson et al., 2008) with clonal cell lines selected and validated (Figure S2; Dattani & Wilkinson, 2019).

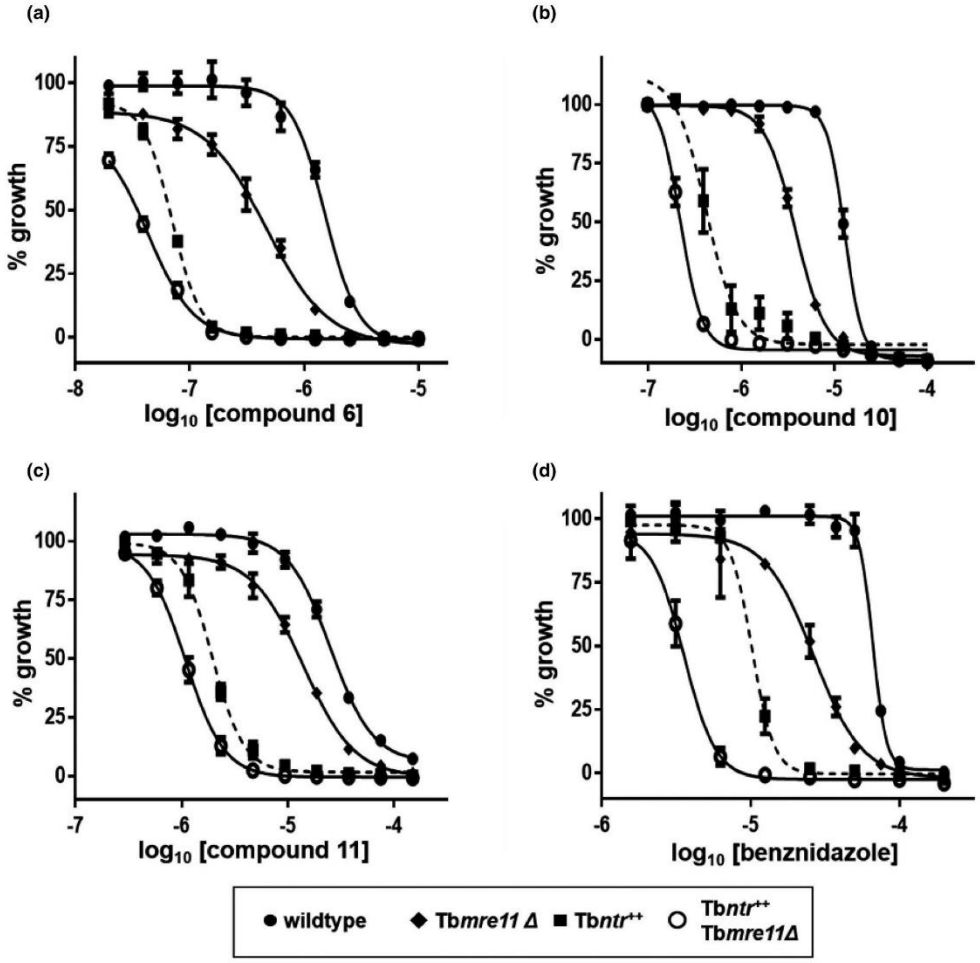
When evaluating the linkage between NTR activation and the HR pathway, an increase in sensitivity of the Tb*mre11* null line and *T. brucei* expressing elevated levels of Tb*ntr* towards the nitroheterocycles tested (**6**, **10**, **11** and benznidazole) was observed relative to controls (Figure 2). This phenotype was further enhanced in Tb*mre11*-deficient trypanosomes expressing an ectopic copy of Tb*ntr*. For example, *T. brucei* lacking Tb*mre11* or over expressing Tb*ntr* was 2.5- or 6.5-fold more



(f)

Cell line	compound				
	6	10	11	bnz	nfx
wild type	1.5 ± 0.1	17.3 ± 0.6	24.9 ± 3.2	66.5 ± 2.4	3.28 ± 0.10
Tb*exo1*Δ	1.3 ± 0.1	19.6 ± 0.8	27.7 ± 1.5	60.7 ± 0.7	3.5 ± 0.1
Tb*snm1*Δ	1.5 ± 0.0	17.3 ± 0.4	28.6 ± 1.3	66.1 ± 2.7	3.23 ± 0.0
Tb*mre11*Δ	0.6 ± 0.2*	4.8 ± 0.7*	9.5 ± 0.4*	26.4 ± 1.4*	2.7 ± 0.05
Tb*csb*Δ	0.3 ± 0.0*	7.2 ± 0.4*	10.4 ± 1.3*	64.3 ± 1.0	3.7 ± 1.8
Tb*mre11*ΔTb*csb*Δ	0.5 ± 0.1*	0.3 ± 0.2*#	2.3 ± 0.7*#	nd	nd

FIGURE 1 Susceptibility of *Trypanosoma brucei* null mutants towards 2-nitroimidazoles. (a–e) Dose response curves of compounds 6 (a), 10 (b), 11 (c), benznidazole (d), and nifurtimox (e) against *T. brucei* wild type (wt), *T. brucei* null for Tb*mre11* (Tb*mre11Δ*) or Tb*csb* (Tb*csbΔ*), and parasites lacking Tb*mre11* and Tb*csb* (Tb*mre11Δ* Tb*csbΔ*). The wild-type curves are representative of graphs obtained when the screens were extended to cells lacking Tb*exo1* or Tb*snm1* (Figure S1). All data points are mean values ± standard deviations from experiments performed in quadruplicate. (f) The susceptibility of the *T. brucei* single and double null mutant lines against the 2-nitroimidazole, as judged by their EC_{50} values (in μM). *Indicates significant differences in susceptibility ($p < .0002$) between the Tb*mre11Δ*, Tb*csbΔ*, and Tb*mre11Δ* Tb*csbΔ* line relative to the wild type, while #indicates significant differences in susceptibility ($p < .0001$) between the Tb*mre11Δ* Tb*csbΔ* line relative to Tb*mre11Δ* and/or Tb*csbΔ* parasites, as assessed by Student's *t* test (GraphPad Software Inc.)



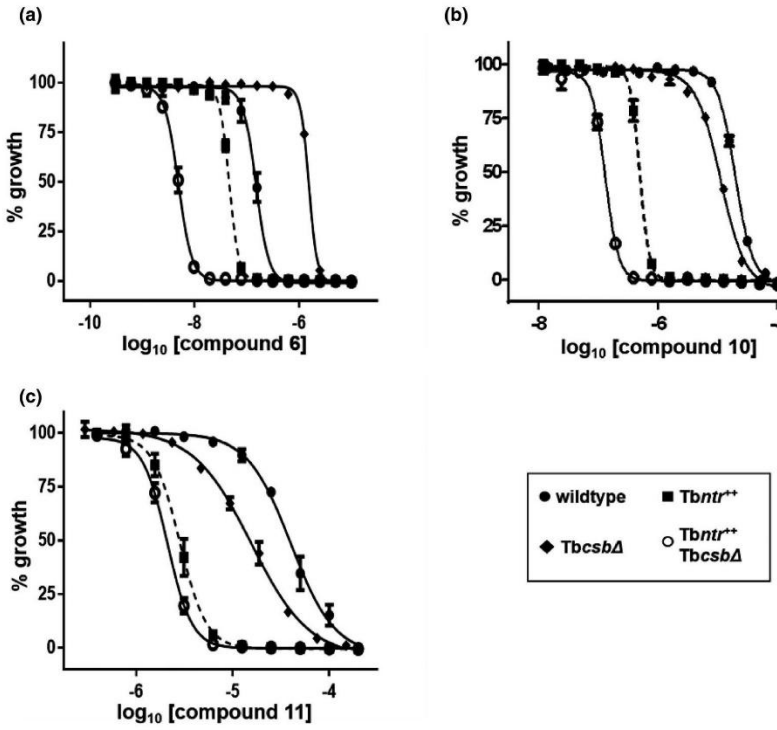
(e)

Cell line	compound 6	10	11	bnz
wild type	1.52 ± 0.06	12.90 ± 0.50	25.20 ± 0.60	66.5 ± 2.4
Tb*mre11Δ*	0.47 ± 0.06*	3.80 ± 0.20*	13.40 ± 0.60*	26.20 ± 0.70*
Tb*ntr*$^{++}$	0.07 ± 0.00*	0.50 ± 0.10*	2.00 ± 0.1*	10.20 ± 1.40*
Tb*ntr*$^{++}$Tb*mre11Δ*	0.04 ± 0.00*,#	0.20 ± 0.00*,#	1.10 ± 0.10*,#	3.50 ± 0.30*,#

FIGURE 2 Linking activation of 2-nitroimidazole prodrugs to DNA damage. (a–d) Growth inhibitory effect of 6 (a), 10 (b), 11 (c), and benznidazole (d) on *Trypanosoma brucei* wild type (wt), *T. brucei* expressing an ectopic copy of Tb*ntr* (Tb*ntr*++), a Tb*mre11* null line (Tb*mre11Δ*) and Tb*mre11* null line expressing elevated levels of Tb*ntr* (Tb*ntr*$^{++}$Tb*mre11Δ*). All data points are mean values ± standard deviations from experiments performed in quadruplicate. (e) The susceptibility of *T. brucei* lines against each 2-nitroimidazole, as judged by their EC_{50} values (in μM). *indicates significant differences in susceptibility ($p < .0001$) between the Tb*ntr*$^{++}$, Tb*mre11*Δ, and Tb*ntr*$^{++}$Tb*mre11Δ* lines relative to the wild type, while #indicates significant differences in susceptibility ($p < .001$ or .0001) between Tb*ntr*$^{++}$Tb*mre11Δ* cells relative to Tb*ntr*$^{++}$ and Tb*mre11*Δ parasites, as assessed by Student's *t* test (GraphPad Software Inc.)

susceptible to benznidazole relative to wild type, respectively, with a 19-fold increase in potency noted to parasites deficient in TbMRE11 that also express elevated levels of TbNTR (Figure 2d). Together, these data indicate that the TbMRE11 repair pathway plays an important role in resolving the DNA damage generated by **6**, **10**, **11**, and benznidazole following their activation by TbNTR.

Extending the investigation to assess for any association between prodrug activation and the TbCSB repair pathway, two of the tested compounds (**6** and **10**) behaved similarly to the Tb*mre11* null lines (Figure 3). Treatment of the Tb*csb* null line and *T. brucei* expressing elevated levels of Tb*ntr* with **6** or **10** resulted in an increased susceptibility towards the two nitroheterocycles, while the Tb*csb* null line that over expresses TbNTR displayed hypersensitivity (Figure 3a,b): TbCSB null trypanosomes were 10- and 2-fold more sensitive to **6** or **10**, respectively, *T. brucei* ectopically expressing elevated level of TbNTR were 33-fold and 40-fold more sensitive to each compound, respectively, while the TbCSB deficient and TbNTR over expressing line exhibited a ~310- and ~200-fold increase in



(d)

	compound		
Cell line	6	10	11
wild type	1.54 ± 0.01	19.60 ± 0.50	37.60 ± 4.80
Tb*csb1Δ*	0.15 ± 0.00*	11.10 ± 0.20*	15.60 ± 2.00*
Tb*ntr*$^{++}$	0.05 ± 0.00*	0.50 ± 0.00*	2.80 ± 0.38*
Tb*ntr*$^{++}$Tb*csbΔ*	0.01 ± 0.00*,#	0.10 ± 0.00*,#	2.10 ± 0.10*,#

FIGURE 3 Linking activation of 2-nitroimidazole prodrugs to DNA damage. (a–c) Growth inhibitory effect of 6 (a), 10 (b), and 11 (c) on *Trypanosoma brucei* wild type (wt), *T. brucei* expressing an ectopic copy of Tb*ntr* (Tb*ntr*$^{++}$), Tb*csb* null mutant (Tb*csbΔ*), and Tb*csb* null mutants expressing elevated levels of Tb*ntr* (Tb*ntr*$^{++}$Tb*csbΔ*) lines. All data points are mean values ± standard deviations from experiments performed in quadruplicate. (d) The susceptibility of each *T. brucei* lines against each 2-nitroimidazole, as judged by their EC_{50} values (in μM). *indicates significant differences in susceptibility ($p < .0001$) between the Tb*ntr*$^{++}$, Tb*csbΔ*, or Tb*ntr*$^{++}$Tb*csbΔ* lines relative to the wild type, while #indicates significant differences in susceptibility ($p < .0001$) between the Tb*ntr*$^{++}$Tb*csbΔ* line relative to Tb*ntr*$^{++}$ and Tb*csbΔ* parasites, as assessed by Student's *t* test (GraphPad Software Inc.)

sensitivity, respectively. When treated with **11**, the *T. brucei* Tb*csb* null line and Tb*ntr* over expressing lines displayed increased susceptibility (2.5- and 13.5-fold, respectively; Figure 3c). However, parasites lacking TbCSB and ectopically expressing TbNTR only exhibited a slight (18-fold) increase in potency relative to lines with elevated levels of the nitroreductase. Together, the above susceptibility data demonstrate a clear link between prodrug activation and DNA repair with the NTR-mediated activation of **6**, **10,** and to a lesser extent **11** promoting formation of lesions that are resolved by a pathway in which TbCSB plays a key role.

2.4 | Benznidazole and the HR repair pathway

In yeast and mammalian cells the localization and/or expression level of HR DNA repair factors changes in response to DNA damage. To determine whether benznidazole promotes similar alterations in *T. brucei*, parasite lines were engineered to express TbMRE11 or TbRAD51 variants tagged at their amino terminal with the fluorescent protein mNeonGreen. Using a PCR-based CRISPR/Cas9 system, amplicons corresponding to gene-specific single guide RNA (sgRNA) and targeting cassettes were cotransformed into BSF *T. brucei* cells expressing T7 RNA polymerase/*Streptococcus pyogenes* Cas9 nuclease genes and the selected clonal recombinant lines validated (Figure S3).

The fluorescent signal of each recombinant line grown in the absence or presence of benznidazole was evaluated (Figure 4a). For wild type and some (~30%) untreated mNeonGreen-TbMRE11 expressing *T. brucei* no signal was detected. For the remaining untreated recombinant cells, a weak signal throughout the nucleus was observed with some (>10%) possessing a few discrete foci. Upon treatment with benznidazole (30 μM for 2 hr) the number of mNeonGreen-TbMRE11 expressing trypanosomes now possessing intense punctate nuclear signal had increased relative to untreated controls, a characteristic also noted in phleomycin treated parasites (5 μg/ml; overnight; Figure 4a,c). These fluorescent centres were found in the nucleoplasm and present in cells at all stages of the cell cycle. This alteration in TbMRE11 expression and subnuclear localization was shown to take place within the first hour of adding benznidazole (30 μM) to the culture medium with the fluorescence signal in cells from treated populations remaining at approximately the same level over the next 2 hr before declining 4-hr posttreatment and reverting to untreated levels by 6 hr (Figure 5a).

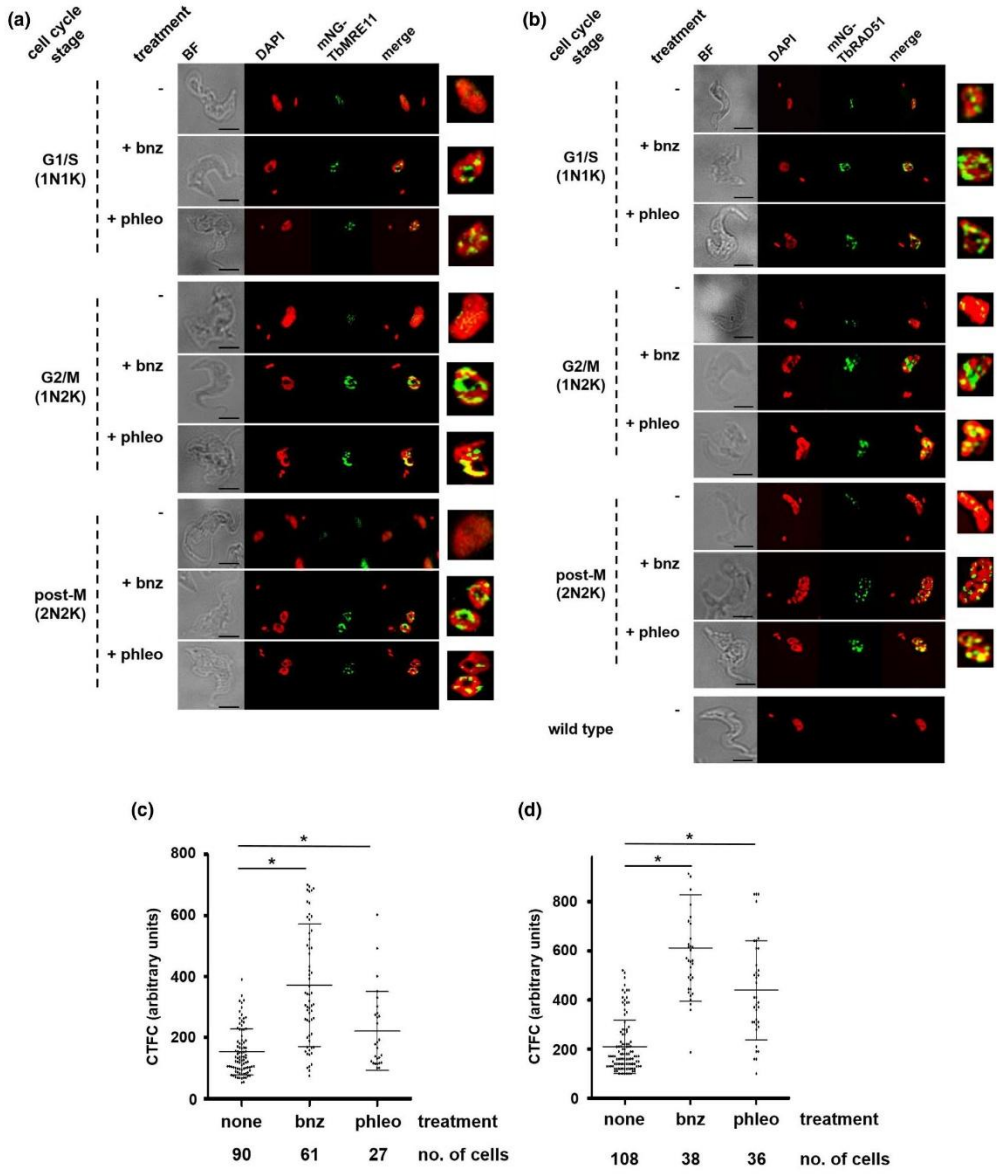
For trypanosomes engineered to express mNeonGreen-TbRAD51 and cultured in untreated growth medium, most (~90%) cells presented a few (1 to 5) discrete foci with no discernible signal observed in the remaining (~10%) parasites (Figure 4b). Following exposure to benznidazole (30 μM for 6 hr) or phleomycin (5 μg/ml; overnight) and in 60%–70% of cases, the number and intensity of the discrete foci increased relative to untreated controls with these fluorescent centers found throughout the nucleoplasm in cells at all stages of the cell cycle (Figure 4b,d). Intriguingly, some (5%–10%) drug-treated trypanosomes presented with a single, intense TbRAD51 spot possibly indicating that the multiple foci previously noted may have coalesced. The above benznidazole-induced alterations to TbRAD51 expression and subnuclear localization were temporal, becoming apparent 4-hr posttreatment, a time point where the TbMRE11 signal had started to decline (Figure 5b).

In trypanosomes, the phosphorylation of H2A can be used as a biomarker for DNA damage including DSB formation (Glover & Horn, 2012). To determine if benznidazole can promote this post-translational modification in *T. brucei*, protein extracts generated from wild-type parasites grown in the presence of the 2-nitroimidazole (30 μM) were probed for γH2A formation using an antiserum against this modified histone (Figure 5c): As a control, enolase was analyzed in parallel. Comparison of the normalized γH2A and enolase band intensities revealed that the γH2A signal increased over the first 4 hr of treatment then declined over the next 4 hr indicating resolution of the DNA damage (Figure 5d). A summary of the combined temporal changes in γH2A formation and in TbMRE11 or TbRAD51 expression in response to benznidazole is presented in Figure 5e.

3 | DISCUSSION

Over the last decade, there has been renewed interest in developing nitroimidazole-based compounds as anti-infectives resulting in pretomanid, delamanid, and fexinidazole gaining approval for treatment of tuberculosis or HAT (Deeks, 2019; Liu et al., 2018; Patterson & Fairlamb, 2019; Thakare et al., 2020). Invariably, such compounds function as prodrugs with how they undergo activation often underlying specificity. Here, and using the anti-Chagasic prodrug benznidazole as lead, the growth inhibitory properties of a 2-nitroimidazole library were tested against BSF *T. brucei* and their mechanism of action assessed. Of the 39 compounds evaluated, 10 displayed activity towards the parasite (EC_{50} values of <30 μM) with the growth inhibitory properties of these being TbNTR1-dependent. Additional phenotyping of these active compounds (and benznidazole) on trypanosome lines lacking the DNA repair factors TbSNM1, TbEXO1, TbMRE11, and/or TbCSB revealed that several (including benznidazole) were able to promote DNA damage in the *T. brucei* nuclear genome. Based on susceptibility profiles, these four compounds could be divided into three groups. We show that cells lacking TbMRE11 were more sensitive to benznidazole than controls while trypanosomes null for TbMRE11 or TbCSB displayed a greater sensitivity to the remaining compounds relative to wild type, with these activities exhibiting epistatic or nonepistatic interactions.

Trypanosomal NTR1s play a key role in the activation of benznidazole with this reduction reaction producing highly reactive nitroso, hydroxylamine, nitrenium, and hydroxy intermediaries (Hall & Wilkinson, 2012; Trochine et al., 2014). Many of these metabolites can readily interact with and covalently modify different macromolecules within the parasite resulting in pleotropic effects within the cell (Goijman & Stoppani, 1985; Trochine et al., 2014). One biological molecule that is particularly prone to the effects of benznidazole is DNA with prodrug-induced base substitution, oxidative lesions, and



DSB all detected in trypanosomal-based assays (Campos et al., 2017; Goijman et al., 1985; Rajao et al., 2014). In *T. cruzi*, this damage results in the accumulation of genome-wide mutations and alterations in gene copy number that can promote resistance with benznidazole and other structurally related and unrelated drugs/prodrugs (Campos et al., 2017). Based on susceptibility studies and expression profiling, we postulate that following treatment with benznidazole, DSBs are generated in the nuclear genome of BSF *T. brucei*, with markers associated with trypanosomal HR-based repair pathways (TbMRE11, γH2A, and TbRAD51) rapidly recruited to lesion sites that cooperate to resolve the prodrug-induced damage (Figures 1, 4, and 5). Precisely how the DSBs are generated in the nuclear genome of the trypanosome remains unclear but given that the benznidazole susceptibility phenotype displayed by *T. brucei* lacking TbMRE11 is exacerbated

FIGURE 4 Effect of benznidazole on TbMRE11 and TbRAD51. (a, b) Parasites expressing mNeonGreen tagged TbMRE11 (a) (mNG-TbMRE11; green) or TbRAD51 (mNG-TbRAD51) treated with benznidazole (30 μM for 2 or 4 hr for mNG-TbMRE11 and mNG-TbRAD51, respectively) (+bnz) or phleomycin (5 μg/ml; overnight) (+phleo) were co-stained with DAPI (DNA; red). The cells were examined by fluorescence microscopy and the brightfield (BF) image captured. The pattern of colocalization (merged; yellow), including a close up image of the nucleus, in treated and untreated trypanosomes at different stages of the cell cycle was compared: The ratio of the nuclear (N) and mitochondrial (known as the kinetoplast [K]) genomes in a single parasites represents a marker for the trypanosomal cell cycle, with *Trypanosoma 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 (Glover & Horn, 2012; Siegel et al., 2008; Woodward & Gull, 1990). The fluorescence signals obtained when analyzing a wild-type cell are also shown. Scale bar = 5 μm. (c, d) The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen tagged TbMRE11 (c) or TbRAD51 (d) following benznidazole (bnz) or phleomycin (phleo) treatment was compared against untreated controls (none). Each datum point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the gray horizontal lines. The number of cells analyzed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($p < .001$) between untreated and treated cells, as assessed by Student's *t* test (GraphPad Software)

when expressing elevated levels of TbNTR1 (Figure 2), metabolites generated by activation are key to promoting the DNA damage. Such metabolites can promote DSB formation directly through interaction with the DNA backbone, or indirectly during the near-simultaneous repair of oxidative base lesion clusters found on opposing DNA strands (Cannan & Pederson, 2016; Edwards, 1993; Kedderis et al., 1989).

