Thesis submitted for the degree of Doctor of Philosophy to
the Department of Chemistry, University of Sheffield

Mechanistic and Structural studies of the Helical Arch of Flap Endonucleases

Supervisor: Dr. Jane Grasby

Nikesh Patel

1/5/2012

Declaration

Except where specific references have been made to other sources, the work in this thesis is the original work of the author, and it has not been submitted, wholly or in part, for any other degree.

Nikesh Patel

Acknowledgements

I would first and foremost like to thank Dr. Jane Grasby, to whom I am hugely indebted, not only for giving me the opportunity to carry out this work, but for the support and encouragement afforded to me during this project. I would also like to thank the group members both past and present for making the last three years so enjoyable: Blanka, Amanda, Dave (and his encyclopaedic knowledge of all things DNA), Chris (for showing me the ropes), John (for buying a multi dispensing pipette and his wonderful organisation of just about everything around the lab) and of course wee Jack Exell for being a constant source of amusement, intentional or not. The other members of the F floor corridor also deserve a mention, even if most were bad influences and forced me to the pub (I'm looking at you, Matt).

I would like to thank Elaine Frary for synthesising my first oligos and kicking this project off, and Jakob Karaffa of DNA technology for his superb substrates. I would be remiss if I didn't also mention the contribution of Simon Thorpe, who ran all the mass spectra for my substrates, and Robb Hanson, who was called on almost weekly to fix the dHPLC.

A large thank you is owed to Isabel, who shouldered the responsibility of basically feeding, and sometimes dressing me admirably. A huge part was also played in keeping me sane and making sure I didn't become a slave to my PhD. Lastly, I'd like to thank my family for their support, my brothers, and of course my parents, to whom this thesis is dedicated.

Abstract

The work covered in this thesis concerns the helical arch of FEN enzymes, which are important in maintaining genomic stability during DNA replication and repair. This thesis scrutinises the proposed mechanisms of reaction concerning this feature of these enzymes.

The helical arch of FEN enzymes has been hypothesised to facilitate, using the active site at its base, substrate selection, which include those that possess 5' unannealed flaps and gapped flaps. To do this, it is thought that it orientates the scissile phosphate of substrates over the active site. The manner in which this occurs has been a contentious point in the reaction mechanism of this enzyme.

The general mechanisms that have been previously posited are, 1) a tracking mechanism, where the ssDNA flap is bound at the 5’ terminus then passed through the arch, 2) a threading mechanism, where a similar but more passive mechanism is employed, 3) a bind then thread mechanism where binding at the region of bifurcation on DNA substrates occurs prior to threading of the ssDNA flap and lastly, 4) a bind then clamp mechanism that is similar to the latter, but differs in the manner of binding the flap, clamping it either side of the helical arch. To test these models, 5’ biotinylated oligonucleotides were manufactured, which possessed a range of lengths of 5’ flap. One of the key experiments in this work compares the rate of hydrolysis of the scissile phosphate between streptavidin (SA) conjugated oligo (blocked substrate) and a preassembled FEN-substrate (ES) complex to ensure threading/ tracking/ clamping, which is then conjugated to SA, potentially trapping ES in complex. This gave a clear indication of two modes of reaction, where trapped ES complexes react at a biologically relevant rate of reaction, whereas 5’ blocked substrates reacted at a reduced rate of reaction. This ruled out tracking and threading mechanisms.

Competition experiments were used to distinguish between the remaining models. Unlabelled competitor oligos were used to challenge SA conjugated FAM labelled oligos within complex with FEN; where substrates were either trapped or blocked in complex with FEN using SA, or simply free and uninhibited. Successful competition was achieved from ES complexes that were uninhibited and blocked but not those that were trapped. This, along with recently solved crystal structures implied a bind then thread mechanism. This model depicts a scenario where binding of the duplex regions of substrate triggers ordering of the helical arch around the ss flap portion of substrates, assembling the active site ready for catalysis.

Fluorescence anisotropy measurements were also carried out in order to determine the dissociation constant of the synthesised oligos bound to FEN enzymes and validate conditions used in the aforementioned competition experiments. These experiments revealed a small dependence on substrate flap length to binding, and showed stimulation of K_D on the addition of divalent metal ions. This was likely due to shielding of 7/8 conserved carboxylates that ligand divalent metal ions within the active site of the enzyme. Measurements with hFEN1 mutants, also highlighted the fact that initial binding of substrates is not affected by the helical arch of FEN enzymes.

Mutagenesis studies were also carried out. Leucine residues were mutated to proline along the helical arch of human FEN, to disrupt proper alpha helical structure and any disorder-to-order transitions. This gave an extremely deficient phenotype with rates extremely low compared to the wild type enzyme, showing a large significance of the helical arch structure on the activity of the enzyme.

Publications Arising from this Work

Patel, N., Atack J. M., Finger, L. D., Exell, J. C., Thompson, P., Tsutakawa, S. E., Classen S. E., Tainer, J. A. & Grasby J. A. *Nucleic Acids Research* **(doi: 10.1093/nar/gks051)** Flap Endonucleases pass 5' flaps through a flexible helical arch using a disorder-order-disorder mechanism to confer specificity for free 5' ends.

Abbreviations

5yt	5x Yeast Tryptone media
Af	Archaeaglobus fulgidis
AGT	O^6 AlkylGuanine – DNA alkyltransferase
Ala / A	Alanine
AP	Apuridinic / pyrimidinic
Arg / R	Arginine
Asn / N	Asparagine
Asp / D	Aspartic acid
B	Biotin
Bp	Base Paired
BSA	Bovine Serine Albumin
CDDP	Cisdiamminedichloroplatin
Cdk	Cyclin Dependant Kinase
Cm	Chloramphenicol
COMP	Competitor
CPS	Carbomyl Phosphate Synthetase
CRN	Caenorhabditis elegans
DF	Double Flap
Dm	Drosophila Melanogaster
DMT	DiMethoxyTrityl
DNA	DeoxyriboNucleicAcid
dNTP	DeoxyriboNucleotide TriPhosphate
ds	Double Stranded
DTT	Dithiotreitol
E	Enzyme
E.coli	Escherichia Coli
EDTA	EthyleneDiamineTetraacetic Acid
EMSA	Electrophoretic Mobility Shift Assay
ES	Enzyme-Substrate
FAM	Fluorescein
FEN	Flap Endonuclease
FPLC	Fast Protein Liquid Chromatography
GEN	Gap Endonuclease
Gln / Q	Glutamine
Glu / E	Glutamic acid
H	Human
HEPES	HydroxyEthylPiperazineEthane Sulphonic acid
His / H	Histidine
HJ	Holliday Junction
HPLC	High Pressure Liquid Chromatography
IMAC	Immobilised Metal ion Affinity Chromatography
IPTG	IsoPropyl β D-I-ThioGalactopyranoside
Kan	Kanamycin
KGly	Potassium Glycinate
LB	Luria Broth
Leu / L	Leucine
lpBER	Long patch Base Excision Repair
Lys / K	Lysine
MALDI ToF	Matrix Assisted Laser Desorption Ionisation – Time of Flight

Met / M	Methionine
Mj	Methanococcus Janaschii
MM	Michaelis Menten
MWCO	Molecular Weight Cut Off
NER	Nucleotide Excision Repair
nt(s)	Nucleotide(s)
OD	Optical Density
ORF	Open Reading Frame
P	Product
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PCNA	Proliferating Cell Nuclear Antigen
PCR	Polymerase Chain Reaction
PEG	Poly Ethylene Glycol linker
Phe / F	Phenylalanine
PIP	PCNA interaction Partner
Pol	Polymerase
Pro / P	Proline
pY	Pseudo Y
Q	Bound Product
Q seph	Q Sepharose
RNA	RiboNucleicAcid
RNAi	RiboNucleicAcid Interference
RNase	RiboNuclease
Rp	Reverse phase
RPA	Replication Factor A
RT	Room Temperature
S	Substrate
SA	Streptavidin
Sc	Sacharomyces Cerevisiae
SDS	Sodium Dodecyl Sulphate
Ser / S	Serine
SNase	Staphylococcal nuclease
SOC media	Super Optimal broth with Catabolic repression
SP Hep	Heparin-Sepharose
ss	Single Stranded
TEAAc	Triethyl ammonium acetate
TEG	Tri Ethylene Glycol linker
Thr / T	Threonine
TNR	Tri Nucleotide Repeats
Tyr / Y	Tyrosine
WRN	Werner Protein
WT	Wild Type
XPG	Xeroderma Pigmentosum G
YEN	Yeast Endonuclease

Contents

Chapter 1: Introduction

1.1: The Phosphoryl Transfer Reaction

Both DNA replication and repair depend upon rapid enzyme catalysed DNA hydrolysis. This phosphoryl transfer reaction occurs via attack of a water nucleophile on the phosphodiester, followed by the displacement of an alkoxide leaving group (figure 1.1.1).

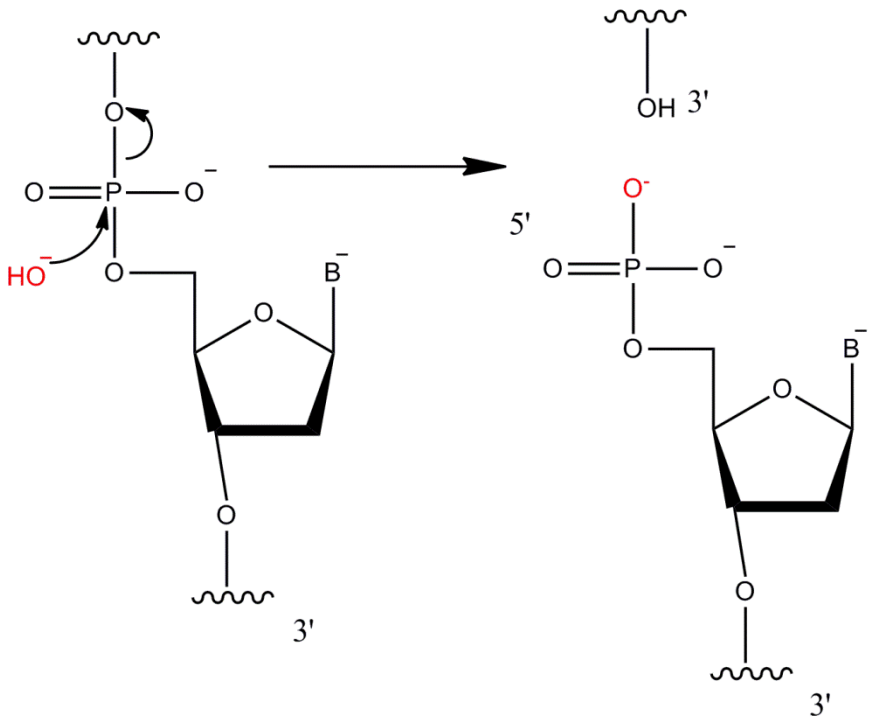


Figure 1.1.1: A typical mechanism for the hydrolysis of DNA within solution. The red nucleophile attacks the phosphodiester, and displaces the leaving group. This generates a 3' OH and 5' phosphate monoester. To produce the alternate products, the direction of attack needs to be on the 3' side instead of the 5' in order to maintain an in-line conformation for an S_N2 reaction.

The uncatalysed hydrolysis of DNA in solution is incredibly slow, with a half-life of approximately 30 million years at 25°C and pH 7.0 (based only on the upper limit of the hydrolysis of the dimethyl phosphate ion) due to the large energetic barrier inherent to this reaction (Schroeder 2005). Biological systems must find a method of performing DNA hydrolysis rapidly under physiological conditions. Nucleases, the enzymes that catalyse phosphate diester hydrolysis of DNA and RNA substrates, have evolved to solve this problem. Nucleases generally fall into one of two groups: exonucleases and endonucleases. Exonucleases perform DNA hydrolysis close to the end of double stranded DNA, while endonucleases cleave within the DNA strand.

Phosphoryl transfer reactions can take place through either associative (S_N2) or dissociative (S_N1) mechanisms. The dissociative reaction occurs via formation of an unstable intermediate (metaphosphate) by elimination of the leaving group. This unstable intermediate would be stabilised by addition of a nucleophile. The associative pathway involves bond formation between the nucleophile and the phosphorous first, followed by or simultaneous to elimination of the leaving group. It is thought that in most nucleases, reaction occurs via the associative route.

Nucleases use one or more of the following strategies to accelerate reaction rates:

- Provide a general base or metal to activate water as an attacking nucleophile.
- Provide an acid to stabilise the leaving group by protonation, or a metal to afford charge neutralisation.
- Provide one or more positively charged moieties, balancing the negative charge build up on the phosphoanion transition state.
- Provide a nucleophilic group to which the phosphoryl moiety is transferred (Jencks 1972; Williams 1998; Schroeder 2005).

Nucleases perform a phosphodiesterase activity. This activity for some enzymes is structure specific and many possess the ability to perform sequence specific cleavages, all of which are fundamental to a great range of processes. The actual 'machinery' required when performing phosphodiester cleavage and preparing a nucleophile for attack, is almost always the same. A covalent nucleophilic attack would require a Ser, Tyr or His residue; the 2' phosphate of DNA/RNA can also be used as an intramolecular nucleophile; some RNases use inorganic phosphates as a nucleophile to lyse RNA, creating 5' diphosphates; or most commonly, the nucleophile is simply a metal bound hydroxide (Saenger 1984; Voet 2004). However, it is the manner in which DNA or RNA substrates are delivered to this chemical machinery that determines the way in which cleavage takes place. This means that the active sites of these enzymes all carry out the same phosphodiesterase activity, but the specificities of these enzymes come from the ability and manner in which the phosphodiester is placed within the active site. This makes the structure of the nucleases, the shape of their active sites and position of binding sites relative to these, most important in the general roles of these families of enzymes (Yang 2011).

Exo - and endonucleases are involved in a large variety of intracellular processes. Examples of these are:

- 5' - 3' exo and endonucleolytic activity, in order to remove RNA primers formed in DNA replication and repair (Kao and Bambara 2003; Shen *et al.* 2005).
- 3' - 5' exonucleolytic activity for proofreading during DNA replication (Reha-Krantz 2010).
- The initiation of DNA recombination and repair (Marti and Fleck 2004; Mimitou and Symington 2009).
- Site-specific recombination (Patel and Steitz 2003; Grindley *et al.* 2006) by topoisomerases.
- RNA processing, maturation and interference (Abelson *et al.* 1998; Chu and Rana 2007; Moore and Proudfoot 2009; Nowotny and Yang 2009).
- RNA and DNA degradation essential for microbial defence (James *et al.* 1996; Tock and Dryden 2005; Sorek *et al.* 2008).
- Programmed cell death (Parrish and Xue 2006).

Most nucleases have an absolute requirement for one or more divalent metal ions to assist in hydrolysis. However, a notable exception is RNase A. RNase A catalyses the two step hydrolysis of RNA using histidine residues. In the first step, His12 acts as a general base, activating the 2' OH of the substrate as a nucleophile allowing it to attack the phosphoryl group in an intramolecular reaction, whilst His119 acts as general base protonating the alkoxide leaving group. In the second step of the reaction, His119 acts as a general base, activating H_2O to attack a cyclic phosphate, with leaving group departure general acid catalysed by His12 (Raines 1998; Raines 2004). The hydrolysis of RNA via a transesterification mechanism involving an internal 2'-OH is much easier than hydrolysis of DNA where attack by an exogenous water molecule is necessary. This is highlighted by the fact that almost all metal ion independent nucleases are RNA specific, although a few examples of metal independent enzymes perform DNA hydrolysis utilising a covalent catalysis strategy to form a protein-phosphoryl intermediate have been observed. One example of such an enzyme is Nuc of *Salmonella typhimurium* which acts in a manner analogous to RNase A, where His residues act as the nucleophile and general acid and a Lys residue would stabilise the leaving group (Stuckey and Dixon 1999). However in general, enzyme catalysed DNA hydrolysis requires metal ion cofactors for catalysis.

Staphylococcal nuclease (SNase) is an example of a one metal nuclease, producing single nucleotides by degrading DNA or RNA from the 3' end. SNase is a nonspecific enzyme, increasing the rate of phosphodiester hydrolysis by 10^{16}. SNase uses Ca^{2+} within its active site. Arginine residues close to the phosphodiester bond of the substrate stabilise the transition state, and a Glu residue close to the calcium bound water, is proposed to act as a base to activate the water as a nucleophile to attack the phosphoryl group (figure 1.1.2) (Kramer 1999).

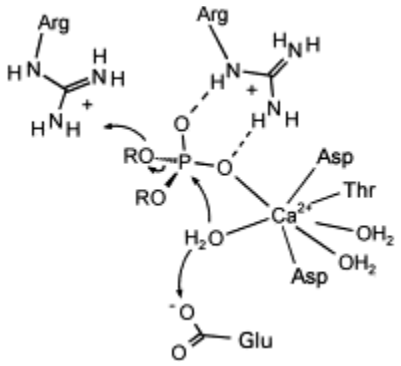


Figure 1.1.2: Reaction mechanism of SNase (taken from (Kramer 1999)). Glu 43 acts as a base, activating the Ca^{2+} bound hydroxide as a nucleophile)

Nucleases proposed to use three metal ion mechanisms are much rarer than nucleases that use fewer metal ions. Endonuclease IV is an example of an enzyme which uses three zinc ions as cofactors. It has been proposed that metal ions Zn1 and Zn2 generate a reactive hydroxide intermediate by deprotonating a nearby water molecule. This activated water attacks the scissile phosphate. All three Zn^{2+} ions are suggested to be involved in stabilising the associative transition state, with Zn3 stabilising the charge generated on the leaving group (figure 1.1.3). Most circumstances of three metal ion mechanisms are engrossed in contentious points, mainly due to the efficient and the well demonstrated nature of two metal ion mechanisms (Ivanov *et al.* 2007).

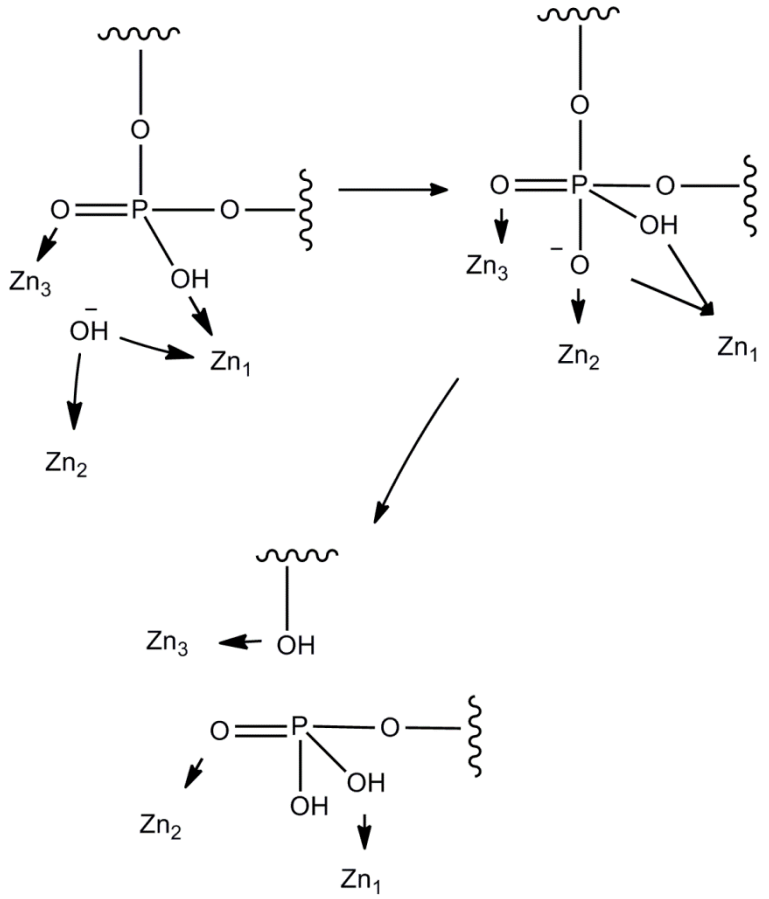


Figure 1.1.3: The proposed mechanism of the 3 metal ion hydrolysis of DNA within Endonuclease IV. Zn3 has accepted the departing O, and Zn2 is now coordinated to a new O atom.

The most thoroughly characterised group of nucleases have two active site divalent metal ions and are generally proposed to operate using a two metal ion mechanism as follows (figure 1.1.4) (Jencks 1972; Beese and Steitz 1991; Williams 1998; Schroeder 2005).

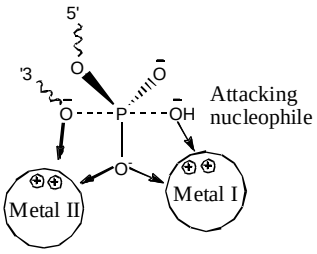


Figure 1.1.4: The general 2 metal ion mechanism. Metal one activates an attacking nucleophile, while metal two stabilises the charge on the leaving group.

Another proposal for the two metal ion mechanism that differs to the one described above, is where the metal ions do not directly contact the scissile phosphate. The metal ions of bovine pancreatic deoxyribonuclease I are proposed to act by stabilising the charge on the bridging oxygen atoms of the two nucleotides concerned which develop during the reaction, effectively lowering the activation energy (Weston *et al.* 1992; Jones *et al.* 1996).

Two metal ion nucleases strive to cleave RNA or DNA but exhibit a wide variety of structures existing in differing multimeric forms, possessing a myriad of protein folds and varied active site motifs. These differences all appear to contribute towards a variety of specific phosphate diester hydrolytic activities vital to the cell cycles of their respective organisms. As these attributes are so diverse, these enzymes will be classified by their function.

1.2 3' - 5' Exonucleases

DNAQ and *E. coli* EXO1

DNAQ or MutD is a subunit of the Klenow fragment that encodes the DEDDh motif. The DEDD motif is absolutely conserved across this class of enzymes that provide the carboxylates required to coordinate the two metal ions required for activity, although one of these ions is thought to aid substrate binding (Hamdan *et al.* 2002). The histidine residue of the active site motif of this enzyme is not conserved over all the exonucleases of this class, but varies between tyrosine and histidine. This difference however, appears to have no effect on activity or specificity. The preferred substrate of this enzyme is ssDNA, proofreading in a 3' to 5' direction during DNA replication (McHenry 1985; Brautigam and Steitz 1998; Brautigam *et al.* 1999).

EXO1 (exonuclease 1) isolated from *E.coli* is a part of this DEDD family, possessing a DEDDh motif. This exonuclease again specifically degrades ssDNA, and does so processively. The crystal structure of the enzyme highlighted the DEDDh motif, but more importantly revealed a C shaped structure, consistent with the processive nature of the enzyme (Breyer and Matthews 2000).

1.3 Holliday Junction Resolvases

T4 endonuclease 7 and T7 endonuclease 1

Holliday junctions (HJ) are structures formed in the process of joining four DNA strands into a branchpoint so that the exchange of strands can occur. This process takes place in strand invasion in recombination, and in DNA replication fork reversal. These junctions are resolved by dimeric enzymes that introduce symmetrical cuts reliant on the arrangement of the dimer interface and subsequent placement of catalytic residues. This shows the dependence of substrate binding on the tertiary or in this case, quaternary structure of the enzyme as shown in figure 1.3.1. T4 endonuclease 7 and T7 endonuclease 1 are examples of recently crystallised HJ resolvases that demonstrate this capability of nucleases to perform the same mode of activity but possess different structures or active sites. In this case the structures are similar, but it is the dimer interface that changes. Hydrolysis of these 4 way junctions occur in close proximity to the junction point, allowing subsequent separation into two DNA duplexes varied from the original genomic blueprint (Declais and Lilley 2008).

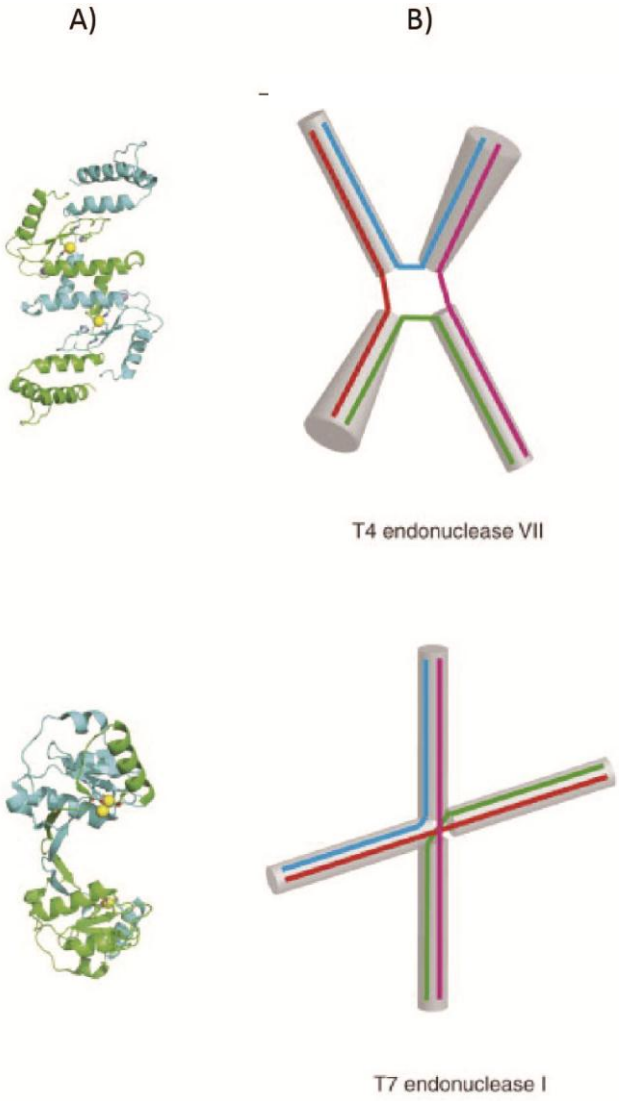


Figure 1.3.1: A) The dimeric structures of T4 endonuclease 7 and T7 endonuclease 1; B) The differing modes of substrate binding dependent on the dimer-dimer interface in T4 endonuclease 7 and T7 endonuclease 1 (taken from (Declais and Lilley 2008)).

1.4 5' - 3' Exonucleases

Phage λ and RecE

These exonucleases and the restriction endonucleases discussed later, share a moderately conserved PD-(D/E)XK motif. This depicts a motif whereby a D residue is first of the catalytic residues on a beta hairpin structure, preceded by a P residue. On the second strand of this hairpin is a D or E residue and K residue separated by hydrophobic residue X. The rest of the structure of these enzymes can vary drastically, displaying monomeric, dimeric and trimeric forms (Kovall and Matthews 1997; Orlowski and Bujnicki 2008; Zhang *et al.* 2009). Phage λ

and RecE share a similar mode of binding and catalysis, in that they are trimeric and tetrameric respectively, both forming a toroidal ring like structure. This forms a 'clamp' and slides down dsDNA, processively cleaving as enclosed structures often do, resulting in a 3' overhang ready for strand invasion and homologous recombination (Kolodner *et al.* 1994; Stahl *et al.* 1997). Dna2 in eukaryotes, essential for removing excessively long primers in DNA replication is posited to act in a similar manner, however its activity requires separate helicase and nuclease domains (Kim *et al.* 2006; Masuda-Sasa *et al.* 2006).

1.5 Endonucleases

Type II Restriction Endonucleases

Restriction endonucleases possess the PD-(D/E)XK motif, and are a superfamily of enzymes that are employed in bacteria and archaea to defend against invading viruses by selectively cleaving virus dsDNA. They are normally dimeric and sequence specific in nature, which when coupled together result in the ability to recognise palindromic sequences. The catalytic residues of these nucleases approach the minor groove of target DNA, whereas the binding of substrates, shown in MutI that recognises hemimethylated specific sequences, is mediated from the major groove side of the DNA, conferring the sequence specific ability upon the enzyme (Lee *et al.* 2005). By changing the dimer interface of these enzymes and adjusting the location of the catalytic domains accordingly, these nucleases can produce blunt ended duplexes, 3' overhangs or 5' overhangs. As seen in the case of HJ resolvases in section 1.3, restriction endonucleases are another superfamily of nucleases that exhibit specificity dependent upon the structure of the enzyme (Ban and Yang 1998a; Ban and Yang 1998b; Newman *et al.* 1998).

Endonucleases RNase HI and HII

RNase H like endonucleases are enzymes conserved in many organisms, from bacteria and retroviruses to humans. They share a topology with the DNAQ like exonucleases, whereby the core of the enzyme is composed of a cluster of 5 β sheets, on which the catalytic motif is normally located, surrounded by many α helices. The DEDD motif is altered in these enzymes with the first Asp residue absolutely conserved, while the last Asp residue is only partially conserved. These residues coordinate two metal ions, which catalyse the cleavage of

RNA/DNA hybrids, removing RNA primers during the synthesis of DNA during replication and sometimes repair. RNase HI degrades primarily RNA primers, but has a necessary substrate requirement of DNA/RNA hybrids spanning at least 4 ribonucleotides. Substrates are bound by two grooves within the tertiary structure of the enzyme, the RNA strand bound close to the divalent metal ions, ready for cleavage (Nowotny *et al.* 2008; Schultz and Champoux 2008). RNase HII appears to perform the same function but with one crucial difference, only requiring a single RNA/DNA pair embedded within a duplex. A substrate bound structure for RNase HII is yet to be solved; however, studies have shown little sequence similarity between HI and HII, a groove in both enzymes possessing most of the conserved residues between the two on one face. This is thought to implement directionality to the binding of the substrate, ensuring cleavage of only the RNA and possibly showing substrate recognition mediated by variations in the normal duplex structure (Crow *et al.* 2006; Cerritelli and Crouch 2009; Crow and Rehwinkel 2009). This is yet another example of a family of enzymes, where the location of active site residues and the substrate binding residues in relation to these, appears to influence the substrate choice and consequently the specificity of each member.

Flap Endonucleases

Flap endonuclease (FEN) is a catalyst of phosphate diester hydrolysis, which is vital during DNA replication and repair. As with most of the two metal ion nucleases, the FEN catalysed hydrolysis of DNA results in products with a 3’ OH and a 5’ phosphate monoester (figure 1.1.1), (Pickering *et al.* 1999). FENs catalyse the structure specific hydrolysis of 5’-flapped DNAs (figure 1.6.1). FENs are present in organisms from all domains of life. The enzymes that have been studied in most detail include *Caenorhabditis elegans* (CRN FEN), T5 exonuclease (T5 FEN), human FEN1 (hFEN1), *Methanococcus jannaschii* (MjFEN), bacteriophage T4 RNase HI, and *Archaeoglobus fulgidis* FEN (AfFEN). In addition, some organisms contain FEN-like proteins such as, in human Xeroderma pigmentosum G (XPG), human Gap endonuclease 1 (hGEN1), human Exonuclease 1 (hEXO1), *Drosophila melanogaster* GEN (DmGEN) and Yeast Endonuclease (YEN).

In summary, nucleases share a commonality, namely, phosphodiesterase activity. The ability to cleave phosphodiester bonds is aided by a shared need for a component (be it divalent metal ions, a histidine, tyrosine or serine residue or even inorganic phosphates,) that can be

used in order to promote nucleophilic attack, stabilise a pentacoordinate transition state or the leaving group of the reaction and/or provide a nucleophilic group for phosphoryl group transfer. The specificity of these nucleases are dependent on the rest of the protein, the structure determines the recognition of substrates, the way in which substrates are bound, and in some cases determining the processivity of enzymes.

1.6 FEN enzymes and their specificities

FEN enzymes all have a 'βα' Rossman fold, whereby the core nuclease domain is made up of 5 β sheets surrounded by many alpha helices, yet all exhibit a large array of precise enzymatic specificities, a source of controversy and intrigue. These activities can include,

- Flap endonuclease (Flap Endo) activity.
- Single stranded exonuclease (ssExo) activity (of ss-ds DNA junctions).
- Double stranded exonuclease (dsExo) activity.
- Gap endonuclease (Gap Endo) activity.

Each of these is illustrated below (figure 1.6.1),

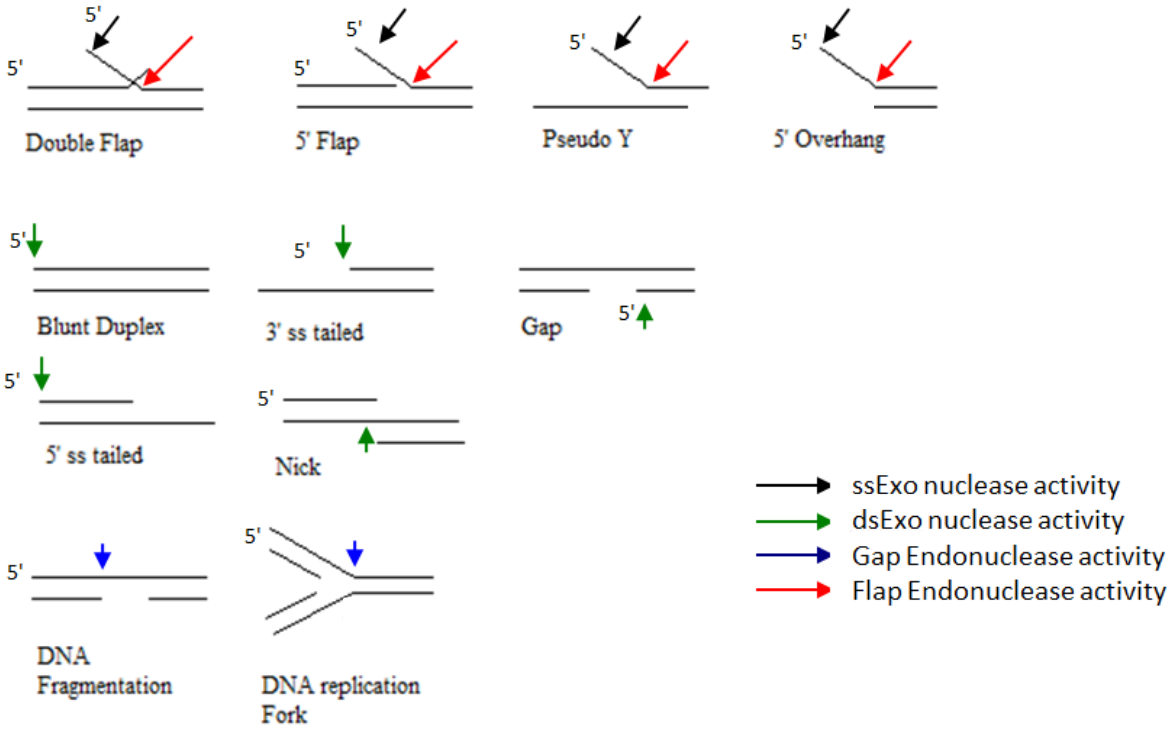


Figure 1.6.1: The many substrates and reactions of the FEN family of enzymes. Arrows indicate the point of cleavage. ss and ds Exo activity occur in most FEN enzymes, the dsEXO on Holliday junctions shown is a function of hFEN1. Gap endonuclease activity is most notably seen in hFEN1 and the hFEN1 paralogue, GEN1. Flap endo activity occurs in all of the well-studied FEN enzymes.

It is this great controversy, mainly centred on the ability of a structurally conserved superfamily of enzymes to possess such a large array of specificities on different substrates both with and without 5' flaps / 3' flaps with such exquisite specificity *in vivo,* which makes flap endonucleases the focus of this study.

Substrates of FENs

The 5’ flap substrate shown above in figure 1.6.1 mimics the structure generated during lagging strand DNA synthesis. Flap substrates are cleaved endonucleolytically by all FENs, and preferentially by hFEN1. The major site of hydrolysis is one nucleotide into the 5' duplex region, termed the downstream binding region (figure 1.6.2), although reactions can occur at other sites around the bifurcation. *In vivo*, 5' flaps are equilibrating structures and can exist as so called ‘double’ 5'-3' flaps because both displaced and nascently synthesised nucleic acid are complementary to the template DNA (figure 1.6.2).

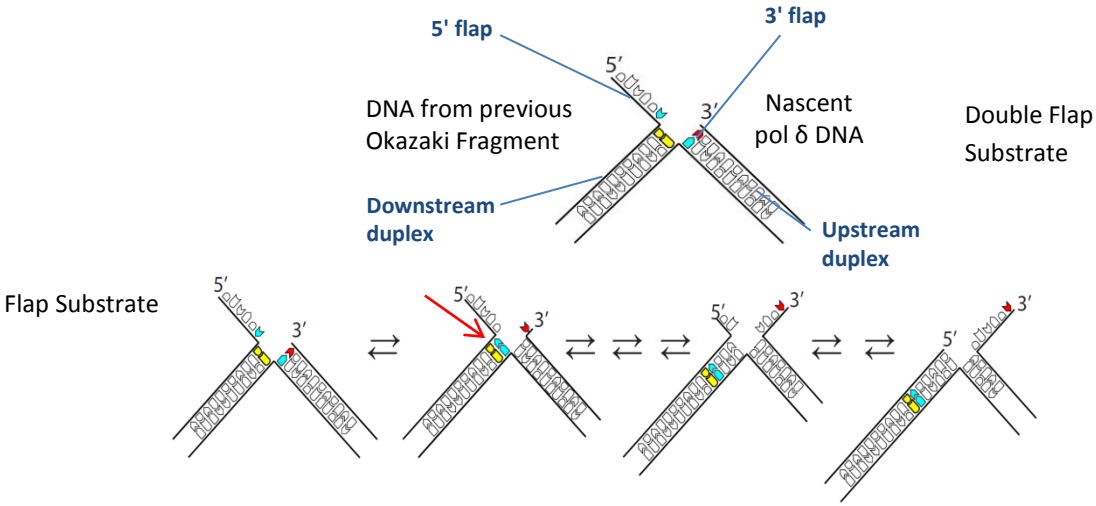


Figure 1.6.2: An equilibrating double flap substrate structure with the important sites highlighted in blue. The red arrow indicates the major site of cleavage despite the ability to equilibrate into the other structures shown.

Double flaps are cleaved by all FENs (Xu *et al.* 1997; Lyamichev *et al.* 1999; Williams *et al.* 2007). In higher organism FENs, substrates/equilibrating substrates with a single nucleotide 3' flap produce a reaction that is completely specific, occurring one nucleotide into the 5' duplex region. It has been noted that this specificity of reaction gives rise to a nicked DNA product that could be immediately joined by DNA ligase without the need for further enzymatic processing by polymerases. In addition to flap structures, FENs will also cleave minimisations of these structures *in vitro*. These include substrates that do not possess upstream duplex (defined in figure 1.6.2) like pseudo Y structures and 5' overhangs. However, the efficiency in binding and catalysis of these minimal substrates in higher organism FENs is reduced. Dervan *et al*. 2001, showed that bacteriophage T5 FEN binds pseudo Y substrates with very high affinity, compared to similar overhang substrates – K_D values of ~5 nM cf. ~90 nM respectively.

FENs can also process structures that do not possess a 5' single stranded flap. These exonucleolytic reactions can occur on nicked DNA, 5' blunt ended duplexes with a 3' flap and duplex DNA. Although apparently a different type of reaction, FENs process these substrates with analogous selectivity for reactions 1 nucleotide into the 5' duplex region. Single stranded exonuclease activity can also sometimes be observed on DNA substrates that present 5' ss flaps, providing single nt products (Sayers and Eckstein 1990; Rumbaugh *et al.* 1999; Garforth *et al.* 2001).

A more controversial aspect of FEN reaction specificity is the so called gap endonuclease reaction. In early literature it was claimed that hybridisation of an oligonucleotide to a 5' ss flap (figure 1.6.3) inhibited FEN activity even when duplex formation left a single stranded region around the site of bifurcation. However, more recent studies have contradicted this (Barnes *et al.* 1996; Bornarth *et al.* 1999; Williams *et al.* 2007; Finger *et al.* 2009; Gloor *et al.* 2010). A systematic study of a double flap substrate to which complementary oligonucleotides were hybridised to the 5' ss flap demonstrated that reaction still occurred, with its efficiency dependant on the gap length (the size of single stranded DNA adjacent to the bifurcation) (Bornarth *et al.* 1999; Williams *et al.* 2007; Finger *et al.* 2009).

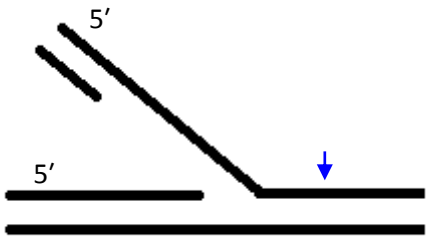


Figure 1.6.3: An illustration of GEN activity on gapped substrates

Gap endonuclease activity is relevant to the situation that may occur *in vivo* when there is secondary structure present in the 5' ss flap. Genetic disorders such as Huntington's disease, colon cancers and Friedrich's ataxia, have been hypothesised to be a result of tri-nucleotide repeat (TNR) sequences within the genome, forming higher order secondary structures within strands, such as hairpins etc. Null hFEN1 mutants have been shown to cause an expansion in the number of repetitive sequences (Henricksen *et al.* 2000). The secondary structures that have arisen from hairpins, tetraplexes and triplexes on flaps can be integrated into the genome, causing TNR expansions. The hFEN1 reaction is inhibited when secondary structure is present within the 5' ss flap. A recent study by Finger *et al.,* 2009 demonstrated at low concentrations of substrate (k_{cat}/K_M conditions) the FEN endonucleolytic reaction was only 8 times slower with a 10 base paired hairpin DNA within the 5' ss flap.

It has also been demonstrated that exclusively endonucleolytic activity one nucleotide into the downstream region takes place when hFEN1 is presented with a double flap GEN substrate that does not present a blunt ended duplex on the 5′ terminus of the substrate, but possesses instead, a fold back duplex or hairpin on the 5′ terminus (Finger *et al.* 2009).

It should be noted that many FEN substrates and particularly those used *in vitro* to study reactions possess multiple potential reaction sites. For example a T5 FEN reaction of a 5' labelled 5' flap substrate can result in both endonucleolytic and ssExo products. Gap endonuclease substrates present opportunities for both endonucleolytic and dsExo reactions (figure 1.6.4).

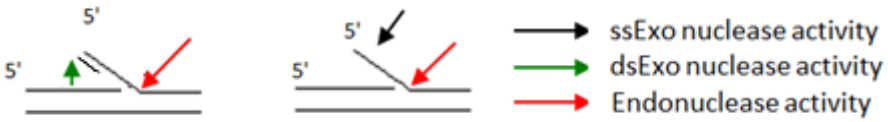


Figure 1.6.4: Illustration of the variety of possible activities available to T5 FEN *in vitro* on minimisations of the DNA substrates found *in vivo*.

Roles of FEN

FENs are thought to be involved in many intracellular processes. Its involvement in DNA replication and repair is well characterised (Qiu *et al.* 2001; Chapados *et al.* 2004; Liu *et al.* 2004; Williams *et al.* 2007; Burgers 2009; Zheng *et al.* 2011). Furthermore, a role in cell death and DNA recombination has been suggested, mainly through participation in a variety of *in vitro* protein-protein interactions with known participants in said processes (Parrish *et al.* 2003; Tseng and Tomkinson 2004; Sakurai *et al.* 2005; Tsutakawa *et al.* 2011; Zheng *et al.* 2011). However, these putative roles are not well characterised.

In DNA replication within human cells the synthesis of new DNA by polymerases must take place with a 5' to 3' direction. As the DNA helix is unwound by a helicase, thought by some to be the MCM2-7 complex (Sclafani and Holzen 2007), replication of the leading strand is a simple matter, with the synthesis of leading strand DNA by pol ε taking place in the required direction. However, this is not the case in the lagging strand; replication would have to take place in a 3' to 5' direction (figure 1.6.5). Therefore, synthesis is carried out by a primase placing a series of RNA primers 5' to 3' complementary to the lagging strand. Pol α and pol δ then synthesize the DNA for these fragments, termed Okazaki fragments. This is thought to

occur at a much slower rate than the leading strand synthesis, in order to prevent the leading strand overtaking the lagging strand which would eventually lead to the replication fork stalling, the 'trombone' method of lagging strand replication is posited to be utilised. When the synthesis of one Okazaki fragment reaches the next fragment, flap displacement occurs, the preceding fragment displacing the next, causing the formation of a 5' unannealed flap. This is called strand displacement synthesis. Human FEN1 would then cleave these displaced flaps endonucleolytically leaving a nick in the lagging strand DNA that is then ligated by DNA ligase I to form one continuous strand (Qiu *et al.* 1999; Chapados *et al.* 2004; Chon *et al.* 2009; Zheng *et al.* 2011). Fifty million flaps are formed per eukaryotic cell cycle, and must be removed for the conservation of genomic stability due to the capability of displaced flaps to form secondary structures and be integrated within the DNA. The importance of this is demonstrated effectively by FEN1 homozygous knockouts in mice displaying embryonic lethality, and heterozygous knockouts exhibiting increased tumour genesis (Larsen *et al.* 2003; Burgers 2009).

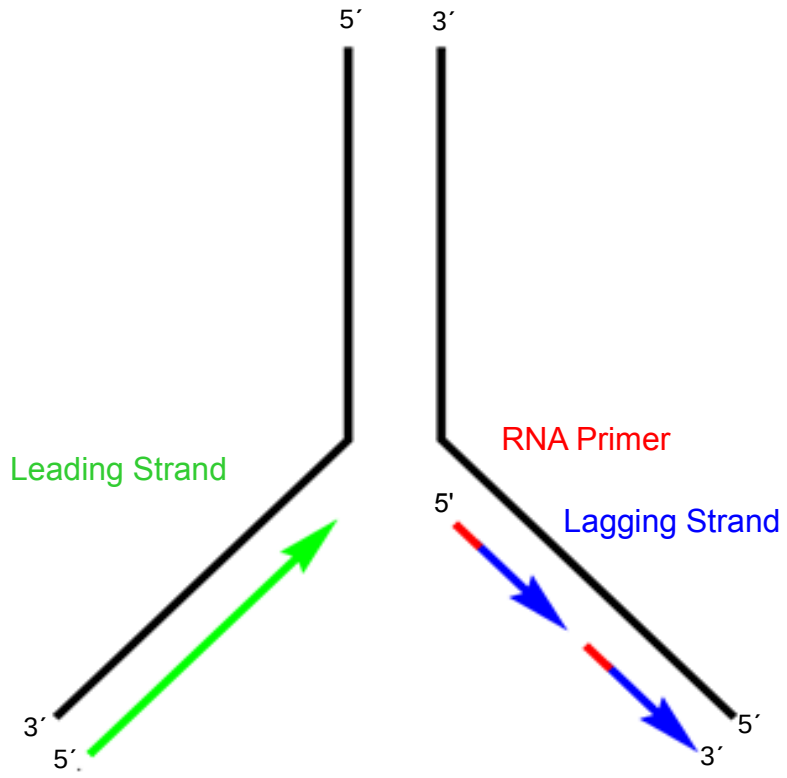


Figure 1.6.5: Diagrammatic depiction of the process of DNA replication. The green leading strand is replicated in a 5' to 3' direction, while the replication of the lagging strand is more complicated, as replication cannot occur 3' to 5', replication must occur via the placement

of a series of RNA primers (red) with the complementary DNA (blue) laid by DNA polymerase δ.

In vivo, longer displaced ssDNA flaps, ~30 nts or more, are not cleaved by hFEN1. Instead, the larger surface area of ssDNA is bound by multiple ssDNA binding proteins, replication factor A (RPA). These factors are refractory to hFEN1 cleavage, so recruitment of the nuclease Dna2 is necessary. Dna2 cleaves this long 5' flap leaving a flap of around 7 nts, ready to be cleaved by hFEN1. These situations are seen more in organisms other than humans which are not as tightly regulated, such as *S. Cerevisiaie* (Budd and Campbell 1997; Kang *et al.* 2010).

PCNA is thought to be instrumental in bringing the elements of DNA replication together. It exists as a trimer, forming a toroidal DNA clamp, and confers processivity to polymerases. It is proposed that each unit of PCNA interacts with a specific protein partner, and in this manner can participate in the hand off of DNA replication factors, delivering pol δ, hFEN1 and DNA ligase I when necessary (Dionne *et al.* 2003; Burgers 2009; Beattie and Bell 2011). The C terminus of hFEN1 is a highly mobile portion of the enzyme. This part of the enzyme is thought to mediate the main interactions with PCNA and is referred to as the PIP box (hFEN1 numbering 331-350). *In vitro,* it has been shown that PCNA stimulates the activity of hFEN1. This is achieved by lowering the K_M and increasing the turnover number, rather than increasing the k_{ST} (single turnover rate constant) (Tom *et al.* 2000; Hutton *et al.* 2010), shown by measurements of hFEN1 activity indicating a mechanism with product release as its rate limiting step (Williams *et al.* 2007; Hutton *et al.* 2008; Finger *et al.* 2009).

There are several pathways by which DNA damage can be resolved, but hFEN1 is only implicated in long patch base excision repair (lpBER). DNA damage is recognised by DNA polymerase δ, possibly by an abnormality in the normal structure of dsDNA, and excised on the 3' side by a DNA glycosylase, breaking the N-glycosidic bond and removing the base to create an abasic residue. AP endonuclease I then removes this abasic residue, leaving a gap in the DNA. This gap can be extended by DNA pol δ, creating a displaced 5' flap similar to those formed in strand displacement synthesis which can be cleaved by hFEN1 to create a ligatable nick (Warner *et al.* 1980; Dianov and Lindahl 1994; Kurthkoti and Varshney 2011). LpBER is thought to be a major component of DNA repair within the mitochondria of cells;

the abundance of oxidative species makes mitochondrial DNA particularly susceptible to damage. This damage results in oxidatively damaged abasic sites within the DNA that can only be repaired by lpBER (Zheng *et al.* 2011). There is conflicting evidence for this, with groups finding evidence of hFEN1 both within and outside of the mitochondria (Szczesny *et al.* 2008; Zheng *et al.* 2008).

There are also some post translation modifications that have been posited for hFEN1. Phosphorylation of the buried Ser187 (Sakurai *et al.* 2005; Tsutakawa *et al.* 2011) at late S phase by Cdk1 and 2 has been shown to inhibit hFEN1 (Henneke *et al.* 2003). Methylation of hFEN1 is thought to have the opposite effect, preventing phosphorylation, and facilitating the binding to PCNA (Zheng *et al.* 2011). Lastly, acetylation is postulated to take place on 4 lysines in the mobile portion of the C terminus of hFEN1. These modifications cause a variety of effects, influencing protein-protein interactions (Sakurai *et al.* 2005); blocking the strand displacement synthesis and influencing cleavage down the Dna2 route; and lastly regulating substrate binding (Tsutakawa *et al.* 2011; Zheng *et al.* 2011).

2 metal ion requirement with Flap Endonucleases

All FENs are thought to use 2 metal ions for catalysis of the hydrolysis of DNA, utilising the mechanism depicted within figure 1.1.4. These divalent metal ions, usually Mg^{2+} are coordinated by 7-8 conserved carboxylates (figure 1.6.6), located within the active site (Ceska *et al.* 1996; Syson *et al.* 2008).

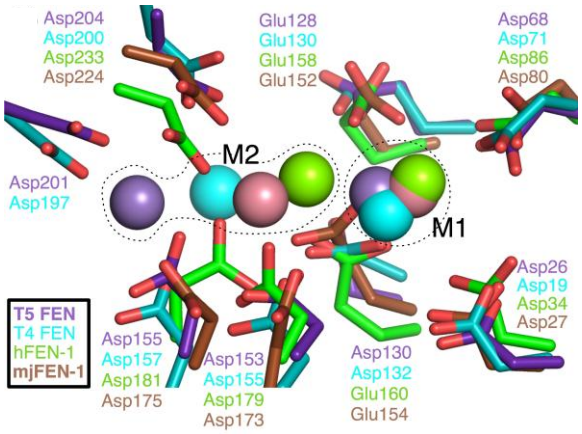


Figure 1.6.6: Conserved residues within the FEN family of enzymes possessing carboxylates dictating the position of metal ion sites. Metal sites 1 and 2 from T5, T4, *Mj* (*Methanococcus janaschii*) and hFEN1 shown, superimposed onto the residues of T5 FEN.

(from B. Chapados, Scripps Institute). Distances between ions are 8 Å in T5 FEN, 3.4 Å in hFEN1, 6.3 Å in T4 FEN and 5 Å in mjFEN.

As shown in figure 1.6.6, in most substrate-free FEN enzymes have two metal ions in the active site that are too far apart to perform the classical two metal ion mechanism depicted in figure 1.1.4. For example the metal-metal distance in T5 FEN (figure 1.6.6, purple) is 8Å. Upon measuring the kinetics of the bacteriophage enzyme T5 FEN with respect to concentration of divalent magnesium ion over several orders of magnitude, a need for 3 ions within the active site was revealed. However, on titrating the inhibitory divalent ion calcium into the T5 FEN active site occupied by divalent magnesium ions, a $1/[Ca^{2+}]^2$ dependence at high concentrations was seen. This highlighted a catalytic requisite for 2 metal ions within the active site. It was therefore posited, that the observed metal ion 2 site in the bacteriophage enzyme is present to encourage substrate binding, and that upon substrate binding a third ion is bound close enough to metal 1 to participate in the two metal ion mediated hydrolysis of the phosphodiester backbone of DNA (Tock *et al.* 2003; Syson *et al.* 2008).

Structures of FEN and related enzymes

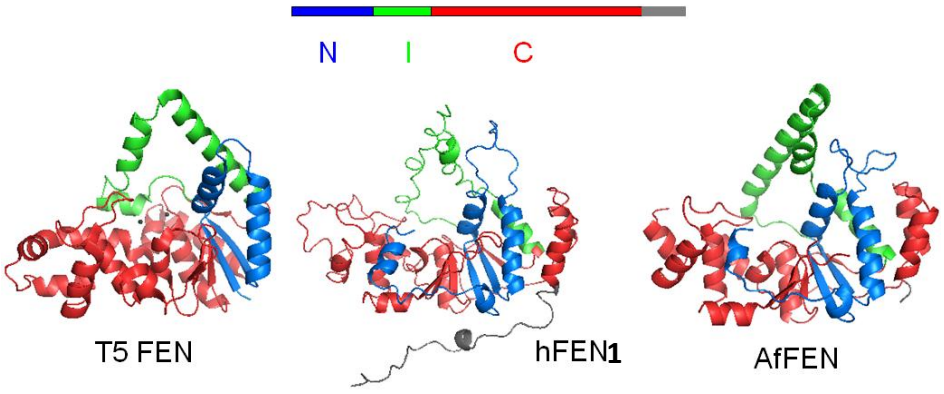


Figure 1.6.7: Schematic diagram of the domains of FEN mapped onto substrate-free structures of T5, human and AfFEN. The grey portion of the hFEN1 is an extension of the C domain not present in T5 FEN.

FEN

A number of X-ray structures of FEN proteins have been solved (Ceska *et al.* 1996; Devos *et al.* 2007; Tsutakawa *et al.* 2011). Three FEN proteins have been crystallised with bound

DNAs. These structures of bacteriophage T4 RNase H bound to a pY substrate, AfFEN bound to a duplex with a 3'-overhang and most recently structures of hFEN1 bound to substrates and products with active site metals; have produced a wealth of information relating to FEN-substrate interactions. FEN proteins can be divided into three domains, the N-terminal domain (N), the C-terminal domain (C), and intermediate (I) linker domain (Figure 1.6.7). In higher organism FENs the C terminus of the protein is extended into a separate section which acts as a site of protein – protein interaction (figure 1.6.7, hFEN1 structure - grey). The active sites of FENs are made up of parts of all 3 domains. Structures of FEN proteins bound to substrates (Devos *et al.* 2007; Tsutakawa *et al.* 2011), together with a number of earlier biochemical analyses (Pickering *et al.* 1999; Dervan *et al.* 2002; Patel *et al.* 2002; Williams *et al.* 2007; Finger *et al.* 2009; Sengerova *et al.* 2010), indicate that the downstream duplex (figure 1.6.2) is accommodated by the C-terminal domain of the protein, whereas the upstream (figure 1.6.2) part of substrates is bound in a binding site formed mainly from the N-terminus. FEN proteins contain a number of substrate binding motifs that have largely been conserved throughout evolution from bacteriophages to humans. These interaction motifs are reviewed below, in the context of hFEN1 (figure 1.6.8, A-Fii).

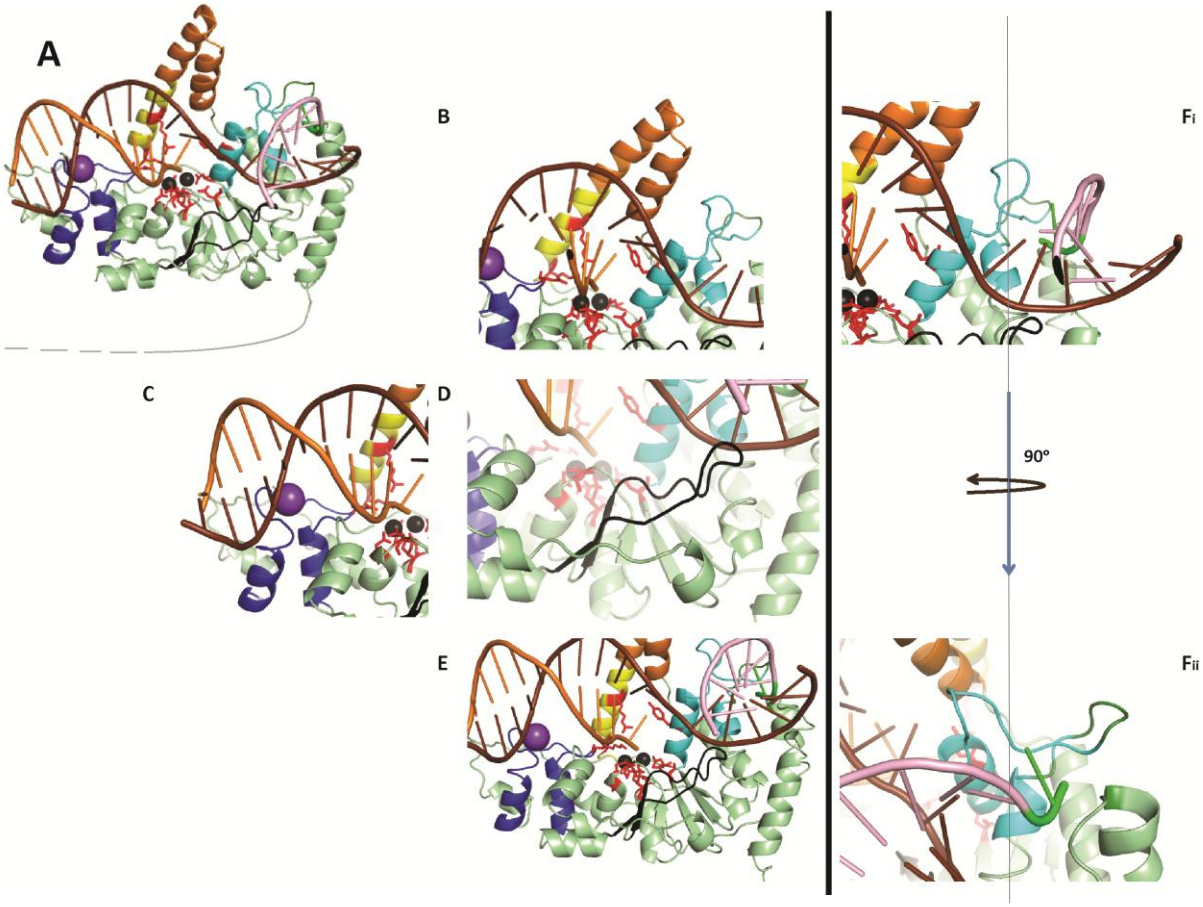


Figure 1.6.8: Crystal structure of hFEN1 defining FEN binding sites to be discussed below: A) Full length hFEN1 protein; B) The helical arch domain and active site: the base of the helical arch and the cap are coloured yellow and orange respectively. Residues that contribute to the active site are coloured red. Divalent metal ions pictured below the helical arch and within the active site are coloured black; **C) The H3TH motif:** coloured in blue with a bound monovalent cation (potassium) shown as a purple sphere. The downstream duplex is shown in this caption bound primarily by this motif; **D) The beta pin:** a positively charged series of conserved residues coloured black; **E) The βα/Rossman fold 'saddle':** a series of beta sheets surrounded by alpha helices encompassing both the upstream (cyan), downstream (blue) and beta pin binding site coloured in pale cyan; **Fi**) **The hydrophobic wedge** is coloured in cyan and when rotated left 90° in **Fii) the acid block and 3' flap of the upstream duplex** is visible in green.

The H2/H3TH Motif

The binding site for the reacting DNA duplex or downstream duplex is dominated by a non-specific duplex DNA interaction known as the helix-two (or in some cases three) turn helix

(H2/H3TH) motif (Figure 1.5.3), conserved among FENs. A number of biochemical studies with FENs from various organisms confirm this region to be the site of downstream DNA binding (Mueser *et al.* 1996; Dervan *et al.* 2002; Liu *et al.* 2006; Devos *et al.* 2007). Recent crystallographic images of a truncated human FEN1 and DNA show a potassium ion bridge in this region. The K^+ ion coordinates the DNA backbone and protein carboxyl oxygens and the hydroxyl group of residue S237 within the H2TH motif. This ion bridge and the four basic residues, (R239, K244, R245 and K267, hFEN1 numbering) provide a surface for interaction with the reacting duplex, largely by contacts to the template strand (figure 1.6.8, C) (Pelletier and Sawaya 1996).

The Upstream Duplex Binding Site and the hydrophobic wedge

Below the helical arch and adjacent to the upstream binding site, a hydrophobic wedge (figure 1.6.8, Fi) forces open the dsDNA-dsDNA or dsDNA-ssDNA junction. This wedge, originally identified in AfFEN is composed of α2 and α3 and the connecting loop (Chapados *et al.* 2004). In substrate structures of both T4 RNase HI and hFEN1, residues from α2 stack on the terminal base pair of the downstream duplex, which is located in front of the active site. In hFEN1 where a second duplex is accommodated in the juxtaposed binding site, the top of helices α2-α3 and their connecting loop interact with the terminal base pair of the upstream duplex bent at a 100^o angle. In T4 RNase HI, this region of the protein interacts with single stranded DNA. Human FEN1 structures both with and without DNA show the hydrophobic wedge undergoes a disorder to order conformational change upon substrate binding. Mutation of several residues of the hydrophobic wedge severely affected activity, further emphasising a key role for this binding site (Chapados *et al.* 2004).

A protein secondary structure element that resembles a 'β pin', which provides positively charged residues to the upstream binding site, was identified within hFEN1 substrate/ product structures. The β pin is conserved throughout evolution and is also present in bacteriophage enzymes. Upon substrate binding, the β-pin appears to be moved backwards (Tsutakawa *et al.* 2011) providing contacts to the template DNA within the upstream binding site (figure 1.6.8, D).

A 3' Flap Binding Site in higher organism FEN1 enzymes

As noted previously the preferred substrates of higher organism FENs is the double flap, which possesses a 1nt 3' flap and a 5' displaced ssDNA flap (Williams *et al.* 2007; Finger *et al.*

2009; Tsutakawa *et al.* 2011). Binding of the single nucleotide 3’ flap within a binding pocket present in higher organism FENs (figure 1.6.8, Fii), stimulates catalysis of substrates through stabilising the enzyme substrate complex, effectively demonstrated by how at lower concentrations of substrate, double flaps are cleaved at faster rates of reaction than single flaps by both hFEN1 proteins and their archaeal counterparts. This binding site is absent in bacteriophage enzymes, which display no stimulation on introduction of a 3’-flap. The molecular basis for specificity for 3’-flaps was first observed in structures of AfFEN. More recently the details of the 3’-flap interaction in the context of total substrate were observed in hFEN1-substrate and hFEN1-product structures. This revealed that one sixth of the binding surface between hFEN1 and its substrates is mediated between the interaction between the 3' flap and the flap binding site in hFEN1. This is discussed further in chapter 3.7. The one nucleotide 3’ flap requirement in higher organism FENs appears to be selected for by an acidic block of residues (E56-E59, hFEN1 numbering). This block creates a negatively charged ‘road block’, (figure 1.6.8, Fii (green)) restricting the 3' flap binding pocket to be able to fit only 1 nucleotide flaps (figure 1.6.8, Fii) (Sengerova *et al.* 2010; Tsutakawa *et al.* 2011).

Double nucleotide unpairing in the active site of FENs

A notable feature of the T4 RNase HI-DNA structure solved in the presence of EDTA is that substrate does not occupy the active site. Based on the metallobiochemistry of FENs and assuming similar positions for active site ions as in the substrate free protein it was suggested that to allow contact between the scissile phosphodiester bond and active site metals the DNA would have to unpair. Moreover, modelling indicated that the unpairing of two nucleotides would be required. In substrate structures of hFEN1, where the substrate DNA forms an exonucleolytic substrate with a 3’-flap, a fully base paired duplex is observed. The base paired terminal nucleotide of duplex is stacked upon Tyr 40 in hFEN1 substrate structures. However, this nt has departed in the product complex. In the product structures, Tyr 40 instead stacks on the unpaired terminal nt of the product, which supports the proposal that two nucleotides of substrate must unpair to position the scissile bond for reaction. On other FENs this residue is variously replaced by phenylalanine (T5 FEN) or isoleucine (T4 RNase HI), although it is maintained as Tyr in archaea. This residue together with conserved helical arch residues Lys 93 and Arg 100 appear to form a capture mechanism for the unpaired DNA (figure 1.6.9).

