Molecular Characterisation of the *Trypanosoma brucei* Flap Endonuclease

By:

Sarah Louise Oates

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Medicine, Dentistry and Health
Department of Infection and Immunity

April 2016

Abstract

Kinetoplastid parasites pose a huge global health burden on both humans and livestock. Current treatments are toxic and difficult to administer. With the emergence of drug-resistant organisms, it is essential to identify new therapeutic targets. African trypanosomiasis, or sleeping sickness, is caused by *Trypanosoma brucei*. This thesis focuses on the characterisation and validation of a potential new drug target in this kinetoplastid parasite. Flap endonucleases are essential enzymes involved in DNA replication and repair. Synthetic constructs, encoding wild type and mutated FEN genes, were produced and expressed in *Escherichia coli*. Purification procedures and quantitative enzyme activity assays were developed. The catalytic and binding activity of the D183K variant were determined and compared to the wild type enzyme. The former was shown to be catalytically inert, yet able to bind a DNA substrate with similar affinity to the wild type enzyme (K_d = 2.22 and 1.37 µM respectively). The D183K variant was crystallised collaboratively and its structure determined by colleagues in the Department of Molecular Biology & Biotechnology. The wild type and D183K variant genes were over-expressed in *T. brucei* bloodstream form cells, using a tetracycline-inducible system. Expression of the catalytically inert protein had a severely detrimental effect on cell growth, and morphological changes were observed 72 hours post-induction. Knock-down assays were attempted using siRNA techniques, but results were inconclusive. The recombinant wild type enzyme was targeted with known FEN inhibitors. Myricetin, a commercially available flavonoid, inhibited *T. brucei* flap endonuclease (IC_{50} = 14.81 µM). *T. brucei* bloodstream form cells were killed upon exposure to myricetin (LD_{50} = 23.63 µM). A high-throughput thermal shift assay was developed to aid fragment-based drug design. This assay can be used to detect binding by myricetin, a known *Tb*FEN inhibitor. Use of this assay will open up further avenues of exploration for targeting the *T. brucei* flap endonuclease.

Acknowledgements

I would like to express my gratitude to my supervisor, Professor Jon Sayers, for providing me with the opportunity to study for my PhD. You have been supportive, and whilst I am sure I have tried your patience on a number of occasions, I appreciate all the help, advice and comments on my work and my thesis.

The members of the Sayer's laboratory and the Department of Infection and Immunity have been a great source of help, solace and cheer throughout my PhD. Of particular mention are Dr Sarbendra Pradhananga, without whom I would have floundered at the very beginning, and Mrs Janine Phipps. I cannot forget to mention the two people who I began my PhD with: Dr Hannah McMellon and Mr Sam Harding. I'm incredibly glad I started with the pair of you, and you, alongside Miss Jenny Ventress, have helped to keep me sane.

I would also like to thank Dr Helen Price, from Keele University, for helping me with the *T. brucei* parasite work. I thoroughly enjoyed working with you. I would also like to thank the late Professor Peter Artymiuk, and his group members Jason Wilson and Dr Faizah Ahmed AlMalki, with whom I collaborated. There are a number of individuals who helped me with experiments, equipment and general discussions. I cannot name you all here, but I appreciate everything all the same.

I would also like to mention my two sports teams: SLRUFC and SULRUFC. I may have disappeared towards the end through injury, time commitments and general 'end-of-PhD' panic, but I thoroughly enjoyed playing for both teams. Thank you for bearing the brunt of my frustrations.

I cannot forget to thank my family. Thanks Mum and Dad for putting up with me (and giving me the occasional kick up the backside). Thanks Dan, Becci, Chrissy and Gran for drinks, food and general support. Thanks Jack and Shelby (the rotties), and Max (the lab) for walks, strokes and general hugs.

Last but not least, I get to my other half, Andrew Berry. You relocated to Sheffield for me, you stuck it out, you supported me throughout and made this process a lot easier. Thank you. Now it's my turn for you.

Contents

List of Figures

List of Tables

Nomenclature

*Hs*FEN	*Homo sapiens* (human) flap endonuclease
427::FEN1	Clones that over-express the wild type *Tb*FEN, under the control of a tetracycline inducible expression system
427::fen1D183K	Clones that over-express the *Tb*FEN D183K variant, under the control of a tetracycline inducible expression system
427‡FEN1	Clones that contain the *Tb*FEN fragment for siRNA knock down
D179K	Residue 179 mutated from an aspartic acid (D) to a lysine (K)
D183K	Residue 183 mutated from an aspartic acid (D) to a lysine (K)
*Tb*FEN	*Trypansoma brucei* flap endonuclease
ΔC	Truncated at the C-terminus
NS	Not statistically significant, $p > 0.05$
*	Statistically significant, where $p \leq 0.05$
**	Statistically significant, where $p \leq 0.01$
***	Statistically significant, where $p \leq 0.001$
****	Statistically significant, where $p \leq 0.0001$

Abbreviations

AAT	Animal African Trypanosomiasis
AP site	apurinic/apyrimidinic site
APE1	AP endonuclease
AT1	Aminopurine transporter
BER	Base excision repair
BLM	Bloom syndrome protein
bp	Base pair
BSA	Bovine serum albumin
BSF	Bloodstream form
CD	Circular dichroism
CDC	Centre for Disease Control
cDNA	Complementary DNA
COFS syndrome	Cerebro-oculo-facio-skeletal syndrome
dA/dT/dC/dG	Deoxynucleotide adenine/thymine/cytosine/guanine
DFMO	Eflornithine
DNA	Deoxyribonucleic acid
DNDi	Drugs for neglected diseases initiative
dNTP	Deoxynucleotide
dsRNA	Double stranded RNA
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EXO1	Exonuclease 1
FAO	Food and Agricultural Organisation
FEN	Flap endonuclease
FRET	Förster resonance energy transfer
FT	Flow through from protein purification column
G4 DNA	G-quadruplex DNA
GEN	Gap endonuclease
H2TH	Helix-two turn-helix
HAPT1	High affinity pentamidine transporter
HAT	Human African Trypanosomiasis
IC_{50}	The concentration at which 50% of enzyme activity is inhibited

IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kbp	Kilobase pair
K_d	Dissociation constant (k_{off}/k_{on})
kDa	Kilodaltons
k_{off}	Dissociation rate constant
k_{on}	Association rate constant
L	Agarose gel 100 bp DNA ladder
LD_{50}	The concentration at which 50% of cells are not viable
Lig I	DNA Ligase I
L_k	Agarose gel 1 Kbp DNA ladder
M	Low range SDS-PAGE protein marker
MMR	Mismatch repair
MMS	Methyl methanesulphonate
mRNA	Messenger RNA
NCI	National Cancer Institute
NECT	Nifurtimox eflornithine combination therapy
NER	Nucleotide excision repair
P	Pellet
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PEI	Polyethyleneimine
PMSF	Phenylmethylsulphonyl fluoride
Pol	Polymerase
POT1	Protection of telomeres protein 1
P-P	Pellet following PEI precipitation
P-SN	Supernatant following PEI precipitation
RFC	Replication factor C
RNA	Ribonucleic acid
RNAi	RNA interference
RPA	Replication protein A
RT qPCR	Reverse transcription quantitative polymerase chain reaction
S	Sample from dialysis bag

S2	Supernatant from dialysis bag
SC	Sample after sonication
SDS	Sodium dodecyl sulphate
siRNA	Short interfering RNA
SN	Supernatant
SOE-PCR	Splicing by overlap extension polymerase chain reaction
SRA	Serum resistance associated gene
TAE	Tris, acetic acid and EDTA
Taq	Thermus aquaticus
TAT1	Alpha tubulin
TBE	Tris, boric acid and EDTA
TBG	Tris, Bicine, EDTA and glycerol
T_M	Melting temperature
TMZ	Temozolamide
TRF2	TTAGGG-repeat factor 2
tRNA	Transfer RNA
TUNEL	Terminal deoxynucleotidyl transferase dUTP nick-end labelling
U	Uncut
Un	Un-induced
UV	Ultra violet
vPBS	Voorheis' modified PBS
VSG	Variable surface glycoprotein
W	Column wash (with loading buffer)
WHO	World Health Organisation
WRN	Werner syndrome protein
XPG	Xeroderma pigmentosum gene

1 Introduction

1.1 African Trypanosomiasis

1.1.1 A Brief History

African trypanosomiasis is caused by kinetoplastid parasites of the *Trypanosoma* genera. This ancient disease was first accurately reported in 1734, although there is evidence indicating its prevalence in Ancient Egypt (Steverding, 2008). The causative agent of the cattle version of the disease (Nagana) was discovered by David Bruce in 1895, and called *Trypanosoma brucei brucei* (Bruce, 1895). It was in the 1920-1940 human African Trypanosomiasis (HAT) epidemic where significant advances in the prevention, control and treatment of the disease occurred. The development of two therapeutics (tryparsamide and suramin) helped in the early stages of the epidemic (Pearce, 1921; Steverding, 2008). The deployment of mobile teams that detected and treated the disease had a significant contribution in controlling both the epidemic and the parasite reservoir (Steverding, 2008). These factors, in combination with the development and use of Pentamidine and Melarsoprol, helped to nearly eliminate HAT in the 1960s (Figure 1-1) (Bray *et al.*, 2003; de Raadt, 2005). The onset of independence from colonial powers resulted in the collapse of many screening and control programmes within African countries. This lead to a re-emergence of the disease that peaked in the 1990s (Figure 1-1) (Brun *et al.*, 2010). Concerted efforts by the World Health Organisation (WHO) and a number of non-governmental organisations has helped to reduce the number of cases by 73% between 2000-2012 (WHO, 2015b). The WHO Neglected Tropical Diseases initiative aims to eliminate HAT by 2020 (WHO, 2015b).

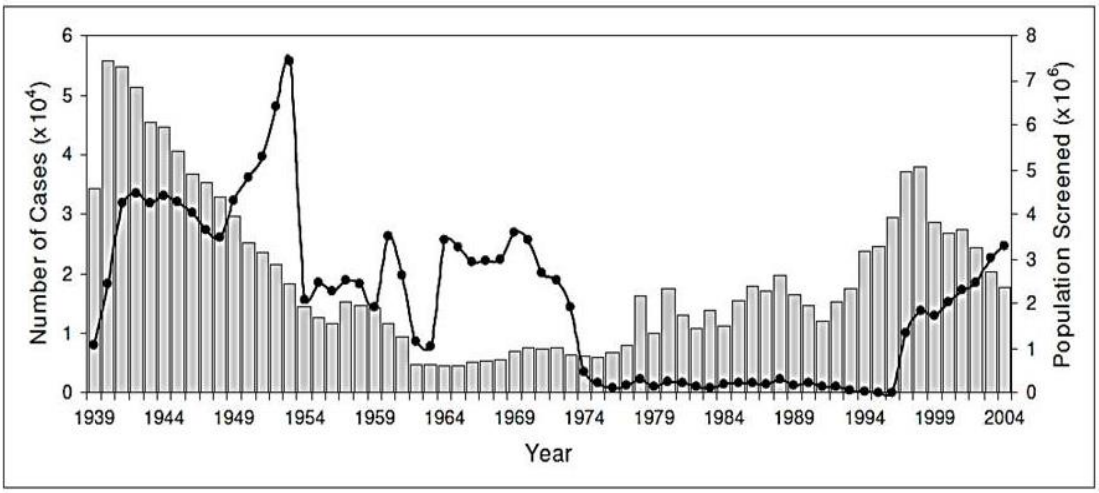


Figure 1-1 Comparison between the number of reported HAT cases and the level of surveillance

Grey columns indicate the number of HAT cases reported between 1939 and 2004. Black circles indicate the level of surveillance against HAT between 1939 and 2004. The graph was reproduced under the creative commons license from Steverding (2008).

1.1.2 Human African Trypanosomiasis

There are two sub-species of *T. brucei* that cause disease in humans: *T. brucei rhodesiense* and *T. brucei gambiense*. *T. brucei gambiense* infections are localised to the western and central regions of sub-Saharan Africa, whilst *T. brucei rhodesiense* infections are localised to the eastern and southern regions (WHO, 2015b). The localisation of the parasites is dependent on the availability of their vector; the tsetse fly (*Glossina spp*). The life cycle of these parasites is complex, as they require two separate hosts in order to develop (Figure 1-2). Briefly, the mammalian host becomes infected by metacyclic trypomastigotes following a bite from an infected tsetse fly (CDC, 2015; Langousis & Hill, 2014). Once in the mammalian blood stream, the parasites transform into bloodstream trypomastigotes, which divide in the blood. These infect the tsetse fly during feeding. Once in the fly midgut, the parasites transform into procyclic trypomastigotes (CDC, 2015; Langousis & Hill, 2014). These transform into epimastigotes and migrate to the salivary glands, where they are ready to infect another mammalian host (CDC, 2015; Langousis & Hill, 2014).

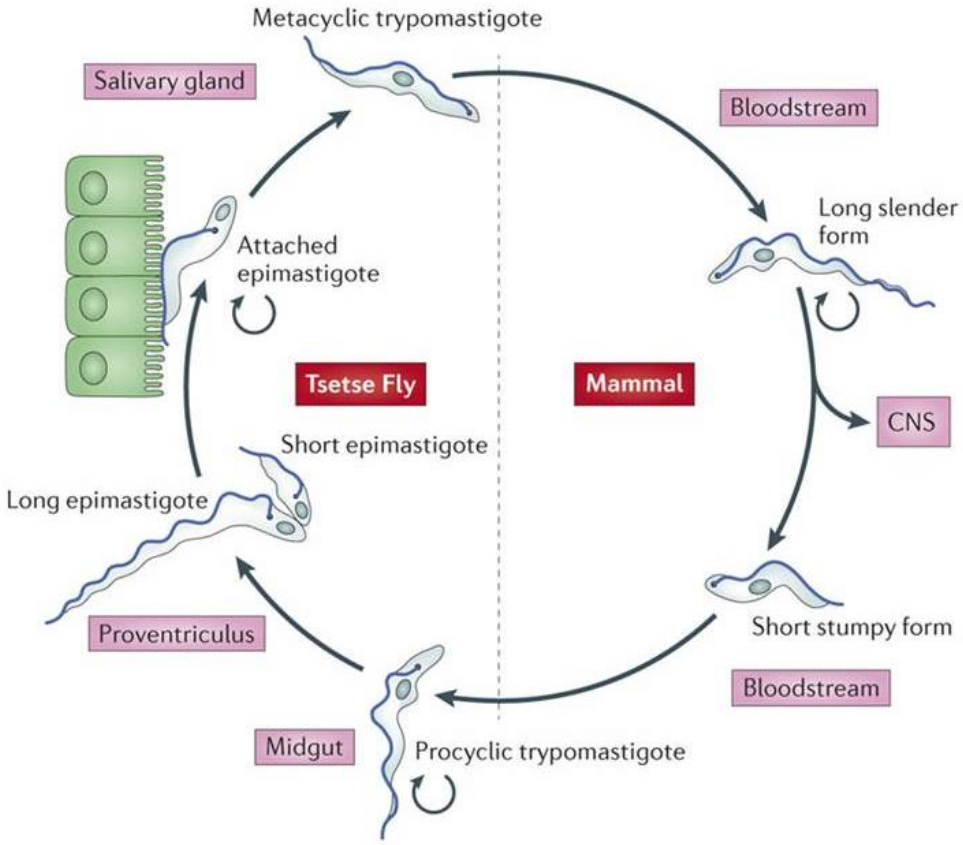


Figure 1-2 Life cycle of the *T. brucei* parasite

The mammalian host becomes infected by metacyclic trypomastigotes following a bite from an infected tsetse fly (CDC, 2015; Langousis & Hill, 2014). Once in the mammalian blood stream, the parasites differentiate into long slender bloodstream trypomastigotes, which divide in the blood (CDC, 2015; Langousis & Hill, 2014). The long slender form differentiates into the short stumpy form, which is adapted for survival in the insect vector (Langousis & Hill, 2014). These infect the tsetse fly during feeding. Once in the fly midgut, the parasites transform into procyclic trypomastigotes (CDC, 2015; Langousis & Hill, 2014). These transform into epimastigotes and migrate to the salivary glands via the proventriculus. It is in the proventriculus where the parasite divides asymmetrically, producing two epimastigotes: one that is long and one that is short (Langousis & Hill, 2014). The short epimastigote attaches to the salivary gland epithelial cells where it replicates to produce metacyclic trypomastigotes (Langousis & Hill, 2014). These are now ready to infect another mammalian host (CDC, 2015; Langousis & Hill, 2014). Image adapted by permission from Macmillan Publishers Ltd: Nature Reviews Microbiology, copyright 2014 (Langousis & Hill, 2014).

The two sub-species of *Trypanosoma brucei* that cause human disease display different disease progressions, but are both characterised by two distinct stages. The first, haemolymphatic stage occurs shortly after infection, and symptoms tend to be non-specific. The second, meningoencephalitic stage occurs when the parasite crosses the blood brain barrier (Brun *et al.*, 2010). Infection by either parasite will ultimately lead to death if left untreated. The *T. b. gambiense* infection is chronic, with an average survival rate of three years. It accounts for 98% of HAT cases (Brun *et al.*, 2010). Neurological symptoms occur after 1-2 years (CDC, 2015). Kieft *et al* showed that the ability of *T. b. gambiense* parasites to infect humans is due to expression of reduced levels of a haptoglobin-related protein receptor. This decreases the uptake of the high density lipoprotein sub-fraction called trypanosome lytic factor (Kieft *et al.*, 2010). The *T. b. rhodesiense* infection is acute, and death usually occurs within months (Brun *et al.*, 2010) (CDC, 2015). The difference in infectivity is because *T. b. rhodesiense* is ill-adapted to human infection. The ability of *T. b. rhodesiense* to infect humans is thought to be due to a deletion mutation within a variable surface glycoprotein (VSG) gene. This produced a serum resistance associated (SRA) gene (Campillo & Carrington, 2003; Gibson, 2005). This SRA gene enables the parasite to avoid killing by trypanosome lytic factor (Oli *et al.*, 2006).

A concerted effort by WHO to eradicate HAT by 2020 has led to a decrease in the number of new reported cases from 37, 991 in 1998 to 3, 796 in 2014 (WHO, 2015a). Whilst this is encouraging, decreases of a similar level have been seen previously, and have been followed by resurgence in the number of cases (Figure 1-1).

1.1.3 Animal African Trypanosomiasis

Whilst HAT cases are on the decline, African Trypanosomiasis is still a major problem in domestic and farm animals. Animal African trypanosomiasis (AAT) affects many domesticated species including dogs, cattle and goats. AAT in cattle is considered to have the most economic importance. Bovine trypanosomiasis, or nagana, threatens an estimated 50 million head of cattle in Africa (FAO, 2015). The economic impact is huge, with an estimated loss of US$1.2 billion per year (FAO, 2015). AAT has a detrimental effect on the general well-being of infected animals. This in turn affects fertility, growth, and yields of milk and meat (Connor, 1992). Diagnosis of the disease requires the identification of parasites in bodily fluids, but this can be difficult, as the parasite burden is often low (Connor, 1992). This is because many animals reach an equilibrium with the parasites (Connor, 1992).

There are ten known animal pathogens within the *Trypanosoma* genus, and these are depicted in Figure 1-3. Of these ten, two are also human pathogens (*T. b. gambiense* and *T. b. rhodesiense*). *T. brucei, T. congolense, T. vivax* and *T. simae* are considered the most important causative agents of nagana, whilst *T. evansi* causes surra (trypanosomiasis of camels) (Anene *et al.*, 2001). *T. vivax* is considered the most important parasite of livestock in West Africa (Silva *et al.*, 2013). The parasite is specific to ungluates (hooved animals), and is found in both wild and domestic animals (Osório *et al.*, 2008). *T. vivax* is unique in that it has moved out of the tsete belt and into South America. This is thought to have occurred in the mid-19th Century through cattle importation (Pimentel *et al.*, 2012). Transmission in the Americas is thought to occur by mechanical means because the insect vector is not found in these locations (Osório *et al.*, 2008). Cases of *T. vivax* have been reported in French Guiana (Desquesnes & Gardiner, 1993), Columbia (Otte *et al.*, 1994), Boliva (Silva *et al.*, 1998) and Brazil (Pimentel *et al.*, 2012). Outbreaks of *T. vivax* disease are characterised by anaemia, emaciation, miscarriage and death (Jones & Dávila, 2001; Osório *et al.*, 2008). The financial impact of *T. vivax* in the Brazilian Pantanal and Bolivian lowlands was estimated to exceed US$160 million for the 11 million head of cattle thought to be farmed in these regions (Seidl *et al.*, 1999).

1.1.4 Current treatments for HAT

The current treatments for HAT come with a number of issues, which are discussed in detail in the following sections. The drugs are either stage or strain specific, with complex treatment regimens and severe side effects. The chemical structures of the current treatments are shown in Figure 1-4.

Suramin is the oldest drug still in use to treat HAT. It was discovered in 1916, by the Bayer Company in Germany (Steverding, 2010). The discovery of Suramin (also called Germanin) was instigated through the idea of using synthetic dyestuff as therapeutics, a concept pioneered by Paul Ehrlich (Greenwood *et al.*, 2007). Suramin is still the first line treatment for early stage *T. b. rhodesiense* infections (Kennedy, 2013). The drug is delivered by intravenous injection, with five injections required every seven days for one month (Brun *et al.*, 2010). The exact mode of action of Suramin is still not known, but it is taken up by the parasites by receptor mediated endocytosis (Vansterkenburg *et al.*, 1993; WHO, 2013).

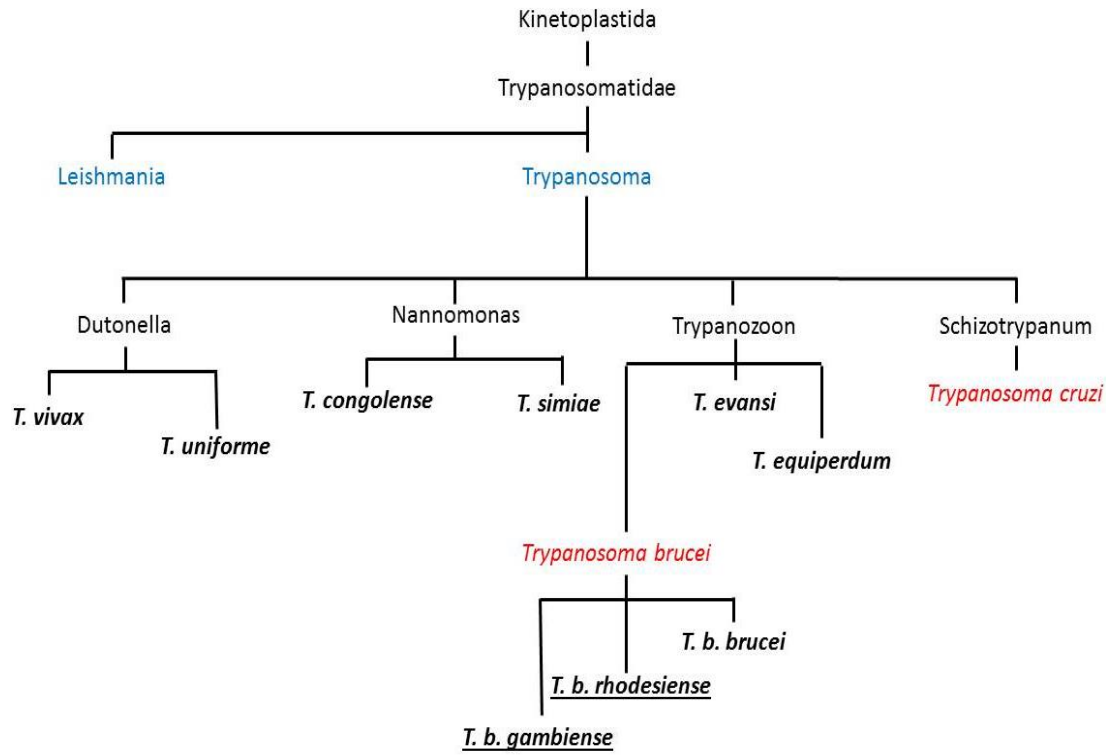


Figure 1-3 Taxonomy of the human pathogenic genera (blue) within the kinetoplastida order, and the animal pathogens within the Trypanosome genera

The taxonomy of the human pathogenic genera is shown alongside the animal pathogenic trypanosomes. The human pathogenic genera are shown in blue, with the human pathogenic *Trypanosoma* species highlighted in red. Animal pathogenic *Trypanosomes* are in **bold**, whilst both human and animal pathogens are in **bold** and underlined (Federhen, 2012; Uilenberg, 1998).

Pentamidine was developed in 1937, and is used to treat first stage *T. b. gambiense* infections (Steverding, 2010). The mode of action is still unknown, but Pentamidine is known to bind to AT-rich DNA (Fox *et al.*, 1990). The drug is actively transported into *T. brucei* blood-stream forms using a number of transporters (Damper & Patton, 1976; WHO, 2013). Pentamidine is delivered intramuscularly for up to ten days. However, clinical trials have begun to determine the effectiveness of a three day regimen (WHO, 2013). Pentamidine treatment is a highly effective therapy, with a cure rate of between 93 and 98% (WHO, 2013).

The organoarsenic Melarsoprol was developed in 1949, and is used to treat the second stage disease caused by both *T. b. rhodesiense* and *T. b. gambiense* (Brun *et al.*, 2010; Steverding, 2010). Melarsoprol is delivered by intravenous injection at 24 hour intervals for ten days (Brun *et al.*, 2010). The active compound is melarsen oxide, which is produced from metabolised Melarsoprol (Barrett *et al.*, 2011). The main issue with Melarsoprol is the

reactive encephalopathy that occurs in 5-10% of cases, with 1-5% of those resulting in death (Kuzoe, 1993). The mechanism of action is still unknown, but Melarsoprol is thought to inhibit dithiol-containing enzymes in both trypanosomes and mammals (WHO, 2013).

Eflornithine (also called DFMO) was originally produced as an anti-cancer therapeutic in the 1970s, but was re-purposed in 1990 for use against *T. b. gambiense* infections (Meyskens & Gerner, 1999; Steverding, 2010). Eflornithine targets ornithine decarboxylase, which is involved in polyamine synthesis (Oredsson *et al.*, 1980). Eflornithine is only active against *T. b. gambiense* because the turnover rate of the ornithine decarboxylase enzyme is approximately 4 times slower in this species then in *T. b. rhodesiense* (18/19 hours and 4 hours respectively) (Iten *et al.*, 1997). Eflornithine is difficult to administer as it requires slow intravenous delivery at six-hour intervals for fourteen days (Brun *et al.*, 2010).

Nifurtimox is used for the treatment of American trypanosomiasis, caused by the related trypanosome *T. cruzi* (WHO, 2013). It is not registered as a treatment for HAT, but is used in combination with Eflornithine (referred to as Nifurtimox Eflornithine combination therapy, or NECT) (WHO, 2013). This combination therapy involves the intravenous administration of Eflornithine every twelve hours for seven days, with Nifurtimox administered orally three times daily for ten days (Barrett & Croft, 2012). This combination not only improves the administration regimen, but it is hoped it will delay the onset of drug resistance (Barrett & Croft, 2012). Nifurtimox is a prodrug, meaning it requires additional processing for it to be trypanocidal. This processing is performed by trypanosomal type I nitroreductases, which reduce nifurtimox to its trypanocidal nitrile product (Hall *et al.*, 2011). The targets of this nitrile product have yet to be identified (Hall *et al.*, 2011).

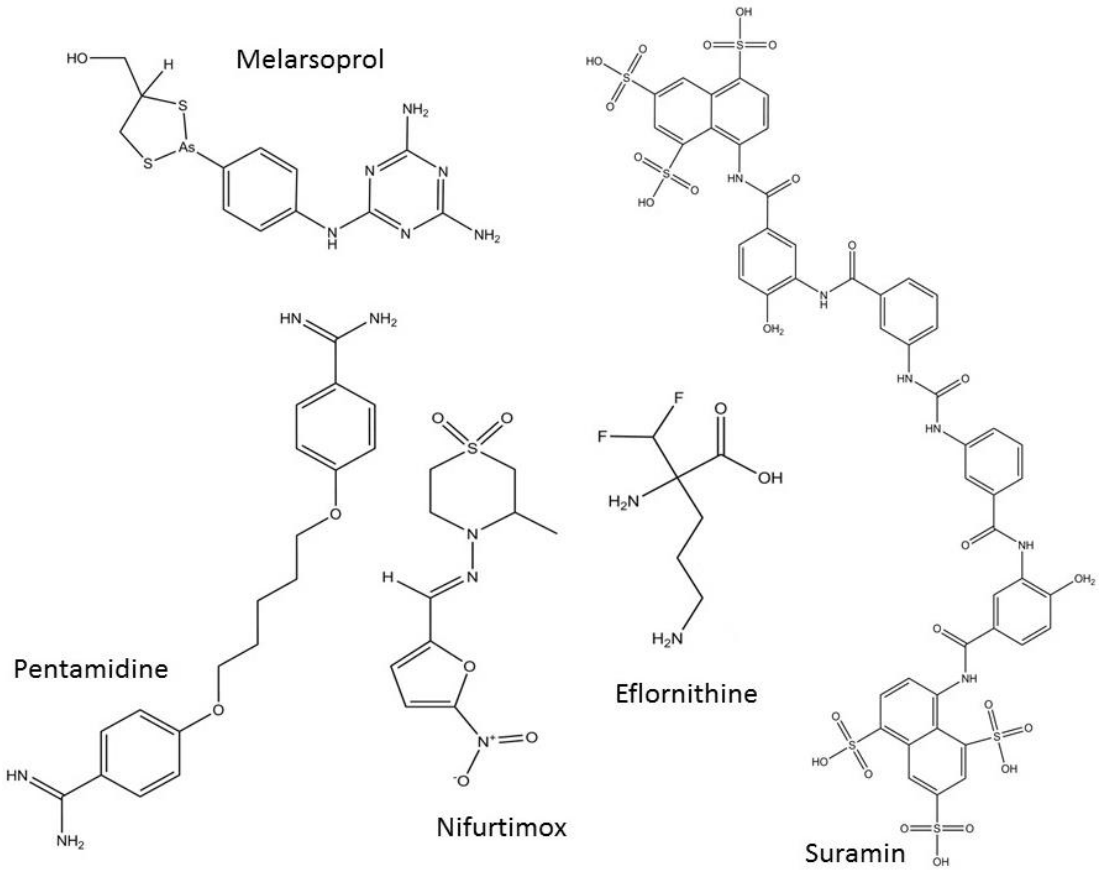


Figure 1-4 Current treatments for HAT

Chemical structures of the compounds currently used to treat HAT. Structures were rendered using Chemdraw 14, and based on the structures depicted in (Steverding, 2010).

The establishment of the Drugs for Neglected Diseases initiative (DNDi) has triggered a massive increase in the amount of research into drugs for diseases such as HAT. The ultimate aim of DNDi (with regards to HAT) is to deliver a treatment against both stages of the disease that is safe, effective when administered orally, and affordable (DNDi, 2016). The side effects, delivery route and length of treatment require drug administration to be performed in either a hospital setting, or under close supervision. (Keating *et al.*, 2015). This significantly increases the cost of treatment, and has been reviewed recently (Keating *et al.*, 2015). Three drugs identified by the DNDi initiative to help address these issues are described below.

Fexinidazole was originally identified in the 1970s, where its trypanocidal properties were first observed (Jennings & Urquhart, 1983; Raether & Seidenath, 1983; Torreele *et al.*, 2010). In the current studies, Fexinidazole is administered orally, and has been shown to be effective against both stages of *T .b. gambiense* and *T. b. rhodesiense* infections (Torreele *et al.*, 2010). Phase I clinical trials began in 2009, with phase II/III clinical trials commencing in

late 2012 (Torreele *et al.*, 2010; WHO, 2013). Whilst the mechanism of action has yet to be elucidated, it is thought that it may be a prodrug that needs to be reduced in order to be activated, like nifurtimox (Torreele *et al.*, 2010).

SCYX-7158 is a member of the benzoxaborole class of compounds (Jacobs *et al.*, 2011). It has been shown to be active against both *T. b. rhodesiense* and *T. b. gambiense in vitro*, and is able to cure both stages of the disease in a murine model (Jacobs *et al.*, 2011). The compound is orally active, and progressed to phase I clinical evaluation in 2012 (Jacobs *et al.*, 2011; WHO, 2013). Safety profiling was successfully completed in 2015, and patient recruitment for trials are due to begin this year (2016) (DNDi, 2015).

DB829 is an orally active aza-analogue of furamidine (Wenzler *et al.*, 2009). Furamidine was under development as a therapeutic against HAT, but it was shown to cause severe liver toxicity (Wenzler *et al.*, 2009). DB829 is more effective against *T. b. gambiense in vivo*, with an intraperitoneal injection curing all infected animals (Wenzler *et al.*, 2013). The same dose cured only 50% of *T. b. rhodesiense* infected mice (Wenzler *et al.*, 2013). DB829 is also curative in a murine model of second stage *T. b. gambiense* infection, curing all infected mice when administered for 10 days (Wenzler *et al.*, 2009). DB829 is under preclinical investigation (WHO, 2013).

1.1.5 Current treatments for AAT

The therapeutic arsenal against AAT is rapidly dwindling due to the emergence of drug resistance (Delespaux *et al.*, 2008). Despite this, the production of new trypanocides is not considered commercially viable, and the last drug was developed in 1961 (Connor, 1992). The current drugs against AAT are described below. As with the drugs against HAT, delivery method tends to be injection by intramuscular, intravenous or subcutaneous routes. Some of these therapeutics are used as prophylactics in animals, although the majority are used as curative agents. Certain HAT therapies are also used against AAT. One example is Suramin, which is used in camels to combat *T. evansi* infections (Connor, 1992). The chemical structures are shown in Figure 1-5.

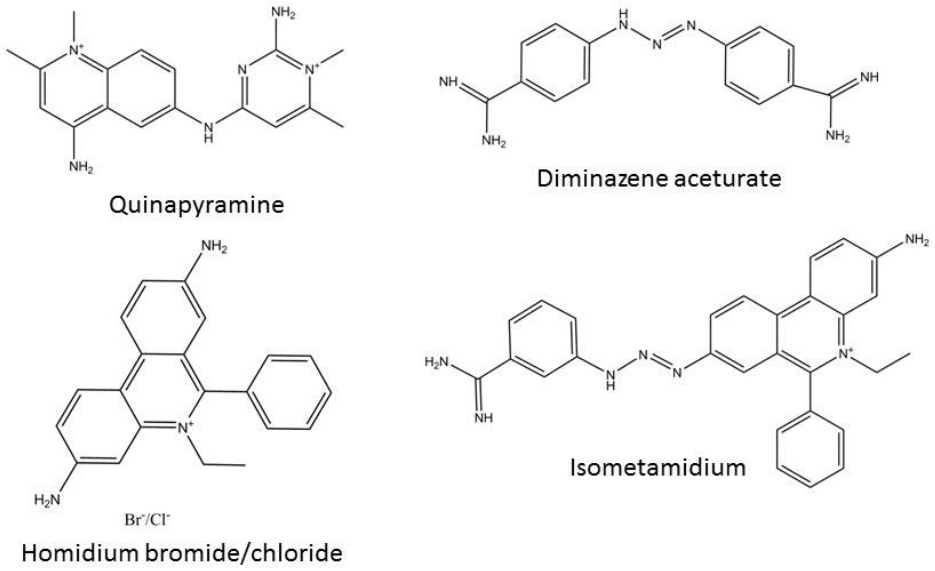


Figure 1-5 Current treatments for AAT

Chemical structures of the compounds currently used to treat AAT. Structures were rendered using Chemdraw 14, and based on the structures depicted in (Melaku & Birasa, 2013).

Diminazene aceturate is an aromatic diamidine that is effective against *T. congolense, T. vivax* and *T. brucei* (Leach & Roberts, 1981). It was introduced in the 1950s as a curative therapeutic that can be delivered either intramuscularly or subcutaneously (Connor, 1992; Leach & Roberts, 1981). Diminazene aceturate is thought to bind to AT rich regions of the kinetoplastid DNA (kDNA), where it can cause the complete loss of this nucleic acid (González *et al.*, 1997; Riou & Benard, 1980). Diminazene aceturate is rapidly excreted and it was thought that this would help prevent the occurrence of drug resistant strains (Leach & Roberts, 1981). However, a naturally occurring resistant *T. vivax* strain was identified in 1967 (Leach & Roberts, 1981).

Homidium is a phenanthridinium compound, and is available in two forms: chloride and bromide (Connor, 1992). Homidium bromide (also called ethidium bromide) was developed in 1952 (Watkins & Woolfe, 1952). Both forms are active against *T. congolense* and *T. vivax* infections, and are used as curative therapies (Connor, 1992; Kinabo, 1993). The drug is delivered via intramuscular injection (Kinabo, 1993). The mechanism of action for both forms of homidium is to intercalate with DNA and RNA and block the actions of DNA polymerases (Kinabo, 1993; Youssif *et al.*, 2010). Because of the non-specific nature of the drug, it is not surprising that it is highly toxic. In goats, the drug is neurotoxic and can cause demyelination, as well as bloody diarrhoea, colic and alopecia (Youssif *et al.*, 2010). The

drug needs to be fully eliminated from the animal before the meat is safe for human consumption (Youssif *et al.*, 2010).

Isometamidium is a phenanthridinium compound like homidium, and is active against *T. congolense, T. vivax, T. brucei* and *T. evansi* (Kinabo, 1993). It was synthesised in 1958, and is delivered intramuscularly (Wragg *et al.*, 1958). It is used as both a therapeutic and a prophylactic treatment (Sahin *et al.*, 2014). Isometamidium is an irritant and can cause tissue damage at the injection site (Connor, 1992; Kinabo, 1993).

Quinapyramine is a bis-quaternary compound that was introduced in the 1950s as both a therapeutic and a prophylactic (Kinabo, 1993). It was effective against *T. congolense, T. vivax, T. brucei, T. evansi, T. equinium* and *T. equiperdum,* but was withdrawn in 1976 due to increased resistance against it (Anene *et al.*, 2001; Leach & Roberts, 1981). Quinapyramine has since been introduced in two forms: quinapyramine prosalt and quinapyramine sulphate. The sulphate form is used against *T. brucei, T.congolense* and *T. vivax* as a therapeutic, and is delivered by subcutaneous injection (Leach & Roberts, 1981). The prosalt form is used as both a therapeutic and prophylactic (Connor, 1992).

1.1.6 Drug Resistance

Drug resistance seen in *Trypanosoma* parasites will be either innate or acquired (Anene *et al.*, 2001). One example of innate resistance can be seen with eflornithine. This treatment is only effective against *T. b. gambiense*, not *T. b. rhodesiense*. This is because the enzyme targeted by eflornithine has a faster turnover rate in *T. b. rhodesiense* (Iten *et al.*, 1997). Acquired resistance requires exposure of the parasite to sub-lethal concentrations of the drug. Factors that increase the likelihood of resistance include over use, failure to complete course of treatment and the use of drugs prophylactically (Barrett *et al.*, 2011). Another factor that plays an important role in AAT is the availability of poor quality counterfeit drugs (Schad *et al.*, 2008).

The mechanism of action for many of the trypanocidal drugs (for both HAT and AAT) is not fully understood. It is therefore difficult to fully understand how the parasites have evolved resistance. However, a decrease in drug uptake or an increase in drug efflux appears to be behind many cases of resistance. In eflornithine resistance, the enzyme turnover rate is not the only resistance mechanism available to the *T. b. rhodesiense* parasites. A decrease in drug accumulation due to the loss of the AAT6 transporter has been reported (Bellofatto *et al.*, 1987; Phillips & Wang, 1987; Vincent *et al.*, 2010).

The P2 adenosine transporter AT1 (encoded by the gene *TbAT1*) has been linked to resistance to Melarsoprol and diamidines, including Pentamidine and Diaminazene (Baker *et al.*, 2013; Carter *et al.*, 1995; Carter & Fairlamb, 1993; Teka *et al.*, 2011). Loss of this transporter is not the only way to decrease drug accumulation within the parasite. Analysis of melarsen oxide (the active metabolite of Melarsoprol) sensitivity in AT1 knock-out strains indicated that, whilst sensitivity was reduced, the parasites were still susceptible to the drug (Matovu *et al.*, 2003). A second transporter, HAPT1, has also been linked to resistance against Melarsoprol and diamidines (Bridges *et al.*, 2007; de Koning, 2008; Paine *et al.*, 2010; Teka *et al.*, 2011). The current theory for Melarsoprol and Pentamidine resistance is that the loss of both AT1 and HAPT1 is required for high resistance levels within the parasites (Barrett *et al.*, 2011). The loss of either transporter may be due to point mutations within the gene, loss of the gene, or the gene not being transcribed (Barrett *et al.*, 2011; Carter *et al.*, 1995; Lanteri *et al.*, 2006; Stewart *et al.*, 2010). Another transporter thought to contribute to Melarsoprol and Pentamidine accumulation within the parasite is aquaglyceroporin 2 (AGP2) (Baker *et al.*, 2012). AGP2 a is a porin that is involved in osmoregulation and metabolism, but causes resistance when knocked out in blood stream form cells (Baker *et al.*, 2012). A second, unrelated mechanism for Melarsoprol resistance has been identified. This involves the overexpression of an efflux pump, specifically MRPA (Lüscher *et al.*, 2006). This causes the log IC_{50} of melarsen to double (Lüscher *et al.*, 2006).

Resistance to nifurtimox also confers resistance to fexinidazole, a new therapy currently in phase III clinical trials (Sokolova *et al.*, 2010). Exactly how this occurs is not fully understood, but may be linked to both compounds requiring activation *in vivo* to produce the trypanocidal species (Sokolova *et al.*, 2010) The idea that resistance to a new therapeutic is being primed before it has even entered the market is concerning.

1.1.7 African Trypanosomiasis – a perspective

The number of cases of HAT have reduced significantly, but the isolation of drug-resistant field strains to all current therapeutics (including those used to treat AAT) is very concerning (Melaku & Birasa, 2013). This highlights the need for continued research in this area. Work on therapies to combat AAT is also sorely needed, not just from an economic point of view, but also with regards to human health. Infected animals (both wild and domestic) act as a reservoir for human infections, so controlling these through therapeutics (as well as additional strategies such as vector control) is vital. As a consequence, potential

new drug targets need to be identified, and work on new therapeutics for HAT and AAT must be a priority.

1.2 Flap Endonucleases

1.2.1 The Cellular Roles of Flap Endonucleases

Flap endonucleases are essential enzymes involved in the replication and maintenance of the genome. Commonly referred to as FENs, 5′ nucleases and 5′-3′ exonucleases, these structure-specific enzymes are characterised by their unique ability to recognise 5′ flaps. FENs are classed as metalloenzymes, as they require the presence of divalent metal cations to process their substrates. The essential roles of these enzymes in DNA replication and repair are discussed below.

1.2.1.1 DNA Replication

Precise replication of DNA is essential for cell viability. This semi-conservative process is highly regulated, ensuring a low error rate of an estimated 10^{-9} mutations per base pair per year in higher animals (Kumar & Subramanian, 2002). DNA replication itself is highly complex, and is summarised in Figure 1-6. Within the eukaryotic cell, replication begins by the assembly of the origin-replication complex at the replication origin (Henneke *et al.*, 2003). This complex contains helicases, which unwind the DNA duplex. This causes the recruitment of Replication Protein A (RPA), which stabilises the single stranded DNA. DNA Polymerase α then synthesises a complementary RNA/DNA primer on the parental DNA strands. This is followed by a 'polymerase switch' from DNA Pol α to DNA Pol δ on the lagging strand (Henneke *et al.*, 2003), and to DNA Pol ε on the leading strand (Zheng & Shen, 2011). This process requires the presence of two proteins: replication factor C (RFC) and PCNA (proliferating cell nuclear antigen). These two proteins, plus DNA Pol δ/ε, form the holoenzyme that extends the RNA/DNA primer. This forms the replication fork, which moves along the DNA strand as replication progresses.

The DNA helix is anti-parallel, so replication can only occur in one direction: 5′ to 3′ (Lodish *et al.*, 2004). On the leading strand this is not an issue, because DNA replication can occur continuously from a single RNA/DNA primer (Figure 1-6). The problem arises on the lagging strand, where replication has to occur in the opposite direction (3′ to 5′) to the moving replication fork (Figure 1-6). The solution is a series of short RNA/DNA primers (synthesised by primase; a heterotetramer of RNA Pol and DNA Pol α) that are produced every few

hundred base pairs (Zheng & Shen, 2011). These are elongated by DNA Pol δ to form Okazaki fragments (short, discontinuous sections of DNA) (Zheng & Shen, 2011). The initial RNA/DNA primer is displaced by DNA Pol δ, and FEN1 is recruited by PCNA to cleave the displaced RNA/DNA primer (Henneke *et al.*, 2003). The final step is the recruitment of DNA Ligase 1 by PCNA, to seal the nick between the DNA strands and maintain genomic integrity (Henneke *et al.*, 2003).

Three possible pathways for the removal of the RNA/DNA primer have been postulated: the RNase H-FEN1, FEN1-Dna2 and FEN1 only pathways. The RNase H-FEN1 pathway is not thought to be the main route in eukaryotes, because deletion of the genes encoding RNase H1 and H2 in yeast does not cause an obvious phenotype (Zheng & Shen, 2011). The FEN1-Dna2 pathway (Figure 1-7) is thought to be important in yeast. Studies in *Saccharomyces cerevisiae* where the Dna2 protein was mutated to prevent its endonuclease activity did not allow the cells to grow (Lee *et al.*, 2000). However, it is thought to be less important in mammals, where DNA2 is primarily associated with mitochondrion, and any that remains in the nucleus is not associated with replication foci (Duxin *et al.*, 2009; Zheng *et al.*, 2008). It is therefore the FEN1-only pathway (Figure 1-7) that is thought to be the dominant route in mammals. This pathway starts with the strand displacement of the RNA/DNA primer by DNA Pol δ. Garg *et al* showed that DNA Pol δ is only capable of displacing 1-2 nucleotides, leaving an ideal substrate for FEN1 cleavage (Garg *et al.*, 2004).

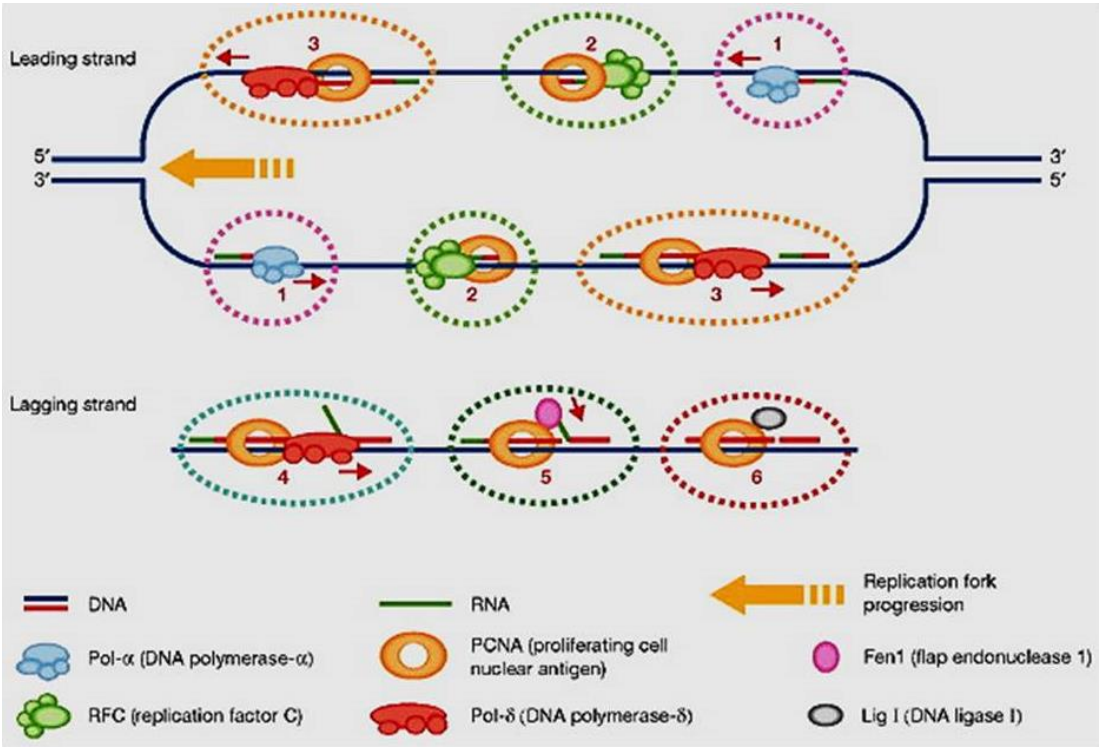


Figure 1-6 An overview of semi-conservative DNA replication in eukaryotes

Leading strand synthesis: **1.** RNA/DNA primer synthesis by DNA Pol α. **2.** DNA Pol α is displaced, and PCNA is recruited by RFC. **3.** The DNA strand is elongated to produce nascent DNA, by DNA Pol ε.

*Lagging strand synthesis***: 1.** RNA/DNA primer synthesis by DNA Pol α**. 2.** DNA Pol α is displaced, and PCNA is recruited by RFC. **3.** The Okazaki fragment is elongated by DNA Pol δ. **4.** The downstream Okazaki fragment is displaced (via strand displacement) by DNA Pol δ. **5.** FEN1 is recruited by PCNA. 6. Cleavage of the 5′ DNA/RNA flap by FEN1. The nick is then ligated by DNA Lig 1 to maintain genomic integrity.

Reproduced by permission from John Wiley and Sons: EMBO Reports (Henneke *et al.*, 2003).

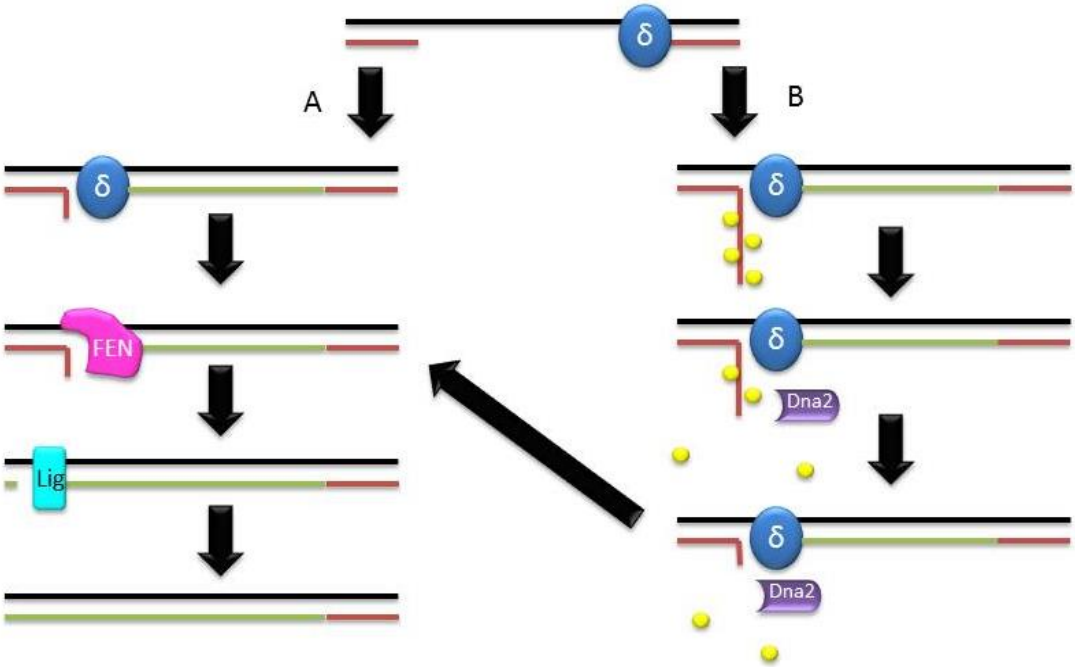


Figure 1-7 Removal of the RNA primer using either the FEN1 only, or FEN1-Dna2 pathways

DNA Pol δ (blue) displaces the 5′ end of the downstream initiator primer to produce a 5′ flap. Production of a small flap (***A;*** 2-12 nucleotides) is cleaved by FEN1 (pink) (Garg *et al.*, 2004). Production of a large flap (***B;*** ≥30 nucleotides) results in the recruitment of RPA (yellow) (Bae & Seo, 2000). Dna2 (purple) is recruited, which displaces RPA and periodically cleaves the single stranded DNA (Bae & Seo, 2000). This produces a 5-6 nucleotide flap that cannot be bound by RPA, but can be cleaved by FEN1 (Bae & Seo, 2000; Zheng & Shen, 2011). DNA Lig 1 (cyan) ligates the two Okazaki fragments together to form a complete nascent DNA strand.

1.2.1.2 Rescuing Stalled Replication Forks

During DNA replication, the replication fork moves along the duplex DNA, unwinding it as replication progresses. Replication forks are required for complete genomic replication, so anything that prematurely halts its progress is detrimental to the cell. In eukaryotic cells, a number of checkpoint kinases are activated when a replication fork stalls (Labib & Hodgson, 2007). Without these, the replication fork will collapse, producing double-stranded DNA breaks which are fatal for the cell (Hodgson *et al.*, 2007; Lopes *et al.*, 2001; Tercero & Diffley, 2001). Replication forks stall for a number of reasons. Some are part of the cell's normal replicative process, and are referred to as replication fork barriers. These include GC-rich regions of DNA, other protein-DNA complexes, polar barriers that form at transfer RNA genes, and bidirectional barriers at centromeres (Hodgson *et al.*, 2007). Other, less favourable, reasons for replication fork stalling include dNTP depletion, and DNA lesions

that form following assault by exogenous factors, for example ultraviolet radiation (Mirkin & Mirkin, 2007).

FEN1 interacts with WRN (Werner Syndrome Protein), a helicase from the RecQ subfamily. These two proteins co-localise to the site of the stalled replication fork, where WRN is thought to produce a single-stranded 5' substrate. RPA associates with the single-stranded DNA, where it is phosphorylated. FEN1 cleaves the single-stranded DNA via its gap endonuclease activity (see section 1.2.3.1), which is stimulated two-fold by the presence of phosphorylated RPA. This allows the initiation of break induced replication, which rescues the replication fork (Sharma *et al.*, 2004a; Zheng *et al.*, 2005).

1.2.1.3 DNA Repair

DNA lesions block DNA replication and transcription. If these are left untreated, or they are repaired incorrectly, they will lead to mutations that could impact cell viability. DNA lesions arise from both exogenous and endogenous agents. Endogenous agents include by-products from oxidative respiration and redox cycling events involving environmental toxins (Jackson & Bartek, 2009). The most common exogenous agent is ultraviolet light, but other examples include ionizing radiation and tobacco products (Jackson & Bartek, 2009). Consequently, the cell has a large number of DNA repair pathways, including mismatch repair, base excision repair and nucleotide excision repair. The repair of DNA is relatively straightforward, because DNA is a duplex comprised of two complementary strands. Any damage in one strand can be repaired by using the second strand as a template. The complexity in this process arises from the number of proteins involved.

Defects in the cellular DNA repair pathways can predispose the organism to a number of diseases. The most commonly associated disease is cancer. Mutations in XPG (a FEN1 family protein involved in nucleotide excision repair) can cause xeroderma pigmentosum. This predisposes the individual to skin cancer (Jackson & Bartek, 2009). Another example involves the BRACA 1+2 genes. These are involved in homologous recombination, and mutations in these predisposes the individual to breast cancer (Jackson & Bartek, 2009). Overexpression of FEN1 has been seen in a number of cancers. The expression of FEN1 in colon, stomach and lung tumours increased 1-2 fold, whilst a 2.5 fold increase in FEN1 expression was seen in breast and uterine samples (Singh *et al.*, 2008). An increase in FEN1 expression can be correlated with increasing severity of breast cancer tumours (Singh *et al.*, 2008). Mutations within the FEN1 protein itself have been observed in a number of human cancer types including non-small cell lung carcinoma, melanoma and breast

adenocarcinoma (Zheng *et al.*, 2007). Another group of diseases associated with defective DNA repair are the neurodegenerative disorders. For example Cockayne syndrome is caused by mutations in genes involved in nucleotide excision repair (including XPG); similarly with cerebro-oculo-facio-skeletal (COFS) syndrome (Jackson & Bartek, 2009).

FEN1 is directly involved in base excision repair. Two FEN1 family members, EXO1 and XPG, are involved in mismatch repair and nucleotide excision repair respectively. The base excision repair (or BER) pathway corrects single base DNA damage caused by alkylation, deamination and oxidative damage (Robertson *et al.*, 2009). There are two BER pathways, short-patch BER and long-patch BER. Short-patch BER is initiated by DNA glycosylase. This enzyme recognises and removes a damaged base by cleaving the N-glycosidic bond (Robertson *et al.*, 2009). This leaves a gap called the apurinic/apyrmidinic (AP) site. DNA AP endonuclease (APE1) then cleaves the DNA backbone to produce a single stranded DNA nick. APE1 processes this to create a single nucleotide gap in the DNA (Robertson *et al.*, 2009). This is filled by DNA Pol β, and the process is completed by DNA Lig III (Robertson *et al.*, 2009). The long-patch BER pathway was initially described by Frosina *et al* (1996), and is summarised in Figure 1-8 (Frosina *et al.*, 1996). This pathway begins in much the same way as the short-patch pathway up until the polymerase recruitment stage. At this step either DNA Pol β or DNA Pol δ is recruited, along with PCNA, FEN1 and DNA Ligase 1. DNA Pol β/δ then displaces the DNA strand adjacent to the AP site, creating a flap substrate. FEN1 cleaves this flap, leaving an end that can be ligated by DNA Lig I (Pascucci *et al.*, 2011).