To evaluate the temporal response of the *T. brucei* HR pathway following exposure to benznidazole, the changes in TbMRE11 and TbRAD51 expression and in γH2A formation were monitored (Figure 5). This demonstrated that TbMRE11, as part of the parasite's DSB sensory MRN complex (Passos-Silva et al., 2010), is rapidly recruited (in the first hour of treatment) to lesion sites, in keeping with the situation observed for mammalian cells (Haince et al., 2008). Expression of this nuclease remained constant over the next 2 hr before declining to pretreatment levels. TbMRE11 recruitment to DSB then promotes phosphorylation of H2A yielding γH2A (Glover & Horn, 2012), a process that peaks 4-hr postbenznidazole treatment before tailing off. Next, and guided by γH2A deposition, TbRAD51 (along with other effector proteins) is recruited to the damage site with the level of this recombinase rising 4 to 6 hr after addition of the prodrug to cultures. Once at the damage site, the effector proteins facilitate crossover events using sister chromatids as template, resolving the DSB. In procyclic form *T. brucei*, the rate of HR-mediated DSB repair after ionizing radiation (IR) is considerably faster than that observed here for benznidazole: γH2A formation and TbRAD51 deposition peaked approximately 2 and 2 to 5 hr post-IR damage, respectively (Marin et al., 2018). The variation in recruitment kinetics shown by the HR factors in DSB repair may reflect differences between the two *T. brucei* developmental stages examined, as observed for other lesions (Vieira-da-Rocha et al., 2019) or the nature of the DSB-inducing treatment itself.

When the phenotyping studies were extended to other 2-nitroimdazoles, three compounds (**6**, **10**, and **11**) were shown to promote TbNTR1-dependent DNA damage in the *T. brucei* nuclear genome, although different repair mechanisms appear to function to resolve the lesions they generate (Figures 1–3). This may reflect the different types of damage the three nitroimidazoles invoke, the cell cycle stage where their effect is most pronounced, and/or pharmacological differences between the parental compounds and their metabolites. Based on the epistatic interaction displayed by TbCSB and TbMRE11 required to repair the lesions induced by **6**, we postulate metabolites derived from **6** can readily access the nucleus and interact with DNA to generate pleiotropic effects including base oxidation and DSB damage, lesions that cause stalling of transcription complexes. By analogy to responses reported in other eukaryotes, TbCSB, a helicase whose activity was thought to be restricted to TC-NER may then be recruited to lesion site where it can invoke transcription-coupled HR repair or BER pathways (Batenburg et al., 2015; Menoni et al., 2012, 2018; Stevnsner et al., 2008; Wei et al., 2016). The BER pathway can inadvertently give rise to DSBs via the near-simultaneous repair of clusters of oxidized bases occur (Cannan & Pederson, 2016). In contrast, the repair of **10**- and **11**-induced lesions requires systems where TbCSB and TbMRE11 operate independently of each other. In these situations, the damage created by these nitroimidazoles may cause: (1) distortion of the DNA backbone where TbCSB functions to guide a TC-NER response (Machado et al., 2014) and (2) DSBs that are repaired by TbCSB-independent and TbMRE11-dependent HR repair pathways (Conway et al., 2002).

Here, we establish that one of the anti-*T. brucei* activities of benznidazole upon activation by TbNTR involves DNA damage and that components of the HR pathway help resolve the induced lesions. The lines developed here will allow us to readily screen if the trypanocidal activities of other nitroheterocycles, including fexinidazole, function through a similar mechanism and by constructing equivalent cells in other trypanosomes such as *T. cruzi* will allow us to determine if this DNA damage effect extends across to related organisms.

4 | EXPERIMENTAL PROCEDURES

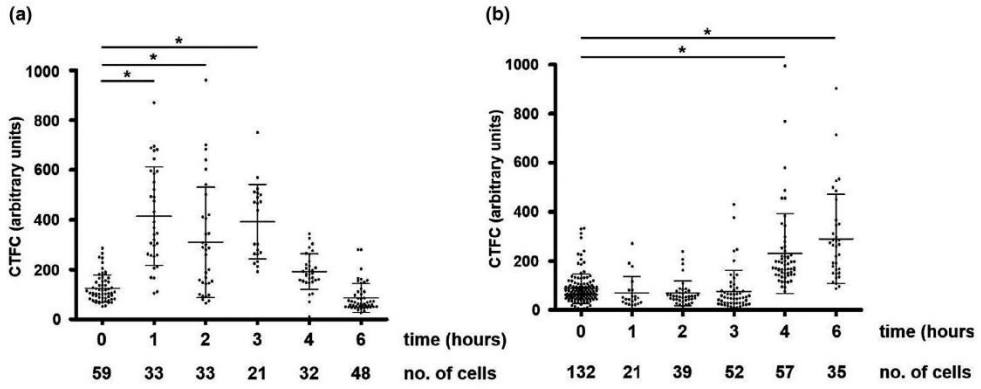
4.1 | Compounds

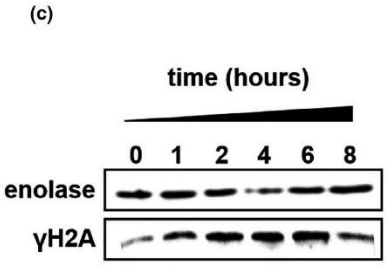
The compounds and their structures used in this study are listed in Table S1. These were sourced from the Drug Synthesis and Chemistry Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, USA,

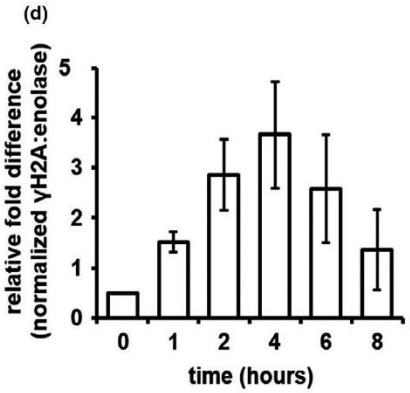
except for **16** and **25**, which were supplied by Prof Cameron Koch (University of Pennsylvania, USA), benznidazole (**30**), and nifurtimox, which were supplied by Prof Simon Croft (School of Hygiene and Tropical Medicine, UK), and **36**, which was obtained from Dr Maria Papadopoulou (NorthShore University Health System, USA). The selective compounds blasticidin, puromycin, G418, and hygromycin were purchased from Melford Laboratories Ltd.

4.2 | Cell culturing

Bloodstream form *Trypanosoma brucei brucei* MITat 427 strain; clone 221a (designated as wild type) and a derivative (2T1) engineered to constitutively express the tetracycline repressor protein were grown in HMI-11 medium at 37°C under a 5% (v/v) CO_2 atmosphere (Alsford et al., 2005; Hirumi & Hirumi, 1989). Derivatives







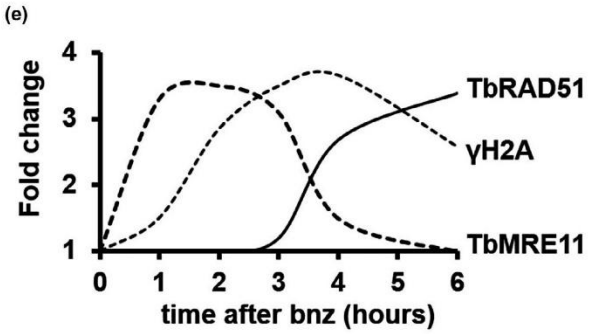


FIGURE 5 Effect of benznidazole on the temporal expression of the *Trypanosoma brucei* homologous recombination DNA repair pathway. (a, b) The corrected total cellular fluorescence (CTFC) of individual *T. brucei* cells expressing mNeonGreen tagged TbMRE11 (a) or TbRAD51 (b) was determined at time intervals following benznidazole (30 μM) treatment. Each data point represents the fluorescence of an individual cell, with the mean fluorescence values per cell ± standard deviation represented by the gray horizontal lines. The number of cells analyzed for each treatment and/or at each time point is given. The asterisk (*) indicates significant differences in the mean fluorescence values per trypanosome ($p < .001$) between untreated and treated cells, as assessed by Student's *t* test (GraphPad Software). (c) *T. brucei* cells were treated with 30-μM benznidazole. Cell lysates were generated at time intervals (0, 1, 2, 4, 6, and 8 hr) and analyzed by western blot using anti-*T. brucei* enolase (loading control) and anti-*T. brucei* γH2A antiserum. (d) The γH2A and enolase signal intensities obtained from western blots containing extracts from three independent benznidazole treated cultures were determined using Image Studio™ Lite (Li-COR Biosciences) and normalized against untreated controls. The data derived from three independent blots are expressed as a mean ± standard deviation relative fold difference between the normalized γH2A and enolase signals. (e) Schematic showing the temporal changes in expression of TbMRE11 or TbRAD51 or γH2A formation following addition of benznidazole (bnz) to cultures

of "221" null for TbCSB, TbEXO1, TbSNM1, and/or TbMRE11 were maintained in the HMI-11 medium containing 2.5-μg/ml hygromycin, 10-μg/ml blasticidin, 2-μg/ml puromycin, and/or 2-μg/ml G418 (Dattani & Wilkinson, 2019; Sullivan et al., 2015), while genetically modified 2T1 lines overexpressing TbNTR and/or null for the DNA repair enzymes TbCSB or TbMRE11 were maintained in growth medium containing 2.5-μg/ml hygromycin, 10-μg/ml blasticidin, 2-μg/ml puromycin, and/or 2-μg/ml G418 as appropriate (Wilkinson et al., 2008). For CRISPR/Cas9 experiments, the *T. b. brucei* SmOx B427 CAS9/pTB011 line, parasites engineered to express the T7 RNA polymerase, and humanized *S. pyogenes* Cas9 nuclease genes (Beneke et al., 2017) were cultured in HMI-11 medium supplemented with 10-μg/ml blasticidin and 2-μg/ml puromycin. Transformed derivatives of this were additionally maintained in 2.5-μg/ml hygromycin.

The human acute monocytic leukemia cell line (THP-1) was grown at 37°C under a 5% (v/v) CO_2 atmosphere in RPMI-1640 medium (GE Healthcare) supplemented with 2-mM pyruvate, 2-mM sodium glutamate, 2.5-U ml^{-1} penicillin, 2.5-μg/ml streptomycin, 20-mM HEPES pH 7.4, and 10% (v/v) fetal calf serum.

4.3 | Construction of recombinant parasite lines

Construction of the DNA vectors used to delete (Tb*csb*) or interrupt (Tb*mre11*; Tb*exo1*) the targeted open reading frame and containing hygromycin B phosphotransferase (*hyg*), neomycin phosphotransferase (*neo*), blasticidin-S-deaminase (*bla*), or puromycin N-acetyltransferase (*pac*) resistance cassette has been described elsewhere (Dattani & Wilkinson, 2019). The constructs were linearized (SacI/KpnI for the *hyg*, *neo*, and *pac* constructs and SacII/KpnI for the *bla* vector) then transformed into BSF *T. brucei* (221 wild type or 2T1 engineered to express Tb*ntr*) using the Human T-cell nucleofection® kit and an Amaxa® Nucleofector™ (Lonza AG) set to program X-001 and clonal recombinant parasites selected using the appropriate drug selection.

The PCR-based CRISPR/Cas9 strategy described by Beneke et al. (2017) was used to generate *T. brucei* lines that express TbMRE11 or TbRAD51 tagged at their amino terminal with mNeonGreen (Beneke et al., 2017). The primer sequences used to amplify DNA fragments corresponding to the sgRNA and targeting cassette were designed using the LeishGEdit primer design (http://www.leishgedit.net/Home.html) software: Primer sequences and combinations are given in Tables S2 and S3, respectively. To generate an amplicon corresponding to the sgRNA, the sgRNA scaffold and a gene-specific primer were mixed in 1× Phusion Master Mix (ThermoFisher Scientific) and subject to 98°C for 30 s followed by 35 cycles of 98°C for 10 s, 60°C for 30 s, and 72°C for 15 s. To amplify the targeting cassette from the pPOTv7-hygro-hygro-mNeonGreen plasmid (Sam Dean, University of Oxford), template DNA and gene-specific upstream forward and upstream reverse primers were mixed in 1×Phusion Master Mix (ThermoFisher Scientific) and sbject to 98°C for 30 s followed by 35 cycles of 98°C for 10 s, 65°C for 30 s, 72°C for 90 s, and a final elongation step at 72°C for 420 s. The two resulting amplicons were combined, purified using a PCR Purification kit (ThermoFisher Scientific), and eluted in Tb-BSF electroporation buffer (Schumann Burkard et al., 2011). DNA fragments were transformed into BSF *T. brucei* SmOx B427 CAS9/pTB011 (10^7) suspended in Tb-BSF using program X-001 on an Amaxa® Nuxleofector™ (Lonza AG) and clonal recombinant parasites selected using hygromycin.

4.4 | Validation of recombinant parasites

To demonstrate that integration of a gene interruption or mNeonGreen-based targeting cassette had occurred at the correct genetic loci and confirm that the *T. brucei* null mutant line was no longer expressing the targeted gene(s), DNA amplification reactions were performed on parasite genomic DNA or cDNA templates. For gene interruption constructs, the primer combinations used generated amplicons specific for the intact targeted gene or the *hyg*-, *neo*-, *pac*-, and/or *bla*-disrupted allele. In the case where a gene of interest was tagged with sequences encoding for the fluorescent protein, a mNeonGreen forward and a gene-specific reverse primer was used to generate amplicons encompassing the fusion event. Fragments amplified from cDNA using Phusion High Fidelity DNA polymerase (ThermoFisher Scientific) were sequenced (Source BioScience) to show that the CRISPR/Cas9 integration had generated the desired in-frame fusion event. To check integrity of cDNA (and hence RNA) and to act as a loading control, additional amplifications reactions aimed at detecting expression of the housekeeping gene Tb*tert* were

performed. The primer sequences and combinations used in cell line validation are listed in Tables S2 and S3.

4.5 | Microscopy

BSF *T. brucei* in the exponential growth phase were fixed in growth medium with an equal volume of 2% (w/v) paraformaldehyde/PBS and washed twice in PBS and aliquots containing approximately 2.5×10^5 cells air dried onto a microscope slide. Trypanosome genomes were stained using Vectashield Mounting Medium containing 4′,6-diamidino-2-phenylindole (DAPI) (Vectorshield Laboratories) and cells visualized using a DeltaVision Elite Deconvolution Microscope (GE Healthcare). All images were processed and corrected fluorescence intensities determined using ImageJ and the statistical significance of any differences in mean fluorescence values per cell assessed using the Student's *t* test calculator (GraphPad Software).

4.6 | Protein analysis

For western blot analysis, protein extracts from approximately 5×10^6 trypanosomes were probed with rabbit anti-γH2A (Dr Richard McCulloch, University of Glasgow) and rabbit anti-enolase (Prof Paul Michels, University of Edinburgh) antisera used at 1:1,000 or 1:10,000 dilution, respectively, followed by goat antirabbit IRDye™ 800CW antibody (LI-COR) diluted at 1:5,000. Detection of the near infrared signal was monitored using an Odyssey® CLx infrared imaging system (Li-COR Biosciences). Band signal intensity measurements were carried out using the Analysis Tool in Image Studio™ Lite (Li-COR Biosciences).

4.7 | Phenotypic screens

All assays were performed in a 96-well plate format. *T. brucei* BSF parasites (200 cells per well) were treated with growth medium containing different concentrations of nitroimidazole. Where appropriate, protein expression was induced by adding tetracycline (1 μg/ml). After incubation at 37°C for 3 days, resazurin (20 μl of 0.125-μg/ml stock in phosphate buffer saline [PBS]) (Sigma Aldrich) was added to each well and the plates incubated for a further 6–8 hr. 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 change in fluorescence resulting from the reduction of resazurin is proportional to the number of live cells. The EC_{50} (a concentration of compound that inhibited parasite growth by 50%) value was established using the nonlinear 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). Controls assessing the susceptibility of wild type and recombinant *T. brucei* towards nonnitroheterocyclic compounds (e.g., tetracycline, melarsprol, G418, and eflornithine) or DNA damaging agents (e.g., UV light, mechlorethamine, phleomycin, hydroxyurea, and methyl methanesulfonate) have been previously reported (Bot et al., 2013; Dattani & Wilkinson, 2019; Meredith et al., 2017; Wilkinson et al., 2008).

To assess mammalian cell cytotoxicity, THP-1 cells (2000 cells per well) were treated with growth medium containing different concentrations of nitroimidazole. After incubation at 37°C for 5 days, resazurin (20 μl of 0.125-μg/ml stock in PBS) (Sigma Aldrich) was added to each well and the plates incubated for a further 8 hr. The cell density of each culture was determined as described above and the EC_{50} value established.

ACKNOWLEDGMENTS

We thank Prof Richard McCulloch and Dr Eva Gluenz (both University of Glasgow) for the anti-*T. brucei* γH2A antisera or CRISPR-competent *T. brucei* lines, respectively. Ambika Dattani and Isatou Drammah were funded by the Queen Mary University of London and Islamic Development Bank Studentships, respectively.

ORCID

Shane R. Wilkinson https://orcid.org/0000-0002-8474-1943

REFERENCES

Alsford, S., Eckert, S., Baker, N., Glover, L., Sanchez-Flores, A., Leung, K.F. et al. (2012) High-throughput decoding of antitrypanosomal drug efficacy and resistance. *Nature*, *482*, 232–236. https://doi.org/10.1038/nature10771

Alsford, S., Kawahara, T., Glover, L. & Horn, D. (2005) Tagging a *T. brucei RRNA* locus improves stable transfection efficiency and circumvents inducible expression position effects. *Molecular and Biochemical Parasitology*, *144*, 142–148. https://doi.org/10.1016/j.molbiopara.2005.08.009

Babokhov, P., Sanyaolu, A.O., Oyibo, W.A., Fagbenro-Beyioku, A.F. & Iriemenam, N.C. (2013) A current analysis of chemotherapy strategies for the treatment of human African trypanosomiasis. *Pathogens and Global Health*, *107*, 242–252. https://doi.org/10.1179/2047773213Y.0000000105

Batenburg, N.L., Thompson, E.L., Hendrickson, E.A. & Zhu, X.D. (2015) Cockayne syndrome group B protein regulates DNA double-strand break repair and checkpoint activation. *EMBO Journal*, *34*, 1399–1416.

Beneke, T., Madden, R., Makin, L., Valli, J., Sunter, J. & Gluenz, E. (2017) A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. *Royal Society Open Science*, *4*, 170095. https://doi.org/10.1098/rsos.170095

Bennett, C., Straily, A., Haselow, D., Weinstein, S., Taffner, R., Yaglom, H. et al. (2018) Chagas disease surveillance activities – seven states, 2017. *MMWR. Morbidity and Mortality Weekly Report*, *67*, 738–741. https://doi.org/10.15585/mmwr.mm6726a2

Blumenstiel, K., Schoneck, R., Yardley, V., Croft, S.L. & Krauth-Siegel, R.L. (1999) Nitrofuran drugs as common subversive substrates of *Trypanosoma cruzi* lipoamide dehydrogenase and trypanothione reductase. *Biochemical Pharmacology*, *58*, 1791–1799. https://doi.org/10.1016/S0006-2952(99)00264-6

Bot, C., Hall, B.S., Álvarez, G., Di Maio, R., González, M., Cerecetto, H. et al. (2013) Evaluating 5-nitrofurans as trypanocidal agents. *Antimicrobial Agents Chemotherapy*, *57*, 1638–1647.

Campos, M.C., Leon, L.L., Taylor, M.C. & Kelly, J.M. (2014) Benznidazole-resistance in *Trypanosoma cruzi*: evidence that distinct mechanisms can act in concert. *Molecular and Biochemical Parasitology*, *193*, 17–19. https://doi.org/10.1016/j.molbiopara.2014.01.002

Campos, M.C., Phelan, J., Francisco, A.F., Taylor, M.C., Lewis, M.D., Pain, A. et al. (2017) Genome-wide mutagenesis and multi-drug resistance in American trypanosomes induced by the front-line drug benznidazole. *Scientific Reports*, *7*, 14407. https://doi.org/10.1038/s41598-017-14986-6

Cannan, W.J. & Pederson, D.S. (2016) Mechanisms and consequences of double-strand DNA break formation in chromatin. *Journal of Cellular Physiology*, *231*, 3–14. https://doi.org/10.1002/jcp.25048

Conway, C., Proudfoot, C., Burton, P., Barry, J.D. & McCulloch, R. (2002) Two pathways of homologous recombination in *Trypanosoma brucei*. *Molecular Microbiology*, *45*, 1687–1700.

Dattani, A. & Wilkinson, S.R. (2019) Deciphering the interstrand crosslink DNA repair network expressed by *Trypanosoma brucei*. *DNA Repair (Amst)*, *78*, 154–166. https://doi.org/10.1016/j.dnarep.2019.04.009

Deeks, E.D. (2019) Fexinidazole: first global approval. *Drugs*, *79*, 215–220. https://doi.org/10.1007/s40265-019-1051-6

Diaz de Toranzo, E.G., Castro, J.A., Franke de Cazzulo, B.M. & Cazzulo, J.J. (1998) Interaction of benznidazole reactive metabolites with nuclear and kinetoplastic DNA, proteins and lipids from *Trypanosoma cruzi*. *Experientia*, *44*, 880–881. https://doi.org/10.1007/BF01941187

Diaz-Viraque, F., Chiribao, M.L., Trochine, A., Gonzalez-Herrera, F., Castillo, C., Liempi, A. et al. (2018) Old yellow enzyme from *Trypanosoma cruzi* exhibits *in vivo* prostaglandin F2alpha synthase activity and has a key role in parasite infection and drug susceptibility. *Frontiers in Immunology*, *9*, 456.

Docampo, R. & Stoppani, A.O. (1979) Generation of superoxide anion and hydrogen peroxide induced by nifurtimox in *Trypanosoma cruzi*. *Archives of Biochemistry and Biophysics*, *197*, 317–321. https://doi.org/10.1016/0003-9861(79)90251-0

Dufernez, F., Yernaux, C., Gerbod, D., Noel, C., Chauvenet, M., Wintjens, R. et al. (2006) The presence of four iron-containing superoxide dismutase isozymes in trypanosomatidae: characterization, subcellular localization, and phylogenetic origin in *Trypanosoma brucei*. *Free Radical Biology and Medicine*, *40*, 210–225. https://doi.org/10.1016/j.freeradbiomed.2005.06.021

Edwards, D.I. (1993) Nitroimidazole drugs–action and resistance mechanisms. I. Mechanisms of action. *Journal of Antimicrobial Chemotherapy*, *31*, 9–20.

Ferreira, R.C. & Ferreira, L.C. (1986) Mutagenicity of nifurtimox and benznidazole in the *Salmonella*/microsome assay. *Brazilian Journal of Medical and Biological Research*, *19*, 19–25.