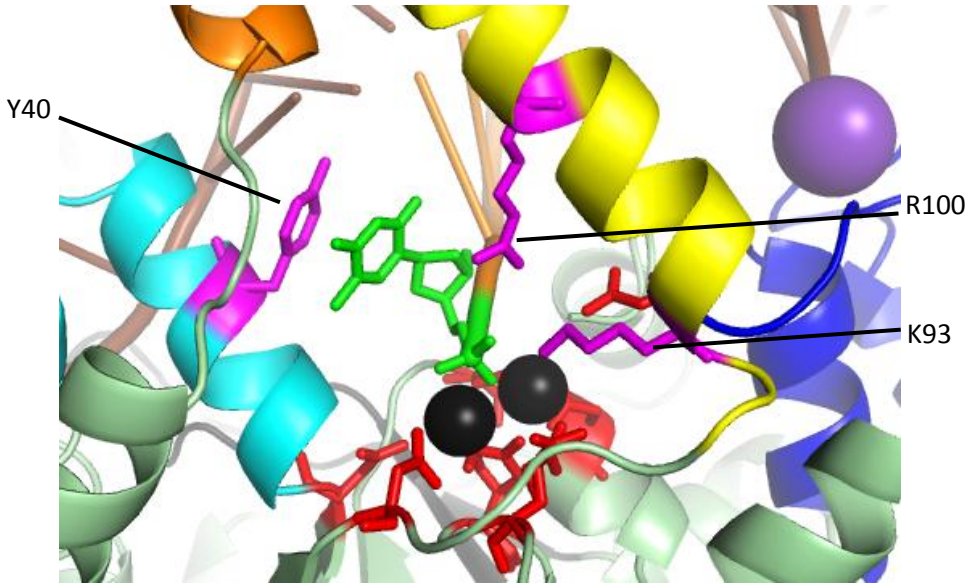


Figure 1.6.9: Stacking of the unpaired terminal nt of the downstream duplex in an hFEN1-product complex as viewed from behind the enzyme: The terminal nt (green), which has been brought into close proximity to the divalent metal ions (black) within the active site is stacked on the Y40 residue on the cyan helix, highlighted in magenta. Unpairing this nucleotide is thought to be encouraged by the other magenta residues on the helical arch (yellow), R100 and K93 (left to right).

Tsutakawa *et al.* propose a model (figure 1.6.10) that implies a disordered to ordered transition upon substrate binding of the helical arch and hydrophobic wedge. The binding of a 3' flap and template strand seems to position the 5' flap in a manner that results in the flap being encircled within the helical arch, between alpha helices 2 and 4 (hFEN1 numbering) on ordering of the enzyme structure.

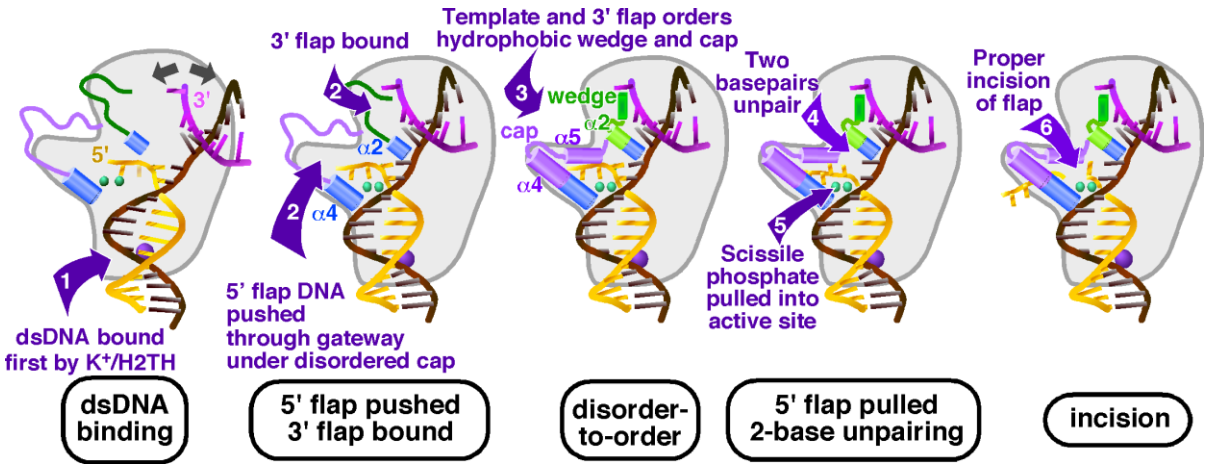


Figure 1.6.10: Cartoon Schematic of a proposed mechanism for endonucleolytic cleavage by hFEN1 on a double flap substrate: N.B. The 'cap' mentioned in this cartoon refers to the top of the helical arch mentioned within the text, the rest of which (termed the 'gateway') is made up by alpha helices 2 and 4. (Taken from (Tsutakawa *et al.* 2011))

The Helical Arch

The intermediate domain of FEN proteins, known as the helical arch (figure 1.6.8, B) was first visualised in structures of T5 FEN. In this bacteriophage FEN the arch is made up of two α helices, one containing hydrophobic residues and the other positively charged residues (Ceska *et al.* 1996). The arch is only big enough, when ordered, for single stranded (ss) DNA to fit through. In substrate-free structures of FENs from a range of organisms, the conformation of the helical arch varies considerably. In the structures of T5 FEN and AfFEN the arch is in a helical ordered conformation. The arch in T4 RNase H is partially disordered so that most of one helix could not be visualised, even with substrate bound, although the substrate in this structure is not positioned in the active site (Devos *et al.* 2007). The helical arches of substrate-free MjFEN and hFEN1 were disordered in comparison to the structure of T5 FEN, where crystal packing effects may produce the observed ordered conformation.

In contrast to substrate-free structures of hFEN1, those with bound substrates or product do adopt a helical conformation. The helical arch of hFEN1 does have some differences to the lower order T5 FEN helical arch; in T5 FEN α5 is a continuous helix that runs behind the main N-terminal domain of the protein, whereas in hFEN1 the equivalent region is two helices with a short linker. Nevertheless, the potential to form a helical aperture is conserved throughout evolution. Notably two basic residues (K93 and R100 in hFEN1 and K83 and R86

in T5 FEN are positionally conserved in all FENs. In structures of hFEN1, with product phosphate monoester bound to active site metals, these two residues interact with the phosphate. Mutations K93A and R100A (hFEN1) and K83A (T5 FEN) severely impair FEN reaction (Finger *et al.* 2009; Sengerova *et al.* 2010). It is suggested that the helical arch may be responsible for the specificity of FEN proteins for substrates that have 5'-termini (discussed in detail below).

Other nucleases related to FEN proteins-the 5'-nuclease superfamily

Originally identified on the basis of conservation of a characteristic large number of active site carboxylates, a superfamily of nucleases with FEN like features exist that play various roles during DNA repair and recombination. Whilst most of these proteins appear to conserve features of the N and C-terminal domains of FENs, some family members have I domains that drastically differ from those found in FENs. It is postulated that this, coupled with changes in the second nucleic acid binding site are the evolutionary alterations that afford substrate specificities.

XPG and bubble structures

XPG is part of a multi-protein complex necessary for nucleotide excision repair (NER). In XPG, the I domain is approximately 600 amino acids whilst in other FEN enzymes this region is about 70 amino acids in length (Staresincic *et al.* 2009). The I domain of XPG, much like in other FEN enzymes appears to be instrumental in substrate specificity, but it also has been speculated to be involved in mediating protein-protein interactions in the assembly of the NER complex (Dunand-Sauthier *et al.* 2005). Nevertheless, sequence alignments demonstrate that residues equivalent to hFEN1 K93 and R100 are conserved throughout XPG enzymes.

Unlike FEN proteins, which demand substrate with free 5'-termini, XPG has been shown to cleave bubble structures. This could be linked to the much larger I domain, giving the enzyme an I domain unlike any other within the FEN family. The Scharer lab recently suggested that bubble structure resolution could be a sequential process whereby another nuclease unrelated to FEN, XPF, hydrolyses a phosphate diester on the 5'-side of the bubble structure, thereby creating a structure that resembles a flap or pseudo-Y structure (Staresincic *et al.* 2009). Thus, XPG may be a flap endonuclease instead of a bubble endonuclease *in vivo*.

EXO1

EXO1 is an enzyme originally purified from *S. pombe*, but is also found in other organisms such as humans. EXO1 processively degrades DNA duplexes (figure 1.6.1) with 5' to 3' polarity to generate mononucleotide products (Szankasi and Smith 1995). An EXO1 deletion strain (EXO-1) of *S. pombe* was found to be deficient in mismatch repair, implicating the protein in DNA repair (Lehmann 1996). Recently human EXO1 was co-crystallised (figure 1.6.11) with substrate and product DNAs and active site ions. The structures closely resemble those of FEN-DNA complexes and highlight the conservation of the nucleic acid binding motifs such as the H2TH dsDNA binding region and the helical arch. The EXO1/product DNA structure showed that the product DNA is also unpaired and similarly positioned in the active site. Like the FEN/product DNA structure the 5'-phosphate monoester of the product DNA is also coordinated by the conserved active site arginine and lysine residues. Additionally, in substrate structures, where the DNA remains paired and has not entered the active site, the terminal nucleobase is stacked upon on a α2 histidine residue mimicking the contacts observed in hFEN1 to residue Y40. These similarities in structure between hEXO1 and hFEN1 are beneficial in understanding results seen where full length enzyme and truncated variants of EXO1 have exhibited flap endonuclease activity. The converse is also true; hFEN1 can perform exonucleolytic activity when presented with the appropriate substrate. Significant differences between hFEN1 and hEXO1 structures include the absence of the 3' flap binding site observed in higher organism FENs, presumably reflecting the differences between function between enzymes; and the product structures show that the unpaired nucleobase is stacked upon α2 tyrosine, located below the aforementioned histidine that stacks the paired duplex, possibly aiding the processive nature of hEXO1 (Orans *et al.* 2011).

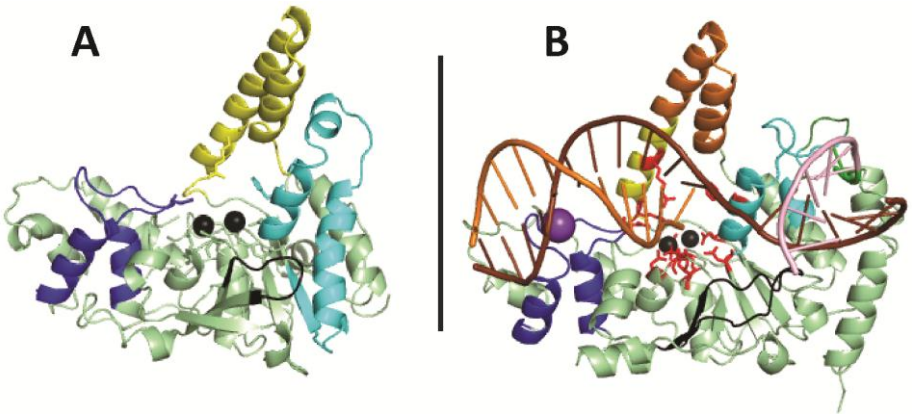


Figure 1.6.11: Crystal structure of A) hEXO1 compared to the structure of B) hFEN1: The important features of the enzymes, the helical arches (yellow), upstream (cyan) and downstream (blue) binding regions and the beta pin (black), have been coloured in an identical fashion to each other. These are also synonymous with the colour scheme utilised in figure 1.6.8.

GEN1

GEN1 proteins have been isolated from *Drosophila melagonaster* (DmGEN), yeast and humans. DmGEN is an 82.5 kDa enzyme, which like FEN proteins contains three domains. DmGEN has 10-30% and 40-50% sequence identity with the N and C domains, respectively, of the FEN superfamily proteins. DmGEN is reported to have a weak flap endonuclease activity but a preference for gapped flaps (Kanai *et al.* 2007). DmGEN was reported to not cleave four way junctions. In contrast, human GEN1 has been proposed to be the major human Holliday junction resolvase in humans because it can process four way junctions *in vitro*. These I domain of the GEN1 proteins is considerably smaller than those of FENs. However, as with other superfamily members, the critical active site lysine and arginine residues are conserved.

Proposed Roles of the Helical Arch

Although there is general agreement that the helical arch is an important feature of FENs and it must somehow interact with substrates, the precise role of the arch has been a controversial feature of FENs. Two general mechanisms have been proposed for interaction with flap substrates (figure 1.6.12). In the first model, FEN acts as a 5'-tracking enzyme, initially interacting with flap substrate by recognition of a free 5'-flap nucleotide and subsequent movement along the flap in a 'ratchet' fashion, where reaction occurs upon encountering duplex DNA (Murante *et al.* 1995). In an alternative proposal, FENs were

posited to initially bind the duplex regions of its substrate and then accommodate the 5'-portion of the substrate to form a cleavage competent complex (Ceska *et al.* 1996; Xu *et al.* 2001; Devos *et al.* 2007; Tsutakawa *et al.* 2011). However, such a mechanism poses the problem of how a ssDNA flap is forced through a small aperture (Orans *et al.* 2011). A further paradox is that FEN superfamily members XPG and GEN1 have been demonstrated to act upon bubbles and 4WJs respectively. These substrates have duplex 5' to the site of enzymatic action and cannot be acted upon by a mechanism requiring threading, implying an implausible scenario where an alternative mechanism would have to be employed for different enzymes throughout this superfamily. As an alternative to threading, the arch regions of FENs have been suggested to function as a clamp (figure 1.6.12) (Hosfield *et al.* 1998; Bornarth *et al.* 1999; Chapados *et al.* 2004; Liu *et al.* 2006). Some evidence in support of the threading hypothesis is provided by the structure of T4 RNase HI bound to pseudo Y DNA where the flap is positioned as though it would pass through. However, in this complex the arch is partially disordered and the substrate does not occupy the active site (Devos *et al.* 2007). Several experimental approaches have been used to interrogate the role of the helical arch and these are summarised below. These have largely focused on testing for the possibility of threading mechanisms.

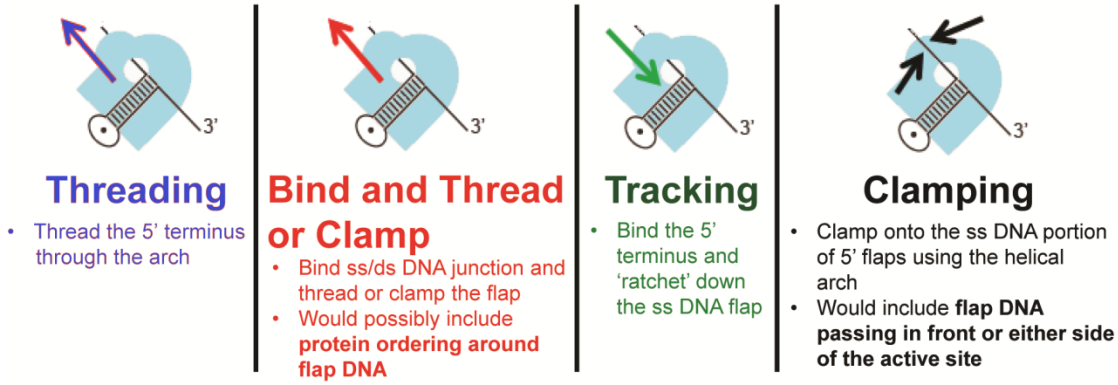


Figure 1.6.12: Proposed models of FEN action. From left to right, the threading, binds then thread/clamp, tracking and clamping models are described. In the threading model the 5' end of flap DNA is threaded through the helical arch. The bind then thread/ clamp model would include binding the ss/ds DNA junction and then thread or clamp the ssDNA portion of the substrate. The tracking and the clamping model involve binding the 5' terminus of flap DNA and tracking down or clamping either side of the helical arch.

Previous Threading Mechanism Tests

Streptavidin

Murante *et al.* claimed that the addition of streptavidin to the end of a 5' ssDNA flap completely abolished the flap endonuclease activity of FEN, presumably by preventing threading through the helical arch. This piece of evidence is in accord with the threading mechanism, although the data in this very publication shows a minor activity in the presence of streptavidin casting doubt on the result (Murante *et al.* 1995). Abolition of activity was claimed to be evidence that FENs operate using a tracking mechanism.

Gapped flap substrates

Early FEN literature claimed that hybridisation of oligonucleotides to 5' ssDNA flaps abolished FEN activity and this was interpreted to indicate a requirements for flaps that were not structured so threading/tracking along the 5' flap could occur (Murante *et al.* 1995). In addition to abolishing activity it was also claimed that preincubation of a 5’ flap substrate with FEN protein and then hybridization of an oligomer to the ss flap could trap the substrate on the enzyme. However more recent studies contradict this and have reported a robust gap endonuclease activity (Zheng *et al.* 2005; Finger *et al.* 2009) for hFEN1. Since gapped flaps contain a region of duplex, they appear to rule out the possibility of any mechanism based on threading substrates through the helical arch, as the structured arch is only large enough to accommodate ss and not ds DNA. Thus gap endonuclease activity has been cited as evidence for a clamping mechanism.

It was suggested that the gap endonuclease activity could be a result of the FEN proteins employing ds Exo activity on the duplex region of the 5' flap until the ss flap was eventually revealed for a conventional flap endonuclease reaction. Although this scenario is likely in some cases of reported gap endonuclease activity, work by Finger *et al.*, demonstrated unambiguous gap endonuclease activity (figure 1.6.13) when hFEN1 was used to cleave a forked gap substrate with a 3’ flap. This suggests that FEN proteins do not recognise ss flaps but rather bind the duplex regions of substrate first before accommodation of the flap region.

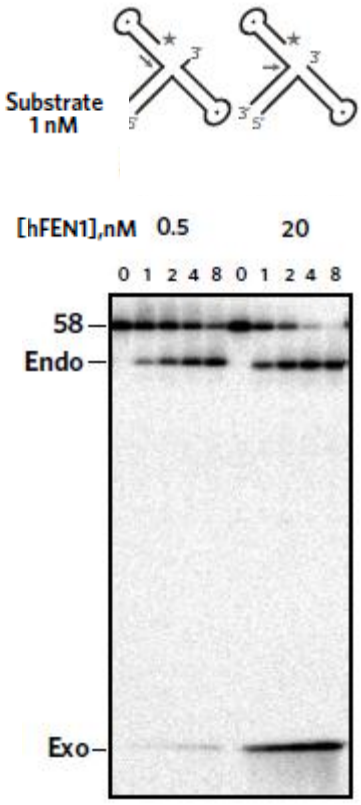


Figure 1.6.13: Gel showing the cleavage products of the reaction between a forked gap substrate with a 3' flap (left), and without (right) (Finger *et al.,* 2009).

Previous work has shown that even the phage enzyme T5 FEN possesses some gap endonuclease-like activity ((Sayers and Eckstein 1991; Garforth *et al.* 2001); Blanka Sengerova; Amanda Beddows, personal communication). T5 FEN was shown to cleave a double stranded circular plasmid with a nick exonucleolytically, creating a gapped substrate which T5 FEN then can cleave in a gap endo manner, cutting the intact strand of the plasmid, as shown below in figure 1.6.14. However, it is worth noting that reaction forcing conditions were used, with high concentrations of T5 FEN with respect to plasmid DNA, and the cofactor used was Mn^{2+}, not Mg^{2+} shown to give a much higher turnover number (Tock *et al.* 2003). This again contradicts the threading and tracking mechanism in a conventional sense and alludes strongly towards the clamp/bind at the ss/ds DNA junction first mechanism.

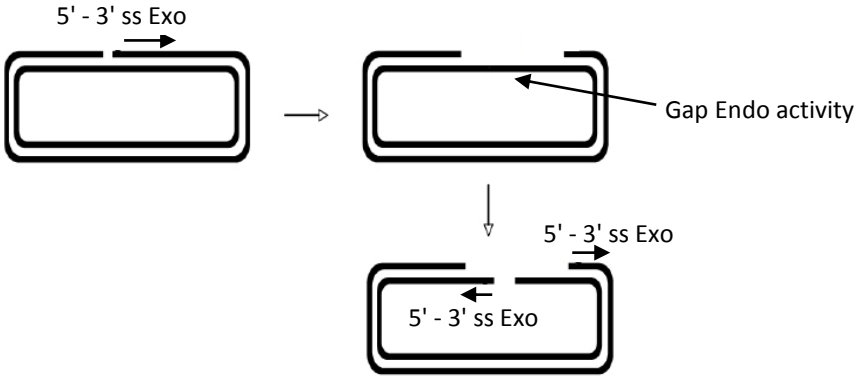


Figure 1.6.14: The locations of T5 FEN cleavage on a nicked circular plasmid substrate.

Branched Substrates

Bornarth *et al.*, 1999 used branches of 4-11 nucleotides in length, 26 and 33 nucleotides along a 60 nucleotide 5' ssDNA flap respectively, to distinguish between threading and tracking mechanisms (figure 1.6.15). Platinated branches were also used, linking adjacent G's and creating an upside down 'A' structure.

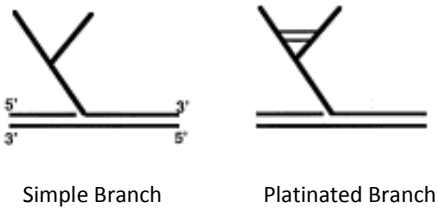


Figure 1.6.15: Examples of branched substrates used

Simple branch structures were cleaved normally by FEN, again conflicting with the threading mechanism, while rigid platinated branches were not cleaved at all. It was suggested that this could fit with the tracking mechanism, assuming the helical arch can become

unstructured in order to accommodate some blockages and not others that exceed the potential space available after disordering (Bornarth *et al.* 1999).

Adducts

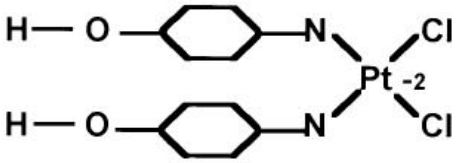


Figure 1.6.16: A cis-diamminedichloroplatin (CDDP) adduct used to add charge and steric bulk to 5' ssDNA flaps

CDDP adducts (figure 1.6.16) were added to 5' ssDNA flaps in an effort to block tracking and threading down to the site of cleavage. These adducts were placed in various positions along the 5' flap, ranging from adjacent to the cleavage site to the very end of the flap. Successful cleavage by FEN occurred in all cases, with the efficiency of cleavage predictably lowered when adducts were placed adjacent to the cleavage site. This was presumably due to the adduct disturbing the normal enzyme substrate interactions. These experiments again contradict the threading and tracking mechanism, but suggest that FEN can cleave despite blockages, possibly through clamping or passage through an unstructured arch. It should be noted that these experiments required reaction forcing conditions (Barnes *et al.* 1996; Bornarth *et al.* 1999).

Flexibility

Bornarth *et al.,* 1999 also incorporated cis-syn thymine-thymine cyclobutane dimers into an 11 nucleotide extension on 5' ssDNA flaps to create a more rigid ssDNA flap. The increased rigidity had little effect on the ability of FEN to cleave, indicating that flexibility of flaps has little effect on the enzymatic ability of FEN.

In summary, these experiments seem to show that blocking the 5' end of flaps using modifications such as streptavidin and platinated branches are inhibitory. This is consistent with threading or tracking mechanisms. However, the fact that some adducts and

modifications are inhibitory and some not (normal branched flap structures, gapped flaps some CDDP adducts, and the faint activity seen in the streptavidin experiments) appear at odds with the concept of threading substrates through a small aperture. The clamping mechanism might be a reasonable alternative (Barnes *et al.* 1996; Bornarth *et al.* 1999). Alternatively, the helical arch could also fold into different conformations, backwards for example, allowing much easier access to the active site of the enzyme. This would make threading not strictly necessary in all circumstances, but whether fast rates of reactions could still be in observed under these circumstances requires investigation (Dervan *et al.* 2002; Williams *et al.* 2007). This is especially so in view of superfamily conservations of arch residues at the base of α4.

Recent structures of hFEN1 and hEXO1 bound to substrate and products show interactions occur mainly with the complementary (during DNA replication termed the template strand) DNA strand of substrates. These observations favour mechanisms where 5'-nucleases bind the main portion of their substrate and not flaps first. Crystallisation of hFEN1 substrate complexes inhibited using Sm^{3+} ions within the active site, have shown that within the crystal environment the binding of substrate appears to be accompanied by the ordering of the hydrophobic wedge and helical arch of hFEN1. This suggests the possibility that flaps could thread through a disordered arch, which could then order, resulting in the helical arch enclosing the unannealed 5' flap. However, neither crystallographic study visualised an unpaired substrate in the active site and positioned for reaction. As a consequence two opposing opinions were offered with the Beese lab favouring a bind and then clamp mechanism and the Tainer lab suggesting a bind and thread mechanism with threading taking place through a disordered archway as the main portion of the substrate becomes accommodated (Orans *et al.* 2011; Tsutakawa *et al.* 2011).

The inconclusive nature of the experiments shown here and the abundance of conflicting hypotheses illustrate the need for further, more in depth investigation to elucidate the proper mechanism for FEN substrate selectivity and accommodation.

1.7 Aims

The precise mechanism by which FEN enzymes both select for and accommodate DNA substrates with a 5' ssDNA flap is still inconclusive. Moreover, the various different mechanisms posited cannot explain all data collected previous to this work.

This thesis aims to further disprove several models using a series of rigorous biochemical experiments and to gather evidence to support a new model that can account for all data. These firstly test the cleavage of DNA substrates with their 5' ends conjugated to streptavidin (SA), analogous to previous work but placing this method of substrate modification under greater scrutiny and the ability to cleave substrates with larger flaps than ssDNA. Secondly the competition of DNA substrates out of a series of enzyme-substrate complexes conjugated to SA in different manners is described. This provides insights into the manner of binding and consequently the method of substrate selection. Additional experiments are presented involving site directed mutagenesis of the human FEN1 enzyme in order to highlight the structural importance of a well ordered helical arch. Finally, quantifying substrate binding using fluorescence anisotropy measurements of interaction with both bacteriophage T5 FEN and hFEN1 highlight the major enzymatic and external influences to binding substrates.

Chapter 2: Methodology

2.1 Over-expression and Purification of FEN Proteins

2.1.1 T5 FEN

Buffers and Media required:

5yT media (1 L) -	40 g Tryptone
	25 g yeast extract
	6 g NaCl
Q-seph buffer A -	25 mM Tris-HCl pH8.0
	1 mM EDTA
	1 mM DTT
	5% (v/v) glycerol
	50 mM NaCl
Q-seph buffer B -	as Buffer A but 1 M NaCl
SP-Hep buffer A -	25 mM $NaPO_4$ pH7.5
	1 mM EDTA
	1 mM DTT
	5% (v/v) glycerol
	50 mM NaCl
SP-Hep buffer B -	as Buffer A but 1 M NaCl

Expression

4 x 5 ml 5yT containing 50 ug/ml carbenicillin were inoculated with glycerol stocks of *E. coli* BL21, which already contained the T5 FEN construct (a gift from Prof. Jon Sayers-University of Sheffield). These were grown for 4 hrs, at 25°C with 225 rpm shaking. 1 ml of this growth was used to inoculate 500 ml of 5yT, and these were grown under the same conditions until the media reached OD_{600} ~ 2.0. The cells were then heat shocked for 2 hrs at 42°C with 225 rpm shaking. The cells were harvested by centrifugation (6000 x g; 4°C; JLA rotor 10,500; 30 min) and the supernatant discarded. Cells were suspended in 400 ml PBS buffer and harvested again, (6000 x g; 4°C; JLA rotor 10,500; 30 min). The resulting pellet was

resuspended in 80 ml lysis buffer (250 mM Tris-HCL pH6.0; 4 mM EDTA; 250 mM NaCl; 5% (v/v) glycerol).

Purification

A spatula full of lysozyme was added to the cells suspended in lysis buffer, and incubated for 45 minutes at room temperature (RT) with occasional stirring. To this, 500 µl protease inhibitor cocktail IV (Calbiochem) was added and incubated for 5 minutes at RT. Then, DTT and sodium deoxycholate were added to a final concentration of 2mM and 0.5 mg/ml, respectively. The cell suspension was mixed gently for 5 minutes at RT. The cell suspension was then sonicated on ice for 4 x 30 s at 15 amp microns. The lysate was clarified by centrifugation to remove cell debris (30000 x g; 4°C; JLA rotor 25,50; 45 min). Protein was then precipitated by adding solid ammonium sulphate to a total concentration of 3.5M to the supernatant. The protein was pelleted by centrifugation (30000 x g; 4°C; JLA rotor 25,50; 30 min). The pellet was resuspended in a minimum volume of Q seph buffer A, and desalted using a HiTrap desalting column at 4°C.

The desalted protein was subjected to anion exchange chromatography using a 5 mL HiTrap Q-Sepharose column (GE Lifesciences). After equilibrating the column in Q-Seph buffer A, the protein was loaded and eluted using a linear gradient from 0%-100% Q-Seph buffer B over 150ml (30 column volumes; CV). The protein eluted between 20-30% Q-Seph B.

The fractions containing T5 FEN were pooled and diluted 1:2 (v:v) in SP-Hep A, and then loaded onto a tandem 5 mL SP-Sepharose and 5 mL Heparin-Sepharose column that had previously been equilibrated in SP-Hep buffer A (Pharmacia Amersham). The SP-Sepharose column was then removed. The protein was eluted from the Heparin-Sepharose column using a 100 mL (10 CV) linear gradient from 0%-100% SP-Hep B buffer. T5 FEN eluted at approximately 60% SP-Hep B buffer. The SP-Sepharose column was then cleaned with 100% SP-Hep B buffer, and the fractions analysed by SDS-PAGE to ensure no T5 FEN was present.

All T5 FEN fractions were analysed by SDS-PAGE, and further purified, if necessary, with a Q-Sepharose column. The pure T5 FEN fractions were then pooled and concentrated using a VivaSpin centrifugal concentrator, (Vivascience) and then diluted 1:1 with 80% (v/v) glycerol (sterile-autoclaved) for storage at -20°C.

2.1.2 Wild Type and mutant human FEN1 Protein Purification

Buffers and Media required:

LB (1L)	10g tryptone, 5g yeast extract, 10g NaCl
SOC media:	20 g tryptone, 5g yeast extract, 0.5g NaCl, 10mL 250 mM KCl, 10mL 1M $MgCl_2$, 10 mL 1M, 20 mM D- glucose
IMAC FF Buffer A:	20mM Tris pH = $7.0^{25ºC}$, 1M NaCl, 0.02% NaN_3, 5mM Imidazole, 5mM 2-mercaptoethanol
IMAC FF Buffer B:	20mM Tris pH = $7.0^{25ºC}$, 500mM NaCl, 0.02% NaN_3, 40mM Imidazole, 0.1% Tween20, 5mM 2-mercaptoethanol
IMAC FF Buffer C:	250mM Imidazole pH = $7.2^{25ºC}$, 500mM NaCl, 0.02% NaN_3, 5mM 2-mercaptoethanol
Hitrap Heparin HP Buffer A:	50mM MES pH = $6.0^{25ºC}$, 1mM EDTA, 0.02% NaN_3, 20 mM 2-mercaptoethanol
Hitrap Heparin HP Buffer B:	50mM MES pH = $6.0^{25ºC}$, 1mM EDTA, 0.02% NaN_3, 1M NaCl, 20 mM 2-mercaptoethanol
HiTrap Phenyl Sepharose HP Buffer A:	20mM Tris pH = $7.4^{25ºC}$, 1.5M $(NH_4)_2SO_4$, 1mM EDTA, 0.02% NaN_3, 20 mM 2-mercaptoethanol
HiTrap Phenyl Sepharose HP Buffer B:	20mM Tris pH = $7.4^{25ºC}$, 1mM EDTA, 0.02% NaN_3, 20 mM 2-mercaptoethanol
Sephacryl GF Buffer A:	100mM HEPES, pH = $7.5^{25ºC}$, 200mM KCl, 0.04% NaN_3, 20 mM 2-mercaptoethanol
Storage Buffer:	50% glycerol, 50mM HEPES, 100mM KCl, 0.02% NaN_3, 5mM Tris(hydroxypropyl)phosphine

25 ml of LB media that was supplemented with 34 µg/ml Chloramphenicol (Cm) and 25 µg/ml Kanamycin (Kan) was inoculated with a single colony of Rosetta-pET28b-hFEN1-L97P/L111P/L130P. The culture was allowed to grow for overnight (~14-16 hours) at 37°C. Four 2 L culture flasks each containing 500 ml LB media supplemented with 34 µg/ml Cm and 25 µg/ml Kan were each inoculated with 5ml of the overnight Rosetta-28b-hFEN1 L97P/L111P/130P culture. The cultures were grown at grown at 37°C until an OD_{600} of 0.6 was achieved, at which point protein expression was induced by adding IPTG to a final concentration of 0.5 mM to each of the cultures. The cultures were then incubated overnight at 18°C.

The cells were pelleted by centrifugation (6000 × g, 30 min, 4°C) in two 1 L centrifuge bottles, and the supernatant was poured off. Each cell pellet was re-suspended in 40 ml of ice-cold 1X PBS. The cells were pelleted again by centrifugation (4000 x g, 20 minutes, 4°C) and the supernatant poured off. The cell pellets were weighed to estimate the total wet cell paste isolated (~1-4 g of wet cell paste from 2L). The two cell pellets (1 per 50 ml Falcon tube) were each suspended in 40 ml of IMAC FF Buffer A. To each suspension, 200 µL of protease cocktail inhibitor VII (Sigma) (50 µL inhibitor cocktail/g of wet cell paste) and 5 ml of lysozyme (10 mg/ml) were added. Cells were incubated on ice for three hours, and then, the cells were then frozen at -20°C overnight. To prepare the cell lysate for protein purification, the suspensions were incubated in cold tap water until it almost completely thawed, at which time the viscous suspensions were placed in ice. The suspensions were sonicated on ice 10 times at 50% power with 10 second bursts with at least 30 seconds between each burst. To each lysate, 5ml of IMAC FF Buffer A containing 1% Tween20 was added. Insoluble cell debris was removed by centrifugation (30,000 × g, 4°C). The presence of the target protein in the cleared lysate was confirmed by SDS-PAGE.

All protein purification steps were conducted in a cold room using an Akta FPLC (GE Lifesciences). A Chelating Sepharose Fast Flow (GE Lifesciences) column (1.6 cm ID, 10 cm length), was charged with Ni^{2+} ions according to the manufacturer's protocol. The column was equilibrated with 5 column volume (CV) of IMAC FF Buffer A. The clarified lysate (~100 ml) was then applied to column, and subsequently, washed with 7 CV of IMAC FF Buffer A

and 5 CV of IMAC FF Buffer B. The target protein was eluted as a single fraction with 5CV of IMAC FF Buffer C. The eluted fraction was then diluted with an equal volume of ice cold water, and then applied directly to HiTrap Heparin HP (3X 5 ml in tandem) using HiTrap Heparin Buffer A. The protein was eluted using a 50 CV linear NaCl gradient (0 to 1 M NaCl) using HiTrap Heparin Buffers A and B and collected as 2.5 ml fractions. Fractions containing hFEN1 mutants were pooled, and the ionic strength of the solution was increased by the slow addition of finely-ground solid $(NH_4)_2SO_4$ on ice with stirring to approximately 38% saturation (~1.5 M). The amount of $(NH_4)_2SO_4$ necessary for this was calculated using the ENCor Biotechnolgy Inc. Ammonium Sulfate Calculator (http://www.encorbio.com/protocols/AM-SO4.htm), which takes into account initial percentage saturation of ammonium sulphate, initial volume of the sample, the partial specific volume of the solid ammonium sulphate added, and temperature at which the procedure is conducted to calculated the mass of ammonium sulfate necessary to reach the desired saturation. The final solution was filtered using a 0.22 µM syringe filter, and then, applied to a HiTrap Phenyl Sepharose HP column (5X 5 ml in tandem) using HiTrap Phenyl Sepharose HP Buffer A. The protein was eluted from the column by an inverse linear salt gradient generated using HiTrap Phenyl Sepharose HP Buffers A and B, and was collected in 10 ml fractions. Protein-containing fractions were pooled and concentrated by ultrafiltration using a 250 ml Amicon Ultrafiltration cell with 10,000 MWCO PES membrane (Millipore) pressurized with N_2 (40 PSI). The volume of pooled fractions was reduced to less than 1ml. The retentate from the Amicon cell was then applied to a Sephacryl S-100 (1.6cm ID x 60 cm) size exclusion column (GE Lifesciences) and isocratically eluted using Sephacryl GF buffer. Protein containing fractions were pooled and concentrated using a 50 ml Amicon Ultrafiltration cell with 10,000 MWCO PES membrane (Millipore) pressurized with N_2 (40 PSI).

Protein concentration was determined by absorbance at 280 nm using a Nanoview spectrophotometer and the calculated extinction coefficient (22,920 M^{-1} cm^{-1}). The volume of the hFEN1 sample was adjusted with the appropriate volume of glycerol, 0.5M THP, and Sephacryl GF buffer supplemented to adjust the final protein concentration to 100µM, the glycerol concentration to 50% (v/v), the THP concentration to 5 mM, and the Sephacryl GF buffer components to 50mM HEPES pH=7.5, 100 mM KCl, 0.02% (w/v) NaN_3 (i.e., storage buffer).

2.2 Determination of protein concentration

Protein concentrations of glycerol stocks were confirmed using the Bradford assay according to the manufacturers protocol (Bradford, 1976). Aliquots (100 µl) of bovine serum albumin (BSA) of various concentrations (0-0.01 mg/ml) were prepared, along with 100 µl samples of hFEN1 at various dilutions to ensure that the signal would be within the range of BSA standards. ddH_2O water (700 µl) and Bradford reagent (200 µl of 1:4 Bradford dye:water) were added together and allowed to stand for 15 minutes.

The Bradford reagent solutions were added to the purified protein and BSA solutions and the A_{595} measured using a Cary Bio UV-Vis spectrophotometer (Varian). Using the BSA samples, a calibration curve of A_{595} against BSA concentration was generated. The protein concentrations of hFEN1 stocks were interpolated from the BSA standard curve.

The wild-type T5 FEN protein batches that were used for the kinetic parameter determinations are from November 2006 (58.4 µM; purified by Karl Syson), December 2007 (38.5 µM; purified by Dr. Sengerova), June 2009 (50 µM purified by myself), April 2010 (42 µM purified by myself) and February 2011 (207 µM; purified by Dr. Atack). Wild-type hFEN1 batches were purified by Dr. Atack, and later Dr. Finger; mutant FEN1 batches were purified by Dr. Finger and I. All hFEN1 stocks were confirmed to be 100 µM.

2.3 Synthesis of Oligonucleotide Substrates

The oligonucleotides used in this work (table 2.3) were synthesised by Elaine Frary (Sheffield, Chemistry) or DNA technology, Denmark. DNA manufactured in house was prepared using an ABI model 394 DNA/RNA synthesiser on a 1 µmol scale. For oligonucleotides with a 5' FAM label (figure 2.3.1), and in some cases, a 5' biotin too (figure 2.3.2), the appropriate phosphoramidites were obtained from Glen Research (Sterling, Virginia). For the oligonucleotide prepared with a 3'-biotin with spacers (figure 2.3.3 A-B), the appropriate phosphoramidites were obtained from Link technologies, Lanarkshire, Scotland. 5'-Dimethoxytrityl-N-benzoyl - 3'-deoxynucleoside - 2'-(2-cyanoethyl-*N*,*N*-diisopropyl) phosphoramidite monomers were used, with a dimethylformamidine group protection for guanine and benzoyl protection for adenine and cytosine. The synthesis procedures were carried out using the manufacturer's instructions.

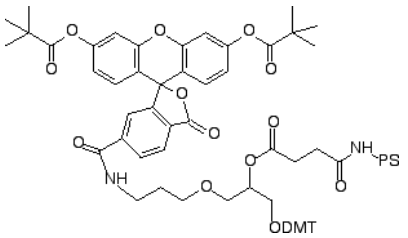


Figure 2.3.1: 6-FAM-ps used in oligonucleotide synthesis.

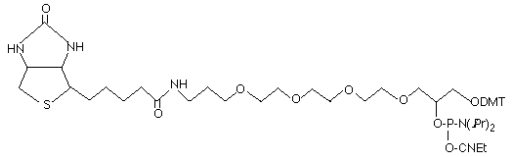


Figure 2.3.2: Biotin TEG phosphoramidite used in oligonucleotide synthesis.

(A)

(B)

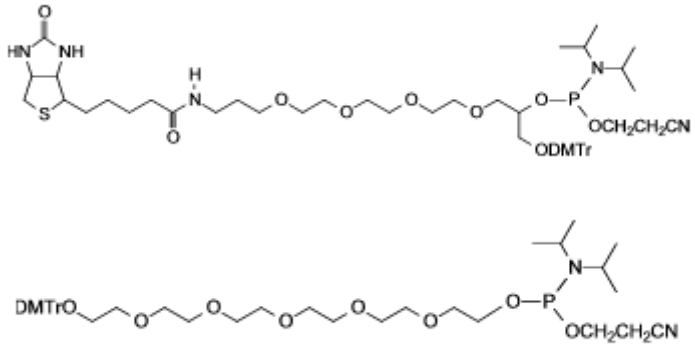


Figure 2.3.3: (A): 3' Biotin TEG phosphoramidite used in oligonucleotide synthesis, (B) Phosphoramidite spacer used in oligonucleotide synthesis.

2.4 Purification of Oligonucleotide Substrates

After deprotection in ammonia solution, the oligonucleotide solutions were evaporated to dryness and then resuspended in 1 ml dH_2O to prepare them for purification via reversed-phase ion pairing high pressure liquid chromatography (rpHPLC). The purification was carried out on a µ-Bondapak C18 column (Waters Chromatography). This was done using a custom gradient, using the following solvent system shown below,

Buffer A: 5% Acetonitrile, 100mM TEAAc, pH6.5

Buffer B: 65% Acetonitrile, 100mM TEAAc, pH6.5 (see below for solvent systems)

Solvent system used for rpHPLC purification of FAM labelled/biotinylated substrates (carried out at 45°C for large DNA substrates, RT for the rest) using the above buffers: 0 min 5% B, 30 mins 20% B, 35 mins 100% B, 45 mins 100% B 50 mins 5% B, 60 mins 5% B.

Solvent system used for rpHPLC purification of DNA substrates manufactured DMT on (carried out at RT) using the above buffers: 0 min 5% B, 30 mins 50% B, 35 mins 100% B, 45 mins 100% B 50 mins 5% B, 60 mins 5% B.

First an analytical (5-10 µl of a total 1 ml solution) run was performed with the UV detector wavelength set at 260 nm and sensitivity at 0.05, followed by preparative runs of 100-150 µl with the UV detector wavelength set to 280 nm and sensitivity set to 2.0. Peaks corresponding to the desired product were pooled, dried and then desalted using a NAP-10 column (Sephadex) according to the manufacturer's direction. After purification by rpHPLC, non-labelled oligonucleotides containing a DMT protecting group on the 5' end were suspended in 1 ml 20% acetic acid solution and incubated for 2 hours to remove the DMT. Oligos were then desalted as described above. All oligonucleotides were characterised using MALDI-ToF spectrometry, by comparing the expected masses with the calculated mass obtained shown in tables 2.1, 2.2 below.

Nucleotide / Modification	**Unit MW / A.U (isotopic averages)**	**E_{260} / $Lmol^{-1}cm^{-1}$**
FAM	537.46	21,000
dA	313.21	15,400
dG	329.21	11,700
dC	289.18	7,300
dT	304.2	8,800
Biotin	569.61	N/A
Pspacer	344.30	N/A

Table 2.1: Masses and extinction coefficients used for the calculating the masses of the synthesised oligonucleotides

Oligonucleotide Code	Calculated Mass	Actual Mass	E_{260} / $Lmol^{-1}cm^{-1}$	T_m at 37°C
pY-7B	13629.21	13630	442.6	82.4
Prod 7	3508.68	3508	95.06	n/a
pY-7	13059.6	13059	442.6	82.4
pY-21B	17867.95	17874	592.7	74.3
Prod 21	7747.42	7760	245.7	n/a
pY-21	17298.34	17323	592.7	74.3
COMP	16760.88	16764	571.7	84.6
COMPB	17322.8	17322	565.2	84.6
pY-21-3'B	18892.77	18903	565.9	74.3
3'OH-6	10920.2	10949	355.9	85.9
3'-FAM-OH-6	11031	11031.001	355.9	75.5
DF-21	11321.8	11266	648.5	50.9
DF-21B	11351	11449	648.5	50.9
DFGENB	13213.1	13217	717.2	46
DFGEN	12643.47	12647	717.2	46
DF-5	7651.64	7651	585.3	74.6
DF-5B	8221.25	8221	585.3	74.6
DF-3	7043.55	7042	568.2	78.3
DF-3B	7613.55	7620	568.2	78.3
DF-COMP5	7114.18	7112	585.3	74.6

Table 2.2: Details of all the oligonucleotides synthesised and purified for this work. Naming system of the oligonucleotides: DF = double flap, pY = pseudo Y, OH = overhang, B = 5'-biotin moiety, number denotes the size of the 5' unannealed nucleotide flap/overhang. T_m stated are calculated (M. Zucker, 2003) for DNA at a final concentration of 5 nM, at 50 mM KCl for pseudo Y substrates and 100 mM KCl for double flap substrates designed for hFEN1.

Oligonucleotide Code	**Sequence**
pY-7B	5'-F-B*–d(CGCTGTCGAACACACACCGCTTGCGGTGTGTGTTCCACAAC)3'
Prod 7	5'-F-B*–d(CGCTGTCG)3'
pY-7	5'-F–d(CGCTGTCGAACACACACCGCTTGCGGTGTGTGTTCCACAAC)3'
pY-21B	5-'F-B*–d(CGCTGTCCTCTCGAACACACAGAACACACACCGCTTGCGG TGTGTGTTCCACAAC)3'
Prod 21	5'-F-B*–d(CGCTGTCCTCTCGAACACACAG)3'
pY-21	5'-F–d(CGCTGTCCTCTCGAACACACAGAACACACACCGCTTGCGGTGTGT GTTCCACAAC)3'
COMP	5'-d(CGCTGTCCTCTCGAACACACAGAACACACACCGCTTGCGGTGTGTGT TCCACAAC)3'
COMPB	5'-B*d(CGCTGTCCTCTCGAACACACAGAACACACACCGCTTGCGGTGTGT GTTCCACAAC)3'
pY-21-3'B	5'-F–d(CGCTGTCCTCTCGAACACACAGAACACACACCGCTTGCGGTGTGT GTTCCACAAC)-Pspacer-Pspacer-Pspacer-B*-3'
3'OH-6	5'-F-P*-ACACACCGCTTGCGGTGTGTGTTCCACAAC-3'
3'-FAM-OH-6	5'-P*-ACACACCGCTTGCGGTGTGTGTTCCACAAC-F-3'
DF-21	5' F-GGATGTTATCTTTATGTTACTTTGAGGCAGAGT 3' Flap strand 5' CCTGCCAAAGTGGCAGAACTCCGTCTCA 3' Template
DF-21B	5' F-GGATGTTATCTTTATGTTACTTTGAGGCAGAGT-B* 3' Flap strand 5' CCTGCCAAAGTGGCAGAACTCCGTCTCA 3' Template
DFGEN	5' TTGGCATAGGGACTATGCCAATTTTTTTTGAGGCAGAGT-F 3' Flap strand 5' CCTGCCAAAGTGGCAGAACTCCGTCTCA 3' Template
DFGENB	5' TTGGCATAGGG(B*)ACTATGCCAATTTTTTTTGAGGCAGAGT-F 3' Flap strand 5' CCTGCCAAAGTGGCAGAACTCCGTCTCA 3' Template
DF-5	5' F-TTTTTGCGACCTGGGCTGTGGAG 3' Flap strand 5' CTCCACAGCCCAGGTCGCGACGGTGAAACCGTCC 3' Template
DF-5B	5' F-B*-TTTTTGCGACCTGGGCTGTGGAG 3' Flap strand 5' CTCCACAGCCCAGGTCGCGACGGTGAAACCGTCC 3' Template

DF-3	5' F-TTTGCGACCTGGGCTGTGGAG 3' Flap strand 5' CTCCACAGCCCAGGTCGCGACGGTGAAACCGTCC 3' Template
DF-3B	5' F-B*-TTTGCGACCTGGGCTGTGGAG 3' Flap strand 5' CTCCACAGCCCAGGTCGCGACGGTGAAACCGTCC 3' Template
DF-COMP5	5' TTTTTGCGACCTGGGCTGTGGAG 3' Flap strand 5' CTCCACAGCCCAGGTCGCGACGGTGAAACCGTCC 3' Template

Table 2.3: Sequences of each oligonucleotide used. F=FAM, B*=Biotin, Pspacer = 18 atom phosphoramidite spacer, P*-Phosphate

The concentration of each oligonucleotide was measured using the equation,

$$A_{260} = \varepsilon_{260}\, c\, l$$

A = the absorbance at 260 nm, ε = the extinction coefficient at 260 nm measured in $Lmol^{-1}cm^{-1}$, c = the concentration in $molL^{-1}$ and l = the path length in cm, each absorbance was measured with a Cary Bio UV-Vis spectrophotometer (Varian).

2.5 Determination of the kinetic parameters of FEN catalysed reactions of altered substrates

5 µM solutions of unimolecular T5 FEN substrates were annealed by heating to 95°C for 90-120 seconds, and cooled on ice for 5 minutes. 5 µM stocks of bimolecular hFEN1 double flap substrates (containing a ratio of 1:1.1 of flap: template strands) were annealed by heating at 95°C for 2 mins and cooling to RT over approximately 30 mins. These steps were carried out in 250 mM KCl, 50 mM HEPES pH7.5 or potassium glycinate (KGly) pH 9.3. For reactions concerning T5 FEN, solutions containing concentrations of FAM labelled unimolecular or bimolecular oligonucleotides, ranging from 1/8 K_M and 10 x K_M were then prepared in solution containing final concentrations of 25 mM HEPES pH7.5 or KGly pH 9.3, 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Mg^{2+}. These reaction mixtures were then pre-incubated for approximately 10 mins at 37°C. Reactions concerning hFEN1 were treated identically; however, final solution concentrations were 50 mM HEPES pH7.5, 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 mM Mg^{2+}.

Enzyme preps were stored in 25 mM HEPES pH7.5, 200 mM KCl, 50% glycerol and 0.04% sodium azide, and for T5 FEN were diluted to a concentration of 0.1-10 nM in a solution of 25 mM HEPES pH7.5 or KGly pH 9.3, 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Mg^{2+}. For hFEN1, enzyme preps were diluted in solutions of 50 mM HEPES pH7.5, 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 mM Mg^{2+}. Reactions were initiated by addition of enzyme to the oligonucleotide substrate solutions at 37˚C. 10-100 µl aliquots (dependent on substrate concentration) of the reaction mixture were removed at 8 time intervals. The aliquots were quenched in 25 µl - 40 µl (dependant on amount of T5 FEN in solution), 250 mM EDTA.

Multiple turnover reactions with streptavidin were carried out in the same manner, with 5 equivalents of streptavidin added after annealing of the oligonucleotide substrate with subsequent incubation on ice for 2 minutes. Aliquots resulting from reactions containing streptavidin bound oligonucleotides were quenched in 8 M Urea and 80 mM EDTA.

pH and Buffer identity	**Range of Substrate/ Enzyme Concentrations**	
	Substrate/ nM	**Enzyme/ pM**
7.5 – HEPES	5-800	60-1000
9.3 – Glycine	5-800	30-250

Table 2.4: Range of concentrations used in multiple turnover experiments to determine kinetic parameters of the reaction between the oligonucleotide substrates and T5 FEN. For reactions with hFEN1, enzyme concentrations used were generally 10x lower, while substrate concentrations remained the same.

A dHPLC (Wave® fragment analysis dHPLC fitted with a fluorescence detector; Transgenomic, Glasgow) was used to separate product and starting material using the FAM fluorescent tag for detection (figure 2.3.1) (excitation wavelength 494 nm, emission, 525 nm). These were separated using a gradient with an acetonitrile or a methanol solvent system on a DNAsep® (Transgenomic, Glasgow) column.

Methanol

Buffer A: 0.1% Methanol, 2 mM EDTA, 2.5 mM tetrabutyl ammonium bromide (tBABr) – dissolved in water

Buffer B: 85% Methanol, 2.5 mM tetrabutyl ammonium bromide – dissolved in methanol

Acetonitrile

Buffer A: 0.1% Acetonitrile, 1 mM EDTA, 2.5 mM tetrabutyl ammonium bromide – dissolved in water

Buffer B: 70% Acetonitrile, 1 mM EDTA, 2.5 mM tetrabutyl ammonium bromide – dissolved in water

Method used with methanol or acetonitrile at 50°C to separate starting oligonucleotides of around 40 nucleotides and products from around 4-22 nucleotides: 0 min 5% B, 5 mins 30% B, 9 mins 50% B, 12 mins 70% B, 13.5 mins 100% B, 14.5 mins 100% B, 14.6 mins 5% B, 17 mins 5% B. Products eluted around 6-13 minutes dependant on the size, starting material at approximately 15-17 minutes.

Method used with methanol or acetonitrile at 50°C to separate starting oligonucleotides of around 54 nucleotides and 21/22mer products: 0 min 30% B, 2.5 mins 57% B, 4 mins 50% B, 17.5 mins 65% B, 19 mins 100% B, 20 mins 100% B, 20.5 mins 5% B, 22.5 mins 5% B. Products eluted at 12 minutes, starting material at 16 minutes.

Initial rates of reaction were determined for each substrate concentration by plotting concentration of product against time, for the first 10% of product formed during the reaction. The initial rate was determined from the gradient of the resulting plot. Kinetic parameters were determined by fitting this data to the Michaelis-Menten equation (equation 1).

$$\frac{v}{[E]} = \frac{k_{cat}[S]}{K_M + [S]} \qquad \textbf{(1)}$$

Curve fitting was carried out using Kaleidagraph software. (Synergy Software, Reading, USA)

2.6 Measurement of FEN-Substrate binding by Fluorescence Anisotropy

Anisotropy (*r*) was measured using a Fluoromax-3 (Horiba, Jobin Yvon, Middlesex, UK) equipped with a polariser accessory. The program Fluoressence (Horiba Scientific, Middlesex, UK) was used to control the spectrometer. Excitation and emission wavelengths were set to 490 nm and 510 nm, respectively. Six measurements were made over 1 min. Slit widths were set at 10 nm. Binding measurements were carried out in a 0.5 ml quartz cuvette. Prior to each experiment, substrates were annealed as described in section 2.5. Solutions of 1-200 nM oligonucleotide were then made in a solution of 25 mM HEPES pH 7.5 or KGly pH 9.3 containing 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 1 mM EDTA or 10 mM $CaCl_2$ for T5 FEN and 50 mM HEPES pH 7.5 containing 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 1 mM EDTA or 10 mM $CaCl_2$ for hFEN1. Calculated molar equivalents of enzyme, competitor or streptavidin (in an identical buffer to above) were added in 0.5-5µl quantities (50-100 nM T5 FEN, 0.1 equivalents of streptavidin or 1-50 nM hFEN1) with subsequent measurement of I_{VH}, I_{VV}, I_{HV}, I_{HH} and I_{TOT}. I_{VH}, I_{VV}, I_{HH} and I_{HV} were used to calculate *r* according to equation 2.

$$r = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}} \qquad \textbf{(2)}$$

The anisotropy data for titrations where hFEN1 / T5 FEN was added to labelled DNA were fitted (Kaleidagraph, Synergy Software, Reading, PA, USA) to equation 3,

$$r = \frac{r_{min} + (r_{max} - r_{min})\left[([S] + [E] + K_D) - \sqrt{([S] + [E] + K_D)^2 - 4[S][E]}\right]}{2[S]} \qquad \textbf{(3)}$$

where *r* is the measured anisotropy at a particular total concentration of hFEN1 / T5 FEN ([*E*]) and fluorescently tagged ligand ([*S*]), r_{min} is the minimum anisotropy of free ligand, and r_{max} the maximum anisotropy when the ligand is saturated with protein). K_D = the dissociation constant of the ligand under investigation.

2.7 Measurement of single turnover rates of FEN catalysed reactions

2.7a Separate mixing experiments

Rapid quench experiments were carried out at 37^{o}C using a RQF-63 quench flow device (Hi-Tech Sci Ltd., Salisbury, UK). An 80µl aliquot of enzyme in reaction buffer (25 mM HEPES pH 7.5 containing 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Mg^{2+} was used for T5 FEN; 50 mM HEPES pH 7.5 containing 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 mM Mg^{2+} for hFEN1) was mixed with an equal volume of oligonucleotide substrate (annealed as described in section 2.5) in the appropriate reaction buffer. Enzyme was used at final concentration of at least 500 nM – 1 µM (~10 x K_M for most substrates), and FAM-labelled oligonucleotide substrates were used at a final concentration of approximately 5 nM (~1/10 K_M for most substrates), down to a minimum of 1 nM to avoid signal detection issues. After a controlled time delay of 6.4 ms to 51.2 s, 80 µl of quench solution (1.5 M NaOH, 60 mM EDTA or 8 M Urea, 80 mM EDTA) was added. Solutions of enzyme, substrate and quench were held in a temperature-controlled water bath set to 37°C within the instrument during the reactions. The quenched reaction mixtures were recovered and then, analysed using denaturing HPLC as described previously in section 2.5.

2.7b Measuring the decay of enzyme substrate complexes

To measure the decay of premixed enzyme substrate complexes, reactions were carried out by pre-incubating enzyme and substrate (annealed as in section 2.5) on ice for 2 minutes when E = T5 FEN and at RT for 2 minutes when E = hFEN1 in reaction buffer: (25 mM HEPES pH 7.5 containing 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 1 mM EDTA for reactions concerning T5 FEN (**EDTA buffer**); 50 mM HEPES pH 7.5 containing 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 2 mM Ca^{2+} for reactions concerning hFEN1 (**Ca buffer**)). To initiate reaction, the pre-incubated mixtures were mixed with an equal volume of 2x Mg^{2+} buffer: (25 mM HEPES pH 7.5 containing 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 20 mM Mg^{2+} for reactions concerning T5 FEN (**T5 Mg buffer**); 50 mM HEPES pH 7.5 containing 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 16 mM Mg^{2+} for reactions concerning hFEN1 (**hFEN Mg buffer**)). Using the RQF-63, aliquots were quenched at the specified times and analysed by dHPLC.

2.7c Measuring the decay of 'trapped' enzyme substrate complexes

Single turnover experiments of enzyme substrate complexes trapped by streptavidin were measured by pre-incubating enzyme and substrate (annealed as in section 2.5) on ice for 2 minutes when E = T5 FEN, and at RT for 2 minutes when E = hFEN1 in reaction buffer (**EDTA buffer** for T5 FEN; **Ca buffer** for reactions concerning hFEN1). This was followed by the addition of 5 equivalents of streptavidin (in a buffer containing 25 mM HEPES-NaOH pH 7.5 and 100 mM KCl), and further incubated at RT for 1 minute. The reactions were then carried out as described above in 2.7b.

2.7d Measuring the decay of 'blocked' enzyme substrate complexes

Single turnover experiments of substrate bound to streptavidin and then, bound to enzyme were measured by incubating 5 equivalents of streptavidin and biotinylated substrate (in a buffer containing 25 mM HEPES pH 7.5 and 100 mM KCl), for 1 minute at RT, followed by the addition of enzyme in reaction buffer (**EDTA buffer** reactions concerning T5 FEN; **Ca buffer** for reactions concerning hFEN1) and further incubation on ice for 2 minutes for T5 FEN and at RT for 2 minutes for hFEN1. The reaction was initiated by adding an equal volume of magnesium buffer (**T5 FEN/hFEN1 Mg buffer** contents described in section 2.7b). Aliquots were taken over a period of 30 minutes and 15 minutes for T5 FEN and hFEN1 reactions, respectively. Reactions were quenched using a solution of 8 M Urea and 80 mM EDTA. These aliquots were analysed using dHPLC.

The product formed over time in all single turnover experiments were fitted to equation 4,

$$P_t = P_\infty (1\text{-}\exp^{-kST.t}) \quad \textbf{(4)}$$

to determine the first order rate of the reaction. Here, P_t is the amount of product at time *t*, P_∞ is the amount of product at time ∞ (end point), and k_{ST} is the single-turnover rate of reaction.

In reactions 2.7a-d the final concentrations of [E] = 500 nM and the final concentrations of [S] = 5 nM.

2.8 Competition experiments

All substrates were treated with chelex beads (5 g/ 100 ml) for one hour with shaking at RT prior to competition experiments. Substrates were diluted to 5-20 μM in 100 mM KCl and 25 mM HEPES pH 7.5. Prior to the use of substrates, an aliquot of sufficient volume was removed and annealed as described in section 2.5.

2.8a General T5 FEN competition scheme

Enzyme (500 nM final concentration) and substrate (5 nM final concentration) were incubated on ice for two minutes in reaction buffer, (**EDTA buffer**). 5 equivalents of streptavidin (in a buffer containing 25 mM HEPES pH 7.5 and 100 mM KCl), were added to this mixture and incubated on ice for 1 minute, prior to the addition of competitor (final conc 2.5 μM). In reactions where streptavidin was added after competitor, 5 equivalents were added and incubated at 37 ^{o}C for 1 minute. Competitor was added at this point if required, and incubated at 37 ^{o}C for 5 minutes. Once the pre-reaction complex had been assembled, and competitor (2.5 μM total concentration) added if necessary, an equal volume of magnesium buffer (**T5 Mg buffer**) was added, using the RQF-63 to quench (8 M Urea, 80 mM EDTA) the reaction after a time controlled delay of 236, 836 and 2360 ms. Aliquots of the reaction were removed and analysed as described in section 2.7a.

2.8b General hFEN1 competition scheme

Enzyme (500 nM final concentration) and substrate (5 nM final concentration) were incubated at RT for two minutes in reaction buffer, (**Ca buffer**). 5 equivalents of streptavidin were added to this mixture and incubated for 1 minute at RT, prior to the addition of competitor (final conc 5 μM). In reactions where streptavidin was added after competitor, 5 equivalents were added and incubated at 37 ^{o}C for 1 minute. Competitor was added at this point if required, and incubated at 37 ^{o}C for 10 minutes. Once the pre-reaction complex had been assembled and competitor added if necessary, an equal volume of magnesium buffer (**hFEN1 Mg buffer**) was added to initiate reaction. The RQF-63 was used to quench (8 M Urea, 80 mM EDTA) the reaction after a time controlled delay of 51, 111, 167, 236, 436 or 836 ms. An aliquot of the reaction was removed and analysed as described in section 2.7a.

For T5 FEN and hFEN1, when streptavidin or competitor was not added, an equal volume of **EDTA / Ca^{2+} buffer** was added in order to keep all reaction conditions constant at all times.

2.9 DNA PAGE gels

200 pM of oligonucleotide substrate was loaded onto gels. Non denaturing DNA PAGE gels (20% acrylamide (w/v) containing a ratio of acrylamide:bis-acrylamide of 19:1) were run at 50 V, until the bromophenol blue dye ran to the bottom of the gel. Denaturing DNA PAGE gels (20% acrylamide (w/v) containing a ratio of acrylamide:bis-acrylamide of 19:1, and 9M urea) were run at 10W, until the bromophenol blue dye ran to the bottom of the gel. Gels were stained using toluidine blue in 5% (v/v) acetic acid for approximately 30 mins, and then destained with 5% (v/v) acetic acid for approximately 1 hour. Alternatively, gels were stained using ethidium bromide for 0.5-1 hour, and visualised using UV light.

2.10 Plasmid Mutation and Isolation

The vector pET-28b-hFEN1-WT (Kan^R), from which hFEN1 with a C-terminal $(His)_6$ affinity tag can be produced, was isolated using a Qiagen miniprep kit. The mutagenesis primers (table 2.5) were designed using www.genomics-agilent.com and ordered from www.invitrogen.com. These primers were used to synthesise the three plasmids that should give the desired proline mutants.

Primer Name	Primer-Template Duplex
L97P	5'-caagtcaggcgagccggccaaacgcagtg-3' ||||||||||||||| ||||||||||||||| cgagttcagtccgctcgaccggtttgcgtcactcg
L97P_antisense	gctcaagtcaggcgagctggccaaacgcagtgagc ||||||||||||||| ||||||||||||||| 3'-gttcagtccgctcggccggtttgcgtcac-5'
L111P	5'-gcagagaagcagccgcagcaggctcag-3' |||||||||||||| |||||||||||||||| ctccgtctcttcgtcgacgtcgtccgagtccga
L111P_antisense	gaggcagagaagcagctgcagcaggctcaggct |||||||||||||| |||||||||||||||| 3'-cgtctcttcgtcggcgtcgtccgagtc-5'
L130P	5'-aaaattcactaagcggccggtgaaggtcactaagc-3' ||||||||||||||||| ||||||||||||||||| ccttttaagtgattcgccgaccacttccagtgattcgtcg
L130P_antisense	ggaaaaattcactaagcggctggtgaaggtcactaagcagc ||||||||||||||||| ||||||||||||||||| 3'-ttttaagtgattcgccggccacttccagtgattcg-5'

Table 2.5: Primer sequences used in order to produce proline mutant plasmids.