Nucleotide excision repair (NER) deals with a wide range of structurally unrelated lesions. These include cyclobutane pyrimidine dimers and 6,4 photoproducts, which are caused by exposure to UV light (Costa *et al.*, 2003). These lesions distort the DNA helix, interfere with DNA base pairing and block DNA transcription (Costa *et al.*, 2003). The human NER pathway is complex because it requires the coordinated action of a large number of proteins. The FEN1 family member involved in NER is XPG. The role of XPG in this process is to cut at the 3' end of the DNA damage to facilitate its removal (de Boer & Hoeijmakers, 2000).

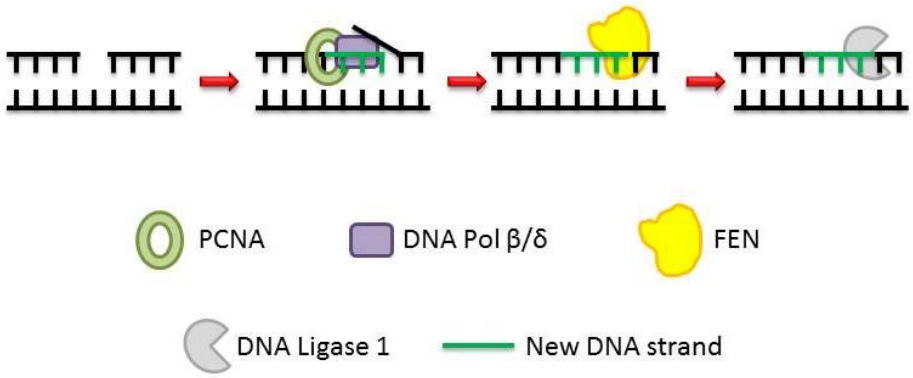


Figure 1-8 The role of FEN1 in long patch base excision repair

The single stranded nick is processed by APE1to create a single nucleotide gap in the DNA (Robertson *et al.*, 2009). DNA Pol β/δ fills the nucleotide gap (red) whilst displacing the adjacent DNA strand. This forms a flap that is cleaved by FEN1 (Pascucci *et al.*, 2011). DNA Lig I then ligates the nascent DNA to the original backbone to complete the repair process (Pascucci *et al.*, 2011).

Mismatch repair (MMR) deals with two types of DNA errors. The first is base-base mismatches, caused by DNA polymerase errors (Jiricny, 2006). The second is insertion/deletion loops that are produced as a result of microsatellite instability (Jiricny, 2006). Microsatellite instability occurs when the number of microsatellite repeats (short, repeated DNA sequences) changes from the original DNA template (NCI, 2015). Without the mismatch repair pathway, these errors are left uncorrected, and ultimately leads to cancer (Jiricny, 2006). The FEN1 family member involved in MMR is EXO1. The role of EXO1 in this process is to degrade the DNA in a 5' to 3' direction to remove the mismatched DNA (Jiricny, 2006).

1.2.1.4 Telomere Maintenance

At the end of each chromosome is a protective cap, known as a telomere. This cap protects chromosomes from recombination and end-to-end fusion (Cong *et al.*, 2002). Telomeres also prevent the ends of chromosomes from being recognised as DNA damage, and they allow complete replication of the chromosomes, specifically in lagging strand synthesis (Cong *et al.*, 2002). Telomeres are comprised of tandem repeats of a hexanucleotide sequence (Blackburn *et al.*, 2006). In mammals, that sequence is comprised of TTAGGG, with a single 3' overhang (Cong *et al.*, 2002). This 3' overhang is able to insert into the duplex telomeric DNA, where it forms a T-loop structure (Cong *et al.*, 2002; Stewart *et al.*, 2012). In addition to the tandem repeats, telomeres are characterised by the presence of specialised multi-protein complexes (Stewart *et al.*, 2012). In mammals, this is known as shelterin (Stewart *et al.*, 2012).

Because of the repetitive nature of the telomeres, their replication is a difficult, but essential, process. Without the telomere, the chromosomes shorten with every round of DNA replication. This leads to either cellular senescence or apoptosis (Gilson & Géli, 2007). Telomeric replication is often hampered by the formation of G quadruplexes (G4), triple helices, four-way junctions, D-loops and T-loops (Gilson & Géli, 2007). The process is reviewed in Gilson and Géli (2007), but the mammalian cell model is described briefly here. The RecQ helicases WRN and BLM are thought to unwind unusual DNA structures formed by the repetitive nature of telomeric DNA (Gilson & Géli, 2007). This activity is stimulated by two proteins: TTAGGG-repeat factor 2 (TRF2) and protection of telomeres protein 1 (POT1). TRF2 also stimulates FEN1 activity on Okazaki fragment maturation (Gilson & Géli, 2007). Mammalian homologues of the RRM3/PIF1 related helicase allows the replication fork to continue on through the nucleoprotein complex (Gilson & Géli, 2007).

More recent work on the role of FEN1 and related enzymes at telomeres has shown that FEN1 can form a complex with telomerase. Telomerase is a reverse transcriptase that counteracts telomere shortening by extending the 3' end of chromosomes (Gilson & Géli, 2007; Sampathi *et al.*, 2009). Whilst this interaction may be indirect, knocking down FEN1 in telomerase-positive cancer cell lines does cause gradual telomere shortening and retarded cell growth (Sampathi *et al.*, 2009). This indicates a role for FEN1 in telomerase regulation (Sampathi 2009). Human FEN1 has also been shown to cleave 5' flaps that contain G4 DNA, indicating its importance in lagging-strand replication at the telomeres (Vallur & Maizels, 2010). FEN1 and its family member EXO1 exhibit exonuclease activity on both DNA strands

of a substrate with a G-rich, 3′ telomeric tail (Vallur & Maizels, 2010). In addition, EXO1, but not FEN1, was able to excise G loops that are formed at transcribed telomeric regions (Vallur & Maizels, 2010). Whilst the exact roles of FEN1 and its family members in telomere maintenance have yet to be fully elucidated, it is clear that both FEN1 and EXO1 play separate but essential roles.

1.2.2 The Essential Nature of Flap Endonucleases

Flap endonucleases are members of the XPG/Rad2 endonuclease family, which is characterised by the presence of two conserved domains; the N-terminal domain and intermediate domain (Figure 1-9) (Harrington & Lieber, 1994a; Moritoh *et al.*, 2005). The prokaryotic versions are composed of these two domains only, whilst the eukaryotic versions also contain a C-terminal domain (Figure 1-9). Flap endonucleases were originally divided into four distinct groups based on those three domains: distinct FEN1 type enzymes (Group 1), bacteriophage 5′-3′ exonucleases (Group 2), prokaryotic DNA Polymerase exonuclease domains (Group 3; Figure 1-10) and nucleotide excision repair proteins (Group 4) (Figure 1-9) (Lieber, 1997). Arguably, a fifth group could now be added to this schematic based on the identification of ExoIX, an *Escherichia coli* flap endonuclease paralogue encoded by the *xni* gene. This gene is common in a number of eubacteria, including *Staphylococcus aureus* and *Yersinia spp* (Anstey-Gilbert *et al.*, 2013). ExoIX lacks the second metal ion binding site, but still retains 3′-5′ exonuclease activity, and 5′ flap endonuclease activity (Anstey-Gilbert *et al.*, 2013; Hodskinson *et al.*, 2007). In eukaryotes, archaea and bacteriophages, a separate and distinct flap endonuclease is encoded. This enzyme associates with DNA polymerases via a protein platform. Within eukaryotes, there are flap endonuclease paralogues, specifically XPG and EXO.

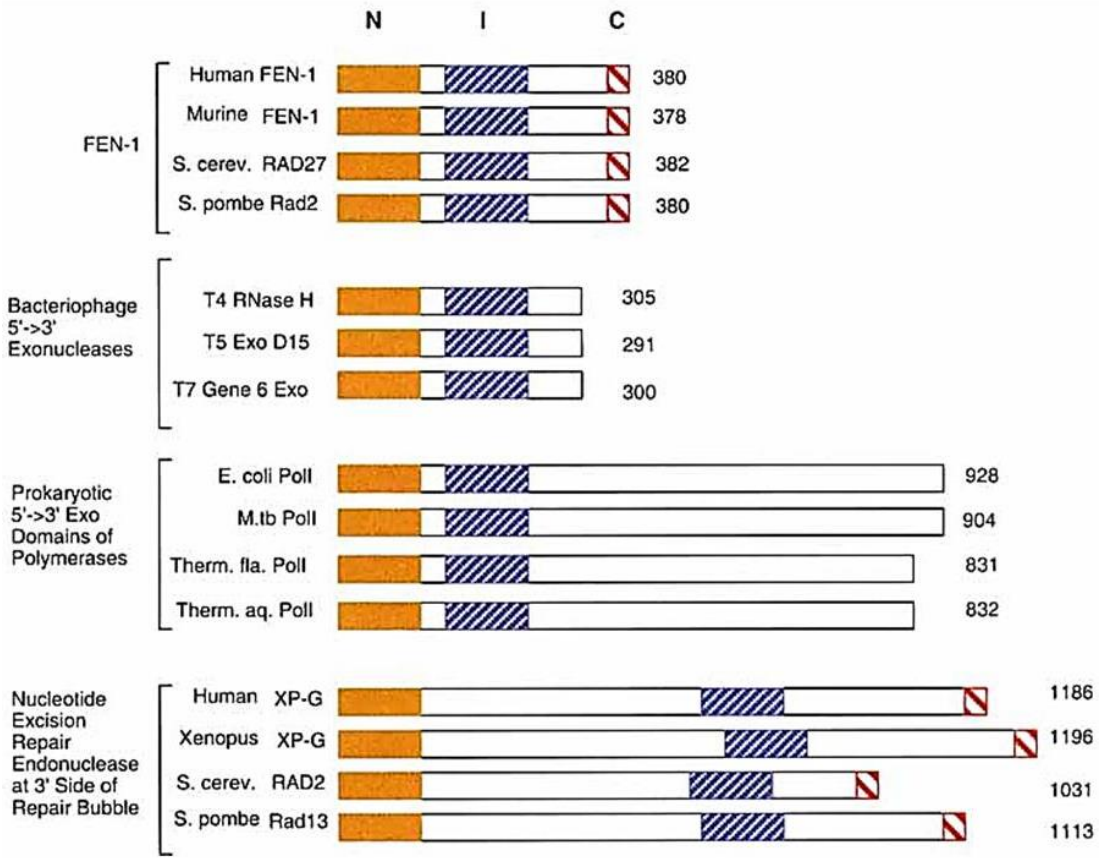


Figure 1-9 Domain homology within the flap endonuclease family

Schematic of FENs from a variety of different organisms, showing the homology domains: the N-terminal domain (yellow), intermediate domain (blue striped) and C-terminal domain (red striped). Only the N and I domains are found in prokaryotic FEN homologues, whilst the N, I and C domains are found in eukaryotic FEN homologues (Harrington & Lieber, 1994a). FEN-like proteins can therefore be split into four groups: distinct FEN1 type enzymes (Group 1), bacteriophage 5′-3′ exonucleases (Group 2), prokaryotic DNA Polymerase exonuclease domains (Group 3) and nucleotide excision repair proteins (Group 4) (Lieber, 1997). The N-terminal domain typically consists of the initial 100 amino acids. Approximately 18 amino acids separates the N-terminal and intermediate domains, except in Group 4 proteins, where this gap can consist of over 640 amino acids (Lieber, 1997). The C-terminal domain is typically composed of basic amino acids (R, K and H), and is located at the C-terminus of the protein (Lieber, 1997). ***M. tb*****:** *Mycobacterium tuberculosis*, ***Therm fla;*** *Thermus flavus,* ***Therm aq;*** *Thermus aquaticus*.

Reproduced by permission from John Wiley and Sons: BioEssays (Lieber, 1997).

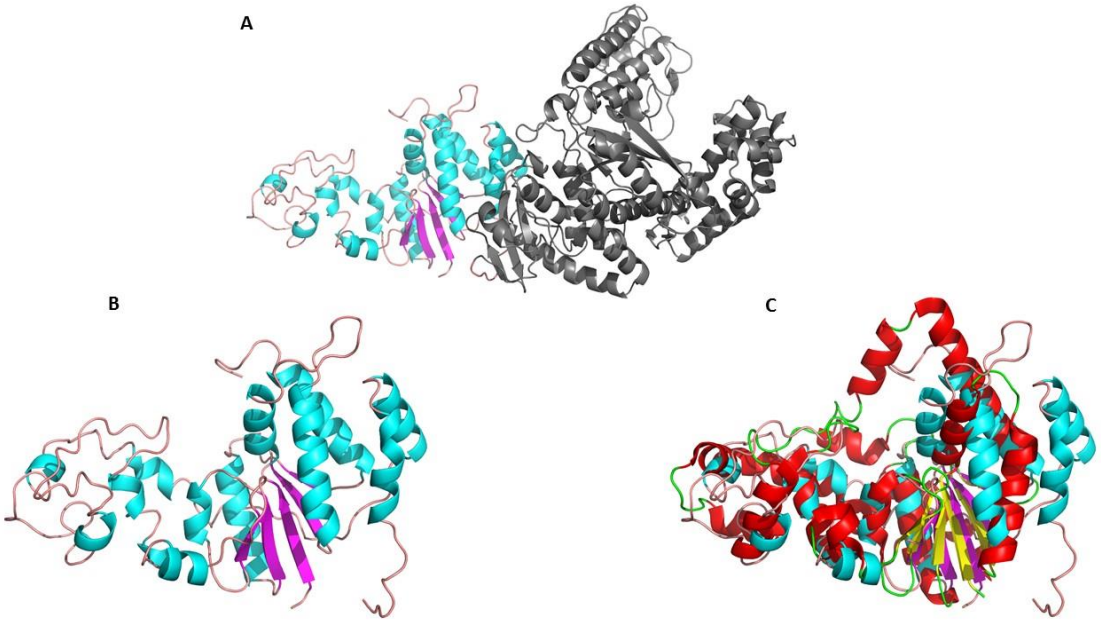


Figure 1-10 Comparison of the *Thermus aquaticus* PolI FEN domain, and the T5 Bacteriophage FEN

(**A**) The crystal structure of the full length *Thermus aquaticus* PolI protein (PDB 1TAQ), displaying its two main domains. The FEN domain is coloured according to secondary structure, where alpha helices are shown in cyan, and beta sheets in purple and connecting residues in pink. The Klenow fragment is coloured grey. (**B**) The *T. aquaticus* PolI FEN domain (residues 1-296) is located at the N-terminus of the full length protein, and is shown separately. (**C**) Alignment of the *T. aquaticus* PolI FEN domain with the T5 bacteriophage FEN (PDB 1UT5) using PyMol (R = 7.746). The T5 bacteriophage is coloured according to secondary structure, where alpha helices are shown in red, beta sheets in yellow and connecting residues in green. All images and alignments were rendered using The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

Flap endonucleases are found in all organisms, and a high level of conservation can be seen between them (see Figure 1-11). This high level of conservation provides an indicator to the essential nature of the enzyme. This has been studied further in both prokaryotes and eukaryotes. Initial studies in *E. coli* showed that knocking out the DNA PolI gene caused cell death when grown on rich media, but cell viability could be restored by adding back the FEN domain (Joyce & Grindley, 1984). Interestingly, knocking out the DNA PolI gene in cells grown on minimal media had no effect on cell viability, which indicates the presence of a 'back-up' protein (Joyce & Grindley, 1984). This has recently been identified as ExoIX, the product of the *xni* gene (Anstey-Gilbert *et al.*, 2013; Fukushima *et al.*, 2007). Double mutants of both the DNA PolI FEN domain and ExoIX are not viable in *Bacillus subtilis* (Fukushima *et al.*, 2007).

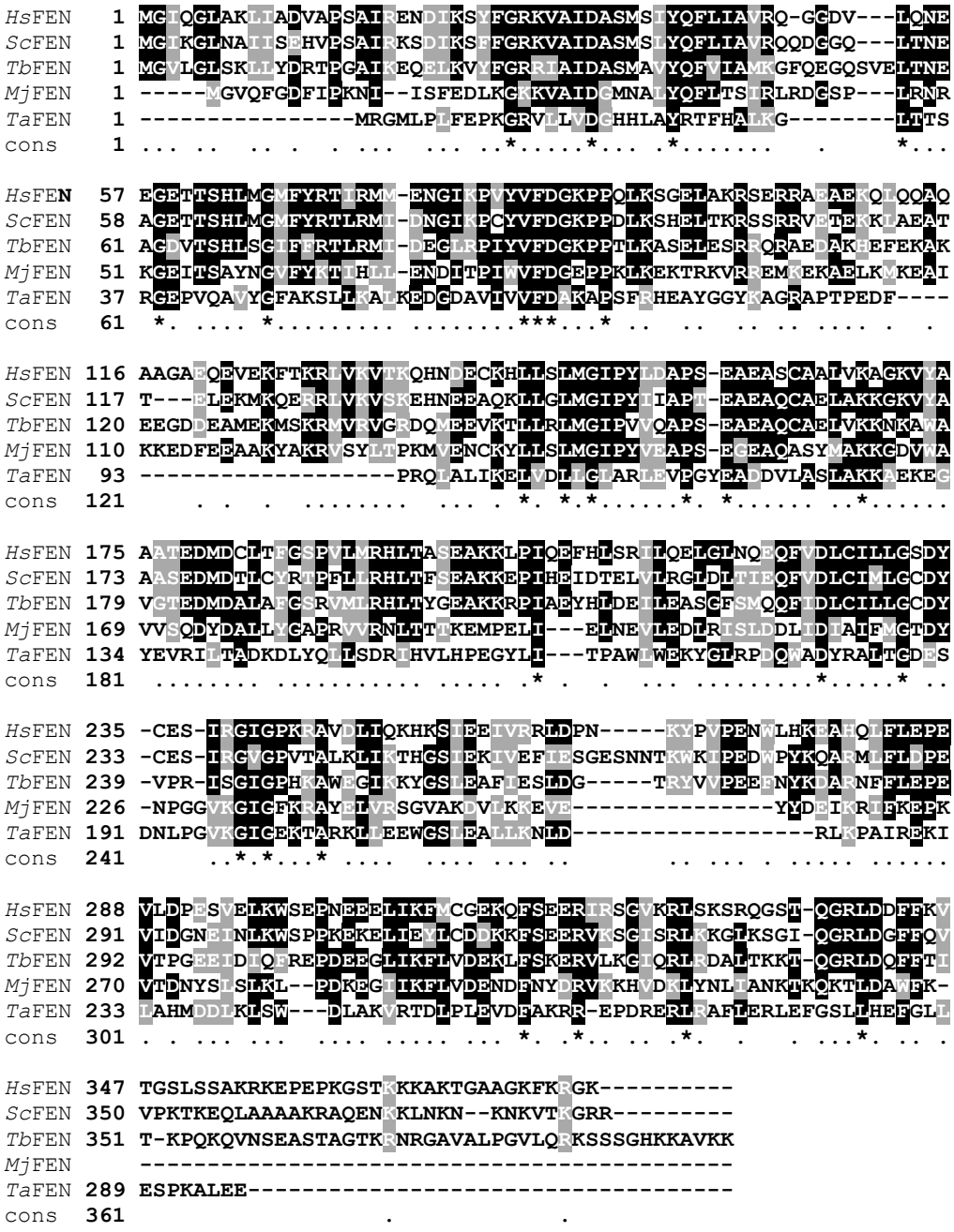


Figure 1-11 Multiple sequence alignment of FEN enzymes from archaea and eukaryotes

The sequences obtained from the NCBI database were aligned using the Multalin alignment tool, and are depicted using BoxShade (Corpet, 1988). Amino acid residues shaded black are identical, whilst residues shaded grey display similar properties. A high level of conservation can be seen.

*Hs*FEN = Homo sapiens (NP_004102.1), *Sc*FEN = Saccharomyces cerevisiae (P26793.1), *Tb*FEN = Trypanosoma brucei (XP_843679.1), *Mj*FEN = Methanocaldococcus jannaschii (Q58839.1), *Ta*FEN = Thermus aquaticus (AAA27507.1).

Initial studies in eukaryotes focused on the budding yeast *Saccharomyces cerevisiae*. The *S. cerevisiae* FEN homologue, Rad27, is a separate protein from the yeast DNA polymerase. *RAD27* null mutants are viable, but show a phenotype characterised by temperature-sensitivity and spontaneous mutations (Tishkoff *et al.*, 1997). Over-expression of ExoI, the flap endonuclease paralogue, in the *RAD27* null mutants can suppress this phenotype (Tishkoff *et al.*, 1997). *EXOI* null mutants do not display the same phenotype as the *RAD27* mutants in that they do not display any growth defects or methylmethanesulphonate sensitivity (Tishkoff *et al.*, 1997). However, the double *EXOI/RAD27* mutations are lethal (Tishkoff *et al.*, 1997). This implies that ExoI is capable of compensating for the loss of Rad27, but only when over-expressed in a laboratory environment is the phenotype fully recovered (Tishkoff *et al.*, 1997).

Five years later, similar studies were carried out in higher eukaryotes by two separate groups. Matsuzaki *et al* analysed flap endonuclease knockouts in chicken DT40 cells. The cells were viable, but displayed slow growth and sensitivity to both methylating agents and hydrogen peroxide when compared to the wild type cells (Matsuzaki *et al.*, 2002). Kucherlapati *et al* used an *in vivo* model to study the effect of a null mutation within the FEN1 gene in mice. Heterozygotes for the null mutation were viable, but they do display a predisposition towards tumour development, with 17% of mice developing non-Hodgkins lymphoma (Kucherlapati *et al.*, 2002). Homozygotes for the null mutation led to embryonic death.

1.2.3 Properties of Flap Endonucleases

1.2.3.1 Substrate Specificity

Flap endonucleases display two types of structure specific endonuclease activity (gap and flap), and exonuclease activity. These properties are used to process DNA/RNA intermediates formed within the cell.

Exonucleases remove nucleotides from the end of a polynucleotide chain. Eukaryotic flap endonucleases display 5'-3' exonuclease activity. This means they cleave nucleotides from the 5' end, releasing mono, di- or tri-nucleotides. The exonuclease activity of FEN1 allows it to process single-stranded DNA, and double-stranded DNA with branched structures, including a 5' flap, nick, gap or recessed end (Figure 1-12) (Harrington & Lieber, 1994b; Murante *et al.*, 1994). The human FEN can process 5'-recessed ends, but cannot process blunt, or 3'-recessed ends (Harrington & Lieber, 1994b). Harrington and Lieber (1994b)

showed that FEN1 was unable to cleave single-stranded DNA. However, earlier work by Goulian *et al* (1990) showed a different result. This work looked at the cleavage capability of cca/exo (later thought to be FEN1) to cleave the single-stranded homopolymers dA, dT and dC (Goulian *et al.*, 1990). The dT homopolymer was a better substrate than the other two tested. However, cleavage of the paired homopolymer to its unlabelled complementary strand was 98% more efficient than the single stranded version (Goulian *et al.*, 1990). This difference between the two groups may be because Harrington and Lieber (1994b) did not use a homopolymer, or may be due to differences in the reaction conditions.

Endonucleases cleave the phosphodiester bond within a polynucleotide chain. This cleavage is often sequence specific. However, flap endonucleases are a unique family of enzymes because they are structure specific (Harrington & Lieber, 1994b). Eukaryotic FENs display optimal cleavage when a 5′ flap and a 1 nucleotide 3′ flap are present. The 3′ flap sits in a pocket in the human FEN (Figure 1-13), where the 3′ hydroxyl group of the nucleotide interacts with the backbone carbonyl of Lys314, and the Thr61 hydroxyl (Tsutakawa *et al.*, 2011). The 5′ flap is thought to fit through a characteristic helical arch, comprised of two alpha helices: α2 and α4. These helices are between 13-15 Å apart, which allows single-stranded DNA to pass through, but is too small for double-stranded DNA, which requires a gap of ~20 Å (Tsutakawa *et al.*, 2011). The pseudo-Y structure (Figure 1-12) can be cleaved by eukaryotic FENs, but the efficiency of the reaction is reduced up to 100-fold when compared to the 5′ flap structure (Harrington & Lieber, 1994b). Additional work shows that all studied eukaryotic FENs are unable to process the following substrates: mismatched bubble, heterologous loops, 3′ flaps, D-loops, Holiday junctions, 3′ overhangs and 5′ overhangs (Alleva & Doetsch, 1998; Bornarth *et al.*, 1999; Lieber, 1997). However, a recent study by Kazak *et al* showed that a mitochondrial *Hs*FEN1 isoform, truncated at the N-terminus, is capable of binding to 5′ RNA flaps, R-loops and, to a lesser extent, D-loops *in vitro* (Kazak *et al.*, 2013).

Gap endonuclease activity is the ability to cleave single-stranded closed-circular DNA. This was initially described for the T5 D15 exonuclease (a bacteriophage T5 FEN family member), and later observed in the human enzyme (Sayers & Eckstein, 1991; Zheng *et al.*, 2005). Zheng *et al* found that *Hs*FEN1 can cleave a DNA bubble structure, which can arise *in vivo* by the stalling of the replication fork following exposure to UV radiation (Zheng *et al.*, 2005). This gap endonuclease activity is often utilised in conjunction with the WRN protein (Zheng *et al.*, 2005).

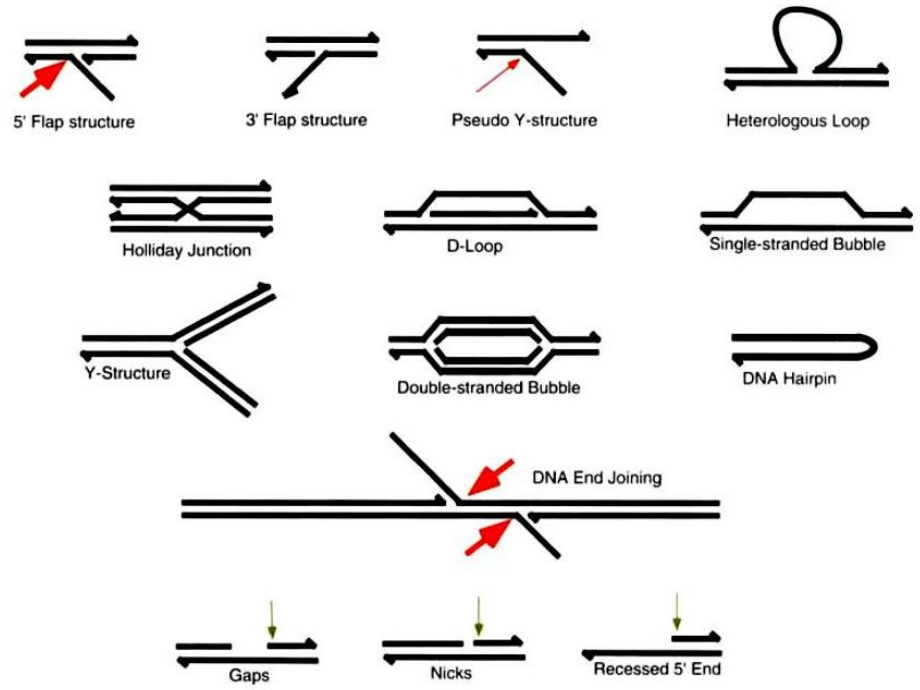


Figure 1-12 Branched DNA structures

A variety of DNA structures available within the cell are depicted, including substrates processed by flap endonucleases. The red arrows indicate the cleavage site for flap endonuclease (FEN) activity. The green arrow indicates the cleavage site for gap endonuclease (GEN) and exonuclease (EXO) activity. Partial arrow heads represent the 3' end of the DNA strand. Reproduced by permission from John Wiley and Sons: BioEssays (Lieber, 1997).

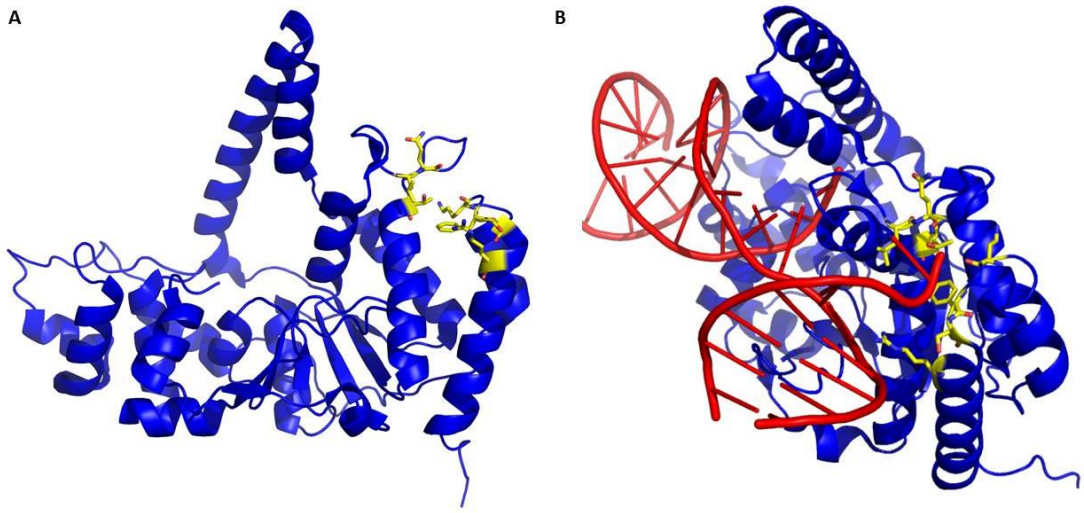


Figure 1-13 The *Hs*FEN 3' nucleotide binding pocket

(**A**) The human FEN (blue) with the 3′ nucleotide pocket highlighted in yellow. Residues L53, Q54, T61, K314, F316, S317 and R320 are shown as sticks (Friedrich-Heineken & Hubscher, 2004). (**B**) An alternative view of the human FEN 3′ nucleotide pocket is shown, with DNA product present (red). The images were produced using RCSB PDB 3Q8L, and rendered using The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

1.2.3.2 The Role of Divalent Metal Ions

Flap endonucleases are members of the metal ion dependent phosphoryl transferase family. This means they catalyse the transfer of a phosphoryl group from a donor molecule onto an acceptor molecule. FENs do this by targeting the phosphodiester bond present between the 3′ oxygen and the phosphorous atom. This bond is then broken by hydrolysis and forms two products: one that terminates in a 5′ phosphate, and a second that commences in a 3′ hydroxyl group.

Until recently, it was thought that the FENs used a two-metal-ion system to facilitate cleavage. In this mechanism, the two metal ions sit either side of the scissile phosphate (Yang, 2008). The first metal ion is situated on the nucleophile side, whilst the second is positioned on the side where the phosphoryl group leaves. The process is summarised in Figure 1-14. Briefly, the first metal ion polarises a water molecule, resulting in deprotonation (Yang, 2008). The nucleophilic water then attacks the phosphorous atom to form a *pentacoordinate intermediate* (Yang, 2008). In this intermediate species, the phosphorous atom is at the centre, with five atoms at the corners (Berg *et al.*, 2002b; Yang, 2008). It is thought that the second metal ion may destabilise the substrate and enhance catalysis (Yang, 2008; Yang *et al.*, 2006). The nucleophile and scissile phosphate then form a phosphoryl bond, and the 3′ leaving group is re-protonated (Yang *et al.*, 2006).

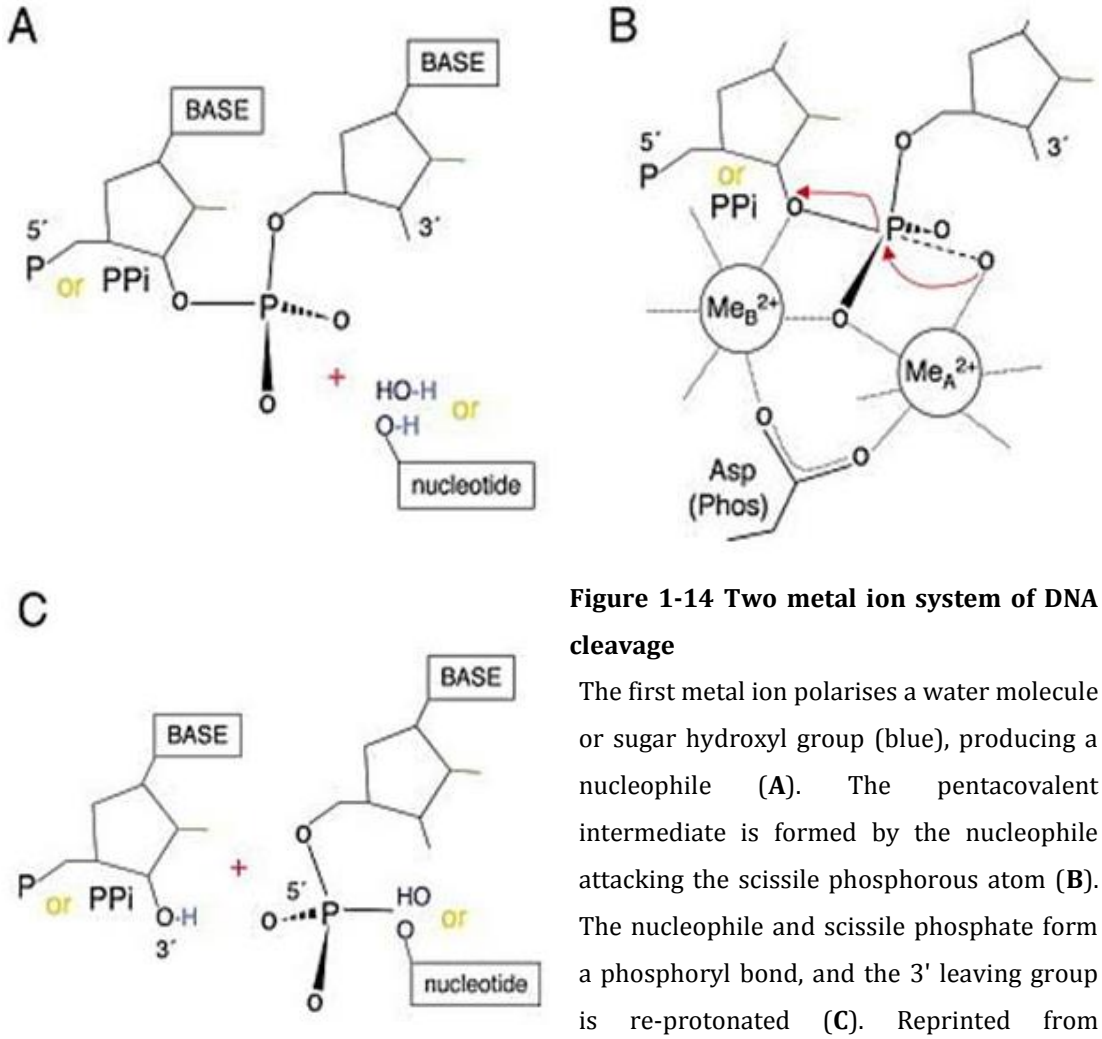


Figure 1-14 Two metal ion system of DNA cleavage

The first metal ion polarises a water molecule or sugar hydroxyl group (blue), producing a nucleophile (**A**). The pentacovalent intermediate is formed by the nucleophile attacking the scissile phosphorous atom (**B**). The nucleophile and scissile phosphate form a phosphoryl bond, and the 3' leaving group is re-protonated (**C**). Reprinted from Molecular Cell (Yang *et al.*, 2006) with permission from Elsevier.

The position of the two metal ions is important, as they are situated very close together. The two magnesium ions can neutralise the negative charge on the intermediate phosphate species. In the two-ion catalysis system, the preference for magnesium as the metal ion is high. This may be because of stringency imposed by the amino acid side chains coordinating the two magnesium ions. Magnesium is the metal ion most often associated with nuclease because of its high natural abundance in cells, small ionic radius (0.65 Å) and redox inertness (Cowan, 1998). However, some enzymes, including FEN family members, are capable of using a wide range of cofactors, as shown in Table 1-1.

Recent work has highlighted the possibility of a three-metal-ion mechanism of action within the T5 FEN. This third divalent metal cation is thought to have a lower association constant for the enzyme in the absence of DNA, and only binds in the presence of the substrate (Syson *et al.*, 2008). The aspartic acid at residue 130 in T5 FEN could be involved in its

coordination (Syson *et al.*, 2008). However, in archaeal and mammalian FENs, the N-terminal region may substitute for one metal ion site (Syson *et al.*, 2008).

Table 1-1 The wide range of cofactors utilised by nucleases

Enzyme	Cofactors	References
TaqI	Mg^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ca^{2+}, Ni^{2+}, Zn^{2+}	(Cao & Barany, 1998).
DNase I	Mg^{2+}, Mn^{2+}, Co^{2+}, Ca^{2+}, Zn^{2+}	(Price, 1975).
BamHI	Mg^{2+}, Mn^{2+}, Co^{2+}, Zn^{2+}, Cd^{2+}	(Viadiu & Aggarwal, 1998).
T5FEN	Mg^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}	(Garforth *et al.*, 2001), (Feng *et al.*, 2004).
EcoRV	Mg^{2+}, Mn^{2+}, Co^{2+}, Ca^{2+}, Ni^{2+}, Zn^{2+}, Cd^{2+}	(Vipond *et al.*, 1995).

1.2.4 Structural Characteristics of Flap Endonucleases

The structure of the human FEN1 (*Hs*FEN) enzyme can be simplified by dividing it into seven distinct structural elements: the helix-two turn-helix (H2TH), β pin, acid block, hydrophobic wedge, helical gateway, active site and cap (Tsutakawa *et al.*, 2011). These elements are described further in Table 1-2, and depicted in Figure 1-15. Seven amino acids in the active site are required for metal ion coordination, and are essential for catalytic function (Tsutakawa *et al.*, 2011). The seven amino acids in *Hs*FEN are D34, D86, E158, E160, D179, D181 and D233. These can be split into two distinct sites: metal ion binding site 1 and 2. Site 1 is comprised of D86 and E160, whilst site 2 is comprised of D179, D181 and E160. The remaining amino acids, D34, D233 and E158, interact with both metal ions via water mediated interactions (see Figure 1-15) (Tsutakawa *et al.*, 2011). Mutational analyses that replaced these seven amino acids with alanine residues has been performed, and in all cases mutations resulted in the loss of *Hs*FEN activity (Shen *et al.*, 1997). However, loss of E158, D179 and D233 also resulted in loss of substrate binding (Shen *et al.*, 1997).

Table 1-2 The seven main structural elements of the human FEN

*Hs*FEN Structural Element	Secondary structures/amino acid residues involved	Function
Helix-two turn-helix (H2TH)	α helices 10 and 11	Binds double-stranded DNA located upstream of the 5′-flap.
β pin	β 6 and 7 loop	Involved in the approximate 100° bend of the DNA template strand that allows the enzyme to recognise 5′-DNA flaps.
Acid block	Residues 56-59	Inhibits longer 3′-flap formation by blocking additional DNA movement (beyond the one base pair pocket) by charge repulsion.
Hydrophobic wedge	α helix 2, and the α2-α3 loop	Breaks the DNA path, and forms the binding site of the 3′-flap.
Helical gateway	α helix 4	Ensures only single stranded DNA/RNA gains access to the active site.
Active site	Residues 34, 86, 158, 160, 179, 181, 233	Coordinates divalent metal ions, binds the 5′-flap and cleaves the scissile phosphodiester bond.
Cap	α helices 4 and 5	Imposes a preference for 5′-termini on the protein.

Information obtained from Tsutakawa *et al* (2011).

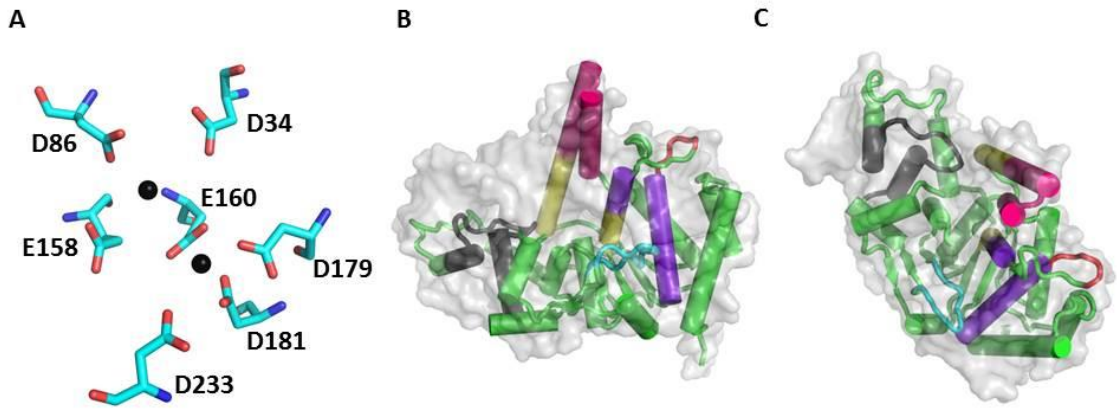


Figure 1-15 The structural elements of the human FEN

(**A**) The seven essential amino acids involved in the coordination of the two divalent metal ions. The five aspartic acid residues and two glutamic acid residues are depicted surrounding two Sm^{3+} ions (black). (**B**) Six of the *Hs*FEN structural elements involved in DNA binding and coordination are depicted. The hydrophobic wedge (purple), helical gateway (yellow), acid block (red), cap (pink), β pin (cyan) and helix-two turn-helix (black) are shown in relation to the enzyme as a whole (grey). The remaining protein structure is shown in green. (**C**) Alternate view of the *Hs*FEN. The images were produced using the RCSB PDB file 3Q8L, and rendered using The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

The other six elements within the *Hs*FEN protein are involved in DNA recognition, and ensure that only the correct substrate is recognised by the enzyme. The H2TH interacts with the double-stranded DNA located upstream from the 5'-flap, whilst the helical cap ensures *Hs*FEN only recognises 5'-termini (Tsutakawa *et al.*, 2011). The hydrophobic wedge and the acid block are involved in the recognition of 3'-flaps. The former produces the binding site for the 3'-flap, whilst the latter ensures that the 3'-flap is no longer than one nucleotide in length (Tsutakawa *et al.*, 2011). Fredrich-Heineken and Hübscher (2004) showed that the extrahelical 3' flap pocket is conserved in FENs, and is comprised of the following amino acids in the human enzyme: L53, Q54, T61, K314, F316, S317 and R320 (Friedrich-Heineken & Hubscher, 2004). Replacing four of these amino acids with alanines (specifically L53, T61, F316 and R320) resulted in severe loss of substrate specificity (Friedrich-Heineken *et al.*, 2003). Replacing all seven amino acids with alanines resulted in complete loss of the 3' flap specificity (Friedrich-Heineken & Hubscher, 2004). However, both mutants retained their general DNA binding ability, but were no longer able to distinguish their 'physiological' substrate from other ones (Friedrich-Heineken & Hubscher, 2004). This conserved hydrophobic pocket binds to the 3' flap and anchors the DNA in a defined orientation that positions the scissile phosphate near the active site (Friedrich-

Heineken & Hubscher, 2004). The helical gateway is the entry point of single-stranded DNA within the enzyme, and it also forms the active site in which the metal ions are located (Tsutakawa *et al.*, 2011). The β pin is involved in bending the template DNA strand 100° (Tsutakawa *et al.*, 2011). All of these elements are involved in the recognition and endonucleolytic cleavage of the DNA/RNA substrate by *Hs*FEN.

The C-terminus is involved in protein-protein interactions and DNA binding. Analysis of four acetylatable lysines in the C-terminus (K354, K375, K377 and K380) was performed by mutating them to alanines (Friedrich-Heineken *et al.*, 2003). This had reduced activity on all flap substrates except for the optimal, double-flap substrate, where the cleavage efficiency remained the same (Friedrich-Heineken *et al.*, 2003). Mutating Arg47 and Arg70 to alanines showed their importance in DNA substrate interactions (Qiu *et al.*, 2002). One of the first proteins shown to interact with *Hs*FEN was PCNA, the processivity factor for the DNA Pol δ and ε. The crystal structure of PCNA, with *Hs*FEN bound to it, is shown in Figure 1-16. *Hs*FEN binds to PCNA by its C-terminus, and 3 molecules of *Hs*FEN bind to one PCNA trimer (Chen *et al.*, 1996). PCNA is therefore thought to sequester *Hs*FEN to its site of cleavage (Li *et al.*, 1995). Within the *Hs*FEN C-terminus, the amino acids TQGRLDDFFK (residues 336-345) were shown to be important for PCNA- *Hs*FEN binding (Warbrick *et al.*, 1997). A report published in the same year described amino acids 328-355 as being required by *Hs*FEN to bind PCNA (Gary *et al.*, 1997). This report went one step further by identifying R339 as being essential for *Hs*FEN -PCNA interaction (Gary *et al.*, 1997). PCNA is not the only protein that interacts with *Hs*FEN. A search performed on the BioGRID interaction database in February 2016 identified sixty one proteins that interact with the human FEN1 (Chatr-Aryamontri *et al.*, 2015). These proteins include those involved in DNA repair, replication and telomere maintenance (Brosh *et al.*, 2002; Dianova *et al.*, 2001; Lee *et al.*, 2011; Sharma *et al.*, 2004a; Sharma *et al.*, 2004b).

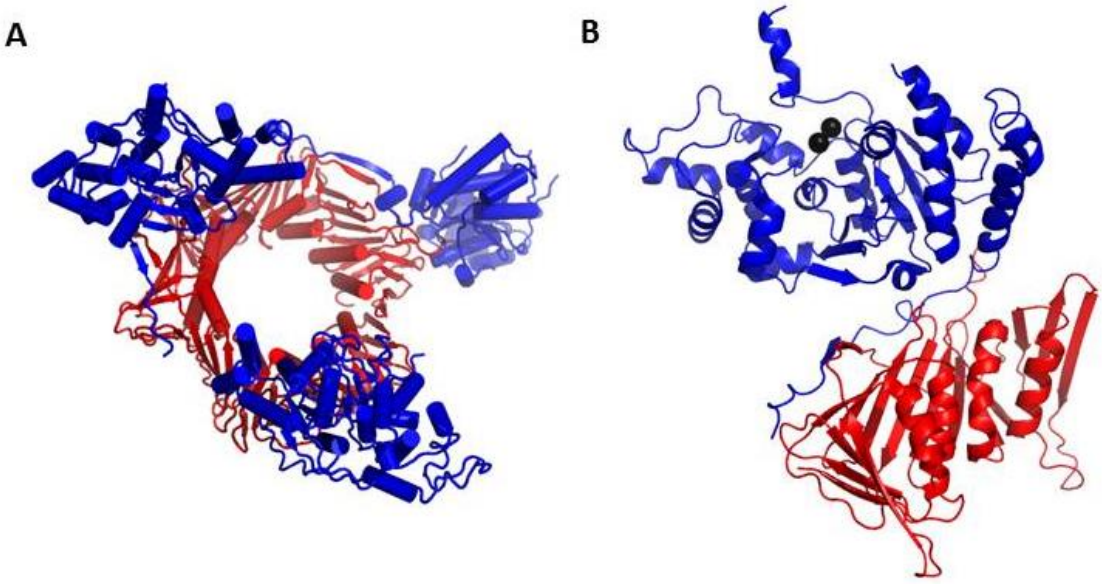


Figure 1-16 Crystal structure of *Hs*FEN bound to PCNA

(**A**) The PCNA trimer (red) bound to three separate *Hs*FEN enzymes (blue). (**B**) A single PCNA monomer attached to a single *Hs*FEN enzyme. Magnesium ions are shown in black. . The images were produced using the RCSB PDB file 1U1L, and rendered using The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

1.2.5 DNA-Protein Binding Models

Until recently, there were two primary models to describe how flap endonucleases bind to 5' flaps. The first model, called the **threading model**, involves the FEN1 enzyme recognising the double-stranded DNA junction. The enzyme then threads the 5' flap through its positively charged helical arch, similar to threading a needle (Xu *et al.*, 2001). The second model, called the **tracking model**, involves FEN1 recognising the 5' end of the 5' flap and tracking along it. Both these models rely on the presence of a free 5' end. Indeed, blocking the 5' end by hairpin formation, annealing primers or the presence of biotin-streptavidin complexes on the flap prevented cleavage (Murante *et al.*, 1995; Robins *et al.*, 1994; Xu *et al.*, 2001). However, smaller modifications to the 5' flap structure, for example small covalent adduct and RNA segments, could be tolerated and did not prevent cleavage (Barnes *et al.*, 1996; Murante *et al.*, 1995). This could not be explained by either model. Much of the data that corroborates the threading model also corroborates the tracking model. Xu *et al* (2001) do provide an interesting theory as to why the threading model is more biologically relevant. This theory was based on the observation that a 30 nucleotide 5'

flap is cleaved with considerably less efficiency then a flap that is 10 nucleotides in length. If the rate-limiting step in this reaction is the formation of the DNA-substrate complex, tracking along three times the amount of nucleotides is not going to take that much longer (Xu *et al.*, 2001). However, threading 30 nucleotides through a gap would take considerably longer than threading 10 nucleotides. A breakthrough came in 2012, when Patel *et al* proposed a new, threading based model. This model, descriptively named the **disorder-thread-order model**, involves the FEN1 enzyme binding the substrate at the double-stranded DNA junction and threading the 5' flap through the disordered helical arch (Patel *et al.*, 2012). Once threading was complete, the arch becomes ordered (Patel *et al.*, 2012). Threading through the disordered arch explains why small 5' flap modifications do not inhibit cleavage, but larger ones do. Binding of the enzyme at the DNA junction appears to orient the 5' flap in such a way that it enters the disordered arch (Patel *et al.*, 2012). Finally, ordering of the helical arch positions the catalytic residues correctly, thus linking reaction site selection to catalysis (Patel *et al.*, 2012).

1.2.6 Targeting the Human Flap Endonuclease

*Hs*FEN is upregulated in a number of cancers, specifically colon, stomach, lung, breast and uterine (Singh *et al.*, 2008). The concept of using *Hs*FEN as an anti-cancer target has been validated *in vitro* by McManus *et al* (2009). They showed that reducing *Hs*FEN activity by either siRNA silencing or inhibition decreases cellular proliferation in human colorectal cancer cells (McManus *et al.*, 2009). The identification and development of specific low molecular weight inhibitors has been of considerable interest (Dorjsuren *et al.*, 2011; McWhirter *et al.*, 2013; Panda *et al.*, 2009; Tumey *et al.*, 2005; Tumey *et al.*, 2004; van Pel *et al.*, 2013). The use of a *Hs*FEN inhibitor alongside chemotherapeutics such as temozolamide (TMZ) would ideally lower the effective dose of TMZ required to kill the cancerous cells (Tumey *et al.*, 2004). This would sensitise the cells to the DNA alkylating agent. Indeed, Panda *et al* (2009) showed that a FEN1 specific small molecular weight inhibitor (NSC-281680) blocked colon cancer growth when administered with TMZ (Panda *et al.*, 2009). This shows the potential of *Hs*FEN inhibitors as anti-cancer therapeutics, either alone or in conjunction with current chemotherapeutic agents. However, *Hs*FEN inhibitors may not be specific to the enzyme in malignant cells, so measures must be in place to both determine and prevent the effects of these inhibitors on healthy cells (Panda *et al.*, 2009).

1.3 Research Hypothesis

There is a wealth of information available regarding flap endonucleases from bacteriophages, bacteria, archaea, yeast and mammals. The flap endonucleases of one category of organism have thus far been ignored: protozoans. Within this category are many medically relevant organisms where additional therapeutic targets are urgently required. The focus of my thesis is on the flap endonuclease from the *Trypanosoma brucei* parasite, the causative agent of African trypanosomiasis. The hypothesis that this thesis will test is:

The *Trypanosoma brucei* flap endonuclease has the potential to be a drug target because the functional protein is essential for parasite survival.

1.4 Project Aims

In order to test my research hypothesis, five project aims were identified. These aims are as follows:

1. **To over-express, purify and begin to characterise the *T. brucei* FEN (Chapter Two).**

Using a lactose-inducible expression system, the wild type enzyme was over-expressed and purified. Characterisation of the enzyme was started, using a number of biochemical techniques.

2. **To produce an active site mutant that can bind DNA, but is catalytically inert in the presence of magnesium ions (Chapter Two).**

As of yet, no substrate: protein complex has been identified where the scissile phosphodiester bond rests in the active site. The most recently published *Hs*FEN structure (PDB 3Q8L) is of the enzyme with the substrate, but the scissile phosphate is 5.2 Å away from the active site, and the ions occupying the divalent metal cation sites are samarium ions (Sm^{3+}) (Tsutakawa *et al.*, 2011). An active site mutant that supports DNA binding, but not cleavage, was produced. This was sent for crystallisation trials with collaborators in the Department of Molecular Biology and Biotechnology, The University of Sheffield.

3. **To determine if the *T. brucei* FEN is a potential therapeutic target (Chapter Three).**

In collaboration with Dr Helen Price at Keele University, the *Trypanosoma brucei* FEN was knocked down, and the wild type and D183K variant were over-expressed within the

parasite. Analysis of the phenotype produced by these manipulations will determine if the enzyme is an appropriate therapeutic target.

4. **To screen previously identified FEN inhibitors against the *T. brucei* and human FENs (Chapter Four).**

Previous work in the Sayers laboratory identified two FEN inhibitors. One inhibited T5FEN only, whilst the second inhibited both the T5 and T7 FENs. These were tested against the human and *Trypanosoma brucei* FENs alongside a known human FEN inhibitor: myricetin.

5. **To develop a high through-put thermal shift assay capable of screening large fragment libraries for FEN binding (Chapter Five).**

Thermal shift assays are increasing in popularity as initial stages in both fragment and compound screens. A commercially available assay was optimised for both the *Trypanosoma brucei* and human FEN, and tested against myricetin.

2 Mutational Analysis and Characterisation of the *Tb*FEN enzyme

2.1 Introduction

The aim of this chapter is to extend the previous work performed in the Sayers laboratory on the T5 exonuclease to a medically relevant organism. This involved analysing active site mutations where the functional amino acids (either aspartic or glutamic acids) were replaced with lysine residues. The aim is to produce catalytically inert proteins that can bind, but not cleave DNA. This work was primarily performed to produce a crystal structure of a FEN protein with its DNA substrate, in the presence of biologically relevant divalent cations.

2.2 Results

2.2.1 Bioinformatics Analysis of the *Tb*FEN

The *Trypanosoma brucei* FEN1 (*Tb*FEN) protein sequence (XP_843679.1) was obtained from NCBI Protein database. This was compared to the *Homo sapiens* FEN1 (*Hs*FEN) protein sequence (UniProtKB P39748.1). Protein sequences were aligned using Multalin, a multiple sequence alignment software (Corpet, 1988), and shaded in BoxShade (Figure 2-1). The proteins display a 51% sequence identity, as determined by BLASTP (Altschul *et al.*, 1990). The *Tb*FEN protein sequence was used to produce a predicted structure with Phyre2 (Kelley *et al.*, 2015). The structure is shown in Figure 2-2, aligned to the crystal structure of the *Hs*FEN (RCSB PDB 3Q8L). Aligning the Phyre2 prediction with PDB 3Q8L in The PyMOL Molecular Graphics System gives an RMS of 0.000. The predicted *Tb*FEN structure displays the characteristic twisted β-sheet core surrounded by α helices. These provide the base for the characteristic helical arch that rises above the enzyme's central beta core. Amino acid substitutions were modelled into the active site *in silico* in order to allow initial evaluation of their likely impact on metal ion binding. DNA constructs capable of expressing the seven active-site variants were produced. Two were purified (D34K and D183K), and the D183K variant analysed in detail (Figure 2-3).

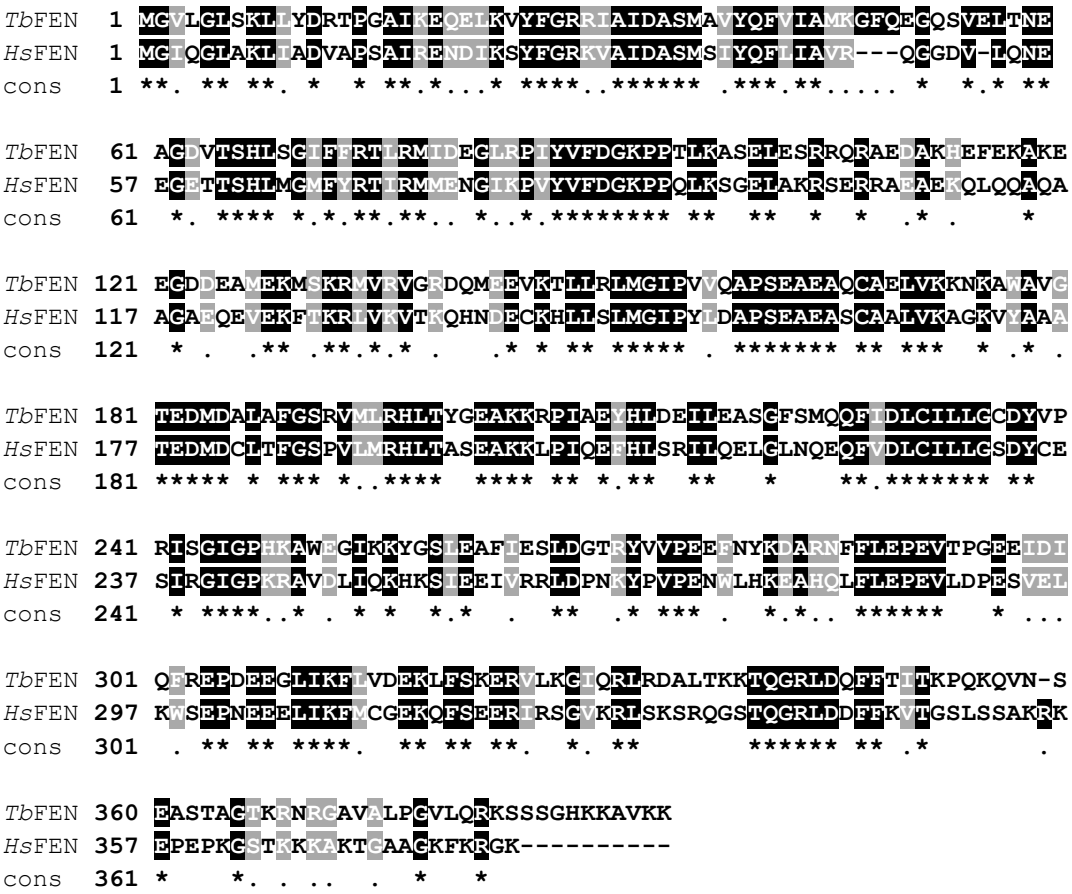


Figure 2-1 Multiple sequence alignment of the *Tb*FEN and *Hs*FEN protein sequences

Black boxes indicate conserved amino acids. Grey boxes indicate similar amino acids. Annotations: ***Tb*FEN;** *Trypanosoma brucei* flap endonuclease, ***Hs*FEN;** human flap endonuclease, **cons;** consensus.