Franco, J.R., Cecchi, G., Priotto, G., Paone, M., Diarra, A., Grout, L. et al. (2020) Monitoring the elimination of human African trypanosomiasis at continental and country level: update to 2018. *PLoS Neglected Tropical Diseases*, *14*, e0008261. https://doi.org/10.1371/journal.pntd.0008261

Garcia-Huertas, P., Mejia-Jaramillo, A.M., Machado, C.R., Guimaraes, A.C. & Triana-Chavez, O. (2017) Prostaglandin F2alpha synthase in *Trypanosoma cruzi* plays critical roles in oxidative stress and susceptibility to benznidazole. *Royal Society Open Science*, *4*, 170773.

Glover, L. & Horn, D. (2012) Trypanosomal histone γH2A and the DNA damage response. *Molecular and Biochemical Parasitology*, *183*, 78–83. https://doi.org/10.1016/j.molbiopara.2012.01.008

Goijman, S.G., Frasch, A.C. & Stoppani, A.O. (1985) Damage of *Trypanosoma cruzi* deoxyribonucleic acid by nitroheterocyclic drugs. *Biochemical Pharmacology*, *24*, 1457–1491. https://doi.org/10.1016/0006-2952(85)90684-7

Goijman, S.G. & Stoppani, A.O. (1985) Effects of nitroheterocyclic drugs on macromolecule synthesis and degradation in *Trypanosoma cruzi*. *Biochemical Pharmacology*, *34*, 1331–1336. https://doi.org/10.1016/0006-2952(85)90514-3

Gomes Passos Silva, D., da Silva Santos, S., Nardelli, S.C., Mendes, I.C., Freire, A.C.G., Repolês, B.M. et al. (2018) The in vivo and in vitro roles of *Trypanosoma cruzi* Rad51 in the repair of DNA double strand breaks and oxidative lesions. *PLoS Neglected Tropical Diseases*, *12*, e0006875. https://doi.org/10.1371/journal.pntd.0006875

Haince, J.F., McDonald, D., Rodrigue, A., Déry, U., Masson, J.Y., Hendzel, M.J. et al. (2008) PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites. *Journal of Biological Chemistry*, *283*, 1197–1208. https://doi.org/10.1074/jbc.M706734200

Hall, B.S., Bot, C. & Wilkinson, S.R. (2011) Nifurtimox activation by trypanosomal type I nitroreductases generates cytotoxic nitrile metabolites. *Journal of Biological Chemistry*, *286*, 13088–13095. https://doi.org/10.1074/jbc.M111.230847

Hall, B.S. & Wilkinson, S.R. (2012) Activation of benznidazole by trypanosomal type I nitroreductases results in glyoxal formation. *Antimicrobial Agents and Chemotherapy*, *56*, 115–123. https://doi.org/10.1128/AAC.05135-11

Hirumi, H. & Hirumi, K. (1989) Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *Journal of Parasitology*, *75*, 985–989. https://doi.org/10.2307/3282883

Kedderis, G.L., Argenbright, L.S. & Miwa, G.T. (1989) Covalent interaction of 5-nitroimidazoles with DNA and protein in vitro: mechanism of reductive activation. *Chemical Research in Toxicology*, *2*, 146–149. https://doi.org/10.1021/tx00009a004

Kubata, B.K., Kabututu, Z., Nozaki, T., Munday, C.J., Fukuzumi, S., Ohkubo, K. et al. (2002) A key role for old yellow enzyme in the metabolism of drugs by *Trypanosoma cruzi*. *Journal of Experimental Medicine*, *196*, 1241–1251. https://doi.org/10.1084/jem.20020885

Lee, B.Y., Bacon, K.M., Bottazzi, M.E. & Hotez, P.J. (2013) Global economic burden of Chagas disease: a computational simulation model. *The Lancet Infectious Diseases*, *13*, 342–348. https://doi.org/10.1016/S1473-3099(13)70002-1

Liu, Y., Matsumoto, M., Ishida, H., Ohguro, K., Yoshitake, M., Gupta, R. et al. (2018) Delamanid: from discovery to its use for pulmonary multidrug-resistant tuberculosis (MDR-TB). *Tuberculosis (Edinb)*, *111*, 20–30. https://doi.org/10.1016/j.tube.2018.04.008

Machado, C.R., Vieira-da-Rocha, J.P., Mendes, I.C., Rajao, M.A., Marcello, L., Bitar, M. et al. (2014) Nucleotide excision repair in *Trypanosoma brucei*: specialization of transcription-coupled repair due to multigenic transcription. *Molecular Microbiology*, *92*, 756–776.

Marin, P.A., da Silva, M.S., Pavani, R.S., Machado, C.R. & Elias, M.C. (2018) Recruitment kinetics of the homologous recombination pathway in procyclic forms of *Trypanosoma brucei* after ionizing radiation treatment. *Scientific Reports*, *8*, 5405.

Mateo, H., Marin, C., Perez-Cordon, G. & Sánchez-Moreno, M. (2008) Purification and biochemical characterization of four iron superoxide dismutases in *Trypanosoma cruzi*. *Memorias do Instituto Oswaldo Cruz*, *103*, 271–276. https://doi.org/10.1590/S0074-02762008000300008

Mejia, A.M., Hall, B.S., Taylor, M.C., Gomez-Palacio, A., Wilkinson, S.R., Triana-Chavez, O. et al. (2012) Benznidazole-resistance in *Trypanosoma cruzi* is a readily acquired trait that can arise independently in a single population. *Journal of Infectious Diseases*, *206*, 220–228. https://doi.org/10.1093/infdis/jis331

Mejia-Jaramillo, A.M., Fernandez, G.J., Palacio, L. & Triana-Chavez, O. (2011) Gene expression study using real-time PCR identifies an NTR gene as a major marker of resistance to benzonidazole in *Trypanosoma cruzi*. *Parasit Vectors*, *4*, 169.

Menoni, H., Hoeijmakers, J.H. & Vermeulen, W. (2012) Nucleotide excision repair-initiating proteins bind to oxidative DNA lesions in vivo. *Journal of Cell Biology*, *199*, 1037–1046. https://doi.org/10.1083/jcb.201205149

Menoni, H., Wienholz, F., Theil, A.F., Janssens, R.C., Lans, H., Campalans, A. et al. (2018) The transcription-coupled DNA repair-initiating protein CSB promotes XRCC1 recruitment to oxidative DNA damage. *Nucleic Acids Research*, *46*, 7747–7756.

Meredith, E.M., Kumar, A., Konno, A., Szular, J., Alsford, S., Seifert, K. et al. (2017) Distinct activation mechanisms trigger the trypanocidal activity of DNA damaging prodrugs. *Molecular Microbiology*, *106*, 207–222. https://doi.org/10.1111/mmi.13767

Mesu, V., Kalonji, W.M., Bardonneau, C., Mordt, O.V., Blesson, S., Simon, F. et al. (2018) Oral fexinidazole for late-stage African *Trypanosoma brucei gambiense* trypanosomiasis: a pivotal multicentre, randomised, non-inferiority trial. *Lancet*, *391*, 144–154. https://doi.org/10.1016/S0140-6736(17)32758-7

Migchelsen, S.J., Buscher, P., Hoepelman, A.I., Schallig, H.D. & Adams, E.R. (2011) Human African trypanosomiasis: a review of non-endemic cases in the past 20 years. *International Journal of Infectious Diseases*, *15*, e517–e524. https://doi.org/10.1016/j.ijid.2011.03.018

Moreno, S.N.J., Docampo, R., Mason, R.P., Leon, W. & Stoppani, A. (1982) Different behaviors of benznidazole as free radical generator with mammalian and *Trypanosoma cruzi* microsomal preparations. *Archives of Biochemistry and Biophysics*, *218*, 585–591. https://doi.org/10.1016/0003-9861(82)90383-6

Passos-Silva, D.G., Rajao, M.A., Nascimento de Aguiar, P.H., Vieira-da-Rocha, J.P., Machado, C.R. & Furtado, C. (2010) Overview of DNA repair in *Trypanosoma cruzi*, *Trypanosoma brucei*, and *Leishmania major*. *Journal of Nucleic Acids*, 1–14. https://doi.org/10.4061/2010/840768

Patterson, S. & Fairlamb, A.H. (2019) Current and future prospects of nitro-compounds as drugs for trypanosomiasis and leishmaniasis. *Current Medicinal Chemistry*, *26*, 4454–4475. https://doi.org/10.2174/0929867325666180426164352

Patterson, S. & Wyllie, S. (2014) Nitro drugs for the treatment of trypanosomatid diseases: past, present, and future prospects. *Trends in Parasitology*, *30*, 289–298. https://doi.org/10.1016/j.pt.2014.04.003

Perez-Molina, J.A. & Molina, I. (2018) Chagas disease. *Lancet*, *391*, 82–94. https://doi.org/10.1016/S0140-6736(17)31612-4

Prathalingham, S.R., Wilkinson, S.R., Horn, D. & Kelly, J.M. (2007) Deletion of the *Trypanosoma brucei* superoxide dismutase gene sodb1 increases sensitivity to nifurtimox and benznidazole. *Antimicrobial Agents and Chemotherapy*, *51*, 755–758. https://doi.org/10.1128/AAC.01360-06

Rajao, M.A., Furtado, C., Alves, C.L., Passos-Silva, D.G., de Moura, M.B., Schamber-Reis, B.L. et al. (2014) Unveiling benznidazole's mechanism of action through overexpression of DNA repair proteins in *Trypanosoma cruzi*. *Environmental and Molecular Mutagenesis*, *55*, 309–321.

Requena-Mendez, A., Moore, D.A., Subira, C. & Munoz, J. (2016) Addressing the neglect: Chagas disease in London, UK. *The Lancet Global Health*, *4*, e231–e233. https://doi.org/10.1016/S2214-109X(16)00047-4

Robinson, N.P., McCulloch, R., Conway, C., Browitt, A. & Barry, J.D. (2002) Inactivation of Mre11 does not affect VSG gene duplication mediated by homologous recombination in *Trypanosoma brucei*. *Journal of Biological Chemistry*, *277*, 26185–26193. https://doi.org/10.1074/jbc.M203205200

Schoneck, R., Billaut-Mulot, O., Numrich, P., Ouaissi, M.A. & Krauth-Siegel, R.L. (1997) Cloning, sequencing and functional expression of dihydrolipoamide dehydrogenase from the human pathogen *Trypanosoma cruzi*. *European Journal of Biochemistry*, *243*, 739–747. https://doi.org/10.1111/j.1432-1033.1997.00739.x

Schumann Burkard, G., Jutzi, P. & Roditi, I. (2011) Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters. *Molecular and Biochemical Parasitology*, *175*, 91–94. https://doi.org/10.1016/j.molbiopara.2010.09.002

Shaw, A.P., Cecchi, G., Wint, G.R., Mattioli, R.C. & Robinson, T.P. (2014) Mapping the economic benefits to livestock keepers from intervening against bovine trypanosomosis in Eastern Africa. *Preventive Veterinary Medicine*, *113*, 197–210.

Siegel, T.N., Hekstra, D.R. & Cross, G.A. (2008) Analysis of the *Trypanosoma brucei* cell cycle by quantitative DAPI imaging. *Molecular and Biochemical Parasitology*, *160*, 171–174. https://doi.org/10.1016/j.molbiopara.2008.04.004

Simarro, P.P., Franco, J.R., Cecchi, G., Paone, M., Diarra, A., Ruiz Postigo, J.A. et al. (2012) Human African trypanosomiasis in non-endemic countries (2000–2010). *Journal of Travel Medicine*, *19*, 44–53. https://doi.org/10.1111/j.1708-8305.2011.00576.x

Stevnsner, T., Muftuoglu, M., Aamann, M.D. & Bohr, V.A. (2008) The role of Cockayne syndrome group B (CSB) protein in base excision repair and aging. *Mechanisms of Ageing and Development*, *129*, 441–448. https://doi.org/10.1016/j.mad.2008.04.009

Sullivan, J.A., Tong, J.L., Wong, M., Kumar, A., Sarkar, H., Ali, S. et al. (2015) Unravelling the role of SNM1 in the DNA repair system of *Trypanosoma brucei*. *Molecular Microbiology*, *96*, 827–838.

Tan, K.S., Leal, S.T. & Cross, G.A. (2002) *Trypanosoma brucei* MRE11 is non-essential but influences growth, homologous recombination and DNA double-strand break repair. *Molecular and Biochemical Parasitology*, *125*, 11–21.

Temperton, N.J., Wilkinson, S.R., Meyer, D.J. & Kelly, J.M. (1998) Overexpression of superoxide dismutase in *Trypanosoma cruzi* results in increased sensitivity to the trypanocidal agents gentian violet and benznidazole. *Molecular and Biochemical Parasitology*, *96*, 167–176. https://doi.org/10.1016/S0166-6851(98)00127-3

Thakare, R., Dasgupta, A. & Chopra, S. (2020) Pretomanid for the treatment of pulmonary tuberculosis. *Drugs Today (Barc)*, *56*, 655–668. https://doi.org/10.1358/dot.2020.56.10.3161237

Trochine, A., Creek, D.J., Faral-Tello, P., Barrett, M.P. & Robello, C. (2014) Benznidazole biotransformation and multiple targets in *Trypanosoma cruzi* revealed by metabolomics. *PLoS Neglected Tropical Diseases*, *8*, e2844. https://doi.org/10.1371/journal.pntd.0002844

Vieira-da-Rocha, J.P., Passos-Silva, D.G., Mendes, I.C., Rocha, E.A., Gomes, D.A., Machado, C.R. et al. (2019) The DNA damage response is developmentally regulated in the African trypanosome. *DNA Repair (Amst)*, *73*, 78–90. https://doi.org/10.1016/j.dnarep.2018.11.005

Viode, C., Bettache, N., Cenas, N., Krauth-Siegel, R.L., Chauviere, G., Bakalara, N. et al. (1999) Enzymatic reduction studies of nitroheterocycles. *Biochemical Pharmacology*, *57*, 549–557. https://doi.org/10.1016/S0006-2952(98)00324-4

Wei, L., Levine, A.S. & Lan, L. (2016) Transcription-coupled homologous recombination after oxidative damage. *DNA Repair (Amst)*, *44*, 76–80.

WHO. (2015) Chagas disease in Latin America: an epidemiological update based on 2010 estimates. *Weekly Epidemiological Record*, *90*, 33–43.

Wilkinson, S.R., Bot, C., Kelly, J.M. & Hall, B.S. (2011) Trypanocidal activity of nitroaromatic prodrugs: current treatments and future perspectives. *Current Topics in Medicinal Chemistry*, *11*, 2072–2084.

Wilkinson, S.R. & Kelly, J.M. (2009) Trypanocidal drugs: mechanisms, resistance and new targets. *Expert Reviews in Molecular Medicine*, *11*, e31. https://doi.org/10.1017/S1462399409001252

Wilkinson, S.R., Prathalingam, S.R., Taylor, M.C., Ahmed, A., Horn, D. & Kelly, J.M. (2006) Functional characterisation of the iron superoxide dismutase gene repertoire in *Trypanosoma brucei*. *Free Radical Biology and Medicine*, *40*, 198–209. https://doi.org/10.1016/j.freeradbiomed.2005.06.022

Wilkinson, S.R., Taylor, M.C., Horn, D., Kelly, J.M. & Cheeseman, I. (2008) A mechanism for cross-resistance to nifurtimox and benznidazole in trypanosomes. *Proceedings of the National Academy of Sciences USA*, *105*, 5022–5027. https://doi.org/10.1073/pnas.0711014105

Woodward, R. & Gull, K. (1990) Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*. *Journal of Cell Science*, *95*, 49–57. https://doi.org/10.1242/jcs.95.1.49

Wyllie, S., Foth, B.J., Kelner, A., Sokolova, A.Y., Berriman, M. & Fairlamb, A.H. (2016) Nitroheterocyclic drug resistance mechanisms in *Trypanosoma brucei*. *Journal of Antimicrobial Chemotherapy*, *71*, 625–634.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

How to cite this article: Dattani A, Drammeh I, Mahmood A, Rahman M, Szular J, Wilkinson SR. Unraveling the antitrypanosomal mechanism of benznidazole and related 2-nitroimidazoles: From prodrug activation to DNA damage. *Molecular Microbiology*. 2021;116:674–689. https://doi.org/10.1111/mmi.14763

10. REFERENCES

Abdullah, U. B., McGouran, J. F., Brolih, S., Ptchelkine, D., El-Sagheer, A. H., Brown, T., & McHugh, P. J. (2017). RPA activates the XPF-ERCC1 endonuclease to initiate processing of DNA interstrand crosslinks. *The EMBO Journal*, *36*(14), 2047–2060. https://doi.org/10.15252/EMBJ.201796664

Aboussekhra, A., & Al-Sharif, I. S. (2005). Homologous recombination is involved in transcription-coupled repair of UV damage in Saccharomyces cerevisiae. The EMBO Journal, 24(11), 1999–2010. https://doi.org/10.1038/sj.emboj.7600665

Akhter, S., & Legerski, R. J. (2008). SNM1A acts downstream of ATM to promote the G1 cell cycle checkpoint. *Biochemical and Biophysical Research Communications*, *377*(1), 236–241. https://doi.org/10.1016/J.BBRC.2008.09.130

Alibu, Vincent P., Lilian Storm, Simon Haile, Christine Clayton, and David Horn. 2005. "A Doubly Inducible System for RNA Interference and Rapid RNAi Plasmid Construction in Trypanosoma Brucei." Molecular and Biochemical Parasitology 139(1):75–82. doi: https://doi.org/10.1016/j.molbiopara.2004.10.002.

Alirol, E., Schrumpf, D., Amici Heradi, J., Riedel, A., De Patoul, C., Quere, M., & Chappuis, F. (2013). Nifurtimox-eflornithine combination therapy for second-stage gambiense human African trypanosomiasis: Médecins Sans Frontières experience in the Democratic Republic of the Congo. Clinical Infectious Diseases, 56, 195–203. https://doi.org/10.1093/cid/cis886

Allerston, C. K., Lee, S. Y., Newman, J. A., Schofield, C. J., Mchugh, P. J., & Gileadi, O. (2015). The structures of the SNM1A and SNM1B/Apollo nuclease domains reveal a potential basis for their distinct DNA processing activities. Nucleic Acids Research, 43(22), 11047–11060. https://doi.org/10.1093/NAR/GKV1256

Angers, S., Li, T., Yi, X., MacCoss, M. J., Moon, R. T., & Zheng, N. (2006). Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. Nature, 443(7111), 590–593. https://doi.org/10.1038/nature05175

Alsford, Sam, Taemi Kawahara, Lucy Glover, and David Horn. 2005. "Tagging a T. Brucei RRNA Locus Improves Stable Transfection Efficiency and Circumvents Inducible Expression Position Effects." Molecular and Biochemical Parasitology 144(2):142–48. doi: https://doi.org/10.1016/j.molbiopara.2005.08.009.

Alsford, S., Glover, L., & Horn, D. (2005). Multiplex analysis of RNA interference defects in Trypanosoma brucei. *Molecular and Biochemical Parasitology*, *139*(1), 129–132. https://doi.org/10.1016/j.molbiopara.2004.11.001

Alsford, S., & Horn, D. (2008). Single-locus targeting constructs for reliable regulated RNAi and transgene expression in Trypanosoma brucei. *Molecular and Biochemical Parasitology*, *161*(1), 76–79. https://doi.org/10.1016/j.molbiopara.2008.05.006

Alsford, S., Turner, D. J., Obado, S. O., Sanchez-Flores, A., Glover, L., Berriman, M., Hertz-Fowler, C., & Horn, D. (2011). High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Research*, *21*(6), 915–924. https://doi.org/10.1101/GR.115089.110

Alsford, S., Eckert, S., Baker, N., Glover, L., Sanchez-Flores, A., Leung, K. F., Turner, D. J., Field, M. C., Berriman, M., & Horn, D. (2012). High-throughput decoding of antitrypanosomal drug efficacy and resistance. *Nature*, *482*(7384), 232–236. https://doi.org/10.1038/NATURE10771

Alsford, S., Kelly, J. M., Baker, N., & Horn, D. (2013). Genetic dissection of drug resistance in trypanosomes. *Parasitology*, *140*(12), 1478–1491. https://doi.org/10.1017/S003118201300022X

Alsford, S. (2014). Increased Trypanosoma brucei cathepsin-L activity inhibits human serum-mediated trypanolysis. *Microbial Cell (Graz, Austria)*, *1*(8), 270–272. https://doi.org/10.15698/MIC2014.08.162

Angers, S., Li, T., Yi, X., MacCoss, M. J., Moon, R. T., & Zheng, N. (2006). Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. Nature, 443(7111), 590–593. https://doi.org/10.1038/nature05175

Apelt, K., Lans, H., Schärer, O. D., & Luijsterburg, M. S. (2021). Nucleotide excision repair leaves a mark on chromatin: DNA damage detection in nucleosomes. Cellular and molecular life sciences : CMLS, 78(24), 7925–7942. https://doi.org/10.1007/s00018-021-03984-7

Araki, M., Masutani, C., Takemura, M., Uchida, A., Sugasawa, K., Kondoh, J., Ohkuma, Y., & Hanaoka, F. (2001). Centrosome Protein Centrin 2/Caltractin 1 Is Part of the Xeroderma Pigmentosum Group C Complex That Initiates Global Genome Nucleotide Excision Repair*. Journal of Biological Chemistry, 276, 18665–18672. https://doi.org/10.1074/jbc.M100855200

Araújo, S. J., Tirode, F., Coin, F., Pospiech, H., Syväoja, J. E., Stucki, M., Hübscher, U., Egly, J. M., & Wood, R. D. (2000). Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. Genes & Development, 14(3), 349–359. https://doi.org/10.1101/GAD.14.3.349

Arora, S., Heyza, J., Zhang, H., Kalman-Maltese, V., Tillison, K., Floyd, A. M., Chalfin, E. M., Bepler, G., & Patrick, S. M. (2016). Identification of small molecule inhibitors of ERCC1-XPF that inhibit DNA repair and potentiate cisplatin efficacy in cancer cells. *Oncotarget*, *7*(46), 75104–75117. https://doi.org/10.18632/oncotarget.12072

Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., Depledge, D. P., Fischer, S., Gajria, B., Gao, X., Gardner, M. J., Gingle, A., Grant, G., Harb, O. S., Heiges, M., Hertz-Fowler, C., Houston, R., Innamorato, F., Iodice, J., … Wang, H. (2009). TriTrypDB: A functional genomic resource for the Trypanosomatidae. Nucleic Acids Research, 38(SUPPL.1). https://doi.org/10.1093/NAR/GKP851

Audebert, M., Salles, B., & Calsou, P. (2004). Involvement of Poly (ADP-ribose) Polymerase-1 and XRCC1/DNA Ligase III in an Alternative Route for DNA Double-strand Breaks Rejoining. Journal of Biological Chemistry, 279(53), 55117–55126. https://doi.org/10.1074/JBC.M404524200

Bacchi, C. J. (2009). Chemotherapy of human African trypanosomiasis. Interdisciplinary Perspectives on Infectious Diseases, 2009, 195040. https://doi.org/10.1155/2009/195040Baddock, H. T., Yosaatmadja, Y., Newman, J. A., Schofield, C. J., Gileadi, O., & McHugh, P. J. (2020). The SNM1A DNA repair nuclease. *DNA Repair*, *95*. https://doi.org/10.1016/j.dnarep.2020.102941

Baddock, H. T., Yosaatmadja, Y., Newman, J. A., Schofield, C. J., Gileadi, O., & McHugh, P. J. (2020). The SNM1A DNA repair nuclease. DNA repair, 95, 102941. https://doi.org/10.1016/j.dnarep.2020.102941

Badjatia, N., Nguyen, T. N., Lee, J. H., & Günzl, A. (2013). Trypanosoma brucei harbours a divergent XPB helicase paralogue that is specialized in nucleotide excision repair and conserved among kinetoplastid organisms. Molecular Microbiology, 90(6), 1293–1308. https://doi.org/10.1111/MMI.12435

Bakari-Soale, M., Ikenga, N. J., Scheibe, M., Butter, F., Jones, N. G., Kramer, S., & Engstler, M. (2021). The nucleolar DExD/H protein Hel66 is involved in ribosome biogenesis in Trypanosoma brucei. Scientific Reports |, 11, 18325. https://doi.org/10.1038/s41598-021-97020-0

Baker, N., Alsford, S., & Horn, D. (2011). Genome-wide RNAi screens in African trypanosomes identify the nifurtimox activator NTR and the eflornithine transporter AAT6. *Molecular and Biochemical Parasitology*, *176*(1), 55–57. https://doi.org/10.1016/j.molbiopara.2010.11.010

Baker, N., de Koning, H. P., Mäser, P., & Horn, D. (2013). Drug resistance in African trypanosomiasis: the melarsoprol and pentamidine story. Trends in Parasitology, 29(3), 110–118. https://doi.org/10.1016/j.pt.2012.12.005

Bakkenist, C. J., & Kastan, M. B. (2003). DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. Nature, 421(6922), 499–506. https://doi.org/10.1038/NATURE01368

Bangs, J. D., Crain, P. F., Hashizume, T., McCloskey, J. A., & Boothroyd, J. C. (1992). Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. *Journal of Biological Chemistry*, *267*(14), 9805–9815. https://doi.org/10.1016/s0021-9258(19)50165-x

Barber, L. J., Ward, T. A., Hartley, J. A., & McHugh, P. J. (2005). DNA interstrand cross-link repair in the Saccharomyces cerevisiae cell cycle: overlapping roles for PSO2 (SNM1) with MutS factors and EXO1 during S phase. Molecular and cellular biology, 25(6), 2297-2309.