PCR reactions contained 5 µl of 10 nM dNTP mix, 1 µl WT hFEN1 template (~100 ng/ul), 5 ul of 3 µM primer pair mix, 33 µl dd H_2O and 1 µl Hs Fusion polymerase and performed in a PCR sprint temperature cycling system (ThermoHybaid, Ashford) as follows,

Stage 1 – 2 min, 95°C → antibody removal

Stage 2 – 1 min, 95°C → melt plasmid DNA

30 sec, 55°C → anneal primers (table 2.5)

1min, 68°C → extension of new plasmids

Cycle 16x for amplification of mutant plasmid

Stage 3 – 1min, 72°C → finishes any gaps in the plasmids

Reaction products were analysed on a 1% agarose gel and visualised using EtBr for confirmation of successful synthesis of amplified mutant plasmid DNA.

Amplified DNA was then incubated with Dpn1 for at least one hour at 37^{o}C, and then, transformed into competent DH5α *E. coli* cells according to protocol. Briefly, after adding plasmid, the cells were incubated on ice for 45 minutes. The cells were heat-shocked at 42°C for 90 seconds, and then, allowed to cool in ice for 5 minutes. Cells were then resuspended in SOC media and incubated for 1 hour at 37°C with shaking for recovery. Cells were pelleted and the supernatant removed. Cell pellets were resuspended in SOC media, and then, were plated on LB (Luria Broth) (Kanamycin (Kan) (25 μg/ml) Chloramphenicol (Cm) (34 μg/ml)) and grown overnight at 37°C. One colony was taken for each mutant grown in 5 ml LB Kan, Cm overnight. Plasmids were isolated using a Qiagen Miniprep kit. Plasmids were sent to Sheffield Medical School sequencing using T7 promoter (forward) and T7 terminator (reverse) primers . Sequences were aligned to WT hFEN1 sequences using Serial Cloner 2.1 (Serial Basics, Freeware), and the best matches were transformed into Rosetta (DE3) BL21 (Kan^{R}; Chloramphenicol (Cm^{R})) *E. Coli* cells. Sequencing data can be found in the Appendices, figures A4-9. These cells were plated on LB (Kan (25 μg/ml), Cm (34 μg/ml)) and single colonies were used for overexpression and purification (section 2.1).

Multiple turnover measurements of hFEN1 proline mutants were performed as described in section 2.5 for hFEN1, for mutant L97P final concentrations of 25 nM E and 500 nM S were used; for mutants L111P and L130P final concentrations of 10 nM E and 500 nM S were used. Single turnover measurements were performed as in 2.7d but without the presence of streptavidin. Final concentrations of E and S were 900 nM and 2.5 nM, respectively.

Chapter 3: Using Streptavidin (SA) to investigate the FEN catalysed reaction

3.1 Introduction

There are several methods by which FEN enzymes are hypothesised to accommodate DNA substrates, as discussed in chapter 1. These can be summarised as follows:

The threading mechanism – This was first proposed by Dahlberg *et al*., as an explanation for FEN specificity for 5’ flaps. FEN enzymes were posited to thread the flap DNA portion of substrates (from the 5' terminus down to the bifurcation in the DNA substrate) through a hole created by protein secondary and tertiary structure (Lyamichev *et al.* 1999). A potential hole of appropriate size was later found to be formed by the helical arch.

The tracking mechanism – FEN enzymes were proposed to initially recognise the 5' termini of substrates by threading or clamping the ssDNA. This was suggested to be followed by sliding down the flap portion of substrates to the region of bifurcation, whereupon cleavage occurs (Murante *et al.* 1995; Barnes *et al.* 1996; Bornarth *et al.* 1999).

Bind then thread mechanism – FENs were proposed to initially bind the dsDNA portion of the bifurcated region of substrates, and then thread the ssDNA flap through the helical arch. Further refinement of the model then suggested that the helical arch undergoes a disorder to order transition around the ssDNA portion of substrates (Devos *et al.* 2007; Tsutakawa *et al.* 2011).

Bind then clamp mechanism – In this proposal FENs bind initially to the dsDNA portion of bifurcated region of substrates, and then clamp the ss 5’-flaps such that it departs on either side of the helical arch without passing through the helical arch (Orans *et al.* 2011).

The aim of this chapter is to investigate which of the proposed mechanisms is consistent with a series of experiments that measure the reactions of substrates that have been preassembled into different enzyme substrate (ES) complexes using biotinylated substrates to which streptavidin (SA) can be conjugated.

SA is a large stable protein that exists as a dimer of dimers (4 x 13 kDa). Each tetramer binds almost irreversibly (K_a ~10^{15} M^{-1}) in a 1:4 complex to biotin (Sano and Cantor 1990). Several factors account for the high affinity of the interaction between biotin and SA. These include the ordering of polypeptide loops on the surface of SA that bury the bound biotin within the protein, and several hydrogen bonds and strong van der Waals forces between the biotin and SA monomers (Weber *et al.* 1989). This extremely strong interaction can be used to differentiate between the threading and clamping mechanisms of FEN enzymes, because it is not possible for a SA bound flap to thread through an ordered or disordered helical arch (figure 3.1.1).

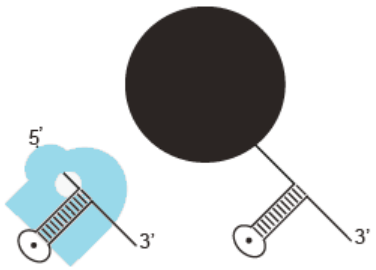


Figure 3.1.1 Illustration of the size of SA, which would block access to the 5' flap: Whereas the size of FEN is ~38000 Da, the SA complex is ~52000 Da. The aperture of the disordered helical arch of hFEN1 is hypothesised to be 42 Å x 42 Å (34 residues within the helical arch, approximately 3.4 Å per disordered residue (Dr. Jane Grasby, personal communication)), while previous work with atomic force microscopy has shown SA possesses a molecular volume of ~1050 Å and a size of approximately 30 Å x 30 Å (Neish *et al.* 2002).

The biotin-SA interaction was used similarly as a 'steric road block,' to determine the 3' to 5' directionality of recruitment of the human DNA repair enzyme *O_6 alkylguanine – DNA alkyltransferase* (AGT), an enzyme that irreversibly repairs methylated DNA by transferring alkyl legions to an active site cysteine residue as shown below (figure 3.1.2).

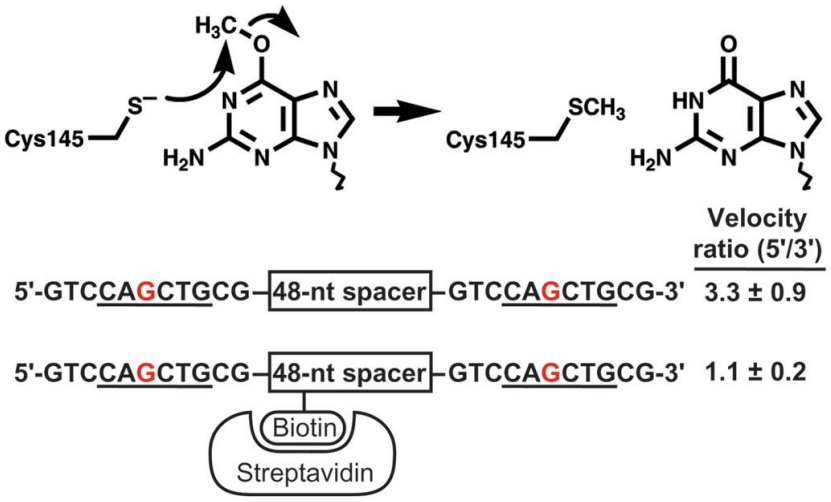


Figure 3.1.2, Top, the manner in which AGT is proposed to repair methylated DNA: Dealkylation of the damaged nucleobase occurs via attack of the methyl group by the sulphur atom of the cysteine residue of AGT. **Bottom, blocking the directionality of AGT:** Once SA is added to the substrate, directionality of AGT is abolished as shown by the relative speed at which each methylguanine (red) legion is repaired (Daniels *et al.* 2004).

Biotin was attached to a long PEG unit spacer that joined two oligonucleotides of identical sequence, each possessing methylguanine lesions equidistant from the PEG units. SA was then conjugated to the biotin. Prior to the incorporation of SA, methylated DNA on the 3' end of substrates was repaired three times quicker than the corresponding methylated DNA on the 5' end. SA removed this bias, suggesting that the recruitment of AGT proteins is prevented by SA, which in this case has sterically inhibited the interaction between enzyme and substrate. To explain the data, AGT binding had to occur from either side of the DNA strand when SA was present, consistent with the equal rate of repairing 3' and 5' methylated DNA (figure 3.1.2, (Daniels *et al.* 2004)).

Experimental design

Previous work has claimed to show that blocking the 5' end of flaps by attaching adducts such as SA and platinated branches (see chapter 1.6) are inhibitory. However, whilst some adducts and modifications are inhibitory others are not. For example, branched flap structures, gapped flaps and some CDDP adducts are cleaved (Murante *et al.* 1995; Barnes *et al.* 1996; Bornarth *et al.* 1999). In contrast to this previous work, which sometimes utilised reaction forcing conditions (e.g., very high concentrations of enzyme with prolonged incubation) and substrate constructs that were not always optimal for the enzyme in question, the reactions proposed in this chapter will use the most suitable substrate for the enzyme in question (i.e. double flap substrates for hFEN1), and through these biochemical assays will aim to rigorously test the mechanism by which FENs accommodate 5' ssDNA flaps and gapped flaps for cleavage.

Overview

To test how FEN accommodates the 5'-portion of flap substrates, biotinylated substrates (S) that could be conjugated to SA were designed, synthesized and purified. If threading of substrates through the helical arch does occur in FEN reactions, the presence of SA on the 5'-flap, as depicted in figure 3.1.1, should adversely affect the rates of reactions. Furthermore, if threading does occur, preassembled complexes of FEN and biotinylated substrate to which SA is subsequently added should form complexes that cannot dissociate when challenged with unlabelled competitor. For clarity, a consistent nomenclature is used throughout this chapter for the following experimental mixtures (figure 3.1.3):

Unmixed: Reaction is initiated by mixing of enzyme and substrate in buffer with subsequent quenching and analysis.

Premixed: Enzyme and substrate are preincubated in reaction conditions that do not support catalysis to form an ES complex; reaction is then initiated by the addition of Mg^{2+} ions and cleavage monitored over time.

Trapped: Enzyme and substrate are preincubated in reaction conditions that do not support catalysis to form an ES complex, which is then potentially "trapped" via conjugation of SA to the 5' flap. Reaction is then initiated by addition of Mg^{2+} and cleavage monitored.

Blocked: SA and substrate are preincubated to form SA-conjugated substrate. This is then preincubated with enzyme under reaction conditions that do not support catalysis, and then, reaction initiated with Mg^{2+}.

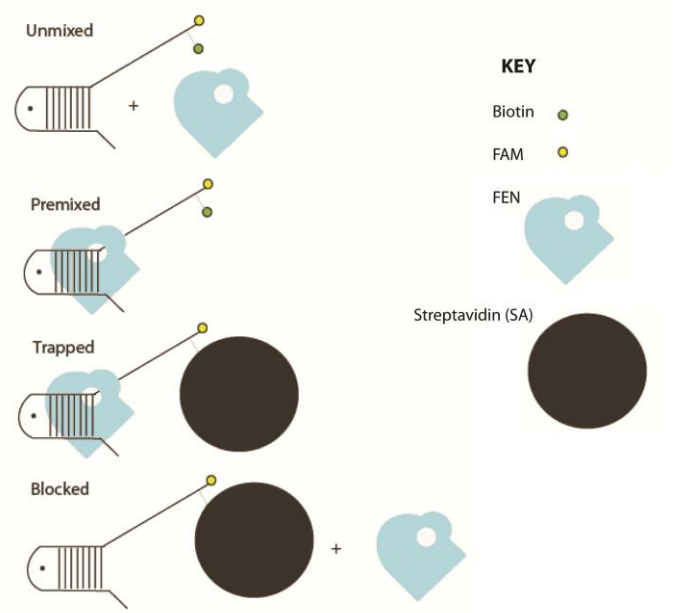


Figure 3.1.3: The four complexes prepared to investigate the accommodation of substrates within FEN. From top - unmixed substrate – enzyme and substrate are added separately and cleavage monitored; premixed substrate – enzyme and substrate are preincubated to form an ES complex and then reaction are initiated and cleavage monitored; trapped substrate – enzyme and substrate are preincubated to form an ES complex, trapped with SA and then reaction is initiated and cleavage monitored; blocked substrate – SA and substrate are preincubated to form SA conjugated substrate. This is preincubated with enzyme, and then reaction is initiated and cleavage monitored.

If hFEN1 accommodates the 5'-flap using a clamping mechanism, the rates of reaction in all four experimental set ups (figure 3.1.3) should be similar. This is because the presence of a large protein on the 5' end of long flap substrates should not obstruct binding of substrates in a clamped manner. In contrast, if threading or tracking were occurring to form ES complexes, the rates of reaction of unmixed, premixed and trapped ES complexes should be very similar to each other. As it is posited threading and tracking occurs via the 5' terminus

of substrates, where SA would be conjugated prior to enzyme binding, reaction should be prevented by blocked substrates.

3.2 Substrate design

Substrates for flap endonucleases can be created from one or more oligomers, designed to fold into a specific structure. The conditions of planned experiments demanded that substrates be stable at 37°C at a concentration of 5 nM in 100 mM monovalent salt to ensure reactions occurs with fully folded substrates. All substrates were designed using stability and structure prediction software, DINAMelt (Markham and Zuker 2005) (for hetero-duplex flap substrates) or MFold (Zucker 2003), to ensure that bimolecular and unimolecular structures were stable under the concentrations and conditions used.

T5 FEN substrates

The preferred substrates of bacteriophage FENs *in vitro* is a pseudo Y (pY) structure, as discussed in chapter 1. Thus, pY oligonucleotide substrates were prepared for the reactions as described in 3.1 for use with T5 FEN. Two 5'-biotinylated substrates were designed for T5 FEN catalysed reactions (figure 3.2.1). Unimolecular pY substrates possessed either a 21 nucleotide (nt) (pY-21B) or 7nt (pY-7B) 5' flap, with a 5' biotin moiety and a 5' FAM label. Non-biotinylated oligonucleotide equivalents were also synthesised for comparison (pY-21, pY-7). Furthermore, the expected products were also prepared with a 5'-biotin moiety and FAM label to verify that reactions occurred without change in reaction site specificity (Prod 21, Prod 7; table 2.3). As a control, a substrate possessing a 3' biotin moiety on the end of a long polyethylene glycol linker (PEG) (pY-21-3'B) was also created. This was made to ensure SA conjugation to oligo was not affecting the reaction catalysed by FEN due to fortuitous interactions between SA and FEN.

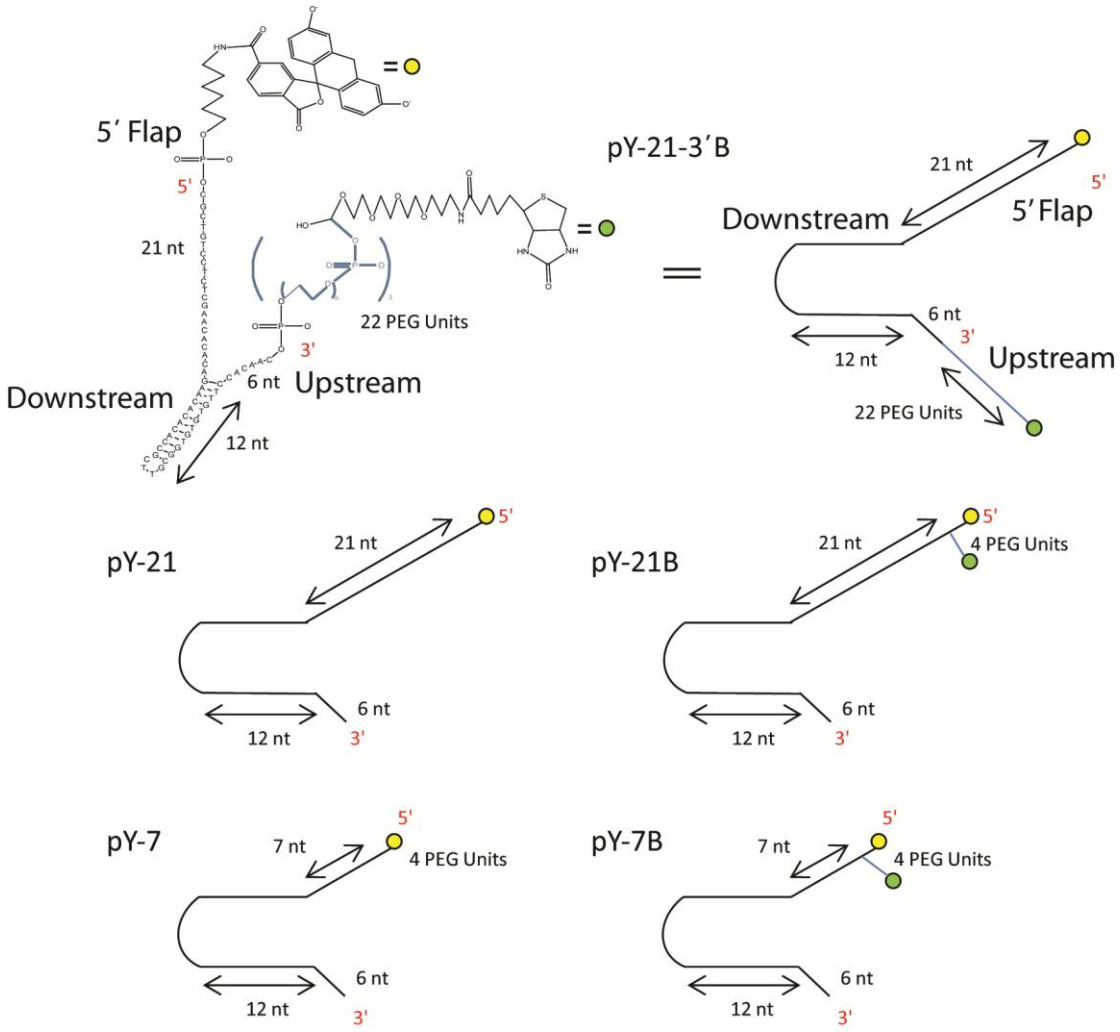


Figure 3.2.1: Structures of the oligonucleotides synthesised for this study, with their main regions defined: pY-21-3'B, a pseudo Y substrate with a 21 nucleotide 5' flap, possessing a 5' FAM and 3' biotin molecule, **pY-21**, a pseudo Y substrate with a 21 nucleotide 5' flap, possessing a 5' FAM, **pY-21B**, a pseudo Y substrate with a 21 nucleotide 5' flap, possessing a 5' FAM and biotin molecule, **pY-7**, a pseudo Y substrate with a 7 nucleotide 5' flap, possessing a 5' FAM and **pY-7B**, a pseudo Y substrate with a 7 nucleotide 5' flap, possessing a 5' FAM and biotin molecule.

As shown in figure 3.2.1, the region of the substrate 3' to the 5' flap is termed the downstream region, whereas the 3'-overhang portion of the substrate adjacent to this is termed the upstream region. All the pY substrates designed for work with T5 FEN possess a 12 nucleotide base paired downstream region capped by a 3 nucleotide hairpin, and a 6 nucleotide upstream region. The lengths of downstream and upstream regions were chosen with reference to the structure of T4 RNase HI (T4FEN) in complex with a pY DNA (Devos *et al.,* 2007) and models that could be created from this structure with T5 FEN.

Human FEN1 substrates

Two sets of optimal double flap (DF) substrates were designed for use with hFEN1. In each case, substrates were created from two oligomers, one which formed the flap strand and another that acted as template and provided the single nucleotide 3'-flap. Those with longer (21 nt) 5' flaps, had a flap strand with a 3' fluorescein label (DF-21), and a 5' biotinylated version was also created (DF-21B). A 3' fluorescein was used for DF-21 as separation of a 3' FAM labelled 11 mer product from a 33 mer starting substrate was much easier than the alternative 5' FAM labelled 22 mer product from a 33 mer starting substrate. A gapped flap substrate, as defined in chapter 1, and shown below (DFGEN, figure 3.2.2) was designed using the same template as the 21 nt 5' flap, with a 9 base paired hairpin within the 5' flap, to test the ability of FEN to thread gapped substrates through the helical arch. The gapped substrate was designed with a biotin on the apex of the 5' hairpin region, between the G and A bases. The addition of a biotin moiety within the duplex on the flap region of the substrate did not change the T_m of the hairpin as confirmed by thermal melting performed by Dr. John Atack (personal communication, (Patel *et al.* 2012)). A second set of substrates had shorter 5' flaps (3 nucleotides (DF-3) and 5 nucleotides (DF-5) long) in order to examine claims made by Gloor *et al.* that smaller flap substrates would not involve threading through the helical arch (Gloor *et al.* 2010). A pair of these substrates were designed, one with a 5' FAM, and another with a 5' biotin as well (DF-3B and DF-5B). This set of substrates possessed an 18 nt duplex in the downstream region as opposed to a 12 nt duplex (12 nt are required to fully occupy the downstream binding region in hFEN1, as seen by crystal structures (Sakurai *et al.* 2005; Tsutakawa *et al.* 2011)) in the longer 5' flap substrates, as this raised the stability of and consequently the predicted T_m of DF substrates from approximately 45^{o}C to approximately 60^{o}C under assay conditions (5 nM S, 100 mM monovalent salt). The template strand of all the DF substrates possess a 6 nucleotide upstream duplex region with a hairpin promoting tri nucleotide sequence to ensure the entire hFEN1 upstream DNA binding site was occupied (Sakurai *et al.* 2005; Tsutakawa *et al.* 2011). The structures of these substrates are shown below in figure 3.2.2.

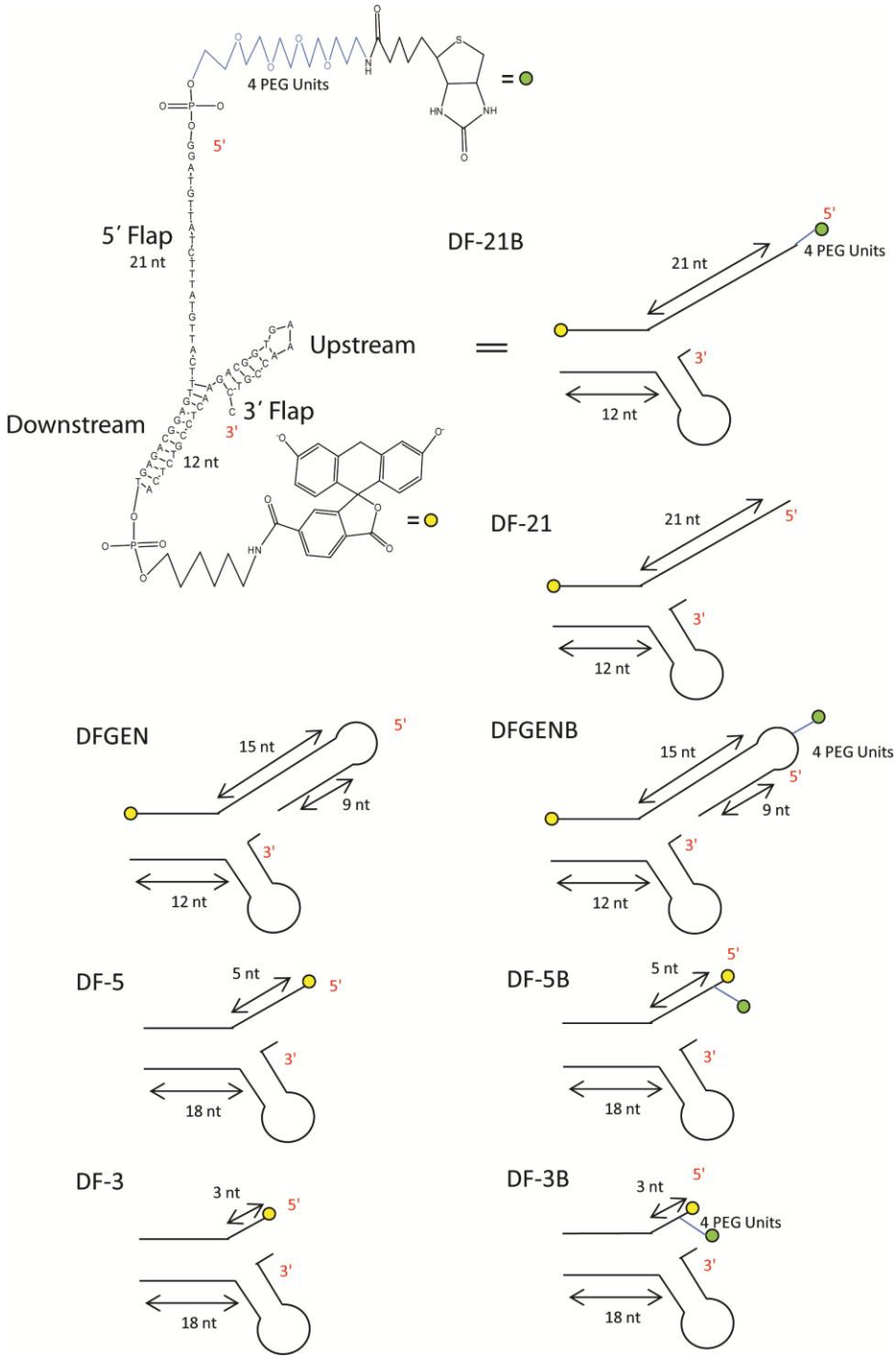


Figure 3.2.2: Structures of the double flap oligonucleotide substrates: DF-21B, a bimolecular substrate with a 21 nucleotide 5' flap and 1 nucleotide 3' flap, possessing a 3' FAM and 5' biotin, **DF-21**, a bimolecular substrate with a 21 nucleotide 5' flap and 1 nucleotide 3' flap, possessing a 3' FAM, **DFGEN**, a bimolecular substrate with a 21 nucleotide 5' flap, on which is a 9 base paired hairpin and a 1 nucleotide 3' flap, possessing a 3' FAM, **DFGENB**, a bimolecular substrate with a 21 nucleotide 5' flap, on which is a 9 base paired

hairpin and a 1 nucleotide 3' flap, possessing a 3' FAM and 5' biotin, **DF-5B**, a bimolecular substrate with a 5 nucleotide 5' flap and 1 nucleotide 3' flap, possessing a 5' FAM and biotin, **DF-5**, a bimolecular substrate with a 5 nucleotide 5' flap and 1 nucleotide 3' flap, possessing a 5' FAM, **DF-3B**, a bimolecular substrate with a 3 nucleotide 5' flap and 1 nucleotide 3' flap, possessing a 5' FAM and biotin, and lastly **DF-3**, a bimolecular substrate with a 3 nucleotide 5' flap and 1 nucleotide 3' flap, possessing a 5' FAM.

3.3 Determination of Michaelis-Menten parameters of biotinylated and non biotinylated substrates

To establish whether the addition of biotin altered the ability of substrates to bind and cleave T5 FEN or hFEN1, Michaelis-Menten kinetic parameters were measured for the substrates pY-7B, pY-21B, DF-21B, DF-5B, DF-3B and DFGENB that possessed the biotin modification and then compared to their counterparts lacking biotin (tables 3.1,2).

The high-throughput assays (figure 3.3.1) used herein utilise a reverse phase ion pairing dHPLC system equipped with a fluorescence detector to separate substrate and product DNA. The resulting peaks can be quantified by integration of the fluorescence chromatogram. Initial studies to determine if the biotin modification and subsequent SA conjugation affected dHPLC separations showed that the retention time of the biotinylated product was approximately a maximum of 30 seconds longer than its unmodified variant and good separation of substrate and products were observed as shown below in figure 3.3.2. Thus, these modifications did not significantly alter the ability to separate and quantify substrate and products formed, even when the flap and consequently, the monitored product were short.

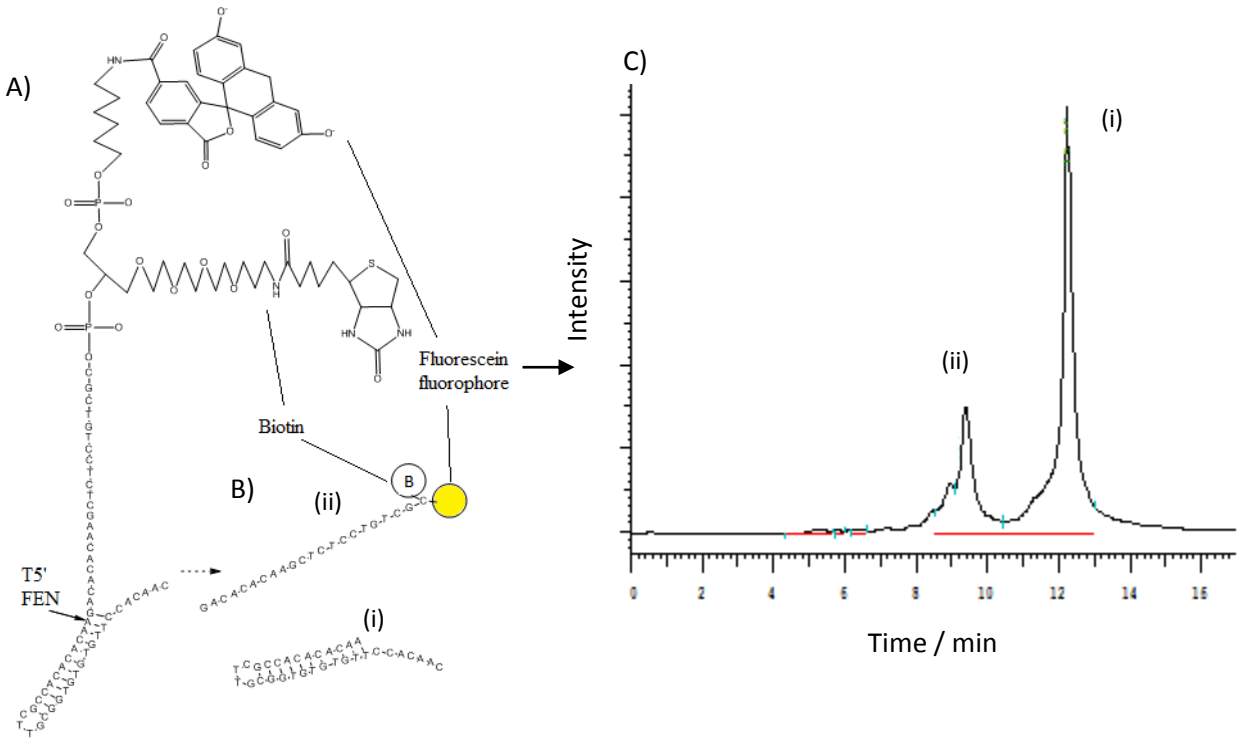


Figure 3.3.1: Fluorescent assay used to investigate T5 FEN-catalysed hydrolysis: A) Fluorescently labelled biotinylated synthetic flap substrate, pY-21B. B) Cleavage of substrate by T5 FEN, with the yellow ball indicating fluorescein. C) Separation of products and substrate by dHPLC, indicating the major product peak (ii) and the substrate peak (i).

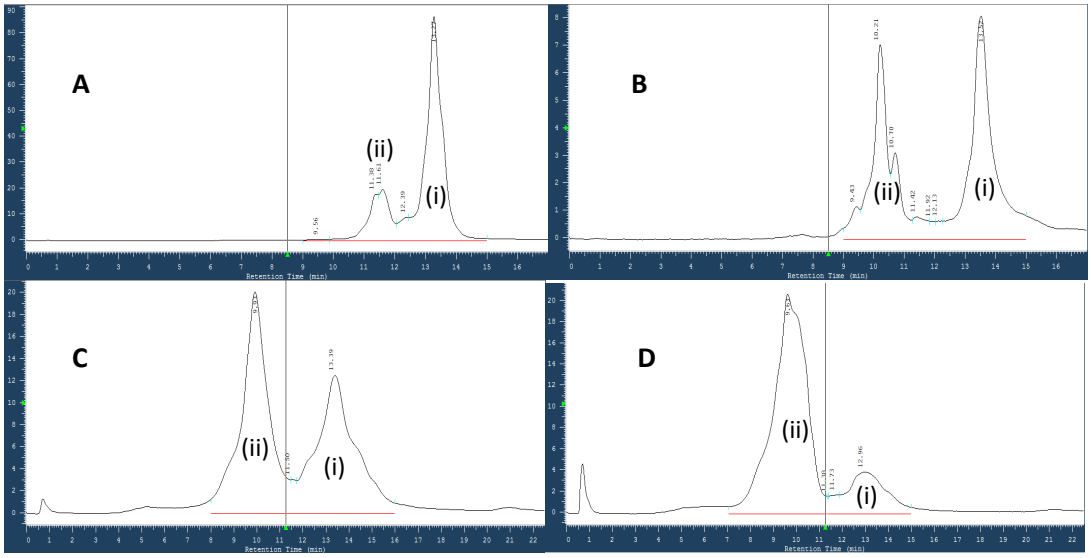


Figure 3.3.2: Sample traces of A: pY-7B and its resulting product, B: pY-7 and its resulting product, C: pY-21B and its resulting product and D: pY-21 and its resulting product. Starting materials (i), and resulting products (ii). All oligonucleotides were 100nM and T5 FEN enzyme 60 pM in starting concentration. Reactions were performed in 25 mM HEPES pH 7.5, 50 mM KCl, 10 mM Mg^{2+} and 0.01 mg/ml BSA at 37^{o}C and quenched in 200mM EDTA. The buffers and gradients used are described in section 2.5.

Fluorescence traces (figure 3.3.2) were integrated using the WAVE™ system software (Wavemaker, Hitachi) to calculate the percentage product at each time point, which could be converted to the amount of product with knowledge of the starting substrate concentration. As exemplified for the T5 FEN catalysed reaction of pY-7B (figure 3.3.3), initial rates of reaction were derived by linear regression of plots of concentration of product versus time. To obtain Michaelis-Menten parameters, initial rates of reaction were measured at varying substrate and enzyme concentrations, from which normalised initial rates of reaction were generated from the quotient of initial rate and enzyme concentration used (figure 3.3.4). The kinetic parameters k_{cat} and K_M were derived by non-linear regression using a Michaelis-Menten model (equation 1). Parameters were measured for T5 FEN and its pseudo Y substrates illustrated in figure 3.2.1. These measurements were plotted as shown in figure 3.3.4, with the corresponding constants reported in table 3.1.

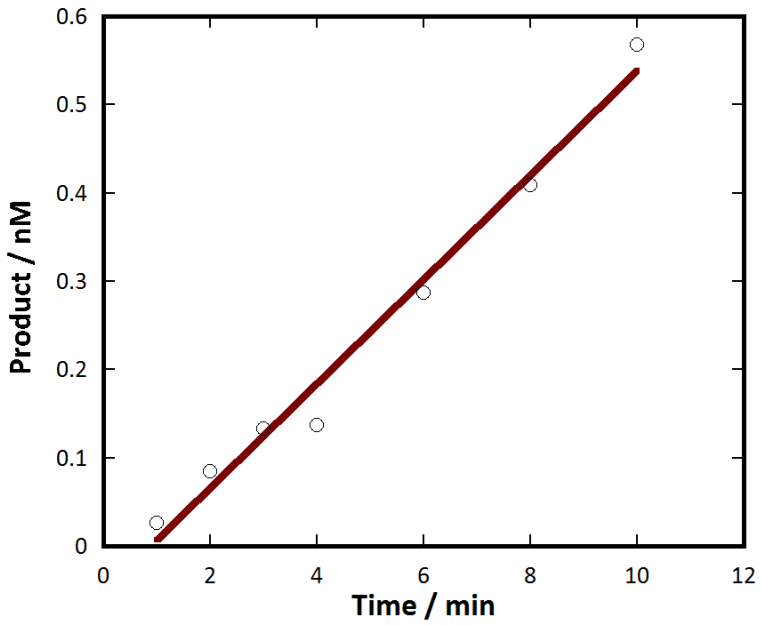


Figure 3.3.3: A sample graph of the initial rate of reaction of 5 nM pY-7B catalysed by 30 pM T5 FEN at 37˚C with product formed plotted against time. The slope of the line is the initial rate of reaction. In every Michaelis-Menten profile measured in this project, the initial rates were measured in this manner. The reaction was carried out in 25 mM pH 9.3 potassium glycine containing 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Mg^{2+}.

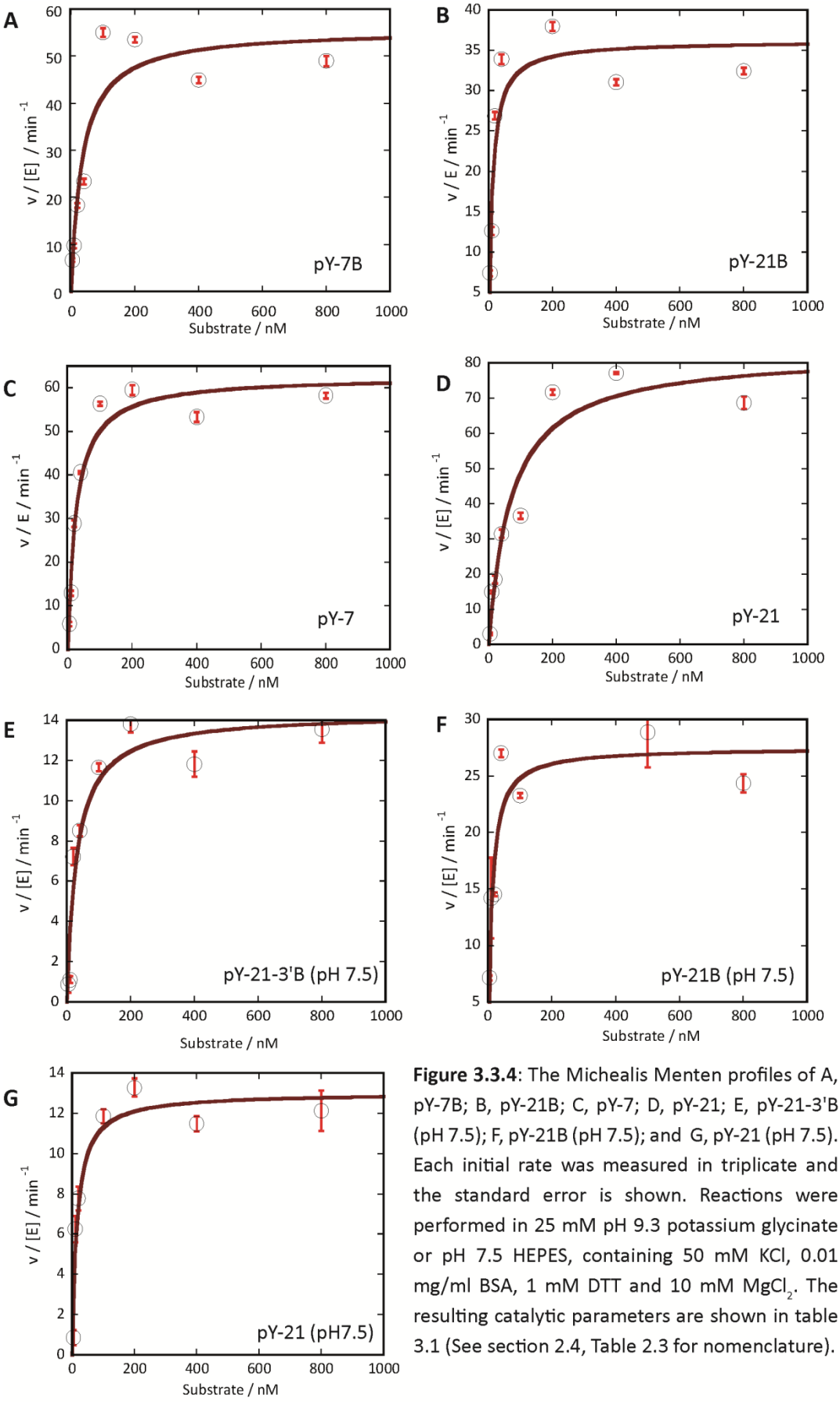


Figure 3.3.4: The Michealis Menten profiles of A, pY-7B; B, pY-21B; C, pY-7; D, pY-21; E, pY-21-3'B (pH 7.5); F, pY-21B (pH 7.5); and G, pY-21 (pH 7.5). Each initial rate was measured in triplicate and the standard error is shown. Reactions were performed in 25 mM pH 9.3 potassium glycinate or pH 7.5 HEPES, containing 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM $MgCl_2$. The resulting catalytic parameters are shown in table 3.1 (See section 2.4, Table 2.3 for nomenclature).

Substrate	k_{cat} (min^{-1})	K_M (nM)	k_{cat}/K_M ($nM^{-1}min^{-1}$)
pY-7B (pH 9.3)	55 ± 6	34 ± 14	1.6
pY-7 (pH 9.3)	38 ± 3	40 ± 13	0.95
pY-21B (pH 9.3)	62 ± 3	24 ± 6	2.6
pY-21 (pH 9.3)	80 ± 8	70 ± 23	1.14
pY-21B (pH 7.5)	27 ± 2	11 ± 4	2.45
pY-21 (pH 7.5)	11 ± 1	30 ± 9	0.36
pY-21-3'B (pH 7.5)	13 ± 1	15 ± 5	0.86

Table 3.1: Michaelis Menten parameters $k_{cat,}$ K_M and k_{cat}/K_M for the pseudo Y substrates calculated in this chapter. All measurements were made, as indicated in 25 mM pH 9.3 potassium glycine or pH 7.5 HEPES containing 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Mg^{2+}.

The biotinylated and non biotinylated substrates in table 3.1 give similar catalytic parameters, showing little or no effect on the addition of biotin to the enzyme-substrate interaction, or on subsequent reaction. The biotin – SA complex is most stable at pH 7.5, the pH at which all the experiments involving SA will be carried out at. Previous work (Sengerová 2009) has shown that a decrease in pH from the optimum ~8.5 leads to a decrease in the turnover number of reaction. In order to rule out any reduction in rate (in reactions with SA coated substrates) being directly due to the decrease of pH, the kinetic profiles of pY-21B, pY-21 and pY-21-3'B at pH 7.5 (figure 3.3.4 E, F, G) were measured for comparison purposes. There is a reduction in k_{cat} of pY-21B at pH 7.5, approximately 2 fold, 69 min^{-1} cf. 27 min^{-1}. With respect to pY-21, going from pH 9.3 to 7.5 lowers the K_M by 2 fold and the k_{cat} by 8 fold. The kinetic parameters of pY-21-3'B are similar to pY-21B and pY-21 at pH 7.5 (table 3.1). The same measurements were then made for hFEN1 and its double flap and gapped substrates at pH 7.5, as shown in figure 3.3.5 and the resulting parameters, k_{cat}, K_M and k_{cat}/K_M reported in table 3.2.

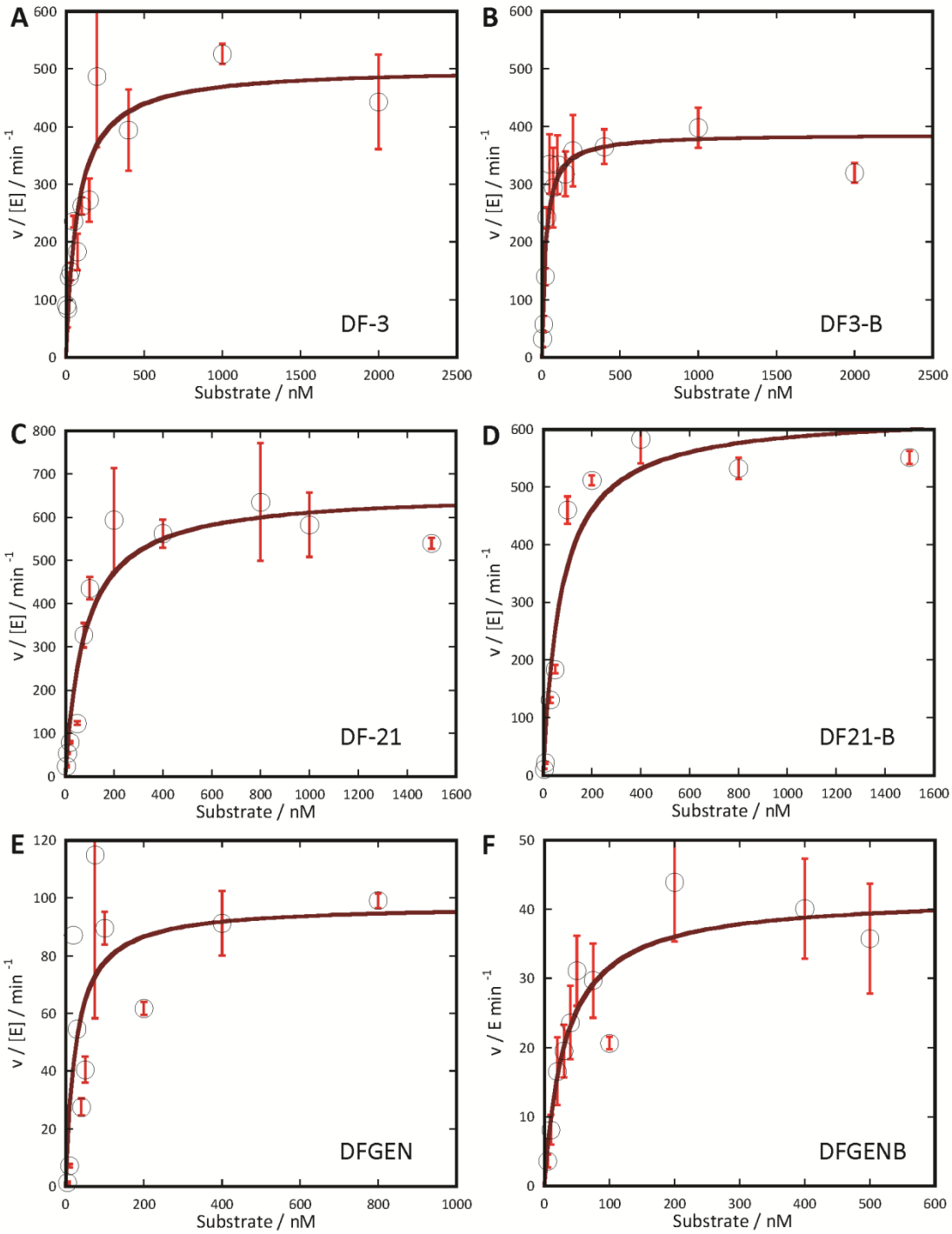


Figure 3.3.5: The Michealis Menten profiles of A, DF-3; B, DF-3B; C, DF-21; D, DF-21B; E, DFGEN; and F, DFGENB. Each initial rate was measured in triplicate and the standard error is shown. Reactions were performed in 50 mM pH 7.5 HEPES, containing 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 mM $MgCl_2$. The resulting catalytic parameters are shown in table 3.2 (See section 2.4, table 2.3 for nomenclature).

Substrate	k_{cat} / min^{-1}	K_M / nM	k_{cat}/K_M / $nM^{-1}min^{-1}$
DF-3	*488 ± 15*	*77 ± 1*	*6.34*
DF-3B	*387 ± 22*	*24 ± 6*	*16.1*
DF-5B	*451 ± 32*	*60 ± 15*	*7.52*
DF-21	659 ± 51	80 ± 24	8.24
DF-21B	629 ± 36	73 ± 36	8.62
DFGEN	98 ± 19	64 ± 19	1.53
DFGENB	42 ± 3	33 ± 11	1.27

Table 3.2: Michaelis Menten parameters, K_M, k_{cat} and k_{cat}/K_M for reactions of biotinylated and non-biotinylated double flap and gapped DF substrates catalysed by FEN1. All measurements were made in 50 mM pH 7.5 HEPES containing 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 mM Mg^{2+}. *Measurements performed by Dr. John Atack.*

Once again, as with the T5 FEN substrates, biotinylation of the hFEN1 substrates does not significantly alter kinetic parameters, with a maximum of approximately two fold difference between biotinylated and non biotinylated substrates. Such two-fold differences are commonly seen between different preps of the same oligo, and are thus, insignificant.

3.4 The effect of mutation K93A on the rate of hFEN1 reactions

As shown above (section 3.3), and reported earlier by Finger *et al*., the introduction of a stable duplex into the 5'-flap (substrates DFGEN and DFGENB) had a relatively modest effect on hFEN1 activity. However, earlier literature reports suggested that formation of duplex in flaps prevented FEN reaction, an observation that was initially cited as support for a tracking mechanism for flap endonucleases (Murante *et al.* 1995; Barnes *et al.* 1996). Thus, the ability of FENs to cleave so-called gapped flaps has been controversial. To verify that cleavage of gapped-flaps was due to FEN proteins and not the result of a contaminating co-purified nuclease, we performed reactions using the hFEN1 mutant K93A.

Lysine 93 is a positionally conserved helical arch residue in FEN family that is proposed to act as an electrostatic catalyst during FEN cleavage (Sengerova *et al.* 2010). This residue is also proposed to assist with capture of unpaired scissile phosphate (figure 3.4.1, (Sengerova *et al.* 2010; Tsutakawa *et al.* 2011)). Human FEN1 mutant K93A lacks this conserved lysine residue, which is replaced by an alanine residue.

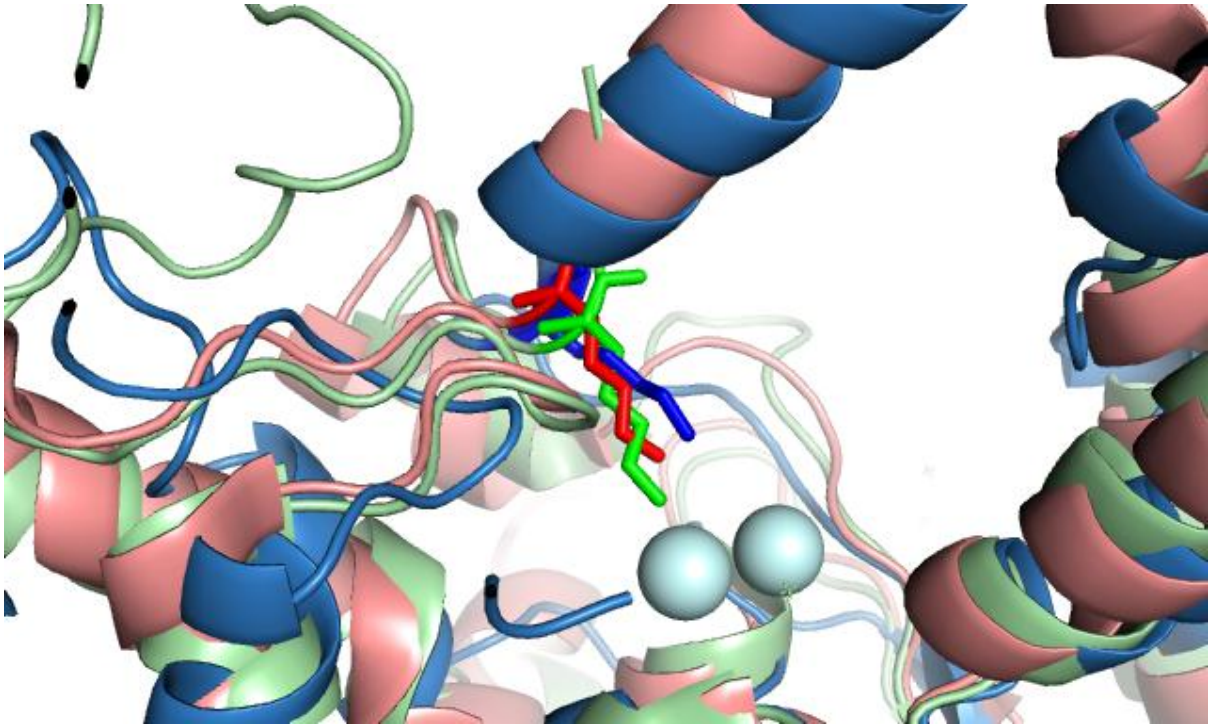


Figure 3.4.1: Overlaid structures of T5 FEN (pale red), T4 FEN (pale green) and hFEN1 (pale blue), showing the positionally conserved and correspondingly coloured K93 residue at the base of alpha helix 5 (T4 and T5 FEN numbering) or alpha helix 4 (FEN1 numbering) as sticks. The metal ions of hFEN1 are shown in cyan as a reference for the base of the helical arch in this figure.

The rates of single turnover reactions were determined with wild type hFEN1 and mutant K93A, both purified using the same rigorous purification procedure. A comparable decrease in activity when cleaving both double flap and gapped double flap substrates would be expected if all reactions were a result of FEN catalysis. As expected, the rate of reaction of both double flap and gapped double flap substrates was decreased by three orders of magnitude by the K93A mutation. Thus, one can conclude that the observed gap endonuclease activity is an activity of hFEN1 protein, and not the result of a co-purified protein, which would not be affected by the mutation. The cleavage of gapped

flap substrates with 5'-duplex also implies that tracking mechanisms requiring the binding of ss 5' termini of substrates is a specious hypothesis.

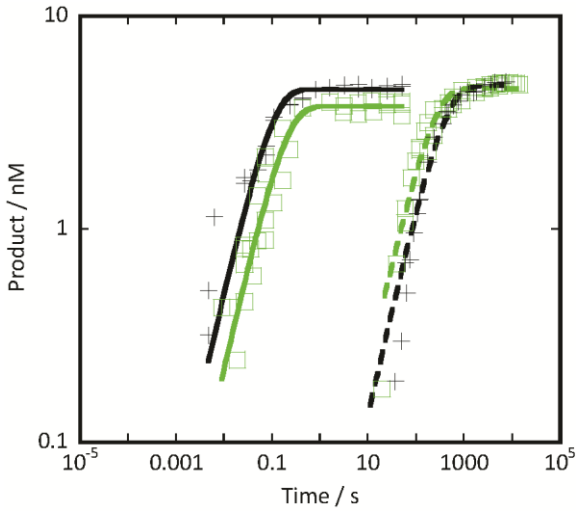


Figure 3.4.2 Plots of log [product] versus log time of the reactions catalysed by hFEN1 mutant K93A (hashed line) and WT hFEN1 (solid line) with the 21 nucleotide double flap substrate DF-21B (black crosses) or gapped substrate, DFGENB (green squares): Reactions were performed in 50 mM HEPES pH 7.5, 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 mM Mg^{2+}, and were initiated by mixing of 500 nM enzyme and 5 nM S. Data were fitted to equation 4, $P_t = P_\infty (1\text{-}\exp^{-kST.t})$ to give 462 ± 40 min^{-1} (WT hFEN1 - premixed DF-21B) 359 ± 54 min^{-1}, (WT FEN1 - premixed DFGENB) 0.3 ± 0.02 min^{-1} (K93A FEN1 – unmixed DF-21B), 0.17 ± 0.01 min^{-1} (K93A FEN1 – unmixed DFGENB).

3.5 Effect of conjugation of substrates to SA

To be certain that SA could be fully conjugated to the biotinylated, substrate, all biotin containing oligonucleotides were measured by fluorescence anisotropy as a function of SA concentration to show that the oligo could be saturated. Fluorophores have a set spectroscopic value termed r_0 (the spectroscopically determined maximal anisotropy (*r*) value for a given fluorophore), which for fluorescein is 0.35. Fluorescence anisotropy is a measure of the tumbling rate of a fluorophore in solution, that has been excited by plane polarized light of appropriate wavelength. Only fluorophores with absorption transition moments aligned along the electric vector of the incident light are preferentially excited (Lundblad *et al.* 1996). Anisotropy (*r*) describes the extent to which the resulting emission is polarised, quantified by measuring the difference between the amount of

emission signal detected in the plane parallel to the excitation plane and the amount of signal detected in the plane perpendicular to the excitation plane (Lakowicz 1999). The slower a molecule tumbles, the more the light remains polarised. The faster a molecule tumbles, the less the light will remain polarised. So, binding of an enzyme, or this case, SA to a substrate molecule will cause the tumbling rate of the FAM fluorophore to decrease, which will be reflected in the *r* value, becoming increasingly larger due to the detection of more polarised emission light (Jameson and Sawyer 1995; Lakowicz 1999; Tomlinson 2010).

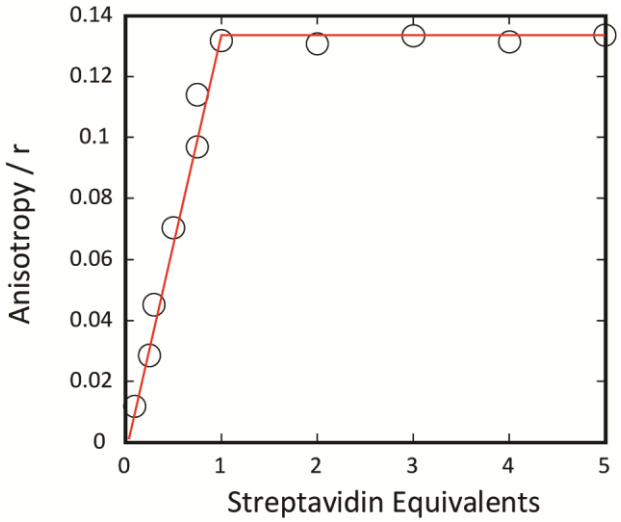


Figure 3.5.1: Plot of the anisotropy change on titration of SA into 100 nM pY-21B: This was performed at 25 mM HEPES pH 7.5 containing 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Mg^{2+} to maintain normal FEN reaction conditions. The excitation wavelength was set to 490nm, the emission at 510 nm recorded. Slit widths were set to 10 nm, with an average of 10 scans taken per reading (fit manually drawn for clarity). An equivalent of SA is the manufacturer's unit.

The resulting plot of SA against *r* shows that there is almost stoichiometric binding occurring up to 1 'equivalent' (as manufacturer's units) of SA to biotinylated substrate (Figure 3.5.1), upon which full saturation of the oligonucleotide substrate occurs. One equivalent of SA is the amount of SA needed to bind all of the biotin moieties present within solution.

Anisotropy could not be used to monitor the binding of oligonucleotides pY-21-3'B, DF-21B and DFGEN to SA because the 3' FAM or 5' FAM moieties were separated from the 3' or 5' biotin

molecules by a PEG linker, and 33-54 nucleotides, respectively. As a result of the increased mobility afforded by this, little change in anisotropy was observed on addition of up to 10 equivalents of SA to these substrates. Instead, a non-denaturing PAGE gel was used to visualise the binding of SA to pY-21-3'B via electrophoretic mobility changes (figure 3.5.2).

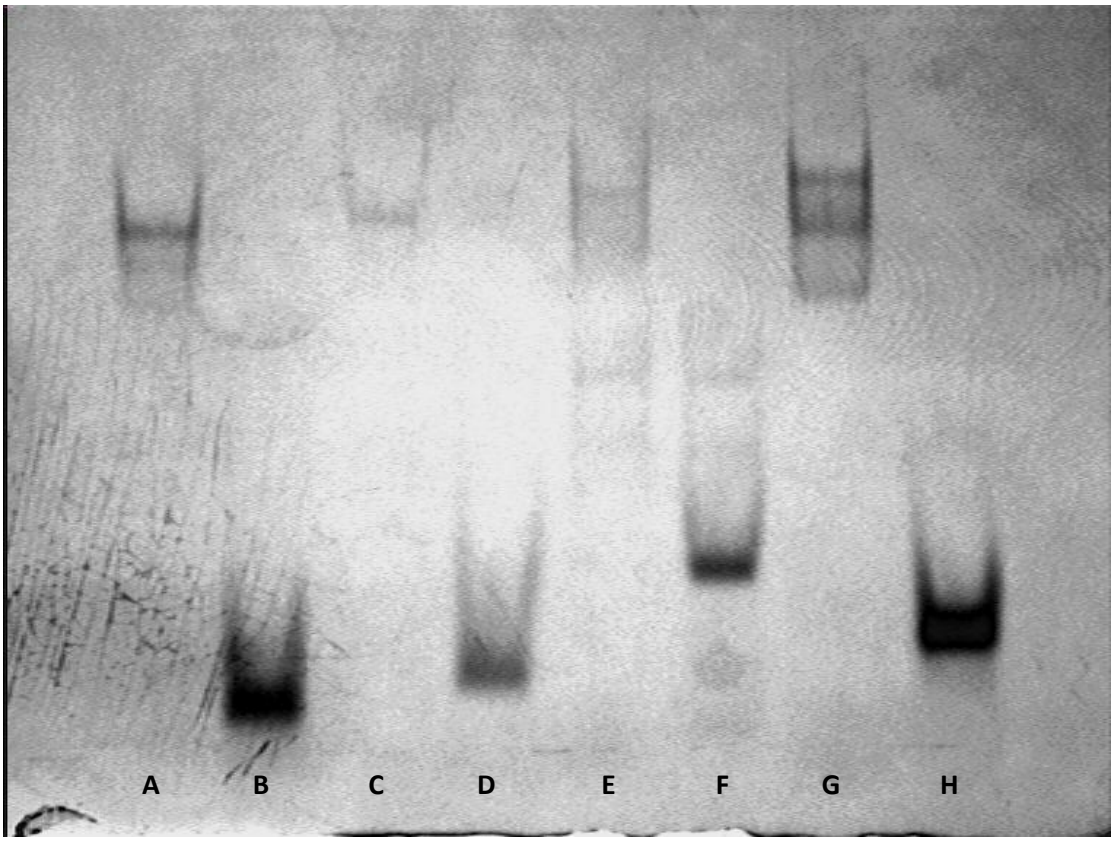


Figure 3.5.2: PAGE gel of A) 200 pM of pY-21-3'B and 1 equivalent of SA, B) 200 pM of pY-21-3'B, C) 200 pM of pY-21 and 1 equivalent of SA, D) 200 pM of pY-21, E) 200 pM of DF-21B and 1 equivalent of SA, F) 200 pM of DF-21B, G) 200 pM of DFGEN and 1 equivalent of SA and H) 200 pM of DFGEN: The gel was run at a constant 50 V for approximately 1 hour. The DNA was visualised by staining for 30 minutes with 5% acetic acid containing toluidine blue, and destaining for an hour using 5% acetic acid.

The results from gel mobility assays were similar to earlier fluorescence polarisation studies and showed that "1 equivalent" of SA was enough to conjugate all the substrate. In the lanes with 1 equivalent of SA, there was evidence of more than one high MW S-SA band. This is presumably due

to the maximum possible 4:1 binding ratio of biotin to SA, leading to more than one biotinylated substrate binding to one SA tetramer. Using atomic force microscopy, Neish *et al.* visualised the binding of biotinylated duplex DNAs to SA and showed under representation of higher occupancy (3:1 and 4:1) states presumably due to steric effects. Sigma Aldrich do not state the amount of moles of SA present in "one equivalent" but instead define the weight of biotin it will conjugate, but it is clearly enough to conjugate all the biotinylated substrate. To be absolutely certain of complete saturation of oligonucleotide with SA, and to disfavour the formation of higher occupancy states, a minimum of 5 equivalents of SA were used in all FEN reactions containing biotinylated oligonucleotides.

Analysing SA Conjugated Oligonucleotides

To test the threading mechanism, T5 FEN and hFEN1 catalysed reactions of SA conjugated oligonucleotides were conducted. The products of these reactions had to be analysed using dHPLC (see section 3.3). An initial reluctance to apply intact SA to dHPLC DNAsep columns led to a series of tests designed to either remove SA from biotinylated oligonucleotides, or denature the SA protein in solution to facilitate analysis via dHPLC. Different solutions were used, 1.5 M HCl, 1 M NaOH and 7 M urea to denature the proteins all with 60 mM EDTA to prevent further reaction. Attempted displacement of SA from modified oligonucleotides was also tested by boiling in water in the presence of excess biotin.

Boiling in 1 M NaOH for 1 minute only resulted in the degradation of the FAM label on the oligonucleotide even without conjugation to SA (Figure 3.5.3).

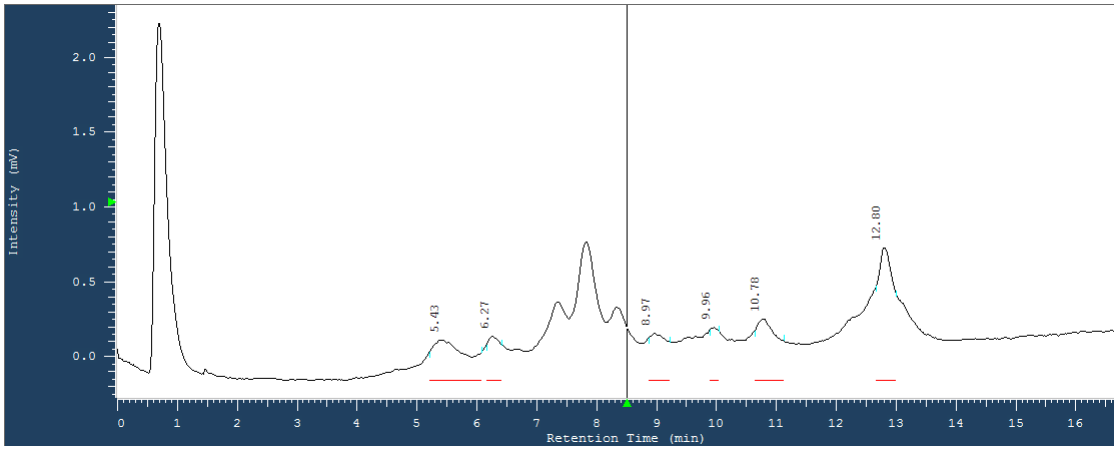


Figure 3.5.3: 100 nM pY-21B boiled in 1 M NaOH, run on a DNA-Sep column (Transgenomic).