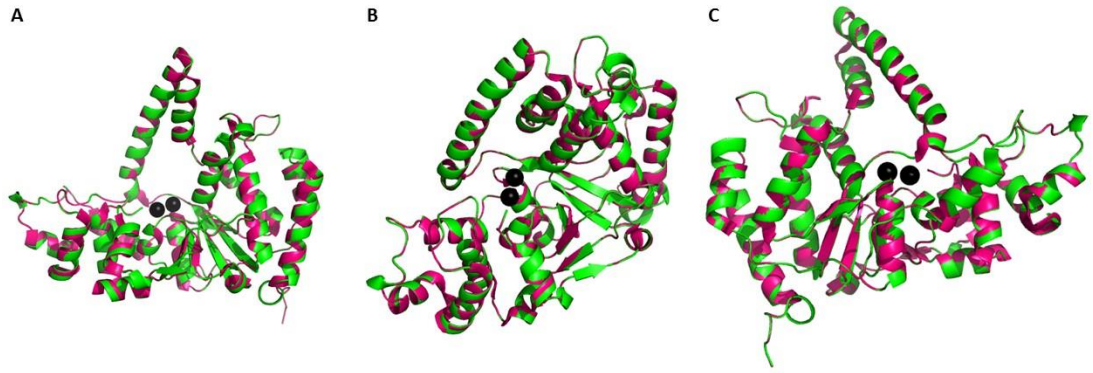


Figure 2-2 Structures of the *Tb*FEN Phyre2 model (green) and *Hs*FEN crystal structure (RCSB PDB 3Q8L; magenta)

Three alternate views of the structures of the *Tb*FEN (green) and *Hs*FEN (magenta). Samarium ions (Sm^{3+}) are depicted as black spheres. The images were rendered by the author using The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

2.2.2 Construct Production and Protein Purification

A synthetic *Tb*FEN sequence, codon optimised for expression in *Escherichia coli*, was designed and supplied by Eurofins MWG Operon. The gene construct was inserted into the lactose inducible pET21a+ vector (Figure 2-4). The protein sequence was analysed using Expasy ProtParam, an online resource that deduces a number of physico-chemical properties from the protein sequence (Gasteiger E., 2005). Knowledge of the theoretical pI (7.66) was used to assist with protein purification. The success of the purification procedure was monitored by SDS-PAGE (Figure 2-7). Zymogram analysis of the wild type enzyme was used to assess nuclease activity of the protein, and to monitor the presence of any contaminating nucleases (Figure 2-7).

A sequence encoding the *Tb*FEN D183K variant was originally synthesised by Eurofins in a truncated form, missing the last 53 amino acids (*Tb*FEN D183K ΔC). An insert encoding the C-terminal 53 amino acids were introduced using an oligonucleotide cassette by standard recombinant DNA technology, producing the *Tb*FEN D183K full-length construct (Figure 2-5).

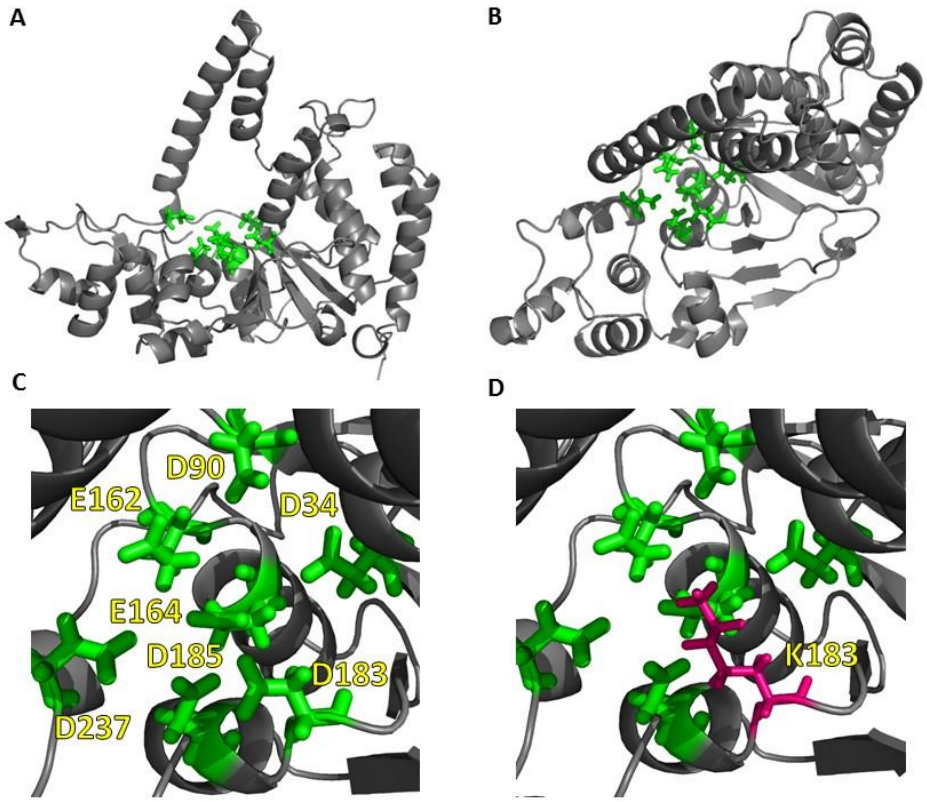


Figure 2-3 The active site of the wild type and D183K *Tb*FEN variant

(**A, B**) Two alternate views of the Phyre2-predicted *Tb*FEN structure, depicting the seven conserved amino acids located in the enzyme active site. Residues are displayed as sticks, and are coloured green (**C**) A close-up of the wild type active site. (**D**) A close-up of the D183K mutation within the active site. The mutated lysine is coloured magenta. Images were produced using the predicted Phyre2 *Tb*FEN structure, and rendered by the author using The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

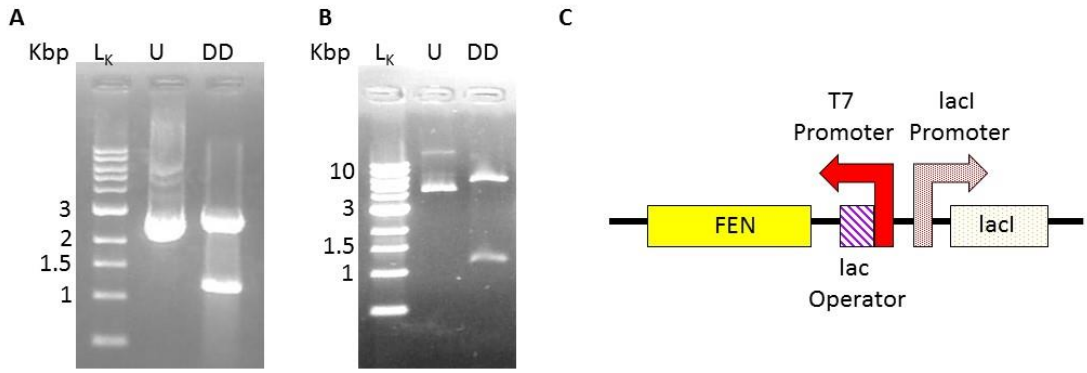


Figure 2-4 Gel electrophoresis of the cloning process to insert the *Tb*FEN construct into a lactose inducible expression vector

(**A**) The uncut pBSiiSK (+)-*Tb*FEN construct (U) was digested by *NdeI* and *HindIII* (DD) to remove the *E. coli* optimised *Tb*FEN gene. (**B**) Successful sub-cloning into the pET21a+ expression vector was monitored by agarose gel electrophoresis of uncut (U) and restriction digested (DD) plasmid and confirmed by DNA sequencing (Medical School Core Genomics Facility). (**C**) A schematic showing part of the pET21a+ expression vector and the arrangement of promoters and repressor. **L_K**; 1 Kbp DNA ladder (NEB).

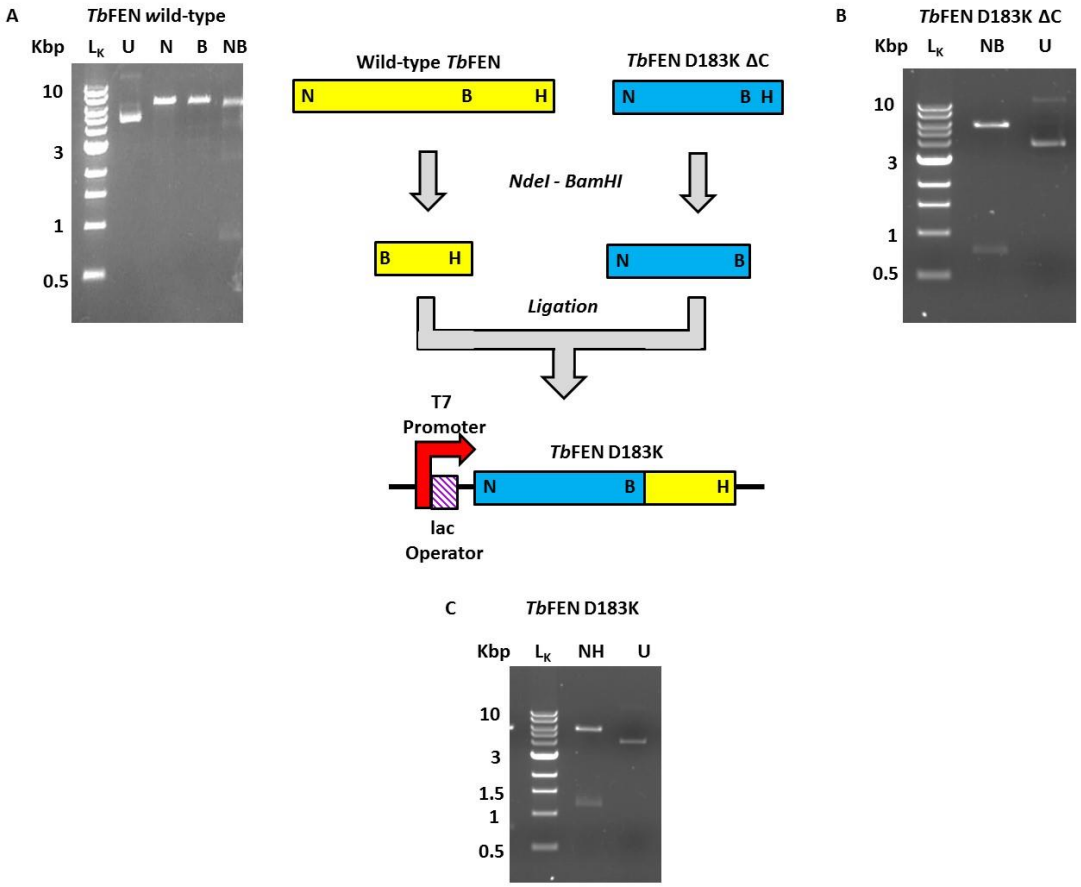


Figure 2-5 Production of the full length *Tb*FEN D183K construct

(**A**) The full length wild type pET-*Tb*FEN construct was digested by NdeI (N) and BamHI (B), and the resulting pET-C-terminus product was purified. (**B**) The pET-*Tb*FEN D183K ΔC construct was digested by NdeI and BamHI (NB), and the resulting cassette was purified. (**C**) The *Tb*FEN D183K cassette was ligated to pET-Cterm, producing the pET-*Tb*FEN D183K full length construct. The success of the procedure was determined by restriction digest with *NdeI* and *HindIII* (NH), and confirmed by DNA sequencing (Medical School Core Genomics Facility). Each step was monitored by gel electrophoresis. **L_K**; 1 Kbp DNA ladder (NEB), **U;** undigested plasmid.

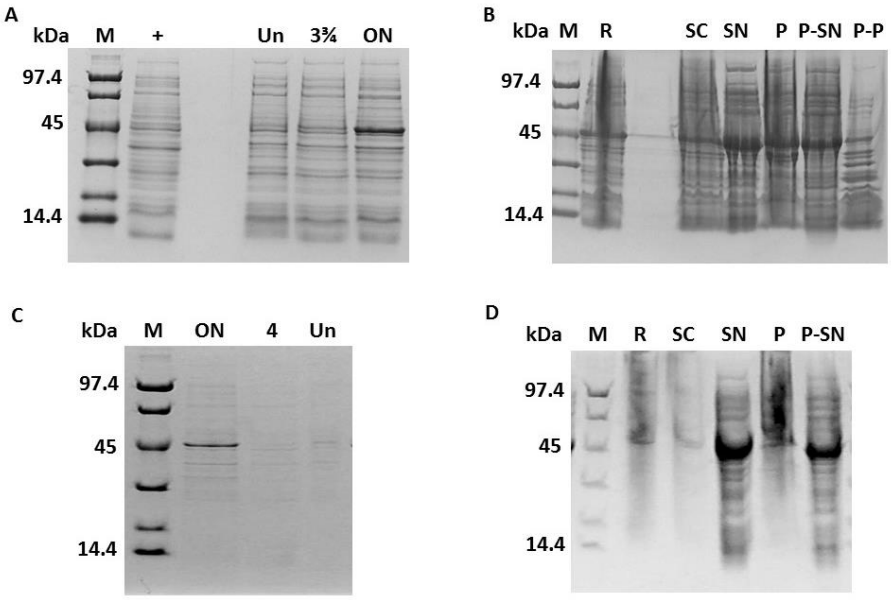


Figure 2-6 SDS-PAGE analysis of the over-expression and lysis of the recombinant wild type *Tb*FEN (A, B) and the *Tb*FEN D183K variant (C, D)

(**A, C**) Analysis of the un-induced (Un) and induced cell pellets at two time points post-induction. Samples were normalised against A_{600}. (**B, D**) Analysis following lysis of the induced cell pellet, as described in the materials and methods. **M**; low range protein marker (Bio-Rad), **+**; control of the empty pET21a+ vector in the same *E. coli* strain, **3 ¾, 4;** time post-induction, **ON;** overnight induction, **R**; sample following re-suspension of the cell pellet, **SC**; sample after sonication, **SN** and **P**; supernatant and pellet respectively, following centrifugation of the lysed, sonicated sample, **P-SN** and **P-P**; supernatant and the pellet respectively, from polyethylenimine precipitation and centrifugation step.

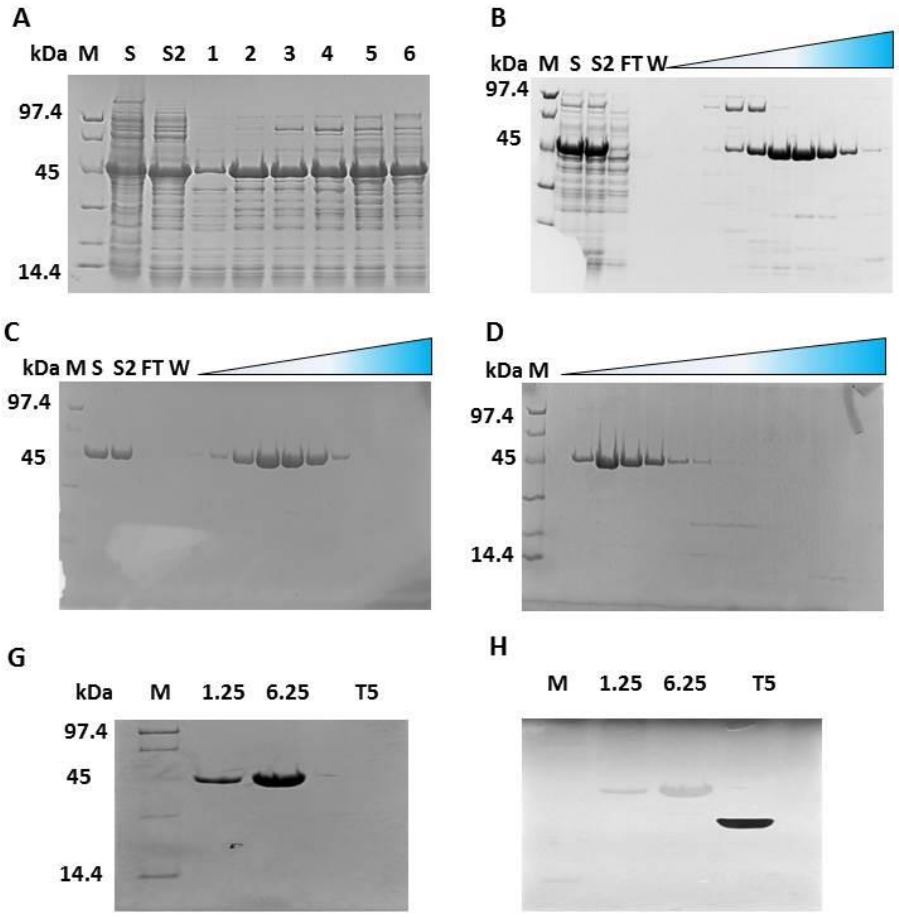


Figure 2-7 SDS PAGE analysis of the wild type *Tb*FEN purification process

Soluble *Tb*FEN was purified using a series of chromatography steps. (**A**) *Tb*FEN was applied to a Bio-Rad High Q chromatography column (pH 8.7) to which it did not bind. (**B**) The flow through from the previous step was injected into a Bio-Rad High S chromatography column (pH 6.7). Fractions were eluted over a 0.05-0.75M salt gradient, and are shown in ascending order (white → blue). (**C**) Pooled fractions F4-9 underwent buffer exchange by dialysis before being applied to a GE Healthcare HiTrap Heparin HP column (pH 6.7). Fractions were eluted over a 0.05-0.75M salt gradient, and are shown in ascending order (white → blue). (**D**) Fractions 4-10 were concentrated and further purified by size exclusion chromatography (GE Healthcare Superdex 200 10/300 GL, pH 6.7). Purified *Tb*FEN was concentrated and analysed by zymogram and stained with both Coomassie blue (**G**) and ethidium bromide (**H**). Annotations: **M;** low range protein marker (Bio-Rad), **S**; sample from the dialysis bag, **S2**; supernatant from the dialysis bag following centrifugation and filtration to remove precipitation, **1-6**; flow through samples collected during column loading, **FT**; flow through collected during column loading, W; column wash with the loading buffer, **T5**; T5Δ192CA positive control. **1.25/6.25;** indicates the amount (in µg) of pure *Tb*FEN loaded per well.

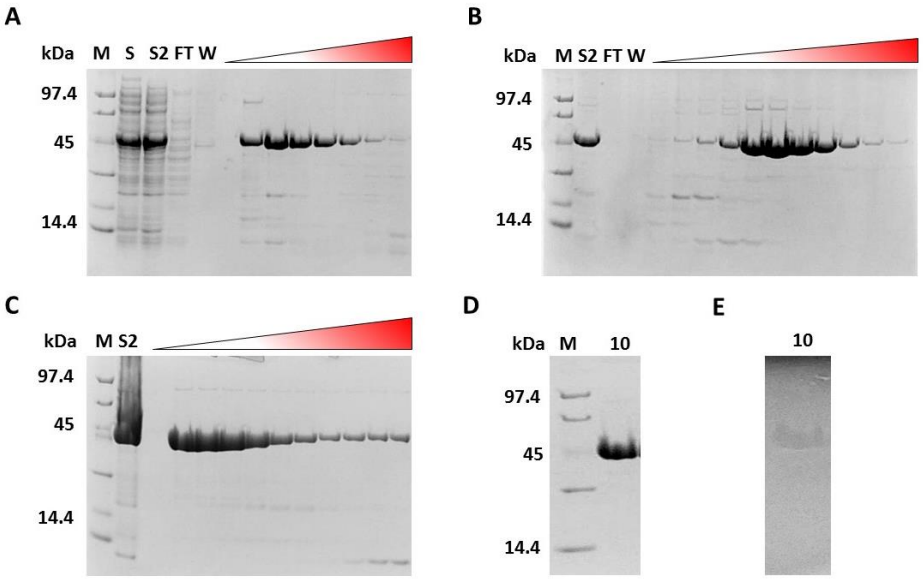


Figure 2-8 SDS PAGE analysis of the *Tb*FEN D183K purification process

Soluble *Tb*FEN D183K was purified using a series of chromatography steps. (**A**) *Tb*FEN D183K was applied to a Bio-Rad High Q chromatography column coupled to a Bio-Rad High S chromatography column. Fractions were eluted from the High S column only over a 0.05-0.75 M salt gradient, and are shown in ascending order (pH 7.5). (**B**) Pooled fractions 7-12 underwent buffer exchange by dialysis before being applied to a GE Healthcare HiTrap Heparin HP column (pH 7.5), and fractions were eluted over a 0.05-0.75 M salt gradient. (**C**) Fractions were pooled based on purity, then concentrated and purified further by size exclusion chromatography (GE Healthcare Superdex 200 10/300 GL, pH 6.7). Purified *Tb*FEN D183K was concentrated and analysed by zymogram and stained with both Coomassie blue (**D**) and ethidium bromide (**E**). Annotations: **M;** low range protein marker (Bio-Rad), **S**; sample from the dialysis bag, **S2**; supernatant from the dialysis bag following centrifugation and filtration to remove precipitation, **FT**; flow through collected during column loading, **W**; column wash with the loading buffer, **10**; indicates the amount (in μg) of pure *Tb*FEN D183K loaded per well.

2.2.3 The Effect of the Active Site Mutation on Structure and Stability

The effect of the D183K mutation on protein structure and stability were compared to the wild type enzyme using two methods: circular dichroism and thermal shift. The buffer used for circular dichroism was initially tested to ensure its suitability. The presence of sodium chloride does affect the spectra, but only at wavelengths below 195 nm (data not shown). The spectra were therefore measured between 195-250 nm, and samples were normalised before analysis. The wild type *Tb*FEN spectra has negative values at 222 nm and 208 nm (-4.8 $x10^{12}$ and -5.5 $x10^{12}$ deg cm^2 $dmol^{-1}$ respectively) (Figure 2-9). The spectra also displays a negative value at 218 nm (-4.8 $x10^{12}$ deg cm^2 $dmol^{-1}$) and a positive value at 195 nm (3.1 $x10^{12}$ deg cm^2 $dmol^{-1}$) (Figure 2-9). The spectra were analysed further using the K2D3 software (http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d3/) (Louis-Jeune *et al.*, 2012). Both constructs yield the same result; specifically the proteins are comprised of 71 % alpha helical regions, and 4.6 % beta sheet regions.

A thermal shift-based assay was used to assess the effect of the active site mutations on the stability of the proteins (Lavinder *et al.*, 2009). The assay conditions were optimised, and are described in Chapter 5. The D183K mutation had a de-stabilising effect on the *Tb*FEN protein (Figure 2-10).

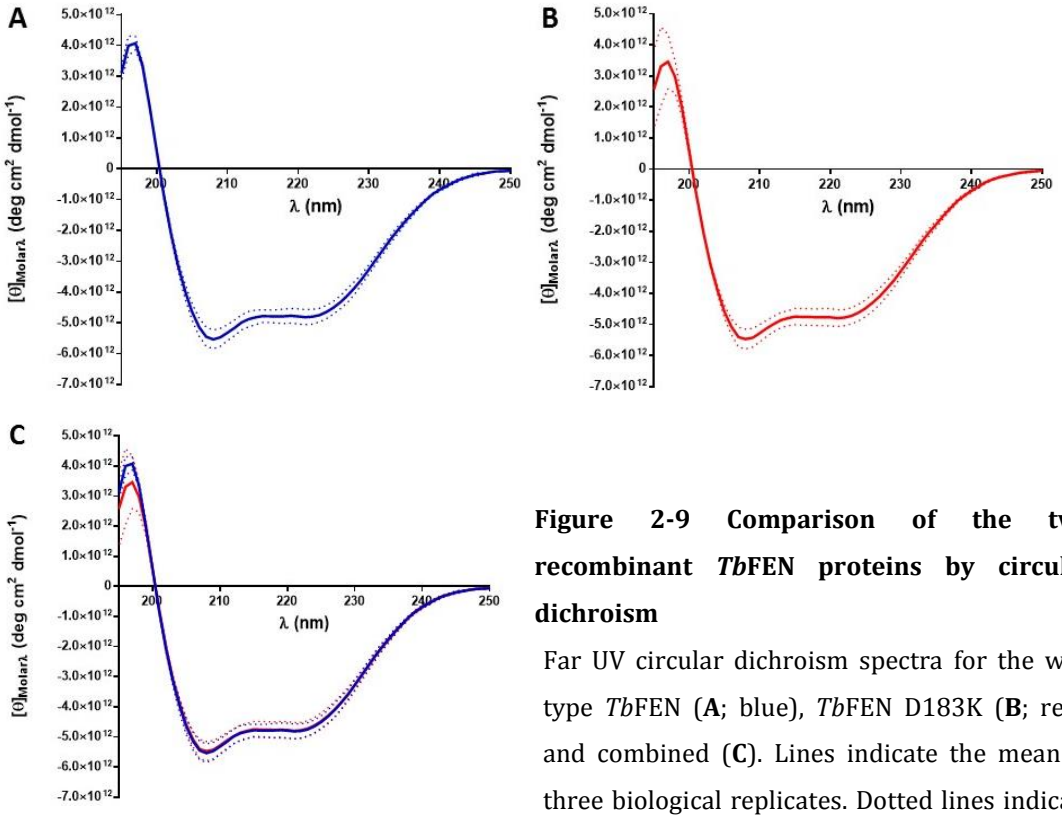


Figure 2-9 Comparison of the two recombinant *Tb*FEN proteins by circular dichroism

Far UV circular dichroism spectra for the wild type *Tb*FEN (**A**; blue), *Tb*FEN D183K (**B**; red), and combined (**C**). Lines indicate the mean of three biological replicates. Dotted lines indicate standard deviation. Spectra are given as molar ellipticity (Θ) at different wavelengths (λ).

2.2.4 The Wild type *Tb*FEN Displays Divalent Cation Specificity

The effect of three different divalent metal cations (Mg^{2+}, Mn^{2+}, Ca^{2+}) on the ability of the wild type *Tb*FEN to cleave a non-specific DNA substrate was assessed. In the presence of Mg^{2+}, wild type *Tb*FEN cleaved DNA with a specific activity of 240 ± 20 U/mg (Figure 2-11). The presence of Mn^{2+} nearly doubled the rate of DNA cleavage when compared to Mg^{2+}, with a specific activity of 430 ± 70 U/mg. Both Mg^{2+} and Mn^{2+} significantly increased the specific activity of the wild type *Tb*FEN compared to the protein in the absence of any divalent metal cations ($p \leq 0.0001$ for both conditions). The presence of Ca^{2+} appeared to inhibit the cleavage ability of the wild type *Tb*FEN. The specific activity in the presence of Ca^{2+} was not significantly different to the *Tb*FEN in the absence of any divalent metal cations ($p = 0.9988$).

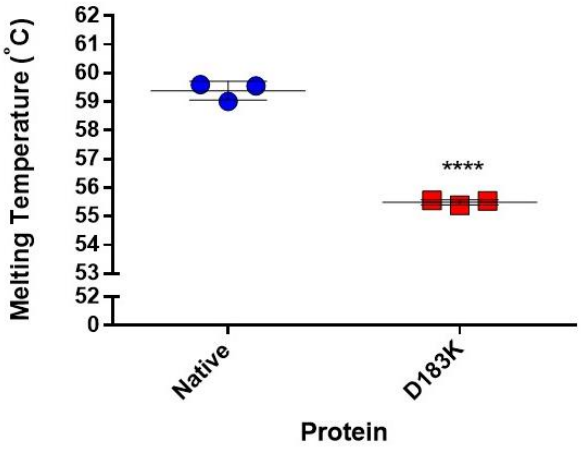


Figure 2-10 The effect of the D183K mutation on structural stability

The melting temperature (T_M) was calculated for each protein using the Boltzmann equation (GraphPad Prism 6.0). Error bars show standard deviation of three biological replicates. The D183K mutation significantly decreases the T_M of the protein by approximately 5˚C (calculated by an unpaired T test in GraphPad Prism 6.0).

2.2.5 The D183K Variant is Catalytically Inert

The *Tb*FEN D183K variant was predicted to be catalytically inert. This was initially assessed using a spectrophotometric UV assay on a non-specific DNA substrate (Figure 2-11). The presence of ten times the amount of the *Tb*FEN D183K variant did not show cleavage of the substrate after two hours in the presence of either Mg^{2+} or Mn^{2+}. This yielded a specific activity close to zero for D183K (3.2 ± 1.5 and 1.5 ± 0.4 U/mg in the presence of either Mg^{2+} or Mn^{2+} respectively). This difference was statistically significant when compared to the activity of the wild type enzyme ($p \leq 0.0001$). The activity displayed by the *Tb*FEN D183K variant in the presence of either Mg^{2+} or Mn^{2+} was not significantly different to the wild type *Tb*FEN in the absence of any divalent metal cations ($p = 0.9978$ and 0.9991 respectively).

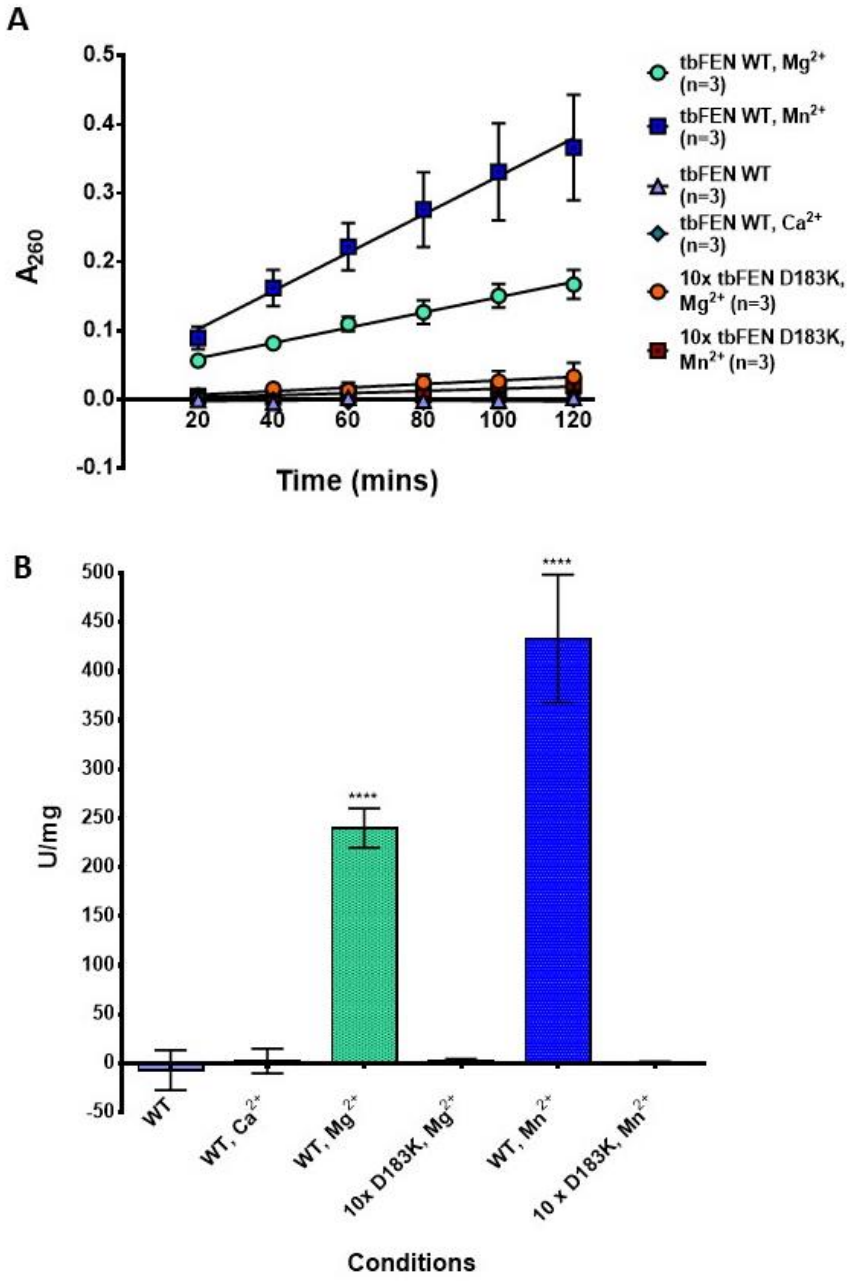


Figure 2-11 Ability of the wild type *Tb*FEN and D183K variant to release acid soluble products from high molecular weight DNA

(A) The effect of three different divalent metal cations, and the D183K active site mutation, on exonuclease activity. Hydrolysis of the high molecular weight DNA substrate was assayed by UV spectroscopy as described in the Materials & Methods. Error bars indicate standard error of the mean of three biological replicates. **(B)** Linear regression analysis was fitted in GraphPad Prism (6.0), and used to calculate the specific activity per mg of protein (U/mg). Error bars indicate standard deviation of the mean. Statistical significance was measured by One Way ANOVA in GraphPad Prism (6.0), and significant results (when compared to the WT) are shown.

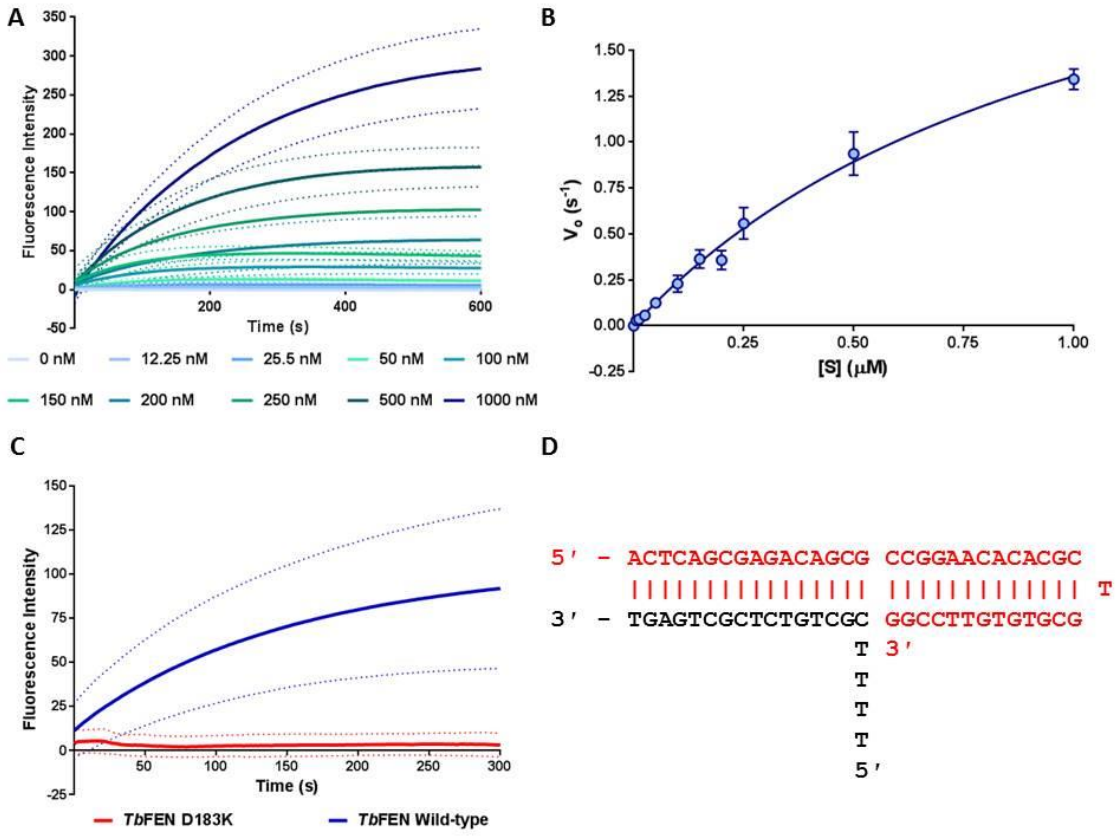


Figure 2-12 Substrate specificity and activity, determined using a Förster resonance energy transfer (FRET) based assay.

(**A**) Reaction progress was monitored by fluorescence increase over time at the substrate concentrations indicated. Curves were normalised against no protein controls. (**B**) The initial rate of reaction (V_0: taken in the first thirty seconds) was plotted against substrate concentration and analysed by non-linear regression of the Michaelis-Menton equation (GraphPad Prism 6.0). (**C**) Comparison of the fluorescence profile of 340 nM of wild type and D183K *Tb*FEN against 250 nM of the 5' flap substrate. (**D**) The structure of the DNA substrate. One oligonucleotide (red) forms a hairpin, whilst the second oligonucleotide (black) is labelled with Cy3 at the 5' end and FAM at the 3' end. The mean of three biological replicates is shown, with error bars displaying standard deviation.

Table 2-1 Catalytic parameters of the wild type TbFEN protein against the 5' flap substrate

Substrate	**k_{cat} (sec^{-1})**	**K_M (µM)**	**k_{cat}/K_M (μM^{-1} sec^{-1})**
Four nucleotide 5' flap	8.142 ± 1.116	1.088 ± 0.2329	7.483

Results depict the mean of three biological replicates ± standard deviation.

A more specific, Förster resonance energy transfer (FRET) based assay was used to assess the substrate specificity of the wild type enzyme. A double stranded substrate with a four nucleotide 5' flap was tested (Figure 2-12D). Initial rates of reaction were measured at different substrate concentrations ranging from 12.25 nM to 1000 nM (Figure 2-12). The catalytic parameters were calculated by non-linear regression of the Michaelis-Menton equation (Table 2-1). The *Tb*FEN D183K variant was also tested using the FRET-based assay (Figure 2-12C). A single concentration of DNA substrate (250 nM) was tested against 340 nM of *Tb*FEN protein, either wild type or D183K. As shown in Figure 2-12 C, incubation of the substrate with the *Tb*FEN D183K variant does not produce the increase in fluorescence that is associated with DNA cleavage. This indicates that the D183K mutation renders the protein catalytically inert.

2.2.6 The Wild type and D183K Variant Bind DNA at a Similar Level

The ability of both the wild type and D183K enzymes to bind to the 5' overhang DNA substrate was measured using a bio-layer interferometry-based binding assay. Biotinylated DNA substrate was attached to a streptavidin probe (Figure 2-13). Binding of the enzyme to the probe: DNA complex was measured, and catalytic parameters calculated using the in-built software (BLItz Pro version 1.1.0.31, FortèBIO) (Figure 2-14). A peak is seen at the start of the dissociation stage in the presence of high concentrations of the wild type *Tb*FEN. The affinity of the *Tb*FEN D183K variant for the DNA substrate is approximately double that for the wild type enzyme (2.22 and 1.37 µM respectively) (Table 2-2). The k_{off} values remain fairly similar at 0.050 and 0.048 1/Ms for the wild type and mutant respectively. The wild type enzyme associates with the DNA substrate with greater affinity than the D183K mutant enzyme (k_{on} = 1.64 x10^7 and 2.20 x10^4 1/s respectively). Despite this, no statistical significance was observed for the K_d, k_{on} or k_{off} between the wild type and D183K variant, calculated using an unpaired T Test in GraphPad Prims (p = 0.295, 0.875 and 0.374 respectively).

Table 2-2 Binding parameters for the wild type and D183K *Tb*FEN enzymes to a 5' overhang DNA substrate.

Protein	K_d (µM)	k_{on} (1/s)	k_{off} (1/Ms)
Wild type	1.37 ± 0.84	0.050 ± 0.018	1.64 x10^7 ± 2.83 x10^7
D183K	2.22 ± 0.89	0.048 ± 0.016	2.20 x10^4 ± 0.36 x10^4

Results depict the mean of three biological replicates ± standard deviation. Parameters were analysed for statistical significance using an unpaired T test (GraphPad Prism 6.0). There was no statistically significant difference between the wild type and D183K variant.

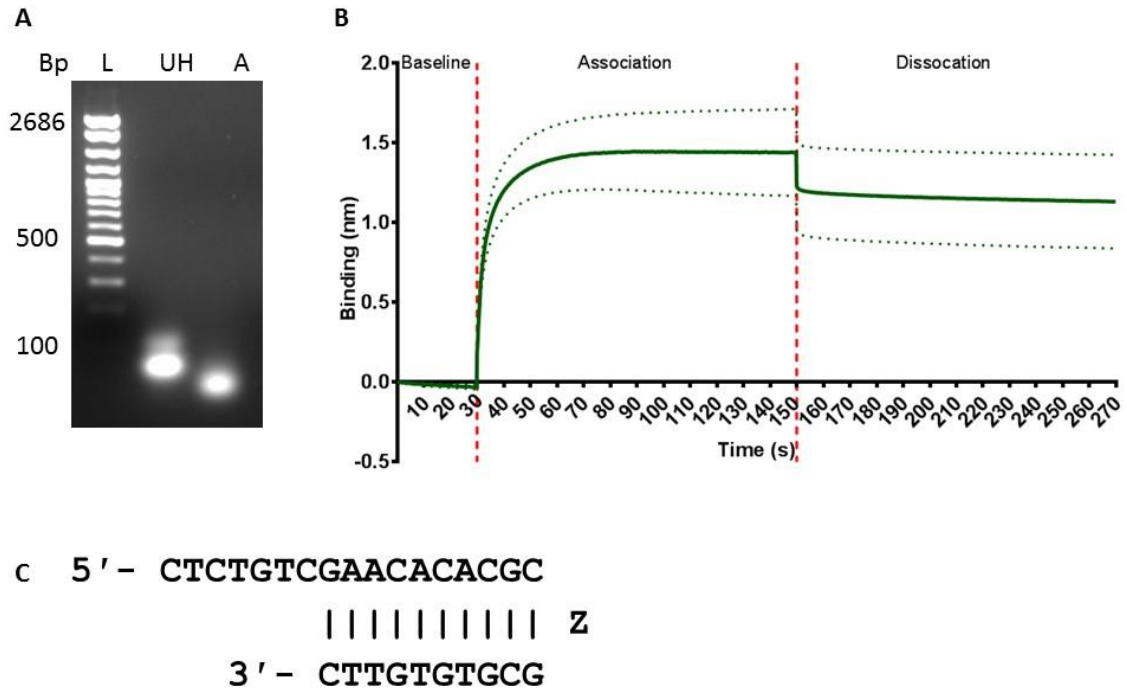


Figure 2-13 Preparation and Association of the OHP2 substrate to the streptavidin probe

(**A**) Gel electrophoresis (2%) of the Bio_OHP2 substrate before (UH) and after (A) heating at 95°C for five minutes. This creates the double-stranded 5' overhang substrate. (**B**) Immobilising the biotinylated OHP2 substrate to the streptavidin probe. Binding buffer was applied for 30 seconds, followed by the annealed substrate for 120 seconds. This was then washed for 120 seconds in binding buffer to remove unbound DNA. The graph depicts the mean association of biotinylated OHP2 from 11 runs. Error bars are depicted as dotted lines, and show standard deviation from the mean. (**C**) The sequence and double-stranded structure of the Bio_OHP2 substrate, which contains a seven nucleotide 5' overhang. The biotin label is located on the hairpin bend (Z).

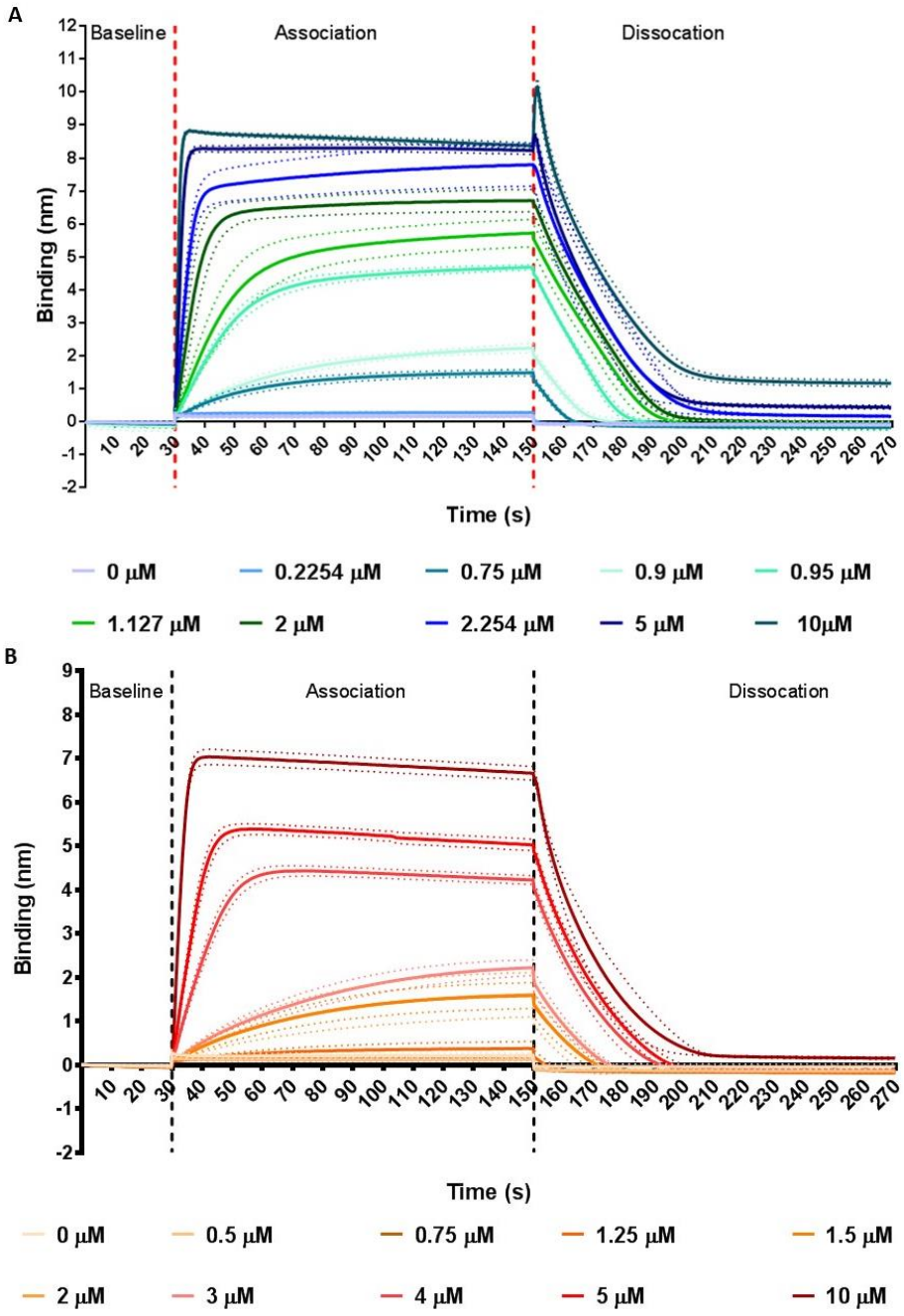


Figure 2-14 Binding of the wild type (A) and D183K variant (B) to the OHP2 substrate

(**A**) Association and dissociation of the wild type *Tb*FEN to the immobilised OHP2 substrate. All protein concentrations were repeated at least in duplicate. Error bars are depicted as dotted lines, and show standard deviation from the mean. (**B**) Association and dissociation of the *Tb*FEN D183K to the immobilised Bio_OHP2 substrate. All protein concentrations were repeated at least in duplicate. Error bars are depicted as dotted lines, and show standard deviation from the mean. All experiments were performed on three biological replicates. Graphs depicted here display average values of one biological replicate.

2.3 Discussion

2.3.1 Bioinformatics Analysis of the *Tb*FEN

Comparing the protein sequence of both the *Tb*FEN and *Hs*FEN enzymes shows they have a sequence identity of 51%. The C-termini display less similarity to each other. This is thought to be where protein-protein interactions occur. *Tb*FEN displays an additional 12 amino acids at the very end of the protein sequence (Figure 2-1). Interestingly, a similar protein sequence (K*$&GH^K#V^K, where * is A/G/S/T, $ is S/T, & is S/A/G, ^ is K/R and # is A/M/V) is seen in flap endonucleases in all trypanosomatidae currently sequenced (data not shown). This may indicate a specific function within this order that is currently unknown, and deserves further investigation.

The structure prediction tool Phyre2 was used to model the *Tb*FEN protein structure. Whilst other tools are available, Phyre2 has a user friendly interface with minimal difference in the accuracy of structural prediction when compared to other software (Kelley *et al.*, 2015). The prediction software works on two assumptions: firstly that protein structure is more conserved than protein sequence, and secondly that there are a finite number of unique protein folds (Kelley *et al.*, 2015). This allows the server to predict the protein structure of unknown proteins based on known structures. For the purposes of producing the hypothetical *Tb*FEN structure, the structure of *Hs*FEN (PDB 3Q8L) was used as a model. It is therefore not surprising that the two align perfectly with a RMS value of 0.00 Å in The PyMol Molecular Graphics System. One limitation with the Phyre2 server is that is cannot accurately determine the effect of point mutations on the structure of the enzyme as a whole, and further analysis would need to be performed (Kelley *et al.*, 2015). The hypothetical structure of *Tb*FEN examined here could inform drug discovery work by predicting binding modes and compound affinities (Breda *et al.*, 2006).

The active-site mutation produced and studied in this chapter was designed specifically to produce a protein that binds to DNA, but is catalytically inert. Seven amino acids within the enzyme's active site are conserved and are required for metal ion coordination and catalytic function. Within *Hs*FEN, these are D34, D86, E158, E160, D179, D181 and D233 (Tsutakawa *et al.*, 2011). These are equivalent to D34, D90, E162, E164, D183, D185 and D237 in the *Tb*FEN. Performing single alanine mutations to these seven amino acids within the *Hs*FEN enzyme caused loss of enzymatic activity in all cases. Substrate binding ability

was also lost in the E158A, D179A and D233A mutations (Shen *et al.*, 1997). The aspartic acid and glutamic acid residues are all negatively charged at physiological pH (Berg *et al.*, 2002a). This allows them to coordinate the divalent metal cations into the active site for substrate cleavage, either directly or by water-mediated interactions (Tsutakawa *et al.*, 2011). Replacing negatively charged amino acids with positively charged lysine residues seemed likely to prevent coordination of one (or both) metal ions. However, it should still allow the negatively charged phosphodiester DNA backbone to enter the active site. Previous work within the Sayers laboratory with the T5 bacteriophage FEN homologue shows that the D155K variant (equivalent to the D183K mutation in *Tb*FEN) was able to bind but not cleave the DNA substrate, and bound with a higher affinity to the mutated active site compared to the wild type (Zhang, 2012).

2.3.2 Construct Production and Protein Purification

Originally, the plasmid encoding the *Tb*FEN D183K variant was supplied in a truncated form, lacking the last 53 amino acids. This was done to aid crystallography of the mutant protein, as the C-terminal amino acid residues encode the flexible PCNA binding domain on the *Hs*FEN homologue. The C-terminus is thought to be disordered when not bound to protein partners, and disordered regions do not form a fixed structure under physiological conditions (Finger *et al.*, 2012; Le Gall *et al.*, 2007). The C-terminus needed to be inserted into the gene because it is important for substrate binding, and was inserted back by restriction digest (Stucki *et al.*, 2001).

A DNA zymogram was used to assess both the purity and nuclease activity of the purified enzymes. Bands of reduced fluorescence normally indicate DNA cleavage. However, these bands are seen for the *Tb*FEN D183K variant, as well as for the wild type enzyme. However, as the work in the rest of this chapter shows, the enzymes are catalytically inert. This may be due to the removal of SDS from the gel, which allows the proteins to re-fold and bind to the DNA. When the DNA is bound to the protein, ethidium bromide can no longer intercalate, thus producing false positive areas of reduced fluorescence. Addition of 1% SDS back into the gel, followed by re-staining with ethidium bromide, removes the band of reduced fluorescence (data not shown). This shows the enzyme is binding to the DNA rather than cleaving it in the zymogram.

2.3.3 The Effect of the Active Site Mutation on Structure and Stability

Circular dichroism is a spectrophotometric method commonly used to determine whether amino acid substitutions affect protein conformation (Greenfield, 2006b). The buffer (20 mM phosphate pH 7.0, 50 mM NaCl, 1 mM DTT, 5% glycerol) was initially tested to ensure its suitability for use. Phosphate buffers are suitable for far UV studies as they do not absorb strongly at these wavelengths (Kelly *et al.*, 2005). Glycerol and NaCl were used for their protective properties, to prevent the protein from precipitating (Kelly *et al.*, 2005). Whilst no tests had been performed to see if this was the case for the *Tb*FEN proteins, they were purified with a minimum of 5% glycerol and 50 mM NaCl. Having these in the buffer was considered a sensible precaution. The reducing agent DTT was added to ensure the side chains of all cysteine residues remained in their reduced state (Kelly *et al.*, 2005). There are three cysteine residues within the *Tb*FEN enzymes, so the presence of DTT was essential. Both DTT and NaCl do absorb strongly in far UV (Kelly *et al.*, 2005), so the buffer was checked for suitability. DTT at the low concentration of 1 mM did not appear to cause interference between 190-250 nm (Figure 2-9). However, NaCl did absorb strongly below 195 nm. Because of this, the proteins were measured between 195 – 250 nm, and readings were normalised against a no protein control.

The wild type *Tb*FEN spectra has negative values at 222 nm and 208 nm (-4.8 $x10^{12}$ and -5.5 $x10^{12}$ deg cm^2 $dmol^{-1}$ respectively) corresponding to alpha helical regions, and a negative value at 218 nm (-4.8 $x10^{12}$ deg cm^2 $dmol^{-1}$) and a positive value at 195 nm (3.1 $x10^{12}$ deg cm^2 $dmol^{-1}$) corresponding to beta sheet regions. These characteristic features of alpha helical and beta sheet regions are also seen the *Tb*FEN D183K protein (Figure 2-9) (Greenfield, 2006b). Both proteins display the same characteristics, specifically defined alpha helical and beta sheet regions that comprised 71.18% and 4.63% of the protein structure respectively. The K2D3 program was used to calculate this because it has been shown to produce improved predictions for beta sheet content within the wavelengths 200-240 nm (Louis-Jeune *et al.*, 2012). The alpha helical, and beta sheet content is to be expected when compared to the hypothetical protein structure produced by Phyre2. A beta-sheet core is surrounded by two alpha helical bundles, and a helical arch or clamp rises above this base (Figure 2-2). The D183K substitution does not appear to have a gross effect on protein conformation.

The thermal shift based assay was used to observe the effect of the D183K mutation on the thermal stability of the enzyme. The assay uses a dye that fluoresces when bound to hydrophobic regions of the protein, and is described in detail in Chapter 5 (Pantoliano *et al.*, 2001). This assay allows a quick and relatively cheap option for analysing the effect of the active site substitution on protein stability. The D183K mutation decreases the melting temperature of the protein, implying decreased thermal stability. This may be partially due to the changes affecting coordination of the divalent metal cations. The presence of Mg^{2+} increases protein stability by over 3% in the WT protein (Chapter 5). It is also likely that the change in charge of the D183K variant has had a negative effect on the protein. Studies have shown that charge-reversal point mutations can either increase or decrease protein stability (Dao-pin *et al.*, 1991; Grimsley *et al.*, 1999; Pace *et al.*, 2000). The effect appears to depend on the ability of the side chains to form hydrogen bonds and ion pairs (Pace *et al.*, 2000). The *Tb*FEN D183K charge reversal point mutation was detrimental in terms of temperature stability, but the exact mechanism remains to be elucidated. It may be due to steric clashes within the active site, resulting in structure destabilisation. Additional, more in-depth experiments should be performed. For example analysis by circular dichroism at different temperatures would allow the thermodynamics of protein unfolding to be determined (Greenfield, 2006a). Another useful technique is differential scanning calorimetry, which has a variety of applications including measuring thermal stability and absolute heat capacity (Johnson, 2013). One benefit of these two techniques is that they do not rely on fluorescent dyes. They also provide considerably more information than the thermal shift assay can.

2.3.4 The Wild type *Tb*FEN Displays Divalent Cation Specificity

The divalent metal cation specificity of the wild type enzyme was tested in the absence and presence of three divalent metal cations; Mg^{2+}, Mn^{2+} and Ca^{2+}. Mn^{2+} has been shown to stimulate cleavage capability of the *Hs*FEN enzyme resulting in an increased rate of reaction (Harrington & Lieber, 1994b). However, for the T5 exonuclease, Mg^{2+} and Mn^{2+} resulted in similar levels of product formation (Garforth *et al.*, 2001). This was explored in the *Tb*FEN enzyme. This work shows that the presence of $MnCl_2$ resulted in a two fold increase in the amount of product formed over a two hour time period, compared to the same concentration of $MgCl_2$. The reason why Mn^{2+} is stimulatory is not fully understood. In EcoRV, it was thought Mn^{2+} enhanced stability and accelerated reaction rate (Baldwin *et al.*, 1999). However, in the T5 FEN, the Mn^{2+} complex was found to be less stable than the Mg^{2+}

complex (Tock *et al.*, 2003). The T5 FEN was also able to catalyse aberrant reactions in the presence of Mn^{2+}, specifically cleavage of double-stranded closed circular plasmid DNA (Garforth *et al.*, 2001). This may explain why the presence of Mn^{2+} causes an increase in acid soluble nucleotides, as it is able to cleave a greater variety of the DNA substrates present within the solution. However, this needs to be analysed further as no investigation into the dependence of reaction rate on metal ion concentration was performed in this thesis. The presence of $CaCl_2$ appears to be inhibitory, which was expected as this has been seen with *Hs*FEN (Harrington & Lieber, 1994b), and T5FEN (Feng *et al.*, 2004). The presence of either Mg^{2+} or Mn^{2+} was required for product formation, as shown by the inability of the enzyme to cleave DNA in the absence of any divalent metal cations. The ionic concentration was kept the same by the addition of more sodium chloride to ensure ionic strength was not a factor in DNA cleavage. This work appears to show that the *Tb*FEN displays similar characteristics to the *Hs*FEN in that Mg^{2+} and Mn^{2+} are absolute requirements for catalysis.