Barnes, R. L., & McCulloch, R. (2007). Trypanosoma brucei homologous recombination is dependent on substrate length and homology, though displays a differential dependence on mismatch repair as substrate length decreases. Nucleic Acids Research, 35(10), 3478–3493. https://doi.org/10.1093/NAR/GKM249

Barrett, M. P., Burchmore, R. J. S., Stich, A., Lazzari, J. O., Frasch, A. C., Cazzulo, J. J., & Krishna, S. (2003). The trypanosomiases. Lancet, 362(9394), 1469–1480. https://doi.org/10.1016/S0140-6736(03)14694-6

Basu, A. K., Marnett, L. J., & Romano, L. J. (1984). Dissociation of malondialdehyde mutagenicity in Salmonella typhimurium from its ability to induce interstrand DNA cross-links. Mutation Research, 129(1), 39–46. https://doi.org/10.1016/0027-5107(84)90121-0

Batenburg, N. L., Mersaoui, S. Y., Walker, J. R., Coulombe, Y., Hammond-Martel, I., Wurtele, H., Masson, J. Y., & Zhu, X. D. (2021). Cockayne syndrome group B protein regulates fork restart, fork progression and MRE11-dependent fork degradation in BRCA1/2-deficient cells. *Nucleic Acids Research*, *49*(22), 12836–12854. https://doi.org/10.1093/nar/gkab1173

Batenburg, N. L., Thompson, E. L., Hendrickson, E. A., & Zhu, X. (2015). Cockayne syndrome group B protein regulates DNA double-strand break repair and checkpoint activation. *The EMBO Journal*, *34*(10), 1399–1416. https://doi.org/10.15252/embj.201490041

Bauer, G. B., & Povirk, L. F. (1997). Specificity and kinetics of interstrand and intrastrand bifunctional alkylation by nitrogen mustards at a G-G-C sequence. Nucleic Acids Research, 25(6).

Bauer, N. C., Corbett, A. H., & Doetsch, P. W. (2015). The current state of eukaryotic DNA base damage and repair. Nucleic Acids Research, 43(21), 10083–10101. https://doi.org/10.1093/NAR/GKV1136

Baumann, P., & West, S. C. (1998). Role of the human RAD51 protein in homologous recombination and double-stranded-break repair. Trends in Biochemical Sciences, 23(7), 247–251. https://doi.org/10.1016/S0968-0004(98)01232-8

Beernink, P. T., Segelke, B. W., Hadi, M. Z., Erzberger, J. P., Wilson, D. M., & Rupp, B. (2001). Two divalent metal ions in the active site of a new crystal form of human apurinic/apyrimidinic endonuclease, ape1: implications for the catalytic mechanism. Journal of Molecular Biology, 307(4), 1023–1034. https://doi.org/10.1006/JMBI.2001.4529

Bernhard, S. C., Nerima, B., Mäser, P., & Brun, R. (2007). Melarsoprol- and pentamidine-resistant Trypanosoma brucei rhodesiense populations and their cross-resistance. International journal for parasitology, 37(13), 1443–1448. https://doi.org/10.1016/j.ijpara.2007.05.007

Beneke, T., Madden, R., Makin, L., Valli, J., Sunter, J., & Gluenz, E. (2017). A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. Royal Society Open Science, 4(5), 1–16. https://doi.org/10.1098/rsos.170095

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., Böhme, U., Hannick, L., Aslett, M. A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U. C. M., Arrowsmith, C., … El-Sayed, N. M. (2005). The genome of the African trypanosome Trypanosoma brucei. Science, 309(5733), 416–422. https://doi.org/10.1126/science.1112642

Beverley, S. M., & Clayton, C. E. (1993). Transfection of Leishmania and Trypanosoma brucei by electroporation. *Methods in molecular biology (Clifton, N.J.)*, *21*, 333–348. https://doi.org/10.1385/0-89603-239-6:333

Bonnet, J., Boudot, C., & Courtioux, B. (2015). Overview of the Diagnostic Methods Used in the Field for Human African Trypanosomiasis: What Could Change in the Next Years? *BioMed Research International*, *2015*. https://doi.org/10.1155/2015/583262

Bot, C., Hall, B. S., Bashir, N., Taylor, M. C., Helsby, N. A., & Wilkinson, S. R. (2010). Trypanocidal activity of aziridinyl nitrobenzamide prodrugs. *Antimicrobial Agents and Chemotherapy*, *54*(10), 4246–4252. https://doi.org/10.1128/AAC.00800-10

Bowles, M., Lally, J., Fadden, A. J., Mouilleron, S., Hammonds, T., & Mcdonald, N. Q. (2012). Fluorescence-based incision assay for human XPF-ERCC1 activity identifies important elements of DNA junction recognition. *Nucleic Acids Research*, *40*(13). https://doi.org/10.1093/NAR/GKS284

Bradsher, J., Auriol, J., de Santis, L. P., Iben, S., Vonesch, J. L., Grummt, I., & Egly, J. M. (2002). CSB is a component of RNA pol I transcription. *Molecular Cell*, *10*(4), 819–829. https://doi.org/10.1016/S1097-2765(02)00678-0

Brun, R., Blum, J., Chappuis, F., & Burri, C. (2010). Human African trypanosomiasis. *The Lancet*, *375*(9709), 148–159. https://doi.org/10.1016/S0140-6736(09)60829-1

Burchmore, R. J., Ogbunude, P. O., Enanga, B., & Barrett, M. P. (2002). Chemotherapy of human African trypanosomiasis. *Current pharmaceutical design*, *8*(4), 256–267. https://doi.org/10.2174/1381612023396159

Bugreev, D. V, & Mazin, A. V. (2004). Ca 2 activates human homologous recombination protein Rad51 by modulating its ATPase activity. www.pnas.orgcgidoi10.1073pnas.0402105101

Burkard, Gabriela, Cristina M. Fragoso, and Isabel Roditi. 2007. "Highly Efficient Stable Transformation of Bloodstream Forms of Trypanosoma Brucei." Molecular and Biochemical Parasitology 153(2):220–23. doi: 10.1016/J.MOLBIOPARA.2007.02.008.

Burri, C. (2010). Chemotherapy against human African trypanosomiasis: is there a road to success? Parasitology, 137(14), 1987–1994. https://doi.org/10.1017/S0031182010001137

Burton, P., McBride, D. J., Wilkes, J. M., Barry, J. D., & McCulloch, R. (2007). Ku heterodimer-independent end joining in Trypanosoma brucei cell extracts relies upon sequence microhomology. *Eukaryotic Cell*, *6*(10), 1773–1781. https://doi.org/10.1128/EC.00212-07

Buschini, A., Ferrarini, L., Franzoni, S., Galati, S., Lazzaretti, M., Mussi, F., Albuquerque, C. N. de, Zucchi, T. M. A. D., & Poli, P. (2009). Genotoxicity Revaluation of Three Commercial Nitroheterocyclic Drugs: Nifurtimox, Benznidazole, and Metronidazole. Journal of Parasitology Research, 2009, 1–11. https://doi.org/10.1155/2009/463575

Buzon, B., Grainger, R. A., Rzadki, C., Huang, S. Y. M., & Junop, M. (2021). Identification of Bioactive SNM1A Inhibitors. *ACS Omega*, *6*(14), 9352–9361. https://doi.org/10.1021/acsomega.0c03528

Caldecott, K. W. (2019). XRCC1 protein; Form and function. DNA Repair, 81, 102664. https://doi.org/10.1016/J.DNAREP.2019.102664

Caldecott, K. W., Tucker, J. D., Stanker, L. H., & Thompson, L. H. (1995). Characterization of the XRCC1-DNA ligase III complex in vitro and its absence from mutant hamster cells. Nucleic Acids Research, 23(23), 4836–4843. https://doi.org/10.1093/NAR/23.23.4836

Cameroni, E., Stettler, K., & Suter, B. (2010). On the traces of XPD: cell cycle matters - untangling the genotype-phenotype relationship of XPD mutations. Cell Division 2010 5:1, 5(1), 1–19. https://doi.org/10.1186/1747-1028-5-24

Campos, M. C., Phelan, J., Francisco, A. F., Taylor, M. C., Lewis, M. D., Pain, A., Clark, T. G., & Kelly, J. M. (2017). Genome-wide mutagenesis and multi-drug resistance in American trypanosomes induced by the front-line drug benznidazole. Scientific Reports 2017 7:1, 7(1), 1–8. https://doi.org/10.1038/s41598-017-14986-6

Carroll, B. L., Zahn, K. E., Hanley, J. P., Wallace, S. S., Dragon, J. A., & Doublié, S. (2021). Caught in motion: human NTHL1 undergoes interdomain rearrangement necessary for catalysis. Nucleic Acids Research, 49(22), 13165–13178. https://doi.org/10.1093/nar/gkab1162

Carvajal-Maldonado, D., Drogalis Beckham, L., Wood, R. D., & Doublié, S. (2021). When DNA Polymerases Multitask: Functions Beyond Nucleotidyl Transfer. Frontiers in Molecular Biosciences, 8, 815845. https://doi.org/10.3389/fmolb.2021.815845

Castillo-Acosta, V. M., Aguilar-Pereyra, F., Vidal, A. E., Navarro, M., Ruiz-Pérez, L. M., & González-Pacanowska, D. (2012). Trypanosomes lacking uracil-DNA glycosylase are hypersensitive to antifolates and present a mutator phenotype. International Journal of Biochemistry and Cell Biology, 44(9), 1555–1568. https://doi.org/10.1016/j.biocel.2012.06.014

Cattell, E., Sengerová, B., & McHugh, P. J. (2010). The SNM1/Pso2 family of ICL repair nucleases: From yeast to man. Environmental and Molecular Mutagenesis, 51(6), 635–645. https://doi.org/10.1002/em.20556

Cayla, M., Rojas, F., Silvester, E., Venter, F., & Matthews, K. R. (2019). African trypanosomes. Parasites & Vectors, 12(1), 190. https://doi.org/10.1186/S13071-019-3355-5

Chang, H. H. Y., Pannunzio, N. R., Adachi, N., & Lieber, M. R. (2017). Non-homologous DNA end joining and alternative pathways to double-strand break repair. In Nature Reviews Molecular Cell Biology (Vol. 18, Issue 8, pp. 495–506). Nature Publishing Group. https://doi.org/10.1038/nrm.2017.48

Charret, K. S., Requena, C. E., Castillo-Acosta, V. M., Ruiz-Pérez, L. M., González-Pacanowska, D., & Vidal, A. E. (2012). Trypanosoma brucei AP endonuclease 1 has a major role in the repair of abasic sites and protection against DNA-damaging agents. DNA Repair, 11(1), 53–64. https://doi.org/10.1016/J.DNAREP.2011.10.006

Checchi, F., Piola, P., Ayikoru, H., Thomas, F., Legros, D., & Priotto, G. (2007). Nifurtimox plus eflornithine for late-stage sleeping sickness in Uganda: A case series. PLoS Neglected Tropical Diseases, 1(2), 1–6. https://doi.org/10.1371/journal.pntd.0000064

Chi, P., Van Komen, S., Sehorn, M. G., Sigurdsson, S., & Sung, P. (2006). Roles of ATP binding and ATP hydrolysis in human Rad51 recombinase function. DNA Repair, 5(3), 381–391. https://doi.org/10.1016/J.DNAREP.2005.11.005

Christmann, M., Tomicic, M. T., Roos, W. P., & Kaina, B. (2003). Mechanisms of human DNA repair: An update. Toxicology, 193(1–2), 3–34. https://doi.org/10.1016/S0300-483X(03)00287-7

Chu, G., & Chang, E. (1988). Xeroderma pigmentosum group E cells lack a nuclear factor that binds to damaged DNA. Science (New York, N.Y.), 242(4878), 564–567. https://doi.org/10.1126/science.3175673

Caldecott, K. W. (2020). Mammalian DNA base excision repair: Dancing in the moonlight. *DNA Repair*, *93*. https://doi.org/10.1016/J.DNAREP.2020.102921

Charret, K. S., Requena, C. E., Castillo-Acosta, V. M., Ruiz-Pérez, L. M., González-Pacanowska, D., & Vidal, A. E. (2012). Trypanosoma brucei AP endonuclease 1 has a major role in the repair of abasic sites and protection against DNA-damaging agents. *DNA Repair*, *11*(1), 53–64. https://doi.org/10.1016/J.DNAREP.2011.10.006

Clayton, C. E. (1999). Genetic manipulation of kinetoplastida. Parasitology Today, 15(9), 372–378. https://doi.org/10.1016/S0169-4758(99)01498-2

Clayton C. E. (2002). Life without transcriptional control? From fly to man and back again. The EMBO journal, 21(8), 1881–1888. https://doi.org/10.1093/emboj/21.8.1881

Clayton, C. (2019). Regulation of gene expression in trypanosomatids: living with polycistronic transcription. *Open Biology*, *9*(6). https://doi.org/10.1098/RSOB.190072

Clayton, C. E. (2016). Gene expression in Kinetoplastids. *Current Opinion in Microbiology*, *32*, 46–51. https://doi.org/10.1016/j.mib.2016.04.018

Connor, T. R. O., & Laval, J. (1989). Physical association of the 2, 6-diamino-4-hydroxy-5N-formamidopyrimidine-DNA glycosylase of Escherichia coli and an activity nicking DNA at apurinic / apyrimidinic sites. 86(July), 5222–5226. https://doi.org/10.1073/pnas.86.14.5222

Conway, C., Proudfoot, C., Burton, P., Barry, J. D., & McCulloch, R. (2002). Two pathways of homologous recombination in Trypanosoma brucei. *Molecular Microbiology*, *45*(6), 1687–1700. https://doi.org/10.1046/J.1365-2958.2002.03122.X

Cooper, S., Wadsworth, E. S., Ochsenreiter, T., Ivens, A., Savill, N. J., & Schnaufer, A. (2019). Assembly and annotation of the mitochondrial minicircle genome of a differentiation-competent strain of Trypanosoma brucei. *Nucleic Acids Research*, *47*(21), 11304–11325. https://doi.org/10.1093/nar/gkz928

Costa, F. C., Francisco, A. F., Jayawardhana, S., Calderano, S. G., Lewis, M. D., Olmo, F., Beneke, T., Gluenz, E., Sunter, J., Dean, S., Kelly, J. M., & Taylor, M. C. (2018). Expanding the toolbox for Trypanosoma cruzi: A parasite line incorporating a bioluminescence-fluorescence dual reporter and streamlined CRISPR/Cas9 functionality for rapid in vivo localisation and phenotyping. PLoS Neglected Tropical Diseases, 12(4). https://doi.org/10.1371/JOURNAL.PNTD.0006388

Cox, F. E. G. (2004). History of sleeping sickness (African trypanosomiasis). Infectious Disease Clinics of North America, 18(2), 231–245. https://doi.org/10.1016/J.IDC.2004.01.004

Cross, M., Kieft, R., Sabatini, R., Dirks-Mulder, A., Chaves, I., & Borst, P. (2002). J-binding protein increases the level and retention of the unusual base J in trypanosome DNA. Molecular microbiology, 46(1), 37–47. https://doi.org/10.1046/j.1365-2958.2002.03144.xDaee, D. L., & Myung, K. (2012). Fanconi-like crosslink repair in yeast. Genome integrity, 3(1), 7. https://doi.org/10.1186/2041-9414-3-7

Cunningham, R. P., & Weiss, B. (1985). Endonuclease III (nth) mutants of Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America, 82(2), 474–478. https://doi.org/10.1073/PNAS.82.2.474

Dasari, S., & Tchounwou, P. B. (2014). Cisplatin in cancer therapy: molecular mechanisms of action. https://doi.org/10.1016/j.ejphar.2014.07.025

da Silva, M. S. (2021). DNA Double-Strand Breaks: A Double-Edged Sword for Trypanosomatids. Frontiers in Cell and Developmental Biology, 9, 669041. https://doi.org/10.3389/fcell.2021.669041. DNA damage. Molecular Microbiology, 116(2), 674–689. https://doi.org/10.1111/MMI.14763

Dattani, Ambika, and Shane R. Wilkinson. 2019. "Deciphering the Interstrand Crosslink DNA Repair Network Expressed by Trypanosoma Brucei." DNA Repair 78(February):154–66. doi: 10.1016/j.dnarep.2019.04.009.

Dattani, A., Drammeh, I., Mahmood, A., Rahman, M., Szular, J., & Wilkinson, S. R. (2021). Unraveling the antitrypanosomal mechanism of benznidazole and related 2-nitroimidazoles: From prodrug activation to DNA damage. Molecular Microbiology 116, 674-689. https://doi.org/10.1111/mmi.14763 (2021).

Davis, C. N., Rock, K. S., & Keeling, M. J. (n.d.). Human African trypanosomiasis: current status and eradication efforts. https://doi.org/10.1079/PAVSNNR202015028

Dean, S., Sunter, J., Wheeler, R. J., Hodkinson, I., Gluenz, E., & Gull, K. (2015). A toolkit enabling efficient, scalable, and reproducible gene tagging in trypanosomatids. Open Biology, 5(1). https://doi.org/10.1098/RSOB.140197

Dean, S., Sunter, J. D., & Wheeler, R. J. (2017). TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. Trends in Parasitology, 33(2), 80–82. https://doi.org/10.1016/j.pt.2016.10.009

de Almeida, L. C., Calil, F. A., Machado-Neto, J. A., & Costa-Lotufo, L. V. (2021). DNA damaging agents and DNA repair: From carcinogenesis to cancer therapy. *Cancer Genetics*, *252–253*, 6–24. https://doi.org/10.1016/j.cancergen.2020.12.002

Deeks, E. D. (2019). Fexinidazole: First Global Approval. Drugs, 79(2), 215–220. https://doi.org/10.1007/s40265-019-1051-6

Demple, B., & Harrison, L. (1994). Repair of oxidative damage to DNA: Enzymology and Biology. www.annualreviews.org

Derheimer, F. A., Hicks, J. K., Paulsen, M. T., Canman, C. E., & Ljungman, M. (2008). Psoralen-induced DNA interstrand cross-links block transcription and induce p53 in an ATR-dependent manner. Molecular Pharmacology Fast Forward, 8. https://doi.org/10.1124/mol.108.051698

Dexheimer, T. S. (2013). DNA Repair Pathways and Mechanisms BT - DNA Repair of Cancer Stem Cells (L. A. Mathews, S. M. Cabarcas, & E. M. Hurt (eds.); pp. 19–32). Springer Netherlands. https://doi.org/10.1007/978-94-007-4590-2_2

Díaz de Toranzo, E. G., Castro, J. A., Franke de Cazzulo, B. M., & Cazzulo, J. J. (1988). Interaction of benznidazole reactive metabolites with nuclear and kinetoplastic DNA, proteins, and lipids from Trypanosoma cruzi. Experientia, 44(10), 880–881. https://doi.org/10.1007/BF01941187

Dianov, G., & Lindahl, T. (1994). Reconstitution of the DNA base excision-repair pathway. Current biology: CB, 4(12), 1069–1076. https://doi.org/10.1016/s0960-9822(00)00245-1).

Dong, H., Nebert, D. W., Bruford, E. A., Thompson, D. C., Joenje, H., & Vasiliou, V. (2015). Update of the human and mouse Fanconi anemia genes. https://doi.org/10.1186/s40246-015-0054-y

Dorléans, A., Li de la Sierra-Gallay, I., Piton, J., Zig, L., Gilet, L., Putzer, H., & Condon, C. (2011). Molecular basis for the recognition and cleavage of RNA by the bifunctional 5'-3' exo/endoribonuclease RNase J. Structure (London, England : 1993), 19(9), 1252–1261. https://doi.org/10.1016/j.str.2011.06.018

Doyle, J. J., H. Hirumi, K. Hirumi, E. N. Lupton, and G. A. M. Cross. 1980. "Antigenic Variation in Clones of Animal-Infective Trypanosoma Brucei Derived and Maintained in Vitro." Parasitology 80(2):359–69. doi: DOI: 10.1017/S0031182000000810

Dronkert, M. L. G., de Wit, J., Boeve, M., Vasconcelos, M. L., van Steeg, H., Tan, T. L. R., Hoeijmakers, J. H. J., & Kanaar, R. (2000). Disruption of mouse SNM1 causes increased sensitivity to the DNA interstrand cross-linking agent mitomycin C. Molecular and Cellular Biology, 20(13), 4553–4561. https://doi.org/10.1128/MCB.20.13.4553-4561.2000

Durrant, J. D., Hall, L., Swift, R. V., Landon, M., Schnaufer, A., & Amaro, R. E. (2010). Novel naphthalene-based inhibitors of Trypanosoma brucei RNA editing ligase 1. PLoS neglected tropical diseases, 4(8), e803. https://doi.org/10.1371/journal.pntd.0000803

Eckert-Boulet, N., Rothstein, R., & Lisby, M. (2011). Chapter 30 Cell Biology of Homologous Recombination in Yeast. https://doi.org/10.1007/978-1-61779-129-1_30

Edfeldt, N. B. F., Harwood, E. A., Sigurdsson, S. T., Hopkins, P. B., & Reid, B. R. (2004). Sequence context effect on the structure of nitrous acid induced DNA interstrand cross-links. Nucleic Acids Research, 32(9), 2795–2801. https://doi.org/10.1093/NAR/GKH607

Eid, J., & Sollner-Webb, B. (1991). Stable integrative transformation of Trypanosoma brucei that occurs exclusively by homologous recombination. Proceedings of the National Academy of Sciences of the United States of America, 88(6), 2118. https://doi.org/10.1073/PNAS.88.6.2118

El-Khamisy, S. F., Saifi, G. M., Weinfeld, M., Johansson, F., Helleday, T., Lupski, J. R., & Caldecott, K. W. (2005). Defective DNA single-strand break repair in spinocerebellar ataxia with axonal neuropathy-1. Nature, 434(7029), 108–113. https://doi.org/10.1038/NATURE03314

El-Sayed, N. M., Myler, P. J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renauld, H., Worthey, E. A., Hertz-Fowler, C., Ghedin, E., Peacock, C., Bartholomeu, D. C., Haas, B. J., Tran, A. N., Wortman, J. R., Alsmark, U. C. M., Angiuoli, S., Anupama, A., … Hall, N. (2005). Comparative genomics of trypanosomatid parasitic protozoa. Science (New York, N.Y.), 309(5733). https://doi.org/10.1126/SCIENCE.1112181

El-Sayed, N. M., Myler, P. J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renauld, H., Worthey, E. A., Hertz-Fowler, C., Ghedin, E., Peacock, C., Bartholomeu, D. C., Haas, B. J., Tran, A. N., Wortman, J. R., Alsmark, U. C. M., Angiuoli, S., Anupama, A., … Hall, N. (2005). Comparative genomics of trypanosomatid parasitic protozoa. Science (New York, N.Y.), 309(5733). https://doi.org/10.1126/SCIENCE.1112181

Emmert, B., Hallier, E., Schön, M. P., & Emmert, S. (2011). [Xeroderma pigmentosum (XP) : A genetic disease sheds light on UV-induced skin cancer]. Der Hautarzt; Zeitschrift Fur Dermatologie, Venerologie, Und Verwandte Gebiete, 62(2), 91–97. https://doi.org/10.1007/S00105-010-2050-4

Enanga, B., Ariyanayagam, M. R., Stewart, M. L., & Barrett, M. P. (2003). Activity of megazol, a trypanocidal nitroimidazole, is associated with DNA damage. *Antimicrobial agents and chemotherapy*,*47*(10), 3368–3370. https://doi.org/10.1128/AAC.47.10.3368-3370.2003

Engstrom, L. M., Brinkmeyer, M. K., Ha, Y., Raetz, A. G., Hedman, B., Hodgson, K. O., Solomon, E. I., & David, S. S. (2014). A Zinc Linchpin Motif in the MUTYH Glycosylase Interdomain Connector Is Required for Efficient Repair of DNA Damage. https://doi.org/10.1021/ja502942d

European Medicines Agency. (2019). Fexinidazole Winthrop | European Medicines Agency: Opinion on medicine for use outside EU. European Medicines Agency. https://www.ema.europa.eu/en/opinion-medicine-use-outside-EU/human/fexinidazole-winthrop

Evans, E., Moggs, J. G., Hwang, J. R., Egly, J.-M., & Wood, R. D. (1997). Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. In The EMBO Journal (Vol. 16, Issue 21).