Boiling the SA conjugated oligonucleotide in the presence of excess biotin dissolved in DMSO resulted in a large loss of FAM labelled material. This could be due to the conjugated material binding to the excess biotin, and precipitating. When this material was removed and the remaining solution analysed by dHPLC, the intensity of the signal representing free substrate was greatly reduced. Adding solutions of SA conjugated oligonucleotides in 1 M NaOH formed a precipitate, partly comprised of the SA conjugated material as shown once again by the loss of intensity in the signal representing free substrate in solution. This test was repeated with solutions of 1.5 M HCl. This, however seems to quench the fluorescence of the FAM label, making quantifying substrates and products almost impossible and unreliable. Adding solutions of SA conjugated oligonucleotides to denaturing solutions of 8 M urea and 80 mM EDTA gave the most encouraging results (figure 3.5.4), as shown below. SA conjugated oligonucleotides (pY-21B and Prod 21 product standard) to which the urea and EDTA were added eluted at similar times to samples to which SA had not been added. It is not known whether addition of urea denatures SA or not; nevertheless, reproducible traces where obtained and the lifetime of dHPLC columns were not significantly decreased. In later experiments with added FEN enzymes where protein was present in excess relative to substrate, it was also noted that urea was required to prevent FEN-substrates complexes persisting during analysis that eluted early in dHPLC traces.

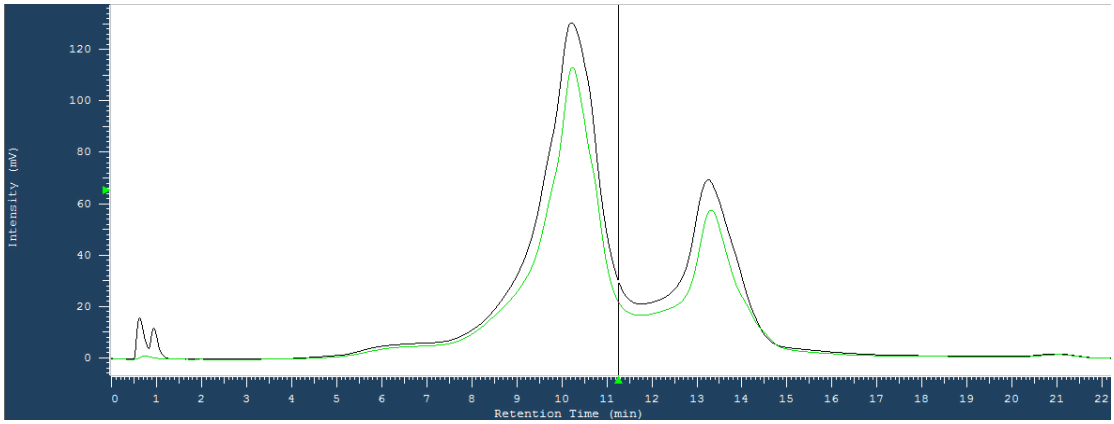


Figure 3.5.4: 100 nM standards pYB-21 and Prod 21 in 250 mM EDTA (black), and pYB-21 and Prod 21 standards conjugated to SA in 8 M urea and 80 mM EDTA (green), run on a DNA-Sep column (Transgenomic).

Using this urea and EDTA solution as a reaction quench, solutions of SA conjugated oligonucleotides were separable on a DNAsep column without loss of signal or major broadening of peaks. Thus, the dHPLC assay was valid for SA conjugated material used in the experiments described below. The effectiveness of this solution as a quench for reactions was tested in experiments that did not contain SA. Results were compared with experiments quenched with the conventional 1.5 M NaOH and 60 mM EDTA mix, which has been widely used as an instant quench in single turnover reactions with FEN enzymes. As shown later when comparing the 'premixed' single turnover rate constants (table 3.4) for the biotinylated and unbiotinylated substrates DF-21B and DF-21, respectively, using the NaOH quench for DF-21B and urea quench for DF-21 resulted in negligible differences between rates of reaction (723 min^{-1} cf. 683 min^{-1} respectively), implying that both are equally effective at stopping reactions.

3.6 Exploring the reaction of SA conjugated substrates using single turnover kinetics

As our experiments planned to potentially form trapped complexes, to which we would add magnesium ions to monitor reaction, the most appropriate comparison would be to use single turnover experiments to obtain rate constant k_{ST} for every substrate with the appropriate FEN enzyme. To determine the maximal rate of single turnover reactions, [E] must be much higher than the dissociation constant of the enzyme-substrate complex, to ensure that substrate will be fully saturated with enzyme. The rate constant k_{ST} can encompass the events involved in creating E-S

complex (dependant on [E]), but when conducted under conditions where rates of reaction do not increase with increasing [E], the rate constant k_{ST} can measure steps after initial ES formation (figure 3.6.1).

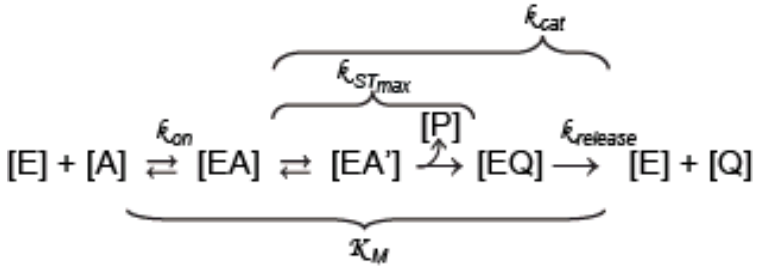


Figure 3.6.1: Proposed reaction scheme for FEN enzymes (courtesy of Dr D. Finger), showing the aspects of the reaction each catalytic parameter is thought to be responsible for. E= FEN enzyme; A= starting material; P= ssDNA product; Q =dsDNA product.

To test the FEN catalysed reactions of SA conjugated substrates, several rapid quench reactions were performed as shown below (figure 3.6.2). Reactions were performed on 'unmixed' enzyme substrate complexes, where enzyme and substrate were mixed from separate syringes using an RQF-63 apparatus. Reactions were carried out on 'premixed' enzyme substrate complexes where enzyme and substrate were preincubated in the absence of divalent magnesium ions and mixed with magnesium ions to initiate the reaction. Reactions were also performed on 'trapped' enzyme substrate (ES) complexes where enzyme and substrate were preincubated in the absence of divalent magnesium ions, SA was added to trap the ES complex, and then mixed with magnesium ions to initiate the reaction. Finally, reactions were performed on 'blocked' enzyme substrate complexes, where substrate was incubated in the presence of 5 equivalents of SA and this was then added to enzyme. After a period of preincubation, divalent magnesium ions were added in order to initiate the reaction.

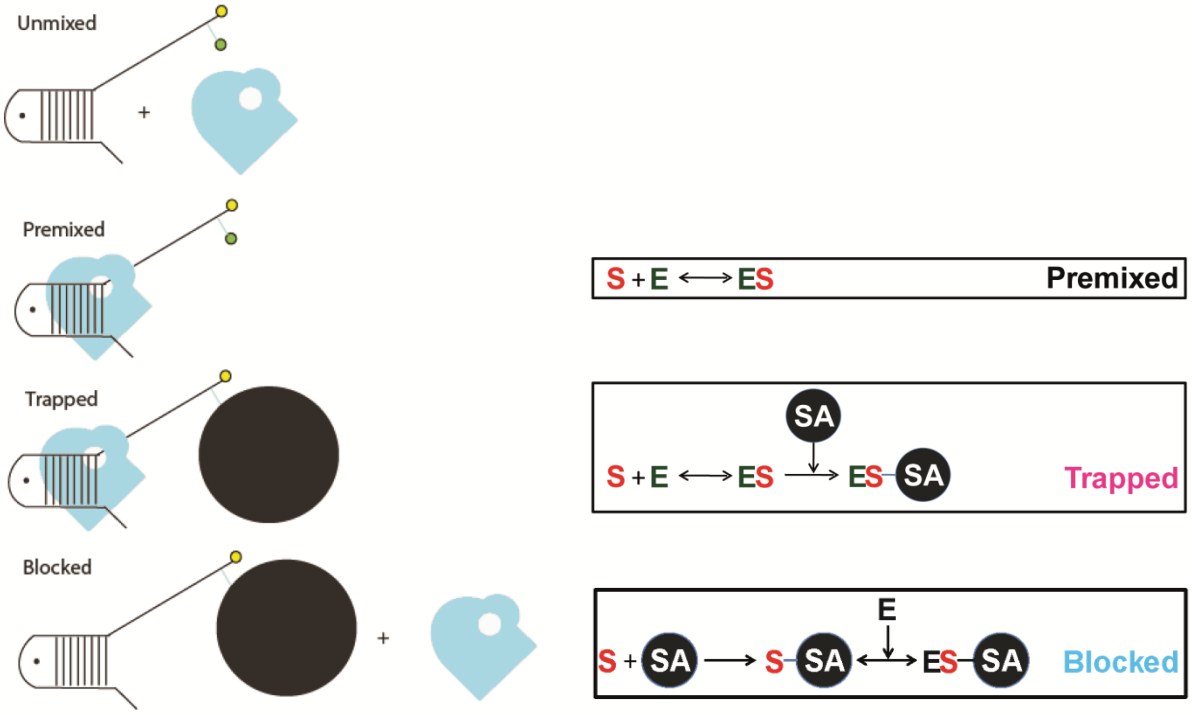


Figure 3.6.2: Diagrams showing the procedures involved in forming 'premixed,' 'trapped' and 'blocked' enzyme-substrate complexes. 5 nM Substrate and 500 nM enzyme, in the case of T5 FEN, were added on ice and incubated for 2 minutes. For hFEN1, E and S preincubation was carried out at room temperature. SA was added on ice, and the mixture incubated at room temperature for one minute. Pre-incubation was carried out in the presence of EDTA, or the catalytically inert Ca^{2+} ions.

Initially slow reaction was observed when enzyme and substrate were mixed without addition of divalent ions. Thus 'trapped' and 'premixed' ES complexes were assembled in the presence of EDTA or the catalytically inert Ca^{2+} ions, due to traces of contaminating divalent metal ions that support cleavage by FEN enzymes. These divalent metal ion contaminants were observed even after treatment of the substrates and buffers with Chelex resins, although reactions observed without addition of any cofactors were considerably reduced after this treatment. Residual ions most likely arise from the enzyme preparations; attempts to treat FEN proteins with Chelex considerably reduced their concentration presumably due to binding to the resin. To make sure any reaction that occurred during pre-incubation was not misinterpreted during analyses, the amounts of any product formed prior to addition of magnesium ions was always analysed (and was plotted at time = 0).

With the conditions for assembling ES complexes successfully determined, and the method of quenching and analysing the results of cleaving SA conjugated substrates ascertained, the ST reactions catalysed by T5 FEN were monitored (figure 3.6.3).

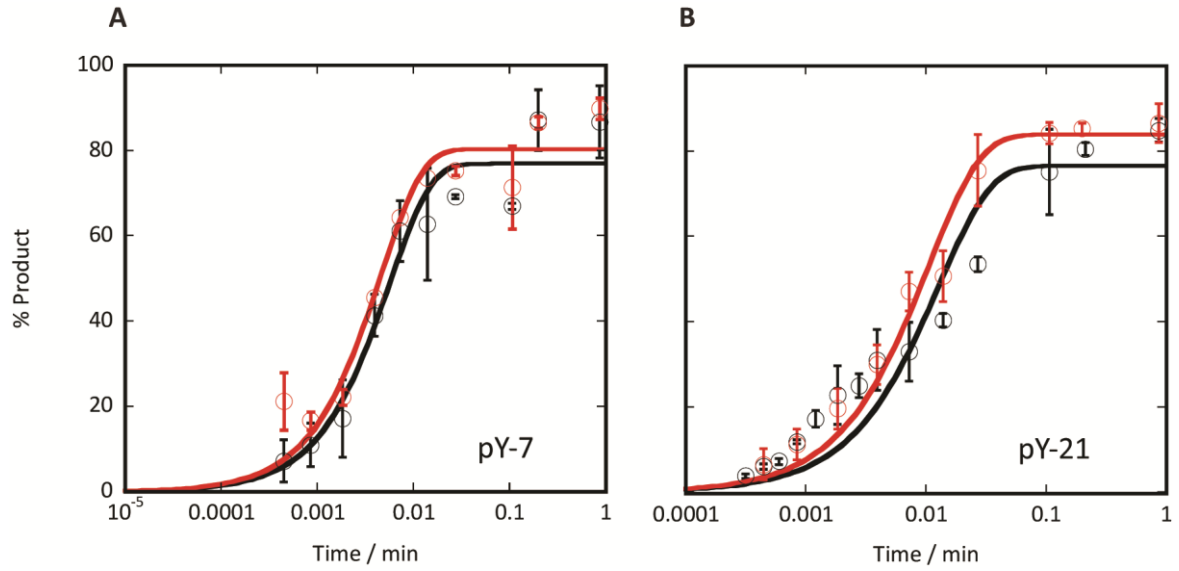


Figure 3.6.3: The unmixed (black) and premixed (red) single turnover profiles of pY-7 (A) and pY-21, B. Each plot with error bars shown were performed in triplicate and the standard error calculated. Reactions were performed in 25 mM HEPES pH 7.5, 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Mg^{2+}. Pre-mixed complexes were pre-incubated in the presence of 1 mM EDTA. Data were fitted to equation 4, $P_t = P_\infty (1\text{-}\exp^{-kST.t})$ to give catalytic parameters shown in table 3.3.

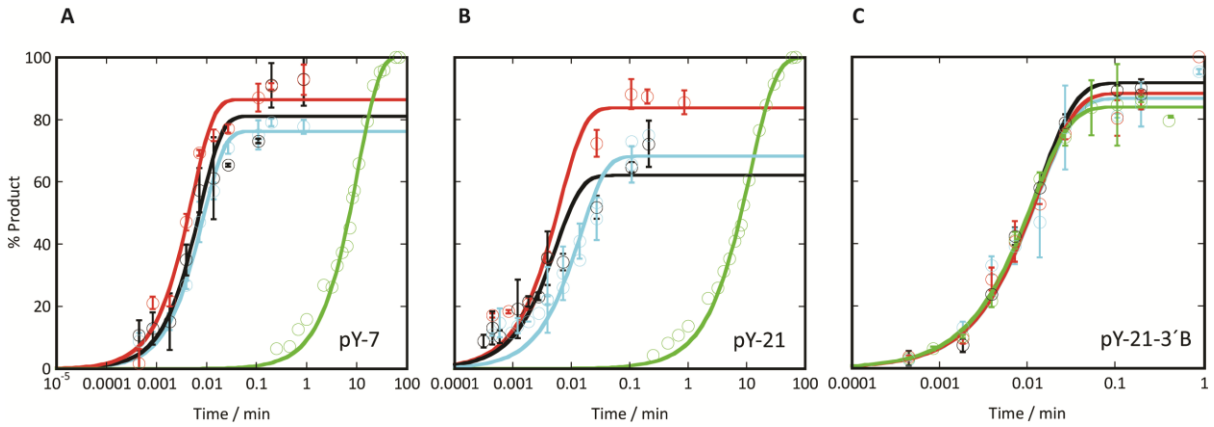


Figure 3.6.4: The unmixed (black), premixed (red), trapped (blue) and blocked (green) single turnover profiles of pY-7B (A), pY-21B (B) and pY21-3'B (C). Each plot with error bars shown were performed in triplicate and the standard error calculated. Reactions were performed in 25 mM HEPES pH 7.5, 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Mg^{2+}. Trapped and pre-mixed complexes were pre-incubated in the presence of 1 mM EDTA. Data were fitted to equation 4 to give catalytic parameters shown in table 3.3.

Comparison of the rate of T5 FEN catalysed reaction of biotinylated and non-biotinylated substrates under single turnover conditions shows that there is no discernible difference in rate between the unmodified and modified substrates (figures 3.6.3 and 3.6.4). Furthermore, there is no difference in the rate of reaction of premixed and unmixed ES complexes. This suggests that the pre-incubation of enzyme and substrate in the absence of divalent metal ions does not contribute a great deal to the reaction; the association of enzyme and substrate, the manner in which the ssDNA flap is accommodated within the active site, and as shown by Finger *et al.,* the chemistry of the reaction are extremely efficient, and some other step in the T5 FEN reaction is rate limiting. A small decrease in rate of reaction occurs upon formation of a trapped ES complex. Additionally, reaction of a blocked 3' biotinylated substrate where SA has been added before mixing with T5 FEN protein occurs at a similar rate to the other ES complexes. However, when SA was used to create 5' blocked substrates (figure 3.6.4), a large drop in rate was observed, reduced by approximately 4 orders of magnitude (table 3.3).

Substrate	Unmixed		Premixed		Trapped		Blocked	
	k (min^{-1})	$t_{1/2}$ (s)	k (min^{-1})	$t_{1/2}$ (s)	k (min^{-1})	$t_{1/2}$ (s)	k (min^{-1})	$t_{1/2}$ (s)
pY-7	178 ± 32	0.2	260 ± 38	0.2	n/a	n/a	n/a	n/a
pY-7B	133 ± 27	0.3	194 ± 26	0.2	117 ± 10	0.4	0.02 ± 0.0003	2100
pY-21	95 ± 25	0.4	110 ± 27	0.4	n/a	n/a	n/a	n/a
pY-21B	167 ± 70	0.2	158 ± 40	0.3	69 ± 10	0.6	0.05 ± 0.01	835
pY-21-3'B	75 ± 6	0.5	76 ± 9	0.5	77 ± 11	0.5	95 ± 3	0.4

Table 3.3: k_{ST} and $t_{1/2}$ for the pseudo Y T5 FEN substrates calculated in this chapter. All measurements were conducted in 25 mM pH 7.5 HEPES, mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Mg^{2+}. Premixed and trapped reactions have 1 mM EDTA present to prevent reaction prior to initiation with Mg^{2+}. $t_{1/2}$ was calculated using the equation, $t_{1/2} = \ln(2) / k_{ST.}$

The lack of difference between the trapped, unmixed and blocked single turnover rates of reaction concerning the 3' biotinylated substrate, pY-21-3'B (table 3.3) suggest that these reactions are all occurring via the same mechanism. Despite differences in the manner in which enzyme and substrate are treated prior to reaction, only a maximum of 2 fold difference in k_{ST} occurs. This all indicates that the large decrease in rate of reaction of 5'-blocked substrates is due to a different rate determining step or reaction mechanism. This must be due to the fact that the 5' flap portion of the substrate cannot bind to T5 FEN in its most productive conformation with SA bound to the 5' terminus of substrates, while the uninhibited dsDNA portions most likely bind optimally. The cleavage of 5' blocked substrates will be discussed in detail in section 3.7.

As a follow on to initial attempts to trap ES complexes resulting in some cases in a low end point, an investigation was carried out in order to elucidate the underlying reasons. Reactions were set up as described in figure 3.6.2, in order to test whether trapped ES complexes and premixed ES complexes were being made successfully prior to the addition of divalent magnesium ions. T5 FEN ES complexes and hFEN1 ES complexes were set up in premixed and trapped conformations in the presence of 2 mM Ca^{2+} and 1 mM EDTA. The reactions were then initiated with magnesium ions, and the formation of product was monitored over the course of 1 minute (figure 3.6.5).

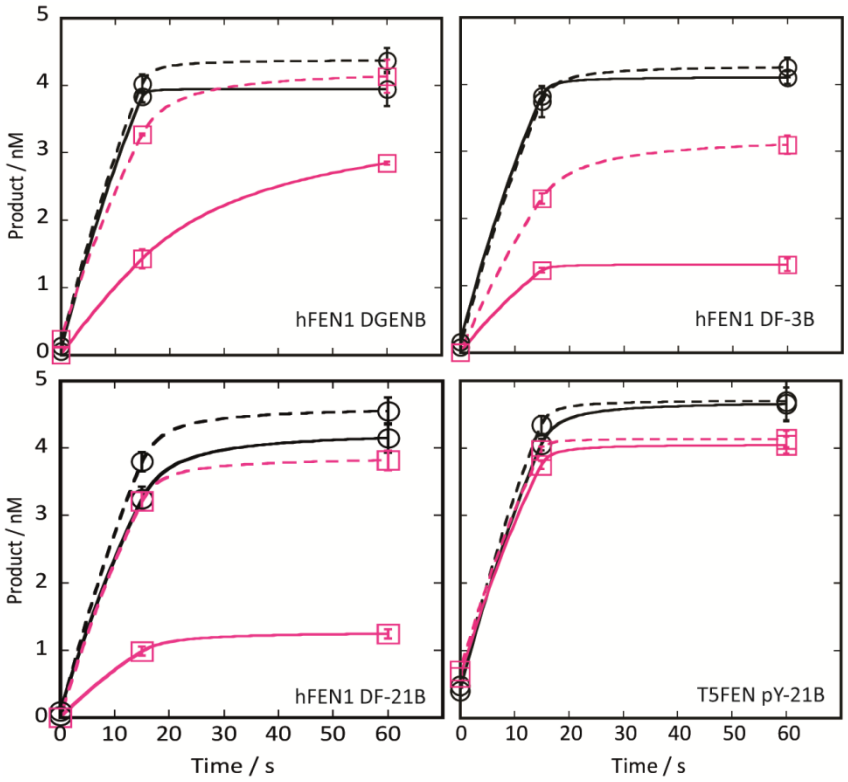


Figure 3.6.5: The levels of FAM product formed when E + S are 'premixed' (hashed line) with T5 FEN or hFEN1 and when E + S are 'trapped' (solid line) with T5 FEN or hFEN1 in the presence of 2 mM Ca^{2+} (black) or 1 mM EDTA (pink) at room temperature: All reactions were performed in 25 or 50 mM HEPES pH7.5 containing 50 or 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 or 10 mM Mg^{2+} (T5 FEN/hFEN1 conditions). The substrates used in these reactions are indicated within the figures. Substrates and enzyme were 'trapped' and 'premixed' as described in figure 3.6.2, [E] and [S] were at a final concentration of 500 nM and 5 nM respectively.

T5 FEN ES complexes reacted as expected when assembled in the presence of EDTA and Ca ions. However, trapped and premixed ES complexes, in the case of hFEN1, only reached the normal endpoint of reaction when preassembled in buffers containing 2 mM Ca^{2+}, instead of 1 mM EDTA. This is shown by the decay of 'trapped' hFEN1 ES complexes possessing a disproportionately low amount of product (20-60% cleavage) when preincubated in the presence of EDTA, however when compared to the analogous reaction after preincubation in divalent Ca ions, the amount of product in the analogous reactions is a lot closer to 80-95% cleavage, as is the case in the 'trapped' reactions with T5 FEN in the presence of EDTA (figure 3.6.5), showing successful trapping of the substrates within the tertiary structure of these enzymes.

The single turnover rates were then measured for hFEN1 and its double flap substrates (figures 3.6.6-7; table 3.4), in the same manner as with T5 FEN, the only difference being the environment in which ES complexes were preassembled as shown in figure 3.6.5. As seen previously; unmixed, premixed and trapped hFEN1-substrate complexes decay at a comparable single turnover rate. In sharp contrast, the 5' blocked ES complex reacts 3 and 4 orders of magnitude slower in the case of the 21 nt 5' flap, gapped flap and 3 nt 5' flap substrate, respectively (table 3.4).

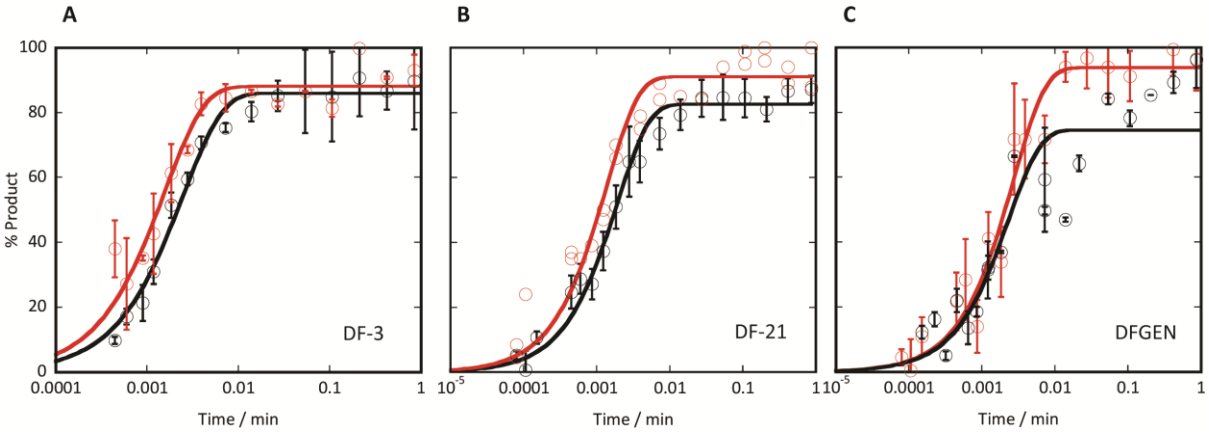


Figure 3.6.6: The unmixed (black) and premixed (red) single turnover profiles of DF-3 (A), DF-21 (B) and DFGEN (C). Each plot with error bars shown were performed in triplicate and the standard error calculated. Reactions were performed in 50 mM HEPES pH 7.5, 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 mM Mg^{2+}. Premixed complexes were preincubated in the presence of 2 mM Ca^{2+}. Data were fitted to equation 4 to give catalytic parameters shown in table 3.5.

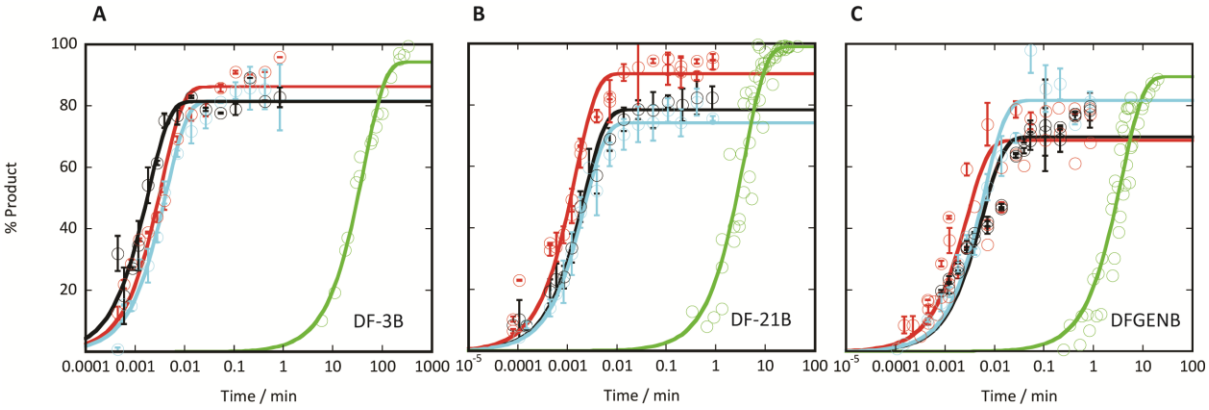


Figure 3.6.7: The unmixed (black), premixed (red), trapped (blue) and blocked (green) single turnover profiles of DF-3B (A), DF-21B (B) and DFGENB (C). Each plot with error bars shown were performed in triplicate and the standard error calculated. Reactions were performed in 50 mM HEPES pH 7.5 containing 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 mM Mg^{2+}. Trapped and premixed complexes were preincubated in the presence of 2 mM Ca^{2+}. Data were fitted to equation 4 to give catalytic parameters shown in table 3.5.

Substrate	Unmixed		Premixed		Trapped		Blocked	
	k (min^{-1})	$t_{1/2}$ (s)	k (min^{-1})	$t_{1/2}$ (s)	k (min^{-1})	$t_{1/2}$ (s)	k (min^{-1})	$t_{1/2}$ (s)
DF-3	401 ± 26	0.1	634 ± 66	0.06	n/a	n/a	n/a	n/a
DF-3B	541 ± 50	0.08	288 ± 35	0.1	235 ± 20	0.2	0.02 ± 0.002	2100
DF-5B	*551 ± 40*	*0.08*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*	*n/a*
DF-21	518 ± 42	0.08	723 ± 78	0.06	n/a	n/a	n/a	n/a
DF-21B	462 ± 40	0.1	683 ± 64	0.06	448 ± 39	0.1	0.3 ± 0.002	139
DFGEN	365 ± 46	0.1	376 ± 71	0.1	n/a	n/a	n/a	n/a
DFGENB	204 ± 28	0.2	359 ± 54	0.1	160 ± 23	0.3	0.27 ± 0.03	154

Table 3.4: k_{ST} and $t_{1/2}$ for the double flap hFEN1 substrates calculated in this chapter. All measurements were made in 50 mM pH 7.5 HEPES containing 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 mM Mg^{2+}. Premixed and trapped reactions have 2 mM Ca^{2+} present to prevent reaction prior to initiation with Mg^{2+}. $t_{1/2}$ was calculated using the equation, $t_{1/2} = \ln(2) / k_{ST}$. *Measurements performed by Dr. John Atack.*

In addition to the dramatic effects of blocking substrates with streptavidin seen with both FEN proteins, two additional trends are observed. Firstly, a comparison of the rates of reaction of hFEN1 and T5 FEN catalysed reactions (tables 3.3-4) reveals T5 FEN k_{ST} values are approximately 6 fold lower than those of hFEN1 values for flap substrates. This is also reflected in the k_{cat} values from tables 3.1-2, implying that overall, hFEN1 is a much more efficient catalyst than its bacteriophage counterpart. This is in part due to the 3' flap binding pocket as detailed in chapter 1.6. Secondly, T5 FEN reactions under multiple turnover conditions are 5-9 times slower than those over single turnover conditions (tables 3.1 and 3.3) implying a rate limiting release of product as suggested

previously (Williams *et al.* 2007). However, for hFEN1 catalysed hydrolysis of double flap substrates k_{cat} values are only slightly lower than the k_{ST} values presented, implying a mechanism where product release is not rate limiting, at odds with previous work reported in (Finger *et al.* 2009). In this work, the hFEN1 substrates are different to those used in earlier work. The upstream and downstream binding regions used by Finger *et al.* are 18 bp and 20 bp respectively, whereas in this work, these regions are 6 bp and 12 or 18 bp, respectively. Substrates with longer upstream regions may be able to interact with the highly positively charged C-terminal tail of hFEN1 (Harrington and Lieber 1995; Gomes and Burgers 2000; Chapados *et al.* 2004; Tsutakawa *et al.* 2011; Zheng *et al.* 2011). Shortening the upstream regions, as in our substrates, may result in easier release of product, accounting for the similarity of k_{ST} and k_{cat}. The phage enzyme does not have a C-terminal tail and the length of upstream pY substrates (6 nt) is therefore optimal in terms of potential for interactions (Dervan *et al.,* 2001).

To test whether differences seen in the measurement of the cleavage of blocked substrates and unmixed/premixed/trapped substrates were not simply an effect of an altered enzyme-substrate binding equilibrium, an experiment analogous to the previous blocked single turnovers, but involving 10 times more enzyme (5 µM E cf. 500 nM E in figure 3.6.4), was designed to confirm the saturation of blocked substrate with hFEN1 (figure 3.6.8). If the previous reactions, in figures 3.6.4 and 7 were set up incorrectly, we would expect an equilibrium shift towards more bound blocked ES complexes that would result in a faster rate of reaction.

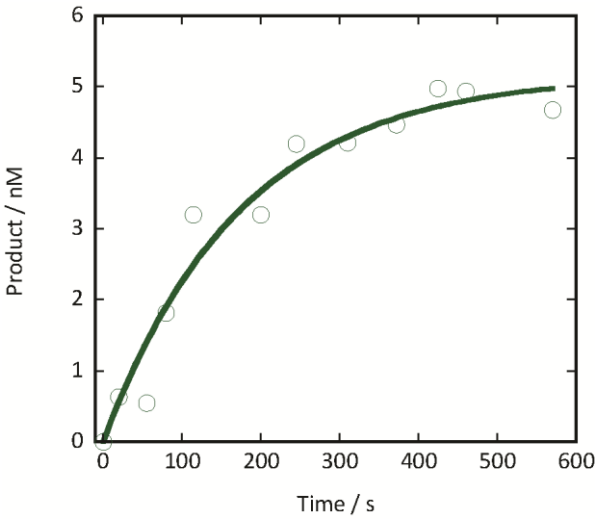


Figure 3.6.8: The single turnover profile 5 nM DF-21B and 5 µM hFEN1. Reaction was performed in 50 mM HEPES pH 7.5 containing 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 mM Mg^{2+}. Blocked ES complex was preformed as described in figure 3.6.2. Data were fitted to equation 4 to give a k_{ST} of 0.34 +/- 0.064 min^{-1}.

The reaction between 5 µM hFEN1 and 5 nM blocked DF-21B, gave a rate of 0.34 min^{-1}, identical within error to the reaction between 500 nM hFEN1 and 5 nM DF-21B, with a rate of 0.30 min^{-1}. This confirms the validity of the reaction conditions concerning blocked substrate single turnover reactions and indicates no significant perturbation of the enzyme-SA-substrate dissociation constant.

3.7 Discussion

To test the mode of interaction between FENs and their oligonucleotide substrates, DNAs with biotin modifications to their 5' ends were synthesised (figure 3.2.1-2). When SA is conjugated to these substrates, the conjugated protein is too large to prevent threading even when the helical arch is disordered. Thus, binding of SA was used to block access to the aperture of the helical arch preventing flaps from threading through the helical arch of FEN enzymes (figure 3.1.2). The expectation was that if threading was an absolute pre-requisite, then reaction would be prevented upon addition of 5'-SA. On the other hand, pre-incubation of enzyme and substrate potentially allowing threading, and subsequent formation of a trapped species, should result in a complex that reacts in a manner and timescale analogous to hFEN1 and a normal double flap substrate. Results demonstrated that trapped species did in fact react quickly. Although 5' blocked substrates were severely inhibitory, slow reaction was observed.

Binding cannot occur at the 5' terminus

The work in this chapter rules out a mechanism where binding occurs initially at the single-stranded 5' terminus of substrates, a model referred to as tracking. The flap strand of DFGEN terminates in a hairpin duplex, yet reaction of DFGEN still results in the expected endonucleolytic product (figure 3.4.2). Furthermore, the similar rate of hFEN1 catalysed reaction of the DFGEN substrate and other substrates with single stranded flaps (table 3.3-4) suggest that these reactions are occurring via the same mechanism. Additionally we show in figure 3.4.2 that this direct incision of the gapped flap is the result of hFEN1 catalysis, as it is drastically impaired by the hFEN1 K93A mutation.

Recent crystal structures of hFEN1 that reveal the bound surface area of each substrate component also support these conclusions (Tsutakawa *et al.* 2011). Whilst these structures contain product or a "flap" strand without a 5'-single-stranded extension, they underscore the importance of binding to the bifurcated junction, rather than the ss part of the flap. The total surface area between enzyme and substrate is 1828 Å (figure 3.7.1). Of this surface area, the template strand accounts for 887.9 Å, the 3' flap accounts for 221.4 Å and the duplex part of the 5' flap strand accounts for 388.1 Å. The flap strand's share of this interaction is approximately 10%, unlike the template strand, which is approximately 50%, suggesting that interaction with the template strand is the most important aspect in binding, as binding affinity is normally correlated with bound surface area (Tsutakawa *et al.* 2011). Examination of individual contacts between the protein and the DNA also emphasize the importance of template strand and single nucleotide 3 ' flap interactions. Biochemical support for an initial binding of the bifurcated junction was also recently provided by Gloor *et al.* who reported that SA blocked flaps were bound to hFEN1 with equal affinity to those without SA present. Additionally, Finger *et al*. also showed that the ds portion of the cleaved product of the hFEN1 reaction and exo substrates that do not possess a 5' ssDNA flap portion are both effective competitive inhibitors of hFEN1.

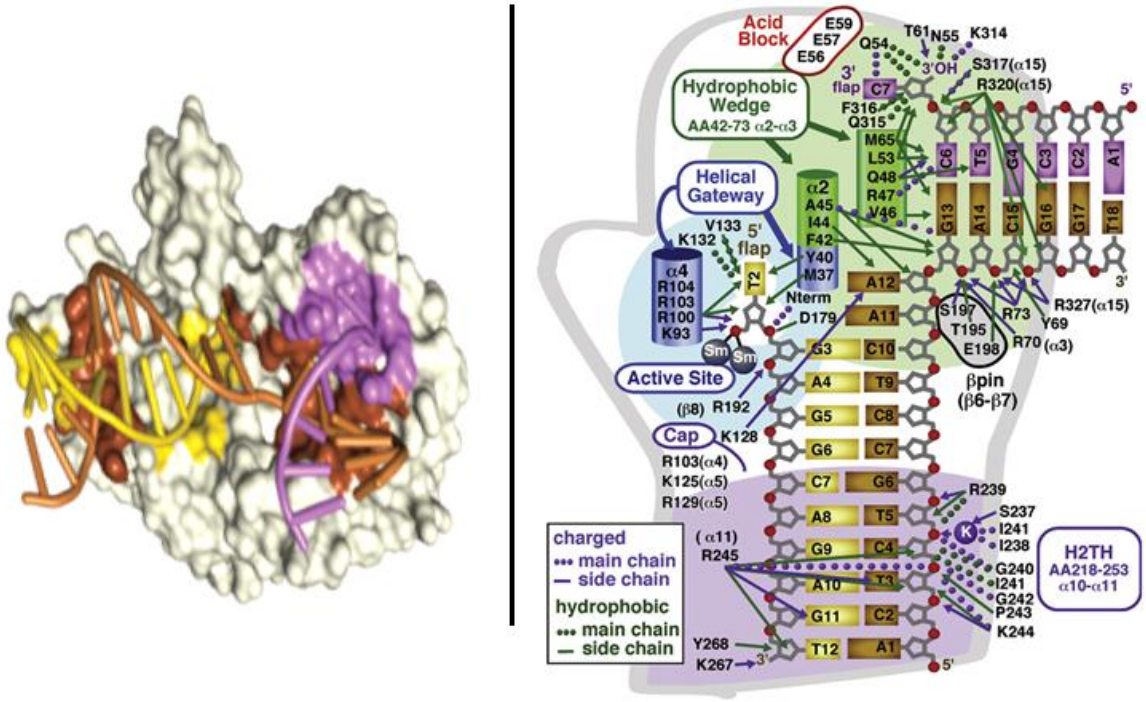


Figure 3.7.1: Left, structure of product bound FEN1 with the surface area between the 5' flap (yellow), template (brown) and 3' flap strand (purple) highlighted (taken from (Tsutakawa *et al.* 2011)). Right, model of hFEN1 with the interactions between the enzyme and the bound DNA product highlighted. The downstream binding area is highlighted in purple, the upstream and 3' flap binding area highlighted in green and the 5' flap binding region is highlighted in cyan. There are approximately 30 contacts between the upstream DNA and hFEN1, approximately 20-30 contacts between the downstream DNA, K^+ ion and hFEN1, and approximately 10 contacts between the 3' flap and hFEN1. In comparison there are 7 contacts to the duplex part of the 5' flap DNA shown.

Threading or Clamping: a reversible process

Rates of reaction of 'unmixed,' 'premixed' and 'trapped' ES complexes of both T5 FEN and hFEN1 are similar, with only a 2-fold decrease observed with 'trapped' vs. 'premixed' reactions, essentially showing all substrates are accommodated identically. The two fold decrease in rate going from 'premixed' to 'trapped' reactions, is expected, due to steric hindrance of the SA protein on the 5' end of substrates possibly affecting the efficient placement of key helical arch residues, and presentation of the scissile phosphate to the active site, but shows the more significant point that substrate with a free 5' terminus can be trapped in complex. In contrast, 'blocked' substrates cannot be optimally bound and cleaved on a biologically relevant timescale.

An interesting observation was that the conditions required to trap substrates with SA to give a large proportion of fast reacting substrate was different for hFEN1 and T5 FEN. For T5 FEN, fast reacting substrate could be trapped in both the presence of EDTA or Ca^{2+} ions. For hFEN1, carrying out the trapping procedure in EDTA resulted in two species, one that reacted quickly and another that reacted on the timescale of a blocked complex. This suggests an equilibrium between a clamped or threaded and unclamped or unthreaded states as depicted graphically for a threaded state below. The addition of SA captures the catalytically competent species (possibly threaded) and isolates a second complex where the 5'-flap portion of the substrate is not optimally positioned (figure 3.7.2).

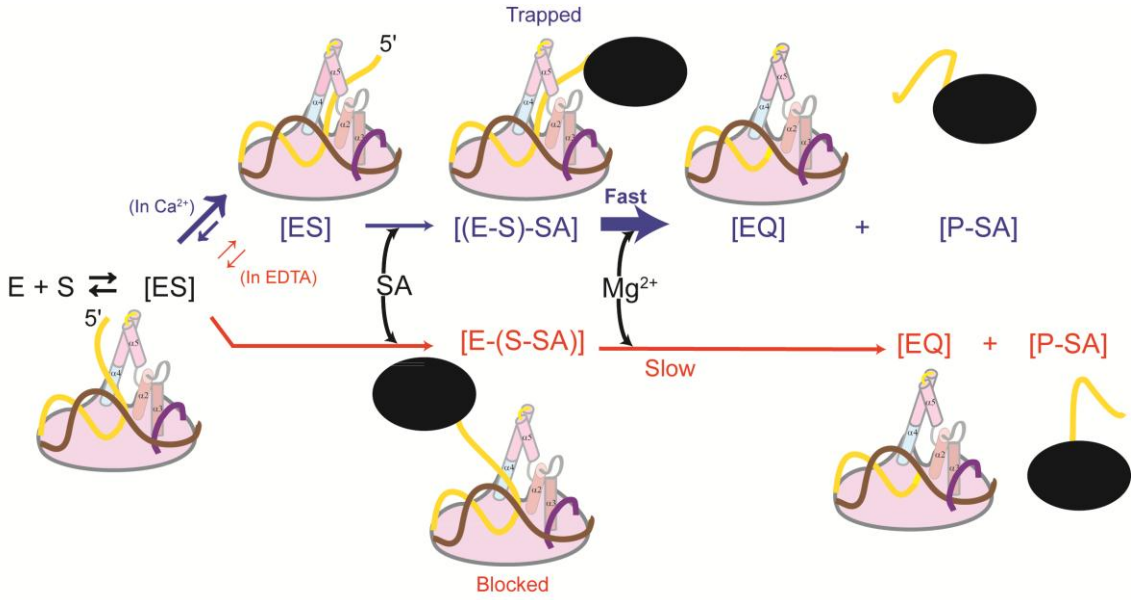


Figure 3.7.2: Schematic diagram of hFEN1 reaction scheme when treated as 'blocked' (red) and 'trapped' (blue) single turnover reactions as discussed in this chapter. The presence of divalent metal ions favours the formation of a catalytically competent ES complex. The phage T5 FEN enzyme would have an almost identical scheme, however T5 FEN does not see the same influence that divalent ions have on the equilibrium between [E-S] and [(E-S)] in hFEN1. Terms are defined in figure 3.6.1.

Addition of Ca^{2+} ions to the mixture of hFEN1 and biotinylated substrates prior to SA trapping produced a large proportion of fast reacting substrate. The results shown in figure 3.6.5 highlight a dynamic equilibrium between a catalytically competent threaded or clamped state and an

unthreaded or unclamped state in the case of hFEN1, but also show that this equilibrium can be manipulated depending on the metal ions present within the active site. This raises the question of why this is not seen in the phage enzyme, T5 FEN. One possibility is that the dissociation constants for substrates with hFEN1 are different in EDTA versus calcium. Whilst this is the case (see chapter 5), we show that the concentration of enzyme was saturating in EDTA. A possible explanation is provided by the structures of these enzymes. When comparing the helical arches of T5 FEN, and hFEN1, it can be seen that one of the helices making up the helical arch, alpha helix 5 differs (figure 3.7.3). In the structure of T4 RNase HI (homologous to the T5 enzyme) interactions between the single-stranded flap and residues of α5 were observed even though the flap is not positioned for reaction due to the absence of divalent ions. Figure 3.7.3 shows several conserved aromatic tyrosine and phenylalanine residues, and positively charged lysine residues on α5 that interact with the flap in the absence of metal ions. These residues are not present in hFEN1, which suggests that the repulsion from the seven conserved metal-binding carboxylates in hFEN1 does not favour accommodation of the flap strand in a position that can readily proceed to reaction without divalent metals present.

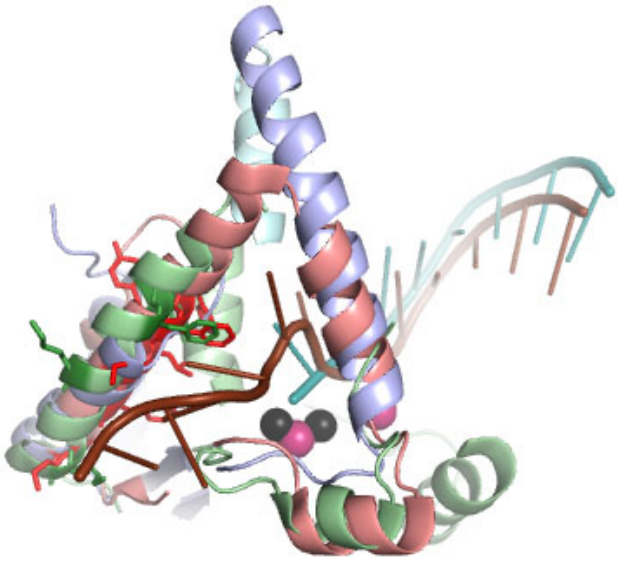


Figure 3.7.3: Overlay of the helical arches of T4 RNase HI (light green), T5 FEN (light red; accompanying substrate DNA - brown) and hFEN1 (light blue/cyan; accompanying substrate DNA - blue). Whereas alpha helix 5 in T4 RNase HI and T5 FEN is strongly conserved (left hand helix), this helix is replaced by alpha helix 2 and an additional 'cap' helix located above (cyan) in hFEN1. Alpha helices 5 from T4 RNase HI and T5 FEN possess many conserved aromatic and charged residues conserved in hFEN1.

The cleavage of 5' blocked substrates

A surprising result was that 5’ blocked substrates still reacted, albeit at a reduced rate. The plots shown in figures 3.6.3-4 and 3.6.6-7 show complete cleavage of every version of ‘blocked’ substrate, from the pseudo Y, to the double flap, flap and gapped substrates. Yet, the rate of reaction of the blocked 3' biotinylated substrate (pY-21-3'B, figure 3.6.3) was unaffected, showing that the lowered rate seen with the 5’ blocked substrates, is not due to a negative interaction between SA and the FEN enzymes.

The fact that monophasic, complete cleavage of 5' blocked substrates is observed shows that reaction of a small portion of non-biotinylated material is not occurring. This was also confirmed by MALDI-ToF data, with peaks within approximately 8 Da of predicted masses for all but 1 of the biotinylated substrates. No evidence of non-biotinylated material was present, as the mass of biotin is 569.61 Da. This makes it very likely that the slow cleavage of blocked substrates is a *bona fide* property of these enzymes. It cannot, however, be ruled out that this reaction is a result of the transient dissociation of SA from substrate, and cleavage occurs as soon as the 5' end of substrate becomes free. This is plausible because catalysis of reactions by hFEN1 approach the rate of diffusion. However, the biotin-SA complex has been shown to have a dissociation rate of $6.8e^{-5}$ s^{-1} at 37°C at pH 7.4 ($t_{1/2}$ of 3 hours or 10100 sec), showing that this would be a slow process (Chivers *et al.* 2010). It should be noted that this is the dissociation of free biotin from SA, not biotinylated DNA, and that the ratio of biotin to SA tetramer is not revealed in this study. As such, this half-life cannot be considered as a direct comparison. The most likely scenario is that the reaction observed is that of the SA blocked species as the SA blocked substrates display half-lives of 154 - 2100 seconds, much lower than that of the dissociation rate of biotin-SA.

There are several possibilities as to how FEN could deal with ‘blocked’ substrates. As seen in figure 3.7.2, the substrate could be clamped rather than threaded. However if clamping was the usual mechanism, it is difficult to understand why SA on a long flap is inhibitory. Alternatively, as threading would be prevented, the substrate could be forced to take an alternative route that does not pass through the arch. This could lose interaction with important residues, such as K93 in hFEN1 resulting in a large decrease in catalytic efficiency. This is suggested by the fact that the rates of both blocked substrates with WT hFEN1 are strikingly similar to the rates with premixed K93A hFEN1, 0.3 cf. 0.3

min^{-1} for DF-21B and 0.17 cf. 0.27 min^{-1} for DFGENB (figure 3.4.2). Another possibility could be that the bulk of SA is blocking the proper ordering of the helical arch that is thought to be seen on substrate binding, which would alter the active site resulting in a similar decrease in the rate of reaction. Lastly the SA could be pushing some portion of the helical arch backwards in order for substrate to be accommodated correctly, once again altering the active site (Tsutakawa *et al.* 2011).

In conclusion the work in this chapter has effectively ruled out all hypotheses based on a 5' first mode of binding, and instead, implies that binding must occur at the region of bifurcation. However there remains some doubt about how ssDNA is accommodated by the helical arch, be it clamping or threading. This is due to the successful and complete albeit slow cleavage of 5' blocked substrates, and the fact that the rate of this reaction alters when the 5' SA is held 3nt away from the active site as opposed to 21nt away. Although these results could be explained by a bind then thread mechanism, a clamping mechanism cannot be completely ruled out.

Chapter 4: Competition Experiments

4.1 Introduction

As discussed in chapters 1 and 3, there are several proposals for how FENs and the related enzyme EXO1 accommodates the scissile phosphate and the 5'-portion of substrates within their active site. One model, known as tracking proposed an initial interaction between enzymes and substrate to be with 5' ss flaps, but several pieces of evidence now discredit this model as detailed in chapter 3. Remaining models involve initial interaction with the main junction part of substrates containing the double-stranded region from which the flap protrudes. However, they differ in the way the 5'-flap is accommodated and cleaved. The first of these is a bind then clamp model. In this model, FENs would bind the downstream and upstream regions of substrates initially, and then clamp the ss region of DNA either side of the helical arch (Orans *et al.* 2011). The 5' flap of substrates are hypothesised to exit the helical arch of FENs by contacting the active site, and passing in front of alpha helix 4 (path 1) or between alpha helix 2/5 (path 2) (figure 4.1.1) .

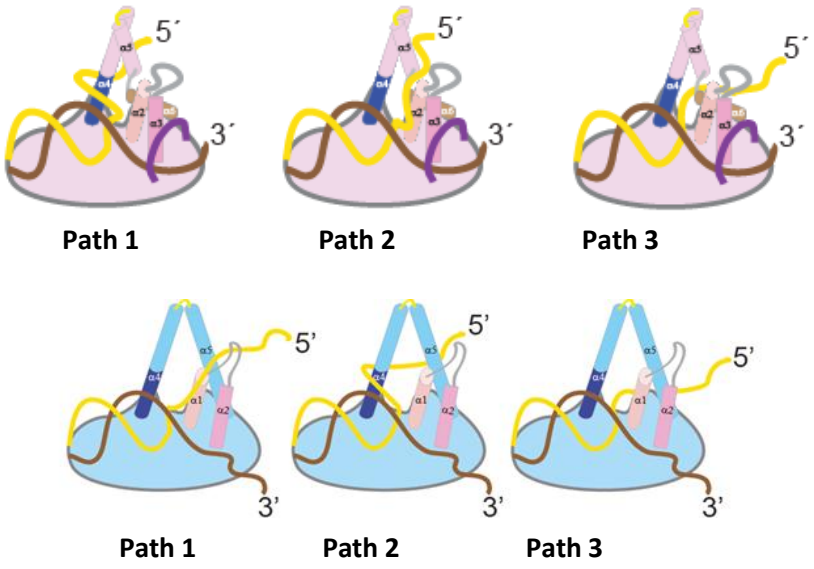


Figure 4.1.1: Top, The pathways postulated by Beese *et al.* (2011) illustrated on a model of hFEN1. Path 1 shows the path taken by the substrate behind the alpha 4 helix; path 2 shows the path taken by substrate if it were to exit the enzyme from between the alpha helices 2 and 5. Path 3 shows the route taken by the threading mechanism (Orans *et al.* 2011). **Bottom, The pathways as shown in the hEXO1 structure, illustrated on models of T5 FEN.** All models shown are courtesy of Dr. David Finger. The substrate strands shown are termed the template (brown) strand, flap (yellow) strand and 3' flap (purple) strand.

An alternative mechanism suggests that the downstream and upstream regions are bound to FENs prior to threading of the ssDNA flap through a structured arch. A variation on this threading mechanism posited suggests that this threading might occur through an unstructured arch, which would then order around the ssDNA portion of substrates (Tsutakawa *et al.* 2011). As the small aperture of the structured helical arch would only allow ssDNA to pass through, a bind then thread through the helical arch mechanism is difficult to reconcile with robust activity observed on gapped substrates that contain duplex within flaps. The version of the bind then thread mechanism which suggests binding at the bifurcated region and then threading of the flap through a structured arch has another major flaw, in that it is hard to envisage the passing of ssDNA through such a small aperture without the aid of an energy source like ATP hydrolysis (Orans *et al.* 2011).

The bind then thread through a disordered arch model involves the binding of the 'template' strand (see figure 4.1.1 for nomenclature) of the DNA substrate at the downstream binding region, departing the surface of the enzyme, and re-joining the enzyme at the upstream binding region, creating an arc. This arc delivers the accompanying flap strand underneath it towards the unstructured aperture of the helical arch. Along with this arced substrate delivery (figure 4.1.2, top), introducing the flap to the helical arch by default, it has been proposed that the binding of the 1 nucleotide 3' flap triggers the ordering of the hydrophobic wedge (figure 4.1.2, bottom), and the ordering of the helical arch around the 5' ssDNA portion of substrates, subsequently helping unpair the two terminal base pairs of substrates ready for cleavage (Tsutakawa *et al.* 2011). The molecular basis of these substrate induced conformational changes is suggested to be tertiary interactions between the 3'-flap site, the wedge and alpha 5 of the helical arch.

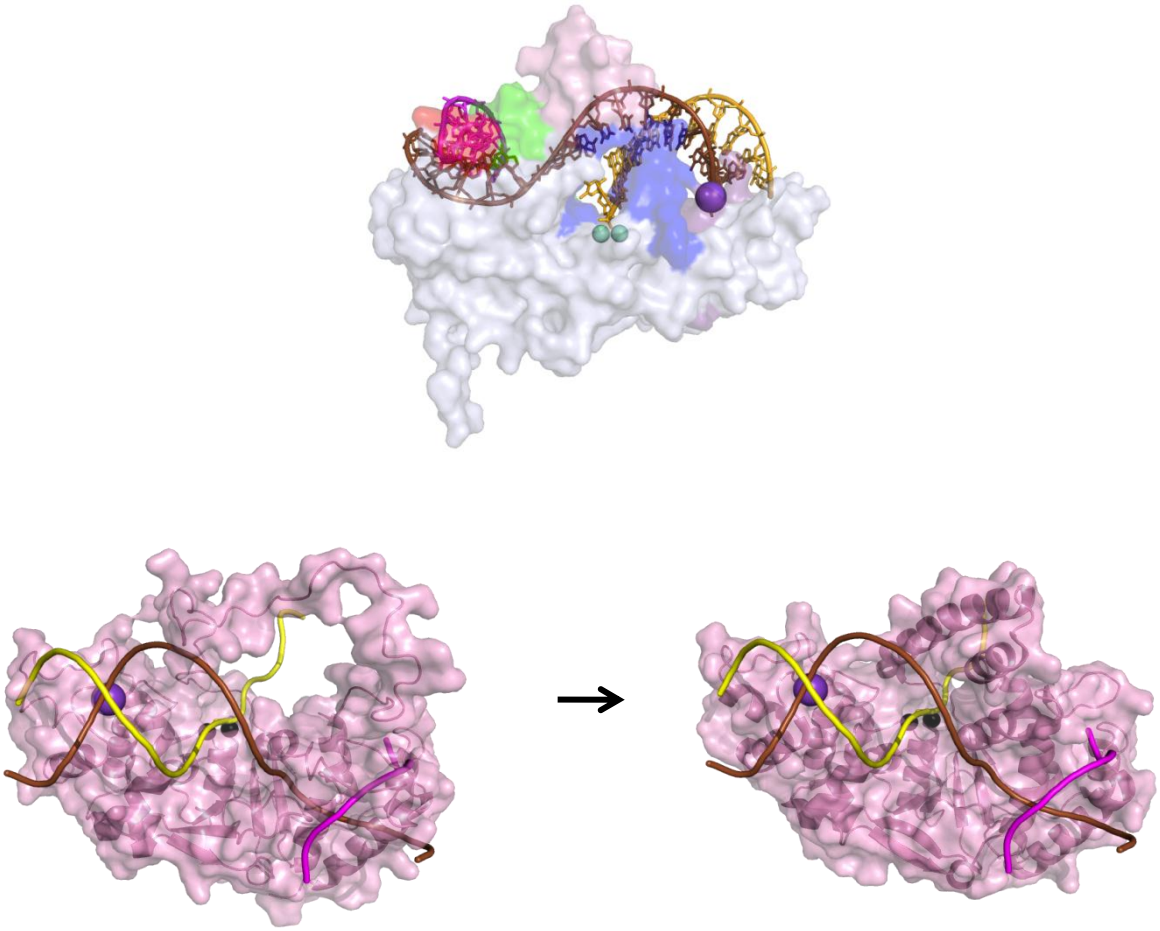


Figure 4.1.2: Illustrations of the possible bind then thread mechanism of FENs: Top (transparent surface representation of hFEN1), the flap strand (yellow) follows the arc of the template strand (brown) as it leaves the surface of hFEN1, bending back around and contacting the downstream binding region. The flap strand would be delivered to the helical arch by default. **Bottom (transparent surface representation of hFEN1, with cartoon secondary structure shown for clarity),** the hypothesised disorder to order transition of the helical arch of hFEN1.

Recently it was suggested that FENs, more specifically hFEN1, employed 2 different mechanisms to cleave substrates dependant on the length of the 5' flap. It was suggested flaps 6 nt and under, were not threaded through the helical arch, while flaps 6 nt and over, were. This proposal was based on differences in the K_Ds derived from gel shift assays without divalent metal ions present (Gloor *et al.* 2010). Trapping experiments described in chapter 3 suggest an equilibrium between threaded and non-threaded exists that is greatly altered by divalent metal ions; hence, the basis of these

conclusions deserves investigation. Thus, whether short and long flaps are processed differently will be tested in the experiments proposed below.

Proposed Experiments

To distinguish between the three pathways outlined above in section 4.2, competition experiments were designed. If threading is the mode of interaction of FEN with the 5'-flap portion of its substrates, then formation of a threaded 5'-biotinylated substrate complex followed by addition of streptavidin should produce a complex that cannot exchange when excess unlabelled substrate is added. In contrast, if the substrate is clamped by the arch, adding SA should not produce a complex that is resistant to competition, assuming formation of the clamped structure is reversible. Similarly, in the absence of SA, complexes should be readily exchanged upon addition of unlabelled substrate providing that any process that accommodates the 5'-flap is reversible. In these experiments, the status of complexes that were assessed by challenging pre-formed complexes with competitor substrate for a fixed amount of time, and then adding magnesium ions to initiate reaction. Successful competition should be characterised by a lack of FAM-labelled product. Predicted outcomes of the experiments are summarised in figure 4.1.4.

Order of Addition of Substrates	Route 1 Threading 	Route 2 Clamping A 	Route 3 Clamping B 
	Successful Competition?		
1) Preassembled ES complex 2) SA 3) Competitor 4) Mg2+	No	Yes	Yes
1) Preassembled ES complex 2) Competitor 3) Mg2+	Yes	Yes	Yes

Figure 4.1.4: Predicted results of competition experiments in each binding model scenario. Left, the general scheme of these competition reactions whereby enzyme and substrate are pre-assembled as described in figure 3.6.2, and reagents are added until the addition of Mg^{2+} to initiate the reaction. Right, the proposed routes and their predicted outcomes.

Two competitor oligonucleotides (figure 4.1.3) were created for these tests. For experiments with T5 FEN, the unlabelled competitor substrate was of identical sequence to the unimolecular substrate pY-21B (figure 3.2.1), but lacked the biotin and FAM moieties (referred to as COMP). The unlabelled competitor for hFEN1 complexes was a double flap assembled from two oligomers like the substrates described in chapter 3, having a 5 nt 5'-ss flap, but no biotin and FAM (referred to as DF-COMP5).

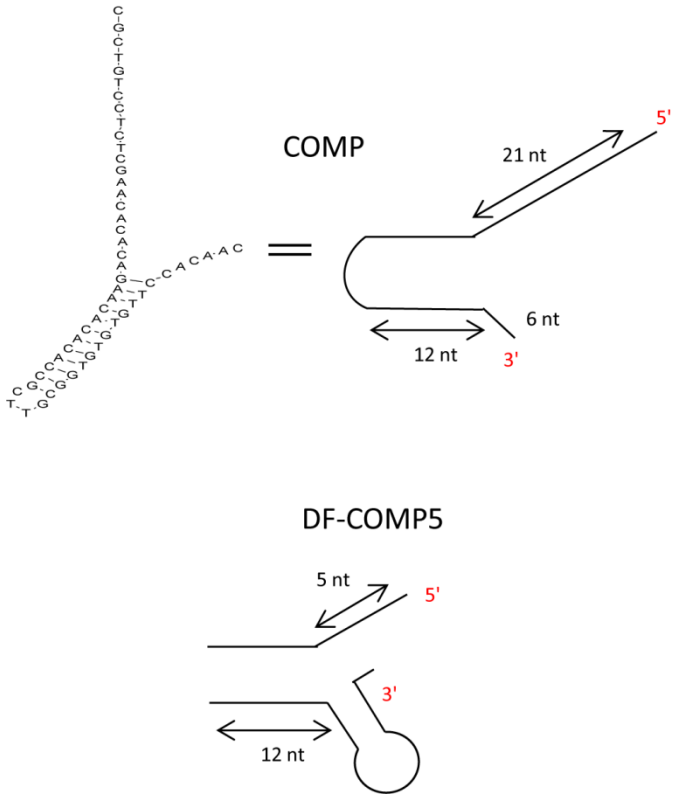


Figure 4.1.3 Competitor substrates used in this chapter. COMP, a substrate with a 21 nucleotide 5' flap, identical to pY-21B. **DF-COMP5**, a bimolecular substrate with a 21 nucleotide 5' flap and 1 nucleotide 3' flap.

4.2 T5 FEN Competition Experiments

Premixed (black), trapped (pink) and blocked (cyan) ES complexes were assembled using T5 FEN and the 21 nt 5' flap pseudo Y substrate (pY-21B) in the presence of EDTA as previously described in chapter 3. The schemes for the construction of premixed, trapped and blocked complexes are repeated below for convenience (figure 3.6.2).