2.3.5 The D183K Variant is Catalytically Inert

The effect of the D183K mutation was initially assessed using the UV spectrophotometric method. The *Tb*FEN D183K variant was tested at a concentration that was ten times that of the wild type enzyme in the presence of both Mg^{2+} and Mn^{2+}. The results show that the *Tb*FEN D183K variant displays activity similar to the wild type enzyme in the absence of any divalent metal ions (Figure 2-11). This inactivity was expected, and is consistent with results previously obtained for the T5FEN (Zhang, 2012). The FRET-based assay allows DNA cleavage to be analysed against a specific substrate. No increase in the fluorescent signal at 520 nm was seen when the *Tb*FEN D183K variant was incubated with 250 nM assay2flap. This indicates that the D183K mutation prevents cleavage of the Cy3 quencher dye from the DNA substrate. This analysis alone does not provide a reason as to why the *Tb*FEN D183K variant cannot cleave DNA. It may be because the enzyme cannot coordinate the divalent metal ions. This would prevent the both formation of the nucleophilic water molecule, and substrate destabilisation, thus preventing DNA cleavage (Yang, 2008; Yang *et al.*, 2006). Another possibility is that the lysine residue is affecting DNA binding by partially or fully blocking the active site. If the scissile phosphodiester is not present in the correct orientation and distance from the active site, cleavage will not occur (Horton *et al.*, 2006). The ability of the enzyme to bind to DNA was explored further in the following section.

2.3.6 The Wild type and D183K Variant Bind DNA at a Similar Level

DNA binding was assessed in the presence of Ca^{2+}. As shown previously, Ca^{2+} is inhibitory towards substrate cleavage, but has been used previously to assess DNA binding interactions with restriction enzymes (Vipond *et al.*, 1995) and T5FEN (Feng *et al.*, 2004; Zhang, 2012). The ability of Ca^{2+} to promote binding, but not cleavage, is thought to be partially due to Ca^{2+} shielding negatively-charged residues located within the active site, thus preventing repulsion of the phosphodiester backbone (Feng *et al.*, 2004).

The substrate that was used to assess DNA binding contains a 5' overhang, and is biotinylated on the hairpin of the substrate (labelled 'Z' in Figure 2-13 C). The hairpin should rest against the streptavidin probe, allowing the 5' overhang to be accessible to the enzyme. Prior to use, the substrate was heated to produce the correct substrate. Samples were run on 2% agarose gels to determine the success of this procedure (Figure 2-13). The sample prior to heating does show as two bands on the gel. This was unexpected, as the substrate should be single stranded, thus preventing any ethidium bromide from intercalating. However, these two bands are likely to be due to the oligonucleotides pairing with each other in an uncontrolled manner. The heating and subsequent cooling of the DNA substrate produces a single band, consistent with the formation of one type, specifically the hairpin with a 5' overhang. In the higher concentrations of wild type *Tb*FEN used in this assay (10 µM, and, to a lesser extent, 5 µM), a slight peak is in seen following the beginning of the dissociation step. This may be due to an excess of protein present.

In order to check the interaction of the protein with the streptavidin probe only, 10 µM of the *Tb*FEN was incubated with the empty probe (data not shown). No interaction was seen, and the results were comparable to the buffer against the probe. A further control was carried out to ensure it is a protein specific interaction with the DNA: probe complex. A known haem binding protein, Gly1-0579 from *Mannheimia haemolytica* (kindly donated by Dr Hannah McMellon) was tested at a concentration of 10 µM. Again, no interaction was seen, similar to that of the *Tb*FEN against the probe only.

The D183K variant is able to bind to the DNA substrate with a similar affinity to the wild type enzyme. The mean K_d value for this interaction is twice as high as that for the wild type enzyme (1.22 µM and 2.22 µM respectively). However, this difference is not statistically significant. This data is different to what has been observed in the T5 exonuclease. The equivalent active site mutation (D155K) binds to DNA tighter than the wild type enzyme (K_d

= 7.7 nM and 13.4 nM respectively) (Zhang, 2012). It is possible that the mutation in the *Tb*FEN is more detrimental than that seen in the T5 enzyme. This is feasible as an alanine mutation of the equivalent amino acid in the human enzyme (D179A) does loose both substrate binding and catalytic activity (Shen *et al.*, 1997). Clearly, the D183K mutation is not as detrimental with regards to DNA binding abilities, but it could still be affecting the enzyme more than in the T5 version.

2.4 Conclusions and Perspectives

Currently, there is no satisfactory structure of any flap endonuclease with the DNA substrate present in the active site of the enzyme in the presence of divalent metal ions. A structure of the *Hs*FEN has been reported, but the cofactors present are Sm^{3+}. The scissile phosphodiester is also located 5.2 Å away from the cleaved phosphate in the *Hs*FEN: product complex solved in the same paper (Tsutakawa *et al.*, 2011). Similarly, the T4 RNase H enzyme was crystallised with the a forked DNA substrate, in the presence of Mg^{2+}, but the DNA was not located in the active site and the enzyme was disordered (Devos *et al.*, 2007). As a consequence, there is a clear need for a FEN: substrate complex with appropriate cofactors and with the scissile phosphodiester located within the enzyme active site. One aim of this chapter was to build on previous work on the T5 FEN, and determine if producing charge replacement mutations within the enzyme active site would assist in obtaining the elusive protein: substrate complex. The *Tb*FEN D183K variant is catalytically inert, and can bind to DNA in the presence of Ca^{2+}. In order to take this further, the binding ability of *Tb*FEN D183K should be tested in the presence of magnesium, and in the absence of any divalent metal cations, to observe these effects.

Attempts to crystallise the *Tb*FEN D183K ΔC mutant by Mr Jason Wilson (Department of Molecular Biology and Biotechnology, The University of Sheffield) have been successful, and a total of ten different DNA substrates have been co-crystallised. However in all cases, no threaded DNA is present (Jason Wilson, personal communication). We are still lacking the elusive FEN: substrate complex.

3 Targeting the *Tb*FEN *in vivo*

3.1 Introduction

The production of a catalytically inert variant provides an opportunity to analyse the effect of an inactive protein on the parasite *in vivo*. Over-expression of the wild type enzyme and the D183K variant was analysed. These were compared to RNAi FEN knock down clones. These recombinant clones were assessed by estimation of growth rate, RT-qPCR, flow cytometry and microscopy. This work highlights the potential of the *Tb*FEN as a therapeutic target.

3.2 Results

3.2.1 Construct Production

Genomic DNA from the *Trypanosoma brucei brucei* strain Lister 427 was provided by Dr Helen Price (University of Keele, UK). The FEN gene was isolated from the genomic DNA and sub-cloned into pUC19 (see Figure 3-1B). An internal *HindIII* site was mutated (AAG → AAA) by full plasmid PCR (Figure 3-1C). A 5' *HindIII* site was inserted, and used (alongside the 3' *BamHI* site) to sub-clone the gene into the p2Mc^{N} vector (provided by Dr Helen Price, originally a gift from D. Horn and S. Alsford). The D183K mutation was inserted by SOE-PCR prior to insertion into the p2Mc^{N} vector **(**Figure 3-2). This vector integrates into the genome of the blood stream form of the *T. brucei brucei* parasite (Frearson *et al.*, 2010).

The oligonucleotides used to isolate a fragment of the *Tb*FEN gene for RNA interference (RNAi) were designed using the online software tool TrypanoFAN: RNAit (http://trypanofan.path.cam.ac.uk/cgi-bin/rnait.cgi). The fragment was isolated by PCR (see Figure 3-3 A) and inserted into the p2T7^{ti} vector (originally from Dr Helen Price). The presence of the gene fragment was checked by restriction digest (see Figure 3-3 B) and confirmed by DNA sequencing (Core Genomics Facility, University of Sheffield Medical School).

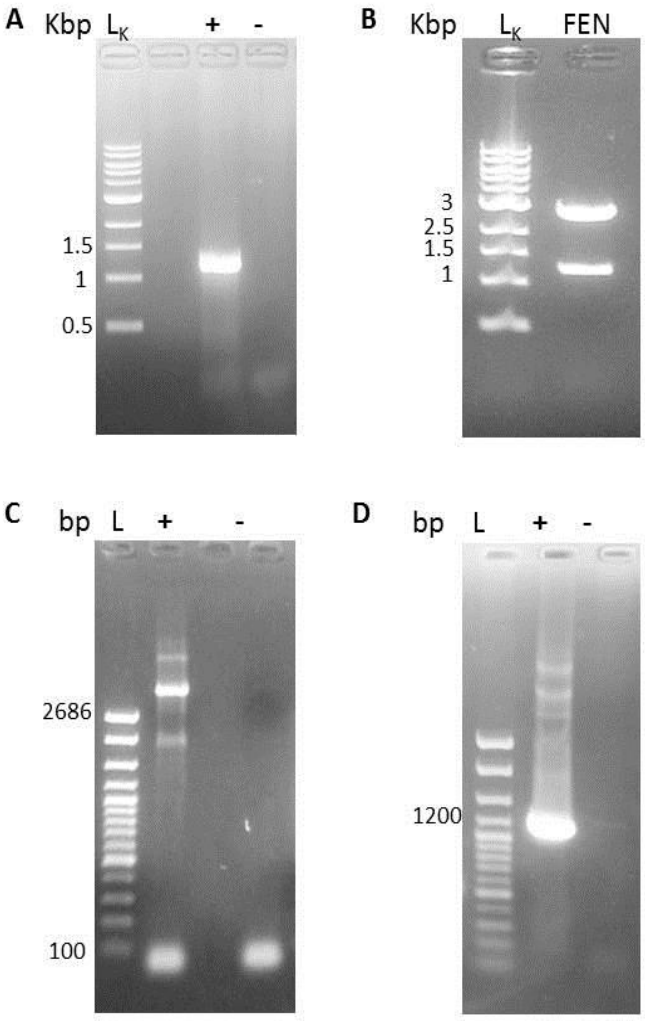


Figure 3-1 Isolation and sub-cloning of the genomic *Tb*FEN gene

(**A**) Isolation of the *Tb*FEN gene from genomic DNA by PCR. (**B**) The isolated FEN gene was sub-cloned into the pUC19 vector by restriction digest. (**C**) Site directed mutagenesis of the internal *HindIII* site (AAG → AAA) by full plasmid PCR. (**D**) An external 5' *HindIII* site was inserted to allow insertion of the gene into the p2Mc^N expression vector. Successful sub-cloning was monitored by gel electrophoresis, and confirmed by in-house DNA sequencing (Core Genomics Facility, The University of Sheffield). Annotations: L_K; 1 Kbp ladder (NEB), *L;* 100 bp marker (Norgen Biotek), **+**; PCR with genomic DNA, **–;** PCR without genomic DNA (negative control).

3.2.2 Effect of RNAi Knock-Down and FEN1 Over-Expression on Cell Growth

All constructs were transfected into *Trypanosoma brucei brucei* strain Lister 427 bloodstream form cells. Both over-expression and RNAi were induced by the addition of 1 µg/mL of tetracycline. Three clones were analysed for over-expression of both the native FEN (427::FEN1) and D183K mutant (427::fen1D183K). Two clones were analysed for the RNAi knockdowns (427‡FEN1). The parental control (427) was grown as a comparison. Results of the growth curve after 24, 48 and 72 hours post-induction are shown in Figure 3-4, and the statistical analysis is summarised in Table 3-1.

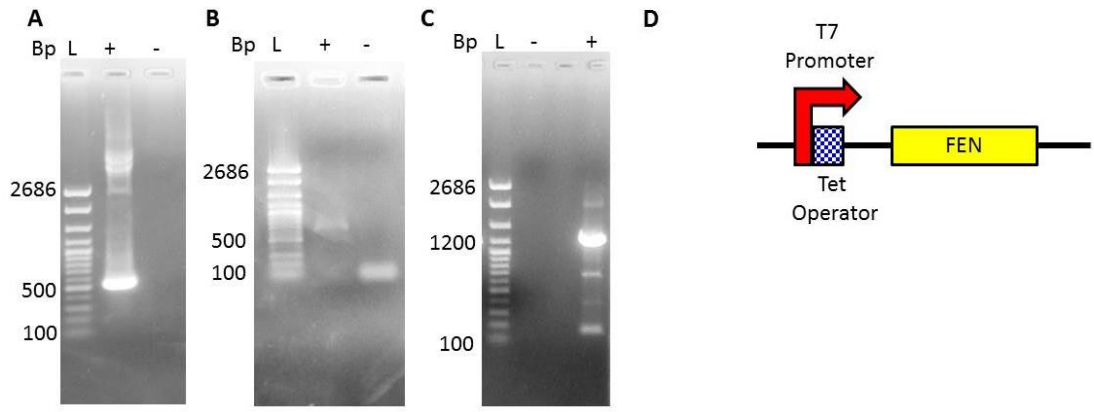


Figure 3-2 Production of the D183K (GAT → AAG) active site variant construct by SOE-PCR

(**A**) PCR using the oligonucleotides Forward_*Hind* and Reverse_D183K. (**B**) PCR using oligonucleotides Forward_D183K and Reverse_*Bam*. (**C**) The final D183K construct was produced using the oligonucleotides Forward_*Hind* and Reverse_*Bam*, and inserted into the p2Mc^{N} vector. Successful sub-cloning was monitored by gel electrophoresis, and confirmed by DNA sequencing (Core Genomics Facility, The University of Sheffield). (**D**) A schematic, showing part of the p2Mc^{N} vector and the arrangement of the promoter. Annotations: *L;* 100 bp DNA ladder (Norgen Biotek). **+;** PCR reaction containing the DNA template, **–;** PCR reaction without the DNA template (negative control).

Over-expression of the wild type FEN does not appear to affect logarithmic cell growth (Figure 3-4). However, statistical analysis of the untransformed data does indicate a statistically significant effect. After 72 hours, the induced cells grew to a cell density that was 29, 23 and 15% lower than their un-induced counterparts (427::FEN1#1, #2 and #3 respectively) (Figure 3-4). This was only statistically significant for 427::FEN1#1 and #2 (Table 3-1). After 48 hours, the induced cells grew to a cell density that was 29, 29 and 30% lower than their un-induced counterparts (427::FEN1#1, #2 and #3 respectively) (Figure 3-4). This was statistically significant for 427::FEN1#1 and #3 only (Table 3-1).

Only one RNAi clone displays any difference compared to the un-induced control: 427‡FEN1#1. Tetracycline induction caused the cells to grow to a cell density that was 45, 62 and 44% lower than the un-induced control at 24, 48 and 72 hours respectively. These differences were significantly different (Table 3-1). No significant difference was seen in the second RNAi clone (427‡FEN1#2; Figure 3-4).

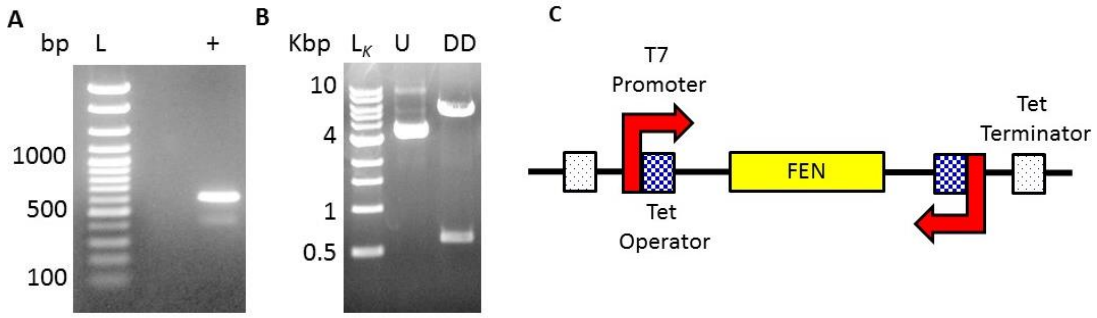


Figure 3-3 Production and sub-cloning the RNA interference construct into p2T7^{ti}

(**A**) A 597 bp fragment from the native *Tb*FEN gene was isolated using the oligonucleotides RNAi_*Xba* and RNAi_*Hind*. (**B**) The fragment was sub-cloned into the p2T7^{ti} vector. Successful sub-cloning was monitored by agarose gel electrophoresis of the uncut (U) and double digest (DD) of the p2T7-RNAi construct by XbaI and HindIII, and confirmed by DNA sequencing (Core Genomics Facility, The University of Sheffield). (**C**) A schematic showing part of the p2T7^{ti} vector, and the arrangement of the regulatory elements. Annotations: *D;* 100 bp DNA ladder (Norgen Biotek). **+;** PCR reaction containing the DNA template, D_K*;* 1 Kbp ladder (NEB).

The most striking difference was observed following over-expression of the D183K FEN variants. All three 427::fen1D183K clones were unable to reach the same (or similar) cell densities as their un-induced controls at 48 and 72 hours post induction. 427::fen1D183K#1 was also significantly lower at 24 hours post induction. 427::fen1D183K#3 appears to be growing slower in the absence of tetracycline than the other two 427::fen1D183K clones. The cell density of 427::fen1D183K#3 reached 241 ± 70.11 x10^4 cells/mL after 72 hours post induction, which was 34% lower than the 427 strain, and 61% and 46% lower than 427::fen1D183K#1 and 427::fen1D183K#2 respectively (Figure 3-4).

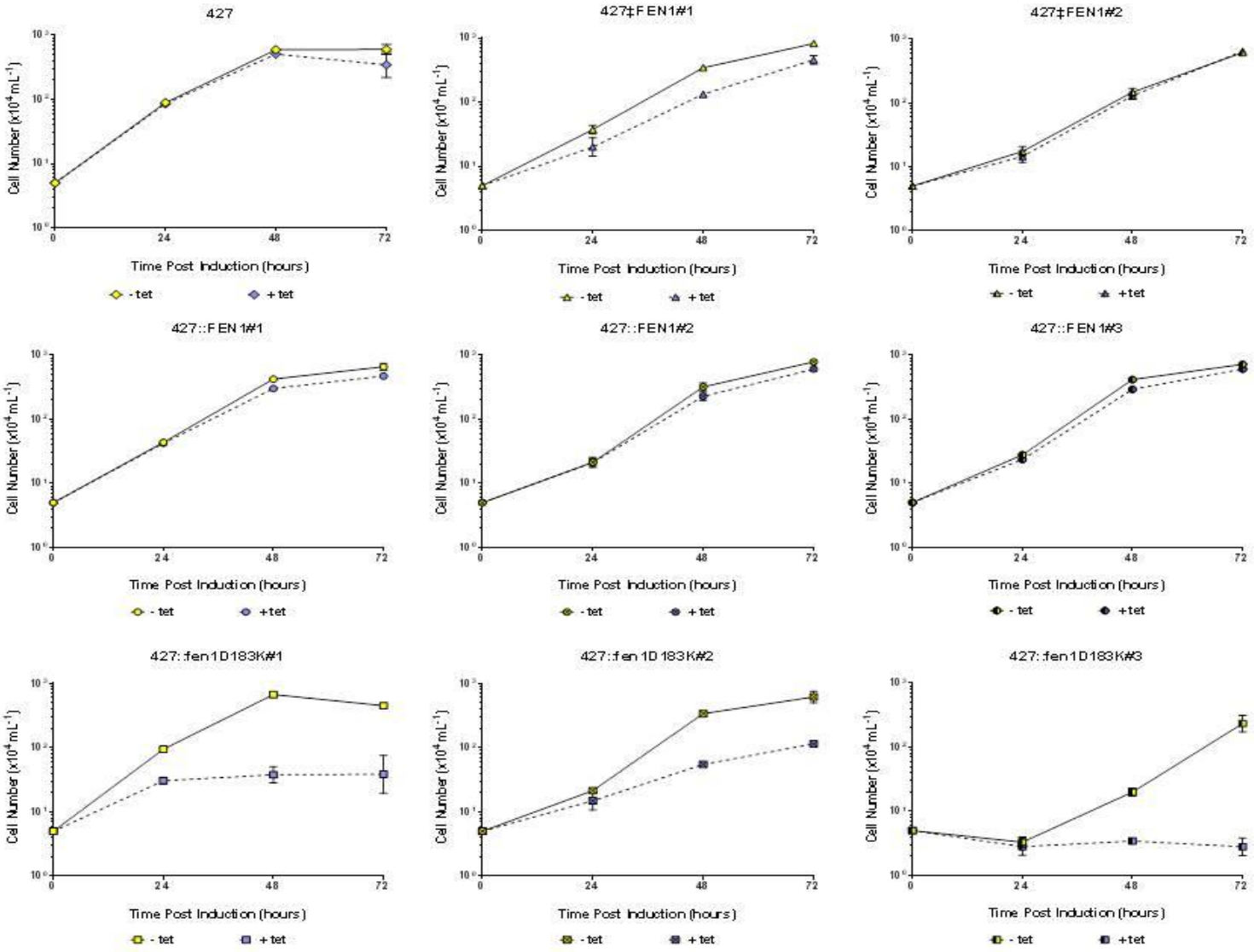


Figure 3-4 Cumulative growth analysis of the transgenic *T. brucei* BSF in the presence and absence of tetracycline

Over-expression/knock down of the FEN1 constructs (427::FEN1, 427::fen1D183K and 427‡FEN1) was induced by the addition of 1 μg/mL tetracycline. Cells were counted in the presence (lilac) and absence (yellow) of tetracycline 24, 48 and 72 hours post induction. Error bars show the standard deviation of the mean of three biological replicates.

Table 3-1 Statistical analysis of the cumulative growth curves

Transgenic Strain	Clone	Time (hours post induction)	Significance	P value
427::FEN1	1	24	NS	0.4294
		48	***	0.0006
		72	*	0.0252
	2	24	NS	0.9380
		48	NS	0.0510
		72	*	0.0347
	3	24	NS	0.0949
		48	**	0.0061
		72	NS	0.0677
427::fen1D183K	1	24	****	6.0×10^{-5}
		48	****	5.3×10^{-5}
		72	****	3.8×10^{-5}
	2	24	NS	0.1284
		48	****	8.7×10^{-5}
		72	***	0.0031
	3	24	NS	0.4902
		48	***	0.0007
		72	**	0.0042
427‡FEN1	1	24	*	0.0239
		48	***	0.0003
		72	*	0.0023
	2	24	NS	0.3414
		48	NS	0.3001
		72	NS	0.4837

Counts in the presence of tetracycline are compared to counts in the absence of tetracycline using an unpaired T test on data prior to logarithmic transformation. Statistical significance was determined using the Holm-Sidak method in GraphPad Prism (6.0).

Quantitative PCR was used as an indicator of the efficiency of the RNAi knockdown and overexpression of the proteins, with α tubulin used as the constitutively expressed control. Oligonucleotides were designed as described by Thornton and Basu (Thornton & Basu, 2011). Quantitative PCR was performed using 40 ng cDNA as a template, and results were normalised against the control. Small, consistent levels of non-specific amplification were present in the 'no template controls' of the FEN1 primers only. Samples were normalised against this to give accurate copy number values.

In the 427 parental controls, the relative level of FEN1 mRNA decreases after 24 and 48 hours post induction compared to the un-induced sample. At 72 hours post induction, the levels increase by 69% (see Figure 3-5 D). In order to compare the transgenic clones with the parental control (427), a repeated-measures mixed ANOVA was performed in IBM SPSS 2.2. The results are summarised in Table 3-2. Briefly, post hoc tests were performed on the over-expressing clones only (427::fen1D183K#1, #2 and #3, and 427::FEN1#1, #2 and #3). This is because these displayed a statistically significant interaction, where $F(3.367, 8.978) = 44.704$, $p \leq 0.0001$ for the 427::fen1D183K clones, and $F(3.234, 8.625) = 6.186$, $p = 0.015$ for the 427::FEN1 clones. The detailed statistical analysis is described in Appendix B.

The 427::FEN1 clones display increased levels of FEN1 mRNA relative to the parental control (see Figure 3-5 B). Clones 1 and 3 follow the same pattern, specifically the FEN1 specific mRNA levels peak at 24 hours post induction, and then decrease by 62% and 46% respectively at 48 hours. Levels in clone 2 peak at 48 hours post induction, and then decreases by 24% by 72 hours post induction. However, these differences were only statistically significant for 427::FEN1#2 at 24 hours post induction, and for 427::FEN1#3 at 48 hours post induction.

The 427::fen1D183K clones all display an increase in FEN1 specific mRNA levels that follows the same pattern, specifically levels peak at 24 hours post induction, then decrease (Figure 3-5). These difference was statistically significant (when compared to the parental control) for all clones at all time points, except 427::fen1D183K#1 at both 0 and 72 hours post induction. Interestingly 427::fen1D183K#2 and #3 both display a significant increase in FEN1 specific mRNA levels when un-induced.

The 427‡FEN1#1 clone displays a slight increase of FEN1 specific mRNA levels compared to the parental control, peaking at 48 hour post induction (see Figure 3-5 C). After 24 hours post induction, the average relative FEN1 mRNA level is half that of the endogenous sample

at the same time point (0.001 and 0.002 respectively). The 427‡FEN1#2 clone does display higher FEN1 levels initially, when compared to the endogenous control. These peak at 24 hours post induction, then decrease by 64% to levels similar to the endogenous control (0.065 and 0.071 respectively) (see Figure 3-5 C).

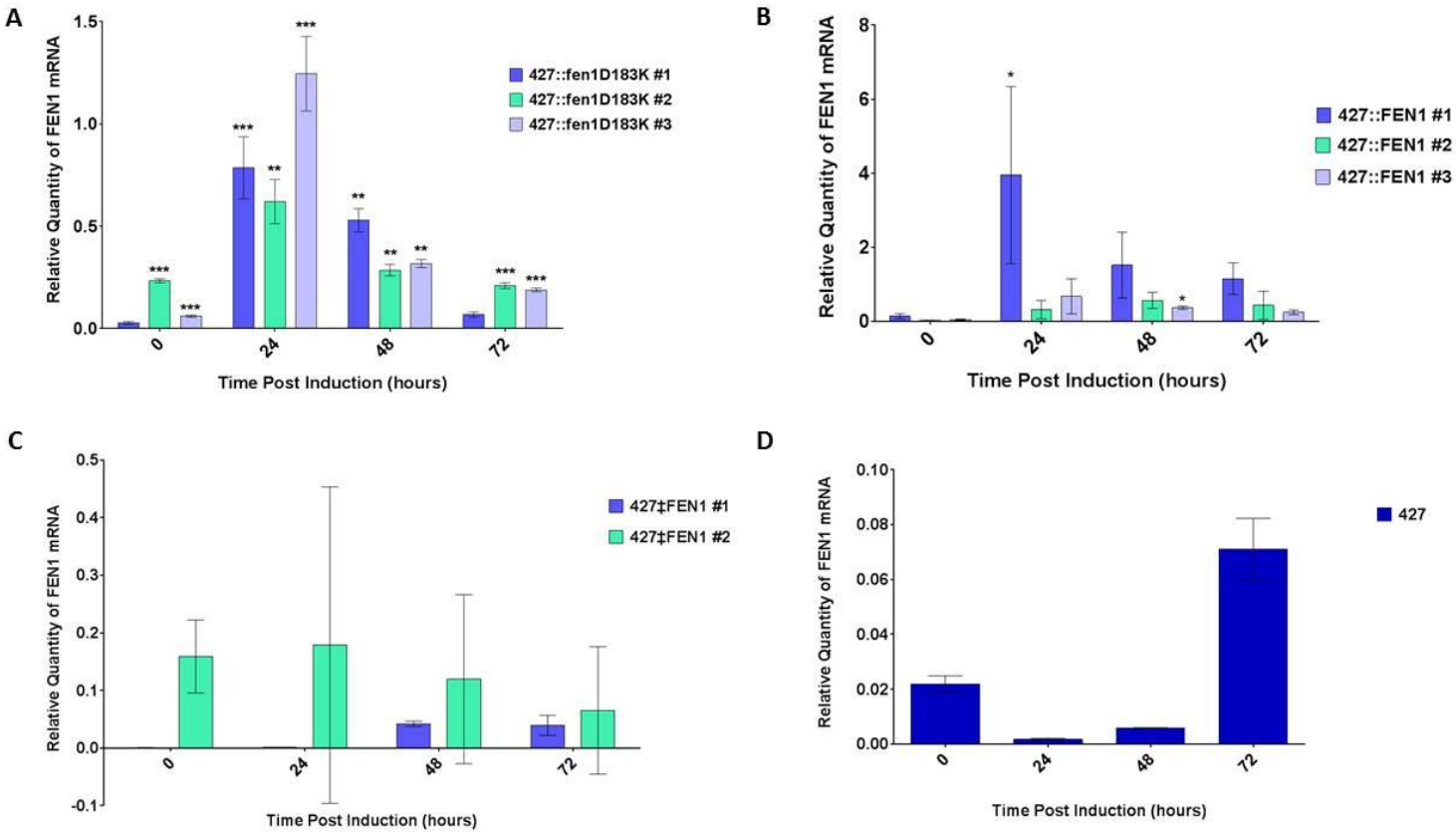


Figure 3-5 The effect of tetracycline on FEN1 specific mRNA levels

FEN1 specific mRNA was measured by RT-qPCR and normalised by RNA quantity, and against a constitutively expressed control. Mean normalised FEN1 specific mRNA levels for 427::fen1D183K clones (**A**), 427::FEN1 clones (**B**) and 427‡FEN1 clones (**C**) are shown. Statistical significance of FEN1 specific mRNA levels, compared to the 427 parental controls (**D**), is shown as stars. Error bars indicate the standard deviation from the mean of three technical replicates.

Table 3-2 Post hoc analysis of the effect of tetracycline on FEN1 specific mRNA levels in the over-expression system

Transgenic Strain	Clone #	Time (hours post induction)	Significance	P value
427::FEN1	1	0	NS	0.246
		24	*	0.022
		48	NS	0.232
		72	NS	0.117
	2	0	NS	0.478
		24	NS	0.362
		48	NS	0.114
		72	NS	0.514
	3	0	NS	0.437
		24	NS	0.304
		48	*	0.012
		72	NS	0.091
427::fen1D183K	1	0	NS	0.708
		24	***	0.0005
		48	**	0.01
		72	NS	0.988
	2	0	***	0.0005
		24	**	0.002
		48	**	0.008
		72	***	0.0005
	3	0	***	0.001
		24	***	0.0005
		48	**	0.003
		72	***	0.0005

FEN1 specific mRNA levels were normalised against a constitutively expressed control and compared to the 427 strain. Values were analysed for statistical significance using a repeated measures mixed ANOVA (IBM SPSS statistics 22).

3.2.3 Transgenic Clones have no effect on Cell Cycle Progression

Propidium iodide was used to stain the DNA of the transgenic *T. brucei* cells to determine whether over-expression or knock down of FEN1 had any effect on cell cycle progression. Figure 3-6 to 3-8 all display the same pattern, specifically two distinct peaks. The first, larger peak corresponds to the G1 phase of the cell cycle. This is followed by a second peak which corresponds to G2/M phase of the cell cycle. No obvious differences are evident between the samples.

3.2.4 D183K Over-Expressing Clones Display Altered Cell Morphology after 72 Hours

Immunofluorescence microscopy of all clones at 0, 24, 48 and 72 hours post induction was performed against FEN1 (green) and TAT1 (red; donated by Dr Helen Price, originally a gift from Professor Keith Gull (Sir William Dunn School of Pathology, University of Oxford, UK)), as shown in Figure 3-9 to 3-11. The parental control (427) displays a faint FEN1 signal at all time points, but it is difficult to distinguish. The 427‡FEN1#1 clone does display FEN1 expression in the un-induced sample, but it is hard to distinguish any in the other three time points. With 427‡FEN1#2, FEN1 expression can be seen, but faintly. No obvious morphological changes are apparent in any of the 427‡FEN1 clones.

The 427::FEN1 clones show increased levels of the FEN1 protein compared to both the un-induced samples, and the parental control (see Figure 3-10). The FEN1 protein is localised in an oval section that appears to correspond to the nucleus. The 427::fen1D183K clones again show increased levels of FEN1 protein at 24, 48 and 72 hours post induction when compared to both the un-induced samples and the parental control. Cell morphology remains relatively similar until 72 hours post induction. Cells still display high levels of FEN1, but now appear rounded.

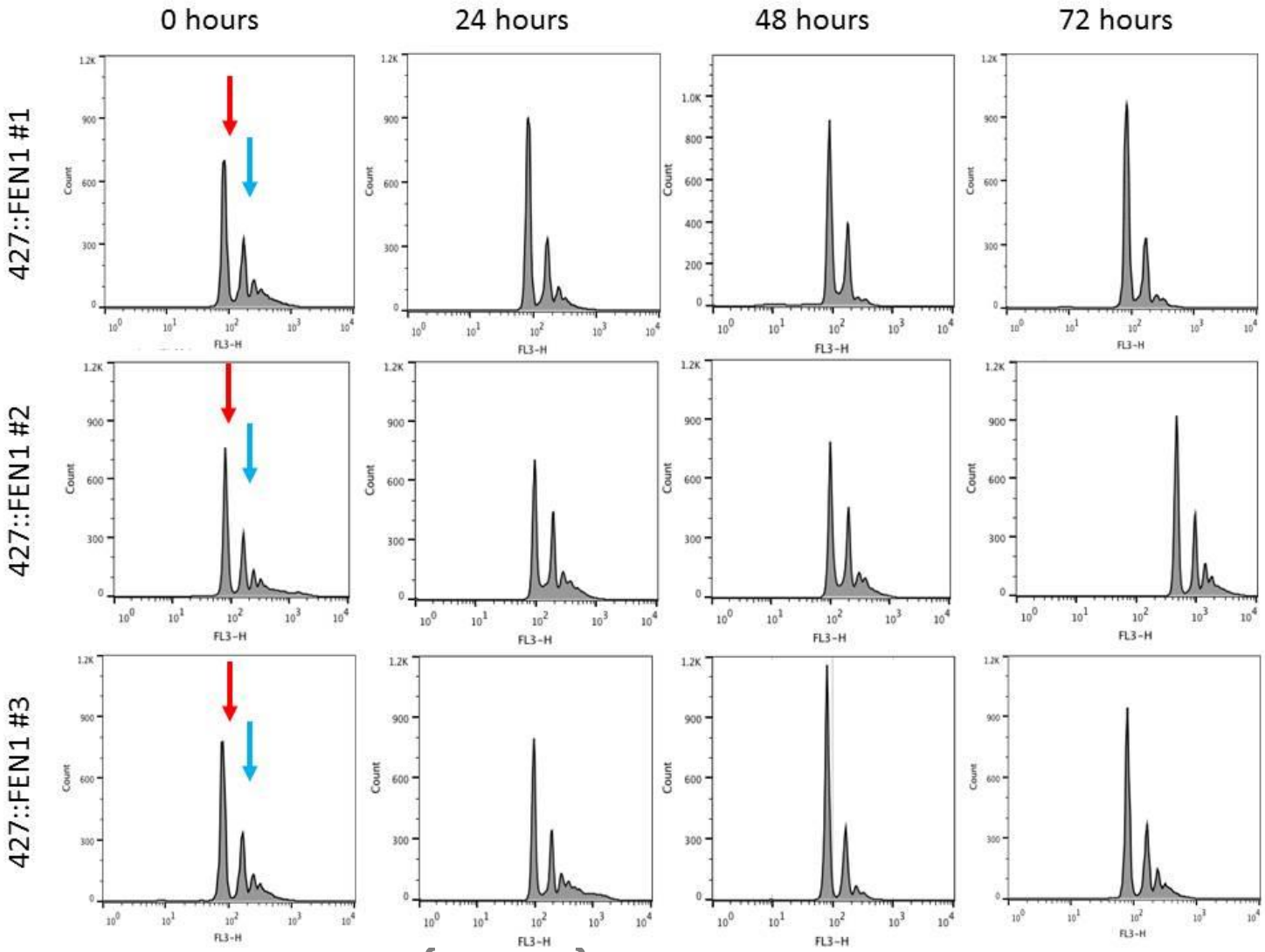


Figure 3-6 Cell cycle progression of the 427::FEN1 clones, measured using propidium iodide staining

427::FEN1 cells were measured by flow cytometry. The first peak (indicated at 0 hours, with the red arrow), represents G1, whilst the second peak (indicated at 0 hours with the blue arrow) represents G2/M phase of the cell cycle.

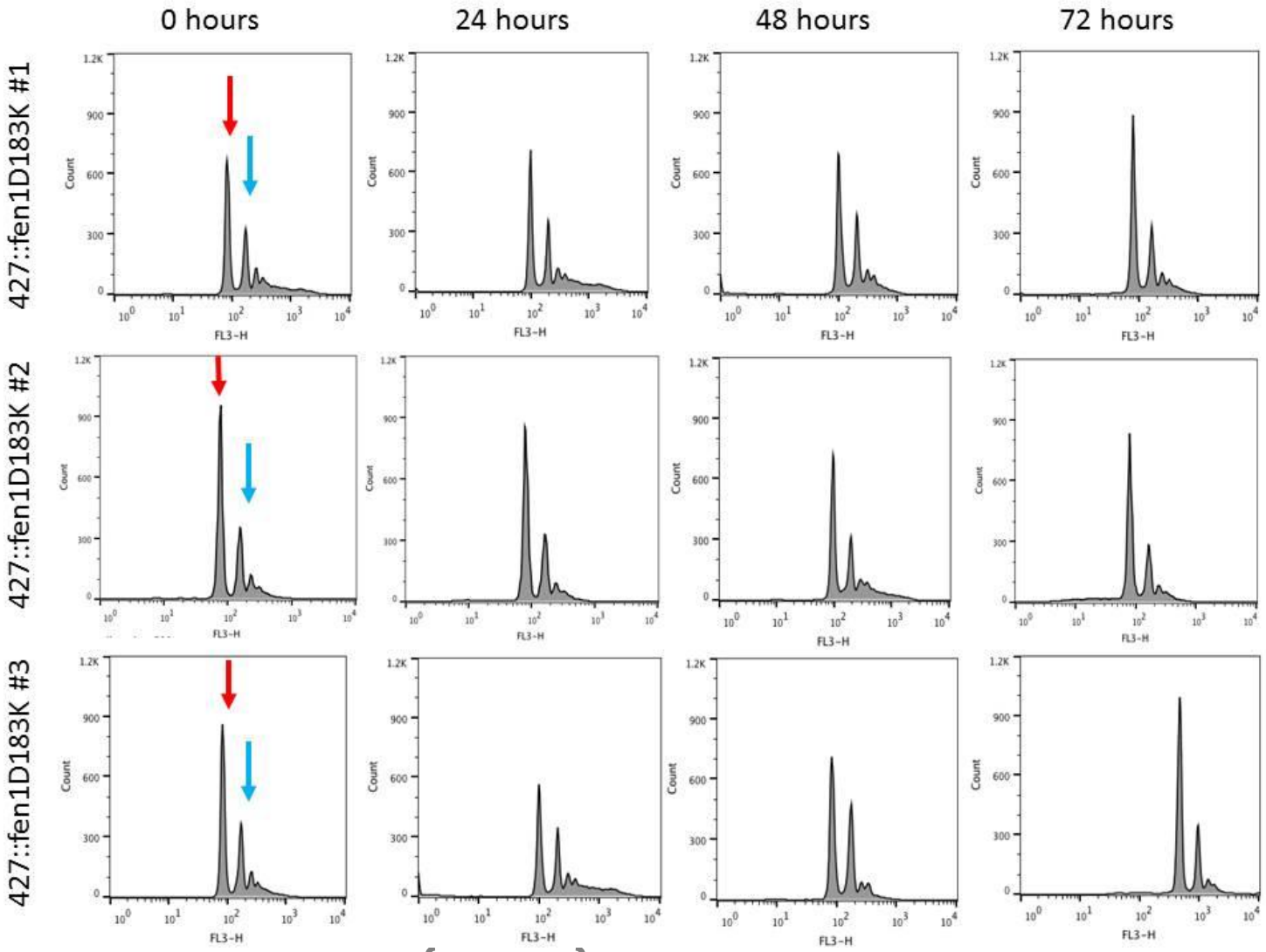


Figure 3-7 Cell cycle progression of the 427::fen1D183K clones, measured using propidium iodide staining

427::fen1D183K cells were measured by flow cytometry. The first peak (indicated at 0 hours, with the red arrow), represents G1, whilst the second peak (indicated at 0 hours with the blue arrow) represents G2/M phase of the cell cycle.

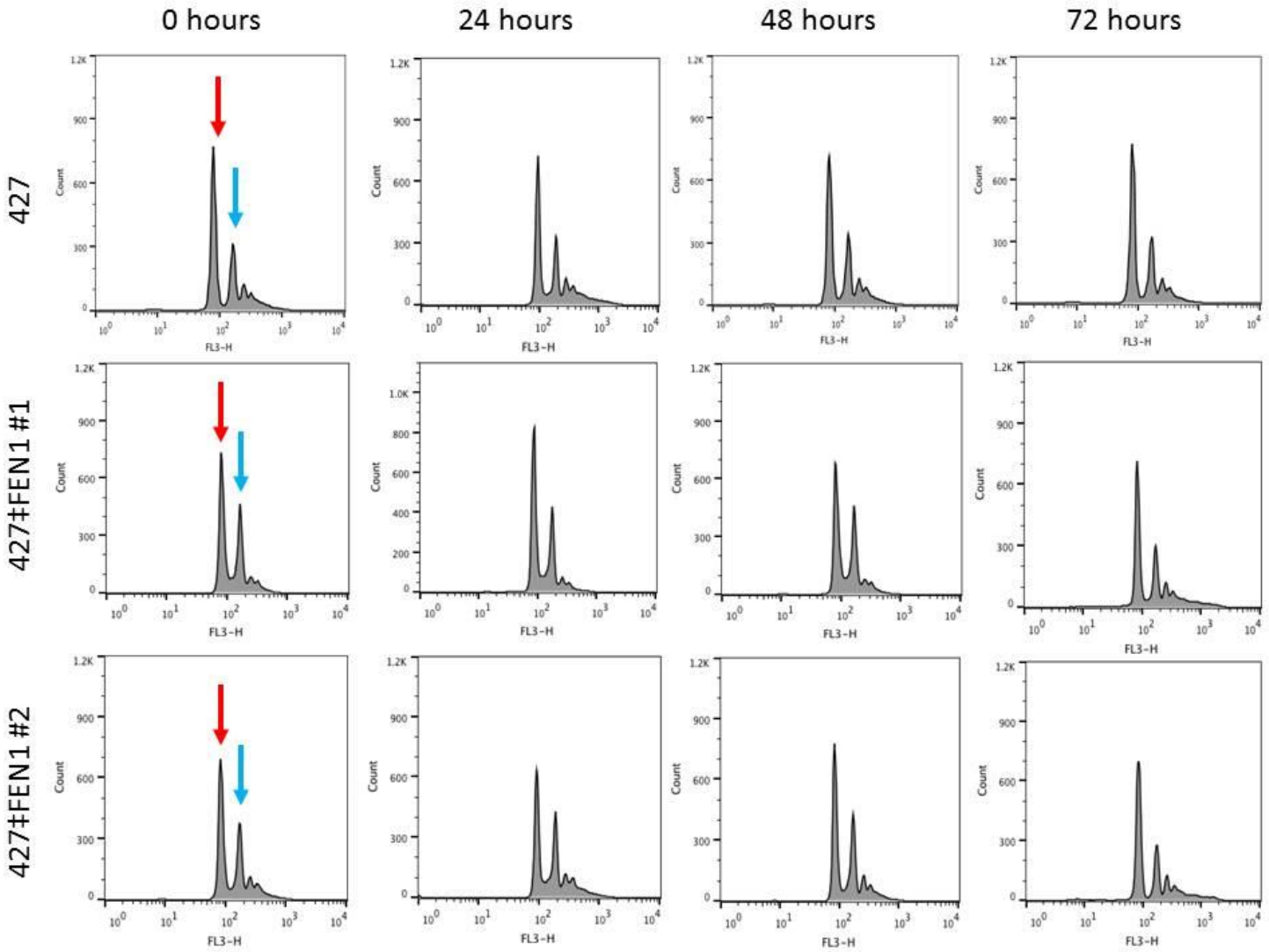


Figure 3-8 Cell cycle progression of the parental control (427) and the 427‡FEN1 clones, measured using propidium iodide staining

427 or 427‡FEN1 cells were measured by flow cytometry. The first peak (indicated at 0 hours, with the red arrow), represents G1, whilst the second peak (indicated at 0 hours with the blue arrow) represents G2/M phase of the cell cycle.

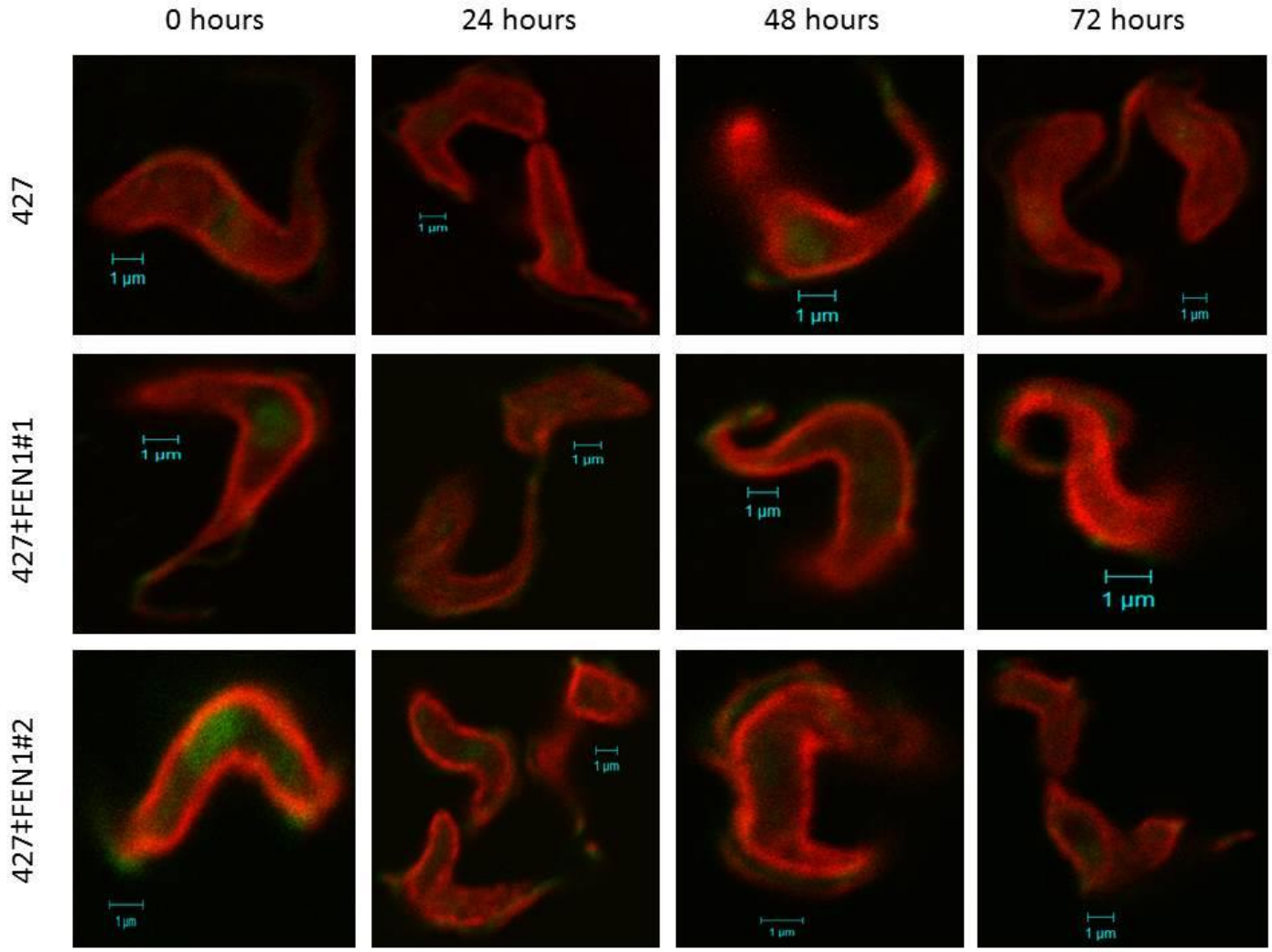


Figure 3-9 Immunofluorescence analysis of the parental control (427) and the two 427‡FEN1 clones
Formaldehyde fixed cells were treated with rabbit anti-FEN1 and mouse anti-TAT1 primary antibodies. Images show merged fluorescence, displaying the FEN1 (green) and TAT1 (red) proteins at 0, 24, 48 and 72 hours post induction. Scale bars represent 1 µm.

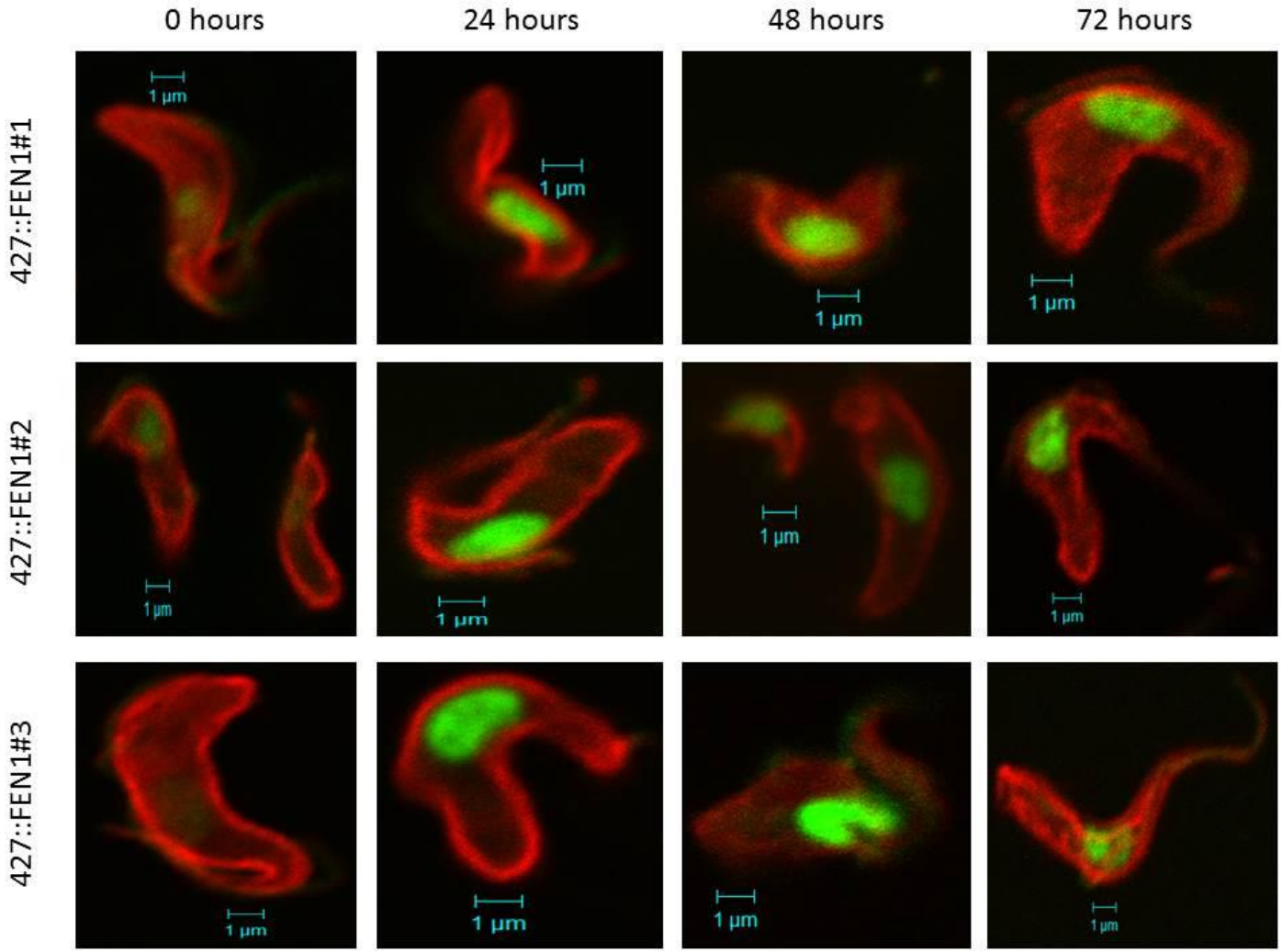


Figure 3-10 Immunofluorescence analysis of the three 427::FEN1 clones

Formaldehyde fixed cells were treated with rabbit anti-FEN1 and mouse anti-TAT1 primary antibodies. Images show merged fluorescence, displaying the FEN1 (green) and TAT1 (red) proteins at 0, 24, 48 and 72 hours post induction. Scale bars represent 1 µm.

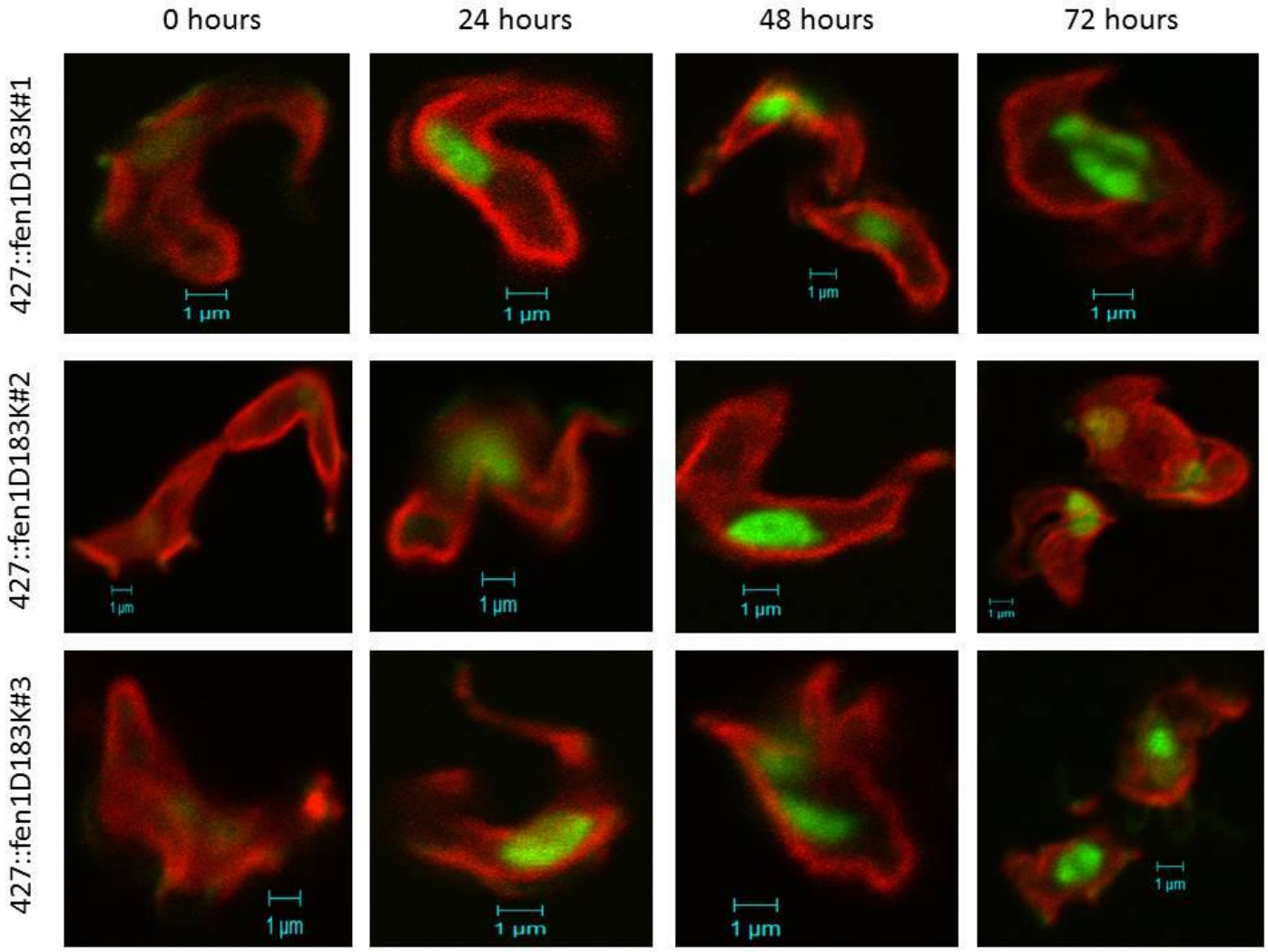


Figure 3-11 Immunofluorescence analysis of the three 427::fen1D183K clones

Formaldehyde fixed cells were treated with rabbit anti-FEN1 and mouse anti-TAT1 primary antibodies. Images show merged fluorescence, displaying the FEN1 (green) and TAT1 (red) proteins at 0, 24, 48 and 72 hours post induction. Scale bars represent 1 µm.

3.3 Discussion

3.3.1 Construct Production

Flap endonucleases play a number of vital roles within the cell, specifically in DNA repair and replication (Balakrishnan & Bambara, 2013; Pascucci *et al.*, 2011; Zheng & Shen, 2011). This, alongside the conserved nature of the enzyme, indicates that it is has an essential role within the cell. Essential enzymes have the potential to be therapeutic targets, providing that interfering with the native protein's function results in a lethal, or detrimental, phenotype.