Fagbemi, A. F., Orelli, B., & Schärer, O. D. (2011). Regulation of endonuclease activity in human nucleotide excision repair. DNA Repair, 10(7), 722–729. https://doi.org/10.1016/J.DNAREP.2011.04.022

Fairlamb, H., Henderson, G. B., & Cerami, A (1989). Trypanothione is the primary target for arsenical drugs against African trypanosomes. Proceedings of the National Academy of Sciences of the United States of America, 86(8), 2607–2611.

Faridounnia, M., Folkers, G. E., & Boelens, R. (2018). Function and interactions of ERCC1-XPF in DNA damage response. Molecules, 23(12). https://doi.org/10.3390/molecules23123205

Feldberg, R. S., Lucas, J. L., & Dannenberg, A. (1982). A damage-specific DNA binding protein. Large scale purification from human placenta and characterization. The Journal of biological chemistry, 257(11), 6394–6401.

Ferreira, R. C. C., & Ferreira, L. C. S. (1986). Mutagenicity of nifurtimox and benznidazole in the Salmonella/microsome assay. Brazilian Journal of Medical and Biological Research = Revista Brasileira de Pesquisas Medicas e Biologicas, 19(1), 19–25. https://europepmc.org/article/med/3542090

Field, M. C., & Carrington, M. (2009). The trypanosome flagellar pocket. Nature reviews. Microbiology, 7(11), 775–786. https://doi.org/10.1038/nrmicro2221

Franco, J., Pere, S., Diarra, Ruiz-Postigo, Samo, & Jannin. (2012). Monitoring the use of nifurtimox-eflornithine combination therapy (NECT) in the treatment of second stage gambiense human African trypanosomiasis. Research and Reports in Tropical Medicine, 93. https://doi.org/10.2147/RRTM.S34399

Franco, J. R., Simarro, P. P., Diarra, A., & Jannin, J. G. (2014). Epidemiology of human African trypanosomiasis. *Clinical Epidemiology*, *6*(1), 257. https://doi.org/10.2147/CLEP.S39728

Franco, J. R., Cecchi, G., Priotto, G., Paone, M., Diarra, A., Grout, L., Simarro, P. P., Zhao, W., & Argaw, D. (2020). Monitoring the elimination of human African trypanosomiasis at continental and country level: Update to 2018. PLoS Neglected Tropical Diseases, 14(5), 1–18. https://doi.org/10.1371/JOURNAL.PNTD.0008261

Franco, J. R., Cecchi, G., Paone, M., Diarra, A., Grout, L., Kadima Ebeja, A., Simarro, P. P., Zhao, W., & Argaw, D. (2022). The elimination of human African trypanosomiasis: Achievements in relation to WHO road map targets for 2020. PLoS neglected tropical diseases, 16(1), e0010047. https://doi.org/10.1371/journal.pntd.0010047

Franco, J. R., Cecchi, G., Priotto, G., Paone, M., Diarra, A., Grout, L., Simarro, P. P., Zhao, W., & Argaw, D. (2020). Monitoring the elimination of human African trypanosomiasis at continental and country level: Update to 2018. *PLoS Neglected Tropical Diseases*, *14*(5), 1–18. https://doi.org/10.1371/JOURNAL.PNTD.0008261

Freudenthal, B. D., Beard, W. A., Cuneo, M. J., Dyrkheeva, N. S., & Wilson, S. H. (2015). Capturing snapshots of APE1 processing DNA damage. 22. https://doi.org/10.1038/nsmb.3105

Friedberg, E. C., Bardwell, A. J., Bardwell, L., Feaver, W. J., Kornberg, R. D., Svejstrup, J. Q., Tomkinson, A. E., & Wang, Z. (1995). Nucleotide excision repair in the yeast Saccharomyces cerevisiae: its relationship to specialized mitotic recombination and RNA polymerase II basal transcription. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 347(1319), 63–68. https://doi.org/10.1098/RSTB.1995.0010

Friedberg, E. C. (2006). The first comprehensive impact factor for DNA Repair. DNA Repair, 5(8), 873–874. https://doi.org/10.1016/j.dnarep.2006.06.004

Fromme, J. C., & Verdine, G. L. (2004). Base excision repair. Advances in Protein Chemistry, 69, 1–41. https://doi.org/10.1016/S0065-3233(04)69001-2

Frosina, G., Fortini, P., Rossi, O., Carrozzino, F., Raspaglio, G., Cox, L. S., Lane, D. P., Abbondandolo, A., & Dogliotti, E. (1996). Two pathways for base excision repair in mammalian cells. Journal of Biological Chemistry, 271(16), 9573–9578. https://doi.org/10.1074/jbc.271.16.9573

Gaillard, P. H. L., & Wood, R. D. (2001). Activity of individual ERCC1 and XPF subunits in DNA nucleotide excision repair. Nucleic Acids Research, 29(4), 872–879. https://doi.org/10.1093/nar/29.4.872

Genois, M.-M., Paquet, E. R., Laffitte, M.-C. N., Maity, R., Rodrigue, A., Ouellette, M., & Masson, J.-Y. (2014). DNA repair pathways in trypanosomatids: from DNA repair to drug resistance. Microbiology and Molecular Biology Reviews: MMBR, 78(1), 40–73. https://doi.org/10.1128/MMBR.00045-13

Gibson, W., & Stevens, J. (1999). Genetic exchange in the trypanosomatidae. *Advances in parasitology*, *43*, 1–46. https://doi.org/10.1016/s0065-308x(08)60240-7

Gilinger, G., & Bellofatto, V. (2001). Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. Nucleic Acids Research, 29(7), 1556–1564.

Glover, L., McCulloch, R., & Horn, D. (2008). Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. *Nucleic Acids Res*, *36*.

Glover, L., & Horn, D. (2009). Site-specific DNA double-strand breaks greatly increase stable transformation efficiency in Trypanosoma brucei. *Molecular and Biochemical Parasitology*, *166*(2), 194. https://doi.org/10.1016/J.MOLBIOPARA.2009.03.010

Glover, L., Jun, J., & Horn, D. (2011). Microhomology-mediated deletion and gene conversion in African trypanosomes. *Nucleic Acids Research*, *39*(4), 1372–1380. https://doi.org/10.1093/NAR/GKQ981

Glover, L., & Horn, D. (2012). Trypanosomal histone γh2A and the DNA damage response. *Molecular and Biochemical Parasitology*, *183*(1), 78–83. https://doi.org/10.1016/J.MOLBIOPARA.2012.01.008

Glover, L., Alsford, S., & Horn, D. (2013). DNA Break Site at Fragile Subtelomeres Determines Probability and Mechanism of Antigenic Variation in African Trypanosomes. *PLOS Pathogens*, *9*(3), e1003260. https://doi.org/10.1371/JOURNAL.PPAT.1003260

Gobbini, E., Cassani, C., Villa, M., Bonetti, D., & Longhese, M. P. (2016). Functions and regulation of the MRX complex at DNA double-strand breaks. OPEN ACCESS | www.Microbialcell.Com 329 Microbial Cell, 3(8). https://doi.org/10.15698/mic2016.08.517

Gomes Passos Silva, D., da Silva Santos, S., Nardelli, S. C., Mendes, I. C., Freire, A. C. G., Repolês, B. M., Resende, B. C., Costa-Silva, H. M., da Silva, V. S., Oliveira, K. A. de, Oliveira, C. F. B., Vilela, L. F. F., Nagem, R. A. P., Franco, G. R., Macedo, A. M., Pena, S. D. J., Tahara, E. B., Sales Junior, P. A., Moreira, D. S., … Machado, C. R. (2018). The in vivo and in vitro roles of Trypanosoma cruzi Rad51 in the repair of DNA double strand breaks and oxidative lesions. PLoS Neglected Tropical Diseases, 12(11). https://doi.org/10.1371/JOURNAL.PNTD.0006875

González-Corrochano, R., Ruiz, F. M., Taylor, N. M. I., Huecas, S., Drakulic, S., Spínola-Amilibia, M., & Fernández-Tornero, C. (2020). The crystal structure of human XPG, the xeroderma pigmentosum group G endonuclease, provides insight into nucleotide excision DNA repair. Nucleic Acids Research, 48(17), 9943–9958. https://doi.org/10.1093/NAR/GKAA688

Gorman, M. A., Morera, S., Rothwell, D. G., De La Fortelle, E., Mol, C. D., Tainer, J. A., Hickson, I. D., & Freemont, P. S. (1997). The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites. The EMBO Journal, 16(21), 6548–6558. https://doi.org/10.1093/EMBOJ/16.21.6548

Gottlieb, T. M., & Jackson, S. P. (1993). The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. Cell, 72(1), 131–142. https://doi.org/10.1016/0092-8674(93)90057-W

Graf, F. E., Ludin, P., Wenzler, T., Kaiser, M., Brun, R., Pyana, P. P., Büscher, P., de Koning, H. P., Horn, D., & Mäser, P. (2013). Aquaporin 2 mutations in Trypanosoma brucei gambiense field isolates correlate with decreased susceptibility to pentamidine and melarsoprol. *PLoS neglected tropical diseases*, *7*(10), e2475. https://doi.org/10.1371/journal.pntd.0002475

Günzl, A., Bruderer, T., Laufer, G., Schimanski, B., Tu, L. C., Chung, H. M., Lee, P. T., & Lee, M. G. S. (2003). RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in Trypanosoma brucei. Eukaryotic Cell, 2(3), 542–551. https://doi.org/10.1128/EC.2.3.542-551.2003

Haince, J. F., McDonald, D., Rodrigue, A., Déry, U., Masson, J. Y., Hendzel, M. J., & Poirier, G. G. (2008). PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites. The Journal of Biological Chemistry, 283(2), 1197–1208. https://doi.org/10.1074/JBC.M706734200

Hall, B. S., Wu, X., Hu, L., & Wilkinson, S. R. (2010). Exploiting the drug-activating properties of a novel trypanosomal nitroreductase. *Antimicrobial Agents and Chemotherapy*, *54*(3), 1193–1199. https://doi.org/10.1128/AAC.01213-09

Hajduk, S., & Ochsenreiter, T. (2010). RNA editing in kinetoplastids. RNA Biology, 7(2), 229–236. https://doi.org/10.4161/RNA.7.2.11393

Hall, B. S., Bot, C., & Wilkinson, S. R. (2011). Nifurtimox activation by trypanosomal type I nitroreductases generates cytotoxic nitrile metabolites. *Journal of Biological Chemistry*, *286*(15), 13088–13095. https://doi.org/10.1074/jbc.m111.230847

Hall, B. S., & Wilkinson, S. R. (2012). Activation of benznidazole by trypanosomal type I nitroreductases results in glyoxal formation. Antimicrobial Agents and Chemotherapy, 56(1), 115–123. https://doi.org/10.1128/aac.05135-11

Halliday, C., de Castro-Neto, A., Alcantara, C. L., Cunha-E-Silva, N. L., Vaughan, S., & Sunter, J. D. (2021). Trypanosomatid Flagellar Pocket from Structure to Function. Trends in parasitology, 37(4), 317–329. https://doi.org/10.1016/j.pt.2020.11.005

Hanawalt, P. C. (2003). Four decades of DNA repair: From early insights to current perspectives. Biochimie, 85(11), 1043–1052. https://doi.org/10.1016/j.biochi.2003.11.007

Hanawalt, P. C., & Spivak, G. (2008). Translesion synthesis Transcription-coupled DNA repair: two decades of progress and surprises. https://doi.org/10.1038/nrm2549

Hashimoto, S., Anai, H., & Hanada, K. (2016). Mechanisms of interstrand DNA crosslink repair and human disorders. Genes and Environment, 38(1), 1–8. https://doi.org/10.1186/S41021-016-0037-9/FIGURES/2

He, H., Chen, Q., & Georgiadis, M. M. (2014). High-Resolution Crystal Structures Reveal Plasticity in the Metal Binding Site of Apurinic/Apyrimidinic Endonuclease I.

Hegde, M. L., Hazra, T. K., & Mitra, S. (2008). Early Steps in the DNA Base Excision/Single-Strand Interruption Repair Pathway in Mammalian Cells. *Cell Research*, *18*(1), 27. https://doi.org/10.1038/CR.2008.8

Heinen, C. D., Schmutte, C., & Fishel, R. (2002). DNA repair and tumorigenesis: Lessons from hereditary cancer syndromes. *Cancer Biology and Therapy*, *1*(5), 477–485. https://doi.org/10.4161/CBT.1.5.160

Herv´, H., Menoni, H., Wienholz, F., Theil, A. F., Janssens, R. C., Lans, H., Campalans, A., Radicella, J. P., Marteijn, J. A., & Vermeulen, W. (2018). The transcription-coupled DNA repair-initiating protein CSB promotes XRCC1 recruitment to oxidative DNA damage. *Nucleic Acids Research*, *46*(15), 7747–7756. https://doi.org/10.1093/nar/gky579

Hide, G. (1999). History of sleeping sickness in East Africa. Clin Microbiol Rev, 12.

Ho, T. V., Guainazzi, A., Derkunt, S. B., Enoiu, M., & Schärer, O. D. (2011). Structure-dependent bypass of DNA interstrand crosslinks by translesion synthesis polymerases. Nucleic Acids Research, 39(17), 7455. https://doi.org/10.1093/NAR/GKR448

Holthausen, J. T., Wyman, C., & Kanaar, R. (2010). Regulation of DNA strand exchange in homologous recombination. In DNA Repair (Vol. 9, Issue 12, pp. 1264–1272). https://doi.org/10.1016/j.dnarep.2010.09.014

Horn, D. (2022). Genome-scale RNAi screens in African trypanosomes. Trends in Parasitology, 38(2), 160–173. https://doi.org/10.1016/J.PT.2021.09.002

Hotez, P. J., Aksoy, S., Brindley, P. J., & Kamhawi, S. (2020). What constitutes a neglected tropical disease?. PLoS neglected tropical diseases, 14(1), e0008001. https://doi.org/10.1371/journal.pntd.0008001

Hotez, P. J., Aksoy, S., Brindley, P. J., & Kamhawi, S. (2020). World neglected tropical diseases day. PLoS neglected tropical diseases, 14(1), e0007999. https://doi.org/10.1371/journal.pntd.0007999

Houtgraaf, J. H., Versmissen, J., & van der Giessen, W. J. (2006). A concise review of DNA damage checkpoints and repair in mammalian cells. Cardiovascular Revascularization Medicine: Including Molecular Interventions, 7(3), 165–172. https://doi.org/10.1016/J.CARREV.2006.02.002

Huang, Y., & Li, L. (2013). DNA crosslinking damage and cancer - a tale of friend and foe. Translational cancer research, 2(3), 144–154. https://doi.org/10.3978/j.issn.2218-676X.2013.03.01

Imlay, J. A. (2003). Pathways of Oxidative Damage. Http://Dx.Doi.Org/10.1146/Annurev.Micro.57.030502.090938, 57, 395–418. https://doi.org/10.1146/ANNUREV.MICRO.57.030502.090938

Iyama, T., Lee, S. Y., Berquist, B. R., Gileadi, O., Bohr, V. A., Seidman, M. M., McHugh, P. J., & Wilson, D. M. (2015). CSB interacts with SNM1A and promotes DNA interstrand crosslink processing. *Nucleic Acids Research*, *43*(1), 247–258. https://doi.org/10.1093/nar/gku1279

Iyama, T., & Wilson, D. M., 3rd (2016). Elements That Regulate the DNA Damage Response of Proteins Defective in Cockayne Syndrome. Journal of molecular biology, 428(1), 62–78. https://doi.org/10.1016/j.jmb.2015.11.020

Iyer, V. N., & Szybalski, W. (1963). A Molecular Mechanism of Mitomycin Action: Linking of Complementary DNA Strands. Proceedings of the National Academy of Sciences

Iyer, L. M., Babu, M. M., & Aravind, L. (2006). The HIRAN domain and recruitment of chromatin remodeling and repair activities to damaged DNA. Cell Cycle (Georgetown, Tex.), 5(7), 775–782. https://doi.org/10.4161/CC.5.7.2629

Jacobs, A. L., & Schär, P. (2012). DNA glycosylases: in DNA repair and beyond. Chromosoma, 121(1), 1–20. https://doi.org/10.1007/S00412-011-0347-4

Jackson, S. P. (2002). Sensing and repairing DNA double-strand breaks. Carcinogenesis, 23(5), 687–696. https://doi.org/10.1093/CARCIN/23.5.687

Jamieson, D., Tung, A. T. Y., Knox, R. J., & Boddy, A. V. (2006). Reduction of mitomycin C is catalysed by human recombinant NRH:quinone oxidoreductase 2 using reduced nicotinamide adenine dinucleotide as

an electron donating co-factor. British Journal of Cancer 2006 95:9, 95(9), 1229–1233. https://doi.org/10.1038/sj.bjc.6603414

Jamieson, E. R., & Lippard, S. J. (1999). Structure, Recognition, and Processing of Cisplatin−DNA Adducts. https://doi.org/10.1021/cr980421nStevens JR, Noyes HA, Schofield CJ, Gibson W (2001). The molecular evolution of Trypanosomatidae. Advances in Parasitology, 48, 1–53. https://doi.org/10.1016/S0065-308X(01)48003-1

Jiang, D., Hatahet, Z., Blaisdell, J. O., Melamede, R. J., & Wallace, S. S. (1997). Escherichia coli endonuclease VIII: cloning, sequencing, and overexpression of the nei structural gene and characterization of nei and nei nth mutants. Journal of Bacteriology, 179(11), 3773–3782. https://doi.org/10.1128/JB.179.11.3773-3782.1997

Jones, M., Beuron, F., Borg, A., Nans, A., Earl, C. P., Briggs, D. C., Snijders, A. P., Bowles, M., Morris, E. P., Linch, M., & McDonald, N. Q. (2020). Cryo-EM structures of the XPF-ERCC1 endonuclease reveal how DNA-junction engagement disrupts an auto-inhibited conformation. Nature Communications, 11(1). https://doi.org/10.1038/S41467-020-14856-2

Jones, M. J. K., & Huang, T. T. (2012). The Fanconi anemia pathway in replication stress and DNA crosslink repair. https://doi.org/10.1007/s00018-012-1051-0

Kaiser, M., Bray, M. A., Cal, M., Trunz, B. B., Torreele, E., & Brun, R. (2011). Antitrypanosomal activity of fexinidazole, a new oral nitroimidazole drug candidate for treatment of sleeping sickness. Antimicrobial Agents and Chemotherapy, 55(12), 5602–5608. https://doi.org/10.1128/AAC.00246-11

Kamath, R. S., & Ahringer, J. (2003). Genome-wide RNAi screening in Caenorhabditis elegans. Methods (San Diego, Calif.), 30(4), 313–321. https://doi.org/10.1016/s1046-2023(03)00050-1

Kavli, B., Otterlei, M., Slupphaug, G., & Krokan, H. E. (2007). Uracil in DNA-General mutagen, but normal intermediate in acquired immunity. DNA Repair, 6(4), 505–516. https://doi.org/10.1016/J.DNAREP.2006.10.014

Kaiser, M., Bray, M. A., Cal, M., Trunz, B. B., Torreele, E., & Brun, R. (2011). Antitrypanosomal activity of fexinidazole, a new oral nitroimidazole drug candidate for treatment of sleeping sickness. Antimicrobial Agents and Chemotherapy, 55(12), 5602–5608. https://doi.org/10.1128/AAC.00246-11

Kaye, J. A., Melo, J. A., Cheung, S. K., Vaze, M. B., Haber, J. E., & Toczyski, D. P. (2004). DNA breaks promote genomic instability by impeding proper chromosome segregation. Current Biology, 14(23), 2096–2106. https://doi.org/10.1016/J.CUB.2004.10.051/ATTACHMENT/F0F0A568-320A-4C7F-91DE-AAE49ACC13DC/MMC3.MPG

Kellum, A. H., Qiu, D. Y., Voehler, M. W., Martin, W., Gates, K. S., & Stone, M. P. (2021). Structure of a Stable Interstrand DNA Cross-Link Involving a β- N-Glycosyl Linkage Between an N6-dA Amino Group and an Abasic Site. Biochemistry, 60(1), 41–52. https://doi.org/10.1021/ACS.BIOCHEM.0C00596

Kelm, J. M., Samarbakhsh, A., Pillai, A., VanderVere-Carozza, P. S., Aruri, H., Pandey, D. S., Pawelczak, K. S., Turchi, J. J., & Gavande, N. S. (2022). Recent Advances in the Development of Non-PIKKs Targeting Small Molecule Inhibitors of DNA Double-Strand Break Repair. Frontiers in Oncology, 12, 850883. https://doi.org/10.3389/fonc.2022.850883

Kennedy, E. E., Caffrey, P. J., & Delaney, S. (2018). Initiating base excision repair in chromatin. *DNA Repair*, *71*, 87–92. https://doi.org/10.1016/j.dnarep.2018.08.011

Klein Douwel, D., Boonen, R. A. C. M., Long, D. T., Szypowska, A. A., Räschle, M., Walter, J. C., & Knipscheer, P. (2014). XPF-ERCC1 acts in unhooking DNA interstrand crosslinks in cooperation with FANCD2 and FANCP/SLX4. *Molecular Cell*, *54*(3), 460. https://doi.org/10.1016/J.MOLCEL.2014.03.015

Klein Douwel, D., Hoogenboom, W. S., Boonen, R. A., & Knipscheer, P. (2017). Recruitment and positioning determine the specific role of the XPF - ERCC 1 endonuclease in interstrand crosslink repair. *The EMBO Journal*, *36*(14), 2034–2046. https://doi.org/10.15252/embj.201695223

Koczor, C. A., Saville, K. M., Andrews, J. F., Clark, J., Fang, Q., Li, J., Al-Rahahleh, R. Q., Ibrahim, M., McClellan, S., Makarov, M. v., Migaud, M. E., & Sobol, R. W. (2021). Temporal dynamics of base excision/single-strand break repair protein complex assembly/disassembly are modulated by the PARP/NAD+/SIRT6 axis. *Cell Reports*, *37*(5). https://doi.org/10.1016/J.CELREP.2021.109917

Korolev, V. G. (2005). Base excision repair: AP endonucleases and DNA polymerases. *Genetika*, *41*(10), 1301–1309. 50(2), 355–362. https://doi.org/10.1073/PNAS.50.2.355

Kozlov, S. V., Graham, M. E., Jakob, B., Tobias, F., Kijas, A. W., Tanuji, M., Chen, P., Robinson, P. J., Taucher-Scholz, G., Suzuki, K., So, S., Chen, D., & Lavin, M. F. (2011). Autophosphorylation and ATM activation: additional sites add to the complexity. The Journal of Biological Chemistry, 286(11), 9107–9119. https://doi.org/10.1074/JBC.M110.204065

Kratz, K., Schöpf, B., Kaden, S., Sendoel, A., Eberhard, R., Lademann, C., Cannavó, E., Sartori, A. A., Hengartner, M. O., & Jiricny, J. (2010). Deficiency of FANCD2-associated nuclease KIAA1018/FAN1 sensitizes cells to interstrand crosslinking agents. Cell, 142, 77–88.