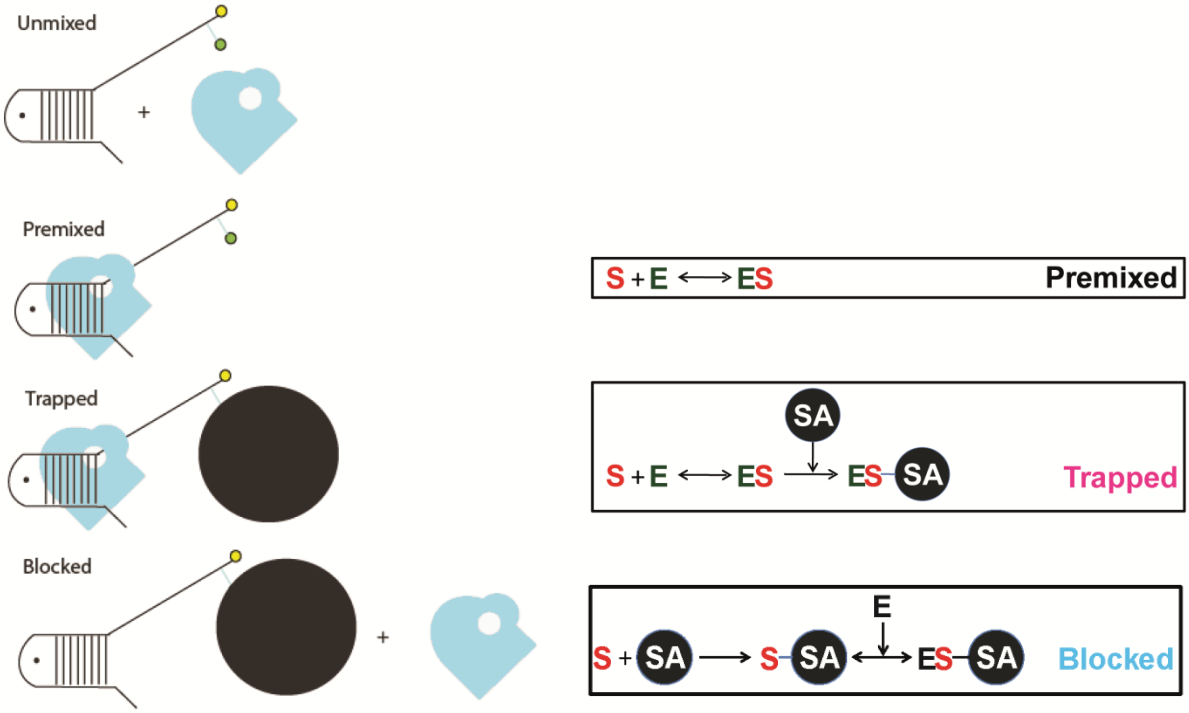


Figure 3.6.2: Diagrams showing the procedures involved in forming 'premixed,' 'trapped' and 'blocked' enzyme-substrate complexes. In the case of T5 FEN, 5 nM Substrate and 500 nM enzyme were added on ice and incubated for 2 minutes in the presence of EDTA. For hFEN1, pre-incubation took place at room temperature for 2 minutes in the presence of catalytically inert Ca^{2+} ions. SA was added on ice, and the mixture incubated at room temperature for one minute.

A 5-fold excess of unlabelled competitor oligo (with respect to enzyme) was added after formation of the respective complexes. After the assembly of complexes, and where appropriate addition of competitor oligonucleotide, complexes were incubated at 37°C for 10 mins. Reaction was then initiated with magnesium ions, and monitored at a series of time points up to 2360 ms, which corresponds to approximately 6 half-lives of an SA trapped pseudo Y complex ($t_{1/2}$ was measured to be 400-600 ms in chapter 3). Although reagents were added in various orders and in some experiments, competitor was not added, all reaction mixes were treated identically. A control experiment quantified the amount of FAM labelled biotinylated product produced by an unchallenged trapped ES complex (magenta). Another control experiment was conducted where competition was attempted from a premixed complex and then SA was added (light grey), which should provide results typical of a successful competition. The results of these experiments are shown in figure 4.2.1.

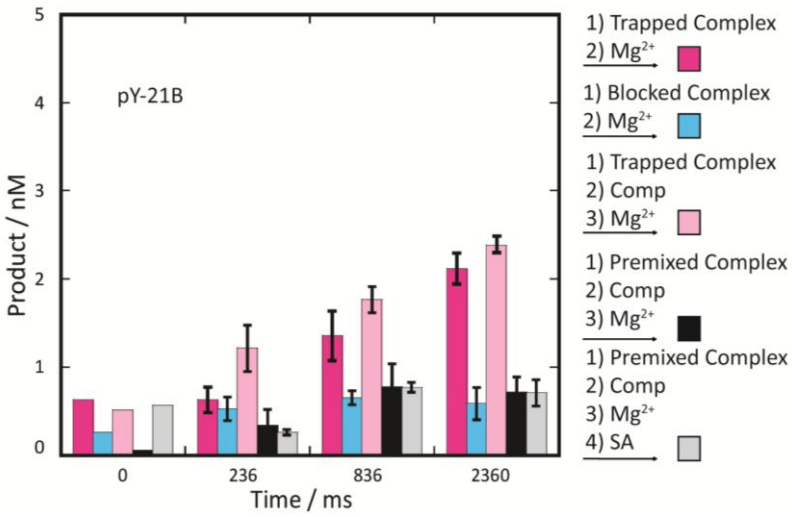


Figure 4.2.1: The effect of pre-addition of non-biotinylated unlabelled competitor substrate on the outcome of reactions of T5 FEN-pY-21B ES complexes. Decay of a SA 'trapped' ES complex, without addition of competitor is shown in magenta. As with experiments where unlabelled competitor substrates were added, the complex was incubated for 10 mins at 37°C prior to reaction. Decay of a SA 'blocked' ES complex, to which competitor substrate (2.5 µM) was added followed by incubation at 37°C for 10 mins prior to reaction, is shown in cyan. Decay of 'trapped' and 'premixed' complexes, to which competitor substrate (2.5 µM) was added followed by incubation at 37°C for 10 mins prior to reaction, are shown in pink and black respectively. Decay of a 'premixed' ES complex, to which competitor substrate (2.5 µM) was added followed by incubation at 37°C for 10 mins, addition of 5 equivalents of SA and incubation at 37°C for a further minute prior to reaction is shown in grey. All complexes are formed as shown in figure 3.6.2, and reactions are initiated by the addition of 10 mM Mg^{2+}. All data points are the result of three independent experiments, with standard errors shown. Reactions were carried out as described in section 2.8a.

When complexes of T5 FEN and 5'-biotinylated substrate were trapped with SA, equivalent amounts of product were produced even when the complex was challenged with unlabelled competitor substrate (compare magenta "trapped" with light pink "trapped" then competed). In these cases the concentration of product formed was approximately 2.5 nM after 2.36 seconds. This is comparable to the levels of product formed after 2.36 s in both unmixed and premixed single turnovers shown in chapter 3 figure 3.6.4. However, when analogous competitions are performed with premixed and blocked complexes shown in cyan, black and grey, the level of FAM labelled product never rises

above 0.9 nM. In most cases, the low concentrations of product are observed before (time zero) and after addition of magnesium ions indicating that a small amount of reaction occurs as the complexes are assembled and pre-incubated. The results show successful competition with premixed and blocked complexes, and unsuccessful competition when complexes are trapped after pre-incubation, consistent with a possible bind then thread mechanism occurring.

A similar set of reactions was also performed where competition was attempted from premixed (magenta), trapped (pink) and blocked (cyan) ES complexes assembled with 3' biotinylated substrate, pY-21-3'B and compared to the decay of a premixed ES complex as normal (black) (figure 4.2.2). The expectation was that competition would be successful from all of these complexes due to any steric 'road block' being on the 3' terminus of substrates, thereby resulting in a decrease in FAM-labelled substrates, which would be observed in all cases except the decay of a premixed ES complex.

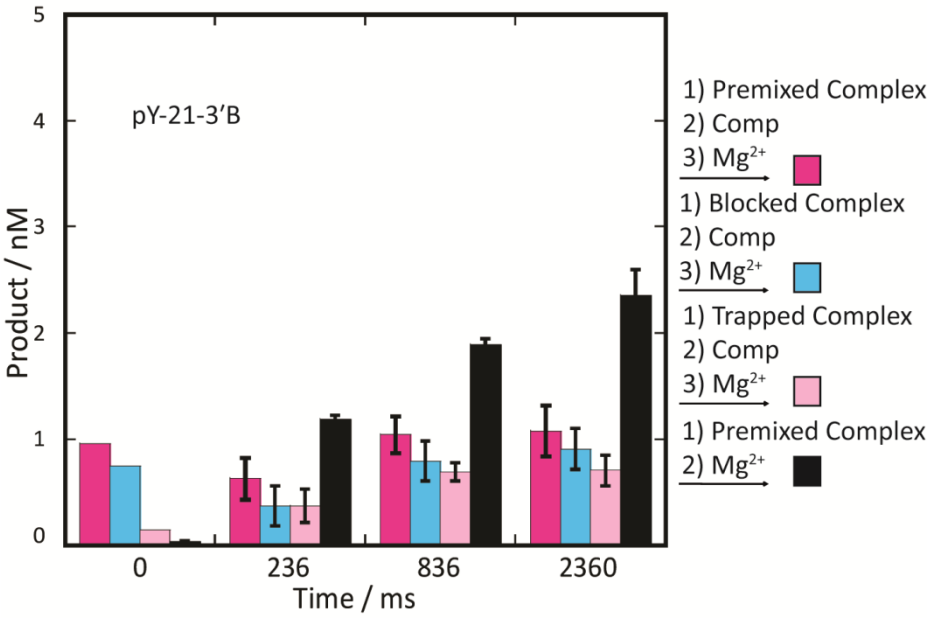


Figure 4.2.2: The effect of addition of non-biotinylated unlabelled competitor substrate on the outcome of reactions of T5 FEN-pY-21-3'B ES complexes. Decay of a SA 'trapped' ES, without addition of competitor is shown in magenta. As with experiments where unlabelled competitor substrates were added, the complex was incubated for 10 mins at 37°C prior to reaction. Decay of a SA 'blocked' ES complex, to which competitor substrate (2.5 µM) was added followed by incubation 37°C for 10 mins prior to reaction is shown in cyan. Decay of a 'trapped' complex to which competitor substrate (2.5 µM) was added followed by incubation at 37°C for 10 mins prior to reaction is shown in pink. Decay of a 'premixed' ES complex, is shown in black. As with experiments where unlabelled competitor substrates were added, the complex was incubated for 10 mins at 37°C prior to reaction. All complexes are formed as shown in figure 3.6.2, and reactions are initiated by the addition of 10 mM Mg^{2+}. All data points are the result of three independent experiments, with standard errors shown. Reactions were carried out as described in section 2.8a.

As expected, competition from the 'premixed', 'trapped' and 'blocked' complexes that involve 3'-biotinylated substrate (magenta, pink and cyan respectively) were successful, with the levels of FAM labelled product never exceeding 1.1 nM. The 'premixed' reaction in black shows normal levels of cleavage for a pseudo Y substrate with a 21nt 5' flap as seen previously in the single turnover reactions of 'premixed' and 'unmixed' pY-21B, pY-21 and pY-21-3'B in chapter 3 (figure 3.6.4). These

experiments with 3'SA-trapped complexes show that the conjugation of SA to oligo is not the reason for the lack of successful competition of the 5'-SA "trapped" complexes. These results are once again supportive of a bind then thread mechanism being employed by T5 FEN.

4.3 hFEN1 Competition Experiments

Competition experiments with hFEN1 were carried out with different length 5' flaps to test the hypothesis set by Gloor *et al.* that flaps less than 6 nt do not thread, whereas those greater than 6 nt do. These were carried out in the same format each time using the 3 nt 5' flap DF-3B (figure 4.3.1), the longer 21 nt 5' flap DF-21B (figure 4.3.2) and the gapped flap, DF-GENB (figure 4.3.3). Competition experiments were attempted with a premixed (black) and trapped (pink) ES complex, preassembled as previously described in chapter 3. If a threading mechanism is also occurring with hFEN1, competition of a labelled biotinylated substrate will be successful from a premixed and blocked ES complex, but not from a trapped ES complex (figure 4.1.4). Moreover, successful competition would be characterised by a lack of FAM-labelled product. For comparative purposes, the concentration of FAM labelled biotinylated product provided by a trapped (magenta) and premixed ES complex (grey) that had gone through the same incubations without addition of competitor were also measured. As a control, addition of competitor to biotinylated substrate from a premixed complex with subsequent addition of SA was tested (light grey). The results of these experiments using the three different hFEN1 substrates are shown in figures 4.3.1-3.

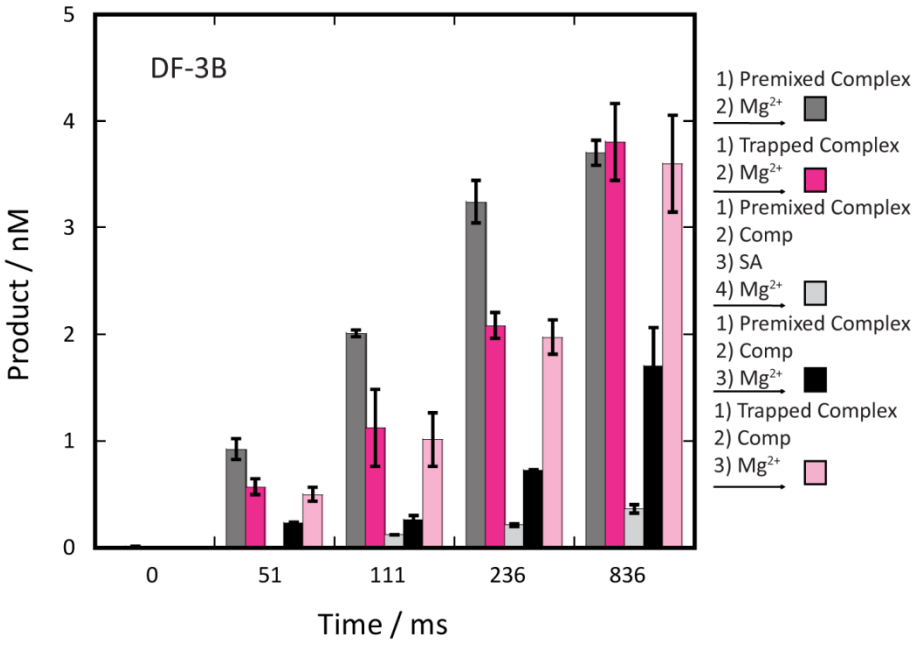


Figure 4.3.1: The effect of pre-addition of non-biotinylated unlabelled competitor substrate on the outcome of reactions of hFEN1-DF-3B ES complexes. Decay of a 'premixed' ES complex, without addition of competitor is shown in dark grey. As with experiments where unlabelled competitor substrates were added, the complex was incubated for 10 mins at 37°C prior to reaction. Decay of a SA 'trapped' ES complex, without addition of competitor is shown in magenta. As with experiments where unlabelled competitor substrates were added, the complex was incubated for 10 mins at 37°C prior to reaction. Decay of a 'premixed' ES complex, to which competitor substrate (5 µM) was added followed by incubation at 37°C for 10 mins, addition of 5 equivalents of SA and incubation at 37°C for a further minute prior to reaction is shown in grey. Decay of 'premixed' and 'trapped' complexes to which competitor substrate (5 µM) was added followed by incubation at 37°C for 10 mins prior to reaction are shown in black and pink, respectively. All complexes are formed as shown in figure 3.6.2, and reactions are initiated by the addition of 8 mM Mg^{2+}. All data points are the result of three independent experiments with standard errors shown. Reactions were carried out as described in section 2.8b.

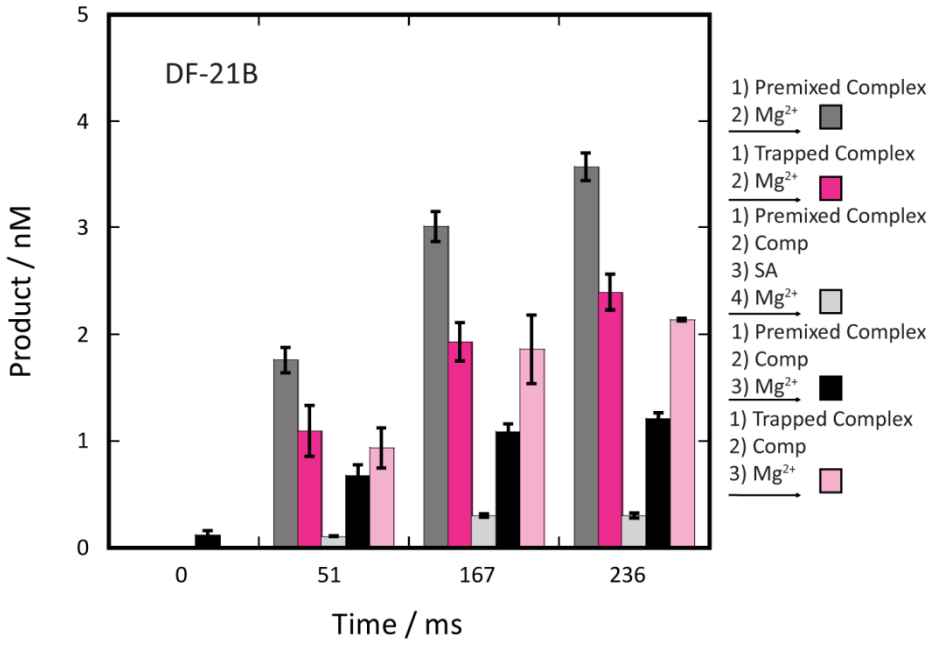


Figure 4.3.2: The effect of pre-addition of non-biotinylated unlabelled competitor substrate on the outcome of reactions of hFEN1-DF-21B ES complexes. Decay of a 'premixed' ES complex, without addition of competitor is shown in dark grey. As with experiments where unlabelled competitor substrates were added, the complex was incubated for 10 mins at 37°C prior to reaction. Decay of a SA 'trapped' ES complex, without addition of competitor is shown in magenta. As with experiments where unlabelled competitor substrates were added, the complex was incubated for 10 mins at 37°C prior to reaction. Decay of a 'premixed' ES complex, to which competitor substrate (5 µM) was added followed by incubation at 37°C for 10 mins, addition of 5 equivalents of SA and incubation at 37°C for a further minute prior to reaction is shown in grey. Decay of 'premixed' and 'trapped' complexes to which competitor substrate (5 µM) was added followed by incubation at 37°C for 10 mins prior to reaction are shown in black and pink, respectively. All complexes are formed as shown in figure 3.6.2, and reactions are initiated by the addition of 8 mM Mg^{2+}. All data points are the result of three independent experiments, with standard errors shown. Reactions were carried out as described in section 2.8b.

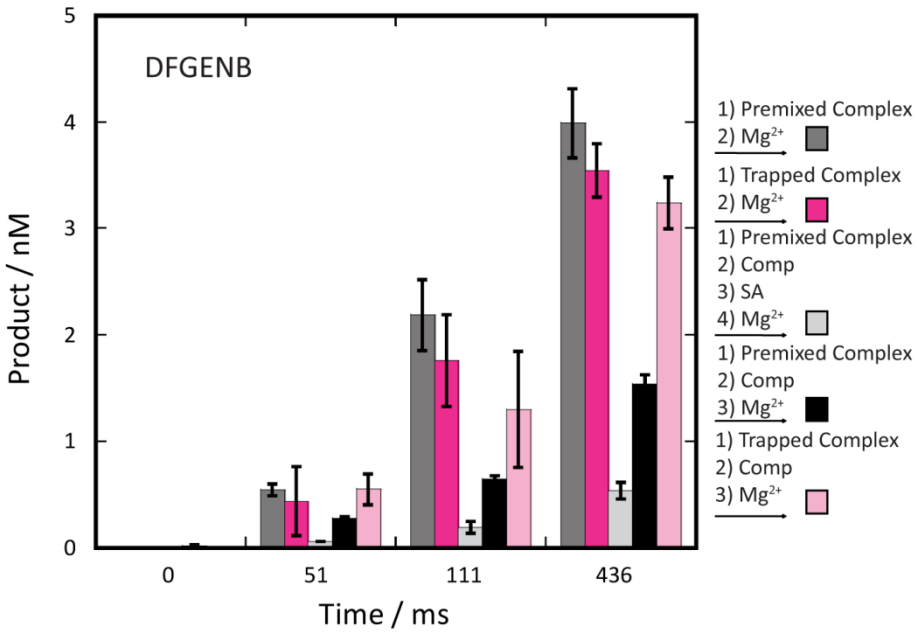


Figure 4.3.3: The effect of pre-addition of non-biotinylated unlabelled competitor substrate on the outcome of reactions of hFEN1-DF-GENB ES complexes. Decay of a 'premixed' ES complex, without addition of competitor is shown in dark grey. As with experiments where unlabelled competitor substrates were added, the complex was incubated for 10 mins at 37°C prior to reaction. Decay of a SA 'trapped' ES complex, without addition of competitor is shown in magenta. As with experiments where unlabelled competitor substrates were added, the complex was incubated for 10 mins at 37°C prior to reaction. Decay of a 'premixed' ES complex, to which competitor substrate (5 µM) was added followed by incubation at 37°C for 10 mins, addition of 5 equivalents of SA and incubation at 37°C for a further minute prior to reaction is shown in grey. Decay of 'premixed' and 'trapped' complexes to which competitor substrate (5 µM) was added followed by incubation at 37°C for 10 mins prior to reaction are shown in black and pink, respectively. All complexes are formed as shown in figure 3.6.2, and reactions are initiated by the addition of 8 mM Mg^{2+}. All data points are the result of three independent experiments, with standard errors shown. Reactions were carried out as described in section 2.8b.

Regardless of substrate flap length or secondary structure, the concentrations of FAM-labelled products produced from a 'premixed' complex (grey) are greater than those produced by 'trapped' complex (magenta) at early time points (figures 4.3.1-3). This is in line with earlier determination of

the rates of decay of these complexes (chapter 3, table 3.4), where premixed complexes reacted approximately two fold faster. However, when monitored for a longer time period (836 ms, essentially to end point in both cases, figure 4.3.1), both complexes produced similar amounts of product. A comparison of concentrations of product produced from 'trapped' complexes (magenta) to 'trapped' complexes to which a large excess of competitor is added reveals them to be identical within experimental error. In contrast, the competition of a 'premixed' complex and the analogous competition experiment with SA added after competitor show large decreases in FAM labelled product implying successful competition in all cases. The results shown in figures 4.3.1-3, in particular, the inability to compete labelled oligonucleotides from a trapped ES complex implicate a threading mechanism for hFEN1 (figure 4.1.4).

In the reactions where unlabelled competitor was added to 'premixed' FEN1-DF21B/DF3B/DFGENB complexes before initiation of the reaction with the addition of Mg^{2+} ions, levels of FAM labelled product were only a minimum 2-fold lower than that of a normal decay of the corresponding 'premixed' or 'trapped' ES complex at the 8.36 s timepoint in figures 4.3.1-3. Despite this, it is apparent by referring to the $t_{1/2}$ of these reactions (table 3.4) that successful competition has occurred in these situations. This can be rationalised by the fact that for every half-life that passes, a portion of FAM labelled substrate would be expected to react. That is, if 10X excess competitor with respect to the enzyme is present, $1/10^{th}$ of the FAM labelled product formed in the normal single turnover reactions would be formed. The further the reaction progresses, a higher portion of FAM labelled product would be formed due to a reduction in intact competitor present. In addition, as these reactions occur over 8-10 half-lives, this amount of FAM labelled product is not unexpected.

4.4 Further Competition Experiments

Due to the significant nature of the competition results, strongly implying one proposed mechanism over the others, several additional experiments were designed to verify the observations shown in sections 4.2-4. Firstly, the ability to compete away complexes formed from each length double flap substrate with hFEN1 was tested in the presence of EDTA instead of catalytically inert Ca^{2+} ions (figure 4.4.1). Despite the hypothesised equilibrium shown in chapter 3 (Figure 3.6.6) where it is suggested that in EDTA a mixture of threaded and non-threaded complexes exist, the results of the competition experiments with preassembly of complexes in either EDTA or Ca^{2+} ions should theoretically be almost identical. This is because the only factor affected when calcium ions are

removed should be the equilibrium between a catalytically and non-catalytically competent ES complex, and both types of complexes should be able to dissociate. Thus, the competitor should still be able to effectively compete the FAM labelled substrate out of a 'premixed' ES complex in EDTA.

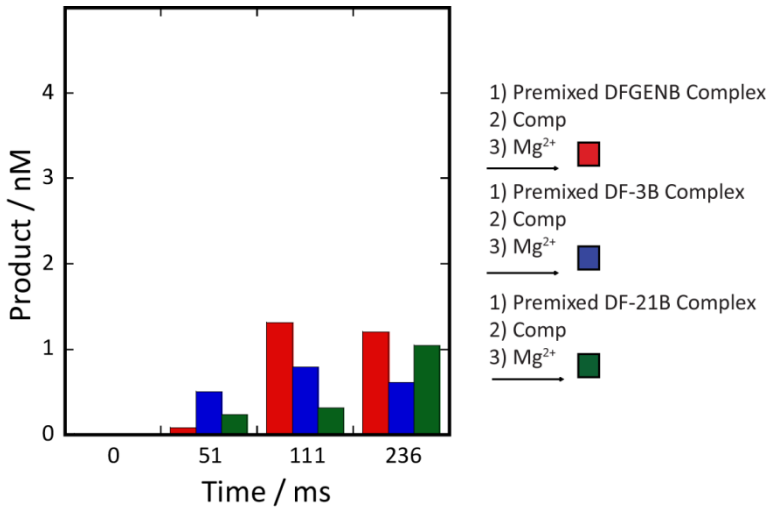


Figure 4.4.1: The effect of pre-addition of non-biotinylated unlabelled competitor substrate on the outcome of reactions of hFEN1-DFGENB (red)/3B (blue)/ 21B (green) ES complexes. Decay of a 'premixed' E-DFGENB complex to which competitor substrate (5 µM) was added followed by incubation at 37°C for 10 mins is shown in red. Decay of a 'premixed' E-DF-3B to which competitor substrate (5 µM) was added followed by incubation at 37°C for 10 mins is shown in blue. Decay of a 'premixed' E-DF-21B complex to which competitor substrate (5 µM) was added followed by incubation at 37°C for 10 mins is shown in green. All complexes are formed as shown in figure 3.6.2, and reactions are initiated by the addition of 10 mM Mg^{2+}. Reactions were carried out as described in section 2.8b; however, buffers contained 1 mM EDTA instead of 2 mM Ca^{2+}.

As seen in figure 4.4.1, a very similar pattern to the results with Ca^{2+} pre-incubated complexes (figure 4.3.1-3 (black)) was observed, with the FAM labelled product rising slowly and never exceeding 1.5 nM in total.

A series of experiments were also performed, shown below in figure 4.4.2 where a mixture of substrate and competitor were pre-incubated together prior to enzyme, to test whether the results

in 4.3.1-3 were reproducible if the order of reagent addition differed slightly. Reactions, where competitor, substrate and SA were preincubated (blue) and where competitor and substrate were preincubated together (green) prior to addition of enzyme, were conducted to replicate the results observed in experiments trying to compete away a blocked complex. These controls should give the same result as the other successful competitions shown in figures 4.2.1 and 4.3.1-3 (black). The reaction with substrate and competitor pre-incubated together prior to addition of enzyme (figure 4.4.2, green) shows a gradual increase in FAM labelled product up to around 0.8 nM at the 236 ms time-point. Comparison with the analogous experiment where ES was made prior to addition of competitor in figure 4.3.1 (black) shows that the amounts of product are in good agreement with one another. The reaction with substrate and competitor pre-incubated together along with SA prior to addition of enzyme (figure 4.4.2, blue) also shows a gradual increase in FAM labelled product up to around 0.1-2 nM at the 236 ms time-point. Comparison to the analogous experiment, where ES was made prior to addition of SA and competitor, (figure 4.3.1, light grey) shows that results are almost identical.

A concern involving the competition experiments was the possibility of SA tetramer bound to more than one substrate, thereby resulting in a high local concentration of substrate in experiments. In the gel shift assay used to assess SA binding to biotin substrates shown in figure 3.3.2, there was evidence for more than one biotinylated oligonucleotide substrate binding to one SA when one "equivalent" (manufacturer's unit) of SA was used. All competition experiments described above used five "equivalents" of SA. Nonetheless, to rule out the possibility that this was a factor that led to a failure to compete away trapped substrates, the competition experiment was also performed with an even larger excess of SA, that should lead to negligible amounts of any species that is not 4:1 SA monomer:substrate. Additionally, dimeric and trimeric species could have been a factor both in inhibiting blocked reactions, or even in facilitating a transient release of substrate that led to observation of a blocked reaction. Competitor was added to a 'premixed' E-DF-3B complex prior to the addition of 20 equivalents of SA, and subsequent initiation of the reaction using Mg^{2+} was carried out (figure 4.4.2, red). Figure 4.4.2 (red) shows an identical amount of FAM labelled product was formed to that previously in figure 4.3.1 (light grey), implying that there are no negative interactions between the enzyme or substrate and SA, as these would be exacerbated in this experiment. Furthermore, any dissociation of substrate from SA is unlikely even in the cases of reactions where the biotin:SA tetramer ratio is closer to 4:1.

Lastly, an experiment was conducted (figure 4.4.2, black) to test the effect of time on the E-S ⇆ E-Comp equilibrium. This was performed in an analogous manner to the reactions outlined in figures 4.3.1-3, where competition was attempted from a premixed ES complex. However, the incubation of E-S and COMP was increased from 10 minutes to 20 minutes. This showed, on comparison to the analogous reaction in figure 4.3.1 (black), that there is a smaller amount of FAM labelled product formed. 1.5 nM FAM labelled product is formed at the endpoint of the control reaction compared to ~2 nM in figure 4.3.1, within error of each other. Such a small difference in values that are within error of each other implies that the E-S ⇆ E-Comp reaches equilibrium after 10 minutes.

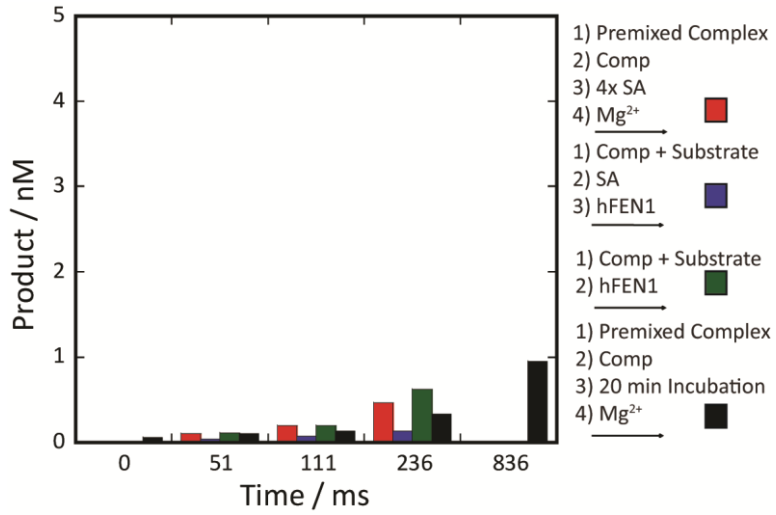


Figure 4.4.2: The effect of pre-addition of non-biotinylated unlabelled competitor substrate on the outcome of reactions of hFEN1-DF-3B ES complexes. Decay of a 'premixed' E-DF-3B complex, and incubated at 20°C for two minutes, after which competitor substrate (5 µM) was added followed by incubation at 37°C for 10 mins, and then 20 equivalents of SA were added prior to further incubation at 37°C for minute prior to reaction is shown in red. Decay of a SA 'blocked' ES complex, to which competitor (5 µM) is added prior to addition of hFEN1 is shown in blue. These sets of data lack an 836 ms timepoint due to a lack of unlabelled competitor substrate. Decay of a 'premixed' ES complex, to which competitor (5 µM) is shown in black. The normal 10 minute incubation is extended to 20 min prior to initiation in this case. All complexes are formed as shown in figure 3.6.2, and reactions are initiated by the addition of 8 mM Mg^{2+}. Reactions were carried out as described in section 2.8b, unless stated.

4.5 Discussion

The results presented in this chapter show that adding competitor substrate to 'trapped' ES complex causes negligible alteration to the concentration of FAM labelled product produced from reactions

catalysed by T5 FEN and hFEN1. However, concentrations of product are severely reduced when the initial complex is only “premixed” (i.e., does not have SA conjugated to the 5’ end). Thus, successful competition occurs from a ‘premixed’ complex (i.e. reversible complex formation) but is unsuccessful from a ‘trapped’ complex (i.e. irreversible complex formation). This strongly implicates the threading mechanism, (figure 4.1.1, path 3) as if the 5' ss portion of flaps are enclosed within the helical arch of FEN, there would be no possible way for the SA to come back through the helical arch, even when disordered, for the substrate to dissociate. However, in the clamping mechanism this would not be the case, as it should always be possible for the substrate to dissociate from FEN, even with SA on the 5' end. It is also worth noting that the results of experiments are identical for both T5 FEN and hFEN1 enzymes showing evolutionary conservation of this mechanism.

The results obtained with 5’-trapped substrates contrast markedly with those where the trapping procedure is carried out with a 3’-biotin moiety (figure 4.2.2). The competition reactions were performed with the 3’-biotinylated substrates to ensure that the mere presence of SA was not the cause of the effects seen above. As expected, the decay of a premixed T5 FEN-pY-21-3'B complex afforded levels of FAM labelled product similar to that of the single turnover profile of premixed T5 FEN-pY-21-3'B shown earlier in figure 3.6.4. However, competition experiments (figure 4.2.2), were all successful, exhibiting very low levels of FAM labelled product when challenging premixed, trapped and blocked ES species. This ruled out the possibility of fortuitous interactions between SA and FEN resulting in a tertiary complex that holds the substrate in place on the enzyme. Furthermore, the 22 PEG unit linker of pY-21-3'B is extremely long and flexible, giving the biotin-SA complex the ability to move freely, including into the region that would be occupied by the analogous biotin-SA complex on pY-21B. However, there was no evidence of 3’ trapped species that were resistant to competition.

Of relevance to the proposal that hFEN1 processes long and short flaps with different mechanisms, experiments with 5’-biotin substrates behaved identically, including those with a 3nt 5' flap. Clearly, the 3 nt flap of DF-3B was trapped in an ES complex that could decay with a biologically relevant rate of reaction. The formation of this complex is irreversible under the conditions tested as addition of unlabelled competitor oligonucleotide did not decrease amounts of FAM labelled product formed. This disproves the proposal that short flaps were not positioned within the helical arch, and implies that all FEN substrates are processed identically (Gloor *et al.,* 2010).

In this chapter, the ability to irreversibly trap a gapped substrate with SA is also shown. This ability to process substrates with duplex present on the 5′ flap rules out a model where DNA passes through a structured arch. However, this would be consistent with accommodation of substrates when the arch is disordered. Work performed by Dr. John Atack, rules out an alternative scenario where resolution of secondary structure of the duplex DNA within the flap is required to bind gapped substrates, as a substrate with an unresolvable cross-linked hairpin within the 5′ flap reacted with comparable efficiency to one with a single stranded flap, (47 +/- 6 min^{-1} cf. 137 +/- 11 min^{-1}, respectively (Patel *et al.* 2012)).

In conclusion, the results presented here; mostly the ability to compete labelled substrate out of a premixed ES complex, and the inability to compete from a 5′ trapped ES complex even when the substrate contains 5′-duplex, strongly implicate the 'bind and thread through a disordered arch' mechanism. As described earlier in section 4.1, these findings would be consistent with a mechanism that involves the arc of the template strand delivering the 5′ flap to a disordered helical arch. Subsequent ordering of the arch would correctly position components of the FEN active site, notably basic arch residues K93 and R100. As described in chapter 1.6 (figure 1.6.9), the downstream region of the substrate needs to unpair in order for the scissile phosphate to contact the active site; the ordering of the helical arch could be coupled to or be a prerequisite for unpairing to occur. An attractive feature of the bind and then thread through a disordered arch mechanism is that it also explains earlier data that appeared at odds with a threading mechanism. Thus, the ability to cleave flaps with small adducts such as cisplatin and branched flaps can be accounted for (Tsutakawa *et al.* 2011).

Chapter 5: Studying FEN-substrate interactions using fluorescence polarisation

5.1 Introduction

Many early studies of flap endonucleases focused on qualitative or quantitative analyses of FEN-substrate interactions using electrophoretic mobility shift assays (EMSA) (Sayers and Eckstein 1991; Murante *et al.* 1995; Barnes *et al.* 1996; Bornarth *et al.* 1999; Dervan *et al.* 2002; Gloor *et al.* 2010). Different potential substrate constructs were employed, sometimes in conjunction with mutated enzymes, to define the various regions of substrates that were recognized by FENs and to identify which parts of the protein interacted with various parts of the substrate. For example, Dervan *et al*. compared pseudo-Y (pY) structures with 5'-overhangs that lack any upstream DNA and examined interactions with mutants of T5 FEN. These studies deduced the correct orientation of pY substrates on the T5 FEN protein and defined the region and important residues of T5 FEN that interacted with duplex DNA (Dervan *et al.* 2002). A related study of a comprehensive set of mutants of hFEN1 and differing substrates later reached similar conclusions about the orientation of flap substrates on this FEN (Qiu *et al.* 2004), and identified the upstream and downstream binding sites. EMSA techniques were also instrumental in confirming the presence of a 3'-flap binding site in higher organism FENs (Finger *et al.* 2009; Tsutakawa *et al.* 2011). All of these studies tested complexes formed in the absence of cofactors (presence of EDTA), usually incubating enzyme and substrate and carrying out electrophoresis at 4°C.

One of the conclusions of these earlier studies was that FENs recognized and presumably had a binding site for 5'-single stranded (ss) flaps (Sayers and Eckstein 1991; Murante *et al.* 1995; Barnes *et al.* 1996; Bornarth *et al.* 1999; Dervan *et al.* 2002). Removal of the 5'-ss flap from substrates was widely reported to have dramatic effects on FEN-substrate affinity, an observation which in part led to the now discredited "tracking" mechanism (Sayers and Eckstein 1990; Rumbaugh *et al.* 1999; Garforth *et al.* 2001). Yet, later studies showed that exonucleolytic substrates devoid of 5'-flaps underwent a robust FEN reaction, implicating cofactor metal ions in enzyme substrate interactions. Indeed, Sengerova *et al.* recently showed that the exonucleolytic hydrolysis, but not endonucleolytic hydrolysis catalysed by T5 FEN are diffusion controlled (Sengerová 2009; Sengerova *et al.* 2010). It therefore could be argued that DNAs without 5'-flaps are better substrates for this enzyme. As detailed in chapter 1, modelling and structural studies suggest two conformations of substrate occur during the FEN catalytic cycle, one with substrate still fully base paired and another where two

nucleotides are unpaired and the scissile bond of the substrate contacts active site metals. Whilst the base paired form of the substrate is a likely pre-requisite for the unpaired state, even in the base paired form substrates bind close to the seven conserved carboxylates in the FEN active site, therefore, addition of divalent metal ions that bind in the active site could possibly alter affinity for DNA even without substrate unpairing (figure 1.6.9, repeated below).

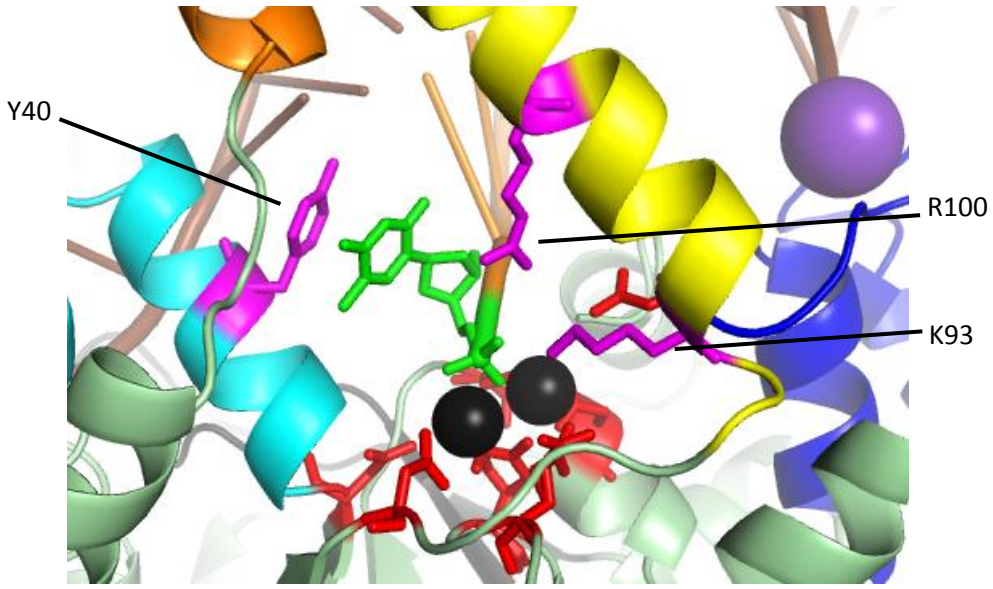


Figure 1.6.9: Stacking of the unpaired terminal nt of the downstream duplex in an hFEN1-product complex as viewed from behind the enzyme: The terminal nt of product (green), which has been brought into close proximity to the divalent metal ions (black) within the active site is stacked on the Y40 residue on the cyan helix, highlighted in magenta. Unpairing this nucleotide is thought to be encouraged by the other magenta residues on the helical arch (yellow), R100 and K93 (left to right).

A very recent publication on the subject of how FEN accommodates the 5'-portion of its substrate used EMSA studies to examine the affinity of substrates with differing length flaps and with streptavidin blocked flaps (Gloor *et al.* 2010). In line with the conclusions reached in Chapter 3, a 5'-biotinylated double flap substrate with a 64 nt 5'-flap was found to have a similar affinity for hFEN1 with and without streptavidin emphasizing the importance of interaction with the bifurcated junction and template DNA and not the 5'-flap. However, in the presence of EDTA, evidence for an increase in dissociation constant on reduction of the length or removal of the 5'-flap was interpreted as support for the hypothesis that long flaps thread through the helical arch, but short ones do not (Gloor *et al.* 2010). In chapter 3 of this thesis, it is shown that in the presence of EDTA, only small

amounts of hFEN1 5'-bioitin complexes are trapped in a fast reacting conformation. Thus, the interpretation of data obtained in the presence of EDTA to give information on the catalytically productive complex may be problematic.

Initially, to determine conditions for formation of the complexes described in chapters 3 and 4 and to rule out large perturbations from addition of biotin, we determined substrate dissociation constants. Later these experiments were extended to investigate the effects of flap length and also include mutant enzymes. Anisotropy measurements were used to monitor formation of complexes (Sevenich *et al.* 1998). Whereas gel shift (EMSA) assays often need to be performed under specific conditions, require buffer re-equilibration when metals ions are included and often involve use of radioactive substrates, conditions can be easily varied during fluorescence measurements. Furthermore, anisotropy allows the determination of binding constants under equilibrium conditions that cannot be mimicked during electrophoresis. Like the work described in Chapters 3 and 4, the properties of complexes of both T5 FEN and hFEN1 were studied.

5.2 Substrate binding to T5 FEN

Using a gel retardation assay, Dervan *et al*. 2001 showed that T5 FEN bound pseudo Y (pY) substrates with very high affinity, especially when compared to a similar 5'-overhang without upstream DNA, with K_D values at 4°C of ~5 nM cf. ~90 nM, respectively. The unimolecular pY substrates synthesised for this work were designed based on these findings, with a 6 nt 3' overhang to occupy the T5 FEN upstream DNA site and a 12-nt bp duplex capped by a 3-nt hairpin turn in the downstream binding region (Figure 3.2.1, Chapter 3). The length of 5'-flap was varied whilst maintaining the same duplex and 3'-overhang. The conditions of experiments were kept as close to those that give optimal rates of reaction as possible, whilst preventing catalysis occuring.

Fluorescence anisotropy experiments were performed with the range of substrates described in chapter 3.2 and T5 FEN to investigate the bacteriophage FEN enzyme. Firstly, binding curves were measured with differing flap lengths of substrates (pY-7 cf. pY-21) with T5 FEN in the presence of EDTA (figure 5.2.1, table 5.1). Dissociation constants were also determined with the different biotinylated substrates used in Chapters 3 and 4 to confirm that the modified substrates share similar binding affinities.

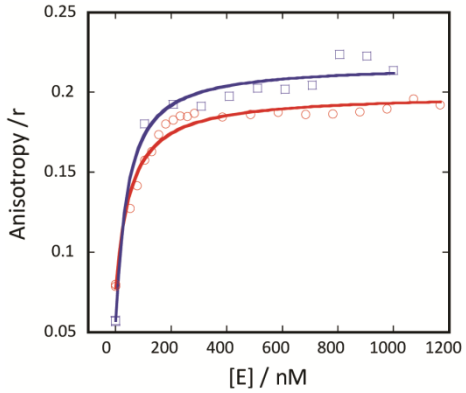


Figure 5.2.1: Example of a binding curve measuring the anisotropy changes on stepwise addition of T5 FEN to 10 nM pY-7 at 20°C. This was performed at 25 mM pH7.5 HEPES, 50 mM KCl and 0.01 mg/ml BSA, 1 mM DTT and 1 mM EDTA to maintain normal reaction conditions without initiating cleavage. The excitation wavelength was set to 490nm, and the emission at 510 nm was recorded. Slit widths were set to 10 nm, with an average of 10 scans taken per reading. Two sets of data were fitted to equation 3 (repeated below), and results are summarised in table 5.1. Aliquots were taken at the endpoint of each curve, and quenched in an equal volume of 8 M Urea and 80 mM EDTA. This was then analysed by dHPLC equipped with a fluorimeter to check for reaction

$$r = r_{min} + (r_{max} - r_{min}) \frac{([S] + [E] + K_D) - \sqrt{([S] + [E] + K_D)^2 - 4[S][E]}}{2[S]} \quad \textbf{(3)}$$

A 5'-flapless exonucleolytic (exo) substrate 3'-OH-6 (figure 5.2.2) (substrate with identical upstream and downstream regions but lacking a 5'flap) was also studied to test whether a 5' flap significantly alters the affinity of enzyme-substrate complexes. Previous results with a 5' FAM label on a different exo substrate indicated very tight binding (Dr. Blanka Sengerova, personal communication), and this could be due to the capture of FAM dye within the helical arch of FEN. Thus, an exo substrate which had a 3' FAM moiety was also tested, (3'-FAM-OH-6; figure 5.2.2). The results of these experiments can be seen in table 5.1.

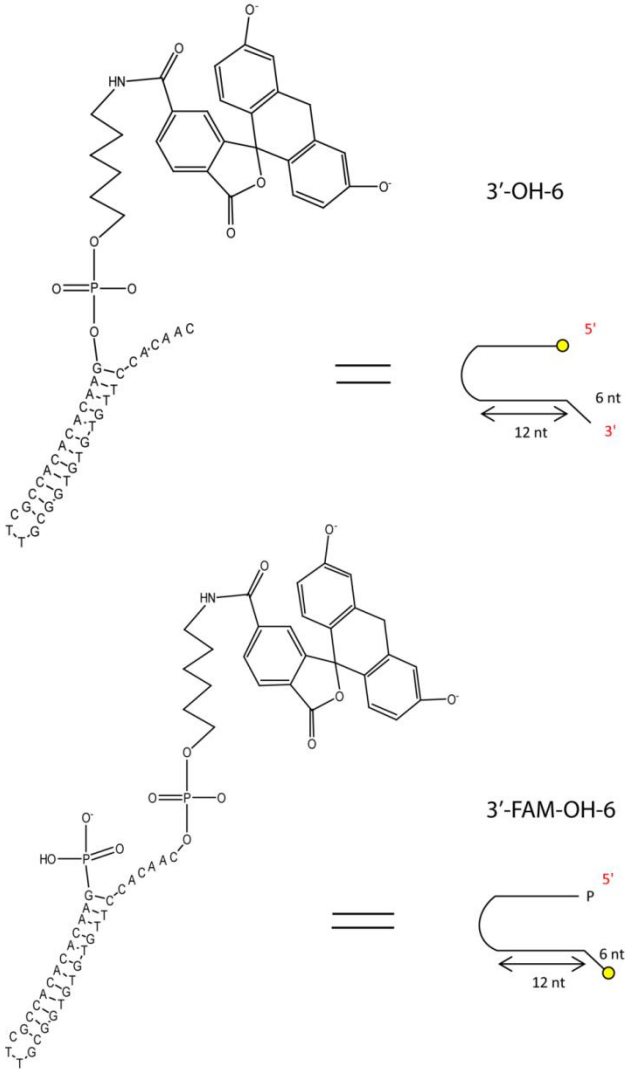


Figure 5.2.2: Structures of the 3' overhang substrates: 3'-OH-6, a unimolecular exo substrate without a 5' flap and a 6 nucleotide 3' overhang, which possesses a 5' FAM; **3'-FAM-OH-6**, a unimolecular exo substrate without a 5' flap and a 6 nucleotide 3' overhang, which possesses a 3' FAM and a 5' phosphate.

In addition, the dissociation constants of substrates were determined in the presence of 10 mM Ca^{2+} (figure 5.2.3). It was reported by Dr. Chris Tomlinson ((Tomlinson 2011), personal communication) that the presence of 10 mM Ca^{2+} would result in the full occupation of the active site of FEN enzymes. Although added divalent ions would generally be detrimental to a DNA-protein interaction, active site carboxylate bound ions neutralize repulsive charge and would presumably permit an unpaired substrate conformation to form.

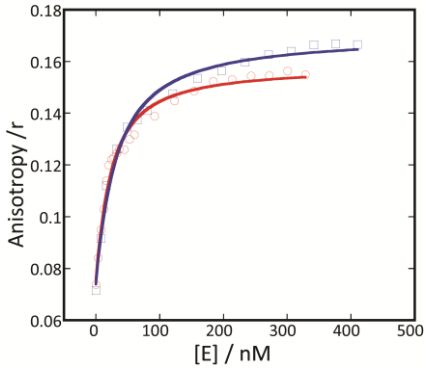


Figure 5.2.3: Example of a binding curve measuring the anisotropy changes on stepwise addition of T5 FEN to 10 nM pY-21 at 20°C in 10 mM Ca^{2+}. This was performed at 25 mM pH7.5 HEPES, 50 mM KCl, 1 mM DTT, 0.01 mg/ml BSA to maintain normal reaction conditions without initiating cleavage. The excitation wavelength was set to 490nm, the emission at 510 nm recorded. Slit widths were set to 10 nm, with an average of 10 scans taken per reading. Two sets of data were fitted to equation 3 to give K_D values in table 5.1. Aliquots were taken at the endpoint of each curve, and quenched in an equal volume of 8 M urea and 80 mM EDTA. This was then analysed dHPLC equipped with a fluorimeter to check for reaction.

After fluorescence anisotropy measurements, 100 µl aliquots were removed and quenched in 8 M urea and 80 mM EDTA. This was then analysed using a dHPLC equipped with a fluorimeter to check for reaction. Measurements made in the presence of EDTA never reacted; however, in the presence of divalent Ca^{2+} ions, it was revealed that ≤ 5% of the substrate present had reacted.

Substrate	**pY-21** 21 nt flap		**pY-21B** 21 nt flap 5'-biotin		**pY-21-3'B** 21 nt flap 3'-biotin
Enzyme	T5 FEN		T5 FEN		T5 FEN
FAM Position	5'		5'		5'
Buffer	EDTA	Ca	EDTA	Ca	EDTA
K_D **/ nM**	309 ± 30	23 ± 4	355 ± 56	60 ± 19	369 ± 28
r_{min}	0.04 ± 0.003	0.07 ± 0.0008	0.04 ± 0.004	0.06 ± 0.0003	0.05 ± 0.004
r_{max}	0.09 ± 0.003	0.16 ± 0.01	0.09 ± 0.003	0.12 ± 0.004	0.1 ± 0.01

Substrate	**pY-7** 7 nt flap	**3'-OH-6** Exo substrate No flap	**3'-FAM-OH-6** Exo substrate No flap 3' FAM moiety	
Enzyme	T5 FEN	T5 FEN	T5 FEN	
FAM Position	5'	5'	3'	
Buffer	EDTA	EDTA	EDTA	Ca
K_D **/ nM**	43 ± 6	6 ± 0.5	107 ± 8	2 ± 0.4
r_{min}	0.07 ± 0.01	0.07 ± 0.001	0.03 ± 0.001	0.04 ± 0.0003
r_{max}	0.2 ± 0.009	0.2 ± 0.0005	0.06 ± 0.0004	0.08 ± 0.001

Table 5.1: K_D parameters and standards errors for 10 nM pY-21, pY-21B, pY-21-3' B, pY-7, 3' -OH-6 and 3' -FAM-OH-6 bound to T5 FEN. All measurements were made in 25 mM pH 7.5 HEPES containing 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Ca^{2+} or 1 mM EDTA as indicated. All measurements were made at 20^{o}C. Accompanying graphs are found in the appendices as figure A1.

T5 FEN and flap lengths

The dissociation constant of the 7 nt flap substrate pY-7 from T5 FEN (table 5.1) was 43 nM, approximately 5 fold lower than that of the analogous 21 nucleotide pseudo Y substrate, pY-21 (figure 5.1. K_D = 309 nM). Both these substrates have 5'-fluorophores. Flap length is reflected in the both the r_{min} value of the free substrate in solution (0.07 for a small flap cf. 0.04 for a large flap), and the r_{max} of the enzyme saturated substrate (0.09 long cf. 0.2 short) implying a more restricted fluorophore on the smaller flap. The other 21 nt flap substrates were also tested in EDTA giving comparable K_D values of 355 nM and 369 nM for pY-21B and pY-21-3' B respectively. These results mirror the kinetic parameters shown in chapter 3 that suggest that biotin modification of the termini of substrates does not drastically alter the stability of FEN-substrate complexes.

For the exo substrates in the presence of 1 mM EDTA, the K_D values were 6 and 108 nM for 3'-OH-6 and 3'-FAM-OH-6 respectively. This shows tight binding of the 5' FAM substrate, 3'OH-6 (table 5.1), and binding with lower affinity to the 3' FAM subtrate, 3'-FAM-OH-6 (table 5.1). The 3' FAM substrate also gave a lowered K_D with respect to the long 21 nt pseudo Y substrate, pY-21. The exo substrate is bound with higher affinity than substrates with a longer more mobile flap, but with comparable affinity to the small flap substrate.

The extremely tight binding of the 3'-OH-6 substrate suggests the 5' FAM moiety interacts with or may be contained within helical arch of T5 FEN. For example, strong interactions between the negatively charged FAM moiety and the positively charged residues that normally position and direct substrate DNA are possible. This is also mirrored in the r_{max} of these substrates, 0.2 (5' FAM) cf. 0.09 (3'-FAM). This is also probably due to the position of 3' FAM relative to the protein, allowing greater rotational freedom and less shielding in general of the fluorescence emission, as shown in the comparable r_{max} values of the long 5' flap substrates, pY-21, 21B etc. The r_{min} values are varied, ranging from 0.03 with a 3' FAM to 0.07 with a 5' FAM. This is probably due to the 5' FAM moiety

being attached to duplex DNA and imparting some motional restriction, while the 3' FAM moiety is attached to a 6 nucleotide 3' overhang which is ssDNA and a more flexible linker.

Testing binding in the presence of divalent ions

As expected, much lower K_D values were obtained in the presence of Ca^{2+} versus EDTA buffers with the substrates pY-21, 21B and 3'-FAM-OH-6 (table 5.1) indicating tighter binding. The r_{min} and r_{max} values both rise in Ca^{2+} buffer as well. This upward trend in r_{min} values infers that the Ca^{2+} in solution restricts dye movements generally due to interaction with either the negatively charged DNA backbone or the fluorophore, which may also in part be responsible for increased r_{max} value as well. The lowered K_D is likely due to Ca^{2+} binding to the conserved carboxylates both lowering repulsion and possibly promoting unpairing of the duplex DNA within the active site. Unpairing could also lead to increased r_{max} due to restricted movement.

5.3 Substrate binding to hFEN1

Substrate binding to hFEN1 was also tested with substrates DF-3B, DF-5B and DF-21B (table 5.2, substrates shown in figures 3.2.2 Chapter 3) to investigate whether trends in substrate dissociation constants observed with the bacteriophage enzyme were also seen with hFEN1. There are some structural differences between the bacterial and human enzymes, such as the 3' flap binding site, which offers 10 extra specific contacts between the hFEN1 enzyme and substrate that is suspected to lower K_D values. The alpha helix 5 in T5 FEN (mentioned in section 3.7) that is not conserved in hFEN1 could also make a difference. The aromatic and positively charged residues that seem to encourage threading of flap substrates in T5 FEN in the absence of divalent metal ions are not present in hFEN1. Thus, it was of interest to see whether this would be reflected in the anisotropy measurements carried out in this section.

Substrate	**DF-3B** 3 nt flap 5'-biotin		**DF-5B** 5 nt flap 5'-biotin		**DF-21B** 21 nt flap 5'-biotin	
Enzyme	hFEN1		hFEN1		hFEN1	
FAM Position	5'		5'		3'	
Buffer	EDTA	Ca	EDTA	Ca	EDTA	Ca
K_D **/ nM**	18 ± 3	5 ± 2.8	44 ± 5	11 ± 2	117 ± 3	2 ± 0.09
r_{min}	0.03 ± 0.0013	0.04 ± 0.001	0.04 ± 0.0001	0.04 ± 0.001	0.027 ± 0.001	0.09 ± 0.001
r_{max}	0.1 ± 0.01	0.1 ± 0.002	0.11 ± 0.01	0.1 ± 0.002	0.15 ± 0.003	0.18 ± 0.01

Table 5.2: K_D parameters and standards errors for 10 nM DF-3B, DF-5B and DF-21B bound to hFEN1. All measurements were made in 50 mM pH 7.5 HEPES containing 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Ca^{2+} or 1 mM EDTA as indicated. All measurements were made at 37^{o}C. Accompanying graphs are found in the appendices as figure A2.

As seen in table 5.2, hFEN1 appears to follow the same trends as T5 FEN in that increasing flap length appears to lead to an in increase in K_D ; values of 18, 44, 117 nM were determined for a 3, 5 and 21 nucleotide 5' flaps respectively. For DF-3B and DF-5B that both have 5'-FAM and biotin moieties, r_{max} values are considerably smaller in comparison to T5 FEN and pY-7 (table 5.1), implying the dye has less conformational restriction in these complexes. This could in part be due to the biotin TEG linker that joins the oligomer to the FAM moiety.

The testing of substrates in the presence of 10 mM Ca^{2+} shows a similar trend as with the bacteriophage enzyme (table 5.1), with much lower K_D values exhibited than in the presence of 1 mM EDTA. All substrates exhibit a stoichiometric interaction when bound in the presence of divalent

metal ions. Although the values should be treated as approximations, the trend on addition of divalent ions is clear.

We also studied the dissociation constants of WT hFEN1, two active site mutants of hFEN1 and the gapped flap substrate that contains a 5'duplex, DFGENB (shown in figure 3.2.2 chapter 3). These mutants were D181A and K93A (figure 5.2.4, table 5.3). In D181A (figure 5.2.4), a conserved active site aspartic residue seen coordinating metal ion 2 within hFEN1 structures is mutated to an alanine. This removes a negatively charged carboxylate in the active site of this hFEN1 mutant. K93A, as discussed in chapter 1 and briefly in chapter 3, has a positionally conserved lysine residue at the base of alpha helix 4 mutated to alanine. Lysine 93 is thought to act as an electrostatic catalyst and to assist with positioning the substrate in an unpaired cleavage competent state (Sengerova *et al.* 2010; Tsutakawa *et al.* 2011). Removing this residue might raise the K_D value for binding substrates due to the absence of a potential positive charge within the active site.

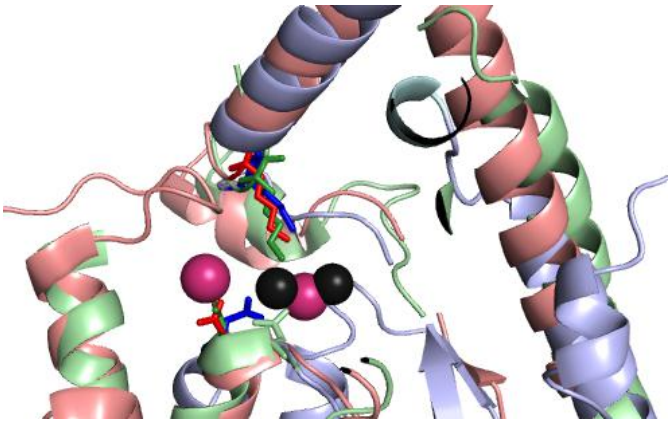


Figure 5.2.4: An overlay of T5 FEN (red), hFEN1 (blue) and T4 RNase HI (green) showing the positionally conserved residues D181 (at the bottom of the figure) and K93 used in this work. The metal ions of T5 FEN (magenta) and hFEN1 (black) are shown with the helical arch shown stretching from the top of the figure and down to the right.

A five nt double flap biotin substrate DF-5B was also tested with hFEN1 mutant Δ336, which is a mutant that lacks the positively charged extended C-terminus of the hFEN1 enzyme (figure 5.2.5).

This region of the protein is a site of protein-protein interaction and is the binding site for the sliding clamp PCNA, which coordinates the activities of hFEN1, polymerase and ligase during replication and repair (Burgers 2009; Finger *et al.* 2009; Zheng *et al.* 2011). On the basis of gel retardation assays, it has been claimed that removal of the C-terminal tails alters the ability to bind substrate. Thus, the dissociation constants of WT hFEN1 and Δ336 mutant were determined with a common substrate, as shown in table 5.3.

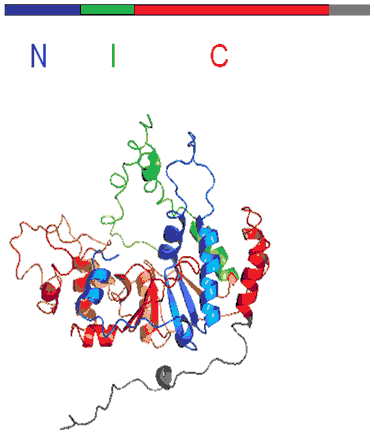


Figure 5.2.5 Model of hFEN1 with the relevant domain colour coding scheme. The C terminus tail, in grey is the portion removed in hFEN1 mutant Δ336.

Substrate	**DF-5B** 5 nt flap 5'-biotin		**DF-5B** 5 nt flap 5'-biotin		***DFGENB*** *Gapped flap 5'-biotin*		
Enzyme	hFEN1		Δ336-hFEN1		*hFEN1*	*K93A*	*D181A*
FAM Position	5'		5'		*5'*	*5'*	*5'*
Buffer	EDTA	Ca	EDTA	Ca	*EDTA*	*EDTA*	*EDTA*
K_D / nM	44 ± 5	11 ± 2	379 ± 59	27 ± 4	*137 ± 12*	*88 ± 16*	*17 ± 0.8*
r_{min}	0.04 ± 0.0001	0.04 ± 0.001	0.03 ± 0.002	0.04 ± 0.002	*N/A*		
r_{max}	0.11 ± 0.01	0.1 ± 0.002	0.08 ± 0.02	0.09 ± 0.01	*N/A*		

Table 5.3: K_D parameters and standards errors for 10 nM DF-5B and DFGENB bound to hFEN1, Δ336-hFEN1, K93A-hFEN1 and D181A-hFEN1. All measurements were made in 50 mM pH 7.5 HEPES containing 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Ca^{2+} or 1 mM EDTA as indicated. All measurements were made at 37°C. Accompanying graphs are found in the appendices as figure A2. *Measurements performed by Jack Exell.*

The work carried out by Jack Exell with the biotinylated gapped substrate DFGENB gave a K_D value of 137 nM, mimicking the trend that is seen in table 5.1 and 5.2 whereby an increase in K_D as the flap length increases is observed. The gapped substrate DFGENB (table 5.3) produced a K_D value higher than that of the 5 nucleotide flap, but comparable to the 21 nucleotide ss flap substrates. It should be noted that the gapped flap could potentially interact with hFEN1 with more than one binding mode; placing either a duplex involving template DNA or the gapped flap in the downstream binding site. D181A and K93A give K_D values of 17 and 88 nM, respectively (table 5.3), with the same gapped substrate. Both values for the D181A and K93A are lower when compared to the WT hFEN1 enzyme (137 nM). The K93A mutant is comparable at 88 nM, whereas the D181A mutant has a much tighter binding affinity for the gapped substrate (10-fold lower). This may be due to the loss of one conserved carboxylate within the active site, reducing the repulsive effect towards the DNA substrate.

In the presence of EDTA, Δ336 showed a K_D value 10-fold higher than that of WT hFEN1 for the 5 nt double flap substrate, which is consistent with the role of the C-terminal tail in enhancing affinity for substrate. Due to the unstructured and mobile nature of this tail, binding may be unspecific in nature. To measure the interaction between Δ336-hFEN1 and substrates in the presence of metal ions, the K_D was measured in the presence of 10 mM Ca^{2+}. This revealed a much tighter binding with a K_D of 27 nM, compared with a value of 11 nM for the full length enzyme. This follows the general trend with metal ions throughout this chapter where in the presence of metal ions, substrate binding with Δ336 is not largely impacted by the loss of the positively charged C terminus tail. This suggests that the effect of negating the repulsive effect of the carboxylates within the active site is of far greater benefit than the C terminal tail of WT hFEN1.

5.4 Discussion

FENs and flap lengths

In the absence of metal ions, K_D values for varying flap length substrates universally adopted a similar trend. The values rise with flap length, implying a general trend in substrate binding whereby flap length is a determining factor in the difficulty of binding a substrate. In the case of T5 FEN, where streptavidin "trapping" data presented in chapter 3 suggests that fast reacting complexes can be formed in EDTA, the increase in K_D with flap length may be due to the increased flexibility of a longer ssDNA flap substrate making it harder catch the flap in the preferred position. A notable and universal observation was the addition of divalent calcium ions stimulated complex formation. Human FEN1 follows the same general trends as T5 FEN, with flap length inversely proportional to the magnitude of the K_D values, but the K_D values are much smaller in general to T5 FEN, which can be attributed to the 3' flap binding pocket which lends 10 extra interactions to a one nucleotide 3' flap on double flap constructs (figure 5.4.1) (Tsutakawa *et al.* 2011).