In order to assess the suitability of the *Tb*FEN as a potential drug target, three constructs were produced from the genomic DNA of the parental strain *T. brucei brucei* Lister 427**.** The wild type gene, and D183K variant, was inserted into the pM2c^{N} vector. This vector allowed the overexpression of constructs either as the native molecules or with an N-terminal Myc tag. The latter approach could not be used here because the N-terminus of the FEN protein is located within the active site of the enzyme. The presence of a tag would therefore disrupt the activity of the enzyme, and could lead to misleading results. Whilst a similar vector that inserts a C-terminal tag was also available (p2Mc^{C}), this was not used because of the importance of the C-terminus in FEN1 protein: protein interactions (Section 1.2.4: (Chen *et al.*, 1996; Gary *et al.*, 1997; Hasan *et al.*, 2001; Warbrick *et al.*, 1997).

The second plasmid, p2T7^{ti}, contains two tetracycline inducible promoters at opposite sides of the DNA fragment (LaCount *et al.*, 2000). This allows both sense and antisense RNA to be generated simultaneously. These anneal to form double-stranded RNA, which activates the RNAi pathway (LaCount *et al.*, 2000). This leads to post transcriptional gene silencing, thus preventing translation of the gene of interest.

3.3.2 Effect of RNAi Knock-Down and FEN1 Over-Expression on Cell Growth

Growth analysis was performed for all constructs in the presence and absence of tetracycline. Over-expression of the wild type protein has an effect on cell growth in all clones at 72 hours post-induction, compared to the un-induced samples (Figure 3-4). In order to assess if this is due to an increase in FEN1 levels, specific mRNA levels were measured by RT-PCR (see Figure 3-5). Samples were normalised against the constitutively

expressed alpha tubulin control. Tubulin has been shown to have relatively stable expression within the different life cycle stages (Brenndörfer & Boshart, 2010). It would have been preferential to use multiple reference genes, for example PFR (paraflagellar rod) or TERT (telomerase reverse transcriptase) alongside the tubulin. This would have decreased the probability of biased normalisation (Brenndörfer & Boshart, 2010). Whilst an increase in FEN1 mRNA levels was seen, this increase was only statistically significant at 24 hours (clone 1) and 48 hours (clone 3) when compared to the endogenous control.

Analysis of the FEN1 specific mRNA levels of the parental strain (427) showed a decrease after 24 and 48 hours, followed by a large increase at 72 hours, when compared to the un-induced control. This could be linked to the morphological change of slender forms transforming to stumpy forms, which occurs above a critical cell density (Reuner *et al.*, 1997). Both the Spliced Leader Trapping (SLT) Database (http://splicer.unibe.ch/cgi-bin/gbrowse/*Tb*rucei_chr3/) and TriTrypDB (http://tritrypdb.org/tritrypdb/) were searched for *Tb*FEN specific results (*Tb*927.3.830, chromosome 3, location 195086-193551). The databases yielded a number of datasets for expression of *Tb*FEN in the different morphological forms. A small increase between low density bloodstream forms (slender) and high density bloodstream forms (stumpy) was apparent, with a larger increase between bloodstream forms and procyclic forms (Queiroz *et al.*, 2009). Interestingly, Kabani *et al* saw the same increase between bloodstream forms and procyclic forms, but not between slender and stumpy bloodstream forms (Kabani *et al.*, 2009). The difference between these two papers is the culture method: Queiroz *et al* used parasites cultured *in vitro*, whilst Kabani *et al* used parasites grown in mice (Kabani *et al.*, 2009; Queiroz *et al.*, 2009).

The presence of the D183K mutation has a hugely detrimental effect on the cells, as they only grew to a maximum of 18% of the cell density of the un-induced (Figure 3-4). The third clone (427::fen1D183K#3) grew at a slower rate than the other two clones (Figure 3-4). However, it still appears to be in logarithmic growth phase at 72 hours post-induction. At least one additional time point (96 hours) should have been analysed to determine if the un-induced cells would ultimately reach stationary phase. The detrimental effect observed on cumulative cell growth caused by the addition of tetracycline correlates with an increase in FEN1 specific mRNA levels (Figure 3-5). A statistically significant increase in FEN1 specific mRNA levels for clones 1 and 2 in the un-induced samples compared to the same time point on the endogenous control was apparent. This indicated that there may be 'leaky' expression of the D183K FEN1 variant, resulting in a slow-growing phenotype (see Figure

3-4). In addition, all clones display a statistically significant increase in FEN1 specific mRNA levels when compared to the endogenous control at 24 and 48 hours post induction. The levels peak at 24 hours post induction. Clone 1 is the only clone that does not display a significant increase in FEN1 specific mRNA levels after 72 hours post induction. This may be caused by spontaneous mutation that yields non-expressing cells. This phenomenon of increased colony expansion following growth inhibition has been documented in RNAi clones, and would also be possible for over-expression clones (Motyka & Englund, 2004).

There was no obviously detrimental effect of the RNAi clones on cell growth (see Figure 3-4). One potential explanation is that the RNAi pathway was not activated, so the FEN1 protein was still being transcribed and expressed. The levels of FEN1-specific mRNA was analysed using RT-qPCR (see Figure 3-5). There was no statistically significant change in the level of FEN1 specific mRNA when compared to the endogenous control. This indicates that the RNAi knockdown was unsuccessful. However, the RT-qPCR for these two clones is not accurate. This is because, despite three separate attempts to harvest high quality RNA from these constructs, the RNA purity level was around 1.7 (A_{260}/A_{280}) compared to the purity of approximately 2.0 for the wild type and D183K constructs. Previous studies in yeast cells that knocked down Rad27 (the yeast FEN1 homologue) show that the yeast null mutants are viable, but are sensitive to temperature changes, and mutate spontaneously (Tishkoff *et al.*, 1997). This is due to the presence of a back-up protein, called EXO1. *T. brucei* does contain an EXO1 protein (XP_847153.1). Further analysis to determine if this synthetic lethality is present in *T. brucei* by using either RT-qPCR, or knock outs of both the *Tb*FEN and *Tb*EXO1, would further establish this relationship. Further analysis on the *Tb*FEN knock out is also needed to see if it, like the yeast version, displays increased temperature sensitivity and a mutator phenotype would be interesting. This does pose the question of whether increasing the mutation rate in a human pathogen is advisable. Whilst a mutator phenotype is likely to adversely affect the parasite, there is the possibility that the mutations may be advantageous. The creation of an advantageous phenotype for parasite survival within the host could have drastic consequences for both the infected individual and for HAT eradication.

Whilst the use of RT-qPCR is an excellent indicator of FEN1-specific mRNA levels, it provides no indication of what occurs at the post-transcriptional level. Whilst there is an increase in mRNA levels for the D183K variant, we have no quantitative data to show that this mRNA increase correlates with an increase in protein expression. Western blot analysis would

provide a quantitative measure of protein expression, and antibodies are available for the *Tb*FEN protein and the TAT1 constitutively expressed control. However, due to time constraints, this work could not be carried out. A qualitative indicator of FEN-specific over-expression can be seen in the confocal microscopy performed on these clones (Figure 3-9 to 3-11).

3.3.3 Transgenic Clones have no effect on Cell Cycle Progression

Due to the role of flap endonucleases in DNA replication and repair, cell cycle analysis was performed to observe the effects of these transgenic clones on cell cycle progression. Methanol-fixed cells were incubated with propidium iodide, a commonly utilised DNA binding dye. This is often used to analyse cell cycle defects, or to determine whether cells are apoptotic (Henry *et al.*, 2013). One initial peak indicates G1. Cells undergoing mitosis will have double the amount of DNA as cells in G1, resulting in a second peak (Henry *et al.*, 2013). Figure 3-6 to 3-8 show that there is no change in the profile for any of the clones when compared to the parental control. It is possible that the 72 hour time point is not long enough to allow full development of a phenotype, but given that affects can be seen on cell growth by 48 hours post induction, this is unlikely. Because the 427::fen1D183k clones are unable to reach the cell densities seen in the un-induced controls (see Figure 3-4), the idea that the cells may be undergoing apoptosis is possible. However, the flow cytometry results do not display a peak before G1. This would correspond to the generation of low molecular weight DNA fragments in apoptotic cells (Wlodkowic *et al.*, 2009). In order to assess this further, the use of specific early apoptotic markers such as annexin V, would yield definitive results, and should be explored in the future. Annexin V has been used as an apoptotic marker in kinetoplastids, with success (Bustos *et al.*, 2015; Grazú *et al.*, 2012; Mamani-Matsuda *et al.*, 2004).

3.3.4 D183K Over-Expressing Clones Display Altered Cell Morphology after 72 Hours

Morphological changes were analysed for each transgenic clone using immunofluorescence microscopy. TAT1 was used as a positive control because tubulin is constitutively expressed and covers the whole cell (Imboden *et al.*, 1987). This makes any morphological changes easier to visualise. The FEN1 protein is localised in an oval section that appears to correspond to the nucleus. This is expected because the *Hs*FEN has a recognisable nuclear

localisation signal (KRK-X_8-KKK) (Qiu *et al.*, 2001). However, the *Tb*FEN does not. This may be due to a lack of information available regarding nuclear localisation signals in trypanosomes. However, the lack of DNA staining prevents this localisation from being definitive.

On a qualitative level, FEN1 expression still appears high, whilst the FEN1 specific mRNA levels have decreased by 72 hours post induction. This may be due to the delay between transcription and translation. What is interesting is that it takes 72 hours for this morphological change to become apparent, but the over-expression affects cell growth after 24-48 hours. The D183K mutant is unable to cleave DNA, but is able to bind to the DNA. This may result in the accumulation of DNA damage that affects cell growth, and ultimately leads to cell death. Whilst the flow cytometry results show no accumulation of low molecular weight DNA, this tends to be a late stage apoptotic marker.

3.4 Conclusions and Perspectives

FEN1 has essential roles in DNA replications and repair. Over-expression of the wild type enzyme may therefore result in an increased level of DNA strand breaks. However, we are not seeing an obvious phenotype, so further analysis is essential. Assays such as TUNEL or Comet (Invitrogen) allow analysis of apoptotic DNA and DNA strand breaks respectively (Collins, 2004). It is possible that the over-expressed wild type FEN1 enzyme is not able to access DNA replication/repair sites because of another limiting factor. For example, the PCNA trimer acts as a processivity factor for *Hs*FEN1, and loads and unloads the enzyme at the correct place, at the correct time (Li *et al.*, 1995). It is possible that an increase in the wild type enzyme will only affect the cells when there is additional PCNA, for example if a DNA damage response has been activated. It would therefore be very interesting to look at the effect of the wild type over-expressing clones when exposed to DNA damaging agents such as UV radiation and methyl methanesulphonate (MMS).

The FEN1 enzyme has a number of binding partners. As a consequence, the detrimental effect seen by over-expression of the D183K variant could be caused by the excess D183K FEN1 preventing these proteins from binding to any other potential binding partners, and therefore affecting their function. For example, FEN1 interacts with the C-terminus of WRN, specifically between residues 949-1092 (Brosh *et al.*, 2001). It is the C-terminus of WRN that also interacts with DNA Pol β, specifically residues 949-1432 (Harrigan *et al.*, 2003). Product formation from strand displacement by DNA Pol β increases from 10% to 25% in the

presence of WRN (Harrigan *et al.*, 2003). Weak strand displacement, such as that seen in the absence of WRN, allows DNA Pol β to expand triplet repeats (Hartenstine *et al.*, 2002), which can lead to detrimental phenotypes and clinical disease (Ashley & Warren, 1995).

Over-expression of the D183K variant has a detrimental phenotype on the parasite, whilst the knock down clones do not. This may be due to insufficient knock down of the FEN1-specific RNA, and based on the data presented here, this cannot be ruled out. It may also be due to functional redundancy between FEN1 and EXO1, as seen in other cell types (Tishkoff *et al.*, 1997). The FEN1 protein may have a longer half-life, so even if the knock down has been successful, there may still be endogenous FEN1 persisting in the cells. The data for the D183K FEN1 construct does indicate that the *Tb*FEN protein has potential as a therapeutic target, but the RNAi knock down work needs to be repeated, and possibly for a longer time course. Clearly the presence of inactive FEN1 is detrimental to the cells, and is expressed at sufficient levels to swamp the endogenous FEN1. As a consequence, inhibiting the native *Tb*FEN seems likely to lead to parasite death. This work provides the first evidence that FEN1 in a medically relevant protozoan pathogen has the potential to be a therapeutic target.

4 Validating Previous FEN Inhibitors

4.1 Introduction

The previous chapter showed that the catalytically inert flap endonuclease D183K variant caused a detrimental phenotype in the *Trypanosoma brucei* bloodstream form cells. In this chapter, the wild type enzyme was tested against three previously identified flap endonuclease inhibitors, from both the Sayers laboratory and the literature. Two inhibitors described previously in the Sayers laboratory (Zhang, 2012), and myricetin, a commercially available flavonol with known anti-FEN properties, were selected (van Pel *et al.*, 2013). In order to assess the inhibitory effects of these compounds, a low through-put gel based assay (designed and optimised in-house) was used. The compounds were tested against both the *Trypanosoma brucei* and human flap endonucleases.

4.2 Results

4.2.1 Construct Production and Purification of the *Hs*FEN

A synthetic *Hs*FEN sequence, codon optimised for expression in *Escherichia coli*, was designed and supplied by Eurofins MWG Operon. This construct lacked the last 44 amino acid residues, and the aspartic acid codon for Asp 179 was mutated to a lysine (*Hs*FEN D179K ΔC). In order to produce the wild type enzyme this construct was cloned into pET21a+, and the codon corresponding to Lys 179 was reverted back to the original aspartic acid by splicing by overlap extension PCR (Materials and Methods). Synthetic DNA encoding the C-terminal 44 amino acids was inserted using an oligonucleotide cassette by standard recombination DNA technology, producing the *Hs*FEN wild type full-length construct (Figure 4-1). This full-length, native construct was sub-cloned into the lactose inducible pET21a+ vector (Figure 4-1), and over-expressed (Figure 4-2). The protein sequence was analysed using Expasy ProtParam, an online resource that deduces a number of physico-chemical properties from the protein sequence only (Gasteiger E., 2005). Knowledge of the theoretical pI (8.80) was used to assist with protein purification. The success of the purification procedure was monitored by Coomassie-stained SDS-PAGE (Figure 4-3). Zymogram analysis of the wild type enzyme was carried out to detect nuclease activity (Figure 4-3).

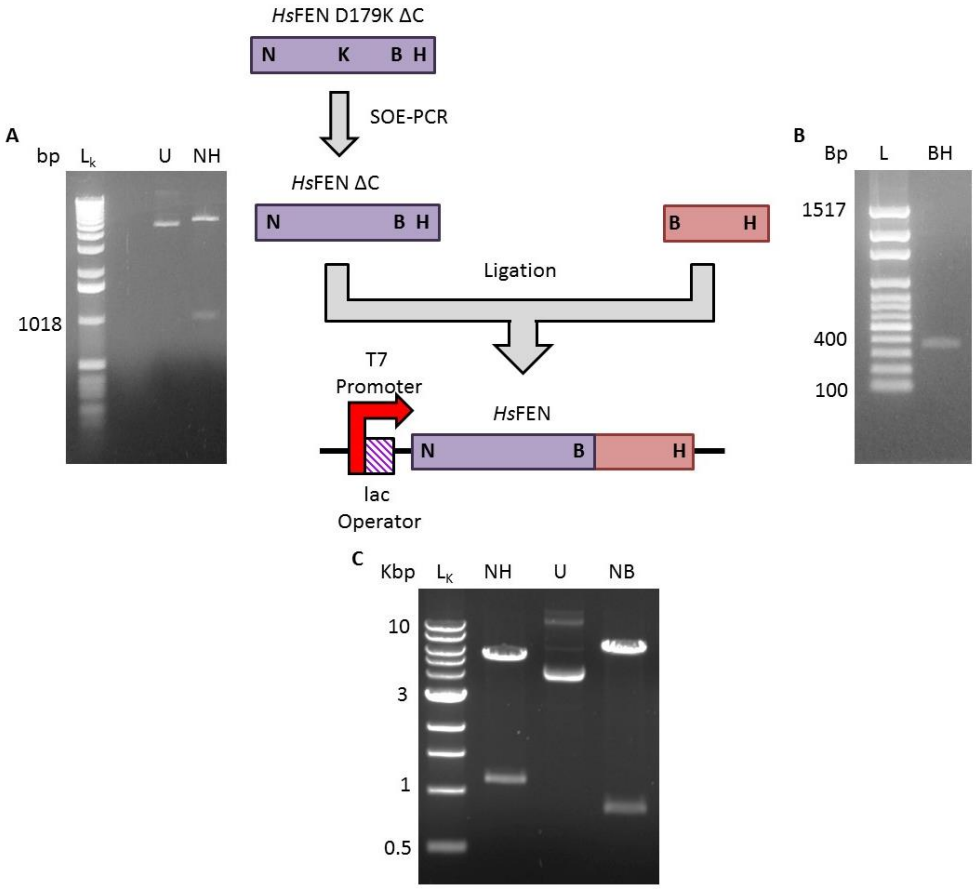


Figure 4-1 Producing the wild type full length *Hs*FEN construct

The lysine residue at 179 was reverted to its original wild type aspartic acid (AAA→GAT) by SOE-PCR. (**A**) The integrity of the wild type *Hs*FENΔC gene was analysed by restriction digest using *NdeI* and *HindIII* (NH). (**B**) The C-terminus of the *Hs*FEN protein was optimised for expression in *E. coli* and ordered from Eurofins. This was removed from its vector backbone by restriction digest using *BamHI* and *HindIII* (BH). The resulting cassette was purified and introduced to pET-*Hs*FEN ΔC, to produce the pET-*Hs*FEN wild type full length construct. (**C**) The success of the procedure was determined by restriction digest with *NdeI* and *HindIII* and *NdeI* and *BamHI* (NB). Each step was monitored by gel electrophoresis (shown), and success was confirmed by DNA sequencing (Medical School Core Genomics Facility). Annotations: **L_K**; 1 Kbp ladder (NEB), **L**; 100bp ladder (Norgen Biotek).

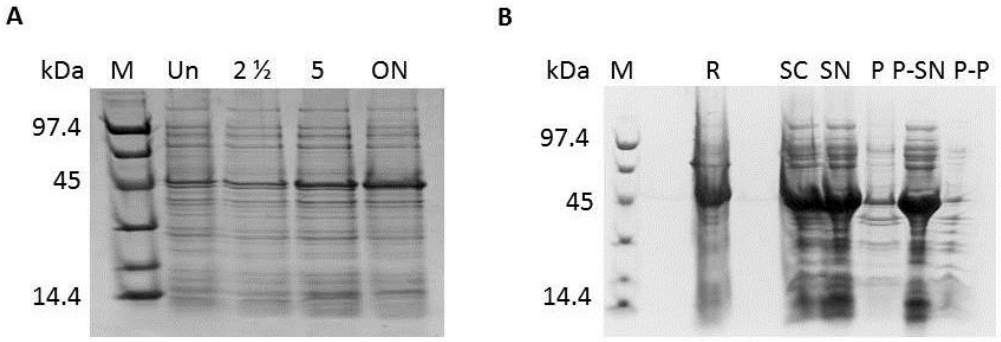


Figure 4-2 SDS-PAGE analysis of the over-expression and cell lysis of the recombinant wild type *Hs*FEN

(**A**) Analysis of the un-induced (Un) and induced cell pellets at three time points post-induction (2 ½ hours, 5 hours and overnight (ON)). Samples were normalised against A_{600}. (**B**) Analysis following lysis of the induced cell pellet, as described in the materials and methods. Annotations: **M**; low range protein marker (Bio-Rad), **+**; control of the empty pET21a+ vector in the same *E. coli* strain, **R**; sample following re-suspension of the cell pellet, **SC**; sample after sonication, **SN** and **P**; supernatant and pellet respectively, following centrifugation of the lysed, sonicated sample, **P-SN** and **P-P**; supernatant and the pellet respectively, from PEI precipitation and centrifugation step.

4.2.2 Solvent Effects and Compound Activity

Previous optimisation of this low-throughput assay showed that the *Hs*FEN cleaved the flap substrate at a lower rate than the *Tb*FEN, and an incubation period three times longer was required for substrate hydrolysis (data not shown). The effect of 5% DMSO (a commonly used solvent in screening protocols) on the rate of substrate cleavage was tested (Figure 4-4). Results were transformed using arcsine before being analysed by repeated measures T test (https://graphpad.com/support/faq/the-arcsin-function/; GraphPad Prism 6.0). Both enzymes displayed a significant increase in product formation in the presence of 5% DMSO. That increase was approximately 50% and 33% for the *Tb*FEN and *Hs*FEN respectively ($p = 0.0122$ and 0.0045 respectively).

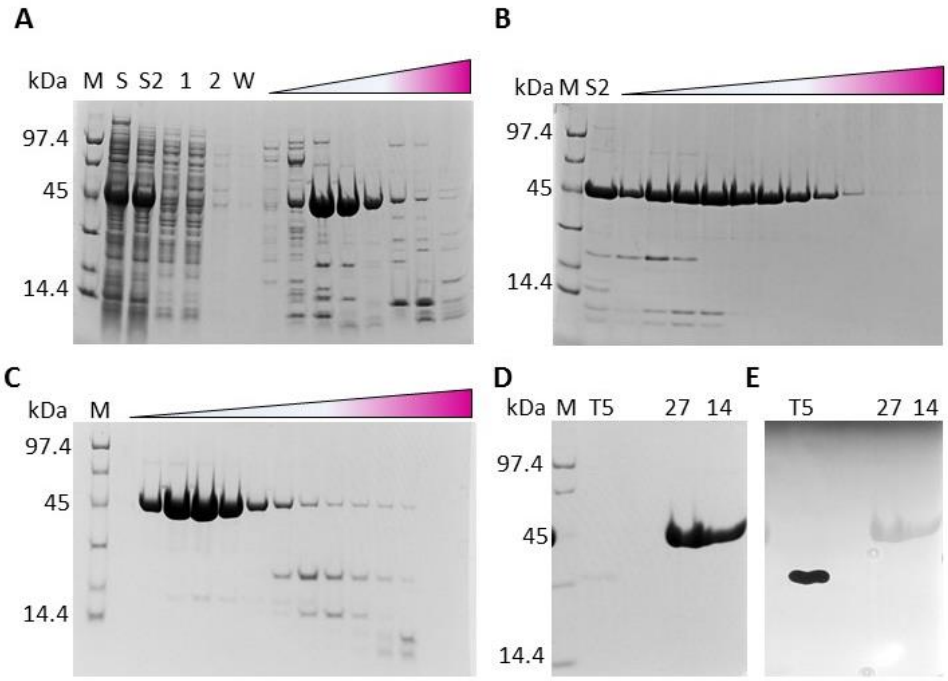


Figure 4-3 SDS-PAGE analysis of the purification of the recombinant, full length *Hs*FEN

Soluble *Hs*FEN was purified using a variety of chromatography techniques. (**A**) *Hs*FEN was eluted over a 0.05-0.75 M salt gradient from a Bio-Rad High S cation exchange chromatography column (pH 7.8). Fractions are shown in ascending order (white → pink). (**B**) Pooled fractions F4-6 were applied to a Bio-Rad High Q chromatography column coupled to a HiTrap heparin HP affinity chromatography column (pH 7.8). Fractions were eluted over a 0.05-0.75 M salt gradient from the heparin column only, and are shown in ascending order (white → pink). (**C**) Pooled fractions 11-22 were concentrated and further purified by size exclusion chromatography (GE Healthcare Superdex 200 10/300 GL, pH 6.7). Purified *Hs*FEN was concentrated and analysed by zymogram and stained with both Coomassie blue (**D**) and ethidium bromide (**E**). Annotations**: M;** low range protein marker (Bio-Rad), **S;** sample from the dialysis bag, **S2;** supernatant from the dialysis bag following centrifugation and filtration to remove precipitation, **1-6;** flow through samples collected during column loading, **FT;** flow through collected during column loading, **W;** column wash with the loading buffer, **T5;** T5Δ192CA positive control. **27/14;** amount (in µg) of pure *Hs*FEN loaded per well.

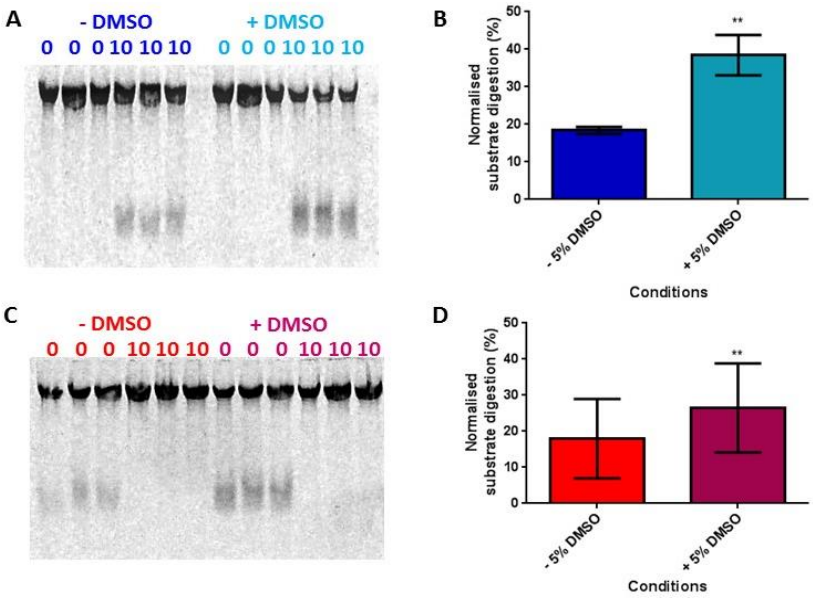


Figure 4-4 DMSO stimulates cleavage in both the *Tb*FEN (top) and *Hs*FEN (bottom) enzymes

The raw gel image is shown for the two enzymes (**A** and **C** for *Tb*FEN and *Hs*FEN respectively). The upper band corresponds to undigested substrate, whilst the lower band corresponds to digested substrate. Densitometry results were converted to percentage substrate digestion, and are shown for both the *Tb*FEN (**B**) and *Hs*FEN (**D**). Error bars indicate standard deviation from three biological replicates. Values were transformed by arcsine before statistical analysis with a repeated measures T test (GraphPad Prism 6.0).

C01 and D11 were previously identified as FEN inhibitors (Zhang, 2012). C01 selectively inhibits the T5 bacteriophage FEN, whilst D11 inhibits both the T5 and T7 bacteriophage FENs (Zhang, 2012). These were tested alongside myricetin, a commercially available *Hs*FEN inhibitor (van Pel *et al.*, 2013). The structures of the three compounds are shown in Figure 4-5. Inhibitors were tested at an initial concentration of 50 µM, and compared against a 5% DMSO control. Results for both enzymes are shown in Figure 4-6. Results were transformed to arcsine before being analysed for statistical significance using a One-Way Repeated Measures ANOVA (https://graphpad.com/support/faq/the-arcsin-function/; GraphPad Prism 6.0). Myricetin causes a statistically significant decrease in product formation ($p = 0.0003$ and 0.0009 for *Hs*FEN and *Tb*FEN respectively). D11 causes a statistically significant decrease in product formation ($p = 0.009$ and 0.0010 for *Hs*FEN and *Tb*FEN respectively).

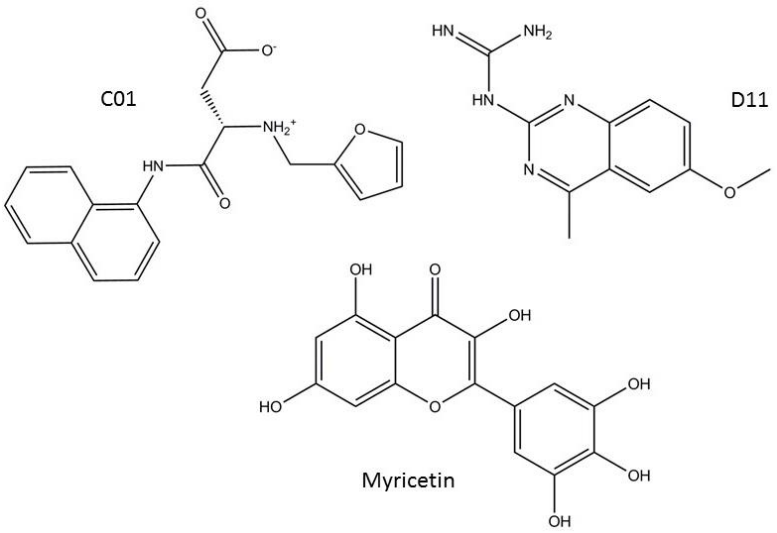


Figure 4-5 Previously identified FEN inhibitors.

D11 inhibits both the T5 and T7 bacteriophage FEN, whilst C01 inhibits the T5 FEN only (Zhang, 2012). Myricetin inhibits the *Hs*FEN (van Pel *et al.*, 2013). Structures were rendered using Chemdraw 14.

The lanes that contain the D11 compound show three bands (Figure 4-6). One correlates to the undigested substrate, whilst the second corresponds to the cleavage product. However, a third band located above the undigested substrate was also present. This low mobility band was also seen even when no DNA or protein was added, indicating that the compound is fluorescent (data not shown). Further analysis on the D11 compound was performed to assess its ability to interact with DNA. The absorbance profile of the compound was measured in the presence and absence of high molecular weight DNA (see Figure 4-7). The presence of DNA shifts the absorbance spectra by 10 nm to the red (Figure 4-7 C). This was compared to two molecules known to intercalate with DNA: ethidium bromide (Figure 4-7A) and methylene blue (Figure 4-7 B). The presence of DNA shifts the absorbance spectra of ethidium bromide by 40 nm to the red and methylene blue by 10 nm to the red.

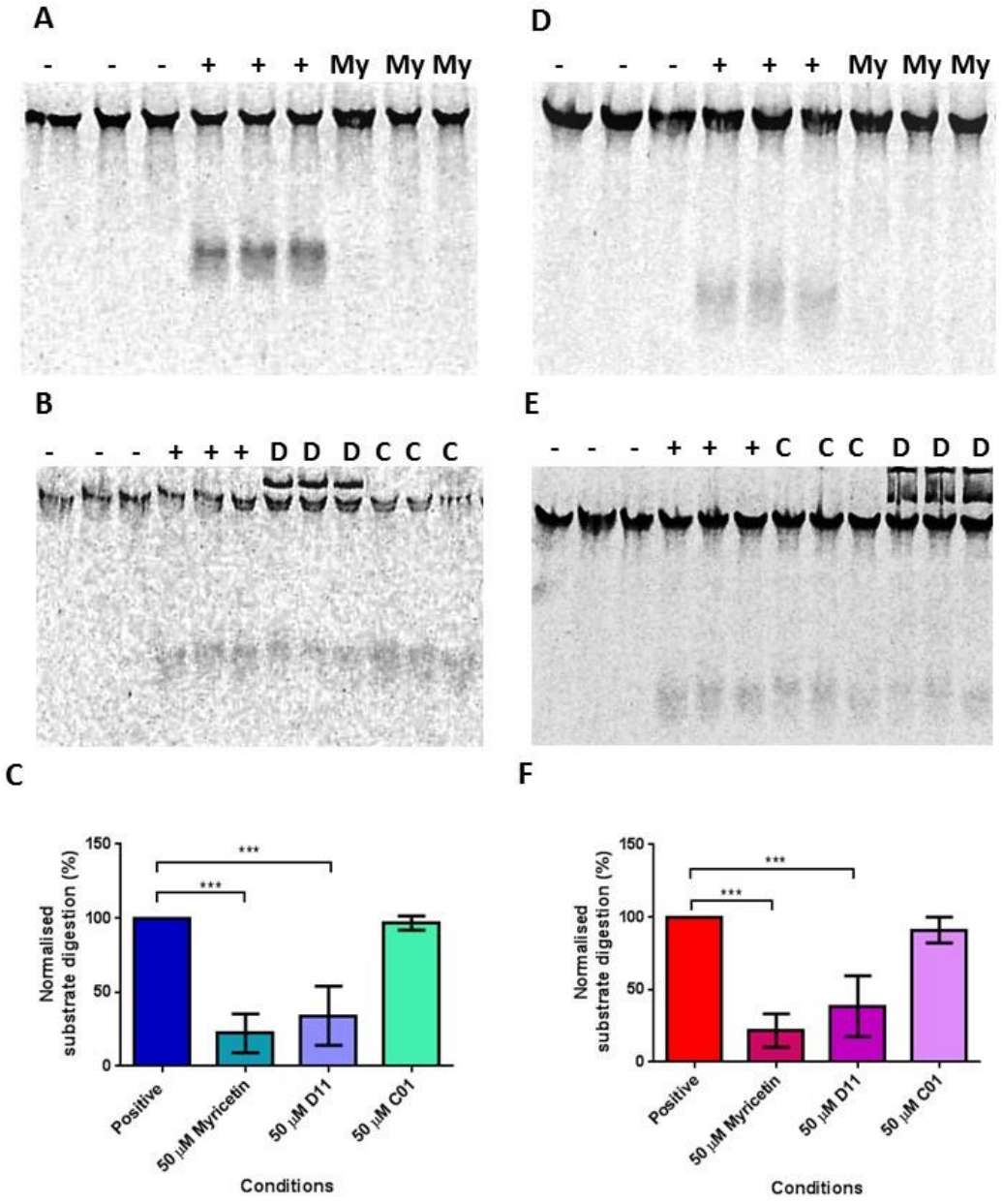


Figure 4-6 Analysis of the inhibitory effect of D11, C01 and myricetin on *Tb*FEN (left) and *Hs*FEN (right)

The effect of 50 μM of each compound on DNA cleavage by 10 μg/mL of protein was compared to the DNA substrate alone (-), and in the presence of 10 μg/mL protein and 5% DMSO (+). (**A, D**) The raw gel image showing the effect of myricetin (My) on the *Tb*FEN and *Hs*FEN enzymes respectively. (**B, E**) The raw gel image showing the effect of D11 (D) and C01 (C) on the activity of the *Tb*FEN and *Hs*FEN enzymes respectively. (**C, F**) Normalised substrate digestion of the DNA substrate by the *Tb*FEN and *Hs*FEN enzymes respectively. Error bars indicate standard deviation from three biological replicates. Values were transformed by arcsine before statistical analysis using a One Way Repeated Measures ANOVA (GraphPad Prism 6.0). Statistical significance is indicated by stars.

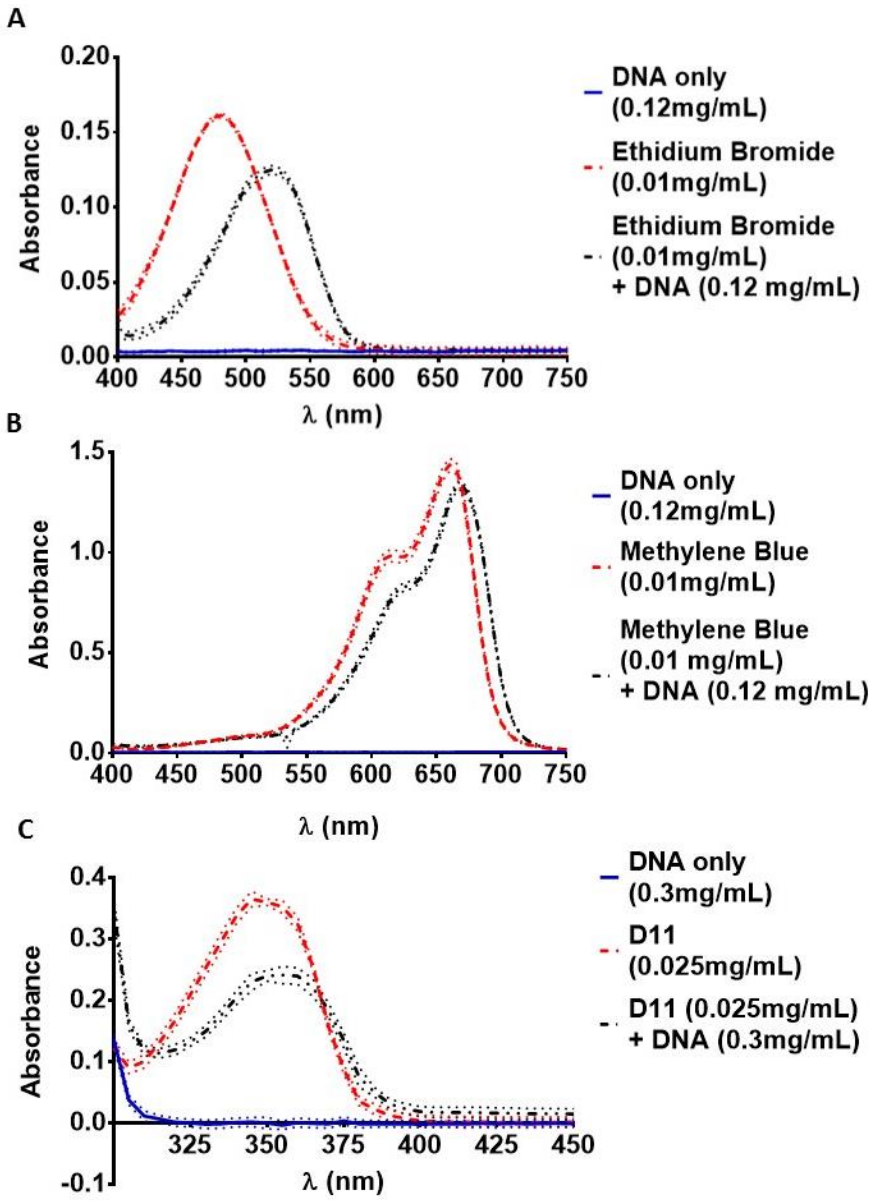


Figure 4-7 Analysis of the DNA intercalation properties of compound D11

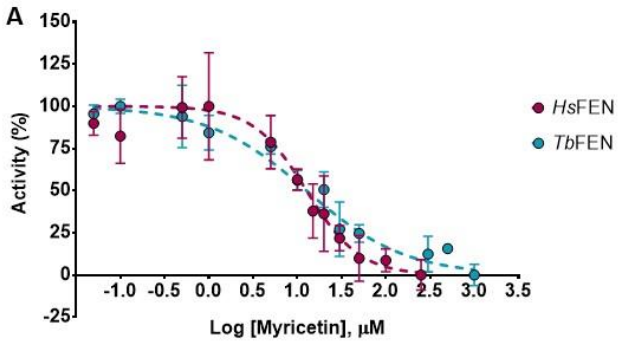
A wavelength scan of the absorbance of each compound in the presence/absence of high molecular weight herring sperm DNA (type XIV; Sigma). (**A**) The presence of DNA with ethidium bromide increased the λ_{max} by approximately 40 nm, from 480 nm to 520 nm. (**B**) The presence of DNA with methylene blue increased the λ_{max} by approximately 10 nm, from 660 nm to 670 nm. (**C**) The presence of DNA with D11 increased the λ_{max} by approximately 10 nm, from 345 to 355 nm. Error bars are depicted as dotted lines, and indicate the standard deviation from the mean of three replicates.

4.2.3 Effect of Myricetin *in vitro* and *in vivo*

The ability of myricetin to inhibit *Hs*FEN is consistent with the previous literature (van Pel *et al.*, 2013). The ability of myricetin to inhibit the *Tb*FEN had not been reported previously, and was explored further. A dose response was performed against both enzymes to determine the IC_{50} (the concentration required to inhibit enzymatic activity by 50%). The IC_{50} for myricetin against the *Tb*FEN was determined as 14.81 µM, compared to 12.03 µM for the human enzyme (see Figure 4-8).

An attempt to identify binding sites for myricetin in both the *Tb*FEN and *Hs*FEN enzyme was performed using SwissDock (Grosdidier *et al.*, 2011a; Grosdidier *et al.*, 2011b). Myricetin (ZINC ID = zinc03874317) was docked onto the *Hs*FEN structure (PDB 3Q8L; without DNA), and the Phyre2-predicted *Tb*FEN structure described in Chapter Two. The top-ranked binding sites for both proteins are depicted in Figure 4-8 and described in Table 4-1. The myricetin binding site on the *Tb*FEN enzyme is located close to the active site of the protein (Figure 4-8 B). The myricetin binding site on the *Hs*FEN enzyme is located towards the back of the enzyme, near the helix-two turn-helix motif.

The next step was to determine whether this inhibition *in vitro* corresponds to parasite cell death. A standard PrestoBlue assay was used to assess cell viability in the presence of different myricetin concentrations compared to a DMSO control (see Figure 4-9). A maximum concentration of 0.5% DMSO could be used before cell viability was significantly affected. A dose response curve was produced to determine the LD_{50} (the dose at which 50% of the parasites are dead). The LD_{50} for myricetin was 23 µM for *T. brucei* blood stream form cells.



B

Protein	IC_{50} ± SE (µM)	R^2
*Hs*FEN	12.03 ± 1.14	0.8553
*Tb*FEN	14.81 ± 1.14	0.9278

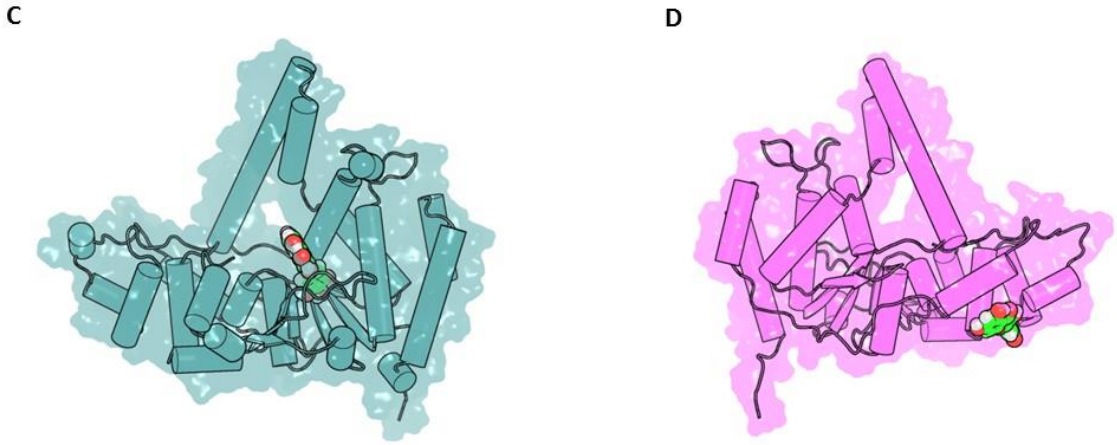


Figure 4-8 Dose dependent inhibition of *Tb*FEN and *Hs*FEN by myricetin

(**A**) Myricetin inhibits both *Tb*FEN and *Hs*FEN in a dose dependent manner. (**B**) IC_{50} values are not significantly different between proteins (p = 0.3189), as determined by non-linear regression (GraphPad Prism 6.0). Molecular docking of myricetin (ZINC ID = zinc03874317) to the Phyre2-predicted structure of *Tb*FEN (**C**) and to a known *Hs*FEN structure (PDB 38QL; **D**) show that the top hit is located at different sites for the two enzymes.

Table 4-1 The top binding sites of myricetin to the two eukaryotic FENs

Protein	Cluster	Cluster Rank	Full Fitness (kcal/mol)	ΔG (kcal/mol)
*Hs*FEN	0	0	-2369.38	-7.35
*Tb*FEN	0	0	-3000.00	-7.46

The Full Fitness and ΔG values for the most favourably ranked complex (Cluster Rank = 0) within the highest ranked cluster (Cluster 0) are shown.

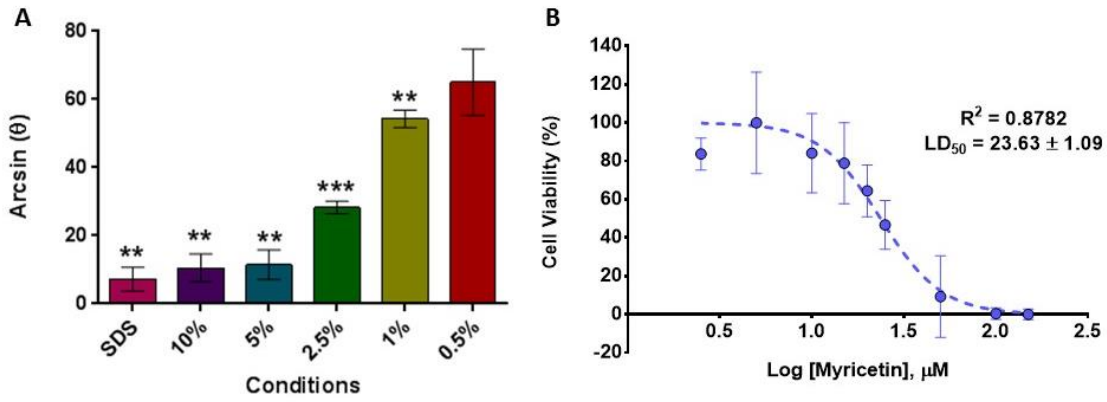


Figure 4-9 Myricetin demonstrates dose dependent killing of the *T. brucei* parasite

The resazurin-based PrestoBlue viability assay was used to determine the effect of myricetin on the *T. brucei* parasite. (**A**) The effect of different DMSO concentrations on cell viability was initially determined. The percentage change in fluorescence compared to a 'cells only control' was converted to arcsine for statistical analysis. A repeated measures one-way ANOVA, with a Dunnet's post-hoc test were performed, and show only a DMSO concentration of 0.5% does not significantly affect cell viability ($p = 0.1203$). (**B**) The percentage cell death was calculated for nine myricetin concentrations. The LD_{50} was calculated by non-linear regression in GraphPad 6.0. The mean ± standard deviation of three biological replicates is shown.

4.3 Discussion

4.3.1 Construct Production and Purification of the *Hs*FEN

The *Hs*FEN D179K ΔC construct contains the equivalent active site substitution to the *Tb*FEN D183K variant described in Chapter Two. This construct, which is missing the C-terminal 44 amino acids that constitute the flexible PCNA binding domain, was originally designed for crystallography work based on the published crystal structure from the Tainer laboratory (Tsutakawa *et al.*, 2011). The lysine was reverted back to the original aspartic acid to create the wild type *Hs*FEN ΔC construct. The *E. coli* codon-optimised C-terminal domain was used in conjunction with the *Hs*FEN ΔC construct to produce the full-length *Hs*FEN1 in the pET21a+ vector. The C-terminus is important in protein-protein interactions and in substrate binding, so creating the full length construct was essential for any drug discovery work (Stucki *et al.*, 2001).

4.3.2 Solvent Effects and Compound Activity

DMSO is a dipolar aprotic molecule commonly used as a solvent (Santos *et al.*, 2003). It is efficient at dissolving water-insoluble compounds, and is considered the solvent of choice for use in drug screens (Santos *et al.*, 2003; Tjernberg *et al.*, 2006). However, DMSO can have huge impacts on proteins, even at the low levels commonly used in drug screening. These effects include changes in binding properties, disordering of protein structures and decrease in thermal stability (Batista *et al.*, 2013; Jackson & Mantsch, 1991; Tjernberg *et al.*, 2006). Knowing how DMSO affects both FEN proteins was crucial to ensure accurate interpretation of results from screening trials. As shown in Figure 4-4, the presence of 5% DMSO significantly increases the activity of both enzymes. Studies on the effect of DMSO on the thermal stability of the enzymes has been analysed briefly in Chapter Five, and indicates that 5% DMSO decreases the thermal stability. This decrease in stability could allow the enzyme to bind to the DNA substrate more easily, resulting in a reduced K_M and/or an increased K_{cat}. Further work needs to be performed to understand how DMSO is affecting both the structure and function of the enzymes. This could include analysing the far-UV circular dichroism spectra of the enzymes in the presence of different DMSO concentrations, and performing a full catalytic screen to obtain the rate constants. For the purpose of this work, the effect of any compounds dissolved in DMSO must be compared to a protein-DMSO control, for accuracy.

Further analysis of the compounds previously identified in the Sayers laboratory has uncovered a false-positive result. Compound D11 is intrinsically fluorescent, and displays DNA intercalation properties. The DNA intercalation assay was tested on two compounds that are known to intercalate with DNA: ethidium bromide and methylene blue. The presence of DNA caused a shift in the λ_{max} of the compounds by 40 nm and 10 nm respectively. These values are consistent with those previously reported in the literature (Hossain & Suresh Kumar, 2009; Waring, 1965). False positives are a major concern in screening methods (Thorne *et al.*, 2010), and this highlights the importance of additional screening steps following the original hit identification. Whilst compound C01 showed promising results in the T5 FEN screen (Zhang, 2012), it is unable to inhibit the activity of either the *Tb*FEN or *Hs*FEN enzymes. Despite this, the fact that it can selectively inhibit the T5 enzyme over the T7 enzyme is still encouraging (Zhang, 2012), and provides optimism that selective inhibition between FEN enzymes can be achieved.

4.3.3 Effect of Myricetin *in vitro* and *in vivo*

SwissDock uses the EADock docking algorithm to identify and rank docking sites for a ligand on a particular protein (Grosdidier *et al.*, 2007). The full fitness function evaluates clusters of complexes, which correspond to binding modes (Grosdidier *et al.*, 2007). The most favourably ranked complex (with a Cluster Rank of 0) is taken as the centre of the first cluster, then the second most favourably ranked complex is chosen for the second cluster, and so on (Grosdidier *et al.*, 2007). A negative ΔG value indicates that the reaction is favourable, or spontaneous, and does not require any external energy input. The Cluster Rank 0 for myricetin binding to the *Tb*FEN is in a different location to that on the *Hs*FEN enzyme. Whilst these binding models are only indicative of the actual binding site, they do imply that despite the *Tb*FEN and *Hs*FEN enzymes having a sequence identity of 51%, there may be slight structural variations that may impact compound interactions. Selective inhibition of FENs has been achieved between species, as shown with compound C01 selectively inhibiting T5 FEN over T7 FEN (Zhang, 2012). It has also been achieved between FEN family members in the same species, as shown with *Hs*FEN and *Hs*XPG (Tumey *et al.*, 2005).

The IC_{50} of myricetin against *Tb*FEN *in vitro* is high at 14 μM. This is slightly higher than that obtained for the *Hs*FEN enzyme using the same assay, but the difference was not significant ($p = 0.3189$). Previously obtained IC_{50} values for *Hs*FEN and myricetin were considerably lower, specifically 0.09 μM (van Pel *et al.*, 2013). This is likely due to assay variation, as van

Pel *et al* used a sensitive fluorescence based spectrophotometric assay. One of the benefits of using the less sensitive gel-based assay is that any abnormalities such as compound fluorescence can readily detected (Figure 4-6).

The toxic effect of DMSO on *T. brucei* blood stream form cells has been documented previously (Sykes & Avery, 2009). The results shown here, where a maximum of 0.5% DMSO can be used, is consistent with those reported (Sykes & Avery, 2009). Testing myricetin on bloodstream form cells showed dose dependent killing of the *T. brucei* parasite. This result is not surprising given the trypanocidal activities of other flavonoids, such as quercetin (Dodson *et al.*, 2011; Grecco *et al.*, 2012; Mamani-Matsuda *et al.*, 2004). Because the IC_{50} of both the *Hs*FEN and *Tb*FEN enzymes is so similar, it is reasonable to assume that myricetin would target and kill human cells as efficiently. Whilst time constraints prevented this from being investigated fully, reviewing the literature has shown some interesting results. Myricetin has been extensively studied in mammalian cell lines (and in mammals themselves), because it displays a number of beneficial properties. These include antioxidant (Wang *et al.*, 2010; Zhang *et al.*, 2011), antimicrobial (reviewed in (Cushnie & Lamb, 2005)), anticancer (reviewed in (Devi *et al.*, 2015)), antiviral (reviewed in (Cushnie & Lamb, 2005)) and neuroprotective effects (Ma *et al.*, 2015; Ma *et al.*, 2007). During these investigations, mice were fed up to 50 mg/kg of myricetin with no ill effects (Ma *et al.*, 2015). This corresponds to approximately 2.7 mM of myricetin (based on a mouse weighing 25 g, with a total blood volume of 1.46 mL) (NC3Rs, 2016).

Because myricetin has such a wide range of beneficial properties, it raises concerns that the results seen against the *T. brucei* parasite are not due to myricetin: *Tb*FEN interactions. A recent review assessed the molecular mechanisms underlying the anticancer effects of myricetin (Devi *et al.*, 2015). Myricetin appears to target a number of proteins, including Akt, VGEF and Hif-I (Devi *et al.*, 2015). This wide range of protein targets may be due to the ability of myricetin to act as a chelating agent (Fernandez *et al.*, 2002). Sequestering metal ions that are required for protein function may explain how myricetin can act on such a wide range of protein targets. This may be the mechanism of action behind myricetin: FEN1 inhibition. However, data derived from the thermal shift screen (Chapter Five) indicates that myricetin is interacting directly with the protein, and not just preventing metal ion binding. In addition, earlier work on the inhibitory effect of myricetin on DNA polymerases (which have a FEN domain) from different organisms lend strength to the *Tb*FEN being an actual target (Ono & Nakane, 1990; Shinozuka *et al.*, 1988). Further work is required to fully understand myricetin and its mechanism of action against *T. brucei.* These experiments

could include over-expressing the *Tb*FEN D183K variant in the parasite (Chapter Three), and observing whether the presence of myricetin could prevent or rescue the detrimental phenotype. Another method would be to produce myricetin-resistant *T. brucei* clones by culturing the cells in increasing concentrations of the compound. The genomic DNA could be harvested and sequenced, allowing the identification of any factors (such as point mutations or loss of a gene) that can identify the target. This technique has been successfully used to produce and characterise nifurtimox and fexinidazole resistant strains (Wyllie *et al.*, 2016). Determining whether or not myricetin interacts directly with the *Tb*FEN enzyme should also be investigated. Isothermal titration calorimetry would enable binding parameters to be determined directly for any such interaction.

The trypanocidal effect of myricetin does validate it as a potential therapeutic. A patent for the use of myricetin against parasitic infections including *T. brucei* is in existence (Prendergast, 2001). The structure of myricetin could serve as a starting point for further structure-activity work aimed at improving selectivity and potency against *Tb*FEN. Another approach is the 'deconstruction-reconstruction' method of fragment-based drug design (Chen *et al.*, 2015). Target groups could be selected from a known inhibitor. These can then be optimised for certain characteristics, such as an ability to cross the blood brain barrier, or optimise its oral bioavailability, which is currently poor in myricetin (Chen *et al.*, 2015; Devi *et al.*, 2015).

4.4 Conclusions and Perspectives

The current data on compound D11 prevents it from being used for further screens. However, this does not detract from the results obtained for the C01 compound (Zhang, 2012). The possibility of selecting a flap endonuclease from one organism over another is possible. Now it needs to be achieved in medically relevant flap endonucleases.

The potential of the deconstruction-reconstruction method, using myricetin as a starting point, should be explored further. As myricetin has been shown to inhibit the *Tb*FEN, and is (to this author's knowledge) the first known compound to do so, it can be used in additional biochemical assays as a positive control.

5 Developing a Thermal Shift-Based High Throughput Screening Assay

5.1 Introduction

Developing compounds that specifically target the *Trypanosoma brucei* FEN over its human orthologue requires a sophisticated approach. One such approach is fragment-based drug design, which involves constructing a novel compound around fragments that bind the target protein. Fragment-based drug design has an excellent record of producing selective inhibitors for difficult targets (Baker, 2013). This chapter focuses on optimising a high throughput thermal shift-based assay for the identification of fragments.

5.2 Assay Rationale

Thermal shift assays require dyes that fluoresce when bound to hydrophobic protein regions (Pantoliano *et al.*, 2001). Mixing the native protein and dye together at physiological temperatures produces low-level fluorescence. Heating the solution causes the protein to denature, exposing its hydrophobic core. The dye binds to the hydrophobic residues and fluorescence increases (Pantoliano *et al.*, 2001). The signal will increase until all the protein is denatured, and the hydrophobic regions are saturated. From this, the melting temperature (T_M) can be calculated. This is equivalent to the temperature at which 50% of the protein is denatured. Compounds (or fragments) that bind to the protein can either stabilise or destabilise it. Stabilising compounds increase the T_M of the protein, whilst destabilising compounds decrease the T_M. For the purpose of this assay, stabilising compounds are considered potential inhibitory compounds, therefore any compounds/fragments that increase the T_M of the protein are considered positive hits (Pantoliano *et al.*, 2001).

5.3 Results

5.3.1 Enzyme Hydrophobicity

The nature of the thermal shift assay requires the protein of interest to have hydrophobic regions that the dye can associate with. In order to assess the hydrophobicity of both the human and the *Trypanosoma brucei* flap endonucleases, hydropathicity plots were produced using the EXPASY ProtScale online tool (http://web.expasy.org/protscale/). Two scales were compared; the Kyte and Doolittle, and Cowan (pH 7.5) hydropathicity scales (Cowan & Whittaker, 1990; Kyte & Doolittle, 1982). Scales were normalised to between 0 and 1, to allow for direct comparison (Figure 5-1). The plot was produced by grouping the protein sequence into clusters of seven consecutive amino acids, and producing an average hydropathicity score for each cluster (Gasteiger E., 2005). Both proteins display hydrophobic peaks, and the results for each scale are consistent.

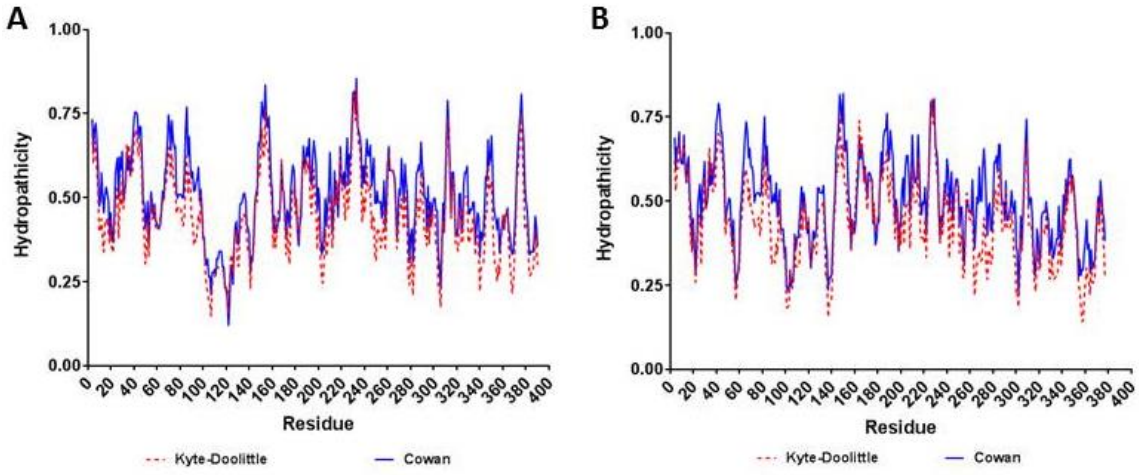


Figure 5-1 Hydropathicity plots for the *Tb*FEN (A) and *Hs*FEN (B)

The hydropathicity values were plotted against the amino acid sequence to generate protein-specific hydropathicity plots. Two scales were used and normalised to allow comparison. Plots were produced using the EXPASY ProtScale online tool (http://web.expasy.org/protscale/) (Gasteiger E., 2005), and visualised using GraphPad Prism (6.0).