Krishnan, V., & Ito, Y. (2017). A regulatory role for RUNX1, RUNX3 in the maintenance of genomic integrity. RUNX Proteins in Development and Cancer, 491-510.

Krokan, H. E., & Bjørås, M. (2013). Base excision repair. Cold Spring Harbor Perspectives in Biology, 5(4), 1–22. https://doi.org/10.1101/CSHPERSPECT.A012583

Krokan, H. E., Standal, R., & Slupphaug, G. (1997). DNA glycosylases in the base excision repair of DNA. The Biochemical Journal, 325 (Pt 1)(Pt 1), 1–16. https://doi.org/10.1042/BJ3250001

Kubota, Y., Nash, R. A., Klungland, A., Schär, P., Barnes, D. E., & Lindahl, T. (1996). Reconstitution of DNA base excision-repair with purified human proteins: Interaction between DNA polymerase β and the XRCC1 protein. EMBO Journal, 15(23), 6662–6670. https://doi.org/10.1002/j.1460-2075.1996.tb01056.x

Kucherlapati, M., Yang, K., Kuraguchi, M., Zhao, J., Lia, M., Heyer, J., Kane, M. F., Fan, K., Russell, R., Brown, A. M. C., Kneitz, B., Edelmann, W., Kolodner, R. D., Lipkin, M., & Kucherlapati, R. (2002). Haploinsufficiency of Flap endonuclease (Fen1) leads to rapid tumor progression. Proceedings of the National Academy of Sciences of the United States of America, 99(15), 9924–9929. https://doi.org/10.1073/PNAS.152321699

Kumar, A. (2018). DNA Interstrand Crosslink Repair in Trypanosoma brucei (Doctoral dissertation, Queen Mary University of London).

LaCount, D. J., Bruse, S., Hill, K. L., & Donelson, J. E. (2000). Double-stranded RNA interference in Trypanosoma brucei using head-to-head promoters. Molecular and Biochemical Parasitology, 111(1), 67–76. https://doi.org/10.1016/S0166-6851(00)00300-5t

LaCount, D. J., & Donelson, J. E. (2001). RNA interference in African trypanosomes. Protist, 152(2), 103–111. https://doi.org/10.1078/1434-4610-00047Lawley, P. D., & Phillips, D. H. (1996). DNA adducts from chemotherapeutic agents. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 355(1–2), 13–40. https://doi.org/10.1016/0027-5107(96)00020-6

Laev, S. S., Salakhutdinov, N. F., & Lavrik, O. I. (2017). Inhibitors of nuclease and redox activity of apurinic/apyrimidinic endonuclease 1/redox effector factor 1 (APE1/Ref-1). Bioorganic and Medicinal Chemistry, 25(9), 2531–2544. https://doi.org/10.1016/j.bmc.2017.01.028

Lam, W., Park, S.-Y., Leung, C.-H., & Cheng, Y.-C. (2006). Apurinic/Apyrimidinic Endonuclease-1 Protein Level Is Associated with the Cytotoxicity of l-Configuration Deoxycytidine Analogs (Troxacitabine and β-l-2′,3′-Dideoxy-2′,3′-didehydro-5-fluorocytidine) but Not d-Configuration Deoxycytidine Analogs (Gemcitab…. Molecular Pharmacology, 69(5).

Lawley, P. D., & Phillips, D. H. (1996). DNA adducts from chemotherapeutic agents. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 355(1–2), 13–40. https://doi.org/10.1016/0027-5107(96)00020-6

Lazzaro, F., Giannattasio, M., Puddu, F., Granata, M., Pellicioli, A., Plevani, P., & Muzi-Falconi, M. (2009). Checkpoint mechanisms at the intersection between DNA damage and repair. DNA Repair, 8(9), 1055—1067. https://doi.org/10.1016/j.dnarep.2009.04.022

LeBowitz, J. H., Smith, H. Q., Rusche, L., & Beverley, S. M. (1993). Coupling of poly(A) site selection and trans-splicing in Leishmania. *Genes & Development*, *7*(6), 996–1007. https://doi.org/10.1101/GAD.7.6.996

Lecordier, L., Devaux, S., Uzureau, P., Dierick, J. F., Walgraffe, D., Poelvoorde, P., Pays, E., & Vanhamme, L. (2007). Characterization of a TFIIH homologue from Trypanosoma brucei. Molecular Microbiology, 64(5), 1164–1181. https://doi.org/10.1111/J.1365-2958.2007.05725.X

Lee, B. Y., Bacon, K. M., Bottazzi, M. E., & Hotez, P. J. (2013). Global economic burden of Chagas disease: a computational simulation model. *The Lancet Infectious Diseases*, *13*(4), 342–348. https://doi.org/10.1016/S1473-3099(13)70002-1

Lee, J. H., Jung, H. S., & Gü Nzl, A. (2009). Transcriptionally active TFIIH of the early-diverged eukaryote Trypanosoma brucei harbors two novel core subunits but not a cyclin-activating kinase complex. Nucleic Acids Research, 37(11), 3811–3820. https://doi.org/10.1093/nar/gkp236

Lehmann, A. R. (2003). DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. Biochimie, 85(11), 1101–1111. https://doi.org/10.1016/J.BIOCHI.2003.09.010

le Page, F., Kwoh, E. E., Avrutskaya, A., Gentil, A., Leadon, S. A., Sarasin, A., & Cooper, P. K. (2000). Transcription-coupled repair of 8-oxoguanine: requirement for XPG, TFIIH, and CSB and implications for Cockayne syndrome. *Cell*, *101*(2), 159–171. https://doi.org/10.1016/s0092-8674(00)80827-2

Legros, D., Ollivier, G., Gastelluetchegorry, M., Paquet, C., Burri, C., Jannin, J., & Buscher, P. (2002). Treatment of human African trypanosomiasis?present situation and needs for research and development. The Lancet Infectious Diseases, 2(7), 437–440. https://doi.org/10.1016/S1473-3099(02)00321-3

Lehoczky´1, P., Lehoczky´1, L., Mchugh, P. J., & Chovanec, M. (n.d.). DNA interstrand cross-link repair in Saccharomyces cerevisiae. https://doi.org/10.1111/j.1574-6976.2006.00046.x

Li, X., & Heyer, W. D. (2008). Homologous recombination in DNA repair and DNA damage tolerance. Cell Research, 18(1), 99. https://doi.org/10.1038/CR.2008.1

Li, J., Wang, Q.-E., Zhu, Q., El-Mahdy, M. A., Wani, G., Praetorius-Ibba, M., & Wani, A. A. (2006). DNA Damage Binding Protein Component DDB1 Participates in Nucleotide Excision Repair through DDB2 DNA-binding and Cullin 4A Ubiquitin Ligase Activity. Cancer Res, 66(17), 8590–8597. https://doi.org/10.1158/0008-5472.CAN-06-1115

Limoli, C. L., Giedzinski, E., Bonner, W. M., & Cleaver, J. E. (2002). UV-induced replication arrest in the xeroderma pigmentosum variant leads to DNA doublestrand breaks, γ-H2AX formation, and Mre11 relocalization. In Proceedings of the National Academy of Sciences of the United States of America (Vol. 99, Issue 1, pp. 233–238). https://doi.org/10.1073/pnas.231611798

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature 1993 362:6422, 362(6422), 709–715. https://doi.org/10.1038/362709a0

Lisby, M., Barlow, J. H., Burgess, R. C., & Rothstein, R. (2004). Choreography of the DNA Damage Response: Spatiotemporal Relationships among Checkpoint and Repair Proteins recruited to sites of DNA damage are coordinated at. In Cell (Vol. 118). http://www.cell.com/cgi/content/full/

Lisby, M., & Rothstein, R. (2009). Choreography of Recombination Proteins during the DNA Damage Response. https://doi.org/10.1016/j.dnarep.2009.04.007

Liu, T., Ghosal, G., Yuan, J., Chen, J., & Huang, J. (2010). FAN1 acts with FANCI-FANCD2 to promote DNA interstrand cross-link repair. Science, 329, 693–696.

Löbrich, M., & Jeggo, P. (2017). A Process of Resection-Dependent Nonhomologous End Joining Involving the Goddess Artemis. Trends in Biochemical Sciences, 42(9), 690–701. https://doi.org/10.1016/J.TIBS.2017.06.011

Lopes, D.deO., Schamber-Reis, B. L., Regis-da-Silva, C. G., Rajão, M. A., Darocha, W. D., Macedo, A. M., Franco, G. R., Nardelli, S. C., Schenkman, S., Hoffmann, J. S., Cazaux, C., Pena, S. D., Teixeira, S. M., & Machado, C. R. (2008). Biochemical studies with DNA polymerase beta and DNA polymerase beta-PAK of Trypanosoma cruzi suggest the involvement of these proteins in mitochondrial DNA maintenance. DNA repair, 7(11), 1882–1892. https://doi.org/10.1016/j.dnarep.2008.07.018

Lopes, A. (2010). Trypanosomatids: Odd Organisms, Devastating Diseases. In The Open Parasitology Journal (Vol. 4, Issue 1, pp. 30–59). https://doi.org/10.2174/1874421401004010030

Lopez-Martinez, D., Liang, C. C., & Cohn, M. A. (2016). Cellular response to DNA interstrand crosslinks: the Fanconi anemia pathway. Cellular and Molecular Life Sciences, 73(16), 3097–3114. https://doi.org/10.1007/s00018-016-2218-x

Lu, D., Yang, N., Wang, S., Liu, W., Zhang, D., Wang, J., Huang, B., & Li, X. (2021). Identifying the predictive role of oxidative stress genes in the prognosis of glioma patients. Medical Science Monitor, 27. https://doi.org/10.12659/MSM.934161

Ludwig, D. L., MacInnes, M. A., Takiguchi, Y., Purtymun, P. E., Henrie, M., Flannery, M., Meneses, J., Pedersen, R. A., & Chen, D. J. (1998). A murine AP-endonuclease gene-targeted deficiency with post-implantation embryonic progression and ionizing radiation sensitivity. Mutation Research/DNA Repair, 409(1), 17–29. https://doi.org/10.1016/S0921-8777(98)00039-1

Lukes, J., Hashimi, H., & Zíková, A. (2005). Unexplained complexity of the mitochondrial genome and transcriptome in kinetoplastid flagellates. Current genetics, 48(5), 277–299. https://doi.org/10.1007/s00294-005-0027-0

Luncsford, P. J., Chang, D.-Y., Shi, G., Bernstein, J., Madabushi, A., Patterson, D. N., Lu, A.-L., Toth, E. A., & Lu, L. (2011). Author manuscript; available in PMC. J Mol Biol, 403(3), 351–370. https://doi.org/10.1016/j.jmb.2010.08.045

Luncsford, P. J., Manvilla, B. A., Patterson, D. N., Malik, S. S., Jin, J., Hwang, B. J., Gunther, R., Kalvakolanu, S., Lipinski, L. J., Yuan, W., Lu, W., Drohat, A. C., Lu, A. L., & Toth, E. A. (2013). Coordination of MYH DNA glycosylase and APE1 endonuclease activities via physical interactions. DNA Repair, 12(12), 1043–1052. https://doi.org/10.1016/j.dnarep.2013.09.007

Machado, C. R., Vieira-Da-Rocha, J. P., Isabela, †, Mendes, C., Rajão, M. A., Marcello, L., Bitar, M., Drummond, M. G., Grynberg, P., Oliveira, D. A. A., Marques, C., Houten, B. Van, & Mcculloch, R. (2014). Nucleotide excision repair in Trypanosoma brucei: specialization of transcription-coupled repair due to multigenic transcription. https://doi.org/10.1111/mmi.12589

MacKay, C., Déclais, A. C., Lundin, C., Agostinho, A., Deans, A. J., MacArtney, T. J., Hofmann, K., Gartner, A., West, S. C., Helleday, T., Lilley, D. M. J., & Rouse, J. (2010). Identification of KIAA1018/FAN1, a DNA Repair Nuclease Recruited to DNA Damage by Monoubiquitinated FANCD2. Cell, 142(1), 65–76. https://doi.org/10.1016/j.cell.2010.06.021

Mahmood, R., Hines, J. C., & Ray, D. S. (1999). Identification of cis andtrans Elements Involved in the Cell Cycle Regulation of Multiple Genes in Crithidia fasciculata. Molecular and Cellular Biology, 19(9), 6174–6182. https://doi.org/10.1128/mcb.19.9.6174

Mair, G., Shi, H., Li, H., Djikeng, A., Aviles, H. O., Bishop, J. R., Falcone, F. H., Gavrilescu, C., Montgomery, J. L., Santori, M. I., Stern, L. S., Wang, Z., Ullu, E., & Tschudi, C. (2000). A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. RNA (New York, N.Y.), 6(2), 163–169. https://doi.org/10.1017/S135583820099229X

Mandel, C. R., Kaneko, S., Zhang, H., Gebauer, D., Vethantham, V., Manley, J. L., & Tong, L. (2006). Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. Nature, 444(7121), 953–956. https://doi.org/10.1038/nature05363

Marin, P. A., da Silva, M. S., Pavani, R. S., Machado, C. R., & Elias, M. C. (2018). Recruitment kinetics of the homologous recombination pathway in procyclic forms of Trypanosoma brucei after ionizing radiation treatment. Scientific Reports, 8(1), 5405. https://doi.org/10.1038/s41598-018-23731-6

Marteijn, J. A., Lans, H., Vermeulen, W., & J Hoeijmakers, J. H. (2014). The integrity of DNA is constantly threatened by endo¬ genously formed metabolic products. https://doi.org/10.1038/nrm3822

Matthews, K. R., Ellis, J. R., & Paterou, A. (2004). Molecular regulation of the life cycle of African trypanosomes. Trends in Parasitology, 20(1), 40–47. https://doi.org/10.1016/j.pt.2003.10.016

Matsumoto, Y., Kim, K., & Bogenhagen, D. F. (1994). Proliferating cell nuclear antigen-dependent abasic site repair in Xenopus laevis oocytes: an alternative pathway of base excision DNA repair. Molecular and Cellular Biology, 14(9), 6187–6197. https://doi.org/10.1128/MCB.14.9.6187-6197.1994

Mathews, K. R. (2005). The developmental cell biology of Trypanosoma brucei. *Journal of Cell Science*, *118*(2), 283–290. https://doi.org/10.1242/JCS.01649

Matthews K. R. (2021). Trypanosome Signaling-Quorum Sensing. Annual review of microbiology, 75, 495–514. https://doi.org/10.1146/annurev-micro-020321-115246

Matsumoto, Y., Kim, K., & Bogenhagen, D. F. (1994). Proliferating cell nuclear antigen-dependent abasic site repair in Xenopus laevis oocytes: an alternative pathway of base excision DNA repair. Molecular and Cellular Biology, 14(9), 6187–6197. https://doi.org/10.1128/MCB.14.9.6187-6197.1994

McAndrew, M., Graham, S., Hartmann, C., & Clayton, C. (1998). Testing Promoter Activity in the Trypanosome Genome: Isolation of a Metacyclic-Type VSG Promoter, and Unexpected Insights into RNA Polymerase II Transcription. Experimental Parasitology, 90(1), 65–76. https://doi.org/10.1006/EXPR.1998.4317

Mcculloch, R., & David Barry, J. (1999). A role for RAD51 and homologous recombination in Trypanosoma brucei antigenic variation. www.genesdev.org

McCullough, A. K., Dodson, M. L., & Lloyd, R. S. (1999). Initiation of base excision repair: glycosylase mechanisms and structures. Annual Review of Biochemistry, 68, 255–285. https://doi.org/10.1146/ANNUREV.BIOCHEM.68.1.255

McNeil, E. M., & Melton, D. W. (2012). DNA repair endonuclease ERCC1-XPF as a novel therapeutic target to overcome chemoresistance in cancer therapy. Nucleic Acids Research, 40(20), 9990–10004. https://doi.org/10.1093/NAR/GKS818

McVey, M. (2010). Strategies for DNA interstrand crosslink repair: insights from worms, flies, frogs, and slime molds. Environmental and Molecular Mutagenesis, 51(6), 646–658. https://doi.org/10.1002/em.20551

Mehnert, A. K., Prorocic, M., Dujeancourt-Henry, A., Hutchinson, S., McCulloch, R., & Glover, L. (2021). The MRN complex promotes DNA repair by homologous recombination and restrains antigenic variation in African trypanosomes. Nucleic Acids Research, 49(3), 1436–1454. https://doi.org/10.1093/nar/gkaa1265

Menoni, H., Hoeijmakers, J. H. J., & Vermeulen, W. (2012). Nucleotide excision repair-initiating proteins bind to oxidative DNA lesions in vivo. *Journal of Cell Biology*, *199*(7), 1037–1046. https://doi.org/10.1083/jcb.201205149

Menoni, H., Wienholz, F., Theil, A. F., Janssens, R. C., Lans, H., Campalans, A., Radicella, J. P., Marteijn, J. A., & Vermeulen, W. (2018). The transcription-coupled DNA repair-initiating protein CSB promotes XRCC1 recruitment to oxidative DNA damage. Nucleic acids research, 46(15), 7747–7756. https://doi.org/10.1093/nar/gky579

Min, J. H., & Pavletich, N. P. (2007). Recognition of DNA damage by the Rad4 nucleotide excision repair protein. Nature, 449(7162), 570–575. https://doi.org/10.1038/nature06155

Mimitou, E. P., & Symington, L. S. (2009). Nucleases and helicases take center stage in homologous recombination. Trends in Biochemical Sciences, 34(5), 264–272. https://doi.org/10.1016/j.tibs.2009.01.010

Miura, M., Nakamura, S., Sasaki, T., Takasaki, Y., Shiomi, T., & Yamaizumi, M. (1996). Roles of XPG and XPF/ERCC1 endonucleases in UV-induced immunostaining of PCNA in fibroblasts. *Experimental Cell Research*, *226*(1), 126–132. https://doi.org/10.1006/excr.1996.0210

Mol, C. D., Arvai, A. S., Sanderson, R. J., Slupphaug, G., Kavli, B., Krokan, H. E., Mosbaugh, D. W., & Tainer, J. A. (1995). Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: Protein mimicry of DNA. Cell, 82(5), 701–708. https://doi.org/https://doi.org/10.1016/0092-8674(95)90467-0

Mol, C. D., Hosfield, D. J., & Tainer, J. A. (2000). Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair: The 3' ends justify the means. Mutation Research - DNA Repair, 460(3–4), 211–229. https://doi.org/10.1016/S0921-8777(00)00028-8

Mony, B. M., MacGregor, P., Ivens, A., Rojas, F., Cowton, A., Young, J., Horn, D., & Matthews, K. (2014). Genome-wide dissection of the quorum sensing signalling pathway in Trypanosoma brucei. *Nature*, *505*(7485), 681–685. https://doi.org/10.1038/nature12864

Mocquet, V., Lainé, J. P., Riedl, T., Yajin, Z., Lee, M. Y., & Egly, J. M. (2008). Sequential recruitment of the repair factors during NER: The role of XPG in initiating the resynthesis step. *EMBO Journal*, *27*(1), 155–167. https://doi.org/10.1038/sj.emboj.7601948

Mony, B. M., & Matthews, K. R. (2015). Assembling the components of the quorum sensing pathway in African trypanosomes. Molecular microbiology, 96(2), 220–232. https://doi.org/10.1111/mmi.12949

Mu, D., Hsu, D. S., & Sancar, A. (1996). Reaction mechanism of human DNA repair excision nuclease. Journal of Biological Chemistry, 271(14), 8285–8294. https://doi.org/10.1074/JBC.271.14.8285

Muñ Oz, I. M., Hain, K., Cile Dé, A.-C., Gardiner, M., Toh, G. W., Sanchez-Pulido, L., Heuckmann, J. M., Toth, R., Macartney, T., Eppink, B., Kanaar, R., Ponting, C. P., Lilley, D. M. J., & Rouse, J. (2009). Molecular Cell Coordination of Structure-Specific Nucleases by Human SLX4/BTBD12 Is Required for DNA Repair. Molecular Cell, 35, 116–127. https://doi.org/10.1016/j.molcel.2009.06.020

Muniandy, P. A., Liu, J., Majumdar, A., Liu, S. T., & Seidman, M. M. (2010). DNA interstrand crosslink repair in mammalian cells: Step by step. Critical Reviews in Biochemistry and Molecular Biology, 45(1), 23–49. https://doi.org/10.3109/10409230903501819

Nakai, K., & Kanehisa, M. (1992). A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics, 14(4), 897–911. https://doi.org/10.1016/s0888-7543(05)80111-9

Nakamura, T., Okabe, K., Hirayama, S., Chirifu, M., Ikemizu, S., Morioka, H., Nakabeppu, Y., & Yamagata, Y. (2021). Structure of the mammalian adenine DNA glycosylase MUTYH: insights into the base excision repair pathway and cancer. Nucleic Acids Research, 49(12), 7154–7163. https://doi.org/10.1093/nar/gkab492

Nakazawa, Y., Hara, Y., Oka, Y., Komine, O., van den Heuvel, D., Guo, C., Daigaku, Y., Isono, M., He, Y., Shimada, M., Kato, K., Jia, N., Hashimoto, S., Kotani, Y., Miyoshi, Y., Tanaka, M., Sobue, A., Mitsutake, N., Suganami, T., … Ogi, T. (2020). Ubiquitination of DNA Damage-Stalled RNAPII Promotes Transcription-Coupled Repair. Cell, 180(6), 1228-1244.e24. https://doi.org/10.1016/J.CELL.2020.02.010

Nepal, M., Che, R., Zhang, J., Ma, C., & Fei, P. (2017). Fanconi Anemia Signaling and Cancer. Trends in Cancer, 3(12), 840–856. https://doi.org/10.1016/J.TRECAN.2017.10.005

Ngo, H., Tschudi, C., Gull, K., & Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in Trypanosoma brucei. Proceedings of the National Academy of Sciences, USA, 95(25), 14687–14692. https://doi.org/10.1073/pnas.95.25.14687