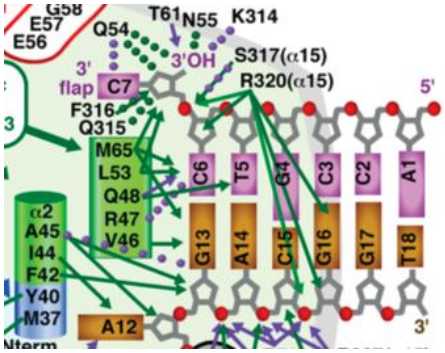


Figure 5.4.1 Taken from (Tsutakawa *et al.* 2011), showing the interactions formed between hFEN1 and the 3' flap (nucleotide C7) of the substrate shown. These interactions originate from residues L53, Q54, N55, T61, M65, K314, Q315, F316, S317 and R320, residues from all over FEN1s structure that come together on folding to form parts of the 3' flap binding pocket and hydrophobic wedge.

The trends observed in this work where decreasing the length of 5'-flap increases the stability of ES complexes is in opposition to the trend reported recently by Gloor *et al*. for hFEN1. Gloor *et al*. reported comparative EMSA data that suggested that flaps with less than 6 nt were bound more weakly, leading to the conclusion they were not threaded in the helical arch. Both sets of experiments did assemble complexes at 37°C. However, the method of analysis (EMSA versus anisotropy) and the conditions of the experiments were different. Notably, the EMSA experiments involved formation of complexes at relatively low ionic strength (30 mM NaCl (EMSA) versus 100 mM KCl (our studies)), which is of additional relevance due to the observation of a potassium ion binding site in hFEN1 structures. The only K_D values reported for hFEN1 in which a titration in the presence of EDTA at 37°C has been examined by EMSA with subsequent fitting of the binding isotherm are for a biotinylated 64 nt double flap substrate, and are 0.96 ± 11 nM and 0.96 ± 0.7 nM with streptavidin. These are quite different to the values observed by us (table 5.2) presumably reflecting these differences in conditions. Additional variation in the lengths of upstream and downstream regions of substrate may also be a factor.

The earlier observation that flapless exo substrates are bound more weakly than those with flaps does not seem to be true when fluorescence polarisation is used to monitor interactions and the substrate has upstream DNA. Using T5 FEN, the measured dissociation constant in EDTA of the 5'-phosphate exo substrate (3'-FAM-OH-6) was twice that of a 7 nt flap, but three times lower than that of a 21 nt flap. However, the pY flap substrates all had a 5'-dye label, whereas for the exo substrate

FAM was conjugated to the 3'. Further experiments using 3' fluorophores would be needed to confirm this result. It is also of interest to note that some anisotropy derived data with hairpin substrates with and without flaps, but both lacking upstream DNA did show large differences in affinity in EDTA, but little difference in calcium ions (Tomlinson 2011). This illustrates that results can alter based on the basic construction of substrates and this factor may be responsible for many apparent discrepancies in earlier work.

Perhaps most importantly, the results in this chapter validate the construction of the ES complexes displayed in both chapter 3 and chapter 4 (figure 3.6.2). For hFEN1, at 37°C in the presence of the inhibitory Ca^{2+} ions, K_D values never exceed 20 nM, exhibiting stoichiometric binding. These results indicate that under the conditions used for hFEN1 in chapter 3 and 4, (20°C, 500 nM E, 5 nM S, 2 mM Ca^{2+}, identical buffer salt and monovalent salt conditions), S should always be completely saturated by E. Interestingly, using EDTA hFEN1 complexes are less stable, but this seems unlikely to account for the failure to "trap" a large proportion of them in fast reacting conformation with streptavidin in the cases of DF-3B and DF-5B (respective K_Ds 18 nM and 44 nM at 37°C), but may have been a contributing factor for DF-21B and DFGENB (respective K_D*s* 117 nM and 137 nM at 37°C). For T5 FEN, the case is similar. The conditions for T5 FEN – S complex assembly differ from hFEN1 (4°C, 500 nM E, 5 nM S, 1 mM EDTA, identical buffer salt and monovalent salt conditions), and the K_D values at 20°C in 1 mM EDTA are approximately 300 nM for 10 nM solutions of 21 nt pY flap substrates and ≤100 nM for everything else. Thus, the use of low temperatures was appropriate for assembly of complexes involving T5 as described in chapters 3 and 4 (figure 3.6.2).

Chapter 6: Studies of the effects of Helical Arch structure using Site Directed Mutagenesis

6.1 Introduction

An assumption of earlier chapters of this thesis and much FEN literature is that the helical arch must order for efficient FEN catalysis. Recent X-ray structures of hFEN1 and hEXO1 in complex with products show arch residues (K93 and R100, hFEN1 numbering) from the base of α4 in contact with the 5'-phosphate monoester product, which is also coordinated to the two active site metal ions. Thus, roles for these residues in chemical catalysis and capture of the scissile phosphate diester on active site metals have been suggested (Sengerova *et al.* 2010; Tsutakawa *et al.* 2011). Evidence for these proposed roles comes from site directed mutagenesis of these residues to alanine, which is severely detrimental to catalysis and implies that this part of the arch at least is likely to be required in structured form (Finger *et al.* 2009; Sengerova *et al.* 2010; Tsutakawa *et al.* 2011) (figure 3.4.1). However, confinement in crystals and crystal packing effects can perturb structure. Additionally, the upper part of α4 and α5 do not appear to be conserved by 5'-nuclease superfamily members XPG and GEN1 implying the intact arch is not a feature of all related enzymes. Furthermore, assuming the threaded state is the catalytically proficient form of FENs as suggested by earlier chapters, the question arises as to how FENs can catalyse the slow reactions of substrates that are blocked. We, therefore, sought to investigate whether an arch that is able to order correctly was required for catalysis.

To disrupt the structure of the helical arch in the active site of hFEN1, site directed mutagenesis was used to insert proline residues into this region of the protein. Some amino acid residues have a higher propensity to fold into alpha helices or beta sheets, which is normally determined by their polarity, size or the availability of their N atom H bond donator and C atom H bond acceptor. In proline, both H bond donor and acceptor are occupied in an intramolecular ring, making itself a unique secondary amide. The rigid structure of proline, that does not lend itself to H bonds, imparts a kink in polypeptide chains of alpha helices which, is approximately 45-50°. This can be accommodated comfortably within the first four residues of an alpha helix (N → C terminus), but can potentially break this secondary structure when proline is further along the helix (Sauer *et al.* 1992).

Previous work by Storici *et al.* highlighted residues that may have some structural significance within the helical arch of hFEN1. Random PCR mutagenesis of the hFEN1 ORF was conducted and then inserted into a yeast expression plasmid under the control of a galactose promoter. Yeast were transformed and selected for on media lacking galactose. The resulting 12000 random point FEN1 mutants were tested for their ability to grow on galactose media, of which 100 mutants failed to grow. These mutants were crossed with different FEN1 *S. cerevisiae* strains; *RAD27*, *Δrad27* (FEN1 knockout) mutant and *rad27-p* (a rad27 mutant lacking its PCNA binding motif). In addition to isolating active site mutants, this screen isolated 3 dominant negative mutants in which three leucines (two in a4 and one in a5) were each mutated to a proline residue. The mutants were L97P, L111P and L130P. L97P was the least viable mutant, not able to grow in any of the clones except a diploid *RAD27*/*Δrad27* clone. L111P and L130P were viable, but these strains grew poorly, apart from the rad27-p and Δrad27 clones. Sequence alignments of FEN1 enzymes reveal that these leucine residues are loosely conserved other eukaryotic organisms (figure 6.1.1).

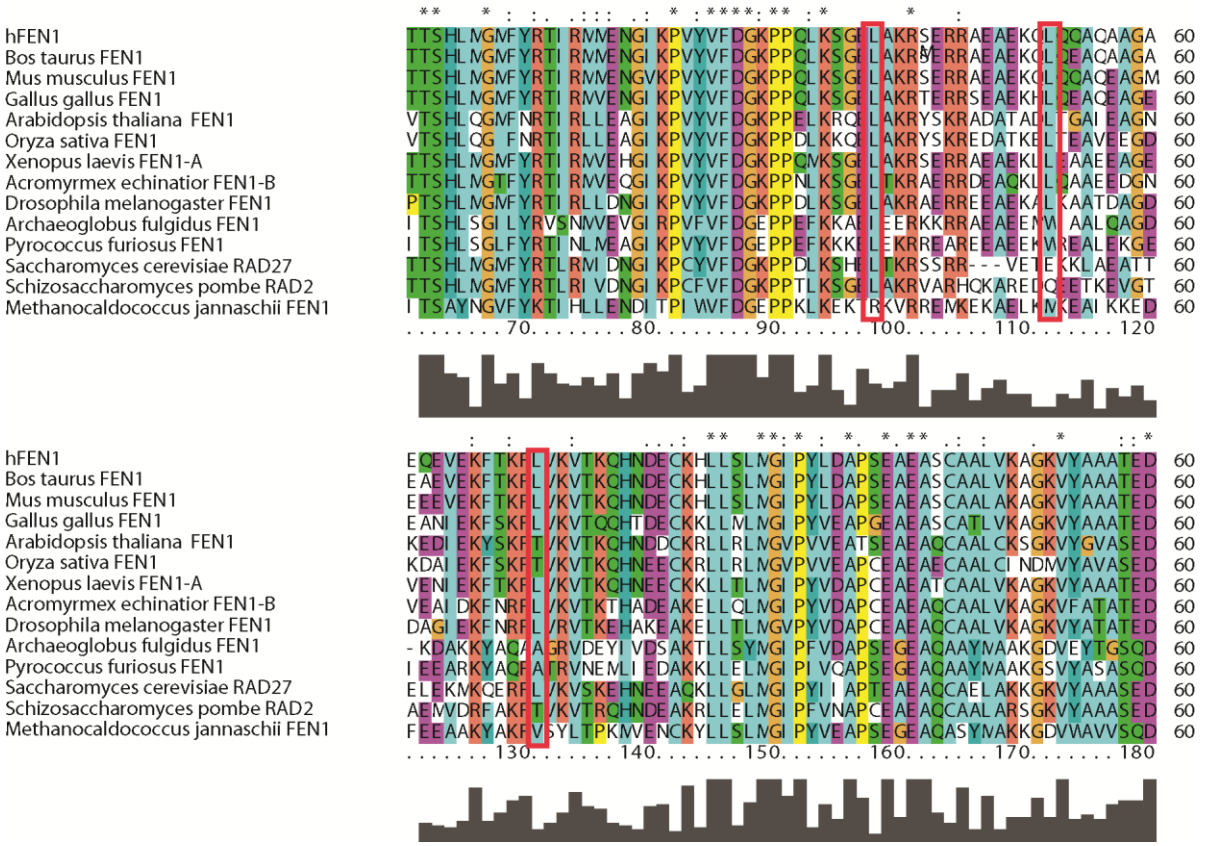


Figure 6.1.1: Sequence alignment of residues 60-180 (hFEN1 numbering) encompassing the helical arch, within various eukaryotic FEN1 enzymes. Residues L97, L111 and L130 are highlighted (red box). These show some conservation over eukaryotic FEN1 enzymes.

Proline, as discussed earlier, is an alpha helix breaker due to an angular constriction, and in conjunction with the results in chapter 4 and the recent work in Tsutakawa *et al.*, gave precedence to fully investigate the impact of these mutants on hFEN1 activity. Using plasmid mutagenesis as described in chapter 2, primers were designed using the Agilent site-directed mutagenesis primer design tool (www.genomics-agilent.com), and purchased commercially (www.invitrogen.com). Using WT hFEN1 plasmids that were isolated using a Qiagen miniprep kit and mutagenesis primer pairs, the plasmids were PCR amplified using a high-fidelity thermophillic polymerase (Pfu). WT DNA was degraded by addition of DpnI. The mixture was then transformed into DH5α *E. coli* for nick repair and plasmid preparation (figure 6.1.2).

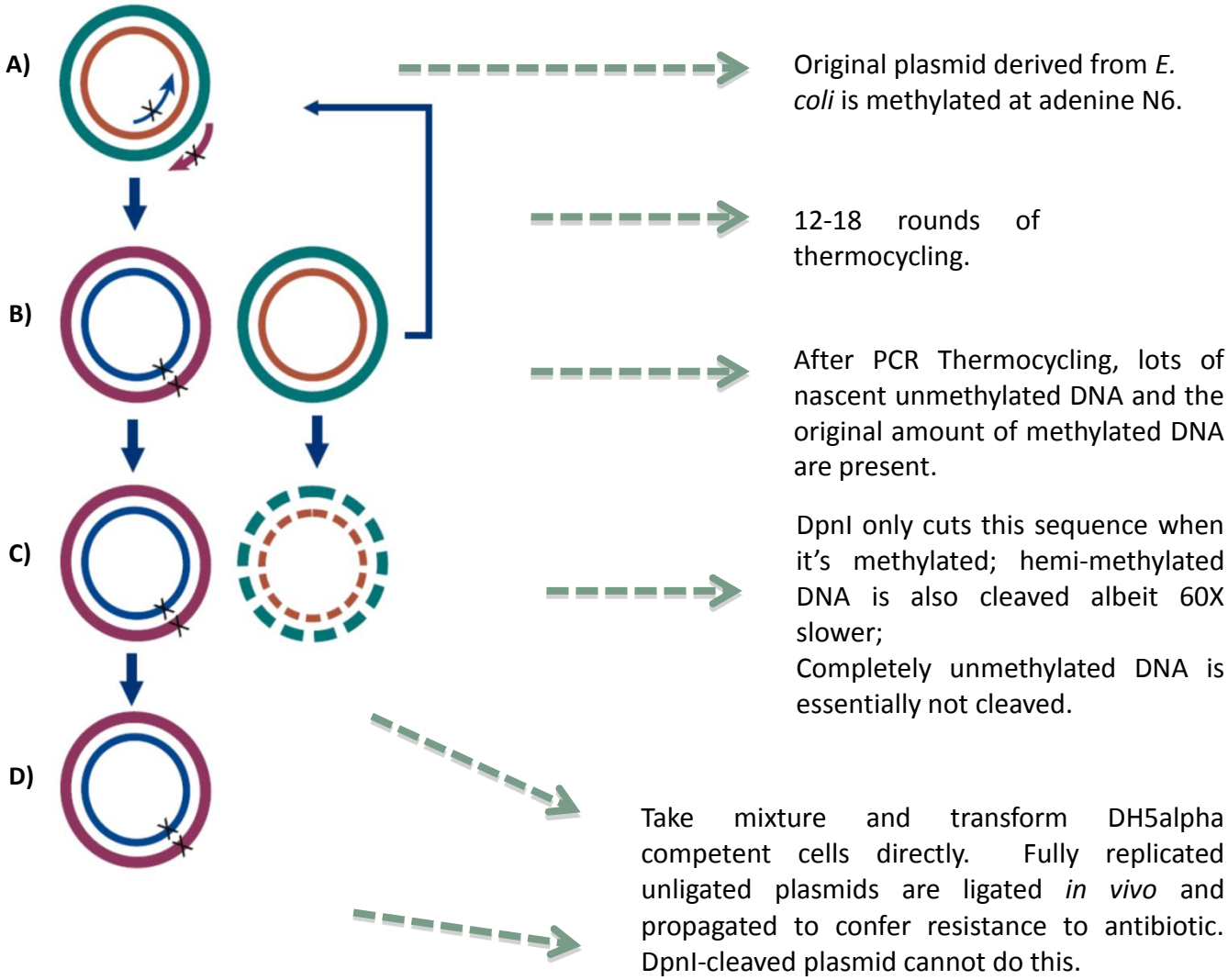


Figure 6.1.2: Stages of PCR mutagenesis (adapted from www.genomics.agilent.com): A) The original methylated plasmid is isolated from *E. coli* and primers designed to introduce the proline mutation at the desired location, which are designed to bind stably to the plasmid despite the small difference in sequence. **B)** After heating the PCR mixture to denature the DNA and to remove the inhibitory antibody of the DNA polymerase (i.e., hot start polymerase), 12-18 rounds of thermo-cycling take place to amplify the desired mutant plasmid. **C)** Dpn1 is used to cleave methylated plasmid DNA originally from *E. coli,* while leaving the mutant plasmid intact due to the absence of N-6 methylated adenines. **D)** Mutant plasmid is transformed into the appropriate strain of *E. coli* for nick repair and plasmid production. Once isolated and sequenced to confirm that the mutation is present, the plasmids are transformed into an E. coli protein expression strain for protein production.

The mutants were then overexpressed *E. coli* (e.g., BL21 (DE3)) and purified. The locations of the mutations introduced are shown below, in figure 6.1.3.

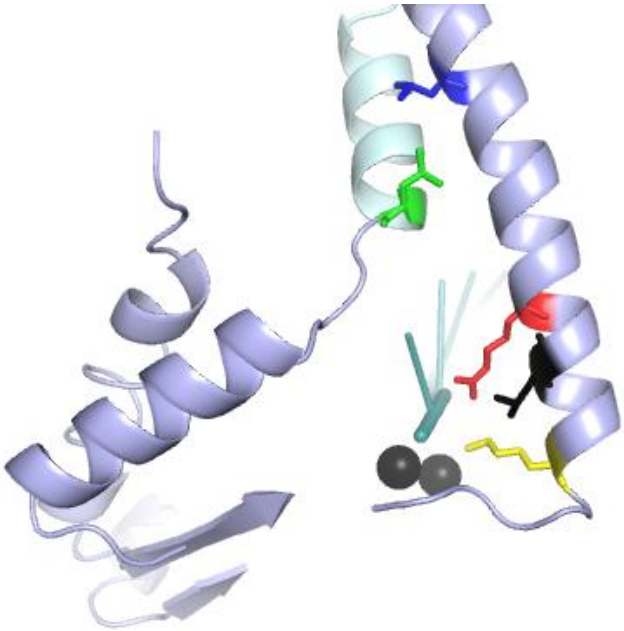


Figure 6.1.3: Alpha helix 4 of hFEN1 (right hand helix), with the positions of K93 (yellow) and R100 (red) highlighted, and the position of L97 shown (black), sandwiched between the two aforementioned residues. L111 (blue) is located higher up on α4, and L130 (green) is located on α5.

As seen in figure 6.1.3, these mutations are in alpha helix 4 and 5 of the hFEN1 structure. The L97P mutation sandwiched between conserved arch residues K93 and R100, which are postulated to be important in DNA unpairing and catalysis, is in the second turn of alpha helix 4. The L111P mutation is between the 5th and 6th turns of alpha helix 4. Finally, the L130P mutation is in the 3rd turn of alpha helix 5, which is suggested by data shown in earlier chapters to act as a cap over the flap portion of bound DNA substrates, stabilising the helical arch (Tsutakawa *et al.* 2011).

As well as the kink introduced by the proline mutants, and the disruption of the H-bonding network of the alpha helices, the dynamics of the alpha helices involved will be altered. The decreased flexibility about these proline residues would affect the ordering of the helical arch, which, as shown in chapter 4, orders around the ssDNA flap portion of substrates to assemble conserved active site

residues in the active site for efficient catalysis. The effect on the dynamics purely by having the proline residue within alpha helix 5 is better shown in previous work by Lovell *et al*. It was shown by Ramachandran plots that the flexibility of the phi angle of proline was restricted to -62°, a sharp contrast to the general angle displayed by the other residues within their alpha helix (figure 6.1.4). Furthermore, on looking at residues preceding proline, it is seen that rotational movement is hindered somewhat. When considering the effect on the dynamics of all three of the mutants in this chapter, a lack of rotational movement by pre-proline and proline residues is likely to hinder the proper ordering of the helical arch as well (Lovell *et al.* 2003).

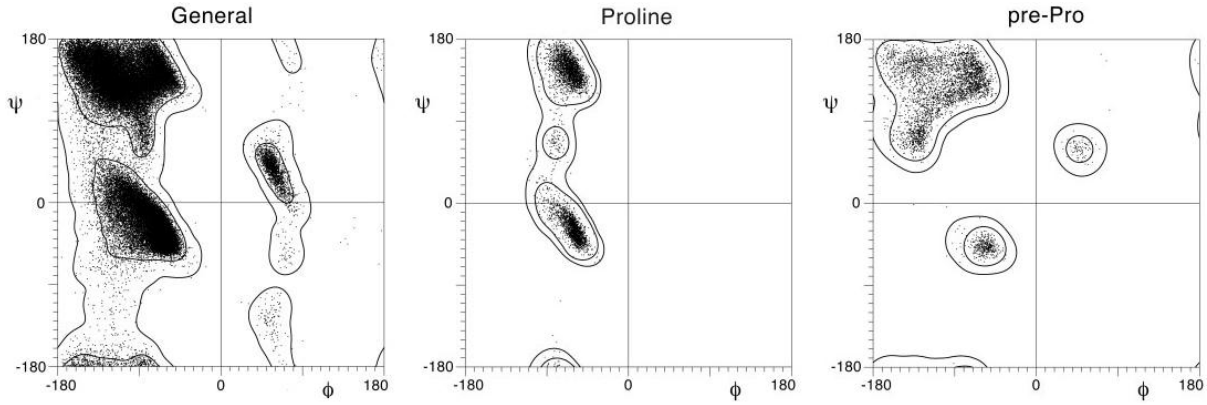


Figure 6.1.4: Ramachandran plots of general residues within an alpha helical structure (left), proline residues within the same alpha helical structure (middle), and pre proline residues within this structure (right). The phi angle of the proline residue is largely restricted to approximately -60-80° while the population of pre pro residues in comparison to the general residues is generally more favourable towards the -130°region for ϕ, and 80° region for ψ angles.

On observation of the positions of these mutations, it can be rationalised that the L97P mutation may have the most deleterious effect on hFEN1 activity due to the prevention of ordering and hindrance of movement for the N-terminal half of alpha helix 4 , where the R100 and K93 residues that are positionally conserved across the FEN superfamily reside (Finger *et al.* 2009; Sengerova *et al.* 2010; Tsutakawa *et al.* 2011). L111P will have a deleterious effect as well, but as it is further away from the N-terminal half of alpha helix 4, one would expect the effect of this mutation to be less severe than the L97P as would be predicted from the severity of the dominant negative phenotype. L130P will perturb the dynamics and structure of alpha helix 5, which is part of the helical cap (Tsutakawa *et al.* 2011).

If as suggested in previous work, the binding of substrate triggers the ordering of the helical arch, these mutations may simply stop these regions from ordering into alpha helices, preventing the optimal configuration of all active site residues of hFEN1. This could impart a similar hindrance upon the catalysis of reaction as the blocked ES reactions reported in chapter 3, due to the fact that the large bulk of streptavidin could also prevent proper ordering of the helical arch region of hFEN1.

This chapter aims to characterise these hFEN1 proline mutants to quantify the magnitude of effects on hFEN1 catalysed reaction, thereby elucidating the importance of a structured helical arch for hFEN1 catalysis.

6.2 Mutant Stability

Because of the potential structure altering nature of proline mutations, initial experiments interrogated the stability of the individual mutated proteins. Thermal denaturation carried out by Ben Ombler within the group showed no significant variation in the T_m of the enzymes with respect to WT hFEN1 (personal communication, Ben Ombler). However, despite this similarity in stability, multiple turnover kinetics (figure 6.2.1) of the L97P mutant never reached completion, exhibiting an early product plateau despite running reactions for over 24 hours (not shown). This suggested some enzyme inactivation occurs during these reactions, especially on a longer timescale. It should be noted that this early product plateau when present, was variable and it cannot be ruled out that enzyme inactivation was due to such long reaction time courses. This means that the characterisation of these mutant enzymes through multiple turnover kinetics would be unreliable. Despite this, it was clear that these mutants are severely retarded in comparison to WT hFEN1 (figure 6.1.2, table 6.2).

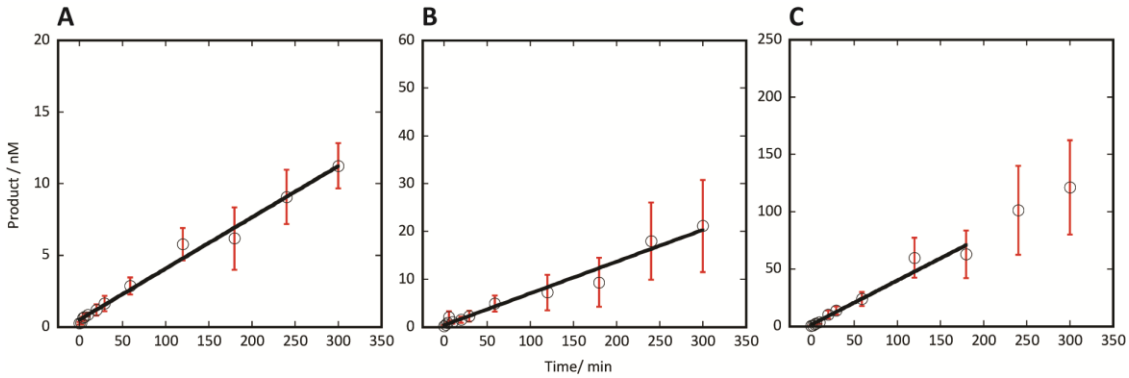


Figure 6.2.1: Multiple turnover graphs of 500 nM DF-5B with A, 25 nM L97P; B, 12.5 nM L111P; and C, 12.5 nM L130P. Each plot with error bars shown was performed in triplicate, and the initial rate plotted. Reactions were performed in 50 mM HEPES pH 7.5, 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 mM Mg^{2+}. L97P gave a rate of 0.00144 +/- 0.00005 min^{-1}, L111P gave a rate of 0.0066 +/- 0.0004 min^{-1} and L130P gave a rate of 0.039 +/- 0.003 min^{-1}.

6.3 Kinetic Analyses

The apparent enzyme inactivation of these mutants under multiple turnover conditions indicated that the activity of these mutants had to be characterised using single turnover kinetics. As the determination of a K_M value for these mutants was impossible, the K_D was measured (table 6.1) using fluorescence anisotropy, shown in detail within chapter 3 and 5. This was performed using "reaction-like" conditions for WT hFEN1 (50 mM HEPES pH 7.5 containing 100mM KCl, 0.01mg/ml BSA, but using 10 mM Ca^{2+} ions to replace the viable Mg2+ ion cofactor). Both long (21 nt) and short (5 nt) 5'-flap double flap substrates were used.

Substrate	**DF-5B** 5 nt flap 5'-biotin		**DF-5B** 5 nt flap 5'-biotin		**DF-5B** 5 nt flap 5'-biotin		**DF-5B** 5 nt flap 5'-biotin	
Enzyme	L97P		L111P		L130P		hFEN1	
FAM Position	5'		5'		5'		5'	
Buffer	EDTA	Ca	EDTA	Ca	EDTA	Ca	EDTA	Ca
K_D **/ nM**	14 ± 8.1	23 ± 5.7	11 ± 1.2	18 ± 7.3	11 ± 0.1	22 ± 0. 9	44 ± 5	11 ± 2
r_{min}	0.03 ± 0.00019	0.05 ± 0.001	0.03 ± 0.0003	0.06 ± 0.012	0.044 ± 0.001	0.06 ± 0.005	0.04 ± 0.0001	0.04 ± 0.001
r_{max}	0.1 ± 0.01	0.13 ± 0.0025	0.12 ± 0.001	0.14 ± 0.011	0.11 ± 0.003	0.13 ± 0.002	0.11 ± 0.01	0.1 ± 0.002

Substrate	**DF-21B** 21 nt flap 5'-biotin		**DF-21B** 21 nt flap 5'-biotin		**DF-21B** 21 nt flap 5'-biotin		**DF-21B** 21 nt flap 5'-biotin	
Enzyme	L97P		L111P		L130P		hFEN1	
FAM Position	3'		3'		3'		3'	
Buffer	EDTA	Ca	EDTA	Ca	EDTA	Ca	EDTA	Ca
K_D **/ nM**	60 ± 1.5	7 ± 1.0	70 ± 5.7	21 ± 4.1	88 ± 4.3	16 ± 0.4	117 ± 3	2 ± 0.09
r_{min}	0.1 ± 0.0005	0.1 ± 0.0005	0.03 ± 0.002	0.08 ± 0.002	0.04 ± 0.001	0.08 ± 0.001	0.027 ± 0.001	0.09 ± 0.001
r_{max}	0.21 ± 0.004	0.21 ± 0.003	0.19 ± 0.006	0.18 ± 0.00002	0.19 ± 0.004	0.18 ± 0.001	0.15 ± 0.003	0.18 ± 0.01

Table 6.1: K_D parameters and standards errors for 10 nM DF-5B and DF-21B bound to L97P, L111P, L130P and WT hFEN1. All measurements were made in 25 mM pH 7.5 HEPES containing 50 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 10 mM Ca^{2+} or 1 mM EDTA as indicated. All measurements were made at 37^{o}C. Accompanying graphs are found in the appendices as figure A3.

The dissociation constants determined in the presence of Ca^{2+} ions by fluorescence anisotropy were ~20 nM using DF-5B as substrate (table 6.1). In the presence of EDTA, dissociation constants were two-fold lower, but within error of the values in the presence of Ca^{2+}. On comparison of these values with WT hFEN1 little difference is seen, indicating that these mutants do not affect enzyme substrate affinity. The dissociation constants of the 21 nt flap were also measured, because the figures for this substrate with WT hFEN1 provided a larger range of K_D values in these different environments (K_D values for a 21nt 5′ flap in the presence of EDTA are approximately 100X greater than in the presence of Ca^{2+}, while the same values for a 5nt 5′ flap are only around 5X greater (table 6.1)). The K_D measurements with DF-21B bound to these proline mutants were comparable to WT hFEN1, (in 1 mM EDTA: ~ 100 nM; in 10 mM Ca^{2+} ~ 10 nM), although a subtle but reproducible trend emerged where dissociation constants in the presence of EDTA were approximately 2 fold lower than the WT enzymes. In contrast, in the presence of Ca^{2+} dissociation constants were 2 fold higher when going from the proline mutants to wild type hFEN1. This shows that mutations do not severely alter the ability of hFEN1 to bind substrates, implying that the helical arch does not contribute to the initial binding of substrates by hFEN1, thereby reinforcing the results shown throughout chapter 3 and 4. Due to the inability to measure binding in the presence of Mg^{2+}ions, the concentrations of enzyme used for single turnover kinetics (figure 6.1.3 A-C) were set at 45x K_D and substrate concentrations were set at 30x < K_D in the presence of EDTA.

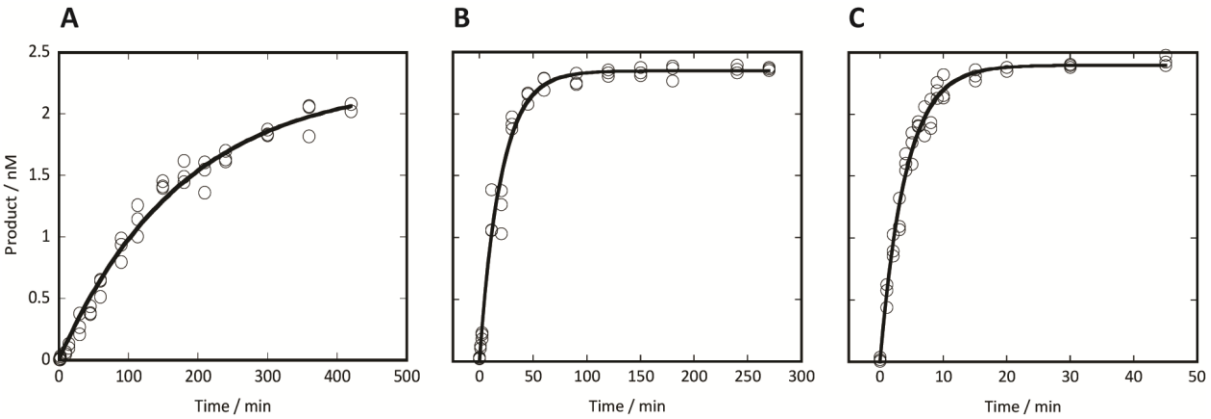


Figure 6.3.1: The unmixed (black) single turnover profiles of 2.5 nM DF-5B with A, 900 nM L97P; B, 900 nM L111P; and C, 900 nM L130P. Each plot was performed in triplicate and plotted on the same graph. Reactions were performed in 50 mM HEPES pH 7.5, 100 mM KCl, 0.01 mg/ml BSA, 1 mM DTT and 8 mM Mg^{2+}.

As shown in table 6.2 below, all the proline mutations dramatically inhibit hFEN1 reaction. The mutation L97P is most detrimental to activity and L130P the least. In comparison to WT hFEN1, the single turnover rates of reaction are approximately 90000-, 12 000- and 2 000-fold lower for L97P, L111P and L130P, respectively. As mentioned previously, the multiple turnover numbers show even larger drops in rate cf. to the single turnover numbers presumably due to enzyme inactivation during the course of the reaction.

	Multiple Turnover (min^{-1})	**Single Turnover (min^{-1})**	**$t_{1/2}$ (s)**
L97P	0.0014 +/- 0.00005	0.0056 +/- 0.00037	7426
L111P	0.0066 +/- 0.0004	0.05 +/- 0.0025	831
L130P	0.039 +/- 0.003	0.26 +/- 0.013	160
WT	*451 +/- 32*	*551 +/- 40*	*0.08*

Table 6.2: Single and multiple turnover rates calculated for the proline mutants investigated in this chapter. $t_{1/2}$ was calculated using the equation, $t_{1/2}$ = ln(2) / k_{ST}. *Measurements performed by Dr. John Atack.*

6.4 Discussion

The three proline mutants produced, purified and partially characterised in this chapter are extremely detrimental to the activity of WT hFEN1. Comparison of single turnover rates of the mutant proteins to WT hFEN1 show rates 90 000-, 12 000- and 2 000-fold lower for mutants L97P, L111P and L130P, respectively (figure 6.3.1). These reductions in rate are more severe (or comparable in the case of L130P) than even the blocked substrates analysed in chapter 3, suggesting an extremely deficient phenotype. These results show that the structure and the correct ordering of the helical arch are extremely important in catalysis.

L130P reacts at a rate comparable to blocked substrates. This is somewhat surprising as the proline residue is not within close enough proximity to affect any residues within the arch that may be critical for efficient catalysis (e.g. K93 and R100 (hFEN1 numbering)). There is a possible effect on the concerted motions of the helical arch, as exemplified in the ramachandran plot in figure 6.1.4, which may indicate a situation where the mutant protein would struggle to adopt a helical arch conformation in a similar position to the WT protein. However, this cannot be inferred from the data presented here, and more investigation would be required.

It is also possible the insertion of the proline residue within a5 reduces the likelihood of forming a well ordered helical arch region, altering the time spent in a disordered or ordered form. If a disordered helical arch is necessary for free 5' terminus recognition and eventually product release, and an ordered helical arch is necessary for key catalytic residue placement and possibly double nucleotide unpairing of substrates, it is likely that time not spent in either form would hinder the reaction catalysed by hFEN1. A similar phenomenon may be reflected by the cleavage of blocked substrates; the presence of SA making it impossible for the helical arch to order around a 5' flap. Testing L130P with streptavidin conjugated (blocked) substrates would be informative. A rate of cleavage of blocked substrates with mutant enzyme L130P that was comparable to WT hFEN1 might reinforce the possibility of hFEN1 having difficulty in ordering the helical arch. Furthermore, if the a4 helix can partially order in blocked substrates and L130P, the catalysis could still take place.

Mutations L97P and L111P produce a larger reduction in rate. They appear to have the most potential to disrupt placement of K93 and R100, which are important for catalysis and potentially DNA unpairing. This unpairing is necessary for cleavage to take place, as it brings the scissile phosphate in proximity of the metals for catalysis. As seen in chapter 1, figure 1.6.9 and below in figure 6.4.1, K93 and R100 are also instrumental in positioning the unpaired duplex within the active site. Therefore L111P, which affects the correct positioning of K93 and R100, and especially L97P, that makes direct contacts with the same residues in question, should be, and as seen in this chapter, are very deficient in catalysis.

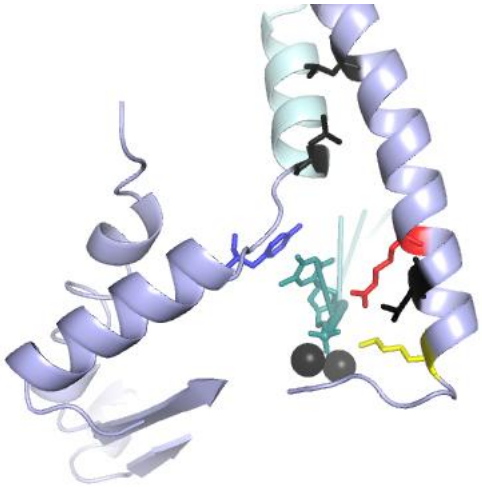


Figure 6.4.1: Stacking of the unpaired terminal nt of the downstream duplex in an hFEN1-product complex as viewed from behind the enzyme: The terminal nt (dark blue), which has been brought into close proximity to the divalent metal ions (black) within the active site is stacked on the Y40 residue (blue). This is thought to be encouraged by the other residues on the helical arch, K93 (yellow) and R100 (red). These residues sandwich residue L97 on alpha helix 4, with L111 further up on this helix. L130 is situated on the end of alpha helix 5 (all black).

The anisotropy measurements for DF-5B and DF-21B with proline mutants in the presence of EDTA and Ca^{2+} show very little difference in comparison to WT hFEN1 (chapter 5 also displayed in table 6.1). This disproves any notion that the reason for such poor rates of reaction are due to decreased substrate binding. However, as mentioned in chapter 3, on closer inspection of the structure of the enzyme, it is clear that the 'saddle' of the enzyme, which binds the downstream and upstream regions of substrates is the most important feature of the enzyme in binding and the helical arch actually contributes very little (chapter 3, figure 3.7.1) (Tsutakawa *et al.* 2011). In this way the proline mutants can bind substrates with comparable efficiency to wild type enzyme, whereas orientation of the flap, along with unpairing 2 nt and loss electrophilic catalysis is the main reason for the reduction of rates of reaction.

Chapter 7: Summary and Conclusions

The mechanism by which the FEN enzymes accommodate the 5'-flapped DNA that it is charged with removing has been a longstanding controversy, with a number of alternative mechanisms posited by several different groups (Murante *et al.* 1995; Devos *et al.* 2007; Orans *et al.* 2011; Tsutakawa *et al.* 2011). Experiments that attempted to provide a 'steric road block' at the 5'-termini of substrates and prevent reaction have been attempted in many different ways, and on occasion, have provided conflicting conclusions (Murante *et al.* 1995; Barnes *et al.* 1996; Bornarth *et al.* 1999).

Dahlberg and coworkers first suggested that FEN specificity for removal of 5'-single stranded nucleic acids could be explained by passing the single stranded part of substrates through a hole in the protein. In 1995, Murante *et al*. reported that adding a biotin moiety to the 5'-flap portion of their DNA substrates to which the large tetrameric protein streptavidin could be conjugated prevented FEN catalysed reaction. As a result of this work, the tracking mechanism was proposed. This mechanism attempted to explain the specificity and reaction site selection in terms of its recognition of the 5'-flap part of substrates. The tracking mechanism hypothesis supposed that the FEN enzyme recognises the 5' terminus of substrate and tracks down the flap, until the enzyme reaches its cleavage site one nucleotide into the 5'-duplex. In 1996, the first structure of a FEN protein that visualised the linking region between the main DNA binding domains showed that it formed a helical arch like structure that was wide enough accommodate ssDNA, but not double-stranded DNA (Ceska *et al.* 1996).

However, the results of some later experiments appeared incompatible with passing ssDNA through an archway. The addition of CDDP (cis-diamminedichloroplatin) adducts (Barnes *et al.* 1996; Bornarth *et al.* 1999) and simple branches on the flaps of substrates (Bornarth *et al.* 1999) contradicted the threading/tracking mechanism, because these flap modifications did not appear to prevent reaction. Furthermore, whereas early experiments claimed that forming short regions of duplex within flaps prevented FEN reactions, later experiments demonstrated that FEN could process such substrates (Murante *et al.* 1995; Finger *et al.* 2009). As many of these modifications would be unable to fit through a structured helical arch, it was suggested that instead of threading through the helical arch, the arch could act as a clamp. In the clamping mechanism, the flap DNA was suggested to pass by one of the sides of the helical arch, and upon reaching dsDNA, the arch would clamp over the 5' flap. This might also include looping of the flap portion of the substrate in

order for it to make contact with the active site of the enzyme, as clamping either side would not naturally take the substrate over the active site as is the case in the threading mechanism, i.e. the scissile phosphate would make contact with the active site, and then sharply turn in order to loop around either side of the helical arch.

The hypothesis that initial interactions occurred with the flap DNA and not with the duplexes from which the flap protrudes was first challenged by Joyce and co-workers, who instead suggested a bind and thread mechanism (Devos *et al.* 2007). In this mechanism, likened to threading a needle, FEN proteins were proposed to first interact with the double stranded part of their substrates before threading the flap ssDNA through the helical arch.

In 2007 a crystal structure of T4 bacteriophage FEN (referred to as T4 RNase HI) , provided the first visualisation of a FEN enzyme actually bound to its DNA substrate, albeit in the absence of metal ions that are necessary for catalysis (Devos *et al*. 2007). In this structure, the negative charge of the conserved carboxylates within the metal-free active site forced the DNA away from this region of the protein. Furthermore, part of the helical arch, possibly due its mobility in the crystal, was not visible in the crystal structure. Therefore, this co-crystal structure did not give a lot of insight into how the 5'-portion of substrates were bound by FENs. Nonetheless, when this bound DNA was superimposed upon the crystal structures of T5 FEN or AfFEN, the flap portion of the DNA substrate travelled through the aperture of the helical arch providing support for a threading mechanism. Nevertheless, a number of biochemical studies continued to provide evidence that flap modified substrates, such as those containing flap duplex that could not pass the arch when structured, were efficiently processed by FENs (Barnes *et al.* 1996; Bornarth *et al.* 1999; Finger *et al.* 2009; Sengerová 2009; Zheng *et al.* 2011).

Whilst this work was in progress, structures of human FEN1 and 5'-nuclease superfamily member hEXO1 bound to DNA were solved (Tsutakawa *et al.* 2011 and Orans *et al.* 2011). Neither of these studies visualised an intact flap substrate positioned for reaction, and opposing conclusions were reached about how the 5'-flap would be accommodated. Both studies rejected a tracking mechanism based on the extensive interactions observed with the continuous template strand, and concluded that the initial interaction with the substrate must involve the double stranded and not the flap DNA. However, Orans *et al.* favoured a clamping mechanism, whereas Tsutakawa *et al.*

preferred a threading mechanism. Building on the bind-then-thread model and observations that the helical arch was disordered in substrate-free structures, it was suggested that binding of the bifurcated region of the substrate would place the 5'-flap in the region of the helical arch that could then order to effect reaction. Whilst the clamping mechanism based on hEXO1 structures also proposed the same initial binding of the upstream and downstream DNA, it was suggested that the flap is instead clamped either side of a well-defined arch. (This is described in more detail in chapter 4, section 1 (figure 4.1.1-2).

The work reported here used the streptavidin-biotin interaction in a manner similar to Murante *et al.,* attempting to block the reactions catalysed by T5 FEN and hFEN1 when streptavidin was added to the 5'-end of substrates. The ability to irreversibly trap productive complexes when streptavidin was added after assembly of FEN-substrate complexes was also investigated. Furthermore, a series of competition experiments were undertaken to verify the irreversible nature of trapped complexes, all in an attempt to elucidate which of the many proposed mechanisms is employed by the FEN family of enzymes.

In chapter 3, preliminary work tested the efficacy of the proposed experiments verifying that biotinylation of substrates had negligible effect on FEN-substrate interactions, and developed methods to analyse reactions mixtures of streptavidin conjugated substrates DNA. When the ability of T5 FEN and hFEN1 to catalyse reactions of blocked streptavidin conjugated substrates was tested, reaction was observed. These observations contrast with previous reports that claimed 5'-conjugation of streptavidin abolished FEN reactions (Murante *et al.* 1995). Importantly the reaction proceeded to completion indicating that it could not be explained by reaction of a small portion of the substrate that was not biotinylated. However, the rate of reaction was severely retarded by approximately 3-4 orders of magnitude, when compared to the substrates without a 5'-block, and therefore, did not occur on a biologically relevant timescale. This result was at odds with both the threading and tracking mechanisms. If reaction occurs with the presence of streptavidin on the 5' terminus of the substrates, then it seems very unlikely that recognition or binding occurs on the end of the 5' flaps. Furthermore, these results demonstrate that threading through the helical arch is not an absolute pre-requisite for FEN catalysed reactions.

However, the fact that these 5'-blocked reactions occur on an extremely slow timescale suggests that blocked substrates may not be accommodated optimally. This could be explained by the fact that the large SA tetramer cannot be passed through or out of a structured arch because of its size. SA cannot pass through the helical arch, even when disordered as shown in chapter 4 (figure 4.1.2). The blocked substrates could pass to the side of the helical arch in a clamping mechanism, somehow preventing the arch adopting the optimal structure required for catalysis. However, this does not seem very likely as one would assume clamping of the blocked substrates would not be inhibited to this degree, especially when SA is bound to very long flap substrates. Alternatively, the presence of SA could force the top of the arch back and prevent optimal placement of the important K93 and R100. This seems the more probable of the two options, as it was shown that the rate of reaction between helical arch mutant K93A and a non-SA bound double flap substrate (figure 3.4.1) is comparable to that of the cleavage of 5' blocked substrates.

Chapter 3 also showed reactions between different enzyme substrate (ES) and streptavidin complexes. These were unmixed, premixed, trapped and blocked reactions formed as below, in a repeat of figure 3.6.2.

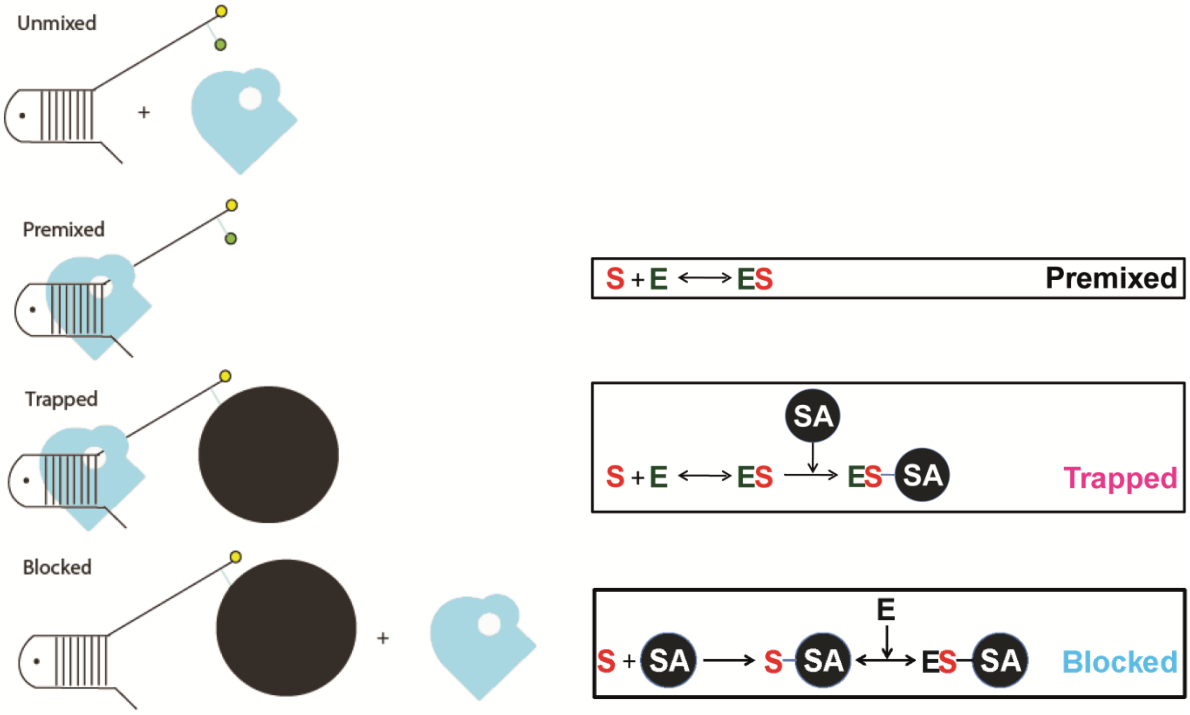


Figure 3.6.2: Diagrams showing the procedures involved in forming 'premixed,' 'trapped' and 'blocked' enzyme-substrate complexes: 5 nM Substrate and 500 nM enzyme, in the case of T5 FEN, were added on ice and incubated for 2 minutes; for hFEN1 E and S preincubation was carried out at room temperature. SA was added on ice, and the mixture incubated at room temperature for one minute. Preincubation was carried out in the presence of EDTA, or the catalytically inert Ca^{2+} ions.

Premixed, trapped and unmixed reactions all produced comparable rates of reaction, which occurred on a biologically relevant timescale, showing that to maintain a normal rate of reaction, a 5' terminus on DNA substrates is necessary. While these experiments rule out tracking mechanism that postulate recognition of the 5' terminus as the initial interaction and instead suggest a mechanism where binding occurs first at the region of bifurcation, they do not fully distinguish between threading and clamping of the flap portion of substrates.

The competition experiments described in chapter 4 were designed to properly elucidate the main mechanism of flap accommodation. These experiments are summarised in figure 4.1.3, repeated below.

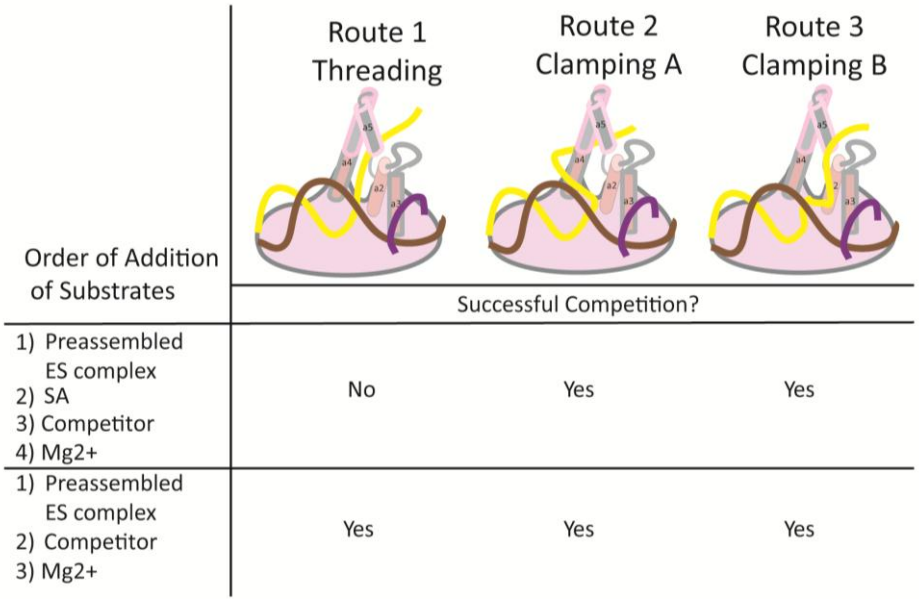


Figure 4.1.3: Predicted results of competition experiments in each binding model scenario, as well as the general scheme of these competition reactions whereby enzyme and substrate are pre-assembled as described in figure 3.6.2, and reagents are added until the addition of Mg^{2+} to initiate the reaction.

The results of these experiments, shown in chapter 4, strongly implicate a bind-then-thread mechanism. This is shown by the ability to compete out of a premixed ES complex, but the consistent inability to compete away a 5'-trapped ES complex of both T5 FEN and hFEN1. Moreover, even substrates with duplex within 5'flaps could be trapped with SA. Because of the dimensions of the structured helical arch this can be explained by a mechanism where the 5'-portion of the substrate is threaded through a disordered arch. The fact that this behaviour is exhibited in both enzymes shows an evolutionarily conserved mechanism.

The bind-then-thread mechanism suggested in this thesis is supported by a recent crystal structure of hFEN1 and its cleaved product, reported by Tsutakawa *et al.,* 2011. The bind-then-thread through a disordered arch mechanism, involves the binding of the 'template' strand of the DNA substrate in a manner that delivers the accompanying flap strand to the aperture of the helical arch (figure 4.1.2, repeated below). The template strand (in brown below) leaves the surface of the enzyme, and contacts it again past the helical arch, forming an arc that forces the flap strand to pass underneath. Along with this arced substrate delivery, the binding of the one nucleotide 3' flap is suggested to

trigger the ordering of the hydrophobic wedge, and facilitate the disorder-order transition of the helical arch around the 5' ssDNA portion of substrates. In substrate-free structures of FEN proteins, the 3'-flap binding site, parts of the wedge and the helical arch are all disordered. Two nucleotides of the reacting duplex are proposed to unpair to bring the scissile phosphate into the proximity of the active site and the active site bound divalent metal ions. This unpairing may be coupled to the structuring of the helical arch described above (Tsutakawa *et al.* 2011).

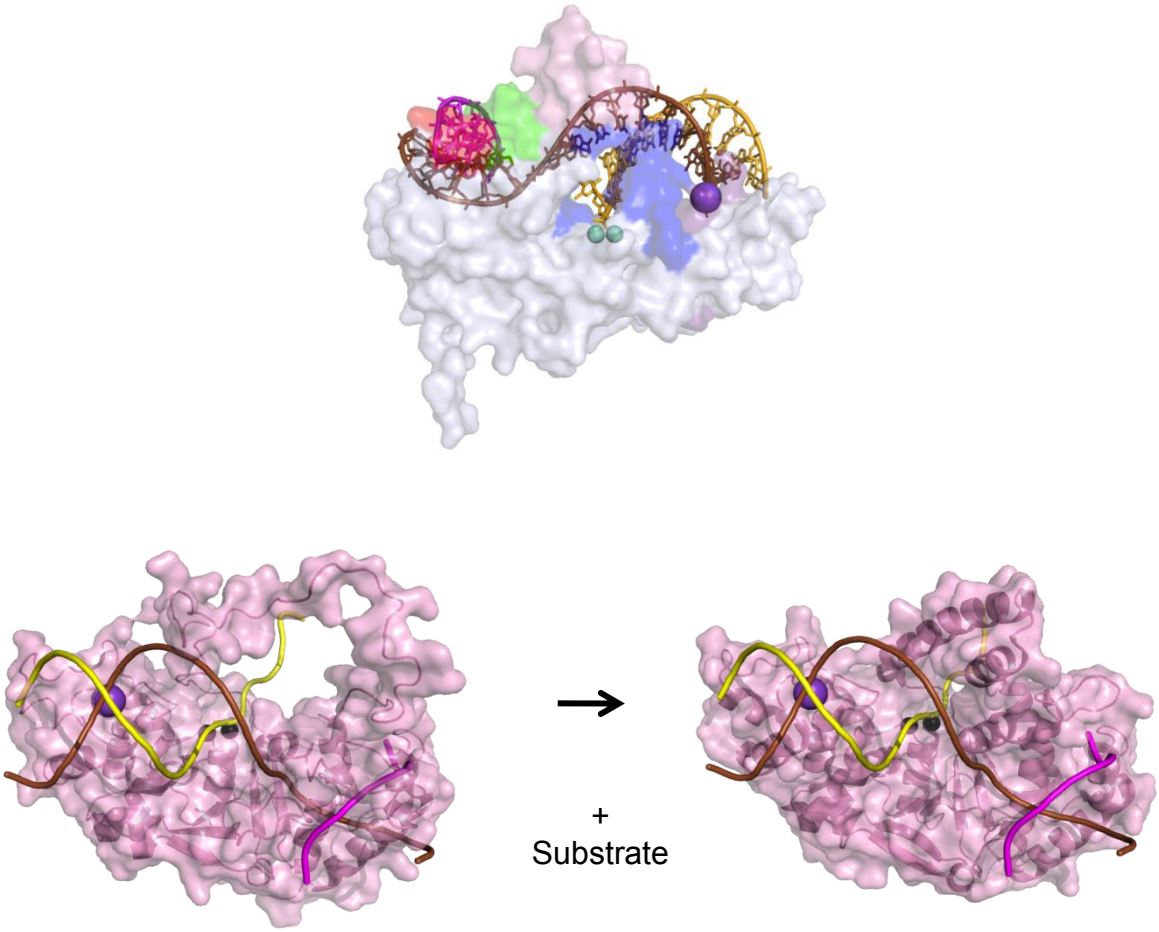


Figure 4.1.2: Top, The flap strand (yellow) follows the arc of the template strand (brown) as it leaves the surface of hFEN1, bending back around and contacting the downstream binding region. The flap strand would be delivered to the helical arch by default. Bottom, the hypothesised disorder to order transition of the helical arch of hFEN1.

It is interesting to note that the disorder-thread-order model posited in this thesis provides an effective method of preserving genomic stability and that this protection would not be afforded if

FENs used a clamping mechanism. As shown below in figure 7.1, a threading mechanism could not productively bind the gapped substrate formed by the parental or template strand and nascent DNA during lagging strand replication, as chapter 3 effectively showed that FENs need a 5' terminus to bind substrates and cleave at a biologically relevant rate. However, a clamping mechanism would not need a 5'-terminus; therefore, processing of the parental or continuous DNA strand could occur during lagging strand replication if clamping is utilised.

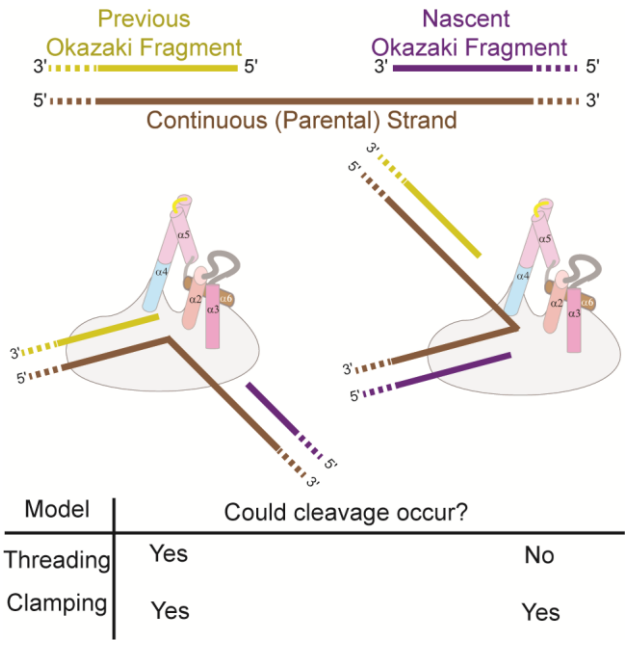


Figure 7.1: Models of hFEN1 and EXO1 substrate interactions and their implications for genome integrity. The gapped junction between Okazaki fragments during lagging strand replication, or an equivalent structure produced as a response to damage, could be bound by hFEN1 and EXO1 in two different ways. Both threading and clamping models could lead to reaction at the free 5' -termini after double nucleotide unpairing, with the 5' -nucleotide contained within or clamped by the arch (left). FENs readily process gapped flaps that contain short regions of duplex by threading them through the disordered arch. However, when the duplex is long or lacks 5' -termini, it could not pass through the arch and so reaction on the continuous strand cannot occur, thereby protecting genome integrity (right). In contrast, if the 5' -portion of the substrate were clamped, the reaction of the continuous strand could occur.

As discussed in chapter 1.6, the FEN superfamily of enzymes act during a number of intracellular processes: (1) the exonucleolytic hydrolyses catalysed by hEXO1 during the double strand break repair, (2) the removal of RNA primers during lagging strand replication catalysed by hFEN1, (3) the ability to cleave bubble structures exhibited by XPG, and (4) cleavage of Holiday junctions by GEN1. While figure 7.1 shows how hFEN1 and hEXO1 do not indiscriminately cleave the healthy parts of the genome during DNA replication, it does not explain the absolute specificity of these enzymes *in vivo*. The structure of hFEN1 solved by Tsutakawa *et al.,* has highlighted two other aspects of FEN which might aid their specificity towards ss DNA flaps with or without small 5'-duplex regions present. The acid block, red in figure 4.1.2 is a group of acidic residues, **E**G**EE**, part of the 3' flap binding pocket, which restricts the 3' flap present on double flap substrates to one nucleotide in length, and importantly, has contacts to the hydroxyl group present on 3' flaps, making it more likely to bind 3' termini. This should prevent inappropriate cleavage that is not one nucleotide into the downstream duplex region. The other more significant factor in the nature of substrates selected is the helical arch alpha helix 5, also known as the helical cap of hFEN1. In the proposed bind-thread-order mechanism proposed above and in chapter 4, the cap once ordered is located at the top of the helical arch as shown in figure 7.2. This cap ensures that there must be a 5′ terminus to the flap portion of the substrate. In related enzymes XPG and GEN1, reported to cleave bubble substrates and Holliday junctions respectively, this helical cap region is either much shorter (GEN1) or very much larger (700+ residues, XPG). If this region of the protein adopts an alternative structure that does not enforce specificity for free 5'-ends in these enzymes, this would help explain how these enzymes accommodate bubble substrates and Holliday junctions while retaining the same active site configuration (figure 7.2).

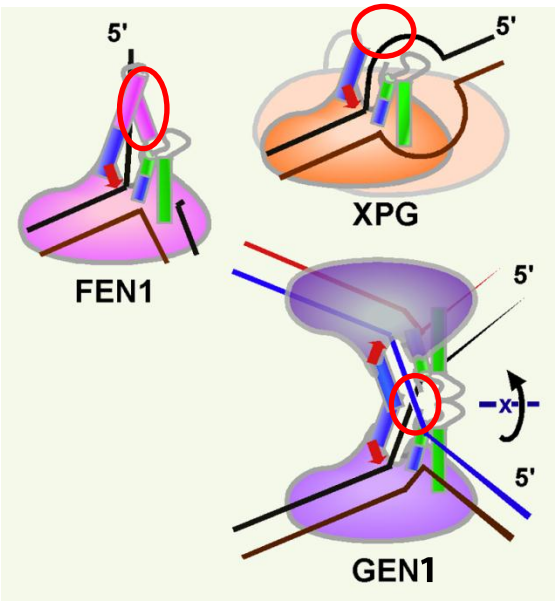


Figure 7.2: Cartoon depictions of enzymes hFEN1, XPG and GEN1, showing the absence of the cap region in hFEN1 (circled), suggesting how the accommodation of these substrates occurs, in correspondence with the mechanism put forward in this thesis. Taken from (Grasby *et al.* 2011) courtesy of Dr. Susan Tsutakawa.

The significance of the helical arch and in particular its structure and mobility was investigated in chapter 6. The arch is proposed to order around the ssDNA of the 5' flap of substrates. The way in which the helical arch orders is proposed to precisely position key catalytic arch residues and in doing so substrates are positioned for efficient catalysis by unpairing two nucleotides into the downstream reacting duplex. Previous work (Storici *et al.* 2002) highlighted three leucine residues, that when mutated to a more rigid proline residue, presented a lethal phenotype. To investigate the cause of this dominant negative phenotype, these mutants were produced and purified to quantify the severity of these mutants on catalysis. Single turnover kinetics were performed with these proteins and compared to WT hFEN1 measurements. This confirmed the drastic nature of these mutations as the rates are 90 000-, 12 000- and 2 000- fold lower than WT with L97P, L111P and L130P, respectively. This catalytic deficiency was not due to altered binding as the measurement of the K_D values using anisotropy revealed no significant change in binding affinity. This is in accord with results presented in chapter 3 and with crystallographic analyses that imply the 'saddle' of the hFEN1 protein, which binds the upstream and downstream portion of substrates, imposes the largest influence on substrate binding. Therefore, enzyme-substrate complex formation is not

affected by the mutation, but structural changes of the helical arch that may be facilitated by substrate binding could be affected. The drastically reduced rates are likely a cumulative effect of the improper placement of important positionally conserved active site residues that help position and unpair the flap of substrates for cleavage. Moreover, disruption of the ability of the helical arch to stably order around the flap of substrates prevents the capture and catalysis. However, due to time constraints, the possibilities were not fully explored and this needs more investigation. Nevertheless, the characterisation of these mutants demonstrates the need for the helical arch for efficient FEN reactions, but not initial interactions with substrates. With more time, full characterisation of the differences between the secondary structure of these mutants and WT FEN1 using CD spectroscopy, both with and without substrate would have been beneficial in understanding the results presented in chapter 6.

Chapter 5 investigates specific substrate elements such as 5' ssDNA flap length and interrogates the importance of such elements in binding to the phage enzyme T5 FEN and hFEN1 using fluorescence anisotropy. Divalent metal ions are shown to influence the affinity of FEN-substrate interactions, although the magnitude of this effect is dependent on the size of the 5’ ssDNA flap portion of substrates. The positively charged divalent ions are thought to shield the repulsive negative charges of the carboxylates within the active site (section 1.6,) which would contribute toward stronger binding of DNA substrates. This observation is important because other studies have claimed to distinguish between models for substrate association based on non-equilibrium affinity measurements in the absence of divalent ions (Gloor *et al.* 2010).