5.3.2 Assay Optimisation

Thermal shift assays are commercially available, allowing easy optimisation. The kit used for this work was the Protein Thermal Shift™ kit from Invitrogen. Use of this kit with *Tb*FEN and *Hs*FEN required optimisation of the buffer conditions, specifically pH and additives, and protein concentration.

5.3.2.1 Protein Concentration

The protein concentration was the first parameter investigated. Assays were initially performed using the neutral, phosphate-based buffer supplied with the kit. Six protein concentrations (11, 2.3, 1.2, 0.6, 0.2 and 0.1 µM) were tested for both FEN proteins (Figure 5-2). The denaturation profile for each enzyme displays three distinct stages, which are shown in Figure 5-2. Because of the presence of this third phase, the curves were truncated prior to fitting with the Boltzmann equation.

Both proteins display a strong signal with a concentration of 11 µM. The *Tb*FEN protein displayed a strong signal at 2.3 and 1.2 µM, with an $R^2 \geq 0.9899$ for the Boltzmann equation. The *Hs*FEN protein had a weaker signal at a concentration of 2.3 µM, but still displayed excellent line fitting ($R^2 \geq 0.9976$). At lower concentrations of the two proteins, the signal quality and the line fitting accuracy is reduced. For some replicates, the Boltzmann equation could not be successfully fitted onto the curve, resulting in ambiguous results that were not included. Based on this, a minimum concentration of 1.2 µM and 2.3 µM of *Tb*FEN and *Hs*FEN (respectively) was used.

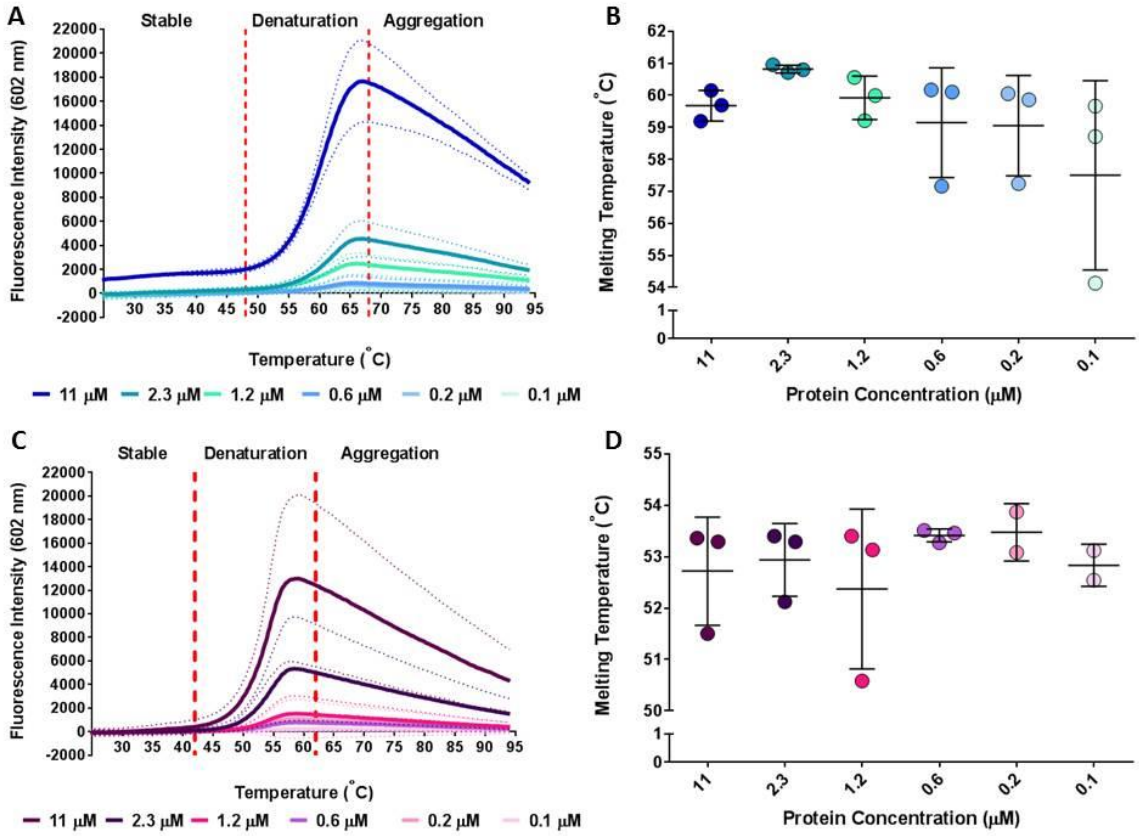


Figure 5-2 Optimising protein concentration for the *Tb*FEN (top) and *Hs*FEN (bottom)

Protein denaturation profiles for the *Tb*FEN (**A**) and *Hs*FEN (**C**) at six different protein concentrations. Each line depicts the mean of three biological replicates, with the dotted lines indicating the standard deviation from the mean. The denaturation profiles were used to calculate the melting temperature for the *Tb*FEN (**B**) and *Hs*FEN (**D**) using the Boltzmann equation (GraphPad Prism 6.0). Error bars indicate standard deviation from the mean.

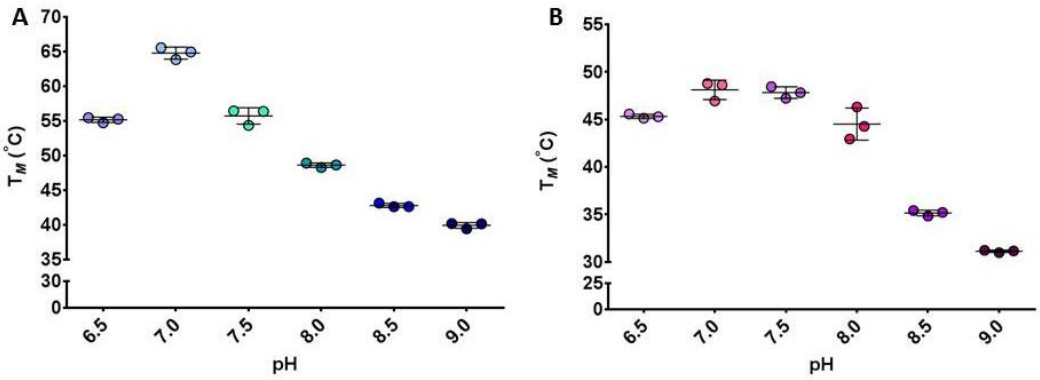


Figure 5-3 Melting temperatures for the *Tb*FEN (A) and *Hs*FEN (B) at six different pH values

Melting temperatures (T_M) were calculated for each pH value using the Boltzmann equation (GraphPad 6.0). Error bars indicate standard deviation of the mean from three technical replicates.

5.3.3.1 Buffer pH

The pH at which the protein was most stable was determined. The *Hs*FEN protein has previously been shown to be most active around pH 8.0 (Harrington & Lieber, 1994b), so four buffers (with a final concentration of 50 mM) at six different pH values were investigated (Figure 5-3). For the *Tb*FEN protein, HEPES-NaOH buffer at pH 7.0 appeared to stabilise the protein the most, with a T_M of 64.79°C ± 0.87. This was significantly higher than for all other pH conditions ($p \leq 0.0001$; determined by one-way ANOVA in GraphPad Prism (6.0)). For the *Hs*FEN protein, HEPES-NaOH buffer at either pH 7.0 or 7.5 appeared to stabilise the protein the most, with a T_M of 48.12°C ± 1.02 and 47.83°C ± 0.61 respectively. The T_M values for these two pH conditions were not significantly different ($p = 0.9993$; determined by one-way ANOVA in GraphPad Prism (6.0)).

Table 5-1 The effect of additives on the melting temperature of each protein

Component	Change in T_M (%)		P value	
	*Hs*FEN	*Tb*FEN	*Hs*FEN	*Tb*FEN
NaCl (50 mM)	-3.14	-3.94	0.1032	0.0031
KCl (100 mM)	-2.73	-8.37	0.1168	<0.0001
$MgCl_2$ (10 mM)	5.92	4.45	0.0011	0.0012
$MnCl_2$ (10 mM)	5.49	7.64	0.0020	<0.0001
DTT (1 mM)	-0.55	2.50	0.9811	0.0561
Tween	NA	NA	NA	NA
Triton X	NA	NA	NA	NA

The melting temperature for each condition was compared to the melting temperature of the protein using a One Way ANOVA in GraphPad Prism (6.0). T_M values were calculated from three technical replicates. NA = not applicable.

5.3.3.2 Buffer Composition

Buffer additives were carefully considered. The ultimate aim is to inhibit the *Tb*FEN enzyme *in vivo*. The buffer conditions must mimic that environment as much as possible, whilst providing good, reproducible readings. The components tested are summarised in Table 5-1. Both Tween and Triton X-100 could not be analysed as they displayed fluorescence in the absence of protein. The presence of both NaCl and KCl decreased the T_M of both proteins, although this decrease was not significant for *Hs*FEN (p = 0.1032 and 0.1168 respectively). Both Mg^{2+} and Mn^{2+} increased the T_M of both proteins significantly. DTT increased the T_M of the *Tb*FEN only although this was not significant (p = 0.0561). A final buffer composition of 50 mM HEPES-NaOH (pH 7.0), 1 mM DTT, 10 mM $MgCl_2$, 50 mM NaCl, 100 mM KCl was used.

Both fragment and compound libraries tend to be dissolved in 100% DMSO, so the effect of 5% DMSO on protein stability was analysed (Figure 5-4). The presence of 5% DMSO significantly decreases the melting temperature of the *Tb*FEN by 2.6%, and *Hs*FEN by 4.75% (p = 0.0220 and 0.0091 respectively, determined using an unpaired T test in GraphPad Prism 6.0).

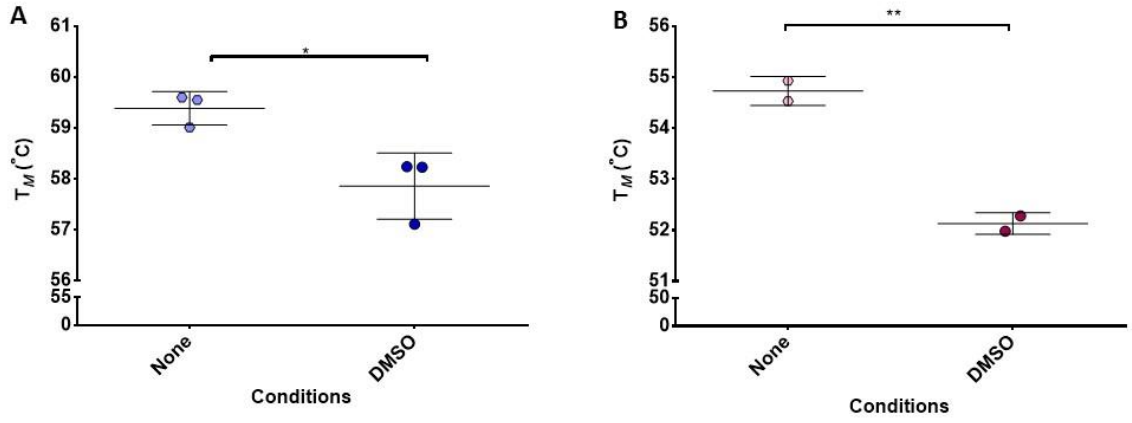


Figure 5-4 Effect of 5% DMSO on the melting temperature of *Tb*FEN (A) and *Hs*FEN (B)
Error bars depict standard deviation from the mean of two (*Hs*FEN) and three (*Tb*FEN) biological replicates. Statistical significance was analysed using an unpaired T test (GraphPad Prism 6.0).

5.3.3.3 Pharmacological Relevance

The sensitivity of the thermal shift assay was assessed using myricetin, a compound known to inhibit both the *Tb*FEN (Chapter Four) and the *Hs*FEN (van Pel *et al.*, 2013). Three myricetin concentrations were tested: 10, 50 and 100 µM. The calculated T_M values were compared to a DMSO control. The presence of myricetin significantly increases the T_M of both the *Tb*FEN and *Hs*FEN (Figure 5-5). For the *Tb*FEN, this increase is statistically significant at 50 and 100 µM ($p = 0.0007$ and $p \leq 0.0001$ respectively). For the *Hs*FEN, this is statistically significant at 100 µM only ($p = 0.0111$). This shows that the assay is capable of detecting known FEN inhibitors.

A statistical test to analyse assay robustness was calculated, and the values described in Table 5-2 (Zhang *et al.*, 1999). This value, called Z', was calculated for the two myricetin concentrations that caused a shift in T_M; 50 and 100 µM. Values between 0.5 and 1.0 indicate that the assay is of sufficient quality, as there is detectable separation between the positive and negative controls (myricetin and DMSO respectively. Only one sample did not reach the required Z' value, and that was one replicate of *Tb*FEN in the presence of 50 µM myricetin.

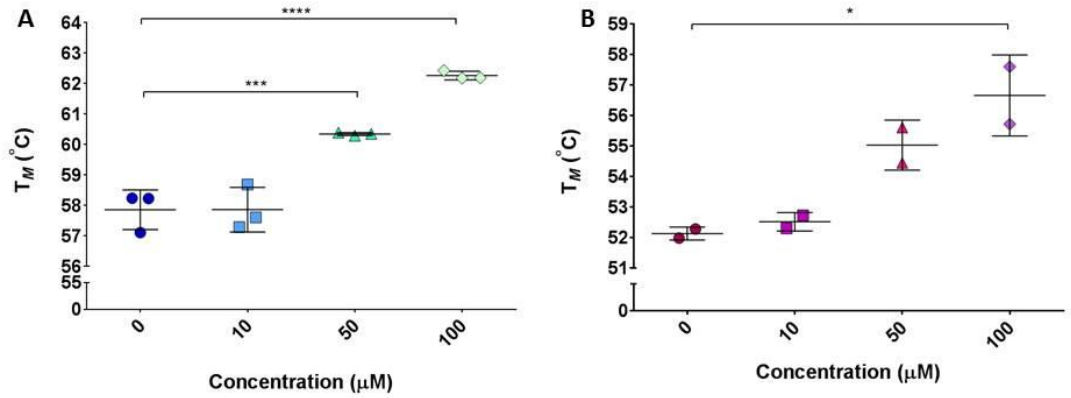


Figure 5-5 Effect of three concentrations of myricetin (10, 50 and 100 μM) on the melting temperature of *Tb*FEN (A) and *Hs*FEN (B) when compared to the DMSO only control (0).
The presence of 50 and 100 μM myricetin significantly increases the melting temperature of the *Tb*FEN. The presence of both 50 and 100 μM myricetin increases the melting temperature of the *h*FEN, but this is only significant with the higher myricetin concentration. Error bars indicate the standard deviation from the mean of three biological replicates for the *Tb*FEN, and two for the *Hs*FEN. Statistical significance was analysed using a One Way ANOVA (GraphPad Prism 6.0).

Table 5-2 Calculated Z' values for the myricetin control

Protein	Replicate	Myricetin Concentration (μM)	Z'
*Tb*FEN	1	50	0.69
		100	0.84
	2	50	0.74
		100	0.86
	3	50	0.17
		100	0.70
*Hs*FEN	1	50	0.69
		100	0.60
	2	50	0.51
		100	0.69

Z’ values were calculated as described in Section 7.8.

5.4 Discussion

5.4.1 Enzyme Hydrophobicity

Both the *Tb*FEN and *Hs*FEN were initially checked for assay compatibility by producing hydropathicity plots. Both enzymes display hydrophobic peaks. This was expected, as 36 and 37% of the amino acid sequence comprised of hydrophobic residues in *Hs*FEN and *Tb*FEN respectively. Other proteins with a similar percentage of hydrophobic residues have been successfully analysed by thermal shift, including human α-oestrogen receptor (37% hydrophobic residues) and bovine dihydrofolate reductase (35% hydrophobic residues) (Pantoliano *et al.*, 2001).

5.4.2 Assay Optimisation

5.4.2.1 Protein Concentration

The dye used in the assay has excitation and emission maxima of ~ 575 and 602nm respectively. This high excitation wavelength is beneficial as it minimises any intrinsic fluorescence properties of the fragments from interfering with the screen (Niesen *et al.*, 2007). Initial assays show that denaturation curves for both enzymes display three distinct stages; a stable stage, a denaturing stage and an aggregation stage (Figure 5-2). In the first stage, the protein is stable, and the fluorescent signal remains constant. In the second stage, the protein becomes denatured, allowing the dye to access and bind to hydrophobic regions in the protein core. This causes an increase in the fluorescence signal. The third stage is where the protein has fully denatured, but the fluorescent signal begins to decrease. This third stage was unexpected, but has been observed previously (Lo *et al.*, 2004; Niesen *et al.*, 2007), and has been attributed to the denatured proteins aggregating together to form plaques. This results in the fluorescence signal being quenched. Because of this, the denaturation plots were truncated prior to fitting with the Boltzmann equation. This results in a better fit for the line equation, as well as a more accurate T_M calculation.

Determining the appropriate protein concentration for the assay is essential. A high enough protein concentration needs to be used for a good, reproducible signal, whilst being low enough to maximise the number of assays per protein batch. Six protein concentrations were picked for analysis (11, 2.3, 1.2, 0.6, 0.2 and 0.1 µM). No statistically significant difference in T_M was calculated. For the *Tb*FEN, a protein concentration below 1.2 µM displayed greater variability between replicates, despite the $R^2 \geq 0.9227$. This is most likely

caused by a lower signal intensity producing more variable readings. For the *Hs*FEN, the signal produced when the protein concentration was less than or equal to 0.2 µM was very low. Because of this, higher protein concentrations were thought to be more accurate, and a minimum of 1.2 µM and 2.3 µM for the *Tb*FEN and *Hs*FEN respectively was picked. For the *Tb*FEN, this protein concentration is similar to those used in the literature (1 µM and 1.5 µM; (Lo *et al.*, 2004; Pantoliano *et al.*, 2001)). The concentration of the *Hs*FEN used is higher, and highlights the importance of optimising the assay on a protein-to-protein basis.

In the neutral phosphate buffer, the T_M of the *Tb*FEN is higher than that of the *Hs*FEN (~ 59°C and ~ 52°C respectively). This increased stability was not expected, but may be a consequence of the range of temperatures the *Trypanosoma brucei* parasite is exposed to throughout its life cycle. It would be interesting to determine if the enzymes are active at higher temperatures. This could be analysed by measuring the effect of temperature on substrate binding and cleavage using the assays described in Chapter Two.

5.4.2.2 Buffer pH

Analysing which pH is optimal for both proteins in this assay format is essential, as pH can have major effects on protein stability (Alber, 1989). The buffering agents picked were MES (pH 6.5), HEPES (pH 7.0, 7.5, 8.0) and Tris (pH 8.5, 9.0). MES, HEPES and Tris are classed as Good's buffers, as they do not make complexes with cations, do not interfere with enzymatic reactions, have minimal absorbance in UV and visible spectra, and changes in temperature have minimal effect on buffer dissociation (Good *et al.*, 1966). The pH of Tris-based buffers can change with temperature, but because of the other beneficial properties, this was still used. All pH values for the buffers were accurate at 37°C. For the *Tb*FEN protein, a pH of 7.0 is the optimal, with the T_M reaching ~65°C. This was interesting as a pH of 7.5 was expected to be the optimal. This is because the intracellular pH of the bloodstream trypomastigotes is 7.4, mimicking the pH of the human blood (Scott *et al.*, 1995). For the *Hs*FEN protein either pH 7.0 or 7.5 were the optimal (Figure 5-3). This was interesting, as the optimal pH of catalytic activity is 8.0. However, this may be due to the additives included in the reaction buffer.

5.4.2.3 Buffer Composition

A KCl concentration of 100 mM was tested because the literature indicates that, whilst KCl decreased *Hs*FEN substrate cleavage (Harrington & Lieber, 1994b), its presence increases DNA binding by ExoIX (Anstey-Gilbert *et al.*, 2013). Potassium ions have been observed in complex with both the *Hs*FEN and *E. coli* ExoIX bound to the helix-two turn-helix motif and

helix-three turn-helix motif respectively (Anstey-Gilbert *et al.*, 2013; Tsutakawa *et al.*, 2011). Binding of K^+ in the *Hs*FEN is coordinated by S237, I238 and I241 (Tsutakawa *et al.*, 2011). Both isoleucine residues are conserved in the *Tb*FEN, but an arginine residue is found in place of the serine. The results shown here indicate that KCl causes a decrease in the T_M of the enzyme, implying that it decreases stability. This may provide a potential reason as to why it increases DNA binding (Anstey-Gilbert *et al.*, 2013). Destabilising the protein may allow the DNA substrate to associate more readily, resulting in an increase in K_d. However, this needs to be analysed further in both the *Hs*FEN and *Tb*FEN, using either the fluorescence DNA anisotropy assay, or the bio-layer interferometry assay detailed in Chapter Two. Sodium chloride was included for protein stabilisation. The presence of 50 mM NaCl does decrease the T_M of both enzymes, but this is only significant for the *Tb*FEN. NaCl was used for its protective properties, to prevent the protein from precipitating (Kelly *et al.*, 2005).

Two divalent metal cations were tested in this system: Mg^{2+} and Mn^{2+}. Previous work in this thesis has shown that the presence of Mn^{2+} increases product formation by two-fold compared to the same concentration of Mg^{2+} (Chapter Two). Interestingly, Mn^{2+} increases the T_M of the *Tb*FEN by 7.6%, whilst Mg^{2+} increases the T_M by 4.5%. For the *Hs*FEN, both metal ions increase the T_M by a similar percentage (5.9 and 5.5% for Mg^{2+} and Mn^{2+} respectively). This increase in T_M may be due to the actual binding of two/three metal ions to the active site of the enzyme itself, therefore stabilising the enzyme further. This raises the question as to why the potential incorporation of the K^+ causes a decrease in protein stability, if this is also incorporated into the *Hs*FEN structure. Despite $MnCl_2$ stabilising the protein more, $MgCl_2$ was incorporated into the buffer. There are two reasons for this. Firstly, Mg^{2+} is the pharmacologically relevant divalent metal cation, and secondly, the background signal is lower (data not shown).

DTT is commonly added to enzymatic reactions to keep the cysteine residues in the reduced state that is typical of the functional protein. Both the *Tb*FEN and *Hs*FEN contain cysteine residues (three and six respectively). With the addition of DTT in the buffer there is no significant change in the T_M of the two proteins. This may be because DTT is present in all buffers throughout the purification process, so there should be no change in oxidation state of the cysteine residues when added to the buffer.

Two non-ionic detergents were tested in this assay format. This was attempted because certain compounds can self-aggregate. This can sequester the target protein onto the

surface of these aggregates, causing false positive results (Thorne *et al.*, 2010). It has been shown that the inclusion of these non-ionic detergents can decrease the number of hits by up to fifteen times when compared to that without detergent (Jadhav *et al.*, 2010). There is past precedent for using detergents in FEN screens: 0.01% v/v Tween-20 was used in a *Hs*FEN screen (McWhirter *et al.*, 2013). However, attempts to incorporate either Tween-20 or Triton X-100 were not successful, as both interacted with the dye, producing significant background noise (data not shown). Based on this, the shift buffer was composed of 50 mM HEPES-NaOH pH 7.0, 1 mM DTT, 50 mM NaCl, 100 mM KCl and 10 mM $MgCl_2$.

Most commercially available fragment/compound libraries are dissolved in 100% DMSO, so analysing the effect of DMSO in this assay is essential. A concentration of 5% was analysed as this was used previously, and shown to stimulate enzyme activity two fold (Chapter Four). DMSO significantly decreases the T_M of both proteins. This negative effect of DMSO has been described previously (Giugliarelli *et al.*, 2012; Tjernberg *et al.*, 2006). As a consequence, the effect of any compound/fragments on the T_M of the protein must be compared to a DMSO-containing control to ensure all potential hits are identified.

5.4.2.4 Pharmacological Relevance

In order to ensure that the assay is able to detect fragment/compound binding, the two proteins were tested using a known inhibitor: myricetin. A dose response can be seen, with an increase in T_M for both proteins with increasing concentrations of myricetin. This data implies that myricetin is interacting directly with the proteins, rather than with any buffer additives. This indicates that the assay can be used to detect a known inhibitor. However, it is important to normalise against a negative, no protein control, as myricetin does display fluorescence in the presence of the dye (data not shown). The Z' statistic was used to assess how robust the assay is by analysing the mean and standard deviation for the T_M values in the presence and absence of myricetin. This statistic was only calculated for the two myricetin concentrations where a shift in T_M was observed (50 µM and 100 µM). As shown in Table 5-2, all Z' values fall within the expected range ($1 > Z' \geq 0.5$) for both proteins, with one exception. One *Tb*FEN replicate in the presence of 50 µM myricetin, had a Z' value of 0.17. This is most likely caused by pipetting errors. Despite this, myricetin should still be considered a good positive control, but at concentrations of 100 µM for either *Tb*FEN or *Hs*FEN.

Whilst the reaction conditions have been optimised, the assay has not been fully validated. Coma *et al* (2009) describe steps to achieve full HTS validation that includes assessing

machine operation by running empty plates , and testing quality of the assay by performing multiple runs of control plates (protein with either 5% DMSO or 100 µM myricetin) (Coma *et al.*, 2009). Whilst the assay appears reproducible on a small scale, this needs to be analysed in detail to ensure the data gathered is robust. Time commitments prevented this from being performed, but should be a priority before any fragment libraries are screened using this method.

5.5 Conclusions and Perspectives

Selective inhibition of the *Tb*FEN over the *Hs*FEN will be difficult to achieve because the enzymes are 51% identical at the sequence level. However, fragment based drug design has yielded selective inhibitors against proteins with a similar level of sequence identity. One such example is selective inhibition of Jak-2 over Jak-3 (Antonysamy *et al.*, 2009). Jak-2 (accession number O60674) and Jak-3 (accession number P52333) display 50% sequence similarity, as determined by (Altschul *et al.*, 1990).

The current assay has been optimised for use with both the *Tb*FEN and *Hs*FEN. Successful optimisation will allow screening of fragment libraries. Screening of a 1000 fragment library (purchased from MayBridge) was attempted, but the concentration used for the fragments was too low to draw any meaningful conclusions from the data set. A minimum of 1 mM should be used in further screens, compared to the 50 µM used initially. Time constraints prevented this work from being repeated with the higher fragment concentration. Any fragment binding identified by this method will need to be analysed further. Potential techniques include nuclear magnetic resonance, surface plasmon resonance and isothermal calorimetry. These techniques will provide further information on the association and dissociation of the fragments with the protein. Successful fragments from these steps should then be taken forward for the labour and time intensive process of X-ray crystallography. Identification of the binding site will allow successful fragments to progress through to the compound production process. In this step, the fragments are joined to produce a compound with the desired properties. These properties include selective inhibition of the *Tb*FEN, ability to cross the blood-brain barrier, and ability to enter the parasite.

6 Summary and Context in *Trypanosoma brucei* Drug Development

6.1 Summary

Flap endonucleases have been intensively studied for nearly fifty years (Klett *et al.*, 1968; Liu *et al.*, 2004). The potential of these enzymes as drug targets has been highlighted for the human homologue relatively recently (Dorjsuren *et al.*, 2011; McManus *et al.*, 2009; McWhirter *et al.*, 2013; Panda *et al.*, 2009; Tumey *et al.*, 2005; Tumey *et al.*, 2004; van Pel *et al.*, 2013). However, research into the potential of flap endonucleases from medically relevant protozoans has been neglected. The work described in this thesis is classed as *'target re-purposing'* , whereby a protein in a different species (*Hs*FEN) is used to identify a parasitic orthologue (*Tb*FEN) (Klug *et al.*, 2016). Using a target-based approach is advantageous over phenotypic assays because the protein that is inhibited is already known (Plouffe *et al.*, 2008). However, this may not correlate to cell death *in vivo* (Plouffe *et al.*, 2008). Despite this, target-based drug design against human African trypanosomiasis (HAT) is widely reported, and has been reviewed recently (Scotti *et al.*, 2016). Of the enzymes reviewed in this paper, only two are involved in nucleic acid processing (RNA editing ligase I and Leucyl-tRNA synthase), whilst none are involved directly in DNA replication (Scotti *et al.*, 2016). This work highlights the potential of a DNA replication and repair protein as potential therapeutic target.

The production of the *Tb*FEN D183K variant was initially performed in order to obtain a protein: substrate complex in the presence of biologically relevant cations. This catalytically inert variant binds to DNA with similar affinity to the wild type enzyme (Chapter Two). The finding that the substitution of aspartic acid 183 with a lysine results in a catalytically inert protein supports work previously performed on the homologous T5 bacteriophage protein (Zhang, 2012). However, the equivalent mutation in the T5FEN protein (D153K) resulted in an approximately 10-fold increase in affinity for the same DNA substrate compared to the wild type T5FEN (Zhang, 2012). Additional investigation needs to be performed to determine if this observation is substrate-specific, or if mutating other active site residues to lysines has a similar effect. It is likely that mutating the aspartic acid residue to lysine has resulted in steric clashes that, whilst they don't affect the protein folding (Chapter Two), may affect the binding ability of the enzyme (Figure 6-1).

Figure 6-1 Potential steric clashes in the D183K active site

The Phyre2 model of *Tb*FEN (described in Chapter 2) was used to visualise the effect of the D183K variant (yellow) on the protein (gray). The PyMol mutagenesis wizard was used to display potential steric clashes (red discs). Two potential conformations are shown. The images were rendered by the author using The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

This inactive variant was hypothesised to cause a detrimental phenotype if over-expressed in *T. brucei*. This was tested and compared to RNAi knockdown of the FEN1 mRNA, and over-expression of the wild type enzyme (Chapter Three). Whilst the success of the RNAi knock down was inconclusive, the effect of the D183K variant on cell growth was striking, with a significant effect on cumulative cell growth observed after 48 hours post-induction, and a morphological change apparent after 72 hours post-induction. However, the phenotype caused by over-expression of the wild type FEN1 was not detrimental. A hypothetical scenario explaining this is described in Figure 6-2, where the effect of FEN1 is limited by loading to PCNA. Disrupting the *Tb*FEN clearly affects the cells, but additional investigation needs to be performed to fully understand the mechanism. Not only might this work be informative from a drug development perspective, it may also provide tools for investigating DNA metabolism.

Nuclear DNA replication within the *T. brucei* parasite has received little attention. That is shifting, but the most recent papers focus on the origin of replication, rather than the replication machinery itself (Marques *et al.*, 2016; Stanojcic *et al.*, 2016). Gaining an understanding of the parasite's nuclear DNA replication processes could highlight novel therapeutic targets, and the work in this thesis provides a starting point for this.

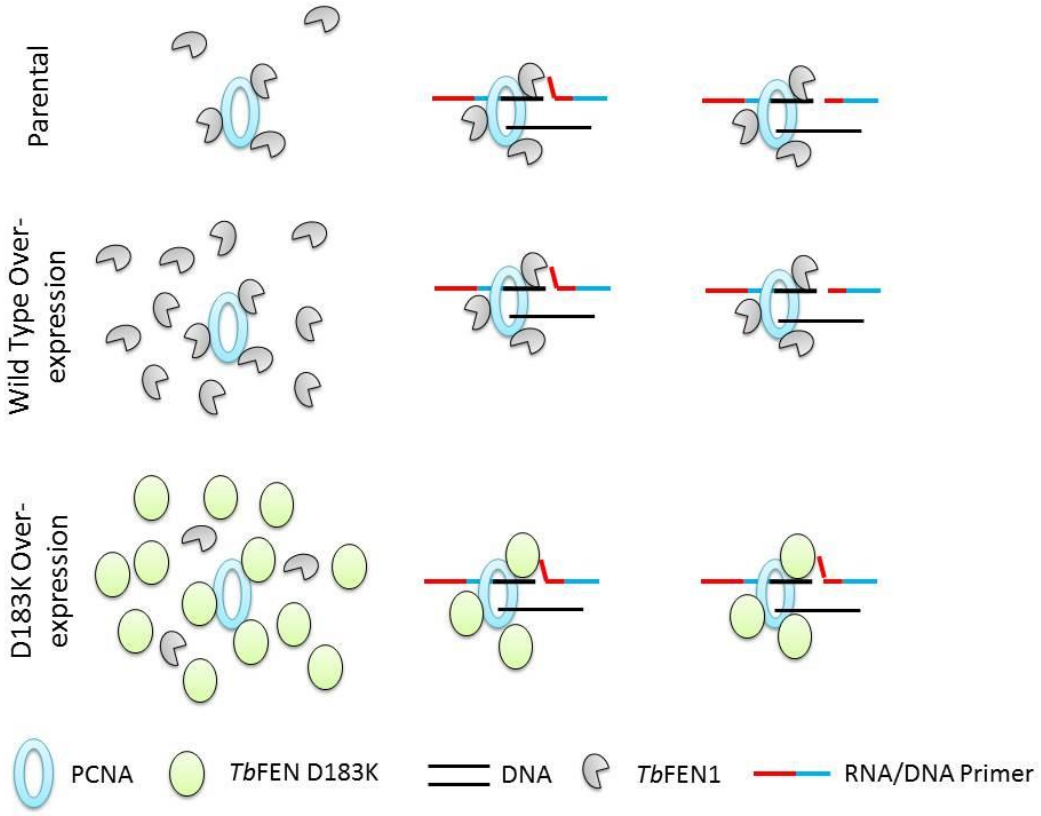


Figure 6-2 Hypothesised effect of FEN1 over-expression in *T. brucei* blood stream form cells

PCNA may act as a limiting factor for *Tb*FEN activity. Three FEN1 proteins can bind to the PCNA trimer, which recruits FEN1 to the site of cleavage (Chen *et al.*, 1996; Li *et al.*, 1995). Any effects caused by over-expression of wild type FEN1 may be limited by this step. Over-expression of the D183K variant may be swamping endogenous FEN1, and become loaded onto PCNA. This prevents successful DNA replication at the lagging strand, resulting in the detrimental phenotype observed.

The results from the cell-based work were promising, so we progressed to drug development, and tested known FEN inhibitors against the *Tb*FEN. Three previously identified compounds (D11, C01 and myricetin) that are known to inhibit other FEN proteins were tested. Only myricetin inhibited the *Tb*FEN, with an IC_{50} similar to the *Hs*FEN in this assay (Chapter Four). Critically, myricetin also caused cell death when tested against the parasite. Whether this is a result of a direct FEN1-myricetin interaction, or a broader, non-specific interaction, remains to be elucidated, and is discussed further in Chapter Four. Whilst it is unlikely that myricetin will be used as a therapeutic, it provides a starting point and a positive control. A typical drug design/development process takes several years, but

having a known FEN inhibitor will aid this, because functional groups can be identified, and potentially used in fragment-based approaches.

Another strategy employed in this thesis was to optimise a thermal shift-based assay to screen a fragment library (Chapter Five). The ability of this assay to detect known FEN inhibitors was tested against myricetin, and yielded positive results (Chapter Five). This type of assay is increasing in popularity as part of fragment-based screening process. It typically acts as the first step, to narrow down the number of fragments that get taken forward to more labour- and consumable-intensive techniques (Mashalidis *et al.*, 2013).

This thesis adds to the body of knowledge by identifying FEN1 as a potential therapeutic target that can be inhibited. A thermal shift-based assay has been optimised and a control compound identified to allow for compound/fragment based drug screening.

6.2 Future Work

6.2.1 Characterisation of the *Tb*FEN

The work presented in Chapter Two has highlighted a number of avenues for further exploration. Lysine substitutions for the remaining six active site amino acids were produced, but not analysed due to time constraints. This should be a priority, as we are still lacking that elusive protein: substrate crystal structure. The enzyme kinetics performed on the *Tb*FEN employed a simple model substrate, so only limited conclusions could be drawn from this in comparison with other FEN enzymes which have been more extensively studied (Anstey-Gilbert *et al.*, 2013; Finger *et al.*, 2009; Friedrich-Heineken & Hübscher, 2004; Garforth *et al.*, 2001; Harrington & Lieber, 1994b; Sayers & Eckstein, 1990; Tsutakawa *et al.*, 2011). However, the low-throughput DNA cleavage assay used in Chapter Three showed that the *Tb*FEN required a shorter incubation time than the *Hs*FEN for the same flap substrate (one hour and three hours respectively). As a consequence, performing a number of comparative kinetic experiments on a variety of different substrates for the *Tb*FEN and *Hs*FEN would be very interesting. In addition, preliminary work performed on the binding specificity of the *Tb*FEN enzyme with different substrate structures showed that the *Tb*FEN can bind to single-stranded DNA, 3’ overhangs and blunt-ended substrates (data not shown). This should be explored further to determine if binding correlates with substrate cleavage. If the *Tb*FEN has greater substrate promiscuity than the *Hs*FEN, it could indicate additional roles for *Tb*FEN that are currently unknown.

6.2.2 Understanding the role of the FEN enzyme in *Trypanosoma brucei* and other kinetoplastids

The work presented in Chapter Three indicates that having a functional flap endonuclease is essential for cell growth. However, initial attempts to elucidate why were not successful. There is no current literature on the role of flap endonucleases within *T. brucei*. If this enzyme is to be considered a drug target, more work needs to be performed to understand how targeting the enzyme will disrupt normal cellular functions, resulting in cell death. Another avenue that should be pursued is the role of the FEN in telomere maintenance (Gilson & Géli, 2007; Sampathi *et al.*, 2009). This could provide insights into telomeres in all organisms, especially as using trypanosomes as model organisms for telomere study was proposed in 2007 (Dreesen *et al.*, 2007). It would also be prudent to consider the potential of FEN1 to be involved with variant surface glycoprotein (VSG) switching. VSG genes are located adjacent to telomeres, and recombination is a key process for producing different VSG variants (Horn, 2014). Rad27 (the FEN1 orthologue in yeast) has been linked to restricting short sequence recombination in *Saccharomyces cerevisiae* (Negritto *et al.*, 2001). Analysis of FENs from other kinetoplastid parasites, specifically *Leishmania spp* and *Trypanosoma cruzi* would also be of interest. A paper published in 2008 highlighted a potential new role for *L. donovanni* FEN in a DNA 'degradesome' involved in programmed cell death, which should be explored further (BoseDasgupta *et al.*, 2008).

6.2.3 Screening Fragment Libraries

The optimised thermal-shift based assay for the *Tb*FEN recombinant protein provides a starting point for a drug development programme. This assay should be further validated, then used to screen fragment libraries (at a concentration of 1-10 mM) against both the *Tb*FEN and *Hs*FEN enzymes. This will allow hits to be identified, and select for those that are specific to the *Tb*FEN. This is essential given the level of conservation seen between the two FEN orthologues. Compounds that interact with the *Tb*FEN and not the *Hs*FEN should be taken forward for further analysis of protein: fragment binding using nuclear magnetic resonance spectroscopy, isothermal titration calorimetry and/or surface plasmon resonance (Dalvit, 2009; Harner *et al.*, 2013; Holdgate & Ward, 2005; Mashalidis *et al.*, 2013). Identification of the binding sites by X-ray crystallography is also essential to identify independent binding site to inform how the fragments can be used to create a novel compound. Given the high sequence identity between flap endonuclease from *T. brucei* and

T. cruzi (76%), and *Leishmania spp* (73%), any compounds produced from developing fragment hits against TbFEN should be tested against the flap endonucleases from these other kinetoplastid parasites.

7 Materials and Methods

7.1 Cloning

Three *Trypanosoma brucei* flap endonuclease (*Tb*FEN) constructs were produced (Table 7-1). These were modelled on the *T. brucei* FEN gene (NCBI Gene ID: 3656016), but were codon optimised for improved expression in *Escherichia coli*. Three additional *T. brucei* constructs were produced for over-expression/knock out *in vivo*. These were modelled on the same gene, but the native *T. brucei* coding sequence was used. Three human flap endonuclease (*Hs*FEN) constructs were used. These were modelled on the FEN gene (NCBI Gene ID: 2237), but were codon optimised for improved expression in *E. coli*.

Table 7-1 Constructs produced and used for this thesis

Name	Experiment	Production
*Tb*FEN-pET	Chapter 2, 4 and 5	Synthesised by Eurofins MWG Operon. Cloned into pET21a+ as part of this thesis.
*Tb*FEN D183K ΔC-pET	Production of *Tb*FEN D183K-pET	Synthesised by Eurofins MWG Operon. Cloned into pET21a+ as part of this thesis.
*Tb*FEN D183K-pET	Chapter 2	Produced from *Tb*FEN-pET and *Tb*FEN D183K ΔC-pET as part of this thesis.
*Tb*FEN-p2Mc^{N}	Chapter 3	Isolated from genomic DNA and cloned into p2Mc^{N} as part of this thesis.
*Tb*FEN D183K-p2Mc^{N}	Chapter 3	Produced from *Tb*FEN-p2Mc^{N} as part of this thesis.
*Tb*FEN-p2T7^{ti}	Chapter 3	Produced from *Tb*FEN-p2Mc^{N} and cloned into p2T7^{ti} as part of this thesis.
*Hs*FEN-pET	Chapter 4	Produced from *Hs*FEN D179K-pET as part of this thesis.
*Hs*FEN D179K-pET	Production of *Hs*FEN-pET	C-terminus cassette was synthesised by Eurofins MWG Operon. Full length *Hs*FEN D179K-pET produced as part of this thesis.
*Hs*FEN D179K ΔC-pET	Production of *Hs*FEN-pET	Synthesised by Eurofins MWG Operon. Cloned into pET21a+ as part of this thesis.

7.1.1 Bacterial Strains and Plasmids

Three *E. coli* strains were used in this work (Table 7-2). The first two strains, XLI blue and DH5α, were used to store recombinant plasmids at -80°C (Bullock *et al.*, 1987; Grant *et al.*, 1990). The third strain, BL21(DE3), was used for the over-expression of the gene of interest (Studier & Moffatt, 1986).

The pET21a+ plasmid (Novagen) was used for expression of all *E. coli*-produced protein. Two plasmids were used for *in vivo* work with *T. brucei* blood stream form cells; $p2Mc^{N}$ and $p2T7^{ti}$ provided by Dr Helen Price, originally from D. Horn and S. Alsford).

Table 7-2 *Escherichia coli* strains used, and their genotypes

Strain	Genotype
XLI Blue	(recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIqZΔM15 Tn10 (Tetr)])
DH5 α	(F^- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 φ80dlacZΔM15 Δ (lacZYA-argF) U169, hsdR17 ($r_K^- m_K^+$), λ^-)
BL21 (DE3)	(B F^- *ompT gal dcm lon* $hsdS_B(r_B^- m_B^-)$ λ(DE3 [*lacI lacUV5-T7 gene 1 ind1 sam7 nin5*]) $[malB^+]_{K\text{-}12}(\lambda^S)$)

7.1.2 Chemically Competent *E. coli* Cells

A single *E. coli* colony was picked from a lysogeny broth (LB) agar plate containing appropriate selection marker/s**.** This was placed in 3 mL of LB media containing appropriate selection marker/s, and incubated overnight at 37°C. The culture was diluted 1:100 in LB media containing appropriate selection marker/s (100 mL). Cells were grown at 37°C until an A_{600nm} of 0.5-0.7 was reached. The culture was centrifuged at 215 *xg* for ten minutes at 4°C. The cell pellet was re-suspended in pre-chilled 100 mM $CaCl_2$ (BDH), in a volume equivalent to 10% of the original LB media (10 mL) and incubated on ice for thirty minutes. The cell suspension was centrifuged as above. The cell pellet was gently re-suspended in pre-chilled 100 mM $CaCl_2$ containing 25% glycerol (v/v), at a volume equivalent to 4% of the initial LB media (4 mL). The cells were stored in 250 µL aliquots at -80°C.

LB-agar: 1 % (w/v) tryptone (Fisher), 0.5% (w/v) yeast extract (Oxoid), 0.5% (w/v) NaCl (Acros Organics), and 1.5% (w/v) of BactoAgar (BD)

LB media: 1.0% (w/v) tryptone, 0.5% (w/v) yeast extract and 0.5% (w/v) NaCl

7.1.3 Preparing Plasmid DNA

Plasmid DNA was isolated and purified from *E. coli* cells using commercial kits (Sigma, Qiagen and E.Z.N.A.). LB agar plates, containing appropriate selection marker/s, were streaked with the appropriate transgenic *E. coli*, and incubated at 37°C overnight. A single colony was isolated and inoculated into 3 mL LB media containing appropriate selection markers, and incubated at 37°C overnight. Overnight culture was diluted 1:50 in LB media, containing appropriate selection marker/s, and incubated at 37°C overnight. Bacterial cells were harvested by centrifugation at 16 000 *xg* for 10 minutes at room temperature, and purified by a modified alkaline-SDS lysis procedure, following the manufacturer's instructions.

7.1.4 Sub-Cloning

Plasmid DNA was digested by appropriate restriction enzymes (NEB or Promega) according to the manufacturer's protocol. Typically, 1 µg of DNA was incubated with 10 units of enzyme (enzyme units are defined by the manufacturer) in the appropriate buffer and incubated for one hour at 37°C. DNA fragments were separated by 1% agarose electrophoresis (see section 7.4.1 for details), and the desired band was excised and purified using the GenElute™ Gel Extraction Kit (Sigma). The DNA fragments were ligated into pET21a+ using T4 DNA Ligase (Thermo Scientific) according to the manufacturer's protocol. An example reaction is described in Table 7-3. Chemically competent *E. coli* cells were thawed on ice, and 50 µL of cells was incubated with 5 µL of the ligation product for one hour on ice. The cells were heat shocked at 42°C for thirty seconds, followed by ten minute incubation on ice. DH5α cells were plated directly onto LB agar plates containing appropriate selection marker/s. XLI blue cells were incubated for one hour at 37°C in fresh LB media (without antibiotic selection) before being plated onto LB agar plates containing appropriate selection marker/s.

The plasmid was isolated from the colonies as described in Section 7.1.3. Glycerol stocks of the bacterial cells were also produced by adding 20% glycerol to bacterial culture to a final volume of 1 mL. Glycerol stocks were stored at -80°C. Plasmid DNA was initially analysed by restriction enzyme digest and sequence identities were confirmed by DNA sequencing (Core Genomics Facility, The University of Sheffield) using appropriate primers (Table 7-7).

Table 7-3 Example ligation reaction

Components	Volume/Quantities
Digested vector	20-100 ng
Digested insert	2: 1 molar ratio (insert: vector)
10x Ligase buffer	1 µL
T4 DNA Ligase	0.5 Unit
Nuclease free water	To 10 µL

7.2 Site Directed Mutagenesis

Polymerase chain reaction (PCR) was used to create the necessary point mutations in all constructs. Two methods were used: splicing by overlap extension (SOE) PCR, and whole plasmid PCR.

7.2.1 Splicing by Overlap Extension PCR

This method was used to produce active site mutations in both the human and *T. brucei* flap endonucleases. The SOE-PCR reaction was performed using the following high fidelity polymerases; Phusion High Fidelity Polymerase (NEB), KAPA high fidelity polymerase (KAPA Biosystems) or Q5 high fidelity polymerase (NEB). Reactions were set up as described in the manufacturer's protocols. Typically 10 ng template DNA was used in a 25 µL reaction, with 0.3 µM of each oligonucleotide primer and 1 unit of polymerase. The PCR annealing temperature was dependent on the oligonucleotide (see Table 7-7). PCR products were analysed on a 1% agarose gel and band-purified using the GenElute^{TM} Gel Extraction Kit (Sigma). The two fragments from the first round of PCR (1 µL of each) were used as the template for the second round. All other reaction conditions remained the same. The PCR product was analysed on a 1% agarose gel and band purified as before. The PCR product was then sub-cloned into pET21a+, and inserted into appropriate *E. coli* cells (Section 7.1.4).

7.2.2 Whole Plasmid Site- Directed Mutagenesis

This technique was used to remove the internal *HindIII* site from the endogenous *T. brucei* FEN in order to insert it into the p2Mc^{N} over-expression vector. This technique is based on the NEB Q5 mutagenesis kit protocol (NEB). Primers were designed using the NEBaseChanger online tool (http://nebasechanger.neb.com/), and are listed in Table 7-7. This creates two oligonucleotide primers, a forward primer that contains the desired

mutation, and a reverse primer that begins on the complementary DNA sequence where the mutation primer begins. PCR was performed using the KAPA high fidelity polymerase (KAPA Biosystems), and set up as described in the manufacturer's protocol. An additional 30 seconds per 1kbp of DNA was used for the elongation steps. PCR products were PCR purified (Qiagen) and treated with *DpnI* (Promega) to remove methylated template DNA, as described by the manufacturer's protocol. *DpnI* was heat inactivated by incubating at 80°C for 20 minutes. The sample was then 5'-phosphorylated with T4 polynucleotide kinase (PNK, supplied by NEB). PNK was heat inactivated by incubating at 65°C for 20 minutes. This was followed by treatment with T4 DNA Ligase (NEB), to ligate the linear plasmid to produce closed circular DNA. This was then inserted into appropriate *E. coli* cells (Section 7.1.4).

7.3 Protein Over-Expression and Purification

7.3.1 Protein Over-Expression

All recombinant genes were expressed from the pET21a+ vector in the *E. coli* BL21(DE3) strain using either Studier's auto-inductive method or by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) (Studier, 2005). For the Studier method, the cells were grown in 500 mL of ZYM-5052 media containing 100 µg/mL of carbenicillin, at 27°C. The culture was incubated until the A_{600nm} plateaued, typically between an A_{600nm} of 4-6. Cells were recovered by centrifugation at 4000 *xg* for twenty minutes at 10°C, thoroughly re-suspended in 25 mM Tris (pH 8.0), 100 mM NaCl and 5 mM EDTA, then centrifuged as before. Typically, 4-8 g of cell paste was harvested per litre of culture. Cells were stored at -80°C.

For the IPTG induction method, cells were grown in 4 L of 2YT media (1.6% (w/v) tryptone, 1% (w/v) yeast and 0.5% (w/v) NaCl) containing 0.1% glucose, at 37°C until the A_{600} reached 0.4. The temperature was then decreased to 25°C, and the cells were induced by adding 0.1 mM IPTG. The cells were grown overnight. Cells were recovered by centrifugation at 4000 *xg* for twenty minutes at 10°C, washed in 25 mM Tris (pH 8.0), 100 mM NaCl and 5 mM EDTA, then centrifuged as before. Typically 5-10 g of cell paste per litre of culture was obtained. Cells were stored at -80°C.

ZYM-5052: 1% (w/v) tryptone, 0.5% yeast, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 0.2 x trace metals, 0.5% glycerol (v/v), 0.05% glucose (w/v), 0.2% α-lactose (w/v).

1000x trace metals: 50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnCl_2$-$4H_2O$, 10 mM $ZnSO_4$-$7H_2O$, 2 mM $CoCl_2$-$6H_2O$, 2 mM $CuCl_2$-$2H_2O$, 2 mM $NiCl_2$-$6H_2O$, 2 mM Na_2MoO_4-$2H_2O$, 2 mM Na_2SeO_3-$5H_2O$, 2 mM H_3BO_3.

2 YT: 1.6 % (w/v) tryptone, 1% (w/v) yeast and 0.5% (w/v) NaCl

7.3.2 Cell Lysis

Frozen cell paste was re-suspended in lysis buffer (50 mM Tris pH 8.0, 2 mM ethylenediaminetetraacetic acid (EDTA), 200 mM NaCl, 5% (v/v) glycerol) at a volume of 5 mL per gram of cells. Lysozyme (Sigma) was added to a final concentration of 200 µg/mL, and incubated at 4°C for twenty minutes. Phenylmethanesulfonyl fluoride (PMSF; Sigma), freshly prepared as a concentrated stock solution, was added to a final concentration of 23 µg/mL. This was followed by a forty minute incubation at 4°C, or until the solution had turned viscous. Sodium deoxycholate (Acros Organic) was added to a final concentration of 500 µg/mL, and incubated at 4°C for twenty minutes. Dithiothreitol (DTT) (1 mM) was added, and incubated at 4°C for five minutes. The solution was sonicated at 20% maximum amplitude for thirty second pulses as required to reduce viscosity and using a Jencons Ultrasonic Processor (GE-50) sonicator. The preparation was kept on ice throughout. Cellular debris was then removed via centrifugation at 75600 *xg* at 4°C for twenty minutes. The soluble fraction was analysed using SDS-PAGE to locate the protein, which was typically found in the supernatant. Ammonium sulphate (0.5 M) was added to the supernatant. The supernatant was stored on ice whilst a 5% polyethylenimine (PEI; Sigma) (w/v) precipitation was performed to precipitate the nucleic acids. The amount of 5% PEI used was determined by serial precipitation performed on 1 mL of supernatant. Initially, 10 µL of 5% PEI was added, vortexed and centrifuged down at 15000 *xg* at room temperature for ten minutes. This was repeated until no further pelleting occurred. This was scaled up for the bulk of the sample, which was stirred for ten minutes at 4°C, then centrifuged at 75600 *xg* for fifteen minutes at 4°C. The supernatant was then removed and placed on ice, whilst the pellet was discarded.

7.3.3 Ion Exchange and Affinity Chromatography

Purification was performed using a combination of cation exchange, anion exchange and affinity chromatography. All columns were run at 4°C. Either a phosphate buffer (20 mM K_2HPO_4, 20 mM KH_2PO_4, 1 mM DTT (Melford), 2 mM EDTA and 5% glycerol (v/v)) or a Tris-based buffer (20 mM Tris pH 8.0, 1 mM DTT, 2 mM EDTA, 5% glycerol) was used, depending on the pH required. All samples were dialysed using 50 mM NaCl, and loaded onto the column in the same buffer. Proteins were then eluted from the packed column bed under a salt gradient in appropriate buffer. Elution was performed over 20x the volume of the column itself. Fractions (one column volume) were collected, and the elution profile was monitored using denaturing SDS-PAGE gel. All proteins were purified to a minimum of 98% purity, as measured by densitometry using Image J.

7.3.4 Size Exclusion Chromatography

Size exclusion chromatography was used to further purify the protein samples, using the Superdex 200 10/300 GL column. Samples were concentrated down to a maximum volume of 500 µL using a centrifugal diafiltration device (Sartorius Stedim Biotech) at 4°C, with a molecular weight cut off of 5 kDa. The concentrated sample was injected into the column using an AKTA Prime system. Samples were eluted in 0.5 mL fractions in 20 mM K_2HPO_4, 20 mM KH_2PO_4, 1 mM DTT, 2 mM EDTA, 150 mM NaCl and 5% glycerol. The elution of the protein from the column was monitored using both UV and denaturing SDS-PAGE.

7.3.5 Protein Storage and Quantification

Purified proteins were concentrated to a maximum volume of 1 mL, using a centrifugal diafiltration device (Sartorius Stedim Biotech) at 4°C, with a molecular weight cut off of 5 kDa. Concentrated protein was stored in 50% glycerol at -20°C. The protein concentration was determined using the Bio-Rad Protein Assay, a colourmetric assay based on the Bradford assay. A standard curve was generated from bovine serum albumin (BSA; Sigma) standards ranging from 20-100 µg/mL (in duplicate). Protein samples and BSA standards were diluted 1:10 in diluted dye reagent (one part dye, four parts deionised water) and incubated at room temperature for 10 minutes. The absorbance at 595 nm was measured. The concentration of the unknown protein sample was calculated using the following equation:

$$\text{Protein concentration (mg/mL)} = \frac{y-c}{m*dilution\ factor} * 1000$$

Where ***y*** is the absorbance at 595 nm, ***c*** is the intercept and ***m*** is the gradient (both c and m are determined from the standard curve).

7.4 Electrophoresis Methods

7.4.1 DNA Electrophoresis

DNA was separated based on size using 1% (w/v) agarose gels. Powdered agarose (Web Scientific) was dissolved in TAE running buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8.0), containing 0.1 µg/mL of ethidium bromide, and set at room temperature. The gel was transferred to an electrophoresis tank containing TAE buffer and run between 2.6 – 4.2 V/cm (equivalent to 50-80 mA) until the DNA samples had migrated sufficiently. This was determined by full separation of the DNA ladder. The DNA bands were then visualised using the DC290 camera and the Kodak 1D 3.6.2 imaging software on a UV trans-illuminator.

7.4.2 Denaturing SDS-PAGE

SDS-PAGE gels were used to visualise proteins by separating them out based on size. Gels containing 10% acrylamide were used unless stated otherwise in the figure legends. Samples were mixed 1:1 with sodium dodecyl sulphate (SDS) loading dye and incubated for three minutes at 95°C. Samples were loaded into wells in the stacking gel and run at 20 V/cm (equivalent to 30 mA) until the dye line reached the edge of the resolving gel. The protein was stained using Coomassie Brilliant Blue stain, followed by a de-staining solution to remove excess dye. The gels were visualised using the DC290 camera and the Kodak 1D 3.6.2 imaging software on the FLA-3000 Imager.

SDS Loading Buffer: 0.2% SDS, 2.5% glycerol, 75 mM Tris-HCl (pH 6.8), 0.5% bromophenol blue (BDH), 5 mM EDTA and 1 mM DTT.