Nishi, R., Okuda, Y., Watanabe, E., Mori, T., Iwai, S., Masutani, C., Sugasawa, K., & Hanaoka, F. (2005). Centrin 2 Stimulates Nucleotide Excision Repair by Interacting with Xeroderma Pigmentosum Group C Protein. Molecular and Cellular Biology, 25(13), 5664–5674. https://doi.org/10.1128/mcb.25.13.5664-5674.2005

Nimonkar, A. V., Genschel, J., Kinoshita, E., Polaczek, P., Campbell, J. L., Wyman, C., Modrich, P., & Kowalczykowski, S. C. (2011). BLM–DNA2–RPA–MRN and EXO1–BLM–RPA–MRN constitute two DNA end resection machineries for human DNA break repair. Genes & Development, 25(4), 350–362. https://doi.org/10.1101/GAD.2003811

Nok, A. J. (2003). Arsenicals (melarsoprol), pentamidine and suramin in the treatment of human African trypanosomiasis. Parasitology Research, 90(1), 71–79. https://doi.org/10.1007/s00436-002-0799-9

Noll, D. M., McGregor Mason, T., & Miller, P. S. (2006). Formation and repair of interstrand cross-links in DNA. In Chemical Reviews (Vol. 106, Issue 2, pp. 277–301). https://doi.org/10.1021/cr040478b

Nozaki, T., Engel, J. C., & Dvorak, J. A. (1996). Cellular and Molecular Biological Analyses of Nifurtimox Resistance in Trypanosoma cruzi. The American Journal of Tropical Medicine and Hygiene, 55(1), 111–117. https://doi.org/10.4269/AJTMH.1996.55.111

Oezguen, N., Schein, C. H., Peddi, S. R., Power, T. D., Izumi, T., & Braun, W. (2007). A "'Moving Metal Mechanism'" for Substrate Cleavage by the DNA Repair Endonuclease APE-1. https://doi.org/10.1002/prot.21397

Odiit, M., Coleman, P. G., Liu, W. C., Mcdermott, J. J., Fevre, E. M., Welburn, S. C., & Woolhouse, M. E. (2005). Quantifying the level of under-detection of Trypanosoma brucei rhodesiense sleeping sickness cases. Tropical Medicine and International Health, 10(9), 840–849. https://doi.org/10.1111/j.1365-3156.2005.01470.x

Ogi, T., Limsirichaikul, S., Overmeer, R. M., Volker, M., Takenaka, K., Cloney, R., Nakazawa, Y., Niimi, A., Miki, Y., Jaspers, N. G., Mullenders, L. H. F., Yamashita, S., Fousteri, M. I., & Lehmann, A. R. (2010). Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. *Molecular Cell*, *37*(5), 714–727. https://doi.org/10.1016/J.MOLCEL.2010.02.009

Oishi, K., Hofmann, S., Diaz, G. A., Brown, T., Manwani, D., Ng, L., Young, R., Vlassara, H., Ioannou, Y. A., Forrest, D., & Gelb, B. D. (2002). Biallelic germline mutations in MYH predispose to multiple colorectal adenoma and somatic G:C-->T:A mutations. *Human Molecular Genetics*, *11*(23), 2961–2967. https://doi.org/10.1093/HMG/11.23.2961

Oksenych, V., & Coin, F. (2010). The long unwinding road: XPB and XPD helicases in damaged DNA opening. Cell Cycle, 9(1), 90–96. https://doi.org/10.4161/cc.9.1.10267

Ormeño, F., Barrientos, C., Ramirez, S., Ponce, I., Valenzuela, L., Sepúlveda, S., Bitar, M., Kemmerling, U., Machado, C. R., Cabrera, G., & Galanti, N. (2016). Expression and the Peculiar Enzymatic Behavior of the Trypanosoma cruzi NTH1 DNA Glycosylase. PloS One, 11(6). https://doi.org/10.1371/JOURNAL.PONE.0157270

Overath, P., & Engstler, M. (2004). Endocytosis, membrane recycling and sorting of GPI-anchored proteins: Trypanosoma brucei as a model system. Molecular microbiology, 53(3), 735–744. https://doi.org/10.1111/j.1365-2958.2004.04224.x

Palenchar, J. B., & Bellofatto, V. (2006). Gene transcription in trypanosomes. Molecular and Biochemical Parasitology, 146(2), 135–141. https://doi.org/10.1016/J.MOLBIOPARA.2005.12.008

Pan, S. S., Andrews, P. A., Glover, C. J., & Bachur, N. R. (1984). Reductive activation of mitomycin C and mitomycin C metabolites catalyzed by NADPH-cytochrome P-450 reductase and xanthine oxidase. The Journal of Biological Chemistry, 259(2), 959–966. https://doi.org/10.1016/S0021-9258(17)43551-4

Parsons, J. L., & Elder, R. H. (2003). DNA N-glycosylase deficient mice: a tale of redundancy. Mutation Research, 531(1–2), 165–175. https://doi.org/10.1016/J.MRFMMM.2003.05.001

Parsons, M. (2004). Glycosomes: parasites and the divergence of peroxisomal purpose. Molecular Microbiology, 53(3), 717–724. https://doi.org/10.1111/j.1365-2958.2004.04203.x

Passos-Silva, D. G., Rajão, M. A., Nascimento de Aguiar, P. H., Vieira-da-Rocha, J. P., Machado, C. R., & Furtado, C. (2010). Overview of DNA Repair in Trypanosoma cruzi, Trypanosoma brucei, and Leishmania major. Journal of Nucleic Acids, 2010, 840768. https://doi.org/10.4061/2010/840768

Patterson, S., & Fairlamb, A. H. (2019). Current and Future Prospects of Nitro-compounds as Drugs for Trypanosomiasis and Leishmaniasis. Current Medicinal Chemistry, 26(23), 4454–4475. https://doi.org/10.2174/0929867325666180426164352

Pays, E. (2005). Regulation of antigen gene expression in Trypanosoma brucei. Trends in Parasitology, 21(11), 517–520. https://doi.org/10.1016/J.PT.2005.08.016

Pearl, L. H., Schierz, A. C., Ward, S. E., Al-Lazikani, B., & Pearl, F. M. G. (2015). Therapeutic opportunities within the DNA damage response. *Nature Reviews Cancer*, *15*(3), 166–180. https://doi.org/10.1038/NRC3891

Palenchar, J. B., & Bellofatto, V. (2006). Gene transcription in trypanosomes. Molecular and Biochemical Parasitology, 146(2), 135–141. https://doi.

Pellegrini, L., Yu, D. S., Lo, T., Anand, S., Lee, M., Blundell, T. L., & Venkitaraman, A. R. (2002). Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. www.nature.com/nature

Peter De Raadt. (n.d.). De Raadt P. The history of sleeping sickness. World ... - Google Scholar. Retrieved October 14, 2021, from https://scholar.google.com/scholar?hl=fr&as_sdt=0%2C5&q=de+Raadt+P.+The+history+of+sleeping+sickness.+World+Health+Org%3B+2005.&btnG=

Petermann, E., Ziegler, M., & Oei, S. L. (2003). ATP-dependent selection between single nucleotide and long patch base excision repair. DNA Repair, 2(10), 1101–1114. https://doi.org/10.1016/S1568-7864(03)00117-4

Petruseva, I. O., Evdokimov, A. N., & Lavrik, O. I. (2014). Molecular Mechanism of Global Genome Nucleotide Excision Repair. 6.

Priotto, G., Kasparian, S., Mutombo, W., Ngouama, D., Ghorashian, S., Arnold, U., Ghabri, S., Baudin, E., Buard, V., Kazadi-Kyanza, S., Ilunga, M., Mutangala, W., Pohlig, G., Schmid, C., Karunakara, U., Torreele, E., & Kande, V. (2009). Nifurtimox-eflornithine combination therapy for second-stage African Trypanosoma brucei gambiense trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. Lancet, 374(9683), 56–64. https://doi.org/10.1016/s0140-6736(09)61117-x

Pizzolato, J., Mukherjee, S., Schärer, O. D., & Jiricny, J. (2015). FANCD2-associated nuclease 1, but not exonuclease 1 or flap endonuclease 1, is able to unhook DNA interstrand cross-links in vitro. Journal of Biological Chemistry, 290(37), 22602–22611. https://doi.org/10.1074/jbc.M115.663666

Rajão, M. A., Furtado, C., Alves, C. L., Passos-Silva, D. G., de Moura, M. B., Schamber-Reis, B. L., Kunrath-Lima, M., Zuma, A. A., Vieira-da-Rocha, J. P., Garcia, J. B. F., Mendes, I. C., Pena, S. D. J., Macedo, A. M., Franco, G. R., de Souza-Pinto, N. C., de Medeiros, M. H. G., Cruz, A. K., Motta, M. C. M., Teixeira,

S. M. R., & Machado, C. R. (2014). Unveiling Benznidazole's mechanism of action through overexpression of DNA repair proteins in Trypanosoma cruzi. *Environmental and Molecular Mutagenesis*, *55*(4), 309–321. https://doi.org/10.1002/EM.21839

Rahimian, E., Amini, A., Alikarami, F., Pezeshki, S. M. S., Saki, N., & Safa, M. (2020). DNA repair pathways as guardians of the genome: Therapeutic potential and possible prognostic role in hematologic neoplasms. DNA Repair, 96, 102951. https://doi.org/10.1016/J.DNAREP.2020.102951

Ralston, K. S., Kabututu, Z. P., Melehani, J. H., Oberholzer, M., & Hill, K. L. (2009). The Trypanosoma brucei flagellum: moving parasites in new directions. Annual Review of Microbiology, 63, 335–362. https://doi.org/10.1146/annurev.micro.091208.073353

Ramrath, D. J. F., Niemann, M., Leibundgut, M., Bieri, P., Prange, C., Horn, E. K., Leitner, A., Boehringer, D., Schneider, A., & Ban, N. (2018). Evolutionary shift toward protein-based architecture in trypanosomal mitochondrial ribosomes. Science, 362(6413). https://doi.org/10.1126/SCIENCE.AAU7735/SUPPL_FILE/AAU7735_RAMRATH_SM.PDF

Ramsden, D. A., & Geliert, M. (1998). Ku protein stimulates DNA end joining by mammalian DNA ligases: a direct role for Ku in repair of DNA double-strand breaks. The EMBO Journal, 17(2), 609. https://doi.org/10.1093/EMBOJ/17.2.609

Ranjha, L., Howard, S. M., & Cejka, P. (2018). Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes. In Chromosoma (Vol. 127, Issue 2, pp. 187–214). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s00412-017-0658-1

Räschle, M., Knipsheer, P., Enoiu, M., Angelov, T., Sun, J., Griffith, J. D., Ellenberger, T. E., Schärer, O. D., & Walter, J. C. (2008). Mechanism of Replication-Coupled DNA Interstrand Crosslink Repair. Cell, 134(6), 969–980. https://doi.org/10.1016/J.CELL.2008.08.030/ATTACHMENT/5C416A9A-2751-4C69-93EA-B6AAB52C474E/MMC1.PDF

Redmond, S., Vadivelu, J., & Field, M. C. (2003). RNAit: An automated web-based tool for the selection of RNAi targets in Trypanosoma brucei. *Molecular and Biochemical Parasitology*, *128*(1), 115–118. https://doi.org/10.1016/S0166-6851(03)00045-8

Rico, E., Jeacock, L., Kovářová, J., & Horn, D. (2018). Inducible high-efficiency CRISPR-Cas9-targeted gene editing and precision base editing in African trypanosomes. *Scientific Reports 2018 8:1*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-26303-w

Rink, S. M., & Hopkins, P. B. (1995). A mechlorethamine-induced DNA interstrand cross-link bends duplex DNA. Biochemistry, 34(4), 1439–1445. https://doi.org/10.1021/BI00004A039

Robbins, J., Dilworth, S. M., Laskey, R. A., & Dingwall, C. (1991). Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. Cell, 64(3), 615–623. https://doi.org/10.1016/0092-8674(91)90245-t

Roberts, J. J., & Friedlos, F. (1987). Quantitative estimation of cisplatin-induced DNA interstrand cross-links and their repair in mammalian cells: Relationship to toxicity. Pharmacology & Therapeutics, 34(2), 215–246. https://doi.org/10.1016/0163-7258(87)90012-X

Robertson, A. B., Klungland, A., Rognes, T., & Leiros, I. (2009). Base excision repair: the long and short of it. https://doi.org/10.1007/s00018-009-8736-z

Robson, C. N., & Hickson, I. D. (1991). Isolation of cDNA clones encoding a human apurinic/apyrimidinic endonuclease that corrects DNA repair and mutagenesis defects in E.coli xth (exonuclease III) mutants. Nucleic Acids Research, 19(20), 5519–5523. https://doi.org/10.1093/NAR/19.20.5519

Rojas, F., & Matthews, K. R. (2019). Quorum sensing in African trypanosomes. *Current Opinion in Microbiology*, *52*, 124–129. https://doi.org/10.1016/j.mib.2019.07.001

Rycenga, H. B., & Long, D. T. (2018). The evolving role of DNA inter-strand crosslinks in chemotherapy. In Current Opinion in Pharmacology (Vol. 41, pp. 20–26). Elsevier Ltd. https://doi.org/10.1016/j.coph.2018.04.004

Rothwell, D. G., & Hickson, L. D. (1996). Asparagine 212 is essential for abasic site recognition by the human DNA repair endonuclease HAP1. Nucleic Acids Research, 24(21), 4217–4221. https://doi.org/10.1093/NAR/24.21.4217

Rouillon, C., & White, M. F. (2011). The evolution and mechanisms of nucleotide excision repair proteins. Research in Microbiology, 162(1), 19–26. https://doi.org/10.1016/J.RESMIC.2010.09.003

Rycenga, H. B., & Long, D. T. (2018). The evolving role of DNA inter-strand crosslinks in chemotherapy. In Current Opinion in Pharmacology (Vol. 41, pp. 20–26). Elsevier Ltd. https://doi.org/10.1016/j.coph.2018.04.004

Russo, M. T., De Luca, G., Degan, P., Parlanti, E., Dogliotti, E., Barnes, D. E., Lindahl, T., Yang, H., Miller, J. H., & Bignami, M. (2004). Accumulation of the oxidative base lesion 8-hydroxyguanine in DNA of tumor-prone mice defective in both the Myh and Ogg1 DNA glycosylases. Cancer Research, 64(13), 4411–4414. https://doi.org/10.1158/0008-5472.CAN-04-0355

Sancar, A., Lindsey-Boltz, L. A., ¨ Nsal-Kaçmaz, K. U., & Linn, S. (2004). MOLECULAR MECHANISMS OF MAMMALIAN DNA REPAIR AND THE DNA DAMAGE CHECKPOINTS. Annu. Rev. Biochem, 73, 39–85. https://doi.org/10.1146/annurev.biochem.73.011303.073723

Sadoughi, F., Mirsafaei, L., Dana, P. M., Hallajzadeh, J., Asemi, Z., Mansournia, M. A., Montazer, M., Hosseinpour, M., & Yousefi, B. (2021). The role of DNA damage response in chemo- and radio-resistance of cancer cells: Can DDR inhibitors sole the problem? DNA Repair, 101, 1568–7864. https://doi.org/10.1016/j.dnarep.2021.103074

Saito, Y., Uraki, F., Nakajima, S., Asaeda, A., Ono, K., Kubo, K., & Yamamoto, K. (1997). Characterization of endonuclease III (nth) and endonuclease VIII (nei) mutants of Escherichia coli K-12. Journal of Bacteriology, 179(11), 3783–3785. https://doi.org/10.1128/jb.179.11.3783-3785.1997

Sarker, A. H., Tsutakawa, S. E., Kostek, S., Ng, C., Shin, D. S., Peris, M., Campeau, E., Tainer, J. A., Nogales, E., & Cooper, P. K. (2005). Recognition of RNA polymerase II and transcription bubbles by XPG, CSB, and TFIIH: Insights for transcription-coupled repair and Cockayne syndrome. *Molecular Cell*, *20*(2), 187–198. https://doi.org/10.1016/J.MOLCEL.2005.09.022

Sartori, A. A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., Baer, R., Lukas, J., & Jackson, S. P. (2007). Human CtIP promotes DNA end resection. 450. https://doi.org/10.1038/nature06337

Schamber-Reis, B. L., Nardelli, S., Régis-Silva, C. G., Campos, P. C., Cerqueira, P. G., Lima, S. A., Franco, G. R., Macedo, A. M., Pena, S. D., Cazaux, C., Hoffmann, J. S., Motta, M. C., Schenkman, S., Teixeira, S. M., & Machado, C. R. (2012). DNA polymerase beta from Trypanosoma cruzi is involved in kinetoplast DNA replication and repair of oxidative lesions. Molecular and biochemical parasitology, 183(2), 122–131. https://doi.org/10.1016/j.molbiopara.2012.02.007

Schärer, O. D. (2013). Nucleotide Excision Repair in Eukaryotes. Cold Spring Harbor Perspectives in Biology, 5(10). https://doi.org/10.1101/CSHPERSPECT.A012609

Schmidt, G. D., & Roberts, L. S. (1977). Foundations of parasitology. CV Mosby Company, 11830 Westline Industrial Drive, St. Louis, Missouri 63141, USA (distributed in UK by Henry Kimpton Publishers, 7 Leighton Place, Leighton Road, London NW52QL)..Robertson, A. B., Klungland, A., Rognes, T., & Leiros, I. (2009). Base excision repair: the long and short of it. https://doi.org/10.1007/s00018-009-8736-z

Schofield, C. J., Jannin, J., & Salvatella, R. (2006). The future of Chagas disease control. Trends in parasitology, 22(12), 583–588. https://doi.org/10.1016/j.pt.2006.09.011

Schumann-Burkard, G., Jutzi, P., & Roditi, I. (2011). Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters. Molecular and Biochemical Parasitology, 175(1), 91–94. https://doi.org/10.1016/j.molbiopara.2010.09.002

Scott, B. R., Pathak, M. A., & Mohn, G. R. (1976). Molecular and genetic basis of furocoumarin reactions. Mutation Research/Reviews in Genetic Toxicology, 39(1), 29–74. https://doi.org/10.1016/0165-1110(76)90012-9

Scrima, A., Koníčková, R., Czyzewski, B. K., Kawasaki, Y., Jeffrey, P. D., Groisman, R., Nakatani, Y., Iwai, S., Pavletich, N. P., & Thomä, N. H. (2008). Structural Basis of UV DNA-Damage Recognition by the DDB1–DDB2 Complex. Cell, 135(7), 1213–1223. https://doi.org/10.1016/J.CELL.2008.10.045

Sedgwick, B., Bates, P. A., Paik, J., Jacobs, S. C., & Lindahl, T. (2007). Repair of alkylated DNA: Recent advances. DNA Repair, 6(4), 429–442. https://doi.org/10.1016/J.DNAREP.2006.10.005

Sengerová, B., Allerston, C. K., Abu, M., Lee, S. Y., Hartley, J., Kiakos, K., Schofield, C. J., Hartley, J. A., Gileadi, O., & McHugh, P. J. (2012). Characterization of the human SNM1A and SNM1B/Apollo DNA

repair exonucleases. Journal of Biological Chemistry, 287(31), 26254–26267. https://doi.org/10.1074/JBC.M112.367243

Sengerová, B., Wang, A. T., & McHugh, P. J. (2011). Orchestrating the nucleases involved in DNA interstrand cross-link (ICL) repair. Cell Cycle, 10(23), 3999–4008. https://doi.org/10.4161/cc.10.23.18385

Sieber, O. M., Lipton, L., Crabtree, M., Heinimann, K., Fidalgo, P., Phillips, R. K. S., Bisgaard, M.-L., Orntoft, T. F., Aaltonen, L. A., Hodgson, S. v., Thomas, H. J. W., & Tomlinson, I. P. M. (2003). Multiple Colorectal Adenomas, Classic Adenomatous Polyposis, and Germ-Line Mutations in MYH. *New England Journal of Medicine*, *348*(9), 791–799. https://doi.org/10.1056/nejmoa025283

Shaner, N. C., Lambert, G. G., Chammas, A., Ni, Y., Cranfill, P. J., Baird, M. A., Sell, B. R., Allen, J. R., Day, R. N., Israelsson, M., Davidson, M. W., Wang, J., & Org, N. (2013). A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. https://doi.org/10.1038/nmeth.2413

Shapiro, T.A, & Englund, P. T. (1990). Selective cleavage of kinetoplast DNA minicircles promoted by antitrypanosomal drugs. Proceedings of the National Academy of Sciences of the United States of America, 87(3), 950–954.

Shapiro, T. A., & Englund, P. T. (1995). The structure and replication of kinetoplast DNA. Annual review of microbiology, 49, 117–143. https://doi.org/10.1146/annurev.mi.49.100195.001001

Sharma, S., & Canman, C. E. (2012). REV1 and DNA polymerase zeta in DNA interstrand crosslink repair. Environmental and Molecular Mutagenesis, 53(9), 725. https://doi.org/10.1002/EM.21736

Sheila S. David1, V. L. O. & S. K. 2. (2007). Base-excision repair of oxidative DNA damage. https://doi.org/10.1038/nature05978

Shi, H., Djikeng, A., Mark, T., Wirtz, E., Tschudi, C., & Ullu, E. (2000). Genetic interference in Trypanosoma brucei by heritable and inducible double-stranded RNA. RNA (New York, N.Y.), 6(7), 1069–1076. https://doi.org/10.1017/s1355838200000297

Siddiqui, M. Q., Rajpurohit, Y. S., Thapa, P. S., Maurya, G. K., Banerjee, K., Khan, M. A., Panda, P., Hasan, S. K., Gadewal, N., Misra, H. S., & Varma, A. K. (2017). Studies of protein-protein interactions in Fanconi anemia pathway to unravel the DNA interstrand crosslink repair mechanism. International journal of biological macromolecules, 104(Pt A), 1338–1344. https://doi.org/10.1016/j.ijbiomac.2017.05.166

Siegel, T. N., Hekstra, D. R., & Cross, G. A. (2008). Analysis of the Trypanosoma brucei cell cycle by quantitative DAPI imaging. Molecular and biochemical parasitology, 160(2), 171–174. https://doi.org/10.1016/j.molbiopara.2008.04.004

Sieber, O. M., Lipton, L., Crabtree, M., Heinimann, K., Fidalgo, P., Phillips, R. K. S., Bisgaard, M.-L., Orntoft, T. F., Aaltonen, L. A., Hodgson, S. v., Thomas, H. J. W., & Tomlinson, I. P. M. (2003). Multiple Colorectal Adenomas, Classic Adenomatous Polyposis, and Germ-Line Mutations in MYH . New England Journal of Medicine, 348(9), 791–799. https://doi.org/10.1056/nejmoa025283

Simpson, L., Sbicego, S., & Aphasizhev, R. (2003). Uridine insertion/deletion RNA editing in trypanosome mitochondria: A complex business. RNA, 9(3), 265–276. https://doi.org/10.1261/RNA.2178403

Singhal, R. K., & Wilson, S. H. (1993). Short gap-filling synthesis by DNA polymerase β is processive. Journal of Biological Chemistry, 268(21), 15906–15911. https://doi.org/10.1016/S0021-9258(18)82338-9

Sliwoski, G., Kothiwale, S., Meiler, J., & Lowe, E.W. (2014). Computational Methods in Drug Discovery. Pharmacological Reviews, 66, 334 - 395.