Longer flaps bind with lower affinity than shorter flaps, as seen by long 21 nt flaps possessing a K_D value of around 100 nM, whereas a shorter 5 nt flap possess a K_D value of around 50 nM in hFEN1. This could be due to the lower range of mobility of smaller flap substrates, less random motion within solution of the ssDNA flap might allow for easier access to the helical arch for threading. This preference for smaller flaps seems to be catered for *in vivo,* where longer flaps are cleaved by Dna2 prior to hFEN1 action (Bae *et al.* 2001; Liu *et al.* 2004). When considering these aspects of binding, it would be prudent to consider that in the grand scheme of FEN enzymes, these binding constants do not contribute to the rate determining step as shown in previous work showing a dependence on the rate of diffusion with flap substrates ((Sengerova *et al.* 2010), Dr. Atack, Dr. Sengerova, personal

communication). However, this work does give us a clearer idea of what constitutes an optimum substrate *in vitro* for future work with FEN enzymes.

In conclusion, the data presented in this thesis strongly imply a mechanism where the region of bifurcation on substrates binds under the helical arch followed by the arch ordering around the flap ssDNA. Either coupled to, or after the ordering of the helical arch, unpairing two base pairs into the downstream duplex occurs to position the scissile phosphate within proximity of the active site metal ions. This mechanism, while lending an absolute requirement of flap endonucleases for substrates with 5′ flaps that possess a region of 5′ ssDNA flap and a free 5′ ss or ds DNA terminus, also shows how FENs role in replication and repair might be controlled in order to avoid jeopardizing genomic stability.

With more time, further investigation of the proline mutations within the helical arch would have been undertaken to understand how these mutants are detrimental to the rate of reaction. Work is also already being undertaken by myself and other members of the Grasby group, investigating the possible unpairing of substrates which, would also greatly aid the understanding of this and other nucleases especially in light of suggestions of similar mechanisms within 3′-5′ exonuclease of archaeal B DNA polymerases and polynucleotide kinase (Russell *et al.* 2009; Coquelle *et al.* 2011). Substrates with a disulphide crosslink between flap strand DNA bases and template strand DNA bases either side of the scissile phosphate, which are unresolvable by FENs are being tested, with the ability to unpair and correctly determine the site of cleavage under scrutiny.

Chapter 8: References

Abelson, J., C. R. Trotta and H. Li (1998). "tRNA splicing." J Biol Chem **273**(21): 12685-12688.

Bae, S. H., K. H. Bae, J. A. Kim and Y. S. Seo (2001). "RPA governs endonuclease switching during processing of Okazaki fragments in eukaryotes." Nature **412**(6845): 456-461.

Ban, C. and W. Yang (1998a). "Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis." Cell **95**(4): 541-552.

Ban, C. and W. Yang (1998b). "Structural basis for MutH activation in E.coli mismatch repair and relationship of MutH to restriction endonucleases." EMBO J **17**(5): 1526-1534.

Barnes, C. J., A. F. Wahl, B. Shen, M. S. Park and R. A. Bambara (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**(47): 29624-29631.

Beattie, T. R. and S. D. Bell (2011). "The role of the DNA sliding clamp in Okazaki fragment maturation in archaea and eukaryotes." Biochem Soc Trans **39**(1): 70-76.

Beese, L. S. and T. A. Steitz (1991). "Structural basis for the 3'-5' exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism." EMBO J **10**(1): 25-33.

Bornarth, C. J., T. A. Ranalli, L. A. Henricksen, A. F. Wahl and R. A. Bambara (1999). "Effect of flap modifications on human FEN1 cleavage." Biochemistry **38**(40): 13347-13354.

Brautigam, C. A. and T. A. Steitz (1998). "Structural principles for the inhibition of the 3'-5' exonuclease activity of Escherichia coli DNA polymerase I by phosphorothioates." J Mol Biol **277**(2): 363-377.

Brautigam, C. A., S. Sun, J. A. Piccirilli and T. A. Steitz (1999). "Structures of normal single-stranded DNA and deoxyribo-3'-S-phosphorothiolates bound to the 3'-5' exonucleolytic active site of DNA polymerase I from Escherichia coli." Biochemistry **38**(2): 696-704.

Breyer, W. A. and B. W. Matthews (2000). "Structure of Escherichia coli exonuclease I suggests how processivity is achieved." Nat Struct Biol **7**(12): 1125-1128.

Budd, M. E. and J. L. Campbell (1997). "A yeast replicative helicase, Dna2 helicase, interacts with yeast FEN-1 nuclease in carrying out its essential function." Mol Cell Biol **17**(4): 2136-2142.

Burgers, P. M. (2009). "Polymerase dynamics at the eukaryotic DNA replication fork." J Biol Chem **284**(7): 4041-4045.

Cerritelli, S. M. and R. J. Crouch (2009). "Ribonuclease H: the enzymes in eukaryotes." FEBS J **276**(6): 1494-1505.

Ceska, T. A., J. R. Sayers, G. Stier and D. Suck (1996). "A helical arch allowing single-stranded DNA to thread through T5 5'-exonuclease." Nature **382**(6586): 90-93.

Chapados, B. R., D. J. Hosfield, S. Han, J. Qiu, B. Yelent, B. Shen and J. A. Tainer (2004). "Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair." Cell **116**(1): 39-50.

Chivers, C. E., E. Crozat, C. Chu, V. T. Moy, D. J. Sherratt and M. Howarth (2010). "A streptavidin variant with slower biotin dissociation and increased mechanostability." Nat Methods **7**(5): 391-393.

Chon, H., A. Vassilev, M. L. DePamphilis, Y. Zhao, J. Zhang, P. M. Burgers, R. J. Crouch and S. M. Cerritelli (2009). "Contributions of the two accessory subunits, RNASEH2B and RNASEH2C, to the activity and properties of the human RNase H2 complex." Nucleic Acids Res **37**(1): 96-110.

Chu, C. Y. and T. M. Rana (2007). "Small RNAs: regulators and guardians of the genome." J Cell Physiol **213**(2): 412-419.

Coquelle, N., Z. Havali-Shahriari, N. Bernstein, R. Green and J. N. Glover (2011). "Structural basis for the phosphatase activity of polynucleotide kinase/phosphatase on single- and double-stranded DNA substrates." Proc Natl Acad Sci U S A **108**(52): 21022-21027.

Crow, Y. J., A. Leitch, B. E. Hayward, A. Garner, R. Parmar, E. Griffith, M. Ali, C. Semple, J. Aicardi, R. Babul-Hirji, C. Baumann, P. Baxter, E. Bertini, K. E. Chandler, D. Chitayat, D. Cau, C. Dery, E. Fazzi, C. Goizet, M. D. King, J. Klepper, D. Lacombe, G. Lanzi, H. Lyall, M. L. Martinez-Frias, M. Mathieu, C. McKeown, A. Monier, Y. Oade, O. W. Quarrell, C. D. Rittey, R. C. Rogers, A. Sanchis, J. B. Stephenson, U. Tacke, M. Till, J. L. Tolmie, P. Tomlin, T. Voit, B. Weschke, C. G. Woods, P. Lebon, D. T. Bonthron, C. P. Ponting and A. P. Jackson (2006). "Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection." Nat Genet **38**(8): 910-916.

Crow, Y. J. and J. Rehwinkel (2009). "Aicardi-Goutieres syndrome and related phenotypes: linking nucleic acid metabolism with autoimmunity." Hum Mol Genet **18**(R2): R130-136.

Daniels, D. S., T. T. Woo, K. X. Luu, D. M. Noll, N. D. Clarke, A. E. Pegg and J. A. Tainer (2004). "DNA binding and nucleotide flipping by the human DNA repair protein AGT." Nat Struct Mol Biol **11**(8): 714-720.

Declais, A. C. and D. M. Lilley (2008). "New insight into the recognition of branched DNA structure by junction-resolving enzymes." Curr Opin Struct Biol **18**(1): 86-95.

Dervan, J. J., M. Feng, D. Patel, J. A. Grasby, P. J. Artymiuk, T. A. Ceska and J. R. Sayers (2002). "Interactions of mutant and wild-type flap endonucleases with oligonucleotide substrates suggest an alternative model of DNA binding." Proc Natl Acad Sci U S A **99**(13): 8542-8547.

Devos, J. M., S. J. Tomanicek, C. E. Jones, N. G. Nossal and T. C. Mueser (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**(43): 31713-31724.

Dianov, G. and T. Lindahl (1994). "Reconstitution of the DNA base excision-repair pathway." Curr Biol **4**(12): 1069-1076.

Dionne, I., R. K. Nookala, S. P. Jackson, A. J. Doherty and S. D. Bell (2003). "A heterotrimeric PCNA in the hyperthermophilic archaeon Sulfolobus solfataricus." Mol Cell **11**(1): 275-282.

Dunand-Sauthier, I., M. Hohl, F. Thorel, P. Jaquier-Gubler, S. G. Clarkson and O. D. Scharer (2005). "The spacer region of XPG mediates recruitment to nucleotide excision repair complexes and determines substrate specificity." J Biol Chem **280**(8): 7030-7037.

Finger, L. D., M. S. Blanchard, C. A. Theimer, B. Sengerova, P. Singh, V. Chavez, F. Liu, J. A. Grasby and B. Shen (2009). "The 3'-flap pocket of human flap endonuclease 1 is critical for substrate binding and catalysis." J Biol Chem **284**(33): 22184-22194.

Garforth, S. J., D. Patel, M. Feng and J. R. Sayers (2001). "Unusually wide co-factor tolerance in a metalloenzyme; divalent metal ions modulate endo-exonuclease activity in T5 exonuclease." Nucleic Acids Res **29**(13): 2772-2779.

Gloor, J. W., L. Balakrishnan and R. A. Bambara (2010). "Flap endonuclease 1 mechanism analysis indicates flap base binding prior to threading." J Biol Chem **285**(45): 34922-34931.

Gomes, X. V. and P. M. Burgers (2000). "Two modes of FEN1 binding to PCNA regulated by DNA." EMBO J **19**(14): 3811-3821.

Grasby, J. A., L. D. Finger, S. E. Tsutakawa, J. M. Atack and J. A. Tainer (2011). "Unpairing and gating: sequence-independent substrate recognition by FEN superfamily nucleases." Trends Biochem Sci.

Grindley, N. D., K. L. Whiteson and P. A. Rice (2006). "Mechanisms of site-specific recombination." Annu Rev Biochem **75**: 567-605.

Hamdan, S., P. D. Carr, S. E. Brown, D. L. Ollis and N. E. Dixon (2002). "Structural basis for proofreading during replication of the Escherichia coli chromosome." Structure **10**(4): 535-546.

Harrington, J. J. and M. R. Lieber (1995). "DNA structural elements required for FEN-1 binding." J Biol Chem **270**(9): 4503-4508.

Henneke, G., S. Koundrioukoff and U. Hubscher (2003). "Phosphorylation of human Fen1 by cyclin-dependent kinase modulates its role in replication fork regulation." Oncogene **22**(28): 4301-4313.

Henricksen, L. A., S. Tom, Y. Liu and R. A. Bambara (2000). "Inhibition of flap endonuclease 1 by flap secondary structure and relevance to repeat sequence expansion." J Biol Chem **275**(22): 16420-16427.

Hosfield, D. J., C. D. Mol, B. H. Shen and J. A. Tainer (1998). "Structure of the DNA repair and replication endonuclease and exonuclease FEN-1: Coupling DNA and PCNA binding to FEN-1 activity." Cell **95**(1): 135-146.

Hutton, R. D., T. D. Craggs, M. F. White and J. C. Penedo (2010). "PCNA and XPF cooperate to distort DNA substrates." Nucleic Acids Res **38**(5): 1664-1675.

Hutton, R. D., J. A. Roberts, J. C. Penedo and M. F. White (2008). "PCNA stimulates catalysis by structure-specific nucleases using two distinct mechanisms: substrate targeting and catalytic step." Nucleic Acids Res **36**(21): 6720-6727.

Ivanov, I., J. A. Tainer and J. A. McCammon (2007). "Unraveling the three-metal-ion catalytic mechanism of the DNA repair enzyme endonuclease IV." Proc Natl Acad Sci U S A **104**(5): 1465-1470.

James, R., C. Kleanthous and G. R. Moore (1996). "The biology of E colicins: paradigms and paradoxes." Microbiology **142 (Pt 7)**: 1569-1580.

Jameson, D. M. and W. H. Sawyer (1995). "Fluorescence anisotropy applied to biomolecular interactions." Methods Enzymol **246**: 283-300.

Jencks, W. P. (1972). "General acid-base catalysis of complex reactions in water." Chem Rev **72**: pp705-718.

Jones, S. J., A. F. Worrall and B. A. Connolly (1996). "Site-directed mutagenesis of the catalytic residues of bovine pancreatic deoxyribonuclease I." J Mol Biol **264**(5): 1154-1163.

Kanai, Y., G. Ishikawa, R. Takeuchi, T. Ruike, R. Nakamura, A. Ihara, T. Ohashi, K. Takata, S. Kimura and K. Sakaguchi (2007). "DmGEN shows a flap endonuclease activity, cleaving the blocked-flap structure and model replication fork." FEBS J **274**(15): 3914-3927.

Kang, M. J., C. H. Lee, Y. H. Kang, I. T. Cho, T. A. Nguyen and Y. S. Seo (2010). "Genetic and functional interactions between Mus81-Mms4 and Rad27." Nucleic Acids Res **38**(21): 7611-7625.

Kao, H. I. and R. A. Bambara (2003). "The protein components and mechanism of eukaryotic Okazaki fragment maturation." Crit Rev Biochem Mol Biol **38**(5): 433-452.

Kim, J. H., H. D. Kim, G. H. Ryu, D. H. Kim, J. Hurwitz and Y. S. Seo (2006). "Isolation of human Dna2 endonuclease and characterization of its enzymatic properties." Nucleic Acids Res **34**(6): 1854-1864.

Kolodner, R., S. D. Hall and C. Luisi-DeLuca (1994). "Homologous pairing proteins encoded by the Escherichia coli recE and recT genes." Mol Microbiol **11**(1): 23-30.

Kovall, R. and B. W. Matthews (1997). "Toroidal structure of lambda-exonuclease." Science **277**(5333): 1824-1827.

Kramer, R. (1999). "Bioinorganic models for the catalytic cooperation of metal ions and functional groups in nuclease and peptidase enzymes." Coordination Chemistry Reviews **182**: pp243-261.

Kurthkoti, K. and U. Varshney (2011). "Base excision and nucleotide excision repair pathways in mycobacteria." Tuberculosis (Edinb) **91**(6): 533-543.

Lakowicz, J. R. (1999). Principles of Fluorescence Spectroscopy.

Larsen, E., C. Gran, B. E. Saether, E. Seeberg and A. Klungland (2003). "Proliferation failure and gamma radiation sensitivity of Fen1 null mutant mice at the blastocyst stage." Mol Cell Biol **23**(15): 5346-5353.

Lee, J. Y., J. Chang, N. Joseph, R. Ghirlando, D. N. Rao and W. Yang (2005). "MutH complexed with hemi- and unmethylated DNAs: coupling base recognition and DNA cleavage." Mol Cell **20**(1): 155-166.

Lehmann, A. R. (1996). "Molecular biology of DNA repair in the fission yeast Schizosaccharomyces pombe." Mutat Res **363**(3): 147-161.

Liu, R., J. Qiu, L. D. Finger, L. Zheng and B. Shen (2006). "The DNA-protein interaction modes of FEN-1 with gap substrates and their implication in preventing duplication mutations." Nucleic Acids Res **34**(6): 1772-1784.

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

Lovell, S. C., I. W. Davis, W. B. Arendall, 3rd, P. I. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson and D. C. Richardson (2003). "Structure validation by Calpha geometry: phi,psi and Cbeta deviation." Proteins **50**(3): 437-450.

Lundblad, J. R., M. Laurance and R. H. Goodman (1996). "Fluorescence polarization analysis of protein-DNA and protein-protein interactions." Mol Endocrinol **10**(6): 607-612.

Lyamichev, V., M. A. Brow, V. E. Varvel and J. E. Dahlberg (1999). "Comparison of the 5' nuclease activities of taq DNA polymerase and its isolated nuclease domain." Proc Natl Acad Sci U S A **96**(11): 6143-6148.

Markham, N. R. and M. Zuker (2005). "DINAMelt web server for nucleic acid melting prediction." Nucleic Acids Res **33**(Web Server issue): W577-581.

Marti, T. M. and O. Fleck (2004). "DNA repair nucleases." Cell Mol Life Sci **61**(3): 336-354.

Masuda-Sasa, T., O. Imamura and J. L. Campbell (2006). "Biochemical analysis of human Dna2." Nucleic Acids Res **34**(6): 1865-1875.

McHenry, C. S. (1985). "DNA polymerase III holoenzyme of Escherichia coli: components and function of a true replicative complex." Mol Cell Biochem **66**(1): 71-85.

Mimitou, E. P. and L. S. Symington (2009). "Nucleases and helicases take center stage in homologous recombination." Trends Biochem Sci **34**(5): 264-272.

Moore, M. J. and N. J. Proudfoot (2009). "Pre-mRNA processing reaches back to transcription and ahead to translation." Cell **136**(4): 688-700.

Mueser, T. C., N. G. Nossal and C. C. Hyde (1996). "Structure of bacteriophage T4 RNase H, a 5' to 3' RNA-DNA and DNA-DNA exonuclease with sequence similarity to the RAD2 family of eukaryotic proteins." Cell **85**(7): 1101-1112.

Murante, R. S., L. Rust and R. A. Bambara (1995). "Calf 5' to 3' exo/endonuclease must slide from a 5' end of the substrate to perform structure-specific cleavage." J Biol Chem **270**(51): 30377-30383.

Neish, C. S., I. L. Martin, R. M. Henderson and J. M. Edwardson (2002). "Direct visualization of ligand-protein interactions using atomic force microscopy." Br J Pharmacol **135**(8): 1943-1950.

Newman, M., K. Lunnen, G. Wilson, J. Greci, I. Schildkraut and S. E. Phillips (1998). "Crystal structure of restriction endonuclease BglI bound to its interrupted DNA recognition sequence." EMBO J **17**(18): 5466-5476.

Nowotny, M., S. M. Cerritelli, R. Ghirlando, S. A. Gaidamakov, R. J. Crouch and W. Yang (2008). "Specific recognition of RNA/DNA hybrid and enhancement of human RNase H1 activity by HBD." EMBO J **27**(7): 1172-1181.

Nowotny, M. and W. Yang (2009). "Structural and functional modules in RNA interference." Curr Opin Struct Biol **19**(3): 286-293.

Orans, J., E. A. McSweeney, R. R. Iyer, M. A. Hast, H. W. Hellinga, P. Modrich and L. S. Beese (2011). "Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family." Cell **145**(2): 212-223.

Orlowski, J. and J. M. Bujnicki (2008). "Structural and evolutionary classification of Type II restriction enzymes based on theoretical and experimental analyses." Nucleic Acids Res **36**(11): 3552-3569.

Parrish, J. Z. and D. Xue (2006). "Cuts can kill: the roles of apoptotic nucleases in cell death and animal development." Chromosoma **115**(2): 89-97.

Parrish, J. Z., C. Yang, B. Shen and D. Xue (2003). "CRN-1, a Caenorhabditis elegans FEN-1 homologue, cooperates with CPS-6/EndoG to promote apoptotic DNA degradation." EMBO J **22**(13): 3451-3460.

Patel, A. A. and J. A. Steitz (2003). "Splicing double: insights from the second spliceosome." Nat Rev Mol Cell Biol **4**(12): 960-970.

Patel, D., M. R. Tock, E. Frary, M. Feng, T. J. Pickering, J. A. Grasby and J. R. Sayers (2002). "A conserved tyrosine residue aids ternary complex formation, but not catalysis, in phage T5 flap endonuclease." J Mol Biol **320**(5): 1025-1035.

Patel, N., J. M. Atack, L. D. Finger, J. C. Exell, P. Thompson, S. Tsutakawa, J. A. Tainer, D. M. Williams and J. A. Grasby (2012). "Flap endonucleases pass 5'-flaps through a flexible arch using a disorder-thread-order mechanism to confer specificity for free 5'-ends." Nucleic Acids Res **40**(10): 4507-4519.

Pelletier, H. and M. R. Sawaya (1996). "Characterization of the metal ion binding helix-hairpin-helix motifs in human DNA polymerase beta by X-ray structural analysis." Biochemistry **35**(39): 12778-12787.

Pickering, T. J., S. J. Garforth, S. J. Thorpe, J. R. Sayers and J. A. Grasby (1999). "A single cleavage assay for T5 5'-->3' exonuclease: determination of the catalytic parameters forwild-type and mutant proteins." Nucleic Acids Res **27**(3): 730-735.

Qiu, J., X. Li, G. Frank and B. Shen (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 Biol Chem **276**(7): 4901-4908.

Qiu, J., R. Liu, B. R. Chapados, M. Sherman, J. A. Tainer and B. Shen (2004). "Interaction interface of human flap endonuclease-1 with its DNA substrates." J Biol Chem **279**(23): 24394-24402.

Qiu, J., Y. Qian, P. Frank, U. Wintersberger and B. Shen (1999). "Saccharomyces cerevisiae RNase H(35) functions in RNA primer removal during lagging-strand DNA synthesis, most efficiently in cooperation with Rad27 nuclease." Mol Cell Biol **19**(12): 8361-8371.

Raines, R. T. (1998). "Ribonuclease A." Chem Rev **98**(3): 1045-1066.

Raines, R. T., Ed. (2004). Active Site of ribonuclease A. Artificial Nucleases. Heidelberg, Germany.

Reha-Krantz, L. J. (2010). "DNA polymerase proofreading: Multiple roles maintain genome stability." Biochim Biophys Acta **1804**(5): 1049-1063.

Rumbaugh, J. A., L. A. Henricksen, M. S. DeMott and R. A. Bambara (1999). "Cleavage of substrates with mismatched nucleotides by Flap endonuclease-1. Implications for mammalian Okazaki fragment processing." J Biol Chem **274**(21): 14602-14608.

Russell, H. J., T. T. Richardson, K. Emptage and B. A. Connolly (2009). "The 3'-5' proofreading exonuclease of archaeal family-B DNA polymerase hinders the copying of template strand deaminated bases." Nucleic Acids Res **37**(22): 7603-7611.

Saenger, W. (1984). "Principles of Nucleic Acid Structure." NY : Springer.

Sakurai, S., K. Kitano, H. Yamaguchi, K. Hamada, K. Okada, K. Fukuda, M. Uchida, E. Ohtsuka, H. Morioka and T. Hakoshima (2005). "Structural basis for recruitment of human flap endonuclease 1 to PCNA." EMBO J **24**(4): 683-693.

Sano, T. and C. R. Cantor (1990). "Cooperative biotin binding by streptavidin. Electrophoretic behavior and subunit association of streptavidin in the presence of 6 M urea." J Biol Chem **265**(6): 3369-3373.

Sauer, U. H., D. P. San and B. W. Matthews (1992). "Tolerance of T4 lysozyme to proline substitutions within the long interdomain alpha-helix illustrates the adaptability of proteins to potentially destabilizing lesions." J Biol Chem **267**(4): 2393-2399.

Sayers, J. R. and F. Eckstein (1990). "Properties of overexpressed phage T5 D15 exonuclease. Similarities with Escherichia coli DNA polymerase I 5'-3' exonuclease." J Biol Chem **265**(30): 18311-18317.

Sayers, J. R. and F. Eckstein (1991). "A single-strand specific endonuclease activity copurifies with overexpressed T5 D15 exonuclease." Nucleic Acids Res **19**(15): 4127-4132.

Schroeder, K., Chetan Lad, Paul Wyman, Nicholas H. Williams, Richard Wolfenden (2005). "The time required for water attack at the phosphorous atom of simple phosphodiesters and of DNA." PNAS **103**(11): pp4052-4055.

Schultz, S. J. and J. J. Champoux (2008). "RNase H activity: structure, specificity, and function in reverse transcription." Virus Res **134**(1-2): 86-103.

Sclafani, R. A. and T. M. Holzen (2007). "Cell cycle regulation of DNA replication." Annu Rev Genet **41**: 237-280.

Sengerová, B. (2009). Mechanistic studies of T5 bacteriophage flap endonuclease, University of Sheffield. **pHD:** 250.

Sengerova, B., C. Tomlinson, J. M. Atack, R. Williams, J. R. Sayers, N. H. Williams and J. A. Grasby (2010). "Bronsted analysis and rate-limiting steps for the t5 flap endonuclease catalyzed hydrolysis of exonucleolytic substrates." Biochemistry **49**(37): 8085-8093.

Sevenich, F. W., J. Langowski, V. Weiss and K. Rippe (1998). "DNA binding and oligomerization of NtrC studied by fluorescence anisotropy and fluorescence correlation spectroscopy." Nucleic Acids Res **26**(6): 1373-1381.

Shen, B., P. Singh, R. Liu, J. Qiu, L. Zheng, L. D. Finger and S. Alas (2005). "Multiple but dissectible functions of FEN-1 nucleases in nucleic acid processing, genome stability and diseases." Bioessays **27**(7): 717-729.

Sorek, R., V. Kunin and P. Hugenholtz (2008). "CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea." Nat Rev Microbiol **6**(3): 181-186.

Stahl, M. M., L. Thomason, A. R. Poteete, T. Tarkowski, A. Kuzminov and F. W. Stahl (1997). "Annealing vs. invasion in phage lambda recombination." Genetics **147**(3): 961-977.

Staresincic, L., A. F. Fagbemi, J. H. Enzlin, A. M. Gourdin, N. Wijgers, I. Dunand-Sauthier, G. Giglia-Mari, S. G. Clarkson, W. Vermeulen and O. D. Scharer (2009). "Coordination of dual incision and repair synthesis in human nucleotide excision repair." EMBO J **28**(8): 1111-1120.

Storici, F., G. Henneke, E. Ferrari, D. A. Gordenin, U. Hubscher and M. A. Resnick (2002). "The flexible loop of human FEN1 endonuclease is required for flap cleavage during DNA replication and repair." EMBO J **21**(21): 5930-5942.

Stuckey, J. A. and J. E. Dixon (1999). "Crystal structure of a phospholipase D family member." Nat Struct Biol **6**(3): 278-284.

Syson, K., C. Tomlinson, B. R. Chapados, J. R. Sayers, J. A. Tainer, N. H. Williams and J. A. Grasby (2008). "Three metal ions participate in the reaction catalyzed by T5 flap endonuclease." J Biol Chem **283**(42): 28741-28746.

Szankasi, P. and G. R. Smith (1995). "A role for exonuclease I from S. pombe in mutation avoidance and mismatch correction." Science **267**(5201): 1166-1169.

Szczesny, B., A. W. Tann, M. J. Longley, W. C. Copeland and S. Mitra (2008). "Long patch base excision repair in mammalian mitochondrial genomes." J Biol Chem **283**(39): 26349-26356.

Tock, M. R. and D. T. Dryden (2005). "The biology of restriction and anti-restriction." Curr Opin Microbiol **8**(4): 466-472.

Tock, M. R., E. Frary, J. R. Sayers and J. A. Grasby (2003). "Dynamic evidence for metal ion catalysis in the reaction mediated by a flap endonuclease." EMBO J **22**(5): 995-1004.

Tom, S., L. A. Henricksen and R. A. Bambara (2000). "Mechanism whereby proliferating cell nuclear antigen stimulates flap endonuclease 1." J Biol Chem **275**(14): 10498-10505.

Tomlinson, C. (2010). University of Sheffield.

Tomlinson, C. G. (2011). An investigation into the mechanism of flap endonuclease catalysis, University of Sheffield**:** 197.

Tseng, H. M. and A. E. Tomkinson (2004). "Processing and joining of DNA ends coordinated by interactions among Dnl4/Lif1, Pol4, and FEN-1." J Biol Chem **279**(46): 47580-47588.

Tsutakawa, S. E., S. Classen, B. R. Chapados, A. S. Arvai, L. D. Finger, G. Guenther, C. G. Tomlinson, P. Thompson, A. H. Sarker, B. Shen, P. K. Cooper, J. A. Grasby and J. A. Tainer (2011). "Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily." Cell **145**(2): 198-211.

Voet, D. V., J. G. (2004). "Biochemistry." NJ : Wiley.

Warner, H. R., B. F. Demple, W. A. Deutsch, C. M. Kane and S. Linn (1980). "Apurinic/apyrimidinic endonucleases in repair of pyrimidine dimers and other lesions in DNA." Proc Natl Acad Sci U S A **77**(8): 4602-4606.

Weber, P. C., D. H. Ohlendorf, J. J. Wendoloski and F. R. Salemme (1989). "Structural origins of high-affinity biotin binding to streptavidin." Science **243**(4887): 85-88.

Weston, S. A., A. Lahm and D. Suck (1992). "X-ray structure of the DNase I-d(GGTATACC)2 complex at 2.3 A resolution." J Mol Biol **226**(4): 1237-1256.

Williams, N. H., Ed. (1998). Phosphate Diesterases and Triesterases. Comprehensive Biological Catalysis. Reaction of Electrophilic Carbon, Phoshorous and Sulphur. London, Academic Press.

Williams, R., B. Sengerova, S. Osborne, K. Syson, S. Ault, A. Kilgour, B. R. Chapados, J. A. Tainer, J. R. Sayers and J. A. Grasby (2007). "Comparison of the catalytic parameters and reaction specificities of a phage and an archaeal flap endonuclease." J Mol Biol **371**(1): 34-48.

Xu, Y., V. Derbyshire, K. Ng, X. C. Sun, N. D. Grindley and C. M. Joyce (1997). "Biochemical and mutational studies of the 5'-3' exonuclease of DNA polymerase I of Escherichia coli." J Mol Biol **268**(2): 284-302.

Xu, Y., O. Potapova, A. E. Leschziner, N. D. F. Grindley and C. M. Joyce (2001). "Contacts between the 5' nuclease of DNA polymerase I and its substrate DNA." Journal of Biological Chemistry **276**: 30167-30177.

Yang, W. (2011). "Nucleases: diversity of structure, function and mechanism." Q Rev Biophys **44**(1): 1-93.

Zhang, J., X. Xing, A. B. Herr and C. E. Bell (2009). "Crystal structure of E. coli RecE protein reveals a toroidal tetramer for processing double-stranded DNA breaks." Structure **17**(5): 690-702.

Zheng, L., J. Jia, L. D. Finger, Z. Guo, C. Zer and B. Shen (2011). "Functional regulation of FEN1 nuclease and its link to cancer." Nucleic Acids Res **39**(3): 781-794.

Zheng, L., M. Zhou, Q. Chai, J. Parrish, D. Xue, S. M. Patrick, J. J. Turchi, S. M. Yannone, D. Chen and B. Shen (2005). "Novel function of the flap endonuclease 1 complex in processing stalled DNA replication forks." EMBO Rep **6**(1): 83-89.

Zheng, L., M. Zhou, Z. Guo, H. Lu, L. Qian, H. Dai, J. Qiu, E. Yakubovskaya, D. F. Bogenhagen, B. Demple and B. Shen (2008). "Human DNA2 is a mitochondrial nuclease/helicase for efficient processing of DNA replication and repair intermediates." Mol Cell **32**(3): 325-336.

Zucker, M. (2003). "Mfold web server for nucleic acid folding and hybridisation prediction " Nucleic acids Research **31**: 3406-3415.

Chapter 9: Appendices

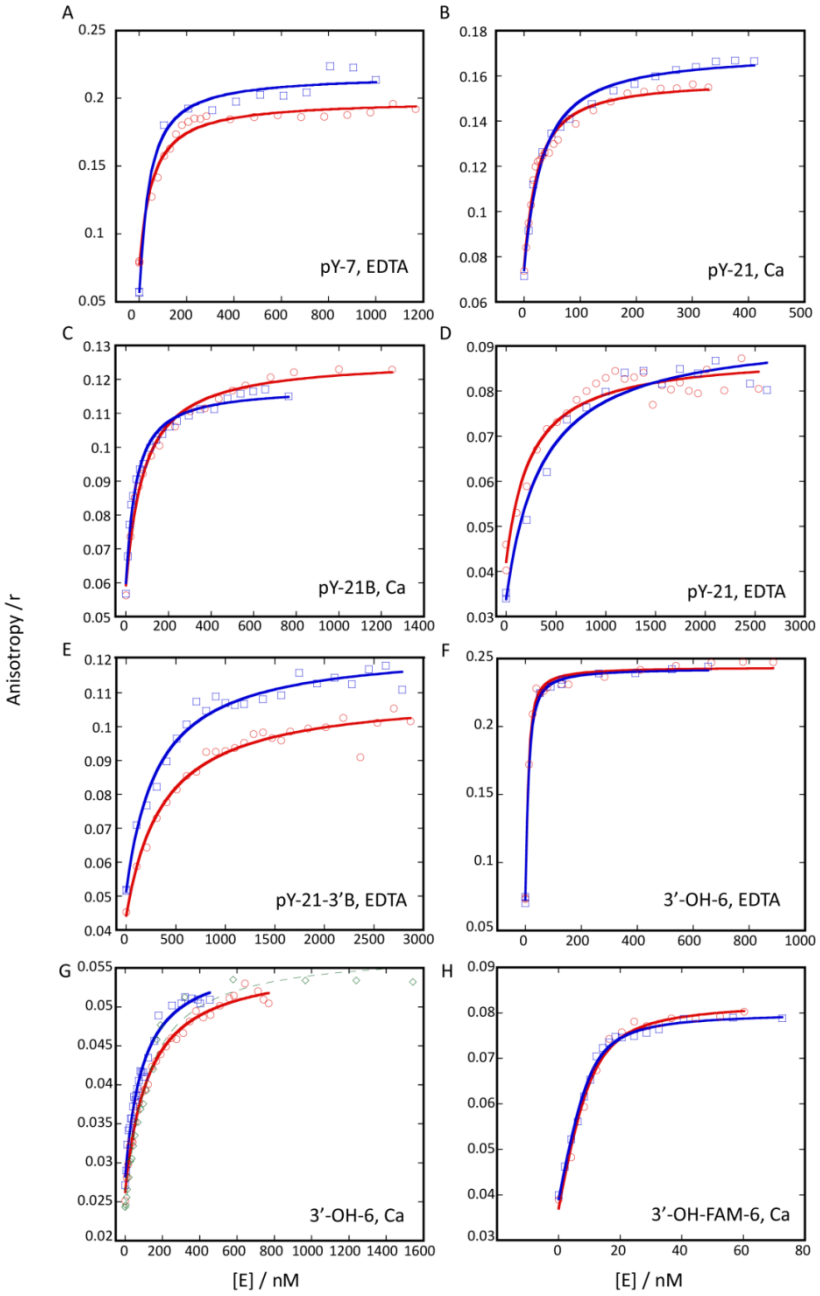


Figure A1: Binding curves measuring the anisotropy changes on stepwise addition of T5 FEN to A) pY-7 in 1 mM EDTA, B) pY-21 in 10 mM Ca^{2+}, C) pY-21B in 10 mM Ca^{2+}, D) pY-21 in 1 mM EDTA, E) pY-21-3'B in 1 mM EDTA, F) 3'-OH-6 in 1 mM EDTA, G) 3'-FAM-OH-6 in 1 mM EDTA, H) 3'-FAM-OH-6 in 10 mM Ca^{2+}, at 37°C. This was performed at 50 mM pH 7.5 HEPES, 100 mM KCl and 0.01 mg/ml BSA and 1 mM DTT to maintain normal reaction conditions without initiating cleavage. The excitation wavelength was set to 490 nm, the emission at 510 nm recorded. Slit widths were set to 10 nm, with an average of 10 scans taken per reading. Two sets of data were fitted to equation 3 (see chapter 2) and results are summarised in table 5.1.

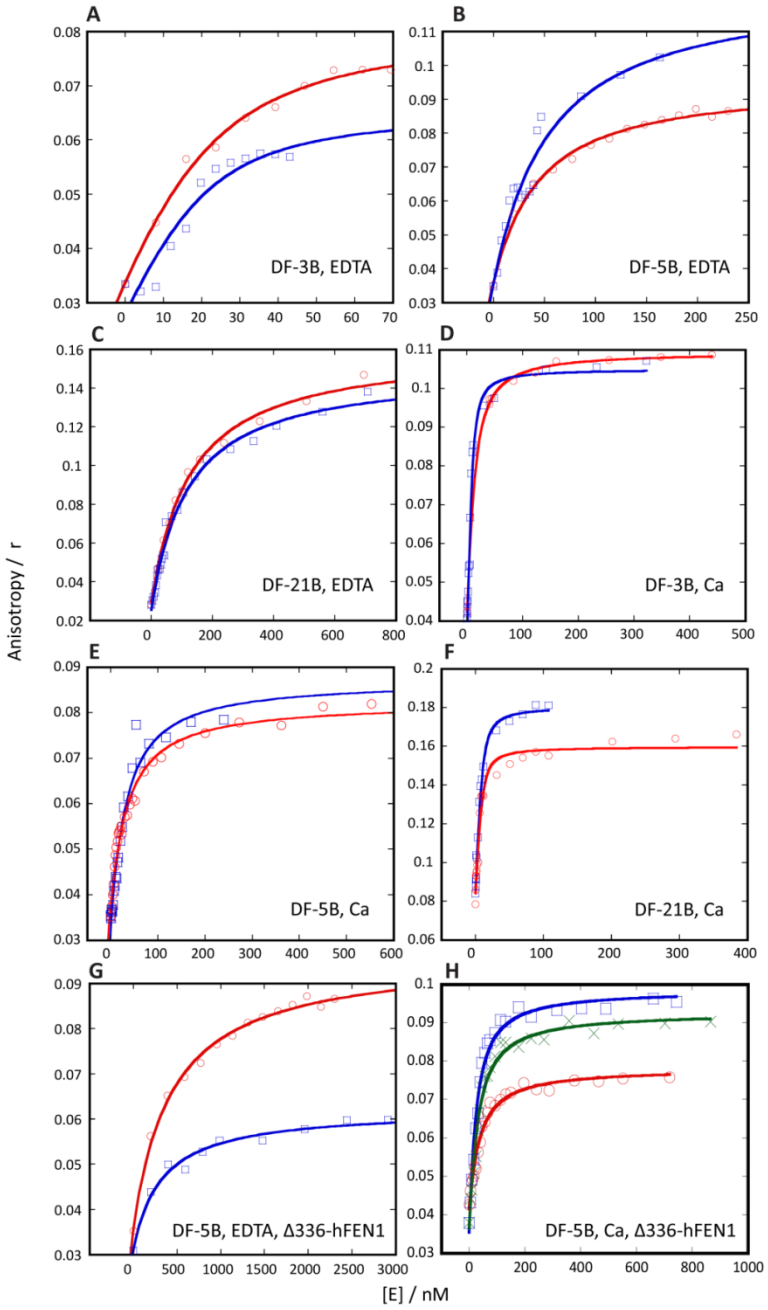


Figure A2: Binding curves measuring the anisotropy changes on stepwise addition of hFEN1 to A) DF-3B in 1 mM EDTA, B) DF-5B in 1 mM EDTA, C) DF-21B in 1 mM EDTA, D) DF-3B in 10 mM Ca^{2+}, E) DF-5B in 10 mM Ca^{2+}, F) DF-21B in 10 mM Ca^{2+} and addition of hFEN1-Δ336 to G) DF-5B in 1 mM EDTA, H) DF-21B in 10 mM Ca^{2+}, at 37°C. This was performed at 25 mM pH 7.5 HEPES, 50 mM KCl and 0.01 mg/ml BSA and 1 mM DTT to maintain normal reaction conditions without initiating cleavage. The excitation wavelength was set to 490 nm, the emission at 510 nm recorded. Slit widths were set to 10 nm, with an average of 10 scans taken per reading. Two sets of data were fitted to equation 3 (see chapter 2) and results are summarised in table 5.2-3.

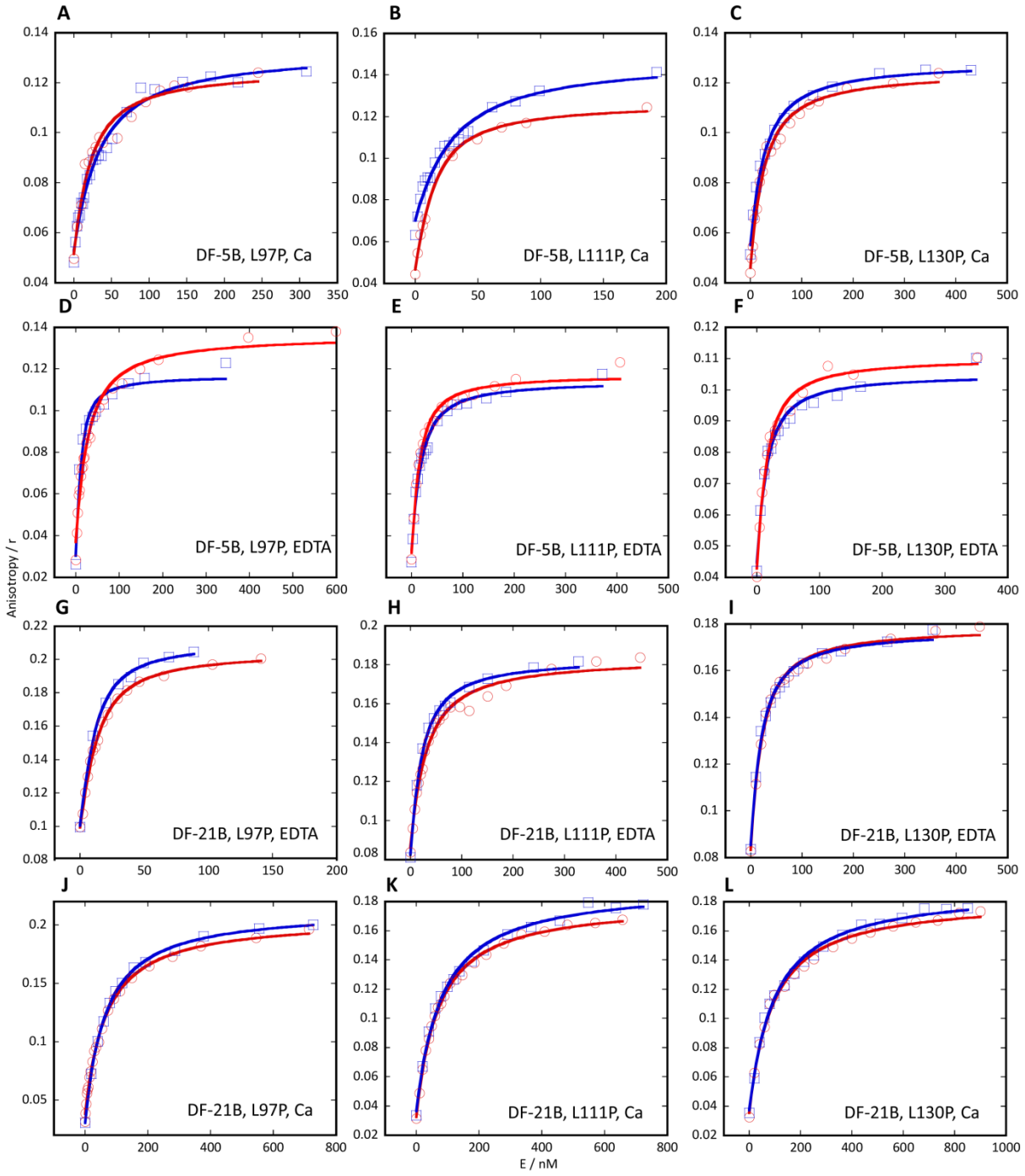


Figure A3: Binding curves measuring the anisotropy changes on stepwise addition of L97P to A) DF-5B in 10 mM Ca^{2+}, D) DF-5B in 1 mM EDTA, G) DF-21B in 1 mM EDTA and J) DF-21B in 10 mM Ca^{2+}; addition of L111P to B) DF-5B in 10 mM Ca^{2+}, E) DF-5B in 1 mM EDTA, H) DF-21B in 1 mM EDTA and K) DF-21B in 10 mM Ca^{2+}; addition of L130P to C) DF-5B in 10 mM Ca^{2+}, F) DF-5B in 1 mM EDTA, I) DF-21B in 1 mM EDTA and L) DF-21B in 10 mM Ca^{2+}, at 37°C. This was performed at 50 mM pH 7.5 HEPES, 100 mM KCl and 0.01 mg/ml BSA and 1 mM DTT to maintain normal reaction conditions without initiating cleavage. The excitation wavelength was set to 490 nm, the emission at 510 nm recorded. Slit widths were set to 10 nm, with an average of 10 scans taken per reading. Two sets of data were fitted to equation 3 (see chapter 2) and results are summarised in table 6.2.

```
Sequence 1: hFEN1 mRNA
Sequence 2: L97P forward

Similarity : 852/1139 (74.80 %)

                                                          M  G  I  Q  G
Seq_1  1     ---------------------------------------------atgggaattcaag  13
                                                          ||||||||||||
Seq_2  1     ***************AATTTT**TTAACTTTAAGAAGGAGA*ATACCATGGGAATTCAAG  60
               X  X  X  X  X  F  X  L  T  L  R  R  R  X  T  M  G  I  Q  G


               L  A  K  L  I  A  D  V  A  P  S  A  I  R  E  N  D  I  K  S
Seq_1  14    gcctggccaaactaattgctgatgtggccccagtgccatccgggagaatgacatcaaga  73
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  61    GCCTGGCCAAACTAATTGCTGATGTGGCCCCCAGTGCCATCCGGGAGAATGACATCAAGA  120
               L  A  K  L  I  A  D  V  A  P  S  A  I  R  E  N  D  I  K  S


               Y  F  G  R  K  V  A  I  D  A  S  M  S  I  Y  Q  F  L  I  A
Seq_1  74    gctactttggccgtaaggtggccattgatgcctctatgagcatttatcagttcctgattg  133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  121   GCTACTTTGGCCGTAAGGTGGCCATTGATGCCTCTATGAGCATTTATCAGTTCCTGATTG  180
               Y  F  G  R  K  V  A  I  D  A  S  M  S  I  Y  Q  F  L  I  A


               V  R  Q  G  G  D  V  L  Q  N  E  E  G  E  T  T  S  H  L  M
Seq_1  134   ctgttcgccagggtggggatgtgctgcagaatgaggagggtgagaccaccagccacctga  193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  181   CTGTTCGCCAGGGTGGGGATGTGCTGCAGAATGAGGAGGGTGAGACCACCAGCCACCTGA  240
               V  R  Q  G  G  D  V  L  Q  N  E  E  G  E  T  T  S  H  L  M


               G  M  F  Y  R  T  I  R  M  M  E  N  G  I  K  P  V  Y  V  F
Seq_1  194   tgggcatgttctaccgcaccattcgcatgatggagaacggcatcaagcccgtgtatgtct  253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  241   TGGGCATGTTCTACCGCACCATTCGCATGATGGAGAACGGCATCAAGCCCGTGTATGTCT  300
               G  M  F  Y  R  T  I  R  M  M  E  N  G  I  K  P  V  Y  V  F


               D  G  K  P  P  Q  L  K  S  G  E  L  A  K  R  S  E  R  R  A
Seq_1  254   ttgatggcaagccgccacagctcaagtcaggcgagctggccaaacgcagtgagcggcggg  313
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Seq_2  301   TTGATGGCAAGCCGCCACAGCTCAAGTCAGGCGAGCCGGCCAAACGCAGTGAGCGGCGGG  360
               D  G  K  P  P  Q  L  K  S  G  E  P  A  K  R  S  E  R  R  A


               E  A  E  K  Q  L  Q  Q  A  Q  A  A  G  A  E  Q  E  V  E  K
Seq_1  314   ctgaggcagagaagcagctgcagcaggctcaggctgctggggccgagcaggaggtggaaa  373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  361   CTGAGGCAGAGAAGCAGCTGCAGCAGGCTCAGGCTGCTGGGGCCGAGCAGGAGGTGGAAA  420
               E  A  E  K  Q  L  Q  Q  A  Q  A  A  G  A  E  Q  E  V  E  K


               F  T  K  R  L  V  K  V  T  K  Q  H  N  D  E  C  K  H  L  L
Seq_1  374   aattcactaagcggctggtgaaggtcactaagcagcacaatgatgagtgcaaacatctgc  433
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  421   AATTCACTAAGCGGCTGGTGAAGGTCACTAAGCAGCACAATGATGAGTGCAAACATCTGC  480
               F  T  K  R  L  V  K  V  T  K  Q  H  N  D  E  C  K  H  L  L


               S  L  M  G  I  P  Y  L  D  A  P  S  E  A  E  A  S  C  A  A
Seq_1  434   tgagcctcatgggcatcccttatcttgatgcacccagtgaggcagaggccagctgtgctg  493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  481   TGAGCCTCATGGGCATCCCTTATCTTGATGCACCCAGTGAGGCAGAGGCCAGCTGTGCTG  540
               S  L  M  G  I  P  Y  L  D  A  P  S  E  A  E  A  S  C  A  A
```

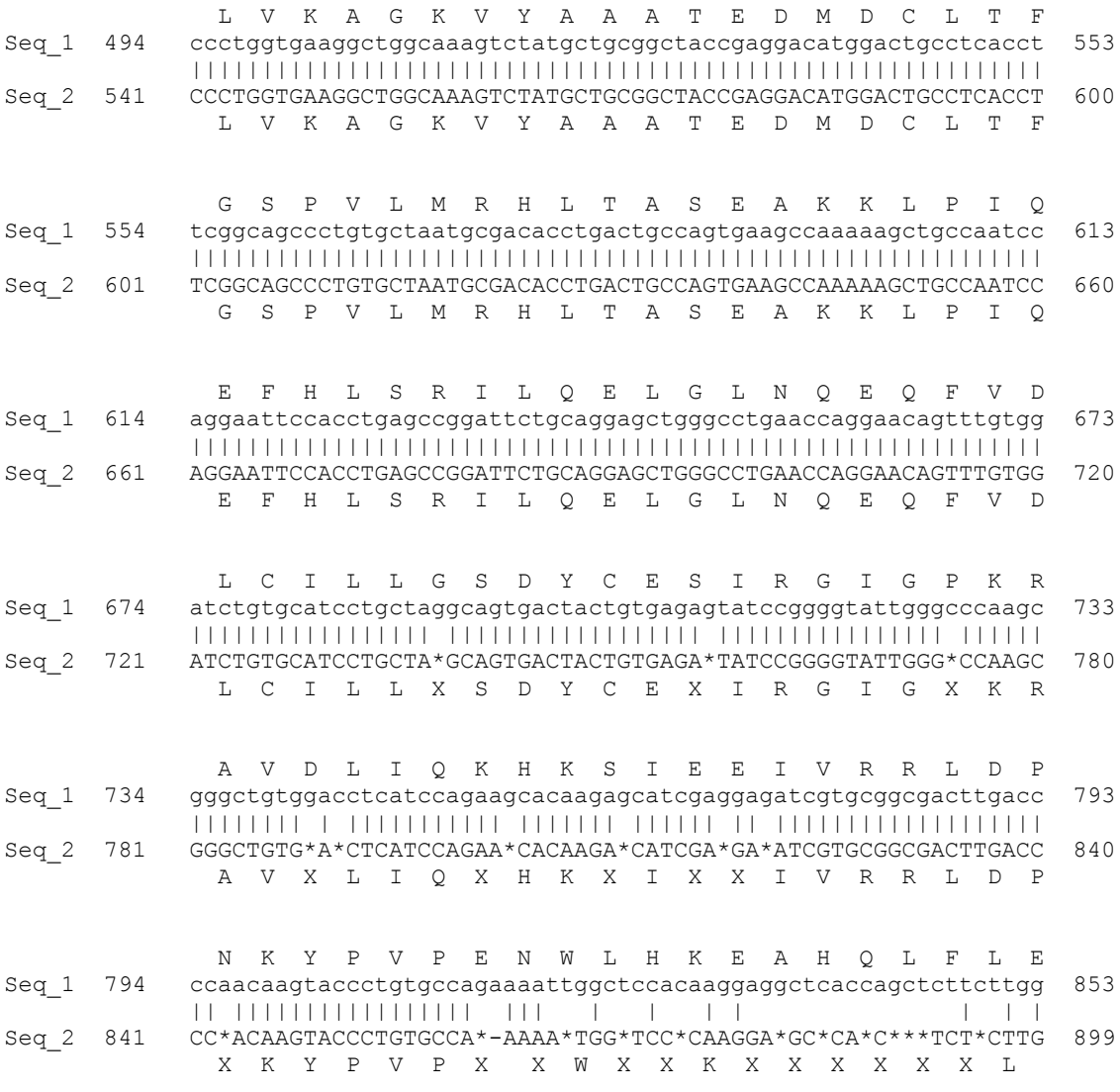


Figure A4: T7F (Forward) sequence alignment of L97P and hFEN1 mRNA: Both T7 forward and reverse sequencing reactions were performed by the University of Sheffield Medical School and subsequent alignment of the ORF of recombinant plasmid L97P and hFEN1 was done using Serial Cloner 2.1 (Serial Basics, Freeware). Mutation is in bold and highlighted, and nucleotides that could not be properly determined are denoted with a *. On occasion the identity of the nucleotide was determined by referring to the accompanying electropherogram.

```
Sequence 1: hFEN1 mRNA
Sequence 2: L97P reverse

Similarity : 859/1126 (76.29 %)


                D  G  K  P  P  Q  L  K  S  G  E  L  A  K  R  S  E  R  R  A
Seq_1  255   tgatggcaagccgccacagctcaagtcaggcgagctggccaaacgcagtgagcggcgggc  314
              |    |  |  |  |   | ||||  |||      | | |||||  ||  ||||||  |||
Seq_2  978   TT*A*G*CA**C**C***G*TCAA*TCA******C*G*CCAAA**CA**GAGCGG*GGG*  919
                X  X  X  X  X  X  S  X  X  X  X  X  P  X  X  X  S  X  G  X


                E  A  E  K  Q  L  Q  Q  A  Q  A  A  G  A  E  Q  E  V  E  K
Seq_1  315   tgaggcagagaagcagctgcagcaggctcaggctgctggggccgagcaggaggtggaaaa  374
              ||||||||   |||||  ||||||  ||  |||  |  |  ||||  |  ||||  |||||||  |||
Seq_2  918   TGAGGCAG***AGCAG*TGCAGC*GG*TCA*G**G*TGGG*C*GAGC*GGAGGTG**AAA  859
                R  Q  X  S  X  C  S  X  X  X  X  X  X  X  S  X  R  X  X  N


                F  T  K  R  L  V  K  V  T  K  Q  H  N  D  E  C  K  H  L  L
Seq_1  375   attcactaagcggctggtgaaggtcactaagcagcacaatgatgagtgcaaacatctgct  434
              ||  ||  ||||||||||||||||||||  |||||||||||||||||||||||||||||  ||
Seq_2  858   AT*CA*TAAGCGGCTGGTGAAGGTCA*TAAGCAGCACAATGATGAGTGCAAACAT*TG*-  800
                X  X  S  G  W  *  R  S  X  S  S  T  M  M  S  A  N  X  X


                S  L  M  G  I  P  Y  L  D  A  P  S  E  A  E  A  S  C  A  A
Seq_1  435   gagcctcatgggcatcccttatcttgatgcacccagtgaggcagaggccagctgtgctgc  494
                    |||||||  ||   |||||||||||||||||||||||||||  ||||||||||||||
Seq_2  799   TGA***CATGGGC*TC**-TATCTTGATGCACCCAGTGAGGCAGAG*CCAGCTGTGCTGC  741
              *  X  H  G  X  X   Y  L  D  A  P  S  E  A  E  X  S  C  A  A


                L  V  K  A  G  K  V  Y  A  A  A  T  E  D  M  D  C  L  T  F
Seq_1  495   cctggtgaaggctggcaaagtctatgctgcggctaccgaggacatggactgcctcacctt  554
              |||||||||||||||||||||||||  |||||||||||||||||||||||||||||  ||
Seq_2  740   CCTGGTGAAGGCTGGCAAAGTCTATG*TGCGGCTACCGAGGACATGGACTGCCTCAC*TT  681
                L  V  K  A  G  K  V  Y  X  A  A  T  E  D  M  D  C  L  X  F


                G  S  P  V  L  M  R  H  L  T  A  S  E  A  K  K  L  P  I  Q
Seq_1  555   cggcagccctgtgctaatgcgacacctgactgccagtgaagccaaaaaagctgccaatcca  614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  680   CGGCAGCCCTGTGCTAATGCGACACCTGACTGCCAGTGAAGCCAAAAAGCTGCCAATCCA  621
                G  S  P  V  L  M  R  H  L  T  A  S  E  A  K  K  L  P  I  Q


                E  F  H  L  S  R  I  L  Q  E  L  G  L  N  Q  E  Q  F  V  D
Seq_1  615   ggaattccacctgagccggattctgcaggagctgggcctgaaccaggaacagtttgtgga  674
              ||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  620   GGAATTCCAC*TGAGCCGGATTCTGCAGGAGCTGGGCCTGAACCAGGAACAGTTTGTGGA  561
                E  F  H  X  S  R  I  L  Q  E  L  G  L  N  Q  E  Q  F  V  D


                L  C  I  L  L  G  S  D  Y  C  E  S  I  R  G  I  G  P  K  R
Seq_1  675   tctgtgcatcctgctaggcagtgactactgtgagagtatccggggtattgggcccaagcg  734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  560   TCTGTGCATCCTGCTAGGCAGTGACTACTGTGAGAGTATCCGGGGTATTGGGCCCAAGCG  501
                L  C  I  L  L  G  S  D  Y  C  E  S  I  R  G  I  G  P  K  R


                A  V  D  L  I  Q  K  H  K  S  I  E  E  I  V  R  R  L  D  P
Seq_1  735   ggctgtggacctcatccagaagcacaagagcatcgaggagatcgtgcggcgacttgaccc  794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  500   GGCTGTGGACCTCATCCAGAAGCACAAGAGCATCGAGGAGATCGTGCGGCGACTTGACCC  441
                A  V  D  L  I  Q  K  H  K  S  I  E  E  I  V  R  R  L  D  P
```

```
                N  K  Y  P  V  P  E  N  W  L  H  K  E  A  H  Q  L  F  L  E
Seq_1  795   caacaagtaccctgtgccagaaaattggctccacaaggaggctcaccagctcttcttgga  854
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  440   CAACAAGTACCCTGTGCCAGAAAATTGGCTCCACAAGGAGGCTCACCAGCTCTTCTTGGA  381
                N  K  Y  P  V  P  E  N  W  L  H  K  E  A  H  Q  L  F  L  E

                P  E  V  L  D  P  E  S  V  E  L  K  W  S  E  P  N  E  E  E
Seq_1  855   acctgaggtgctggacccagagtctgtggagctgaagtggagcgagccaaatgaagaaga  914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  380   ACCTGAGGTGCTGGACCCAGAGTCTGTGGAGCTGAAGTGGAGCGAGCCAAATGAAGAAGA  321
                P  E  V  L  D  P  E  S  V  E  L  K  W  S  E  P  N  E  E  E

                L  I  K  F  M  C  G  E  K  Q  F  S  E  E  R  I  R  S  G  V
Seq_1  915   gctgatcaagttcatgtgtggtgaaaagcagttctctgaggagcgaatccgcagtggggt  974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  320   GCTGATCAAGTTCATGTGTGGTGAAAAGCAGTTCTCTGAGGAGCGAATCCGCAGTGGGGT  261
                L  I  K  F  M  C  G  E  K  Q  F  S  E  E  R  I  R  S  G  V

                K  R  L  S  K  S  R  Q  G  S  T  Q  G  R  L  D  D  F  F  K
Seq_1  975   caagaggctgagtaagagccgccaaggcagcacccagggccgcctggatgatttcttcaa  1034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  260   CAAGAGGCTGAGTAAGAGCCGCCAAGGCAGCACCCAGGGCCGCCTGGATGATTTCTTCAA  201
                K  R  L  S  K  S  R  Q  G  S  T  Q  G  R  L  D  D  F  F  K

                V  T  G  S  L  S  S  A  K  R  K  E  P  E  P  K  G  S  T  K
Seq_1  1035  ggtgaccggctcactctcttcagctaagcgcaaggagccagaacccaagggatccactaa  1094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  200   GGTGACCGGCTCACTCTCTTCAGCTAAGCGCAAGGAGCCAGAACCCAAGGGATCCACTAA  141
                V  T  G  S  L  S  S  A  K  R  K  E  P  E  P  K  G  S  T  K

                K  K  A  K  T  G  A  A  G  K  F  K  R  G  K
Seq_1  1095  gaagaaggcaaagactggggcagcagggaagtttaaaaggggaaaa--------------  1140
             ||||||||||||||||||||||||||||||||||||||||||||
Seq_2  140   GAAGAAGGCAAAGACTGGGGCAGCAGGGAAGTTTAAAAGGGGAAAACATCATCATCATCA  81
                K  K  A  K  T  G  A  A  G  K  F  K  R  G  K  H  H  H  H  H
```

Figure A5: T7R (Reverse) sequence alignment of L97P and hFEN1 mRNA: Both T7 forward and reverse sequencing reactions were performed by the University of Sheffield Medical School and subsequent alignment of the ORF of recombinant plasmid L97P and hFEN1 was done using Serial Cloner 2.1 (Serial Basics, Freeware). Mutation is in bold and highlighted, and nucleotides that could not be properly determined are denoted with a *. On occasion the identity of the nucleotide was determined by referring to the accompanying electropherogram.

```
Sequence 1: hFEN1 mRNA
Sequence 2: L111P forward

Similarity : 775/1050 (73.81 %)

                                                                M  G  I  Q
Seq_1  1     ----------------------------------------------------atgggaattca  11
                                                                |||||||||||
Seq_2  1     **********************T***TTAACTTTAAGAAGGAGA*ATACCATGGGAATTCA  60
              X  X  X  X  X  X  X  X  L  T  L  R  R  R  X  T  M  G  I  Q


              G  L  A  K  L  I  A  D  V  A  P  S  A  I  R  E  N  D  I  K
Seq_1  12    aggcctggccaaactaattgctgatgtggccccagtgccatccgggagaatgacatcaa  71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  61    AGGCCTGGCCAAACTAATTGCTGATGTGGCCCCCAGTGCCATCCGGGAGAATGACATCAA  120
              G  L  A  K  L  I  A  D  V  A  P  S  A  I  R  E  N  D  I  K


              S  Y  F  G  R  K  V  A  I  D  A  S  M  S  I  Y  Q  F  L  I
Seq_1  72    gagctactttggccgtaaggtggccattgatgcctctatgagcatttatcagttcctgat  131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  121   GAGCTACTTTGGCCGTAAGGTGGCCATTGATGCCTCTATGAGCATTTATCAGTTCCTGAT  180
              S  Y  F  G  R  K  V  A  I  D  A  S  M  S  I  Y  Q  F  L  I


              A  V  R  Q  G  G  D  V  L  Q  N  E  E  G  E  T  T  S  H  L
Seq_1  132   tgctgttcgccagggtggggatgtgctgcagaatgaggagggtgagaccaccagccacct  191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  181   TGCTGTTCGCCAGGGTGGGGATGTGCTGCAGAATGAGGAGGGTGAGACCACCAGCCACCT  240
              A  V  R  Q  G  G  D  V  L  Q  N  E  E  G  E  T  T  S  H  L


              M  G  M  F  Y  R  T  I  R  M  M  E  N  G  I  K  P  V  Y  V
Seq_1  192   gatgggcatgttctaccgcaccattcgcatgatggagaacggcatcaagcccgtgtatgt  251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  241   GATGGGCATGTTCTACCGCACCATTCGCATGATGGAGAACGGCATCAAGCCCGTGTATGT  300
              M  G  M  F  Y  R  T  I  R  M  M  E  N  G  I  K  P  V  Y  V


              F  D  G  K  P  P  Q  L  K  S  G  E  L  A  K  R  S  E  R  R
Seq_1  252   ctttgatggcaagccgccacagctcaagtcaggcgagctggccaaacgcagtgagcggcg  311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  301   CTTTGATGGCAAGCCGCCACAGCTCAAGTCAGGCGAGCTGGCCAAACGCAGTGAGCGGCG  360
              F  D  G  K  P  P  Q  L  K  S  G  E  L  A  K  R  S  E  R  R


              A  E  A  E  K  Q  L  Q  Q  A  Q  A  A  G  A  E  Q  E  V  E
Seq_1  312   ggctgaggcagagaagcagct gcagcaggctcaggctgctggggccgagcaggaggtgga  371
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Seq_2  361   GGCTGAGGCAGAGAAGCAGCCGCAGCAGGCTCAGGCTGCTGGGGCCGAGCAGGAGGTGGA  420
              A  E  A  E  K  Q  P  Q  Q  A  Q  A  A  G  A  E  Q  E  V  E


              K  F  T  K  R  L  V  K  V  T  K  Q  H  N  D  E  C  K  H  L
Seq_1  372   aaaattcactaagcggctggtgaaggtcactaagcagcacaatgatgagtgcaaacatct  431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  421   AAAATTCACTAAGCGGCTGGTGAAGGTCACTAAGCAGCACAATGATGAGTGCAAACATCT  480
              K  F  T  K  R  L  V  K  V  T  K  Q  H  N  D  E  C  K  H  L


              L  S  L  M  G  I  P  Y  L  D  A  P  S  E  A  E  A  S  C  A
Seq_1  432   gctgagcctcatgggcatcccttatcttgatgcacccagtgaggcagaggccagctgtgc  491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  481   GCTGAGCCTCATGGGCATCCCTTATCTTGATGCACCCAGTGAGGCAGAGGCCAGCTGTGC  540
              L  S  L  M  G  I  P  Y  L  D  A  P  S  E  A  E  A  S  C  A
```

```
                A  L  V  K  A  G  K  V  Y  A  A  A  T  E  D  M  D  C  L  T
Seq_1  492   tgccctggtgaaggctggcaaagtctatgctgcggctaccgaggacatggactgcctcac  551
             ||||||||||||||||||||||||||||||||||      ||||||||||||||||||
Seq_2  541   TGCCCTGGTGAAGGCTGGCAAAGTCTATGCTGCG******GAGGA*ATGGACTGCCTCAC  600
                A  L  V  K  A  G  K  V  Y  A  A  X  X  E  X  M  D  C  L  T

                F  G  S  P  V  L  M  R  H  L  T  A  S  E  A  K  K  L  P  I
Seq_1  552   cttcggcagccctgtgctaatgcgacacctgactgccagtgaagccaaaaagctgccaat  611
             ||||||||||||||||||||||||||||||||||||||||||||||||| | ||||||||||
Seq_2  601   CTTCGGCAGCCCTGTGCTAATGCGACACCTGACTGCCAGTGAAGCC*A*AAGCTGCCAAT  660
                F  G  S  P  V  L  M  R  H  L  T  A  S  E  A  X  K  L  P  I

                Q  E  F  H  L  S  R  I  L  Q  E  L  G  L  N  Q  E  Q  F  V
Seq_1  612   ccaggaattccacctgagccggattctgcaggagctgggcctgaaccaggaacagtttgt  671
             ||||||||||||||||  ||  |||||||||  ||  ||||||||||  ||||||||||||||
Seq_2  661   CCAGGAATTCCACCTGA*CC*GATTCTGCA*GA*CTGGGCCTGAA*CAGGAACAGTTTGT  720
                Q  E  F  H  L  X  X  I  L  X  X  L  G  L  X  Q  E  Q  F  V

                D  L  C  I  L  L  G  S  D  Y  C  E  S  I  R  G  I  G  P  K
Seq_1  672   ggatctgtgcatcctgctaggcagtgactactgtgagagtatccggggtattgggccc-a  730
             ||||||||||||||||||||||||| |||||||||||||||||||||||||||||| |
Seq_2  721   GGATCTGTGCATCCTGCTAGGCAGTG*CTACTGTGAGAGTATCCGGGGTATTGGGCCC*A  780
                D  L  C  I  L  L  G  S  X  Y  C  E  S  I  R  G  I  G  P  X
```