SDS-PAGE Stacking Gel: 120 mM Tris-HCl (pH 6.9), 10% acrylamide (30:1 acrylamide: bisacrylamide), 0.1% SDS, 0.5 mg/mL ammonium persulphate (APS) and 0.003% (v/v) tetramethylethylenediamine (TEMED).

SDS-PAGE Resolving Gel: 100 mM Tris-Bicine (pH 8.0), 10% acrylamide (30:1 acrylamide: bisacrylamide) (National Diagnostics), 0.1% SDS, 0.5 mg/mL APS (BDH) and 0.15% TEMED.

SDS-PAGE Running Buffer: 50 mM Tris-Bicine (pH 8.3) and 0.1% SDS.

Coomassie Brilliant Blue Stain: 40% (v/v) methanol (BDH), 10% (v/v) acetic acid and 2 mg/mL Coomassie Brilliant Blue R250 (Sigma).

SDS-PAGE De-stain Solution: 20% (v/v) methanol and 10% (v/v) acetic acid.

7.4.3 SDS-PAGE Substrate Gel

Nuclease activity of the FEN proteins (and any contaminants) was visualised using an SDS-PAGE substrate gel. The gel was prepared in the same way as a standard SDS-PAGE gel (see section 7.4.2), but 200 µg of Type XIV Herring sperm DNA (Sigma) was added to the resolving gel. Samples were prepared and run as described in section 7.4.2. The gel was washed thrice with 100 mL of TBG buffer for thirty minutes in total. The gel was incubated overnight in 200 mL of incubation buffer. The nuclease reaction was terminated by the removal of the incubation buffer, and washed once with 100 mL TBG. The gel was washed once with 100 mL TBG supplemented with 0.25 µg/mL of ethidium bromide, then washed in 100 mL TBG. Gels were visualised using the DC290 camera and the Kodak 1D 3.6.2 imaging software on a UV trans-illuminator. Nuclease activity was depicted as an area of reduced fluorescence. The substrate gel was stained with Coomassie Brilliant Blue and de-stained as a typical SDS-PAGE gel (Section 7.4.2) to visualise the protein.

TBG: 100 mM Tris-Bicine pH 8.0, 0.04 mM EDTA and 10% glycerol (v/v).

Incubation Buffer: 100 mM Tris-Bicine pH 8.0, 50 mM NaCl, 1 mM DTT, 10 mM $MgCl_2$, 100 mM KCl and 10% glycerol (v/v).

7.4.4 Low Through-Put Inhibitor Assay

This assay was developed by Dr Sarbendra L Pradhananga, a post-doctoral researcher in the Sayers laboratory. This assay allows the visualisation of a specific, single-labelled DNA substrate, from which the cleavage capability of the enzyme can be determined in the presence of potential inhibitors. These gels contain 25% acrylamide unless otherwise stated. The flap endonuclease of interest was mixed with assay buffer to give a final

concentration of 10 µg/mL. The candidate compounds were added to a final concentration of 50 µM (the DMSO percentage was kept constant at 5%). Compounds C01 and D11 were originally purchased on a "fee-for-service" basis from Sygnature Discovery Ltd (Nottingham), whilst Myricetin was purchased from Sigma Aldrich. Samples were incubated for five minutes at 37°C in a light protected amber tube (Eppendorf). The DNA substrate (Table 7-8) was annealed by diluting to a final concentration of 2.5 pmol/µL in assay buffer, and incubating at 95°C for five minutes, followed by five minute incubation on ice. The annealed substrate was added to the enzyme/inhibitor solution to a final concentration of 0.6 pmol/µL. This solution was incubated at 37°C for either one or three hours for the *Tb*FEN or *Hs*FEN respectively. This was diluted 1:1 in DNA loading dye, and incubated for five minutes at 70°C. Samples were loaded onto the 25% DNA-PAGE gel . and run at 22 V/cm (equivalent to 150 V) for ninety minutes in the dark. Gels were visualised using the DC290 camera and the Kodak 1D 3.6.2 imaging software on a UV trans-illuminator.

Assay Buffer: 25 mM HEPES, 0.5 mM EDTA, 100 mM NaCl, 10 mM $MgCl_2$, 100 mM KCl, 2 mM DTT, 0.1 mg/ml BSA, pH 7.0.

DNA Loading Dye: 30% (v/v) glycerol, 0.2% (w/v) xylene cyanol, 1% (w/v) SDS and 250 mM EDTA.

25% DNA Gel: 5 mL 30% acrylamide/bisacrylamide (19:1), 0.6 mL 10x TBE, 0.3 mL water, 0.1 mL 10% (w/v) APS and 0.01 mL TEMED.

7.5 Antibody Production

Antibodies against the N-terminus of the *Tb*FEN protein were raised in rabbits by a contract research organisation, Generon (www.generon.co.uk). Briefly, one New Zealand rabbit was injected subcutaneously with 450 µg of purified protein in complete Freund's adjuvant on days 1, 21, 51 and 81. Test sera were gathered on days 28 and 58, with final sera collection on day 88. The antibodies were purified using Zeba desalting columns (Thermo Scientific), buffer exchanged into PBS, and stored at -20°C.

7.6 Spectroscopy-Based Methods

7.6.1 UV Exonuclease Assay

The exonuclease activity of the enzyme was determined using a spectrophotometric method (Sayers & Eckstein, 1990). Reactions containing 25 mM HEPES-NaOH (pH 7.5), 10 mM $MgCl_2$, 100 mM KCl, 1 mM DTT and 667 µg/mL of herring sperm type XIV DNA were pre-warmed at 37°C for ten minutes. The recombinant FEN1 proteins were added and incubated at 37°C. Aliquots (100 µL) were taken during over a time course and quenched with equal volumes of 6% $HClO_4$. These were incubated on ice for ten minutes and centrifuged at 14000 *xg* for five minutes. The supernatant (150 µL) was added to 850 µL of H_2O and the absorbance measured at 260 nm in a 1 cm path length quartz cuvette. Increase in absorbance were plotted against time, and fitted by linear regression. One unit is defined as the amount of enzyme required to release 1 nmole of acid soluble nucleotides in thirty minutes at 37°C (1 A_{260} corresponds to 1.2 = 100 nmole nucleotide).

7.6.2 Circular Dichroism

Proteins were exchanged into circular dichroism (CD) buffer (20 mM K_2HPO_4, 20 mM KH_2PO_4, 1 mM DTT, 50 mM NaCl and 5% glycerol, pH 7.0) using a diafiltration device (Sartorius Stedim Biotech) with a molecular weight cut off of 5 kDa. Proteins were quantified by near UV absorption, as described by Kelly, Jess and Price (2005). Briefly, proteins were diluted into 6 M guanidine hydrochloride, pH 7.1, and the absorbance at 280 nm was recorded. The absorbance at 280 nm of a 1 mg/mL solution was defined as (Kelly *et al.*, 2005):

$$A280\ (1\ mg/mL) = \frac{5690_w + 1280_y + 60_c}{Mass\ (Da)}$$

Where ***w*** is the number of tryptophan residues present in the protein, ***y*** the number of tyrosine residues and ***c*** the number of cysteine residues (Kelly *et al.*, 2005).

The measured absorbance at A_{280} was then divided by the calculated A_{280} value for 1 mg/mL, to give the actual protein concentration, in mg/mL. Samples were diluted to between 0.1-0.5 mg/mL for analysis. Samples were run in the JASCO J-810 spectrometer. Run parameters are described in Table 7-4. Readings were measured in ellipticity (commonly called millidegrees). This was converted to molar ellipticity ($[\Theta]_{Molar\lambda}$) using the following equation (Kelly *et al.*, 2005):

$$[\theta]_{Molar\lambda} = \frac{100 * \theta_{\lambda}}{m * d}$$

Where $\boldsymbol{\theta_{\lambda}}$ is ellipticity (or millidegrees) at wavelength λ, ***m*** is the molar concentration of the protein and ***d*** is the path length (cm) (Kelly *et al.*, 2005).

Table 7-4 Parameters used on the JASCO J-810 spectrometer

Parameter	Setting
Start wavelength (nm)	250
End wavelength (nm)	195
Data pitch (nm)	1
Scanning mode	Continuous
Scanning speed (nm/minute)	20
Response (seconds)	8
Band width	1
Accumulation	3

7.6.3 Förster Resonance Energy Transfer Cleavage Assay

Förster resonance energy transfer (FRET) occurs when the emission spectra of a donor fluorophore overlaps with the absorption spectra of a quencher (Lakowicz, 2006). In this assay, a dual-labelled DNA substrate is incubated with enzyme. Cleavage of the quencher dye allows the increased fluorescence of the excited fluorophore to be detected.

Dual fluorescently-labelled oligonucleotides were used as the template for FRET substrates. Unlabelled, partially complementary oligonucleotides were used to create the flap substrates preferred by the FEN enzymes. The oligonucleotides used are shown in Table 7-8. Double-stranded oligonucleotide substrates were prepared to a final concentration of 0.5 µM in 25 mM HEPES-NaOH (pH 7.5), 1 mM EDTA and 0.5 M NaCl. The solution was heated to 85°C for five minutes, and was then left to cool slowly to room temperature. Annealed products were then stored at -20°C.

Substrate at the required concentration was prepared in 25 mM HEPES-NaOH pH 7.5, 100 mM KCl, 0.5 mM EDTA, 2 mM DTT and 0.1 mg/mL BSA. The enzyme was diluted to the appropriate concentration in the same buffer. The enzyme and substrate solution were pre-warmed to 37°C, and the reaction initiated by the addition of 10 mM $MgCl_2$. The fluorescence intensity data collection began ~ 10 seconds after the addition of $MgCl_2$ at one

second intervals using a HITACHI F-2500 FL spectrophotometer at 496 nm/519 nm (excitation/emission).

7.7 Bio-Layer Interferometry

This assay uses the principle of optical interferometry to detect protein binding to an immobilized DNA substrate. This binding causes a change in optical thickness, causing a wavelength shift (Sultana & Lee, 2015). This allows interactions to be monitored in real-time. The oligonucleotide substrate (Table 7-8) was diluted to a final concentration of 5 pmol/µL in binding buffer (25 mM HEPES-NaOH pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1 mg/mL BSA, 5 mM $CaCl_2$). The annealed substrate was produced by incubating the diluted oligonucleotide at 95°C for five minutes, then incubating on ice for five minutes. Binding experiments were performed on the FortèBIO BLItz machine, using streptavidin based probes. The process is described in Table 7-5. Probes were hydrated for a minimum of ten minutes in 200 µL of reaction buffer and 20 pmol of annealed biotinylated oligonucleotide was loaded over 120 seconds, and washed for a further 120 seconds. Baseline measurements were obtained using 250 µL of reaction buffer. Protein-DNA association was measured by loading 4 µL of freshly diluted protein onto the DNA bound probe over 120 seconds. Dissociation rates were measured by placing the protein-DNA bound probe into 250 µL of reaction buffer for 120 seconds. Data was corrected for system artefacts between the association and dissociation steps, and was analysed using global (1:1) curve fitting. This calculated the rate constants for association (k_{on}) and dissociation (k_{off}), and the dissociation constant (K_d).

Table 7-5 Run conditions for binding experiments on the Forte Bio BLItz machine

Stage	Time (mins)	Volume (µL)
Hydration	10	200
Initial Baseline	0.5	250
Association	2	4
Dissociation	2	250

7.8 Thermal Shift Assay

This assay uses a dye that fluoresces when it binds to hydrophobic regions of proteins. Denaturing the protein by increasing the temperature produces a melting curve, from which the melting temperature (T_M) can be calculated. The T_M is the temperature at which 50% of the protein is unfolded.

The assay was set up as described by the manufacturer (Life Technologies). Briefly, protein was diluted to a minimum of 1.2 µM and 2.3 µM (*Tb*FEN or *Hs*FEN respectively) in shift buffer (50 mM HEPES-NaOH pH 7.0, 1 mM DTT, 50 mM NaCl, 100 mM KCl, 10 mM $MgCl_2$ and 1x dye). Samples were loaded on to a MicroAmp® Optical 384 well plate (Applied Biosystems). Plates were sealed with a MicroAmp® Optical Adhesive Film (applied Biosystems) and centrifuged for one minute at 1000 rpm in a Sorvall® Legend T centrifuge. Plates were loaded onto an Applied Biosystems 7900 Real Time PCR machine. Using the SDS software (version 2.1 or 2.4), a dissociation stage was set up to increase from 25°C to 95°C, with a ramp rate of 2%. Data was exported into Microsoft Excel, and the melt curve was inserted into GraphPad Prism (version 6.0). Data was normalised against the negative controls, and analysed using the Boltzmann equation in GraphPad Prism (6.0). The Z' factor, a statistical test to assess assay robustness, was calculated using the following equation (Zhang *et al.*, 1999):

$$Z' = 1 - \frac{3(\sigma_p + \sigma_n)}{(\mu_p - \mu_n)}$$

Where $\boldsymbol{\mu_p}$ and $\boldsymbol{\sigma_p}$ are the mean and standard deviation of the positive control, and $\boldsymbol{\mu_n}$ and $\boldsymbol{\sigma_n}$ are the mean and standard deviation of the negative control (Zhang *et al.*, 1999).

7.9 *In Vivo* Parasitology

This work was performed in collaboration with Dr Helen Price, at Keele University.

7.9.1 Parasite Growth and Maintenance

The *Trypanosoma brucei brucei* bloodstream form (BSF), strain Lister 427 (Single Marker Lister) was maintained in HMI9 media (18.3 mg/mL HM19 powder (Gibco), 36 mM sodium hydrogen carbonate (Sigma), 14.3 µL/L β-mercaptoethanol (Sigma), 2 mM L-glutamine (Gibco), 100 U/mL penicillin (Gibco) and 100 µg/mL streptomycin) supplemented with 10% heat-inactivated tetracycline-free foetal calf serum (Bioline). Neomycin (2.5 µg/mL) was added to maintain the phage-derived T7 RNA polymerase and tetracycline repressor. BSF cells were grown at 37°C with 5% CO_2.

7.9.2 Transfection of *T. brucei* blood stream forms

BSF cells were transfected by electroporation. Briefly, 6 x10^5 cells/mL were collected by centrifugation at 800 *xg* at 20°C for ten minutes. Cells were washed in 10 mL of PBS and centrifuged as before. Cells were washed in 1 mL PBS and centrifuged at 4000 *xg* for five minutes at 20°C. The cell pellet was re-suspended in 100 µL T-cell Nucleofector ® solution (Lonza), and transferred to an Amaxa cuvette. Electroporation was performed using programme X-001 on the Amaxa Nucleofector ® machine (Lonza), with 4 µg, 3.5 µg and 4.5 µg of *NotI* digested p2T7^{ti}-FEN, p2Mc^{N}-FEN and p2Mc^{N}-FEND183K respectively. Transfected cells were diluted into 20 mL of pre-warmed HMI9 media, and plated out onto 24-well plates neat, or after a 1:10 dilution. Stable transfectants were selected for by the addition of 2.5 µg/mL phelomycin (Sigma) for the p2T7^{ti} plasmid, or 10 µg/mL hygromycin for the p2Mc^{N} plasmid, 16 hours post-transfection.

7.9.3 Cumulative Growth Curves

Mid log phase BSF cells (transformed or parental) were grown and then counted using a haemocytometer. Fresh flasks of HMI9 media with the appropriate additives were seeded with 5 x10^4 cells/mL. Expression of double stranded RNA, or the FEN1 wild type or D183K variant proteins, were induced by the addition of 1 µg/mL tetracycline. Cumulative growth was plotted against time (hours post induction).

7.9.4 Total RNA Extraction and Quantification

Cells were harvested by centrifugation at 800 *xg* at 20°C for ten minutes, washed in PBS (equal to one fifth of the original culture volume) and centrifuged at 800 *xg* at 20°C for ten

minutes. Cells were washed in PBS (equal to one hundredth of the original culture volume) and centrifuged at 4000 *xg* at 20°C for five minutes. The cell pellet was re-suspended in an equal volume of TRIsure and incubated at 20°C for five minutes. Chloroform was added (0.2 mL/1 mL TRIsure), mixed vigorously, incubated at room temperature for three minutes, then centrifuged for fifteen minutes at 12 000 *xg* at 4°C. The top aqueous layer was harvested, isopropanol added (0.5 mL/ 1 mL of TRIsure), and mixed by inversion. Samples were incubated at 20°C for ten minutes, then centrifuged for fifteen minutes at 12 000 *xg* at 4°C. The supernatant was discarded, and the pellet washed in 75% ethanol (1 mL per mL of TRIsure). Samples were vortexed and centrifuged at 7500 *xg* for five minutes at 4°C. The supernatant was discarded and the pellet air dried. RNA was re-suspended in 50 µL nuclease-free water and incubated at 60°C until the pellet had dissolved. The RNA concentration and purity was determined using the Nanodrop 1000, then stored at -80°C.

7.9.5 Reverse Transcription Quantitative PCR (RT-qPCR)

Two-step RT-qPCR was carried out on harvested RNA (Section 7.9.4). The first step converts the RNA to cDNA (reverse transcription), whilst the second step involves absolute quantification of the total RNA by comparing threshold cycles (C^T) against a standard curve.

RNA was treated with DNase (Ambion) to remove genomic DNA, according to the manufacturer's protocol. Briefly, 4 µg RNA was treated with 1 unit of DNase for fifteen minutes at 37°C. Inactivation buffer (equal to 10% of the reaction volume) was added, and incubated at 20°C for two minutes. The sample was centrifuged at 10 000 *xg* for one minute at 20°C, and the supernatant harvested. Reverse transcription (RT) was performed using the Omniscript RT kit (Qiagen). Two reactions were set up per RNA sample, one with and one without the Omniscript RT enzyme. Reactions were performed according to the manufacturer's protocol. Briefly, 4 µg DNase-treated RNA was incubated with 0.5 U/µL of RNaseOUT ™ (Invitrogen) inhibitor, 0.2 U/µL of Omniscript RT enzyme, 20 ng/µL oligo-dT (Qiagen) and 0.5 mM each dNTP (dA, dG, dC and dT). Samples were incubated at 37°C for sixty minutes to produce cDNA. Samples were stored at -20°C until required for PCR.

Quantitative PCR was performed using Applied Biosystems 7900HT thermocycler, with the PowerUp™ SYBR® Green Master Mix (Life Technologies). Oligonucleotides were designed as described by Thornton and Basu (2011), except PerlPrimer was used instead of Beacon Designer™ (Thornton & Basu, 2011). The FEN1 specific oligonucleotide pair used is described in Table 7-8. Standards of p2Mc^N- FEN1 were produced, beginning with 1x10^9

molecules per µL, then serially diluted 1:10 down to $1x10^3$ molecules per µL. Reactions were set up according to the manufacturer's protocol. Briefly FEN1-specific oligonucleotides were diluted to a final concentration of 125 nM, whilst α-tubulin specific oligonucleotides (kindly provided by Dr Helen Price) were diluted to 300 nM. A final concentration of 40 ng of cDNA was used (assuming 100% conversion and recovery of RNA to cDNA). MicroAmp® optical 384 well plates (Life Technologies) were sealed with a MicroAmp® Optical Adhesive Film and centrifuged for one minute at 1000 rpm in a Sorvall® Legend T centrifuge. The computer programme SDS (version 2.1 or 2.4) was used with the Applied Biosystems 7900HT thermocycler, and was programmed as described in Table 7-6.

Table 7-6 qPCR parameters

Stage Number	Temperature (˚C)	Time (seconds)	Number of Repeats
One	50	120	1
Two	95	300	1
Three	95	15	40
	60	60	
Four	95	15	1
	60	30	
	95	15	

7.9.6 Immunofluorescent Confocal Microscopy

Mid log phase BSF cells were counted using a haemocytometer. Cells ($5 x10^6$ cells) were harvested by centrifugation at 800 *xg* for ten minutes at 20˚C. The cell pellet was re-suspended in 10 mL vPBS (137 mM NaCl, 3 mM KCl, 16 mM Na_2HPO_4, 3 mM HK_2PO_4, 46 mM sucrose, 10 mM glucose) and centrifuged, as before. The cell pellet was re-suspended in 4% paraformaldehyde/PBS, and incubated on ice for sixty minutes. The paraformaldehyde was diluted 1:10 in ice cold PBS, and centrifuged at 4˚C for ten minutes at 800 *xg*. The cell pellet was re-suspended in 1 mL PBS, and centrifuged at 3000 *xg* for three minutes at 20˚C. The cell pellet was re-suspended in 50 µL of PBS, applied to a polysine slide and left at room temperature for twenty minutes. Excess liquid was removed from the slides, cells were permeabilised in 0.1% Triton-X 100 in PBS, and incubated at room temperature for ten minutes. Slides were washed once in PBS, Image iT FX signal enhancer (Life Technologies) was applied per sample and incubated at room temperature for thirty minutes. Excess liquid was removed and replaced with 100 µL of diluted primary antibody (rabbit α-FEN1; 1:100 and mouse α-TAT1; 1:200, provided by Dr Helen Price, originally a gift from Professor

Keith Gull). Slides were incubated at room temperature for one hour. Excess liquid was removed, and slides washed thrice in PBS for five minutes each time, under gentle agitation. Slides were dried, and 100 µL of diluted secondary antibody (goat α-mouse IgG Alexa Fluor 488 and goat α-rabbit IgG Alexa Fluor 594, both 1:250 (Life Technologies)) was added, and incubated in the dark at room temperature for one hour. Excess liquid was removed and slides washed as before, but in the dark. Excess liquid was removed and the coverslip was added using ProShield mounting medium (Life Technologies). Slides were left overnight in the dark, at room temperature, to allow the mounting medium to set, then stored at 4°C until visualised. Slides were examined using the Zeiss LSM 510 NLO inverted confocal microscope with a 63x/1.4 oil lens. Images were acquired using the LSM 510 software.

7.9.7 Flow Cytometry

Propidium iodide is a DNA intercalating agent that enters dead cells. Using this on dead cells allows the cell cycle progression of the cell population to be analysed, and any defects in the cell cycle, can be detected. Cells were harvested by centrifugation at 800 *xg* at 20°C for ten minutes, and then washed once in PBS. Cells were re-suspended in 70% methanol, 30% PBS and incubated at 4°C overnight. Cells were washed once in 10 mL cold PBS, the centrifuged at 400 *xg* at 20°C for twenty minutes. Cells were re-suspended in 1 mL PBS containing 10 µg/mL propidium iodide (Sigma) and 10 µg/mL RNase A (Thermo Scientific) at a cell density of 5 $x10^6$ cells/mL. Samples were incubated at 37°C for thirty minutes. Cell cycle analysis was performed on the Becton Dickinson FACSCalibur, using the FL3-A detector. Data for 10 000 cells was collected per sample. Data was analysed using FlowJo, version 10.0.7v2.

Table 7-7 Oligonucleotides used in all sub-cloning, site directed mutagenesis and sequencing protocols

Name	Sequence (5' - 3')	Technique
For_*Xba*	ttt*TCTAGA*ATGGGTGTTCTTGGCCTTTCG	Sub-cloning
Rev_*Bam*	ttt*GGATCC*CTACTTTTTGACAGCCTTTTTGTGTCC	Sub-cloning
For_KK	TGGCCTTTCGaaaCTTCTGTACG	Site directed mutagenesis
Rev_KK	AGAACACCCATCTAGAGTCG	Site directed mutagenesis
For_*Hind*	ttt*AAGCTT*ATGGGTGTTCTTGGCCTTTCG	Sub-cloning
For_D183K	GGAACAGAG**AAG**ATGGATGGAT	Site directed mutagenesis
Rev_D183K	GGCATCCAT**CTT**CTCTGTTCCTAC	Site directed mutagenesis
RNA_*Xba*_for	ttgt*TCTAGA*GGATCAGATGGAGGAGGTGA	Sub-cloning
RNA_*Hind*_rev	ttgt*AAGCTT*TTCGTAAGTGCGTCACGAAG	Sub-cloning
D34K_for	CGTATCGCGATTAAAGCCAGTATGGCCG	Site directed mutagenesis
M13_for	AGGGTTTTCCCAGTCACGACGTT	Site directed mutagenesis
M13_rev	GAGCGGATAACAATTTCACACAGG	Site directed mutagenesis
K179D_for	TGCAGCAGCCACCGAA**GAT**ATGGATTGTCTGACCTTT	Site directed mutagenesis
K179D_rev	CCAAAGGTCAGACAATCCAT**ATC**TTCGGTGGCTGCTGCA	Site directed mutagenesis
D34K_rev	CGGCCATACTGGCTTTAATCGCGATACGACG	Site directed mutagenesis
New_for	TATACATATGGGCGTCCTTGGGCTGFTCTAAACTGTTGTATG	Site directed mutagenesis
New_rev	CGCAAGCTTATCATTTTTTCACTGCTTTTTTATGG	Site directed mutagenesis
T7_for	TAATACGACTCACTATAGGG	Sequencing
T7_term	GCTAGTTATTGCTCAGCGG	Sequencing

Codons involved in site directed mutagenesis are underlined and in bold. Restriction enzyme sequences are shown in italics.

Table 7-8 Oligonucleotides used in analytical techniques

Name	Sequence (5' - 3')	Modification	Technique
Bio_OHP2	CTCTGTCGAACACACGCZTGCGTGTGTTC	6-BDT	BLI
Aniso_sub	TAAACGCTGTCTCGCTGAAAGCGAGACAGCGAAAGACGCTCGT	6 FLUDT	Inhibitor
Aniso_flap	ACGAGCGTCTTT	NA	Inhibitor
Assay3	TTTTCGCTGTCTCGCTGAGT	5' CY3, 3'FLU	FRET
Assay3_flap	ACTCAGCGAGACAGCGCCGGAACACACGCTGCGTGTGTTCCGG	NA	FRET
RT_FEN1_for	TGGCCGTATATCAGTTCGTC	NA	RT-qPCR
RT_FEN1_rev	CGTAACATCCCCAGCTTCAT	NA	RT-qPCR
RT_TUB_for	CGTGAGGCTATCTGCATCCA	NA	RT-qPCR
RT_TUB_rev	CCCAGCAGGCGTTACCAA	NA	RT-qPCR

8 References

Alber, T. (1989). Mutational effects on protein stability. *Annu Rev Biochem,* 58, 765-798. doi: 10.1146/annurev.bi.58.070189.004001.

Alleva, J. L. and Doetsch, P. W. (1998). Characterization of Schizosaccharomyces pombe Rad2 protein, a FEN-1 homolog. *Nucleic Acids Res,* 26, 3645-3650.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol,* 215, 403-410. doi: 10.1016/S0022-2836(05)80360-2.

Anene, B. M., Onah, D. N. and Nawa, Y. (2001). Drug resistance in pathogenic African trypanosomes: what hopes for the future? *Vet Parasitol,* 96, 83-100.

Anstey-Gilbert, C. S., Hemsworth, G. R., Flemming, C. S., Hodskinson, M. R., Zhang, J., Sedelnikova, S. E., Stillman, T. J., Sayers, J. R. and Artymiuk, P. J. (2013). The structure of Escherichia coli ExoIX--implications for DNA binding and catalysis in flap endonucleases. *Nucleic Acids Res,* 41, 8357-8367. doi: 10.1093/nar/gkt591.

Antonysamy, S., Hirst, G., Park, F., Sprengeler, P., Stappenbeck, F., Steensma, R., Wilson, M. and Wong, M. (2009). Fragment-based discovery of JAK-2 inhibitors. *Bioorg Med Chem Lett,* 19, 279-282. doi: 10.1016/j.bmcl.2008.08.064.

Ashley, C. T. and Warren, S. T. (1995). Trinucleotide repeat expansion and human disease. *Annu Rev Genet,* 29, 703-728. doi: 10.1146/annurev.ge.29.120195.003415.

Bae, S. H. and Seo, Y. S. (2000). Characterization of the enzymatic properties of the yeast dna2 Helicase/endonuclease suggests a new model for Okazaki fragment processing. *J Biol Chem,* 275, 38022-38031. doi: M006513200 [pii] 10.1074/jbc.M006513200.

Baker, M. (2013). Fragment-based lead discovery grows up. N*at Rev Drug Discov,* 12, 5-7. doi: 10.1038/nrd3926.

Baker, N., de Koning, H. P., Mäser, P. and Horn, D. (2013). Drug resistance in African trypanosomiasis: the melarsoprol and pentamidine story. T*rends Parasitol,* 29, 110-118. doi: 10.1016/j.pt.2012.12.005.

Baker, N., Glover, L., Munday, J. C., Aguinaga Andrés, D., Barrett, M. P., de Koning, H. P. and Horn, D. (2012). Aquaglyceroporin 2 controls susceptibility to melarsoprol and pentamidine in African trypanosomes. P*roc Natl Acad Sci U S A,* 109, 10996-11001. doi: 10.1073/pnas.1202885109.

Balakrishnan, L. and Bambara, R. A. (2013). Flap endonuclease 1. A*nnu Rev Biochem,* 82, 119-138. doi: 10.1146/annurev-biochem-072511-122603.

Baldwin, G. S., Sessions, R. B., Erskine, S. G. and Halford, S. E. (1999). DNA cleavage by the EcoRV restriction endonuclease: roles of divalent metal ions in specificity and catalysis. J *Mol Biol,* 288, 87-103. doi: 10.1006/jmbi.1999.2672.

Barnes, C. J., Wahl, A. F., Shen, B., Park, M. S. and Bambara, R. A. (1996). Mechanism of tracking and cleavage of adduct-damaged DNA substrates by the mammalian 5'- to 3'-exonuclease/endonuclease RAD2 homologue 1 or flap endonuclease 1. J *Biol Chem,* 271, 29624-29631.

Barrett, M. P. and Croft, S. L. (2012). Management of trypanosomiasis and leishmaniasis. B*r Med Bull,* 104, 175-196. doi: 10.1093/bmb/lds031.

Barrett, M. P., Vincent, I. M., Burchmore, R. J., Kazibwe, A. J. and Matovu, E. (2011). Drug resistance in human African trypanosomiasis. F*uture Microbiol,* 6, 1037-1047. doi: 10.2217/fmb.11.88.

Batista, A. N., Batista, J. M., Bolzani, V. S., Furlan, M. and Blanch, E. W. (2013). Selective DMSO-induced conformational changes in proteins from Raman optical activity. P*hys Chem Chem Phys,* 15, 20147-20152. doi: 10.1039/c3cp53525h.

Bellofatto, V., Fairlamb, A. H., Henderson, G. B. and Cross, G. A. (1987). Biochemical changes associated with alpha-difluoromethylornithine uptake and resistance in Trypanosoma brucei. M*ol Biochem Parasitol,* 25, 227-238.

Berg, J. M., Tymoczko, J. L. and Stryer, L. (2002a). Proteins Are Built from a Repertoire of 20 Amino Acids. In B*iochemistry* W H Freeman, New York.

Berg, J. M., Tymoczko, J. L. and Stryer, L. (2002b). Restriction Enzymes: Performing Highly Specific DNA-Cleavage Reactions. In B*iochemistry* W H Freeman, New York.

Blackburn, E. H., Greider, C. W. and Szostak, J. W. (2006). Telomeres and telomerase: the path from maize, Tetrahymena and yeast to human cancer and aging. N*at Med,* 12, 1133-1138. doi: 10.1038/nm1006-1133.

Bornarth, C. J., Ranalli, T. A., Henricksen, L. A., Wahl, A. F. and Bambara, R. A. (1999). Effect of flap modifications on human FEN1 cleavage. B*iochemistry,* 38, 13347-13354.

BoseDasgupta, S., Das, B. B., Sengupta, S., Ganguly, A., Roy, A., Dey, S., Tripathi, G., Dinda, B. and Majumder, H. K. (2008). The caspase-independent algorithm of programmed cell death in Leishmania induced by baicalein: the role of LdEndoG, LdFEN-1 and LdTatD as a DNA 'degradesome'. C*ell Death Differ,* 15, 1629-1640. doi: 10.1038/cdd.2008.85.

Bray, P. G., Barrett, M. P., Ward, S. A. and de Koning, H. P. (2003). Pentamidine uptake and resistance in pathogenic protozoa: past, present and future. T*rends Parasitol,* 19, 232-239.

Breda, A., Valadares, N., Norberto de Souza, O. and Garratt, R. (2006). Protein Structure, Modelling and Applications. In B*ioinformatics in Tropical Disease Research: A Practical and Case-Study Approach* (eds. Gruber, A., Durham, A., Huynh, C., and a*l.*, e.*)*, National Centre for Biotechnology Information, Bethesda.

Brenndörfer, M. and Boshart, M. (2010). Selection of reference genes for mRNA quantification in Trypanosoma brucei. M*ol Biochem Parasitol,* 172, 52-55. doi: 10.1016/j.molbiopara.2010.03.007.

Bridges, D. J., Gould, M. K., Nerima, B., Mäser, P., Burchmore, R. J. and de Koning, H. P. (2007). Loss of the high-affinity pentamidine transporter is responsible for high levels of cross-resistance between arsenical and diamidine drugs in African trypanosomes. M*ol Pharmacol,* 71, 1098-1108. doi: 10.1124/mol.106.031351.

Brosh, R. M., Jr., Driscoll, H. C., Dianov, G. L. and Sommers, J. A. (2002). Biochemical characterization of the WRN-FEN-1 functional interaction. B*iochemistry,* 41, 12204-12216.

Brosh, R. M., Jr., von Kobbe, C., Sommers, J. A., Karmakar, P., Opresko, P. L., Piotrowski, J., Dianova, I., Dianov, G. L. and Bohr, V. A. (2001). Werner syndrome protein interacts with human flap endonuclease 1 and stimulates its cleavage activity. E*MBO J,* 20, 5791-5801. doi: 10.1093/emboj/20.20.5791.

Bruce, D. (1895). P*reliminary report on the tsetse fly disease or nagana in Zululand, B*ennett and Davis, Durban.

Brun, R., Blum, J., Chappuis, F. and Burri, C. (2010). Human African trypanosomiasis. L*ancet,* 375, 148-159. doi: 10.1016/S0140-6736(09)60829-1.

Bullock, W., Fernandex, J. and Short, J. (1987). XL1-Blue: a high efficiency

plasmid transforming recA Escherichia coli strain with β-galactosidase selection. B*iotechniques,* 5, 376-378.

Bustos, P. L., Perrone, A. E., Milduberger, N., Postan, M. and Bua, J. (2015). Oxidative stress damage in the protozoan parasite Trypanosoma cruzi is inhibited by Cyclosporin A. P*arasitology,* 142, 1024-1032. doi: 10.1017/S0031182015000232.

Campillo, N. and Carrington, M. (2003). The origin of the serum resistance associated (SRA) gene and a model of the structure of the SRA polypeptide from Trypanosoma brucei rhodesiense. M*ol Biochem Parasitol,* 127, 79-84.

Cao, W. and Barany, F. (1998). Identification of TaqI endonuclease active site residues by Fe2+-mediated oxidative cleavage. J *Biol Chem,* 273, 33002-33010.

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

Carter, N. S. and Fairlamb, A. H. (1993). Arsenical-resistant trypanosomes lack an unusual adenosine transporter. N*ature,* 361, 173-176. doi: 10.1038/361173a0.

CDC (2015). Parasites - African Trypanosomiasis (also known as Sleeping Sickness).

Chatr-Aryamontri, A., Breitkreutz, B. J., Oughtred, R., Boucher, L., Heinicke, S., Chen, D., Stark, C., Breitkreutz, A., Kolas, N., O'Donnell, L., Reguly, T., Nixon, J., Ramage, L., Winter, A., Sellam, A., Chang, C., Hirschman, J., Theesfeld, C., Rust, J., Livstone, M. S., Dolinski, K. and Tyers, M. (2015). The BioGRID interaction database: 2015 update. N*ucleic Acids Res,* 43, D470-478. doi: 10.1093/nar/gku1204.

Chen, H., Zhou, X., Wang, A., Zheng, Y., Gao, Y. and Zhou, J. (2015). Evolutions in fragment-based drug design: the deconstruction-reconstruction approach. D*rug Discov Today,* 20, 105-113. doi: 10.1016/j.drudis.2014.09.015.

Chen, U., Chen, S., Saha, P. and Dutta, A. (1996). p21Cip1/Waf1 disrupts the recruitment of human Fen1 by proliferating-cell nuclear antigen into the DNA replication complex. P*roc Natl Acad Sci U S A,* 93, 11597-11602.

Collins, A. R. (2004). The comet assay for DNA damage and repair: principles, applications, and limitations. M*ol Biotechnol,* 26, 249-261. doi: 10.1385/MB:26:3:249.

Coma, I., Clark, L., Diez, E., Harper, G., Herranz, J., Hofmann, G., Lennon, M., Richmond, N., Valmaseda, M. and Macarron, R. (2009). Process validation and screen reproducibility in high-throughput screening. J *Biomol Screen,* 14, 66-76. doi: 10.1177/1087057108326664.

Cong, Y. S., Wright, W. E. and Shay, J. W. (2002). Human telomerase and its regulation. M*icrobiol Mol Biol Rev,* 66, 407-425, table of contents.

Connor, R. (1992). The Diagnosis, Treatment and Prevention of Animal Trypanosomiasis Under Field Conditions. In P*rogramme for the control of African animal trypanosomiasis and related development: Ecological and technical aspects* Food and Agriculture Organisation of the United Nations.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. N*ucl. Acids Res.,* 16.

Costa, R. M., Chiganças, V., Galhardo, R. a. S., Carvalho, H. and Menck, C. F. (2003). The eukaryotic nucleotide excision repair pathway. B*iochimie,* 85, 1083-1099.

Cowan, J. A. (1998). Metal Activation of Enzymes in Nucleic Acid Biochemistry. C*hem Rev,* 98, 1067-1088.

Cowan, R. and Whittaker, R. G. (1990). Hydrophobicity indices for amino acid residues as determined by high-performance liquid chromatography. P*ept Res,* 3, 75-80.

Cushnie, T. P. and Lamb, A. J. (2005). Antimicrobial activity of flavonoids. I*nt J Antimicrob Agents,* 26, 343-356.

Dalvit, C. (2009). NMR methods in fragment screening: theory and a comparison with other biophysical techniques. D*rug Discov Today,* 14, 1051-1057. doi: 10.1016/j.drudis.2009.07.013.

Damper, D. and Patton, C. L. (1976). Pentamidine transport and sensitivity in brucei-group trypanosomes. J *Protozool,* 23, 349-356.

Dao-pin, S., Söderlind, E., Baase, W. A., Wozniak, J. A., Sauer, U. and Matthews, B. W. (1991). Cumulative site-directed charge-change replacements in bacteriophage T4 lysozyme suggest that long-range electrostatic interactions contribute little to protein stability. J *Mol Biol,* 221, 873-887.

de Boer, J. and Hoeijmakers, J. H. (2000). Nucleotide excision repair and human syndromes. C*arcinogenesis,* 21, 453-460.

de Koning, H. P. (2008). Ever-increasing complexities of diamidine and arsenical crossresistance in African trypanosomes. T*rends Parasitol,* 24, 345-349. doi: 10.1016/j.pt.2008.04.006.

de Raadt, P. (2005). The history of sleeping sickness.

Delespaux, V., Geysen, D., Van den Bossche, P. and Geerts, S. (2008). Molecular tools for the rapid detection of drug resistance in animal trypanosomes. T*rends Parasitol,* 24, 236-242. doi: 10.1016/j.pt.2008.02.006.

Desquesnes, M. and Gardiner, P. R. (1993). [Epidemiology of bovine trypanosomiasis (Trypanosoma vivax) in French Guiana]. R*ev Elev Med Vet Pays Trop,* 46, 463-470.

Devi, K. P., Rajavel, T., Habtemariam, S., Nabavi, S. F. and Nabavi, S. M. (2015). Molecular mechanisms underlying anticancer effects of myricetin. L*ife Sci,* 142, 19-25. doi: 10.1016/j.lfs.2015.10.004.

Devos, J. M., Tomanicek, S. J., Jones, C. E., Nossal, N. G. and Mueser, T. C. (2007). Crystal structure of bacteriophage T4 5' nuclease in complex with a branched DNA reveals how flap endonuclease-1 family nucleases bind their substrates. J *Biol Chem,* 282, 31713-31724. doi: 10.1074/jbc.M703209200.

Dianova, I., Bohr, V. A. and Dianov, G. L. (2001). Interaction of human AP endonuclease 1 with flap endonuclease 1 and proliferating cell nuclear antigen involved in long-patch base excision repair. B*iochemistry,* 40, 12639-12644.

DNDi (2015). Oxaborole SCYX-7158 (HAT). In H*uman African Trypanosomiasis: Translation,* Vol. 2016.

DNDi (2016). DNDi Strategy.

Dodson, H. C., Lyda, T. A., Chambers, J. W., Morris, M. T., Christensen, K. A. and Morris, J. C. (2011). Quercetin, a fluorescent bioflavanoid, inhibits Trypanosoma brucei hexokinase 1. E*xp Parasitol,* 127, 423-428. doi: 10.1016/j.exppara.2010.10.011.

Dorjsuren, D., Kim, D., Maloney, D. J., Wilson, D. M. and Simeonov, A. (2011). Complementary non-radioactive assays for investigation of human flap endonuclease 1 activity. N*ucleic Acids Res,* 39, e11. doi: gkq1082 [pii]

10.1093/nar/gkq1082.

Dreesen, O., Li, B. and Cross, G. A. (2007). Telomere structure and function in trypanosomes: a proposal. Na*t Rev Microbiol, 5*, 70-75. doi: 10.1038/nrmicro1577.

Duxin, J. P., Dao, B., Martinsson, P., Rajala, N., Guittat, L., Campbell, J. L., Spelbrink, J. N. and Stewart, S. A. (2009). Human Dna2 is a nuclear and mitochondrial DNA maintenance protein. Mo*l Cell Biol, 2*9, 4274-4282. doi: MCB.01834-08 [pii]

10.1128/MCB.01834-08.

FAO (2015). The Disease. In Pro*gramme Against African Trypanosomiasis (PAAT).*

*F*ederhen, S. (2012). The NCBI Taxonomy database. Nuc*leic Acids Res, 40*, D136-143. doi: 10.1093/nar/gkr1178.

Feng, M., Patel, D., Dervan, J. J., Ceska, T., Suck, D., Haq, I. and Sayers, J. R. (2004). Roles of divalent metal ions in flap endonuclease-substrate interactions. Nat *Struct Mol Biol, 11*, 450-456. doi: 10.1038/nsmb754.

Fernandez, M. T., Mira, M. L., Florêncio, M. H. and Jennings, K. R. (2002). Iron and copper chelation by flavonoids: an electrospray mass spectrometry study. J I*norg Biochem, 92*, 105-111.

Finger, L. D., Atack, J. M., Tsutakawa, S., Classen, S., Tainer, J., Grasby, J. and Shen, B. (2012). The wonders of flap endonucleases: structure, function, mechanism and regulation. Sub*cell Biochem, 62*, 301-326. doi: 10.1007/978-94-007-4572-8_16.

Finger, L. D., Blanchard, M. S., Theimer, C. A., Sengerova, B., Singh, P., Chavez, V., Liu, F., Grasby, J. A. and Shen, B. (2009). The 3'-flap pocket of human flap endonuclease 1 is critical for substrate binding and catalysis. J B*iol Chem, 28*4, 22184-22194. doi: 10.1074/jbc.M109.015065.

Fox, K. R., Sansom, C. E. and Stevens, M. F. (1990). Footprinting studies on the sequence-selective binding of pentamidine to DNA. FEB*S Lett, 26*6, 150-154.

Frearson, J. A., Brand, S., McElroy, S. P., Cleghorn, L. A., Smid, O., Stojanovski, L., Price, H. P., Guther, M. L., Torrie, L. S., Robinson, D. A., Hallyburton, I., Mpamhanga, C. P., Brannigan, J. A., Wilkinson, A. J., Hodgkinson, M., Hui, R., Qiu, W., Raimi, O. G., van Aalten, D. M., Brenk, R., Gilbert, I. H., Read, K. D., Fairlamb, A. H., Ferguson, M. A., Smith, D. F. and Wyatt, P. G. (2010). N-myristoyltransferase inhibitors as new leads to treat sleeping sickness. Nat*ure, 46*4, 728-732. doi: 10.1038/nature08893.

Friedrich-Heineken, E., Henneke, G., Ferrari, E. and Hubscher, U. (2003). The acetylatable lysines of human Fen1 are important for endo- and exonuclease activities. J M*ol Biol, 32*8, 73-84.

Friedrich-Heineken, E. and Hubscher, U. (2004). The Fen1 extrahelical 3'-flap pocket is conserved from archaea to human and regulates DNA substrate specificity. Nuc*leic Acids Res, 32*, 2520-2528. doi: 10.1093/nar/gkh576.

Friedrich-Heineken, E. and Hübscher, U. (2004). The Fen1 extrahelical 3'-flap pocket is conserved from archaea to human and regulates DNA substrate specificity. Nuc*leic Acids Res, 32*, 2520-2528. doi: 10.1093/nar/gkh576.

Frosina, G., Fortini, P., Rossi, O., Carrozzino, F., Raspaglio, G., Cox, L. S., Lane, D. P., Abbondandolo, A. and Dogliotti, E. (1996). Two pathways for base excision repair in mammalian cells. J B*iol Chem, 27*1, 9573-9578.

Fukushima, S., Itaya, M., Kato, H., Ogasawara, N. and Yoshikawa, H. (2007). Reassessment of the in vivo functions of DNA polymerase I and RNase H in bacterial cell growth. J B*acteriol, 18*9, 8575-8583. doi: JB.00653-07 [pii] 10.1128/JB.00653-07.

Garforth, S. J., Patel, D., Feng, M. and Sayers, J. R. (2001). Unusually wide co-factor tolerance in a metalloenzyme; divalent metal ions modulate endo-exonuclease activity in T5 exonuclease. Nucl*eic Acids Res, 29,* 2772-2779.

Garg, P., Stith, C. M., Sabouri, N., Johansson, E. and Burgers, P. M. (2004). Idling by DNA polymerase delta maintains a ligatable nick during lagging-strand DNA replication. Gene*s Dev, 18,* 2764-2773. doi: 10.1101/gad.1252304.

Gary, R., Ludwig, D. L., Cornelius, H. L., MacInnes, M. A. and Park, M. S. (1997). The DNA repair endonuclease XPG binds to proliferating cell nuclear antigen (PCNA) and shares sequence elements with the PCNA-binding regions of FEN-1 and cyclin-dependent kinase inhibitor p21. J Bi*ol Chem, 272*, 24522-24529.

Gasteiger E., H. C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A (2005). Prot*ein Identification and Analysis Tools on the ExPASy Server. In* The *Proteomics Protocols Handbook (ed*. Walker, J.), pp. 571-607. Humana Press.

Gibson, W. C. (2005). The SRA gene: the key to understanding the nature of Trypanosoma brucei rhodesiense. Para*sitology, 131*, 143-150.

Gilson, E. and Géli, V. (2007). How telomeres are replicated. Nat *Rev Mol Cell Biol, 8,* 825-838. doi: 10.1038/nrm2259.

Giugliarelli, A., Paolantoni, M., Morresi, A. and Sassi, P. (2012). Denaturation and preservation of globular proteins: the role of DMSO. J Ph*ys Chem B, 116*, 13361-13367. doi: 10.1021/jp308655p.

González, V. M., Pérez, J. M. and Alonso, C. (1997). The berenil ligand directs the DNA binding of the cytotoxic drug Pt-berenil. J In*org Biochem, 68,* 283-287.

Good, N. E., Winget, G. D., Winter, W., Connolly, T. N., Izawa, S. and Singh, R. M. (1966). Hydrogen ion buffers for biological research. Bioc*hemistry, 5,* 467-477.

Goulian, M., Richards, S. H., Heard, C. J. and Bigsby, B. M. (1990). Discontinuous DNA synthesis by purified mammalian proteins. J Bi*ol Chem, 265*, 18461-18471.

Grant, S. G., Jessee, J., Bloom, F. R. and Hanahan, D. (1990). Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. Proc *Natl Acad Sci U S A, 87,* 4645-4649.

Grazú, V., Silber, A. M., Moros, M., Asín, L., Torres, T. E., Marquina, C., Ibarra, M. R. and Goya, G. F. (2012). Application of magnetically induced hyperthermia in the model protozoan Crithidia fasciculata as a potential therapy against parasitic infections. Int *J Nanomedicine, 7,* 5351-5360. doi: 10.2147/IJN.S35510.

Grecco, S. o. S., Reimão, J. Q., Tempone, A. G., Sartorelli, P., Cunha, R. L., Romoff, P., Ferreira, M. J., Fávero, O. A. and Lago, J. H. (2012). In vitro antileishmanial and antitrypanosomal activities of flavanones from Baccharis retusa DC. (Asteraceae). Exp *Parasitol, 130*, 141-145. doi: 10.1016/j.exppara.2011.11.002.

Greenfield, N. J. (2006a). Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nat *Protoc, 1,* 2527-2535. doi: 10.1038/nprot.2006.204.

Greenfield, N. J. (2006b). Using circular dichroism spectra to estimate protein secondary structure. Nat *Protoc, 1,* 2876-2890. doi: 10.1038/nprot.2006.202.

Greenwood, D., Finch, R., Davey, P. and Wilcox, M. (2007). Historical Introduction. In Anti*microbial Chemotherapy pp.* 1-10. Oxford University Press, New York.

Grimsley, G. R., Shaw, K. L., Fee, L. R., Alston, R. W., Huyghues-Despointes, B. M., Thurlkill, R. L., Scholtz, J. M. and Pace, C. N. (1999). Increasing protein stability by altering long-range coulombic interactions. Prot*ein Sci, 8,* 1843-1849. doi: 10.1110/ps.8.9.1843.

Grosdidier, A., Zoete, V. and Michielin, O. (2007). EADock: docking of small molecules into protein active sites with a multiobjective evolutionary optimization. Prot*eins, 67,* 1010-1025. doi: 10.1002/prot.21367.

Grosdidier, A., Zoete, V. and Michielin, O. (2011a). Fast docking using the CHARMM force field with EADock DSS. J Co*mput Chem. do*i: 10.1002/jcc.21797.

Grosdidier, A., Zoete, V. and Michielin, O. (2011b). SwissDock, a protein-small molecule docking web service based on EADock DSS. Nucl*eic Acids Res, 39,* W270-277. doi: 10.1093/nar/gkr366.

Hall, B. S., Bot, C. and Wilkinson, S. R. (2011). Nifurtimox activation by trypanosomal type I nitroreductases generates cytotoxic nitrile metabolites. J Bi*ol Chem, 286*, 13088-13095. doi: 10.1074/jbc.M111.230847.

Harner, M. J., Frank, A. O. and Fesik, S. W. (2013). Fragment-based drug discovery using NMR spectroscopy. J Bi*omol NMR, 56,* 65-75. doi: 10.1007/s10858-013-9740-z.

Harrigan, J. A., Opresko, P. L., von Kobbe, C., Kedar, P. S., Prasad, R., Wilson, S. H. and Bohr, V. A. (2003). The Werner syndrome protein stimulates DNA polymerase beta strand displacement synthesis via its helicase activity. J Bi*ol Chem, 278*, 22686-22695. doi: 10.1074/jbc.M213103200.

Harrington, J. J. and Lieber, M. R. (1994a). Functional domains within FEN-1 and RAD2 define a family of structure-specific endonucleases: implications for nucleotide excision repair. Gene*s Dev, 8,* 1344-1355.

Harrington, J. J. and Lieber, M. R. (1994b). The characterization of a mammalian DNA structure-specific endonuclease. EMBO *J, 13,* 1235-1246.

Hartenstine, M. J., Goodman, M. F. and Petruska, J. (2002). Weak strand displacement activity enables human DNA polymerase beta to expand CAG/CTG triplet repeats at strand breaks. J Bi*ol Chem, 277*, 41379-41389. doi: 10.1074/jbc.M207013200.

Hasan, S., Stucki, M., Hassa, P. O., Imhof, R., Gehrig, P., Hunziker, P., Hubscher, U. and Hottiger, M. O. (2001). Regulation of human flap endonuclease-1 activity by acetylation through the transcriptional coactivator p300. Mol *Cell, 7,* 1221-1231.

Henneke, G., Koundrioukoff, S. and Hübscher, U. (2003). Multiple roles for kinases in DNA replication. EMBO *Rep, 4,* 252-256. doi: 10.1038/sj.embor.embor774.

Henry, C. M., Hollville, E. and Martin, S. J. (2013). Measuring apoptosis by microscopy and flow cytometry. Meth*ods, 61,* 90-97. doi: 10.1016/j.ymeth.2013.01.008.

Hodgson, B., Calzada, A. and Labib, K. (2007). Mrc1 and Tof1 regulate DNA replication forks in different ways during normal S phase. Mol *Biol Cell, 18,* 3894-3902. doi: 10.1091/mbc.E07-05-0500.

Hodskinson, M. R., Allen, L. M., Thomson, D. P. and Sayers, J. R. (2007). Molecular interactions of Escherichia coli ExoIX and identification of its associated 3'-5' exonuclease activity. Nucl*eic Acids Res, 35,* 4094-4102. doi: 10.1093/nar/gkm396.

Holdgate, G. A. and Ward, W. H. (2005). Measurements of binding thermodynamics in drug discovery. Drug *Discov Today, 10,* 1543-1550. doi: 10.1016/S1359-6446(05)03610-X.

Horn, D. (2014). Antigenic variation in African trypanosomes. Mol *Biochem Parasitol, 195*, 123-129. doi: 10.1016/j.molbiopara.2014.05.001.

Horton, H., Moran, L., Scrimgeor, K., Perry, M. and Rawn, J. (2006). Chemical Modes of Enzymatic Catalysis. In Prin*ciples of Biochemistry (ed*. Carlson, G.), pp. 162-167. Pearson Education Ltd, London.

Hossain, M. and Suresh Kumar, G. (2009). DNA intercalation of methylene blue and quinacrine: new insights into base and sequence specificity from structural and thermodynamic studies with polynucleotides. Mol *Biosyst, 5,* 1311-1322. doi: 10.1039/b909563b.

Imboden, M. A., Laird, P. W., Affolter, M. and Seebeck, T. (1987). Transcription of the intergenic regions of the tubulin gene cluster of Trypanosoma brucei: evidence for a polycistronic transcription unit in a eukaryote. Nucl*eic Acids Res, 15,* 7357-7368.

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

Jackson, M. and Mantsch, H. H. (1991). Beware of proteins in DMSO. Bioc*him Biophys Acta, 107*8, 231-235.

Jackson, S. P. and Bartek, J. (2009). The DNA-damage response in human biology and disease. Natu*re, 461*, 1071-1078. doi: 10.1038/nature08467.

Jacobs, R. T., Nare, B., Wring, S. A., Orr, M. D., Chen, D., Sligar, J. M., Jenks, M. X., Noe, R. A., Bowling, T. S., Mercer, L. T., Rewerts, C., Gaukel, E., Owens, J., Parham, R., Randolph, R., Beaudet, B., Bacchi, C. J., Yarlett, N., Plattner, J. J., Freund, Y., Ding, C., Akama, T., Zhang, Y. K., Brun, R., Kaiser, M., Scandale, I. and Don, R. (2011). SCYX-7158, an orally-active benzoxaborole for the treatment of stage 2 human African trypanosomiasis. PLoS *Negl Trop Dis, 5,* e1151. doi: 10.1371/journal.pntd.0001151.

Jadhav, A., Ferreira, R. S., Klumpp, C., Mott, B. T., Austin, C. P., Inglese, J., Thomas, C. J., Maloney, D. J., Shoichet, B. K. and Simeonov, A. (2010). Quantitative analyses of aggregation, autofluorescence, and reactivity artifacts in a screen for inhibitors of a thiol protease. J Me*d Chem, 53,* 37-51. doi: 10.1021/jm901070c.

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

Jiricny, J. (2006). The multifaceted mismatch-repair system. Nat *Rev Mol Cell Biol, 7,* 335-346. doi: 10.1038/nrm1907.

Johnson, C. M. (2013). Differential scanning calorimetry as a tool for protein folding and stability. Arch *Biochem Biophys, 531*, 100-109. doi: 10.1016/j.abb.2012.09.008.

Jones, T. W. and Dávila, A. M. (2001). Trypanosoma vivax--out of Africa. Trend*s Parasitol, 17,* 99-101.

Joyce, C. M. and Grindley, N. D. (1984). Method for determining whether a gene of Escherichia coli is essential: application to the polA gene. J Ba*cteriol, 158*, 636-643.

Kabani, S., Fenn, K., Ross, A., Ivens, A., Smith, T. K., Ghazal, P. and Matthews, K. (2009). Genome-wide expression profiling of in vivo-derived bloodstream parasite stages and dynamic analysis of mRNA alterations during synchronous differentiation in Trypanosoma brucei. BMC *Genomics, 10,* 427. doi: 10.1186/1471-2164-10-427.

Kazak, L., Reyes, A., He, J., Wood, S. R., Brea-Calvo, G., Holen, T. T. and Holt, I. J. (2013). A cryptic targeting signal creates a mitochondrial FEN1 isoform with tailed R-Loop binding properties. PLoS *One, 8,* e62340. doi: 10.1371/journal.pone.0062340.

Keating, J., Yukich, J. O., Sutherland, C. S., Woods, G. and Tediosi, F. (2015). Human African trypanosomiasis prevention, treatment and control costs: a systematic review. Acta *Trop, 150*, 4-13. doi: 10.1016/j.actatropica.2015.06.003.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. and Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat *Protoc, 10,* 845-858. doi: 10.1038/nprot.2015.053.

Kelly, S. M., Jess, T. J. and Price, N. C. (2005). How to study proteins by circular dichroism. Bioc*him Biophys Acta, 175*1, 119-139. doi: 10.1016/j.bbapap.2005.06.005.

Kennedy, P. G. (2013). Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). Lanc*et Neurol, 12,* 186-194. doi: 10.1016/S1474-4422(12)70296-X.

Kieft, R., Capewell, P., Turner, C. M., Veitch, N. J., MacLeod, A. and Hajduk, S. (2010). Mechanism of Trypanosoma brucei gambiense (group 1) resistance to human trypanosome lytic factor. Proc *Natl Acad Sci U S A, 107*, 16137-16141. doi: 10.1073/pnas.1007074107.