Smeaton, M. B., Hlavin, E. M., Mason, T. M. G., Noronha, A. M., Wilds, C. J., & Miller, P. S. (2008). Distortion-dependent unhooking of interstrand cross-links in mammalian cell extracts. Biochemistry, 47(37), 9920. https://doi.org/10.1021/BI800925E

Smogorzewska, A., Desetty, R., Saito, T. T., Schlabach, M., Lach, F. P., Sowa, M. E., Clark, A. B., Kunkel, T. A., Harper, J. W., Colaiácovo, M. P., & Elledge, S. J. (2010). A genetic screen identifies FAN1, a Fanconi anemia associated nuclease necessary for DNA interstrand crosslink repair. Molecular Cell, 39(1), 36. https://doi.org/10.1016/J.MOLCEL.2010.06.023

Sobol, R. W., Horton, J. K., Kühn, R., Gu, H., Singhal, R. K., Prasad, R., Rajewsky, K., & Wilson, S. H. (1996). Requirement of mammalian DNA polymerase-beta in base-excision repair. Nature, 379(6561), 183–186. https://doi.org/10.1038/379183A0

Steverding, D. (2008). The history of African trypanosomiasis. Parasites & Vectors, 1(1), 3. https://doi.org/10.1186/1756-3305-1-3

Stevens, J. R., Noyes, H. A., Schofield, C. J., & Gibson, W. (2001). The molecular evolution of trypanosomatidae. In *Advances in Parasitology* (Vol. 48, pp. 1–53). Academic Press. https://doi.org/10.1016/S0065-308X(01)48003-1

Stevnsner T, Muftuoglu M, Aamann MD, Bohr VA. The role of Cockayne Syndrome group B (CSB) protein in base excision repair and aging. Mechanisms of Ageing and Development. 2008 Jul-Aug;129(7-8):441-448. DOI: 10.1016/j.mad.2008.04.009. PMID: 18541289; PMCID: PMC2538557.

Sternberg, J., & Tait, A. (1990). Genetic exchange in African trypanosomes. *Trends in Genetics : TIG*, *6*(10), 317–322. https://doi.org/10.1016/0168-9525(90)90252-2

Stich, A., Abel, P. M., & Krishna, S. (2002). Human African trypanosomiasis. British Medical Journal (BMJ), 325(July), 203–206.

Stone, M. P., Cho, Y. J., Huang, H., Kim, H. Y., Kozekov, I. D., Kozekova, A., Wang, H., Minko, I. G., Lloyd, R. S., Harris, T. M., & Rizzo, C. J. (2008). Interstrand DNA cross-links induced by α,β-unsaturated aldehydes derived from lipid peroxidation and environmental sources. Accounts of Chemical Research, 41(7), 793–804. https://doi.org/10.1021/ar700246x

Story, R. M., & Steitz, T. A. (1992). Structure of the recA protein-ADP complex. In NATURE · (Vol. 355).

Stork, C. M., & Schreffler, S. M. (2022). Cyclophosphamide. Encyclopedia of Toxicology: Third Edition, 1111–1113. https://doi.org/10.1016/B978-0-12-386454-3.00720-X

Stuart, K, Brun, R., Croft, S., Fairlamb, A., Gürtler, R., McKerrow, J., Reed, S., & Tarleton, R. (2008). Kinetoplastids: related protozoan pathogens, different diseases. The Journal of Clinical Investigation, 118(4), 1301–1310. https://doi.org/10.1172/JCI33945

Sugasawa, K. (2019). Mechanism and regulation of DNA damage recognition in mammalian nucleotide excision repair. Enzymes, 45, 99–138. https://doi.org/10.1016/bs.enz.2019.06.004

Sugasawa, K., Okamoto, T., Shimizu, Y., Masutani, C., Iwai, S., & Hanaoka, F. (2001). A multistep damage recognition mechanism for global genomic nucleotide excision repair. Genes and Development, 15(5), 507–521. https://doi.org/10.1101/GAD.866301

Sullivan, J. A., Tong, J. L., Wong, M., Kumar, A., Sarkar, H., Ali, S., Hussein, I., Zaman, I., Meredith, E. L., Helsby, N. A., Hu, L., & Wilkinson, S. R. (2015). Unravelling the role of SNM1 in the DNA repair system of Trypanosoma brucei. Molecular Microbiology, 96(4), 827–838. https://doi.org/10.1111/mmi.12973

Tan, K. S. W., Leal, S. T. G., & Cross, G. A. M. (2002). Trypanosoma brucei MRE11 is non-essential but influences growth, homologous recombination, and DNA double-strand break repair. Molecular and Biochemical Parasitology, 125(1–2), 11–21. https://doi.org/10.1016/S0166-6851(02)00165-2

Tapias, A., Auriol, J., Forget, D., Enzlin, J. H., Schärer, O. D., Coin, F., Coulombe, B., & Egly, J. M. (2004). Ordered Conformational Changes in Damaged DNA Induced by Nucleotide Excision Repair Factors. Journal of Biological Chemistry, 279(18), 19074–19083. https://doi.org/10.1074/JBC.M312611200

Taladriz, S., Hanke, T., Ramiro, M. J., García-Díaz, M., García De Lacoba, M., Blanco, L., & Larraga, V. (2001). Nuclear DNA polymerase beta from Leishmania infantum. Cloning, molecular analysis and developmental regulation. Nucleic acids research, 29(18), 3822–3834. https://doi.org/10.1093/nar/29.18.3822

Tan, W., & Deans, A. J. (2017). A defined role for multiple Fanconi anemia gene products in DNA-damage-associated ubiquitination. *Experimental hematology*, *50*, 27–32. https://doi.org/10.1016/j.exphem.2017.03.001

Taylor, M. C., Kaur, H., Blessington, B., Kelly, J. M., & Wilkinson, S. R. (2008). Validation of spermidine synthase as a drug target in African trypanosomes. Biochemical Journal, 409(2), 563–569. https://doi.org/10.1042/BJ20071185

Tebbs, R. S., Flannery, M. L., Meneses, J. J., Hartmann, A., Tucker, J. D., Thompson, L. H., Cleaver, J. E., & Pedersen, R. A. (1999). Requirement for the Xrcc1 DNA base excision repair gene during early mouse development. Developmental Biology, 208(2), 513–529. https://doi.org/10.1006/DBIO.1999.9232

Ten Asbroek, A. L. M. A., Ouellette, M., & Borst, P. (1990). Targeted insertion of the neomycin phosphotransferase gene into the tubulin gene cluster of Trypanosoma brucei. Nature, 348(6297), 174–175. https://doi.org/10.1038/348174a0

Thakare, R., Dasgupta, A., & Chopra, S. (2020). Pretomanid for the treatment of pulmonary tuberculosis. Drugs of today (Barcelona, Spain : 1998), 56(10), 655–668. https://doi.org/10.1358/dot.2020.56.10.3161237

Thivolle, A., Mehnert, A. K., Tihon, E., McLaughlin, E., Dujeancourt-Henry, A., & Glover, L. (2021). DNA double strand break position leads to distinct gene expression changes and regulates VSG switching pathway choice. PLoS Pathogens, 17(11). https://doi.org/10.1371/journal.ppat.1010038

Thomas, J., Baker, N., Hutchinson, S., Dominicus, C., Trenaman, A., Glover, L., Alsford, S., & Horn, D. (2018). Insights into antitrypanosomal drug mode-of-action from cytology-based profiling. *PLoS Neglected Tropical Diseases*, *12*(11). https://doi.org/10.1371/JOURNAL.PNTD.0006980

Thongthip, S., Bellani, M., Gregg, S. Q., Sridhar, S., Conti, B. A., Chen, Y., Seidman, M. M., & Smogorzewska, A. (2016). Fan1 deficiency results in DNA interstrand cross-link repair defects, enhanced tissue karyomegaly, and organ dysfunction. *Genes & Development*, *30*(6), 645–659. https://doi.org/10.1101/gad.276261.115

Tiwari, V., Baptiste, B. A., Okur, M. N., & Bohr, V. A. (2021). Current and emerging roles of Cockayne syndrome group B (CSB) protein. *Nucleic Acids Research*, *49*(5), 2418–2434. https://doi.org/10.1093/nar/gkab085

Trenaman, A., Hartley, C., Prorocic, M., Passos-Silva, D. G., Hoek, M. Van Den, Nechyporuk-Zloy, V., MacHado, C. R., & McCulloch, R. (2013). Trypanosoma brucei BRCA2 acts in a life cycle-specific genome stability process and dictates BRC repeat number-dependent RAD51 subnuclear dynamics. Nucleic Acids Research, 41(2), 943. https://doi.org/10.1093/NAR/GKS1192

Trochine, A., Alvarez, G., Corre, S., Faral-Tello, P., Durán, R., Batthyany, C. I., Cerecetto, H., González, M., & Robello, C. (2014). Trypanosoma cruzi chemical proteomics using immobilized benznidazole. Experimental Parasitology, 140(1), 33–38. https://doi.org/10.1016/j.exppara.2014.03.013

Truong, L. N., Li, Y., Shi, L. Z., Hwang, P. Y. H., He, J., Wang, H., Razavian, N., Berns, M. W., & Wu, X. (2013). Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America, 110(19), 7720–7725. https://doi.org/10.1073/PNAS.1213431110/-/DCSUPPLEMENTAL

Tsodikov, O. V., Enzlin, J. H., Schärer, O. D., & Ellenberger, T. (2005). Crystal structure and DNA binding functions of ERCC1, a subunit of the DNA structure-specific endonuclease XPF-ERCC1. Proceedings of the National Academy of Sciences of the United States of America, 102(32), 11236–11241. https://doi.org/10.1073/pnas.0504341102

Tsutakawa, S. E., Shin, D. S., Mol, C. D., Izumi, T., Arvai, A. S., Mantha, A. K., Szczesny, B., Ivanov, I. N., Hosfield, D. J., Maiti, B., Pique, M. E., Frankel, K. A., Hitomi, K., Cunningham, R. P., Mitra, S., & Tainer, J. A. (2013). Conserved Structural Chemistry for Incision Activity in Structurally Non-homologous Apurinic/Apyrimidinic Endonuclease APE1 and Endonuclease IV DNA Repair Enzymes*. Journal of Biological Chemistry, 288(12), 8445–8455. https://doi.org/https://doi.org/10.1074/jbc.M112.422774

Tweats, D., Bourdin Trunz, B., & Torreele, E. (2012). Genotoxicity profile of fexinidazole—a drug candidate in clinical development for human African trypanomiasis (sleeping sickness). *Mutagenesis*, *27*(5), 523–532. https://doi.org/10.1093/mutage/ges015

Ubeda, J. M., Raymond, F., Mukherjee, A., Plourde, M., Gingras, H., Roy, G., Lapointe, A., Leprohon, P., Papadopoulou, B., Corbeil, J., & Ouellette, M. (2014). Genome-Wide Stochastic Adaptive DNA Amplification at Direct and Inverted DNA Repeats in the Parasite Leishmania. PLoS Biology, 12(5). https://doi.org/10.1371/JOURNAL.PBIO.1001868

Ullu, E., Matthews, K. R., & Tschudi, C. (1993). Temporal order of RNA-processing reactions in trypanosomes: rapid trans splicing precedes polyadenylation of newly synthesized tubulin transcripts. Molecular and Cellular Biology, 13(1), 720–725. https://doi.org/10.1128/MCB.13.1.720-725.1993

Vangipuram, R., & Feldman, S. (2015). Ultraviolet phototherapy for cutaneous diseases: a concise review. Wiley Online Library, 22(4), 253–259. https://doi.org/10.1111/odi.12366

Vermeulen, W., Rademakers, S., Jaspers, N. G. J., Appeldoorn, E., Raams, A., Klein, B., Kleijer, W. J., Kjærsgård Hansen, L., & Hoeijmakers, J. H. J. (2001). A temperature-sensitive disorder in basal transcription and DNA repair in humans. Nature Genetics 2001 27:3, 27(3), 299–303. https://doi.org/10.1038/85864

Vickerman, K. (1965). Polymorphism and mitochondrial activity in sleeping sickness. Nature., 208(5012), 762–766. https://doi.org/10.1038/208762a0

Vidal, A. E., Harkiolaki, M., Gallego, C., Castillo-Acosta, V. M., Ruiz-Pérez, L. M., Wilson, K., & González-Pacanowska, D. (2007). Crystal Structure and DNA Repair Activities of the AP Endonuclease from Leishmania major. Journal of Molecular Biology, 373(4), 827–838. https://doi.org/10.1016/J.JMB.2007.08.001

Vieira-da-Rocha, J. P., Passos-Silva, D. G., Mendes, I. C., Rocha, E. A., Gomes, D. A., Machado, C. R., & McCulloch, R. (2019). The DNA damage response is developmentally regulated in the African trypanosome. *DNA Repair*, *73*, 78–90. https://doi.org/10.1016/j.dnarep.2018.11.005

Vink, C., Rudenko, G., & Seifert, H. S. (2012). Microbial antigenic variation mediated by homologous DNA recombination. FEMS Microbiology Reviews, 36(5), 917–948. https://doi.org/10.1111/J.1574-6976.2011.00321.X

Vispé, S., Cazaux, C., Lesca, C., & Defais, M. (1998). Overexpression of Rad51 protein stimulates homologous recombination and increases resistance of mammalian cells to ionizing radiation. In *Nucleic Acids Research* (Vol. 26, Issue 12).

Volker, M., Moné, M. J., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., Hoeijmakers, J. H. J., van Driel, R., van Zeeland, A. A., & Mullenders, L. H. F. (2001). Sequential assembly of the nucleotide excision repair factors in vivo. *Molecular Cell*, *8*(1), 213–224. https://doi.org/10.1016/S1097-2765(01)00281-7

Walker, L. J., Craig, R. B., Harris, A. L., & Hickson, I. D. (1994). A role for the human DNA repair enzyme HAP1 in cellular protection against DNA damaging agents and hypoxic stress. *Nucleic Acids Research*, *22*(23), 4884–4889. https://doi.org/10.1093/NAR/22.23.4884

Walters, K. J., Lech, P. J., Goh, A. M., Wang, Q., & Howley, P. M. (2003). DNA-repair protein hHR23a alters its protein structure upon binding proteasomal subunit S5a. Proceedings of the National Academy of Sciences of the United States of America, 100(22), 12694–12699. https://doi.org/10.1073/PNAS.1634989100

Wang, A. T., Sengerová, B., Cattell, E., Inagawa, T., Hartley, J. M., Kiakos, K., Burgess-Brown, N. A., Swift, L. P., Enzlin, J. H., Schofield, C. J., Gileadi, O., Hartley, J. A., & McHugh, P. J. (2011). Human SNM1a and XPF-ERCC1 collaborate to initiate DNA interstrand cross-link repair. *Genes and Development*, *25*(17), 1859–1870. https://doi.org/10.1101/GAD.15699211

Wang, H., Perrault, A. R., Takeda, Y., Qin, W., Wang, H., & Iliakis, G. (2003). Biochemical evidence for Ku-independent backup pathways of NHEJ. *Nucleic Acids Research*, *31*(18), 5377. https://doi.org/10.1093/NAR/GKG728

Wang, R., Persky, N. S., Yoo, B., Ouerfelli, O., Smogorzewska, A., Elledge, S. J., & Pavletich, N. P. (2014). DNA repair. Mechanism of DNA interstrand cross-link processing by repair nuclease FAN1. *Science (New York, N.Y.)*, *346*(6213), 1127–1130. https://doi.org/10.1126/science.1258973

Wang, W. (2007). Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. *Nature Reviews Genetics*, *8*(10), 735–748. https://doi.org/10.1038/NRG2159

Wang, Z., Wu, X., & Friedberg, E. C. (1993). DNA Repair Synthesis during Base Excision Repair In Vitro Is Catalyzed by DNA Polymerase E and Is Influenced by DNA Polymerases a and 8 in Saccharomyces cerevisiae. *MOLECULAR AND CELLULAR BIOLOGY*, *13*(2), 1051–1058.

Wang, Z., Morris, J. C., Drew, M. E., & Englund, P. T. (2000). Inhibition of Trypanosoma brucei Gene Expression by RNA Interference Using an Integratable Vector with Opposing T7 Promoters *. *Journal of Biological Chemistry*, *275*(51), 40174–40179. https://doi.org/10.1074/JBC.M008405200

Wang, Z., & Englund, P. T. (2001). RNA interference of a trypanosome topoisomerase II causes progressive loss of mitochondrial DNA. *The EMBO Journal*, *20*(17), 4674–4683. https://doi.org/10.1093/EMBOJ/20.17.4674

Ward, I. M., & Chen, J. (2001). Histone H2AX Is Phosphorylated in an ATR-dependent Manner in Response to Replicational Stress *. Journal of Biological Chemistry, 276(51), 47759–47762. https://doi.org/10.1074/JBC.C100569200

Ward, T. A., Dudášová, Z., Sarkar, S., Bhide, M. R., Vlasáková, D., Chovanec, M., & McHugh, P. J. (2012). Components of a Fanconi-like pathway control Pso2-independent DNA interstrand crosslink repair in yeast. PLoS genetics, 8(8), e1002884. https://doi.org/10.1371/journal.pgen.1002884

Warmerdam, D. O., & Kanaar, R. (2010). Dealing with DNA damage: Relationships between checkpoint and repair pathways. Mutation Research/Reviews in Mutation Research, 704(1–3), 2–11. https://doi.org/10.1016/J.MRREV.2009.12.001

Wei, L., Levine, A. S., & Lan, L. (2016). Transcription-coupled homologous recombination after oxidative damage. DNA repair, 44, 76–80. https://doi.org/10.1016/j.dnarep.2016.05.009

Weilbeer, C., Jay, D., Donnelly, J. C., Gentile, F., Karimi-Busheri, F., Yang, X., Mani, R. S., Yu, Y., Elmenoufy, A. H., Barakat, K. H., Tuszynski, J. A., Weinfeld, M., & West, F. G. (2022). Modulation of ERCC1-XPF Heterodimerization Inhibition via Structural Modification of Small Molecule Inhibitor Side-Chains. *Frontiers in Oncology*, *12*. https://doi.org/10.3389/FONC.2022.819172

Wendy J. Cannan and David S. Pederson. (2017). Mechanisms and Consequences of Double-strand DNA Break Formation in Chromatin. Physiology & Behavior, 176(5), 139–148. https://doi.org/10.1002/jcp.25048.Mechanisms

WHO. (2021). Trypanosomiasis, human African (sleeping sickness). https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness)

Wiederhold, L., Leppard, J. B., Kedar, P., Karimi-Busheri, F., Rasouli-Nia, A., Weinfeld, M., Tomkinson, A. E., Izumi, T., Prasad, R., Wilson, S. H., Mitra, S., & Hazra, T. K. (2004). AP Endonuclease-Independent DNA Base Excision Repair in Human Cells. Molecular Cell, 15(2), 209–220. https://doi.org/10.1016/J.MOLCEL.2004.06.003

Wilkinson, S. R., Taylor, M. C., Horn, D., Kelly, J. M., & Cheeseman, I. (2008). A mechanism for cross-resistance to nifurtimox and benznidazole in trypanosomes. Proceedings of the National Academy of Sciences of the United States of America, 105(13), 5022–5027. https://doi.org/10.1073/PNAS.0711014105

Wilkinson, S. R., & Kelly, J. M. (2009). Trypanocidal drugs: Mechanisms, resistance and new targets. Expert Reviews in Molecular Medicine, 11(October 2009), 1–24. https://doi.org/10.1017/S1462399409001252

Wirtz, E., Hartmann, C., & Clayton, C. (1994). Gene expression mediated by bacteriophage T3 and T7 RNA polymerases in transgenic trypanosomes. *Nucleic Acids Research*, *22*(19), 3887. https://doi.org/10.1093/NAR/22.19.3887

Wirtz, E., & Clayton, C. (1995). Inducible gene expression in trypanosomes mediated by a prokaryotic repressor. *Science*, *268*(5214), 1179–1183. https://doi.org/10.1126/science.7761835

Wirtz, E., Hoek, M., & Cross, G. A. M. (1998). Regulated processive transcription of chromatin by T7 RNA polymerase in Trypanosoma brucei. *Nucleic Acids Research*, *26*(20), 4626–4634. https://doi.org/10.1093/nar/26.20.4626

Wirtz, E., Leal, S., Ochatt, C., & Cross, G. M. (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. Molecular and biochemical parasitology, 99(1), 89-101.

Wiser, M. F. (2011). Protozoa and Human disease. Garland Science.

Wood, R. D. (2010). Nucleotide excision repair proteins and interstrand crosslink repair. https://doi.org/10.1002/em.20569

Woodward, R., & Gull, K. (1990). Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of Trypanosoma brucei. Journal of Cell Science, 95(1), 49–57. https://doi.org/10.1242/JCS.95.1.49

Xie, Y., Yang, H., Cunanan, C., Okamoto, K., Shibata, D., Pan, J., Barnes, D. E., Lindahl, T., Mcilhatton, M., Fishel, R., & Miller, J. H. (2004). Deficiencies in Mouse Myh and Ogg1 Result in Tumor Predisposition and G to T Mutations in Codon 12 of the K-Ras Oncogene in Lung Tumors. In CANCER RESEARCH (Vol. 64). http://aacrjournals.org/cancerres/article-pdf/64/9/3096/2529951/zch00904003096.pdf

Xu, J., Lahiri, I., Wang, W., Wier, A., Cianfrocco, M. A., Chong, J., Hare, A. A., Dervan, P. B., DiMaio, F., Leschziner, A. E., & Wang, D. (2017). Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. Nature, 551(7682), 653–657. https://doi.org/10.1038/nature24658

Xanthoudakis, S., Smeyne, R. J., Wallace, J. D., & Curran, T. (1996). The redox/DNA repair protein, Ref-1, is essential for early embryonic development in mice. Proceedings of the National Academy of Sciences of the United States of America, 93(17), 8919–8923. https://doi.org/10.1073/PNAS.93.17.8919

Yamamoto, K. N., Kobayashi, S., Tsuda, M., Kurumizaka, H., Takata, M., Kono, K., Jiricny, J., Takeda, S., & Hirota, K. (2011). Involvement of SLX4 in interstrand cross-link repair is regulated by the Fanconi anemia pathway. Proceedings of the National Academy of Sciences of the United States of America, 108(16), 6492–6496. https://doi.org/10.1073/pnas.1018487108

Yosaatmadja, Y., Baddock, H. T., Newman, J. A., Bielinski, M., Gavard, A. E., Mukhopadhyay, S. M. M., Dannerfjord, A. A., Schofield, C. J., McHugh, P. J., & Gileadi, O. (2021). Structural and mechanistic insights into the Artemis endonuclease and strategies for its inhibition. *Nucleic Acids Research*, *49*(16), 9310–9326. https://doi.org/10.1093/nar/gkab693

Yun, O., Priotto, G., Tong, J., Flevaud, L., & Chappuis, F. (2010). NECT is next: implementing the new drug combination therapy for Trypanosoma brucei gambiense sleeping sickness. *PLoS Neglected Tropical Diseases*, *4*(5). https://doi.org/10.1371/JOURNAL.PNTD.0000720

Zhang, J., Dewar, J. M., Budzowska, M., Motnenko, A., Cohn, M. A., & Walter, J. C. (2015). DNA interstrand cross-link repair requires replication-fork convergence. Nature Structural & Molecular Biology 2015 22:3, 22(3), 242–247. https://doi.org/10.1038/nsmb.2956

Zhang, J., & Walter, J. C. (2014). Mechanism and regulation of incisions during DNA interstrand cross-link repair. DNA Repair, 19, 135–142.

Zheng, H., Wang, X., Warren, A. J., Legerski, R. J., Nairn, R. S., Hamilton, J. W., & Li, L. (2003). Nucleotide excision repair-and polymerase η-mediated error-prone removal of mitomycin C interstrand cross-links. Molecular and cellular biology, 23(2), 754-761.

Zhao, Q., Wang, Q. E., Ray, A., Wani, G., Han, C., Milum, K., & Wani, A. A. (2009). Modulation of nucleotide excision repair by mammalian SWI/SNF chromatin-remodeling complex. Journal of Biological Chemistry, 284(44), 30424-30432.

.