Kinabo, L. D. (1993). Pharmacology of existing drugs for animal trypanosomiasis. Acta *Trop, 54,* 169-183.

Klett, R. P., Cerami, A. and Reich, E. (1968). Exonuclease VI, a new nuclease activity associated with E. coli DNA polymerase. Proc *Natl Acad Sci U S A, 60,* 943-950.

Klug, D. M., Gelb, M. H. and Pollastri, M. P. (2016). Repurposing strategies for tropical disease drug discovery. Bioo*rg Med Chem Lett. do*i: 10.1016/j.bmcl.2016.03.103.

Kucherlapati, M., Yang, K., Kuraguchi, M., Zhao, J., Lia, M., Heyer, J., Kane, M. F., Fan, K., Russell, R., Brown, A. M., Kneitz, B., Edelmann, W., Kolodner, R. D., Lipkin, M. and Kucherlapati, R. (2002). Haploinsufficiency of Flap endonuclease (Fen1) leads to rapid tumor progression. Proc *Natl Acad Sci U S A, 99,* 9924-9929. doi: 152321699 [pii] 10.1073/pnas.152321699.

Kumar, S. and Subramanian, S. (2002). Mutation rates in mammalian genomes. Proc *Natl Acad Sci U S A, 99,* 803-808. doi: 10.1073/pnas.022629899.

Kuzoe, F. A. (1993). Current situation of African trypanosomiasis. Acta *Trop, 54,* 153-162.

Kyte, J. and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. J Mol *Biol, 157,* 105-132.

Labib, K. and Hodgson, B. (2007). Replication fork barriers: pausing for a break or stalling for time? EMBO *Rep, 8, 3*46-353. doi: 10.1038/sj.embor.7400940.

LaCount, D. J., Bruse, S., Hill, K. L. and Donelson, J. E. (2000). Double-stranded RNA interference in Trypanosoma brucei using head-to-head promoters. Mol B*iochem Parasitol, 111,* 67-76.

Lakowicz, J. R. (2006). Princ*iples of fluorescence spectroscopy, Third* edn. Springer Science and Business Media, LLC.

Langousis, G. and Hill, K. L. (2014). Motility and more: the flagellum of Trypanosoma brucei. Nat R*ev Microbiol, 12,* 505-518. doi: 10.1038/nrmicro3274.

Lanteri, C. A., Stewart, M. L., Brock, J. M., Alibu, V. P., Meshnick, S. R., Tidwell, R. R. and Barrett, M. P. (2006). Roles for the Trypanosoma brucei P2 transporter in DB75 uptake and resistance. Mol P*harmacol, 70,* 1585-1592. doi: 10.1124/mol.106.024653.

Lavinder, J. J., Hari, S. B., Sullivan, B. J. and Magliery, T. J. (2009). High-throughput thermal scanning: a general, rapid dye-binding thermal shift screen for protein engineering. J Am *Chem Soc, 131,* 3794-3795. doi: 10.1021/ja8049063.

Le Gall, T., Romero, P. R., Cortese, M. S., Uversky, V. N. and Dunker, A. K. (2007). Intrinsic disorder in the Protein Data Bank. J Bio*mol Struct Dyn, 24,* 325-342. doi: 10.1080/07391102.2007.10507123.

Leach, T. M. and Roberts, C. J. (1981). Present status of chemotherapy and chemoprophylaxis of animal trypanosomiasis in the Eastern hemisphere. Pharm*acol Ther, 13,* 91-147.

Lee, K. H., Kim, D. W., Bae, S. H., Kim, J. A., Ryu, G. H., Kwon, Y. N., Kim, K. A., Koo, H. S. and Seo, Y. S. (2000). The endonuclease activity of the yeast Dna2 enzyme is essential in vivo. Nucle*ic Acids Res, 28,* 2873-2881.

Lee, O. H., Kim, H., He, Q., Baek, H. J., Yang, D., Chen, L. Y., Liang, J., Chae, H. K., Safari, A., Liu, D. and Songyang, Z. (2011). Genome-wide YFP fluorescence complementation screen identifies new regulators for telomere signaling in human cells. Mol C*ell Proteomics, 10,* M110 001628. doi: 10.1074/mcp.M110.001628.

Li, X., Li, J., Harrington, J., Lieber, M. R. and Burgers, P. M. (1995). Lagging strand DNA synthesis at the eukaryotic replication fork involves binding and stimulation of FEN-1 by proliferating cell nuclear antigen. J Bio*l Chem, 270,* 22109-22112.

Lieber, M. R. (1997). The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. Bioes*says, 19,* 233-240. doi: 10.1002/bies.950190309.

Liu, Y., Kao, H. I. and Bambara, R. A. (2004). Flap endonuclease 1: a central component of DNA metabolism. Annu *Rev Biochem, 73,* 589-615. doi: 10.1146/annurev.biochem.73.012803.092453.

Lo, M. C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M. and Ellestad, G. (2004). Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. Anal *Biochem, 332,* 153-159. doi: 10.1016/j.ab.2004.04.031.

Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., Zipursky, S. L. and Darnell, J. (2004). Molecular Cell Biology. pp. 131-137. W.H. Freeman and Company.

Lopes, M., Cotta-Ramusino, C., Pellicioli, A., Liberi, G., Plevani, P., Muzi-Falconi, M., Newlon, C. S. and Foiani, M. (2001). The DNA replication checkpoint response stabilizes stalled replication forks. Natur*e, 412,* 557-561. doi: 10.1038/35087613.

Louis-Jeune, C., Andrade-Navarro, M. A. and Perez-Iratxeta, C. (2012). Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. Prote*ins, 80,* 374-381. doi: 10.1002/prot.23188.

Lüscher, A., Nerima, B. and Mäser, P. (2006). Combined contribution of TbAT1 and TbMRPA to drug resistance in Trypanosoma brucei. Mol B*iochem Parasitol, 150,* 364-366. doi: 10.1016/j.molbiopara.2006.07.010.

Ma, Z., Wang, G., Cui, L. and Wang, Q. (2015). Myricetin Attenuates Depressant-Like Behavior in Mice Subjected to Repeated Restraint Stress. Int J *Mol Sci, 16,* 28377-28385. doi: 10.3390/ijms161226102.

Ma, Z. G., Wang, J., Jiang, H., Liu, T. W. and Xie, J. X. (2007). Myricetin reduces 6-hydroxydopamine-induced dopamine neuron degeneration in rats. Neuro*report, 18,* 1181-1185. doi: 10.1097/WNR.0b013e32821c51fe.

Mamani-Matsuda, M., Rambert, J., Malvy, D., Lejoly-Boisseau, H., Daulouède, S., Thiolat, D., Coves, S., Courtois, P., Vincendeau, P. and Mossalayi, M. D. (2004). Quercetin induces apoptosis of Trypanosoma brucei gambiense and decreases the proinflammatory response of human macrophages. Antim*icrob Agents Chemother, 48,* 924-929.

Marques, C. A., Tiengwe, C., Lemgruber, L., Damasceno, J. D., Scott, A., Paape, D., Marcello, L. and McCulloch, R. (2016). Diverged composition and regulation of the Trypanosoma brucei origin recognition complex that mediates DNA replication initiation. Nucle*ic Acids Res. doi*: 10.1093/nar/gkw147.

Mashalidis, E. H., Śledź, P., Lang, S. and Abell, C. (2013). A three-stage biophysical screening cascade for fragment-based drug discovery. Nat P*rotoc, 8, 2*309-2324. doi: 10.1038/nprot.2013.130.

Matovu, E., Stewart, M. L., Geiser, F., Brun, R., Mäser, P., Wallace, L. J., Burchmore, R. J., Enyaru, J. C., Barrett, M. P., Kaminsky, R., Seebeck, T. and de Koning, H. P. (2003). Mechanisms of arsenical and diamidine uptake and resistance in Trypanosoma brucei. Eukar*yot Cell, 2, 1*003-1008.

Matsuzaki, Y., Adachi, N. and Koyama, H. (2002). Vertebrate cells lacking FEN-1 endonuclease are viable but hypersensitive to methylating agents and H2O2. Nucle*ic Acids Res, 30,* 3273-3277.

McManus, K. J., Barrett, I. J., Nouhi, Y. and Hieter, P. (2009). Specific synthetic lethal killing of RAD54B-deficient human colorectal cancer cells by FEN1 silencing. Proc *Natl Acad Sci U S A, 106,* 3276-3281. doi: 10.1073/pnas.0813414106.

McWhirter, C., Tonge, M., Plant, H., Hardern, I., Nissink, W. and Durant, S. T. (2013). Development of a high-throughput fluorescence polarization DNA cleavage assay for the identification of FEN1 inhibitors. J Bio*mol Screen, 18,* 567-575. doi: 10.1177/1087057113476551.

Melaku, A. and Birasa, B. (2013). Drugs and Drug Resistance in African Animal Trypanosomosis: A Review. In Europ*ean Journal of Applied Sciences, Vol*. 5 pp. 84-91.

Meyskens, F. L. and Gerner, E. W. (1999). Development of difluoromethylornithine (DFMO) as a chemoprevention agent. Clin *Cancer Res, 5, 9*45-951.

Mirkin, E. V. and Mirkin, S. M. (2007). Replication fork stalling at natural impediments. Micro*biol Mol Biol Rev, 71,* 13-35. doi: 10.1128/MMBR.00030-06.

Moritoh, S., Miki, D., Akiyama, M., Kawahara, M., Izawa, T., Maki, H. and Shimamoto, K. (2005). RNAi-mediated silencing of OsGEN-L (OsGEN-like), a new member of the RAD2/XPG nuclease family, causes male sterility by defect of microspore development in rice. Plant *Cell Physiol, 46,* 699-715. doi: 10.1093/pcp/pci090.

Motyka, S. A. and Englund, P. T. (2004). RNA interference for analysis of gene function in trypanosomatids. Curr *Opin Microbiol, 7, 3*62-368. doi: 10.1016/j.mib.2004.06.004.

Murante, R. S., Huang, L., Turchi, J. J. and Bambara, R. A. (1994). The calf 5'- to 3'-exonuclease is also an endonuclease with both activities dependent on primers annealed upstream of the point of cleavage. J Bio*l Chem, 269,* 1191-1196.

Murante, R. S., Rust, L. and Bambara, R. A. (1995). Calf 5' to 3' exo/endonuclease must slide from a 5' end of the substrate to perform structure-specific cleavage. J Bio*l Chem, 270,* 30377-30383.

NC3Rs (2016). Decision tree: How much blood does a mouse have? In Blood *Sampling: Mouse.*

NCI (2015). Cancer Terms. In NCI D*ictionary.*

*Neg*ritto, M. C., Qiu, J., Ratay, D. O., Shen, B. and Bailis, A. M. (2001). Novel function of Rad27 (FEN-1) in restricting short-sequence recombination. Mol C*ell Biol, 21,* 2349-2358. doi: 10.1128/MCB.21.7.2349-2358.2001.

Niesen, F. H., Berglund, H. and Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat P*rotoc, 2, 2*212-2221. doi: 10.1038/nprot.2007.321.

Oli, M. W., Cotlin, L. F., Shiflett, A. M. and Hajduk, S. L. (2006). Serum resistance-associated protein blocks lysosomal targeting of trypanosome lytic factor in Trypanosoma brucei. Eukar*yot Cell, 5, 1*32-139. doi: 10.1128/EC.5.1.132-139.2006.

Ono, K. and Nakane, H. (1990). Mechanisms of inhibition of various cellular DNA and RNA polymerases by several flavonoids. J Bio*chem, 108,* 609-613.

Oredsson, S., Anehus, S. and Heby, O. (1980). Inhibition of cell proliferation by DL-alpha-difluoromethylornithine, a catalytic irreversible inhibitor of ornithine decarboxylase. Acta *Chem Scand B, 34,* 457-458.

Osório, A. L., Madruga, C. R., Desquesnes, M., Soares, C. O., Ribeiro, L. R. and Costa, S. C. (2008). Trypanosoma (Duttonella) vivax: its biology, epidemiology, pathogenesis, and introduction in the New World--a review. Mem I*nst Oswaldo Cruz, 103,* 1-13.

Otte, M. J., Abuabara, J. Y. and Wells, E. A. (1994). Trypanosoma vivax in Colombia: epidemiology and production losses. Trop *Anim Health Prod, 26,* 146-156.

Pace, C. N., Alston, R. W. and Shaw, K. L. (2000). Charge-charge interactions influence the denatured state ensemble and contribute to protein stability. Prote*in Sci, 9, 1*395-1398. doi: 10.1110/ps.9.7.1395.

Paine, M. F., Wang, M. Z., Generaux, C. N., Boykin, D. W., Wilson, W. D., De Koning, H. P., Olson, C. A., Pohlig, G., Burri, C., Brun, R., Murilla, G. A., Thuita, J. K., Barrett, M. P. and Tidwell, R. R. (2010). Diamidines for human African trypanosomiasis. Curr *Opin Investig Drugs, 11,* 876-883.

Panda, H., Jaiswal, A. S., Corsino, P. E., Armas, M. L., Law, B. K. and Narayan, S. (2009). Amino acid Asp181 of 5'-flap endonuclease 1 is a useful target for chemotherapeutic development. Bioch*emistry, 48,* 9952-9958. doi: 10.1021/bi9010754.

Pantoliano, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B. A., Lane, P. and Salemme, F. R. (2001). High-density miniaturized thermal shift assays as a general strategy for drug discovery. J Bio*mol Screen, 6, 4*29-440. doi: 10.1089/108705701753364922.

Pascucci, B., D'Errico, M., Parlanti, E., Giovannini, S. and Dogliotti, E. (2011). Role of nucleotide excision repair proteins in oxidative DNA damage repair: an updating. Bioch*emistry (Mosc), 76,* 4-15. doi: BCM76010008 [pii].

Patel, N., Atack, J. M., Finger, L. D., Exell, J. C., Thompson, P., Tsutakawa, S., Tainer, J. A., Williams, D. M. and Grasby, J. A. (2012). Flap endonucleases pass 5'-flaps through a flexible arch using a disorder-thread-order mechanism to confer specificity for free 5'-ends. Nucle*ic Acids Res, 40,* 4507-4519. doi: 10.1093/nar/gks051.

Pearce, L. (1921). STUDIES ON THE TREATMENT OF HUMAN TRYPANOSOMIASIS WITH TRYPARSAMIDE (THE SODIUM SALT OF N-PHENYLGLYCINEAMIDE-p-ARSONIC ACID). J Exp *Med, 34,* 1-104.

Phillips, M. A. and Wang, C. C. (1987). A Trypanosoma brucei mutant resistant to alpha-difluoromethylornithine. Mol B*iochem Parasitol, 22,* 9-17.

Pimentel, D. e. S., Ramos, C. A., Ramos, R. A., de Araújo, F. R., Borba, M. L., Faustino, M. A. and Alves, L. C. (2012). First report and molecular characterization of Trypanosoma vivax in cattle from state of Pernambuco, Brazil. Vet P*arasitol, 185,* 286-289. doi: 10.1016/j.vetpar.2011.10.019.

Plouffe, D., Brinker, A., McNamara, C., Henson, K., Kato, N., Kuhen, K., Nagle, A., Adrián, F., Matzen, J. T., Anderson, P., Nam, T. G., Gray, N. S., Chatterjee, A., Janes, J., Yan, S. F., Trager, R., Caldwell, J. S., Schultz, P. G., Zhou, Y. and Winzeler, E. A. (2008). In silico activity profiling reveals the mechanism of action of antimalarials discovered in a high-throughput screen. Proc *Natl Acad Sci U S A, 105,* 9059-9064. doi: 10.1073/pnas.0802982105.

Prendergast, P. (2001). Use of flavones, coumarins and related compounds to treat infections.

Price, P. A. (1975). The essential role of Ca2+ in the activity of bovine pancreatic deoxyribonuclease. J Bio*l Chem, 250,* 1981-1986.

Qiu, J., Bimston, D. N., Partikian, A. and Shen, B. (2002). Arginine residues 47 and 70 of human flap endonuclease-1 are involved in DNA substrate interactions and cleavage site determination. J Bio*l Chem, 277,* 24659-24666. doi: 10.1074/jbc.M111941200.

Qiu, J., Li, X., Frank, G. and Shen, B. (2001). Cell cycle-dependent and DNA damage-inducible nuclear localization of FEN-1 nuclease is consistent with its dual functions in DNA replication and repair. J Bio*l Chem, 276,* 4901-4908. doi: 10.1074/jbc.M007825200.

Queiroz, R., Benz, C., Fellenberg, K., Hoheisel, J. D. and Clayton, C. (2009). Transcriptome analysis of differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons. BMC G*enomics, 10,* 495. doi: 10.1186/1471-2164-10-495.

Raether, W. and Seidenath, H. (1983). The activity of fexinidazole (HOE 239) against experimental infections with Trypanosoma cruzi, trichomonads and Entamoeba histolytica. Ann T*rop Med Parasitol, 77,* 13-26.

Reuner, B., Vassella, E., Yutzy, B. and Boshart, M. (1997). Cell density triggers slender to stumpy differentiation of Trypanosoma brucei bloodstream forms in culture. Mol B*iochem Parasitol, 90,* 269-280.

Riou, G. and Benard, J. (1980). Berenil induces the complete loss of kinetoplast DNA sequences in Trypa*nosoma equiperdum. In* Bioch*emical and Biophysical Research Communications, Vol*. 96 pp. 350-354.

Robertson, A. B., Klungland, A., Rognes, T. and Leiros, I. (2009). DNA repair in mammalian cells: Base excision repair: the long and short of it. Cell *Mol Life Sci, 66,* 981-993. doi: 10.1007/s00018-009-8736-z.

Robins, P., Pappin, D. J., Wood, R. D. and Lindahl, T. (1994). Structural and functional homology between mammalian DNase IV and the 5'-nuclease domain of Escherichia coli DNA polymerase I. J Bio*l Chem, 269,* 28535-28538.

Sahin, A., Asencio, C., Izotte, J., Pillay, D., Coustou, V., Karembe, H. and Baltz, T. (2014). The susceptibility of Trypanosoma congolense and Trypanosoma brucei to isometamidium chloride and its synthetic impurities. Vet P*arasitol, 203,* 270-275. doi: 10.1016/j.vetpar.2014.04.002.

Sampathi, S., Bhusari, A., Shen, B. and Chai, W. (2009). Human flap endonuclease I is in complex with telomerase and is required for telomerase-mediated telomere maintenance. J Bio*l Chem, 284,* 3682-3690. doi: 10.1074/jbc.M805362200.

Santos, N. C., Figueira-Coelho, J., Martins-Silva, J. and Saldanha, C. (2003). Multidisciplinary utilization of dimethyl sulfoxide: pharmacological, cellular, and molecular aspects. Biochem *Pharmacol, 65,* 1035-1041.

Sayers, J. R. and Eckstein, F. (1990). Properties of overexpressed phage T5 D15 exonuclease. Similarities with Escherichia coli DNA polymerase I 5'-3' exonuclease. J Biol *Chem, 265,* 18311-18317.

Sayers, J. R. and Eckstein, F. (1991). A single-strand specific endonuclease activity copurifies with overexpressed T5 D15 exonuclease. Nucle*ic Acids Res, 19,* 4127-4132.

Schad, G. J., Allanson, A., Mackay, S. P., Cannavan, A. and Tettey, J. N. (2008). Development and validation of an improved HPLC method for the control of potentially counterfeit isometamidium products. J Pha*rm Biomed Anal, 46,* 45-51. doi: 10.1016/j.jpba.2007.08.026.

Scott, D. A., Moreno, S. N. and Docampo, R. (1995). Ca2+ storage in Trypanosoma brucei: the influence of cytoplasmic pH and importance of vacuolar acidity. Biochem *J, 310* (Pt 3), 789-794.

Scotti, L., Mendonça, F. J., da Silva, M. S. and Scotti, M. T. (2016). Enzymatic Targets in Trypanosoma brucei. Curr *Protein Pept Sci, 17,* 243-259.

Seidl, A., Dávila, A. M. and Silva, R. A. (1999). Estimated financial impact of Trypanosoma vivax on the Brazilian pantanal and Bolivian lowlands. Mem I*nst Oswaldo Cruz, 94,* 269-272.

Sharma, S., Otterlei, M., Sommers, J. A., Driscoll, H. C., Dianov, G. L., Kao, H. I., Bambara, R. A. and Brosh, R. M., Jr. (2004a). WRN helicase and FEN-1 form a complex upon replication arrest and together process branchmigrating DNA structures associated with the replication fork. Mol B*iol Cell, 15,* 734-750. doi: 10.1091/mbc.E03-08-0567.

Sharma, S., Sommers, J. A., Wu, L., Bohr, V. A., Hickson, I. D. and Brosh, R. M., Jr. (2004b). Stimulation of flap endonuclease-1 by the Bloom's syndrome protein. J Biol *Chem, 279,* 9847-9856. doi: 10.1074/jbc.M309898200.

Shen, B., Nolan, J. P., Sklar, L. A. and Park, M. S. (1997). Functional analysis of point mutations in human flap endonuclease-1 active site. Nucle*ic Acids Res, 25,* 3332-3338. doi: gka532 [pii].

Shinozuka, K., Kikuchi, Y., Nishino, C., Mori, A. and Tawata, S. (1988). Inhibitory effect of flavonoids on DNA-dependent DNA and RNA polymerases. Exper*ientia, 44,* 882-885.

Silva, R. A., Morales, G., Eulert, E., Montenegro, A. and Ybañez, R. (1998). Outbreaks of trypanosomosis due to Trypanosoma vivax in cattle in Bolivia. Vet P*arasitol, 76,* 153-157.

Silva, T. M., Olinda, R. G., Rodrigues, C. M., Câmara, A. C., Lopes, F. C., Coelho, W. A., Ribeiro, M. F., Freitas, C. I., Teixeira, M. M. and Batista, J. S. (2013). Pathogenesis of reproductive failure induced by Trypanosoma vivax in experimentally infected pregnant ewes. Vet R*es, 44,* 1. doi: 10.1186/1297-9716-44-1.

Singh, P., Yang, M., Dai, H., Yu, D., Huang, Q., Tan, W., Kernstine, K. H., Lin, D. and Shen, B. (2008). Overexpression and hypomethylation of flap endonuclease 1 gene in breast and other cancers. Mol C*ancer Res, 6, 1*710-1717. doi: 10.1158/1541-7786.mcr-08-0269.

Sokolova, A. Y., Wyllie, S., Patterson, S., Oza, S. L., Read, K. D. and Fairlamb, A. H. (2010). Cross-resistance to nitro drugs and implications for treatment of human African trypanosomiasis. Antim*icrob Agents Chemother, 54,* 2893-2900. doi: 10.1128/AAC.00332-10.

Stanojcic, S., Sollelis, L., Kuk, N., Crobu, L., Balard, Y., Schwob, E., Bastien, P., Pagès, M. and Sterkers, Y. (2016). Single-molecule analysis of DNA replication reveals novel features in the divergent eukaryotes Leishmania and Trypanosoma brucei versus mammalian cells. Sci R*ep, 6, 2*3142. doi: 10.1038/srep23142.

Steverding, D. (2008). The history of African trypanosomiasis. In Paras*ites and vectors, Vol*. 1.

Steverding, D. (2010). The development of drugs for treatment of sleeping sickness: a historical review. Paras*it Vectors, 3, 1*5. doi: 10.1186/1756-3305-3-15.

Stewart, J. A., Chaiken, M. F., Wang, F. and Price, C. M. (2012). Maintaining the end: roles of telomere proteins in end-protection, telomere replication and length regulation. Mutat *Res, 730,* 12-19. doi: 10.1016/j.mrfmmm.2011.08.011.

Stewart, M. L., Burchmore, R. J., Clucas, C., Hertz-Fowler, C., Brooks, K., Tait, A., Macleod, A., Turner, C. M., De Koning, H. P., Wong, P. E. and Barrett, M. P. (2010). Multiple genetic mechanisms lead to loss of functional TbAT1 expression in drug-resistant trypanosomes. Eukar*yot Cell, 9, 3*36-343. doi: 10.1128/EC.00200-09.

Stucki, M., Jónsson, Z. O. and Hübscher, U. (2001). In eukaryotic flap endonuclease 1, the C terminus is essential for substrate binding. J Bio*l Chem, 276,* 7843-7849. doi: 10.1074/jbc.M008829200.

Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. Prote*in Expr Purif, 41,* 207-234.

Studier, F. W. and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol *Biol, 189,* 113-130. doi: 0022-2836(86)90385-2 [pii].

Sultana, A. and Lee, J. E. (2015). Measuring protein-protein and protein-nucleic Acid interactions by biolayer interferometry. Curr *Protoc Protein Sci, 79,* 19.25.11-19.25.26. doi: 10.1002/0471140864.ps1925s79.

Sykes, M. L. and Avery, V. M. (2009). Development of an Alamar Blue viability assay in 384-well format for high throughput whole cell screening of Trypanosoma brucei brucei bloodstream form strain 427. Am J *Trop Med Hyg, 81,* 665-674. doi: 10.4269/ajtmh.2009.09-0015.

Syson, K., Tomlinson, C., Chapados, B. R., Sayers, J. R., Tainer, J. A., Williams, N. H. and Grasby, J. A. (2008). Three metal ions participate in the reaction catalyzed by T5 flap endonuclease. J Bio*l Chem, 283,* 28741-28746. doi: 10.1074/jbc.M801264200.

Teka, I. A., Kazibwe, A. J., El-Sabbagh, N., Al-Salabi, M. I., Ward, C. P., Eze, A. A., Munday, J. C., Mäser, P., Matovu, E., Barrett, M. P. and de Koning, H. P. (2011). The diamidine diminazene aceturate is a substrate for the high-affinity pentamidine transporter: implications for the development of high resistance levels in trypanosomes. Mol P*harmacol, 80,* 110-116. doi: 10.1124/mol.111.071555.

Tercero, J. A. and Diffley, J. F. (2001). Regulation of DNA replication fork progression through damaged DNA by the Mec1/Rad53 checkpoint. Natur*e, 412,* 553-557. doi: 10.1038/35087607.

Thorne, N., Auld, D. S. and Inglese, J. (2010). Apparent activity in high-throughput screening: origins of compound-dependent assay interference. Curr *Opin Chem Biol, 14,* 315-324. doi: 10.1016/j.cbpa.2010.03.020.

Thornton, B. and Basu, C. (2011). Real-time PCR (qPCR) primer design using free online software. Bioch*em Mol Biol Educ, 39,* 145-154. doi: 10.1002/bmb.20461.

Tishkoff, D. X., Boerger, A. L., Bertrand, P., Filosi, N., Gaida, G. M., Kane, M. F. and Kolodner, R. D. (1997). Identification and characterization of Saccharomyces cerevisiae EXO1, a gene encoding an exonuclease that interacts with MSH2. Proc *Natl Acad Sci U S A, 94,* 7487-7492.

Tjernberg, A., Markova, N., Griffiths, W. J. and Hallén, D. (2006). DMSO-related effects in protein characterization. J Bio*mol Screen, 11,* 131-137. doi: 10.1177/1087057105284218.

Tock, M. R., Frary, E., Sayers, J. R. and Grasby, J. A. (2003). Dynamic evidence for metal ion catalysis in the reaction mediated by a flap endonuclease. EMBO *J, 22,* 995-1004. doi: 10.1093/emboj/cdg098.

Torreele, E., Bourdin Trunz, B., Tweats, D., Kaiser, M., Brun, R., Mazué, G., Bray, M. A. and Pécoul, B. (2010). Fexinidazole--a new oral nitroimidazole drug candidate entering clinical development for the treatment of sleeping sickness. PLoS *Negl Trop Dis, 4, e*923. doi: 10.1371/journal.pntd.0000923.

Tsutakawa, S. E., Classen, S., Chapados, B. R., Arvai, A. S., Finger, L. D., Guenther, G., Tomlinson, C. G., Thompson, P., Sarker, A. H., Shen, B., Cooper, P. K., Grasby, J. A. and Tainer, J. A. (2011). Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. Cell, *145,* 198-211. doi: 10.1016/j.cell.2011.03.004.

Tumey, L. N., Bom, D., Huck, B., Gleason, E., Wang, J., Silver, D., Brunden, K., Boozer, S., Rundlett, S., Sherf, B., Murphy, S., Dent, T., Leventhal, C., Bailey, A., Harrington, J. and Bennani, Y. L. (2005). The identification and optimization of a N-hydroxy urea series of flap endonuclease 1 inhibitors. Bioor*g Med Chem Lett, 15,* 277-281. doi: 10.1016/j.bmcl.2004.10.086.

Tumey, L. N., Huck, B., Gleason, E., Wang, J., Silver, D., Brunden, K., Boozer, S., Rundlett, S., Sherf, B., Murphy, S., Bailey, A., Dent, T., Leventhal, C., Harrington, J. and Bennani, Y. L. (2004). The identification and optimization of 2,4-diketobutyric acids as flap endonuclease 1 inhibitors. Bioor*g Med Chem Lett, 14,* 4915-4918. doi: 10.1016/j.bmcl.2004.07.028.

Uilenberg, G. (1998). A field guide for the diagnosis, treatment and prevention of african animal trypanosomiasis. Food and Agriculture Organisation (FAO) of the United Nations, Rome, Italy.

Vallur, A. C. and Maizels, N. (2010). Distinct activities of exonuclease 1 and flap endonuclease 1 at telomeric g4 DNA. PLoS *One, 5, e*8908. doi: 10.1371/journal.pone.0008908.

van Pel, D. M., Barrett, I. J., Shimizu, Y., Sajesh, B. V., Guppy, B. J., Pfeifer, T., McManus, K. J. and Hieter, P. (2013). An evolutionarily conserved synthetic lethal interaction network identifies FEN1 as a broad-spectrum target for anticancer therapeutic development. PLoS *Genet, 9, e*1003254. doi: 10.1371/journal.pgen.1003254.

Vansterkenburg, E. L., Coppens, I., Wilting, J., Bos, O. J., Fischer, M. J., Janssen, L. H. and Opperdoes, F. R. (1993). The uptake of the trypanocidal drug suramin in combination with low-density lipoproteins by Trypanosoma brucei and its possible mode of action. Acta *Trop, 54,* 237-250.

Viadiu, H. and Aggarwal, A. K. (1998). The role of metals in catalysis by the restriction endonuclease BamHI. Nat S*truct Biol, 5, 9*10-916. doi: 10.1038/2352.

Vincent, I. M., Creek, D., Watson, D. G., Kamleh, M. A., Woods, D. J., Wong, P. E., Burchmore, R. J. and Barrett, M. P. (2010). A molecular mechanism for eflornithine resistance in African trypanosomes. PLoS *Pathog, 6, e*1001204. doi: 10.1371/journal.ppat.1001204.

Vipond, I. B., Baldwin, G. S. and Halford, S. E. (1995). Divalent metal ions at the active sites of the EcoRV and EcoRI restriction endonucleases. Bioch*emistry, 34,* 697-704.

Wang, Z. H., Ah Kang, K., Zhang, R., Piao, M. J., Jo, S. H., Kim, J. S., Kang, S. S., Lee, J. S., Park, D. H. and Hyun, J. W. (2010). Myricetin suppresses oxidative stress-induced cell damage via both direct and indirect antioxidant action. Envir*on Toxicol Pharmacol, 29,* 12-18. doi: 10.1016/j.etap.2009.08.007.

Warbrick, E., Lane, D. P., Glover, D. M. and Cox, L. S. (1997). Homologous regions of Fen1 and p21Cip1 compete for binding to the same site on PCNA: a potential mechanism to co-

ordinate DNA replication and repair. Oncog*ene, 14,* 2313-2321. doi: 10.1038/sj.onc.1201072.

Waring, M. J. (1965). Complex formation between ethidium bromide and nucleic acids. J Mol *Biol, 13,* 269-282.

Watkins, T. I. and Woolfe, G. (1952). Effect of changing the quaternizing group on the trypanocidal activity of dimidium bromide. Natur*e, 169,* 506-507.

Wenzler, T., Boykin, D. W., Ismail, M. A., Hall, J. E., Tidwell, R. R. and Brun, R. (2009). New treatment option for second-stage African sleeping sickness: in vitro and in vivo efficacy of aza analogs of DB289. Antim*icrob Agents Chemother, 53,* 4185-4192. doi: 10.1128/AAC.00225-09.

Wenzler, T., Yang, S., Braissant, O., Boykin, D. W., Brun, R. and Wang, M. Z. (2013). Pharmacokinetics, Trypanosoma brucei gambiense efficacy, and time of drug action of DB829, a preclinical candidate for treatment of second-stage human African trypanosomiasis. Antim*icrob Agents Chemother, 57,* 5330-5343. doi: 10.1128/AAC.00398-13.

WHO (2013). Control and surveillance of human African trypanosomiasis: report of a WHO expert committee. In WHO t*echnical report series Worl*d Health Organisation.

WHO (2015a). Global Health Observatory Data Repository. In Human *African Trypanosomiasis, Vol*. 2016.

WHO (2015b). Trypanosomiasis, human African (sleeping sickness). pp. Fact Sheet No 259.

Wlodkowic, D., Skommer, J. and Darzynkiewicz, Z. (2009). Flow cytometry-based apoptosis detection. Metho*ds Mol Biol, 559,* 19-32. doi: 10.1007/978-1-60327-017-5_2.

Wragg, W. R., Washbourn, K., Brown, K. N. and Hill, J. (1958). Metamidium: a new trypanocidal drug. Natur*e, 182,* 1005-1006.

Wyllie, S., Foth, B. J., Kelner, A., Sokolova, A. Y., Berriman, M. and Fairlamb, A. H. (2016). Nitroheterocyclic drug resistance mechanisms in Trypanosoma brucei. J Ant*imicrob Chemother, 71,* 625-634. doi: 10.1093/jac/dkv376.

Xu, Y., Potapova, O., Leschziner, A. E., Grindley, N. D. and Joyce, C. M. (2001). Contacts between the 5' nuclease of DNA polymerase I and its DNA substrate. J Bio*l Chem, 276,* 30167-30177. doi: 10.1074/jbc.M100985200.

Yang, W. (2008). An equivalent metal ion in one- and two-metal-ion catalysis. Nat S*truct Mol Biol, 15,* 1228-1231. doi: 10.1038/nsmb.1502.

Yang, W., Lee, J. Y. and Nowotny, M. (2006). Making and breaking nucleic acids: two-Mg2+-ion catalysis and substrate specificity. Mol C*ell, 22,* 5-13. doi: 10.1016/j.molcel.2006.03.013.

Youssif, F., Mohamed, O., Gameel, A. and Hassan, T. (2010). Efficacy and toxicity of homidium bromide (ethidium) in goats infected with T. vi*vax. In* Small *Ruminant Research, Vol*. 89 pp. 36-41.

Zhang, J. (2012). Biochemical Studies on T5 Exonuclease. In Medic*al School, Vol*. Doctor of Philosophy pp. 154. The University of Sheffield, Sheffield.

Zhang, J. H., Chung, T. D. and Oldenburg, K. R. (1999). A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. J Bio*mol Screen, 4, 6*7-73.

Zhang, K., Ma, Z., Wang, J., Xie, A. and Xie, J. (2011). Myricetin attenuated MPP(+)-induced cytotoxicity by anti-oxidation and inhibition of MKK4 and JNK activation in MES23.5 cells. Neuro*pharmacology, 61,* 329-335. doi: 10.1016/j.neuropharm.2011.04.021.

Zheng, L., Dai, H., Zhou, M., Li, M., Singh, P., Qiu, J., Tsark, W., Huang, Q., Kernstine, K., Zhang, X., Lin, D. and Shen, B. (2007). Fen1 mutations result in autoimmunity, chronic inflammation and cancers. Nat M*ed, 13,* 812-819. doi: nm1599 [pii]

10.1038/nm1599.

Zheng, L. and Shen, B. (2011). Okazaki fragment maturation: nucleases take centre stage. J Mol *Cell Biol, 3, 23*-30. doi: mjq048 [pii] 10.1093/jmcb/mjq048.

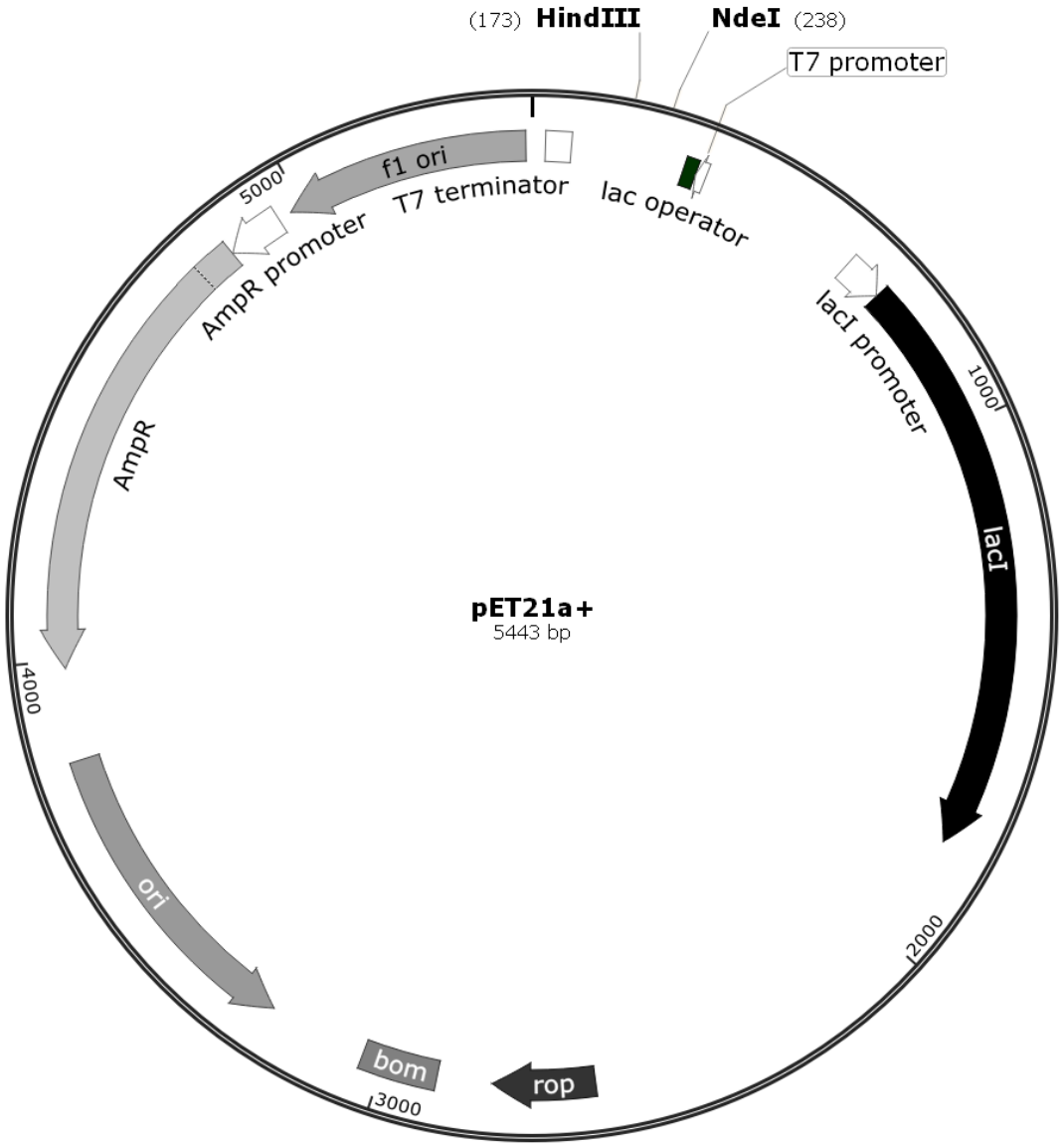
Zheng, L., Zhou, M., Chai, Q., Parrish, J., Xue, D., Patrick, S. M., Turchi, J. J., Yannone, S. M., Chen, D. and Shen, B. (2005). Novel function of the flap endonuclease 1 complex in processing stalled DNA replication forks. EMBO Re*p, 6, 83*-89. doi: 7400313 [pii] 10.1038/sj.embor.7400313.

Zheng, L., Zhou, M., Guo, Z., Lu, H., Qian, L., Dai, H., Qiu, J., Yakubovskaya, E., Bogenhagen, D. F., Demple, B. and Shen, B. (2008). Human DNA2 is a mitochondrial nuclease/helicase for efficient processing of DNA replication and repair intermediates. Mol Cell*, 32, 325*-336. doi: 10.1016/j.molcel.2008.09.024.

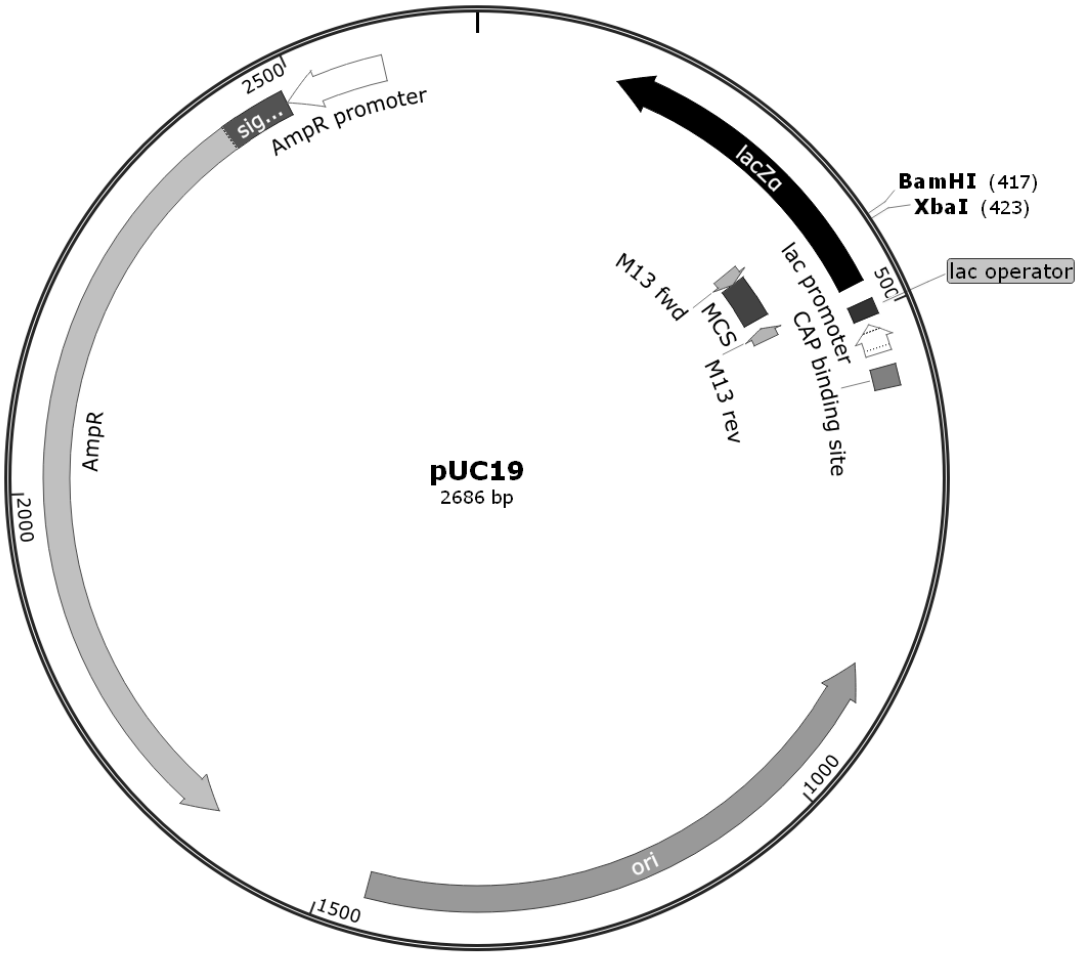
Appendix A – Vector Maps

All vector maps were created using SnapGene® Viewer 2.8.3.

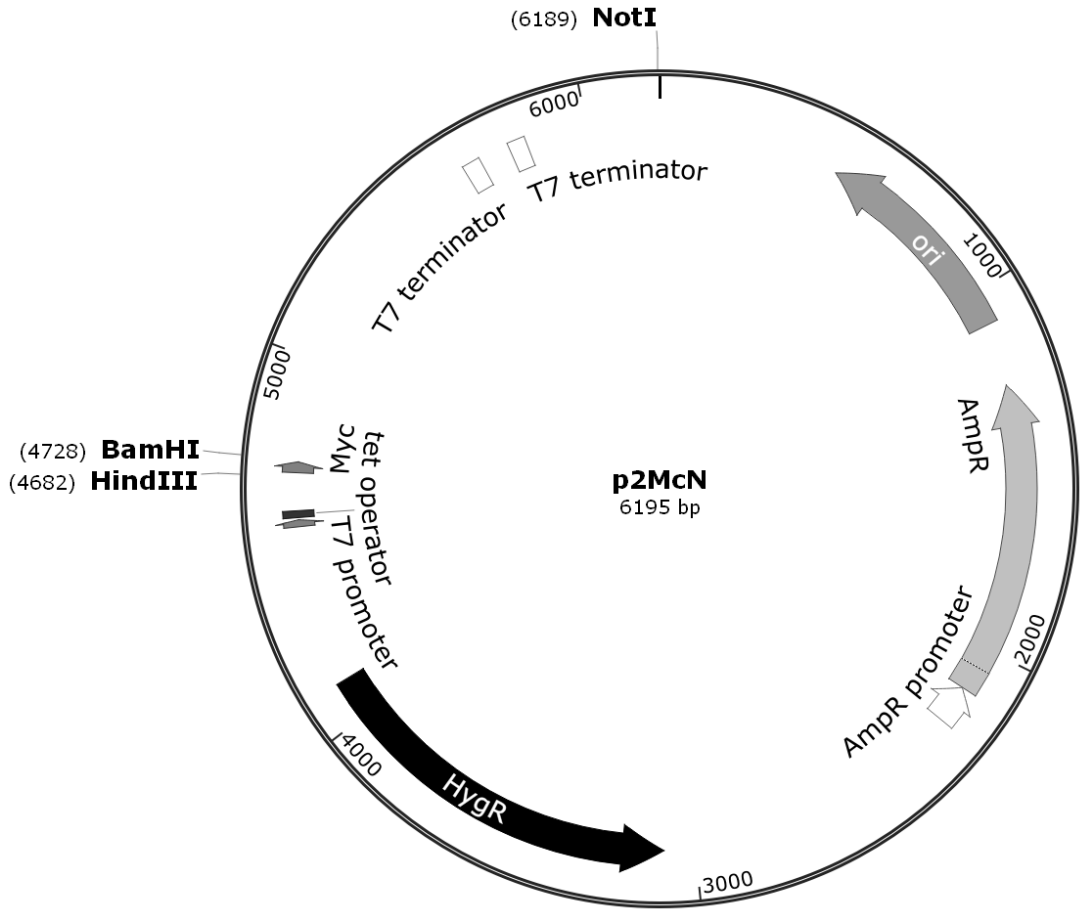
pET21a+ Vector Map



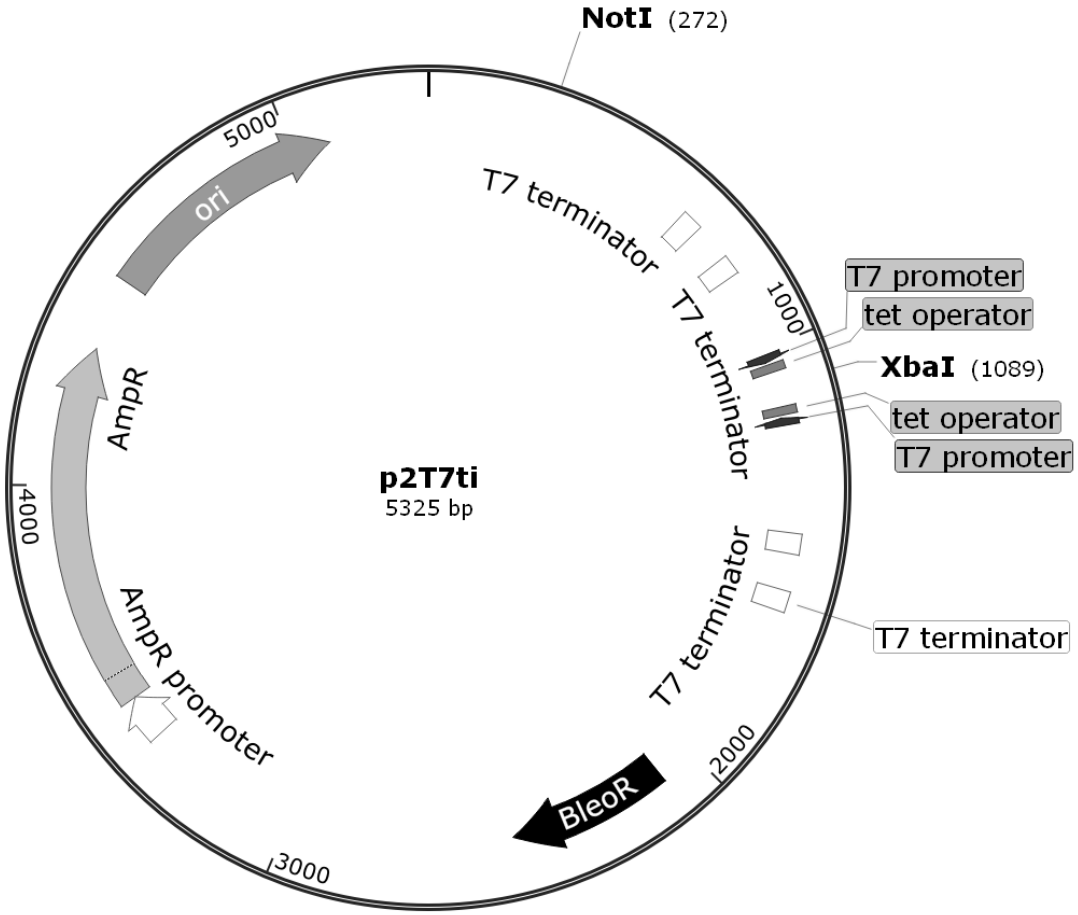
pUC19 Vector Map



p2Mc^{N} Vector Map



p2T7^{ti} Vector Map



Appendix B – Detailed Statistical Analysis for RT-qPCR

Table B-1 Detailed Statistical Analysis Performed on FEN1-specific mRNA Levels

Clone	Time	Homogeneity of Variance	Sphericity	Greenhouse Geisser Correction?	Test Statistic	Post Hoc test
427::fen1 D183K	0	Yes; p = 0.240	Violated; p = 0.0005	Yes	F (3.4, 9.0) = 44.7 p = 0.0005 partial η^2 = 0.944	Tukey
	24	Yes; p = 0.056				Tukey
	48	No; p = 0.023				Games-Howell
	72	Yes; p = 0.770				Tukey
427:: FEN1	0	No; p = 0.008	Not violated; p = 0.344	No	F (9, 21) = 20.93 p = 0.0005 partial η^2 = 0.900	Games-Howell
	24	No; p = 0.032				Games-Howell
	48	No; p = 0.007				Games-Howell
	72	No; p = 0.008				Games-Howell
427‡ FEN1	0	No; p = 0.042	Violated; p = 0.0005	Yes	F (2.1, 6.3) = 2.062 p = 0.204 partial η^2 = 0.407	NA
	24	No; p = 0.004				NA
	48	No; p = 0.004				NA
	72	No; p = 0.009				NA

Transgenic clones were compared with the 427 parental controls using a Repeated Measures mixed ANOVA (IBM SPSS statistics 22).

Appendix C – Line Fitting for Thermal Shift Assays

Table C-1 Melting temperature and R^2 values for protein concentration checks

Protein	Repeat	Concentration	Tm	Mean	SD	R^2
*Tb*FEN	1	5	60.15	59.68	0.48	0.9983
	2		59.19			0.9995
	3		59.69			0.9985
	1	1	60.95	60.82	0.12	0.9967
	2		60.71			0.9966
	3		60.8			0.9965
	1	0.5	60.56	59.92	0.68	0.9936
	2		59.99			0.9963
	3		59.21			0.9898
	1	0.25	60.17	59.14	1.72	0.9898
	2		57.16			0.9879
	3		60.1			0.9905
	1	0.1	60.05	59.05	1.57	0.9869
	2		57.24			0.9624
	3		59.86			0.9900
	1	0.05	59.66	57.50	2.96	0.9621
	2		54.13			0.9227
	3		58.71			0.9789
*h*FEN	1	5	53.36	52.72	1.05	0.9948
	2		51.5			0.994
	3		53.29			0.9957
	1	1	53.29	52.94	0.71	0.9937
	2		52.12			0.9983
	3		53.4			0.9944
	1	0.5	53.4	52.37	1.56	0.9945
	2		50.58			0.9812
	3		53.13			0.9946
	1	0.25	52.88	52.91	0.24	0.9936
	2		52.68			0.7492
	3		53.16			0.9965
	1	0.1	53.87	53.48	0.56	0.9496
	2		Amb			0.1575
	3		53.08			0.9944
	1	0.05	53.12	55.38	4.43	0.9773
	2		60.48			0.6005
	3		52.54			0.9914

Amb; ambiguous

pH Optimisation

Table C-2 Melting temperature and R^2 values for pH comparison

Protein	Repeat	pH	Tm	Mean	SD	R^2
*Tb*FEN	1	6.5	54.73	55.15	0.38	0.9972
	2		55.46			0.9977
	3		55.26			0.9975
	1	7.0	63.85	64.79	0.87	0.9924
	2		65.57			0.9930
	3		64.96			0.9867
	1	7.5	56.44	55.72	1.18	0.9968
	2		56.36			0.9986
	3		54.35			0.9973
	1	8.0	48.91	48.61	0.32	0.9980
	2		48.28			0.9976
	3		48.63			0.9921
	1	8.5	43.13	42.79	0.30	0.9975
	2		42.61			0.9973
	3		42.62			0.9974
	1	9.0	40.16	39.90	0.42	0.9989
	2		40.13			0.9975
	3		39.42			0.9968
	1	9.5	43.19	43.06	0.18	0.9954
	2		43.14			0.9951
	3		42.86			0.9953
*h*FEN	1	6.5	45.28	45.33	0.22	0.9957
	2		45.13			0.9949
	3		45.57			0.9941
	1	7.0	46.94	48.12	1.02	0.9954
	2		48.63			0.9983
	3		48.78			0.9980
	1	7.5	47.23	47.83	0.61	0.9971
	2		47.81			0.9966
	3		48.45			0.9963
	1	8.0	42.93	44.51	1.70	0.9980
	2		44.27			0.9986
	3		46.31			0.9983
	1	8.5	35.41	35.15	0.30	0.9981
	2		35.21			0.9966
	3		34.82			0.9981
	1	9.0	31.22	31.11	0.13	0.9992
	2		31.15			0.9981
	3		30.97			0.9985
	1	9.5	32.96	32.69	0.27	0.9965
	2		32.42			0.9994
	3		32.70			0.9985

Buffer Additives

Table C-3 Melting temperature and R^2 values for additive optimisation

Protein	Repeat	pH	Tm	Mean	SD	R^2
*Tb*FEN	1	None	61.95	61.55	0.48	0.9993
	2		61.02			0.9993
	3		61.68			0.9993
	1	50 mM NaCl	58.60	59.12	0.64	0.9986
	2		58.93			0.9991
	3		59.84			0.9989
	1	100 mM KCl	56.37	56.40	0.03	0.9984
	2		56.43			0.9984
	3		56.40			0.9912
	1	10 mM $MgCl_2$	64.43	64.29	0.67	0.9990
	2		63.56			0.9987
	3		64.87			0.9984
	1	10 mM $MnCl_2$	66.82	66.25	0.56	0.9986
	2		65.70			0.9971
	3		66.23			0.9953
	1	1 mM DTT	63.07	63.09	1.12	0.9971
	2		61.98			0.9977
	3		64.21			0.9950
*h*FEN	1	None	51.00	51.23	0.20	0.9970
	2		51.35			0.9962
	3		51.34			0.9954
	1	50 mM NaCl	49.82	49.62	0.28	0.9912
	2		49.42			0.9905
	3					
	1	100 mM KCl	48.47	49.83	1.26	0.9923
	2		50.08			0.9914
	3		50.95			0.9915
	1	10 mM $MgCl_2$	55.02	54.26	0.84	0.9908
	2		54.41			0.9851
	3		53.36			0.9839
	1	10 mM $MnCl_2$	54.21	54.04	0.21	0.9716
	2		54.10			0.9676
	3		53.81			0.9578
	1	1 mM DTT	50.81	50.95	0.56	0.9874
	2		50.47			0.9844
	3		51.56			0.9895

Table C-4 Melting temperature and R^2 values for DMSO and myricetin

Protein	Repeat	pH	Tm	Mean	SD	R^2
*Tb*FEN	1	Buffer	59.60	59.39	0.33	0.9826
	2		59.55			0.9639
	3		59.01			0.8600
	1	5% DMSO	58.24	57.86	0.65	0.8520
	2		58.23			0.9642
	3		57.11			0.8530
	1	10 μM Myricetin	57.61	57.86	0.73	0.9427
	2		57.29			0.9013
	3		58.69			0.9769
	1	50 μM Myricetin	60.35	60.34	0.05	0.9861
	2		60.29			0.9924
	3		60.39			0.9671
	1	100 μM Myricetin	62.43	62.27	0.14	0.9850
	2		62.19			0.9962
	3		62.18			0.9376
*h*FEN	1	Buffer	54.93	54.73	0.28	0.9268
	2		54.53			0.8619
	1	5% DMSO	51.98	52.13	0.21	0.9534
	2		52.28			0.9754
	1	10 μM Myricetin	52.73	52.52	0.30	0.8626
	2		52.30			0.9868
	1	50 μM Myricetin	55.61	55.03	0.82	0.9818
	2		54.45			0.9710
	1	100 μM Myricetin	57.60	56.66	1.33	0.9621
	2		55.72			0.9879