Figure A6: T7F sequence alignment of L111P and hFEN1 mRNA: Both T7 forward and reverse sequencing reactions were performed by the University of Sheffield Medical School and subsequent alignment of the ORF of recombinant plasmid L111P and hFEN1 was done using Serial Cloner 2.1 (Serial Basics, Freeware). Mutation is in bold and highlighted, and nucleotides that could not be properly determined are denoted with a *. On occasion the identity of the nucleotide was determined by referring to the accompanying electropherogram.

```
Sequence 1: hFEN1 mRNA
Sequence 2: L111P reverse

Similarity : 879/1131 (77.72 %)

              L  K  S  G  E  L  A  K  R  S  E  R  R  A  E  A  E  K  Q  L
Seq_1  274   ctcaagtcaggcgagctggccaaacgcagtgagcggcgggctgaggcagagaagcagct g  333
                 |            | | ||||||||||||  |  ||| ||||||||  |||||||| |
Seq_2  966   TCAA*TC*G*CGAGC*GG*C*AAACGCAGTGAG*G**GGG*TGAGGCAG*GAAGCAGC*G  907
               Q  X  X  R  X  X  X  N  A  V  X  X  G  X  R  Q  X  S  S  X

              Q  Q  A  Q  A  A  G  A  E  Q  E  V  E  K  F  T  K  R  L  V
Seq_1  334   cagcaggctcaggctgctggggccgagcaggaggtggaaaaattcactaagcggctggtg  393
              ||||  ||  |||||||||||||  ||||||  ||||||||||||||||||  |||||  | |||||
Seq_2  906   CAGC*GG*TCAGGCTGCTGGG*CCGAGC*GGAGGTGGAAAAATTCA*TAAGC*G*TGGTG  847
               S  X  X  R  L  L  X  P  S  X  R  W  K  N  S  X  S  X  W  *

              K  V  T  K  Q  H  N  D  E  C  K  H  L  L  S  L  M  G  I  P
Seq_1  394   aaggtcactaagcagcacaatgatgagtgcaaacatctgctgagcctcatgggcatccct  453
              |||||||  |||||||||||||||||||||||||  | ||  ||||||||||||||  | ||
Seq_2  846   AAGGTCA*TAAGCAGCACAATGATGAGTGCAAAC*T*TG*TGAGCCTCATGGGC*T*CC*  787
               R  S  X  S  S  T  M  M  S  A  N  X  X  *  A  S  W  X  X  X

              Y   L  D  A  P  S  E  A  E  A  S  C  A  A  L  V  K  A  G  K
Seq_1  454   ta-tcttgatgcacccagtgaggcagaggccagctgtgctgccctggtgaaggctggcaa  512
              |  |||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Seq_2  786   T**TCTTGATGCACCCAGTGAGGCAGAG*CCAGCTGTGCTGCCCTGGTGAAGGCTGGCAA  727
               X  L  D  A  P  S  E  A  E  X  S  C  A  A  L  V  K  A  G  K

              V  Y  A  A  A  T  E  D  M  D  C  L  T  F  G  S  P  V  L  M
Seq_1  513   agtctatgctgcggctaccgaggacatggactgcctcaccttcggcagccctgtgctaat  572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  726   AGTCTATGCTGCGGCTACCGAGGACATGGACTGCCTCACCTTCGGCAGCCCTGTGCTAAT  667
               V  Y  A  A  A  T  E  D  M  D  C  L  T  F  G  S  P  V  L  M

              R  H  L  T  A  S  E  A  K  K  L  P  I  Q  E  F  H  L  S  R
Seq_1  573   gcgacacctgactgccagtgaagccaaaaagctgccaatccaggaattccacctgagccg  632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||
Seq_2  666   GCGACACCTGACTGCCAGTGAAGCCAAAAAGCTGCCAATCCAGGAATTCCAC*TGAGCCG  607
               R  H  L  T  A  S  E  A  K  K  L  P  I  Q  E  F  H  X  S  R

              I  L  Q  E  L  G  L  N  Q  E  Q  F  V  D  L  C  I  L  L  G
Seq_1  633   gattctgcaggagctgggcctgaaccaggaacagtttgtggatctgtgcatcctgctagg  692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  606   GATTCTGCAGGAGCTGGGCCTGAACCAGGAACAGTTTGTGGATCTGTGCATCCTGCTAGG  547
               I  L  Q  E  L  G  L  N  Q  E  Q  F  V  D  L  C  I  L  L  G

              S  D  Y  C  E  S  I  R  G  I  G  P  K  R  A  V  D  L  I  Q
Seq_1  693   cagtgactactgtgagagtatccggggtattgggcccaagcgggctgtggacctcatcca  752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  546   CAGTGACTACTGTGAGAGTATCCGGGGTATTGGGCCCAAGCGGGCTGTGGACCTCATCCA  487
               S  D  Y  C  E  S  I  R  G  I  G  P  K  R  A  V  D  L  I  Q

              K  H  K  S  I  E  E  I  V  R  R  L  D  P  N  K  Y  P  V  P
Seq_1  753   gaagcacaagagcatcgaggagatcgtgcggcgacttgaccccaacaagtaccctgtgcc  812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  486   GAAGCACAAGAGCATCGAGGAGATCGTGCGGCGACTTGACCCCAACAAGTACCCTGTGCC  427
               K  H  K  S  I  E  E  I  V  R  R  L  D  P  N  K  Y  P  V  P
```

```
                E  N  W  L  H  K  E  A  H  Q  L  F  L  E  P  E  V  L  D  P
Seq_1  813   agaaaattggctccacaaggaggctcaccagctcttcttggaacctgaggtgctggaccc  872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  426   AGAAAATTGGCTCCACAAGGAGGCTCACCAGCTCTTCTTGGAACCTGAGGTGCTGGACCC  367
                E  N  W  L  H  K  E  A  H  Q  L  F  L  E  P  E  V  L  D  P

                E  S  V  E  L  K  W  S  E  P  N  E  E  E  L  I  K  F  M  C
Seq_1  873   agagtctgtggagctgaagtggagcgagccaaatgaagaagagctgatcaagttcatgtg  932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  366   AGAGTCTGTGGAGCTGAAGTGGAGCGAGCCAAATGAAGAAGAGCTGATCAAGTTCATGTG  307
                E  S  V  E  L  K  W  S  E  P  N  E  E  E  L  I  K  F  M  C

                G  E  K  Q  F  S  E  E  R  I  R  S  G  V  K  R  L  S  K  S
Seq_1  933   tggtgaaaagcagttctctgaggagcgaatccgcagtggggtcaagaggctgagtaagag  992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  306   TGGTGAAAAGCAGTTCTCTGAGGAGCGAATCCGCAGTGGGGTCAAGAGGCTGAGTAAGAG  247
                G  E  K  Q  F  S  E  E  R  I  R  S  G  V  K  R  L  S  K  S

                R  Q  G  S  T  Q  G  R  L  D  D  F  F  K  V  T  G  S  L  S
Seq_1  993   ccgccaaggcagcacccagggccgcctggatgatttcttcaaggtgaccggctcactctc  1052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  246   CCGCCAAGGCAGCACCCAGGGCCGCCTGGATGATTTCTTCAAGGTGACCGGCTCACTCTC  187
                R  Q  G  S  T  Q  G  R  L  D  D  F  F  K  V  T  G  S  L  S

                S  A  K  R  K  E  P  E  P  K  G  S  T  K  K  K  A  K  T  G
Seq_1  1053  ttcagctaagcgcaaggagccagaacccaagggatccactaagaagaaggcaaagactgg  1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  186   TTCAGCTAAGCGCAAGGAGCCAGAACCCAAGGGATCCACTAAGAAGAAGGCAAAGACTGG  127
                S  A  K  R  K  E  P  E  P  K  G  S  T  K  K  K  A  K  T  G

                A  A  G  K  F  K  R  G  K                    *
Seq_1  1113  ggcagcagggaagtttaaaaggggaaaa------------------taa-----------  1143
              ||||||||||||||||||||||||||                    |||
Seq_2  126   GGCAGCAGGGAAGTTTAAAAGGGGAAAACATCATCATCATCATCACTAAGCTTGCGGCCG  67
                A  A  G  K  F  K  R  G  K  H  H  H  H  H  H  *  A  C  G  R
```

Figure A7: T7R sequence alignment of L111P and hFEN1 mRNA: Both T7 forward and reverse sequencing reactions were performed by the University of Sheffield Medical School and subsequent alignment of the ORF of recombinant plasmid L111P and hFEN1 was done using Serial Cloner 2.1 (Serial Basics, Freeware). Mutation is in bold and highlighted, and nucleotides that could not be properly determined are denoted with a *. On occasion the identity of the nucleotide was determined by referring to the accompanying electropherogram.

```
Sequence 1: hFEN1 mRNA
Sequence 2: L130P forward

Similarity : 865/1132 (76.41 %)

                                                          M  G  I  Q  G
Seq_1  1     ------------------------------------------------atgggaattcaagg  14
                                                             |||||||||||||
Seq_2  1     *******C*******ATTT***TTAACTTTAAGAAGGAGATATACCATGGGAATTCAAGG  60
               X  X  X  X  X  X  X  *  L  *  E  G  D  I  P  W  E  F  K  X


              L  A  K   L  I  A  D  V  A  P  S  A  I  R  E  N  D  I  K  S
Seq_1  15    cctggcca-aactaattgctgatgtggccccccagtgccatccgggagaatgacatcaaga  73
             |       ||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  61    C********AACTAATTGCTGATGTGGCCCCCAGTGCCATCCGGGAGAATGACATCAAGA  120
               X  X  X  L  I  A  D  V  A  P  S  A  I  R  E  N  D  I  K  S


                 Y  F  G  R  K  V  A  I  D  A  S  M  S  I  Y  Q  F  L  I  A
Seq_1  74    gctactttggccgtaaggtggccattgatgcctctatgagcatttatcagttcctgattg  133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  121   GCTACTTTGGCCGTAAGGTGGCCATTGATGCCTCTATGAGCATTTATCAGTTCCTGATTG  180
               Y  F  G  R  K  V  A  I  D  A  S  M  S  I  Y  Q  F  L  I  A


                 V  R  Q  G  G  D  V  L  Q  N  E  E  G  E  T  T  S  H  L  M
Seq_1  134   ctgttcgccagggtggggatgtgctgcagaatgaggagggtgagaccaccagccacctga  193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  181   CTGTTCGCCAGGGTGGGGATGTGCTGCAGAATGAGGAGGGTGAGACCACCAGCCACCTGA  240
               V  R  Q  G  G  D  V  L  Q  N  E  E  G  E  T  T  S  H  L  M


                 G  M  F  Y  R  T  I  R  M  M  E  N  G  I  K  P  V  Y  V  F
Seq_1  194   tgggcatgttctaccgcaccattcgcatgatggagaacggcatcaagcccgtgtatgtct  253
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  241   TGGGCATGTTCTACCGCACCATTCGCATGATGGAGAACGGCATCAAGCCCGTGTATGTCT  300
               G  M  F  Y  R  T  I  R  M  M  E  N  G  I  K  P  V  Y  V  F


                 D  G  K  P  P  Q  L  K  S  G  E  L  A  K  R  S  E  R  R  A
Seq_1  254   ttgatggcaagccgccacagctcaagtcaggcgagctggccaaacgcagtgagcggcggg  313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  301   TTGATGGCAAGCCGCCACAGCTCAAGTCAGGCGAGCTGGCCAAACGCAGTGAGCGGCGGG  360
               D  G  K  P  P  Q  L  K  S  G  E  L  A  K  R  S  E  R  R  A


                 E  A  E  K  Q  L  Q  Q  A  Q  A  A  G  A  E  Q  E  V  E  K
Seq_1  314   ctgaggcagagaagcagctgcagcaggctcaggctgctggggccgagcaggaggtggaaa  373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  361   CTGAGGCAGAGAAGCAGCTGCAGCAGGCTCAGGCTGCTGGGGCCGAGCAGGAGGTGGAAA  420
               E  A  E  K  Q  L  Q  Q  A  Q  A  A  G  A  E  Q  E  V  E  K


                 F  T  K  R  L  V  K  V  T  K  Q  H  N  D  E  C  K  H  L  L
Seq_1  374   aattcactaagcggct ggtgaaggtcactaagcagcacaatgatgagtgcaaacatctgc  433
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Seq_2  421   AATTCACTAAGCGGCCGGTGAAGGTCACTAAGCAGCACAATGATGAGTGCAAACATCTGC  480
               F  T  K  R  P  V  K  V  T  K  Q  H  N  D  E  C  K  H  L  L


                 S  L  M  G  I  P  Y  L  D  A  P  S  E  A  E  A  S  C  A  A
Seq_1  434   tgagcctcatgggcatcccttatcttgatgcacccagtgaggcagaggccagctgtgctg  493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  481   TGAGCCTCATGGGCATCCCTTATCTTGATGCACCCAGTGAGGCAGAGGCCAGCTGTGCTG  540
               S  L  M  G  I  P  Y  L  D  A  P  S  E  A  E  A  S  C  A  A
```

```
                L  V  K  A  G  K  V  Y  A  A  A  T  E  D  M  D  C  L  T  F
Seq_1  494   ccctggtgaaggctggcaaagtctatgctgcggctaccgaggacatggactgcctcacct  553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  541   CCCTGGTGAAGGCTGGCAAAGTCTATGCTGCGGCTACCGAGGACATGGACTGCCTCACCT  600
                L  V  K  A  G  K  V  Y  A  A  A  T  E  D  M  D  C  L  T  F

                G  S  P  V  L  M  R  H  L  T  A  S  E  A  K  K  L  P  I  Q
Seq_1  554   tcggcagccctgtgctaatgcgacacctgactgccagtgaagccaaaaagctgccaatcc  613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  601   TCGGCAGCCCTGTGCTAATGCGACACCTGACTGCCAGTGAAGCCAAAAAGCTGCCAATCC  660
                G  S  P  V  L  M  R  H  L  T  A  S  E  A  K  K  L  P  I  Q

                E  F  H  L  S  R  I  L  Q  E  L  G  L  N  Q  E  Q  F  V  D
Seq_1  614   aggaattccacctgagccggattctgcaggagctgggcctgaaccaggaacagtttgtgg  673
              |||||||||||||| |||||||||||| ||||||||||||||||||||||||||||||
Seq_2  661   AGGAATTCCACCTGA*CCGGATTCTGCA*GAGCTGGGCCTGAACCAGGAACAGTTTGTGG  720
                E  F  H  L  X  R  I  L  X  E  L  G  L  N  Q  E  Q  F  V  D

                L  C  I  L  L  G  S  D  Y  C  E  S  I  R  G  I  G  P  K  R
Seq_1  674   atctgtgcatcctgctaggcagtgactactgtgagagtatccggggtattgggcccaagc  733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Seq_2  721   ATCTGTGCATCCTGCTAGGCAGTGACTACTGTGAGAGTATCCGGGGTATTGGGCCC*AGC  780
                L  C  I  L  L  G  S  D  Y  C  E  S  I  R  G  I  G  P  X  R

                A  V  D  L  I  Q  K  H  K  S  I  E  E  I  V  R  R  L  D  P
Seq_1  734   gggctgtggacctcatccagaagcacaagagcatcgaggagatcgtgcggcgacttgacc  793
              |||||||| |||||||||| || | | ||| |||| |||| ||||||||||||||||||
Seq_2  781   GGGCTGTG*ACCTCATCCA*AA*C*C*AGA*CATC*AGGA*ATCGTGCGGCGACTTGACC  840
                A  V  X  L  I  X  X  X  X  X  I  X  X  I  V  R  R  L  D  P

                N  K  Y  P  V  P  E  N  W  L  H  K  E  A  H  Q  L  F  L  E
Seq_1  794   ccaacaagtaccctgtgccagaaaattggctccacaaggaggctcaccagctcttcttgg  853
              || ||| |||||||||||||| |||||||||||| || | ||||||||||||||||||||
Seq_2  841   CC*ACA*GTACCCTGTGCCA*AAAATTGGCTCC*CA*G*AGGCTCACCAGCTCTTCTTGG  900
                X  X  Y  P  V  P  X  N  W  L  X  X  X  A  H  Q  L  F  L  E
```

Figure A8: T7F sequence alignment of L130P and hFEN1 mRNA: Both T7 forward and reverse sequencing reactions were performed by the University of Sheffield Medical School and subsequent alignment of the ORF of recombinant plasmid L130P and hFEN1 was done using Serial Cloner 2.1 (Serial Basics, Freeware). Mutation is in bold and highlighted, and nucleotides that could not be properly determined are denoted with a *. On occasion the identity of the nucleotide was determined by referring to the accompanying electropherogram.

```
Sequence 1: hFEN1 mRNA
Sequence 2: L130P reverse

Similarity : 881/1129 (78.03 %)


                D  G  K  P  P  Q  L  K  S  G  E  L  A  K  R  S  E  R  R  A
Seq_1  254   ttgatggcaagccgccacagctcaagtcaggcgagctggccaaacgcagtgagcggcggg  313
              |       |     |        |||||||     ||| ||  ||||||||||||||||||
Seq_2  985   T*A*G***A*C*GCC*C******AAGTCAG*****CTG*CC*AACGCAGTGAGCGGCGG*  926
              X  X  X  X  A  X  X  X  S  Q  X  X  X  X  N  A  V  S  G  X


                E  A  E  K  Q  L  Q  Q  A  Q  A  A  G  A  E  Q  E  V  E  K
Seq_1  314   ctgaggcagagaagcagctgcagcaggctcaggctgctggggccgagcaggaggtggaaa  373
               |||  ||||||||  |||  |||||||||  ||||||||||||||||||  ||||||||||||||
Seq_2  925   *TGA*GCAGAGAA*CAG*TGCAGCAGG*TCAGGCTGCTGGGGCCGA*CAGGAGGTGGAAA  866
              X  X  Q  R  X  X  C  S  R  X  R  L  L  G  P  X  R  R  W  K


                F  T  K  R  L  V  K  V  T  K  Q  H  N  D  E  C  K  H  L  L
Seq_1  374   aattcactaagcggctggtgaaggtcactaagcagcacaatgatgagtgcaaacatctgc  433
              ||||||  |||||   |  ||||||||||  ||||||||||||||||||||||||||  |  ||
Seq_2  865   AATTCA*TAAGC**CCGGTGAAGGTCA*TAAGCAGCACAATGATGAGTGCAAAC*T*TG*  806
              N  S  X  S  X  R  *  R  S  X  S  S  T  M  M  S  A  N  X  X


                S  L  M  G  I  P  Y     L  D  A  P  S  E  A  E  A  S  C  A
Seq_1  434   tgagcctcatgggcatccctt-atcttgatgcacccagtgaggcagaggccagctgtgct  492
              |||||  ||||||||||  ||  |  ||||||||||||||||||||||||||||||||||||
Seq_2  805   TGAGC*TCATGGGCAT*CC*T*ATCTTGATGCACCCAGTGAGGCAGAGGCCAGCTGTGCT  746
              *  X  S  W  A  X  X  X  L  D  A  P  S  E  A  E  A  S  C  A


              A  L  V  K  A  G  K  V  Y  A  A  A  T  E  D  M  D  C  L  T
Seq_1  493   gccctggtgaaggctggcaaagtctatgctgcggctaccgaggacatggactgcctcacc  552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  745   GCCCTGGTGAAGGCTGGCAAAGTCTATGCTGCGGCTACCGAGGACATGGACTGCCTCACC  686
              A  L  V  K  A  G  K  V  Y  A  A  A  T  E  D  M  D  C  L  T


              F  G  S  P  V  L  M  R  H  L  T  A  S  E  A  K  K  L  P  I
Seq_1  553   ttcggcagccctgtgctaatgcgacacctgactgccagtgaagccaaaaaagctgccaatc  612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  685   TTCGGCAGCCCTGTGCTAATGCGACACCTGACTGCCAGTGAAGCCAAAAAGCTGCCAATC  626
              F  G  S  P  V  L  M  R  H  L  T  A  S  E  A  K  K  L  P  I


              Q  E  F  H  L  S  R  I  L  Q  E  L  G  L  N  Q  E  Q  F  V
Seq_1  613   caggaattccacctgagccggattctgcaggagctgggcctgaaccaggaacagtttgtg  672
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  625   CAGGAATTCCAC*TGAGCCGGATTCTGCAGGAGCTGGGCCTGAACCAGGAACAGTTTGTG  566
              Q  E  F  H  X  S  R  I  L  Q  E  L  G  L  N  Q  E  Q  F  V


              D  L  C  I  L  L  G  S  D  Y  C  E  S  I  R  G  I  G  P  K
Seq_1  673   gatctgtgcatcctgctaggcagtgactactgtgagagtatccggggtattgggcccaag  732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  565   GATCTGTGCATCCTGCTAGGCAGTGACTACTGTGAGAGTATCCGGGGTATTGGGCCCAAG  506
              D  L  C  I  L  L  G  S  D  Y  C  E  S  I  R  G  I  G  P  K


              R  A  V  D  L  I  Q  K  H  K  S  I  E  E  I  V  R  R  L  D
Seq_1  733   cgggctgtggacctcatccagaagcacaagagcatcgaggagatcgtgcggcgacttgac  792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  505   CGGGCTGTGGACCTCATCCAGAAGCACAAGAGCATCGAGGAGATCGTGCGGCGACTTGAC  446
              R  A  V  D  L  I  Q  K  H  K  S  I  E  E  I  V  R  R  L  D
```

```
            P  N  K  Y  P  V  P  E  N  W  L  H  K  E  A  H  Q  L  F  L
Seq_1  793  cccaacaagtaccctgtgccagaaaattggctccacaaggaggctcaccagctcttcttg  852
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  445  CCCAACAAGTACCCTGTGCCAGAAAATTGGCTCCACAAGGAGGCTCACCAGCTCTTCTTG  386
            P  N  K  Y  P  V  P  E  N  W  L  H  K  E  A  H  Q  L  F  L

            E  P  E  V  L  D  P  E  S  V  E  L  K  W  S  E  P  N  E  E
Seq_1  853  gaacctgaggtgctggacccagagtctgtggagctgaagtggagcgagccaaatgaagaa  912
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  385  GAACCTGAGGTGCTGGACCCAGAGTCTGTGGAGCTGAAGTGGAGCGAGCCAAATGAAGAA  326
            E  P  E  V  L  D  P  E  S  V  E  L  K  W  S  E  P  N  E  E

            E  L  I  K  F  M  C  G  E  K  Q  F  S  E  E  R  I  R  S  G
Seq_1  913  gagctgatcaagttcatgtgtggtgaaaagcagttctctgaggagcgaatccgcagtggg  972
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  325  GAGCTGATCAAGTTCATGTGTGGTGAAAAGCAGTTCTCTGAGGAGCGAATCCGCAGTGGG  266
            E  L  I  K  F  M  C  G  E  K  Q  F  S  E  E  R  I  R  S  G

            V  K  R  L  S  K  S  R  Q  G  S  T  Q  G  R  L  D  D  F  F
Seq_1  973  gtcaagaggctgagtaagagccgccaaggcagcacccagggccgcctggatgatttcttc  1032
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  265  GTCAAGAGGCTGAGTAAGAGCCGCCAAGGCAGCACCCAGGGCCGCCTGGATGATTTCTTC  206
            V  K  R  L  S  K  S  R  Q  G  S  T  Q  G  R  L  D  D  F  F

            K  V  T  G  S  L  S  S  A  K  R  K  E  P  E  P  K  G  S  T
Seq_1  1033 aaggtgaccggctcactctcttcagctaagcgcaaggagccagaacccaagggatccact  1092
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  205  AAGGTGACCGGCTCACTCTCTTCAGCTAAGCGCAAGGAGCCAGAACCCAAGGGATCCACT  146
            K  V  T  G  S  L  S  S  A  K  R  K  E  P  E  P  K  G  S  T

            K  K  K  A  K  T  G  A  A  G  K  F  K  R  G  K
Seq_1  1093 aagaagaaggcaaagactggggcagcagggaagtttaaaaggggaaaa----------  1140
            ||||||||||||||||||||||||||||||||||||||||||||||
Seq_2  145  AAGAAGAAGGCAAAGACTGGGGCAGCAGGGAAGTTTAAAAGGGGAAAACATCATCATCAT  86
            K  K  K  A  K  T  G  A  A  G  K  F  K  R  G  K  H  H  H  H
```

Figure A9: T7R sequence alignment of L130P and hFEN1 mRNA: Both T7 forward and reverse sequencing reactions were performed by the University of Sheffield Medical School and subsequent alignment of the ORF of recombinant plasmid L130P and hFEN1 was done using Serial Cloner 2.1 (Serial Basics, Freeware). Mutation is in bold and highlighted, and nucleotides that could not be properly determined are denoted with a *. On occasion the identity of the nucleotide was determined by referring to the accompanying electropherogram.

Published online 8 February 2012 *Nucleic Acids Research, 2012, Vol. 40, No. 10* ***4507–4519***
doi:10.1093/nar/gks051

Flap endonucleases pass 5′-flaps through a flexible arch using a disorder-thread-order mechanism to confer specificity for free 5′-ends

Nikesh Patel[1], John M. Atack[1], L. David Finger[1], Jack C. Exell[1], Peter Thompson[1], Susan Tsutakawa[2], John A. Tainer[2,3,4], David M. Williams[1] and Jane A. Grasby[1,*]

[1]Centre for Chemical Biology, Department of Chemistry, Krebs Institute, University of Sheffield, Sheffield S3 7HF, UK, [2]Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [3]Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA, USA and [4]Skaggs Institute for Chemical Biology, La Jolla, CA 92037, USA

Received October 21, 2011; Revised January 9, 2012; Accepted January 13, 2012

ABSTRACT

Flap endonucleases (FENs), essential for DNA replication and repair, recognize and remove RNA or DNA 5′-flaps. Related to FEN specificity for substrates with free 5′-ends, but controversial, is the role of the helical arch observed in varying conformations in substrate-free FEN structures. Conflicting models suggest either 5′-flaps thread through the arch, which when structured can only accommodate single-stranded (ss) DNA, or the arch acts as a clamp. Here we show that free 5′-termini are selected using a disorder-thread-order mechanism. Adding short duplexes to 5′-flaps or 3′-streptavidin does not markedly impair the FEN reaction. In contrast, reactions of 5′-streptavidin substrates are drastically slowed. However, when added to premixed FEN and 5′-biotinylated substrate, streptavidin is not inhibitory and complexes persist after challenge with unlabelled competitor substrate, regardless of flap length or the presence of a short duplex. Cross-linked flap duplexes that cannot thread through the structured arch react at modestly reduced rate, ruling out mechanisms involving resolution of secondary structure. Combined results explain how FEN avoids cutting template DNA between Okazaki fragments and link local FEN folding to catalysis and specificity: the arch is disordered when flaps are threaded to confer specificity for free 5′-ends, with subsequent ordering of the arch to catalyze hydrolysis.

INTRODUCTION

Structure sensing 5′-nucleases are vital for DNA replication, repair, and recombination. Operating without regard to sequence, 5′-nucleases recognize defined nucleic acid junctions and catalyze the hydrolysis of specific phosphate diester bonds (1–3). Exemplary junctions for 5′-nuclease cleavage are formed during lagging strand DNA replication and long-patch base excision repair (lpBER), where 5′-extensions (flaps) occur at adjacent duplexes (Okazaki fragments and lpBER intermediates) as a consequence of polymerase strand displacement synthesis. Divalent metal ion-dependent flap endonucleases (FENs), the prototypical 5′-nuclease family members, are the enzymes that catalyze removal of 5′-flaps. This hydrolytic processing yields 5′-phosphorylated-nicked DNAs for subsequent ligation and during human replication must take place at least 50 million times per cell cycle. The importance of 5′-flap elimination is demonstrated by the lethality of *fen1*$^{(-/-)}$ knockouts in mammals (4). FENs endonucleolytically remove 5′-flaps, thereby avoiding repetitive exonucleolytic processing. Even before structures of FEN proteins became available, it was suggested that FEN specificity for junctions with free 5′-termini, and discrimination against other junctions lacking this feature that occur at replication forks, could be achieved by threading the 5′-flap DNA through a hole in the protein (5). Yet this proposal has remained controversial, and the basis for end specificity has remained enigmatic.

Subsequent structural studies did indeed reveal a hole in FEN proteins formed by helices linking the main DNA-binding domains straddling the active site (Figure 1A) (6,7). Known as the helical arch, this subdomain is partially disordered in some X-ray

*To whom correspondence should be addressed. Tel: +44 11 42 229478; Fax: +44 11 42 229346; Email: j.a.grasby@sheffield.ac.uk
Present address:
Peter Thompson, NIHR Trainees Coordinating Centre, Leeds Innovation Centre, 103 Clarendon Road, Leeds LS2 9DF, UK.

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

structures (Figure 1B and Supplementary Figure S1A) (8–10). In structured form, the arch is only large enough to accommodate single- but not double-stranded (ds) DNA, appearing to account for FEN specificity. Support for a threading mechanism for end specificity came from biochemical experiments that suggested that forming a duplex within the 5′-single-stranded (ss) flap or binding of proteins to this region of substrates prevented the FEN reaction (11,12). Structural studies of bacteriophage T4FEN bound to a pseudo-Y (pY) DNA substrate did show a 5′-flap within the arch region (9). However, in this complex, the DNA did not occupy the divalent metal ion-free active site and one helix of the arch was disordered (Figure 1C and Supplementary Figure S1B). Nevertheless, models can be created using this structure by overlay with FENs crystallized with ordered arches showing the flap DNA passing through, although not yet positioned in the active site for reaction (Figure 1D), furthering controversy regarding a possible threading mechanism for specificity.

In contrast, several studies have challenged the hypothesis that the helical arch enforces FEN specificity. Apparently in conflict with earlier literature, human FEN1 (hFEN1) has been demonstrated to endonucleolytically process so-called gapped flap substrates (13). Gapped flaps contain a short 5′-region of duplex that cannot pass through a structured helical arch. Other biochemical studies on the question of FEN 5′-flap accommodation have also produced results that are apparently at odds with a threading model (14,15). Thus, as an alternative to passage of substrate through the arch, this subdomain has been suggested instead to act as a clamp (3,7,14,16). One possible explanation of FEN specificity known as tracking in which FENs were proposed to initially interact with ss flaps either by threading or clamping and slide along these until junctions were encountered has been discredited (2,3,11–13,16–18).

Deciphering the origins of FEN1 specificity for 5′-flaps is made more complex by other 5′-nucleases that are sequence related to FENs but have differing specificities (1–3,19). In humans, EXO1, the mismatch and resection 5′-nuclease is most closely related to FEN1. EXO1 catalyses the processive hydrolyses of the 5′-termini of gapped, nicked and blunt duplex DNAs. Like FEN1, EXO1 can also endonucleolytically remove 5′-flaps (20). Another superfamily member XPG, the 5′-nuclease of nucleotide excision repair (NER), acts upon bubble substrates (21). The major human Holliday junction resolvase is suggested to be GEN1, another superfamily member (22). However, neither NER bubbles nor four-way junctions possess free 5′-termini *in vivo*. The 5′-portion of these substrates could therefore not be passed through an arch.

Recent structures of hFEN1 and hEXO1 bound to substrates and products in the presence of active site metal

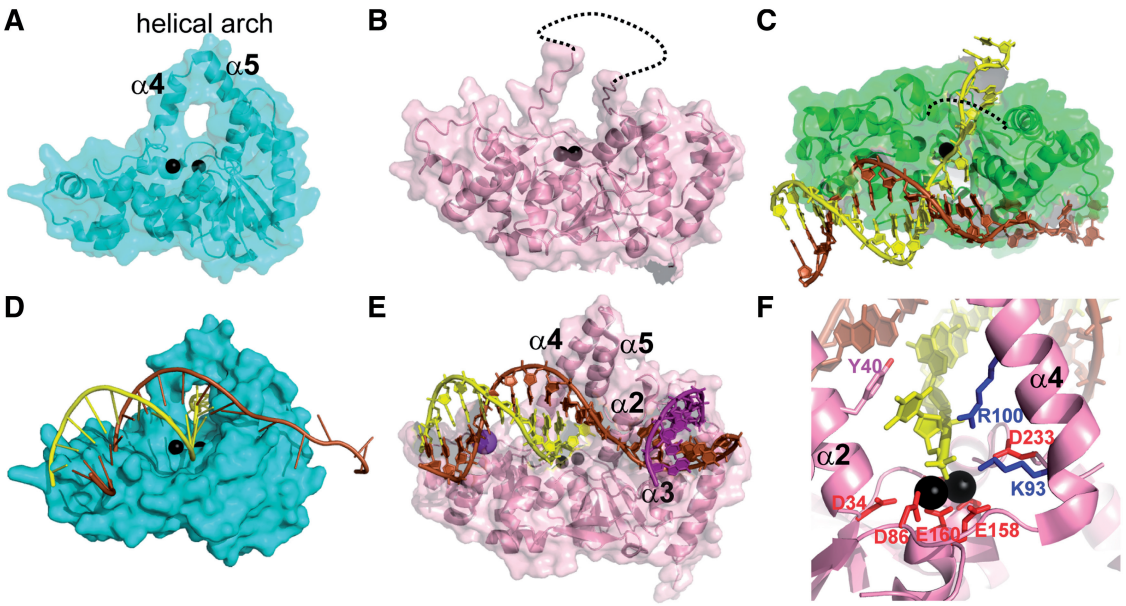


Figure 1. Structures of T5, T4 and human FENs with and without DNA. (**A**) Structure of T5FEN (1UT5) with transparent surface to highlight the helical arch and resulting hole above the active site bound divalent metals (black spheres). (**B**) Structure of hFEN1 (1UL1, X chain; pink) with transparent surface representations showing the disordered arch (missing arch residues, dotted lines; active site metal ions, black). (**C**) T4FEN structure in complex with a pseudo-Y (pY) substrate (2IHN) without metal ions. Based on alignment with a substrate-free T4FEN structure (1TFR), the location of active site divalent metal ions (black spheres) is shown along with template (brown) and 5′-flap strands (yellow) of the pY and disordered residues (dotted lines). (**D**) Model of T5FEN (1UT5) in complex with a pY substrate with active site divalent metals, protein and DNA colored as in (A) and (C) based on alignment with the T4FEN-DNA structure (2IHN) shows that the 5′-flap could go through the helical arch. Some steric clashes are observed suggesting conformational changes. (**E**) Structure of hFEN1 in complex with the product of reaction of a double-flap substrate (3Q8K). Template DNA (brown), the cleaved 5′-flap DNA strand (yellow), and 3′-flap DNA (purple) are shown with active site metal ions (gray) and a K^+ ion (purple). (**F**) Active site of the hFEN1-product DNA complex (3Q8K) showing the 5′-phosphate monoester product interacting with active site divalent metal ions (black spheres). Note, this nucleobase is not paired with the template. 5′-Nuclease superfamily conserved active site carboxylates (red) and helical arch α4 Lys93 and Arg100 (blue) are shown. A tyrosine residue (Tyr40) from α2 stacks upon the unpaired nucleobase.

ions highlight the similarities between 5′-nuclease superfamily members (Figure 1E) (2,3,23). Despite analogous structures for hFEN1 and hEXO1 complexes with product DNAs, which include conserved contacts between the cleaved 5′-phosphate and helical arch residues (Figure 1F), differing interpretations for the requirement for threading versus clamping indicate that key questions regarding the basis for substrate specificities within the FEN-like 5′-nucleases remain. Here, using functional studies with modified DNAs, we resolve how FENs accommodate the 5′-region of substrates, demonstrate that processing of 5′-gapped flaps is an hFEN1 activity that proceeds by the same mechanism and propose a universal model for departure of DNA from the active sites of 5′-nuclease superfamily members. Moreover, our results explain how FENs can function to rapidly remove flaps during replication without risk of destroying template DNA between Okazaki fragments.

MATERIALS AND METHODS

Over-expression and purification of T5FEN and hFEN1

T5FEN and hFEN1 (wild-type and K93A) were over-expressed and purified as described (2,24).

Synthesis and purification of oligonucleotide substrates

Oligonucleotides (ODNs) were synthesized using an ABI model 394 DNA/RNA synthesizer or by DNA Technology A/S (Risskov, Denmark) using 5′-fluorescein-CE- phosphoramidite (6-FAM) or 3′-(6-FAM)-CPG to incorporate 5′-FAM or 3′-FAM, respectively, and biotin TEG phosphoramidite to add biotin (Link Technologies, Lanarkshire, UK). The long tether 3′-biotin substrate [21 nt pY-3′B] was constructed using 3′-biotin TEG followed by three additions of spacer-CE-phosphoramidite-18. ODNs were purified by reverse-phase (RP) HPLC (Waters × bridge 10 × 250 mm C-18 column) using triethylammonium acetate buffers pH 6.5 with a gradient of acetonitrile. Purified ODNs were desalted using NAP-10 columns and subjected to MALDI-TOF mass spectrometry. Residual divalent metal ion contaminants were removed by treatment with Chelex resin. Experimental MWs were all within 3 Da of calculated. A complete list of ODNs is contained in Supplementary Figure S2.

Determination of the rate of decay of enzyme substrate complexes

Substrates were annealed as described (13,17). Enzyme and substrate were pre-incubated at 20°C (hFEN1) or on ice (T5FEN) for 2 min in 25 mM HEPES, pH 7.5, 50 mM KCl, 2 mM $CaCl_2$, 1 mM DTT and 0.1 mg/ml BSA (hFEN; calcium buffer) or 25 mM HEPES, pH 7.5, 50 mM KCl, 1 mM EDTA, 1 mM DTT and 0.1 mg/ml BSA (T5FEN; EDTA buffer) to form 'premixed' complexes. If required, five equivalents (with respect to [S]) of streptavidin (SA) were added before ('blocked' reactions) or after ('trapped' reactions) addition of enzyme and incubated accordingly for 5 min. Increasing the concentration of streptavidin did not alter the outcome. Samples were warmed at 37°C, and the reaction was initiated by mixing with an equal amount of magnesium buffer as above but containing 16 mM $MgCl_2$ instead of $CaCl_2$, (hFEN1) or 20 mM $MgCl_2$ instead of EDTA (T5FEN). The final concentrations of enzyme and substrate were 500 nM and 5 nM, respectively. For 'trapped' and 'premixed' reactions, sampling was carried out using quench flow apparatus (RQF-63 quench flow device, Hi-Tech Sci Ltd., Salisbury, UK). After time delays of 6.4 ms to 51.2 s, quench (8 M Urea containing 80 mM EDTA) was added. 'Blocked' reactions were sampled manually. Reactions were analyzed by dHPLC equipped with a fluorescence detector (Wave® system, Transgenomic, UK) as described (13,17,25,26). After quenching the presence of SA did not alter the dHPLC retention time with tetrabutyl ammonium bromide as the ion-pairing reagent (Supplementary Figure S3). The formation of product formed over time (P_t) was fitted to Equation (1), to determine the first-order rate constant (k) where P_∞ is the amount of product at end point:

$$P_t = P_\infty\left(1 - e^{-kt}\right) \quad (1)$$

Competition experiments

Competitor ODNs were pY or double-flap substrates without FAM label or biotin (Supplementary Figure S2). Enzyme and FAM–biotin–substrate were incubated at 20°C for 2 min (hFEN1) or on ice for 2 min (T5FEN) in either calcium buffer (hFEN1) or EDTA buffer (T5FEN) as above. For 'trapped' reactions, five equivalents of SA were added followed by incubation for a further 1 min. Competitor substrate was then added, and the mixtures were incubated for 10 min at 37°C. Increasing this time to 20 min had no impact on the outcome. An equal volume of magnesium buffer (as above) was added to initiate reaction producing final concentrations of enzyme (500 nM), FAM-labeled substrate (5 nM) and competitor (2.5 μM, T5FEN; 5 μM, hFEN1). Reactions were sampled, quenched and the amount of product determined as above. In experiments where streptavidin and/or competitor were not added an equal volume of appropriate buffer was, and all samples underwent identical incubations.

Preparation of azide-alkyne ODN

9 -(5′-*O*-Dimethoxytrityl-2′-deoxyribofuranosyl)-*N*2-[(dimethylamino) methylidene]-2-amino-6-methylsulfonyl purine-3′-(2-cyanoethyl-*N,N*-diisopropyl)-phospho ramidite (27) was used to construct an ODN with 3′-FAM and a 5′-alkyne (6-Hexyn-1-yl-(2-cyanoethyl)-(*N,N*-diisopropyl)-phosphoramidite (Glen Research) using mild/fast deprotection phopshoramidites for dC, dA and dG. Following 1 μmol scale synthesis, the CPG-bound ODN was treated with 200 μl of 11-Azido-3,6,9-trioxaundecan-1-amine:acetonitrile:DBU at a ratio of 9:9:2 at 37°C for 48 h with gentle mixing. Concentrated NH_3(aq) (1 ml) was then added and the mixture left for

a further 72 h at room temperature. After evaporation to dryness, the residue was suspended in water (150 µl) and extracted with diethyl ether (3 × 500 µl). The aqueous layer was removed and then purified by HPLC as described above. MW (AA-HP) 11 173.56 calculated; 11 176 found.

Preparation of triazole ODN

Reactions contained AA-HP (10 nmol) mixed with $CuSO_4.5H_2O$ (750 nmol) sodium ascorbate (30 µmol) and tris-3-hydroxypropyltriazolylamine (28) (21 µmol) in a total volume of 1 ml with NaCl (final concentration 0.2 M) and were incubated at room temperature overnight with gentle mixing. The reaction mixture was desalted (NAP-10) and purified by RPHPLC under denaturing conditions (as for other ODNs but at 55°C). The triazole ODN (Z-HP) eluted 3.3 min earlier than AA-HP (Figure 5C).

Determination of the rates of the reaction of triazole cross-linked gapped substrates

Kinetic analysis was carried out using GAP DF-AA and GAP DF-Z substrates at a concentration of 50 nM, with 5 pM WT hFEN1 or 2.5 nM K93A at 37°C in 50 mM HEPES, pH 7.5, 100 mM KCl, 8 mM $MgCl_2$, 1 mM DTT and 0.1 mg/ml BSA. Samples of reaction mixture were quenched in an equal volume of 250 mM EDTA, pH 8.0. Reactions were analyzed as above. Initial rates of reaction were obtained from plots of amount of product versus time.

RESULTS

A 5′-block inhibits FEN catalyzed reactions

To test whether FENs use a threading or clamping mechanism, we used biotinylated substrates to which streptavidin (SA) could be bound (Figures 2A, 3A and 4A; Supplementary Figure S2). The 53 kDa SA tetramer forms stable complexes with biotin [$t_{1/2}$ (our reaction conditions) >> 2 h (29)] and is too large to pass through the helical arch even when it is disordered. However, 5′-SA should not prevent clamping at the base of a 5′-flap. The addition of 3′- or 5′-biotin label to T5FEN substrates with 21 nt 5′-flaps did not significantly alter the rate of enzyme catalyzed reactions (Figure 2A and Supplementary Table S1). Neither did adding SA to reactions of non-biotinylated substrates. However, the rates of Mg^{2+} initiated reaction of T5FEN-substrate 'blocked' complexes formed by adding SA-conjugated substrates to enzyme differed by three orders of magnitude depending on the biotinylation site (Figure 2B and C). 5′-Blocked complexes reacted very slowly (Supplementary Figure S3), whereas 3′-SA complexes had a similar rate to those premixed with no block. Similar drastic decreases in the rate of reaction are observed under multiple turnover conditions (Supplementary Table S2).

5′-Substrate trapping does not inhibit reaction

To test whether the 5′-block prevented substrates threading through the helical arch and to rule out that 5′-streptavidin itself was inhibitory, we formed 5′-trapped complexes (Figure 2B). A trapped complex contained the same ingredients as a blocked complex, but the order of the addition of reagents differed. Without viable cofactor present, T5FEN and biotinylated substrate were premixed to form a complex, potentially allowing threading to take place and then excess SA was added. Upon initiation of reaction with Mg^{2+}, the rate of decay of 5′-trapped complexes was only 2-fold slower than premixed complexes where no SA was added and was 2000 times faster than 5′-blocked complexes.

5′-Streptavidin addition creates a non-exchangeable-trapped complex

To further test that 5′-SA-trapped complexes were the result of threading, we challenged FAM-labeled substrate complexes with excess unlabeled substrate prior to initiation of reaction. Threaded 5′-trapped complexes would be predicted to have a dissociation rate commensurate with the biotin–streptavidin interaction and should therefore be essentially irreversible and resistant to competition on shorter time scales. Labeled substrate was competed from non-SA premixed complexes; the amount of FAM-product observed upon initiation of reaction with Mg^{2+} was consistent with ratio of labeled to unlabeled material and the concentration of enzyme (Figure 2C and D; Supplementary Figure S4). However, 5′-SA-trapped complex could not be competed away; the amount of product observed in this case was similar to that produced by the non-challenged trapped complex. In contrast, with the 3′-biotinylated substrate, the SA 'trapping' procedure resulted in a readily exchanged complex (Figure 2C and D). Combined blocking, trapping and competition experiments imply the 5′-portion of T5FEN substrates needs to be threaded through the helical arch for optimal enzyme activity.

hFEN1 uses the same mechanism for short and long 5′-flaps

hFEN1 catalyzed reactions have been suggested to proceed by two different mechanisms, dependent on 5′-flap length (16). Longer flaps are posited to thread through the helical arch, whereas substrates with shorter 5′-flaps (<6 nt) are supposed to not thread. Thus, to determine whether hFEN1 does indeed support two different mechanisms, we conducted experiments analogous to those described for T5FEN using 5′-biotinylated double flap (DF) substrates having 3 and 21 nt 5′-flaps (Figure 3A). As with T5FEN, biotinylation of substrates had negligible impact on hFEN1 catalytic parameters (Supplementary Table S1).

Premixed, blocked and trapped hFEN1 substrate complexes were created in the presence of catalytically inert Ca^{2+} (Figure 2B) (30,31). Results are analogous to those obtained with T5FEN. Rates of reactions of 5′-blocked complexes were decreased by three or four

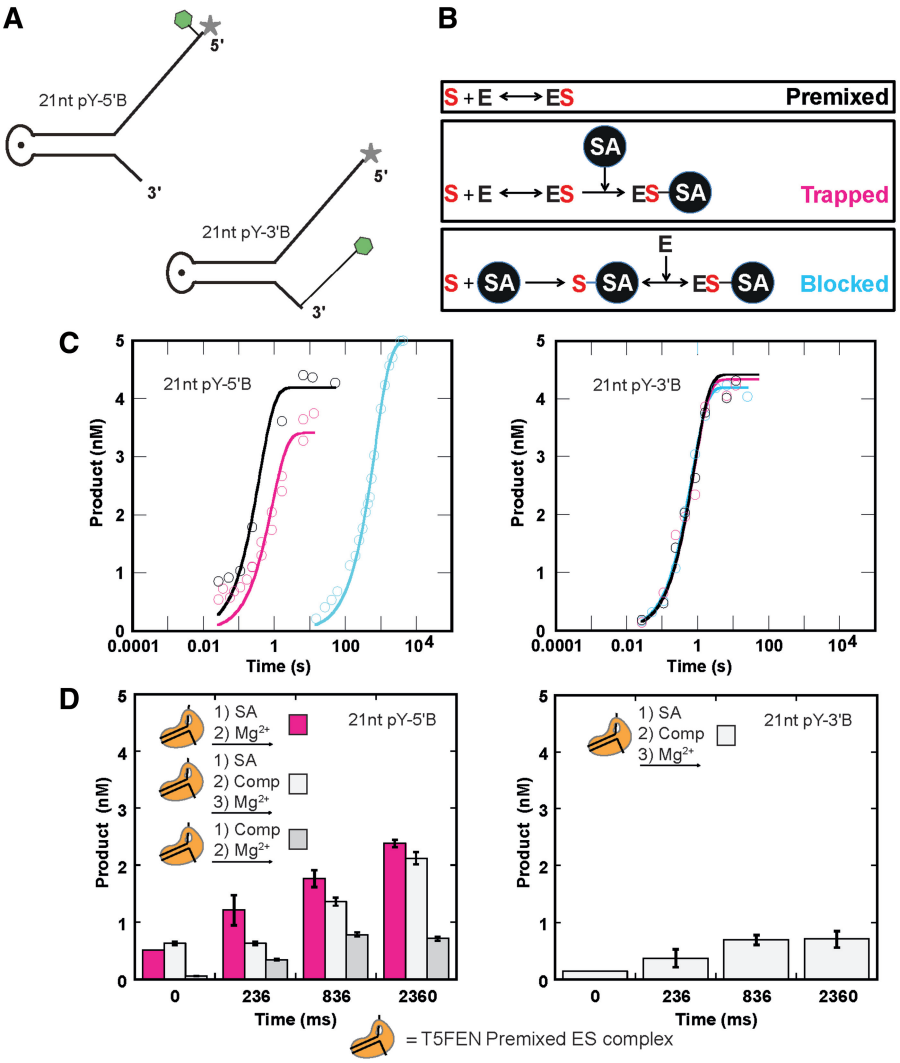


Figure 2. The rate of reaction of complexes of T5FEN and streptavidin substrates is dependent on the order of the addition of reagents and the site of substrate biotinylation. (**A**) Biotinylated substrates for T5FEN with 21 nt 5′-flaps. Green hexagon and gray star indicate biotin and fluorescein (FAM), respectively. (**B**) Order of the addition of reagents schematic illustrating how premixed, blocked and trapped complexes of FEN and substrates were prepared as detailed in 'Materials and Methods' section. (**C**) The effect of the SA addition on rates of T5FEN reactions containing 5′- (left) and 3′- (right) biotinylated substrates shown as plots of product versus time. Note: the *X*-axis (time) is in log format. Premixed (black), blocked (cyan, see Supplementary Figure S3) and trapped (pink) reactions were initiated at 37°C at final concentrations 500 nM T5FEN, 5 nM substrate, 10 mM Mg^{2+}, 50 mM KCl, 1 mM EDTA, 1 mM DTT and 0.1 mg/ml BSA at pH 7.5. All data points are the result of three independent experiments (error bars omitted for clarity). Rate constants are summarized in Supplementary Table S3A. (**D**) The concentrations of product formed at various time intervals after initiation of reactions of trapped and premixed complexes that were challenged with excess competitor (Comp; non-biotinylated and unlabeled 21 nt pY substrate, final concentration 2.5 μM). Trapped and premixed complexes were formed and incubated at 37°C for 10 min with 5-fold excess Comp before initiation of reaction with magnesium buffer. Left, the amounts of product formed from 5′-SA trapped 21 nt pY-5′B complexes with and without the addition of Comp are shown in white and pink, respectively, whereas that formed from premixed 21 nt pY-5′B is shown in gray. Right, the amount of product formed from a 3′-SA trapped 21 nt pY-3′B to which Comp was added is shown in white. All data points are the result of three independent experiments with standard errors shown. See also Supplementary Figure S3.

orders of magnitude dependent on flap length (Figure 3B). Similar reductions in reaction rates were also observed under multiple turnover conditions (Supplementary Table S2). In contrast, 5′-SA-trapped hFEN1 complexes behaved like premixed complexes containing non-SA substrates (Figure 3B). When challenged with excess unlabeled substrate prior to initiation of reaction, premixed enzyme–substrate complexes were readily competed (Figure 3C and Supplementary Figure S4). In contrast, 5′-SA-trapped complexes persisted and underwent hydrolysis. Importantly, similar outcomes are observed regardless of the length of the 5′-flap, and all data are in accord with a threading but against a clamping mechanism for all lengths of 5′-flap.

Unlike T5FEN and despite saturating conditions for the hFEN1–substrate interaction, the proportion of complex that was trapped (i.e. cleaved at 60 s, approximately 500 half-lives for trapped substrates) was altered by the conditions of pre-incubation (Figure 3D). Although some hFEN1–substrate complex was trapped in the

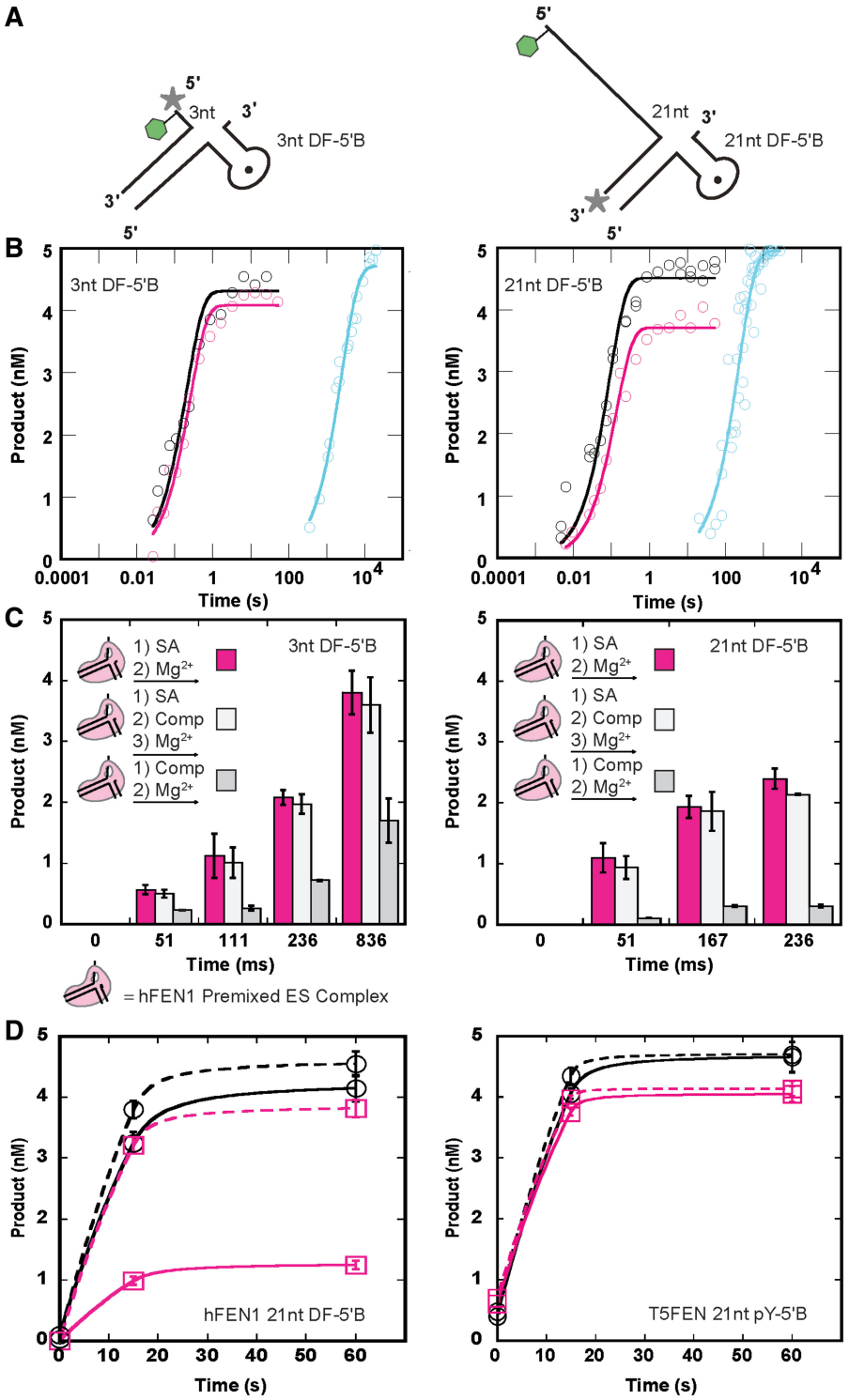


Figure 3. The rates of reactions of hFEN1 catalyzed hydrolysis of 5′-biotinylated substrates in the presence of SA are affected by the order of the addition of reagents. (**A**) Biotinylated DF substrates for hFEN1. A 3 nt 5′-flap DF (left) and 21 nt 5′-flap DF (right) with 5′-biotin and fluorescent label indicated with a green hexagon and gray star, respectively. (**B**) Product versus time plots of hFEN1 reactions of 3 nt DF-5′B (left) and 21 nt DF-5′B (right) illustrating the effects of order of the SA addition on the rates of reaction. Note: the *X*-axis (time) is in log format. Premixed (black), blocked (cyan) and trapped (pink) were assembled as in Figure 2B and then reaction initiated at 37°C by addition of Mg^{2+} at final concentrations 500 nM hFEN1, 5 nM substrate, 8 mM Mg^{2+}, 100 mM KCl, 1 mM EDTA, 1 mM DTT and 0.1 mg/ml BSA at pH 7.5. All data points are the result of three independent experiments (error bars omitted for clarity). Rate constants are summarized in Supplementary Table S3B. (**C**) The concentrations of product formed at various time intervals after initiation of reactions of trapped and premixed complexes that were challenged with excess competitor (Comp; non-biotinylated and unlabeled 5 nt DF substrate, final concentration 5 μM). (**D**) Trapped and premixed complexes of hFEN1 and 3 nt DF-5′B (left) and 21 nt DF-5′B (right) were formed and incubated at 37°C for 10 min with 10-fold excess competitor before the addition of magnesium buffer. Concentrations of product formed from trapped complexes with and without the addition of competitor are shown in white and pink, respectively. Concentrations of product formed from premixed complexes challenged with competitor are shown in gray. All data points are the result of three independent experiments with standard errors shown. See also Supplementary Figure S3.

presence of EDTA (15–20%), a greater proportion of substrate was trapped in the presence of Ca^{2+} ions (75–80%). Altering the temperature at which SA was added did not significantly alter the outcome. The portion of complex that did not react on a subsecond time scale decayed with a similar rate to 5′-SA blocked species. Thus, in EDTA the hFEN1–substrate complex is in equilibrium between threaded and non-threaded complexes both of which can be captured by 5′-SA. The amount of fast-reacting product reflects the equilibrium state between these two forms. When the trapping procedure is carried out with non-catalytic Ca^{2+} ions present, this equilibrium was altered in favor of the more catalytically proficient species (Supplementary Figure S5). These differences between higher and lower evolutionary FENs can be accounted for by their respective structures. As revealed by the T4FEN–PsY complex charged and aromatic residues of the partially structured arch interact with the flap in a state where it is not active site positioned due to lack of divalent ions (9). These residues are largely functionally conserved in T5FEN, but not in hFEN1. In hFEN1, the considerable repulsion afforded by the metal-free seven carboxylate active site, together with the possible influence of active site metals on arch conformation, presumably results in a requirement for divalent ions for flap accommodation.

Gap and flap endonuclease activities occur via the same mechanism

Structured helical arches are not large enough for duplex DNAs to pass through. Despite this and somewhat controversially, hFEN1 supports endonucleolytic reactions of gapped flaps containing short duplexes (11,13). To determine whether reactions of gapped substrates are a bona fide activity of hFEN1 instead of a co-purifying activity from *Escherichia coli*, we compared rates of hydrolysis of gapped and flap substrates with wild-type and K93A mutant hFEN1. Both proteins were rigorously purified using the same procedure. In hFEN1 and hEXO1 structures, this helical arch lysine, known to be important for catalysis *in vitro* and *in vivo*, contacts the 5′-phosphate monoester of product DNA positioned on active site metal ions and is suggested to act as an electrostatic catalyst (Figure 1F) (2,3,32,33). The gapped DF substrate contained a short stable hairpin within the 5′-flap (Figure 4A). The rate of reactions with the 21 nt DF and the gapped DF were both severely decreased with the K93A mutant, emphasizing that hFEN1 is responsible for the cleavage and that this reaction is mechanistically similar to the endonucleolytic incision of flaps lacking secondary structure (Figure 4B).

To ascertain whether gapped substrates are passed through the hFEN1 helical arch, a 5′-hairpin DF with a biotin label added to the hairpin turn was used (Figure 4A). Duplex stability was unaffected by biotin addition (Supplementary Figure S6A and S6B). Results were identical to those of 5′-ss-flap substrates. 5′-Blocking the gapped DF with SA severely decreases the rate of reaction (Figure 4C). Similarly, trapped complexes can be formed when SA is added after pre-incubation of biotinylated substrate with hFEN1; the rate of decay of these complexes is analogous to those premixed in the absence of SA. Additionally, although premixed gapped biotin DF–hFEN1 complex was readily competed by unlabeled substrate, it was not competed when 5′-trapped with SA (Figure 4D and Supplementary Figure S4). Thus even a short duplex contained within a 5′-flap is able to pass through the helical arch.

hFEN1 processes gapped flaps without resolving secondary structure

A possible mechanism for accommodation of 5′-gapped-flap substrates within the helical arch involves passage of a transient ss species (13,16). To test if 5′-duplexes must become ss for FEN processing, we formed a hairpin where the secondary structure within the 5′-flap could not be resolved due to a covalent cross-link. The triazole cross-link was formed by Cu(I) catalyzed Huisgen [3 + 2] cycloaddition reaction ('Click chemistry') of an alkyne and azide (Figure 5). Introduction of an azide was achieved using the convertible nucleoside sulfone (27) (Figure 5A), which after ODN assembly was reacted with an amino-PEG-azide before removal from the CPG support. The resultant azido functionalized 2,6-diaminopurine derivative base paired with the 5′-terminal dT of the hairpin (Figure 5B). An alkyne was introduced to the 5′-terminus using commercially available 5′-hexynyl phosphoramidite. Formation of the cross-linked hairpin was monitored by reversed-phase HPLC (Figure 5B and C). The resultant cross-linked oligomer, which no longer exhibited a temperature-dependent melting transition (Supplementary Figure S6C and D), was hybridized to a template strand to yield a locked-hairpin gapped DF substrate. hFEN1-catalyzed cleavage of this non-resolvable substrate proceeded with only a modest 2-fold reduction in rate (Figure 5D and E). Thus, accommodation of substrates with short duplex regions within the 5′-flap does not require the resolution of secondary structure.

Implications of threading or clamping models

To understand the implications of threading and clamping models for 5′-nucleases, we examined hFEN1 and hEXO1 product DNA structures. Both complexes show the respective product 5′-phosphate monoesters bound by two active site divalent metal ions and superfamily conserved Lys93 and Arg100 residues (hFEN1 numbering) from the first helix (α4) of the helical arch (Figure 1F) (2,3). Three routes are possible for ssDNA extending from the 5′-phosphate, depicted schematically in Figure 6A. Following a simple linear path threads the DNA through the helical arch so that it encloses the first added nucleotide joined to the scissile bond (Figure 6A, route 1). For flapless DNAs that are substrates for both FENs and EXO1, this corresponds to the terminal 5′-nucleotide of the substrate. Introducing a turn at this nucleotide was necessary to create clamping models that allowed substrate to depart either side of the arch. Passing in front of α4 is unhindered but requires the phosphate

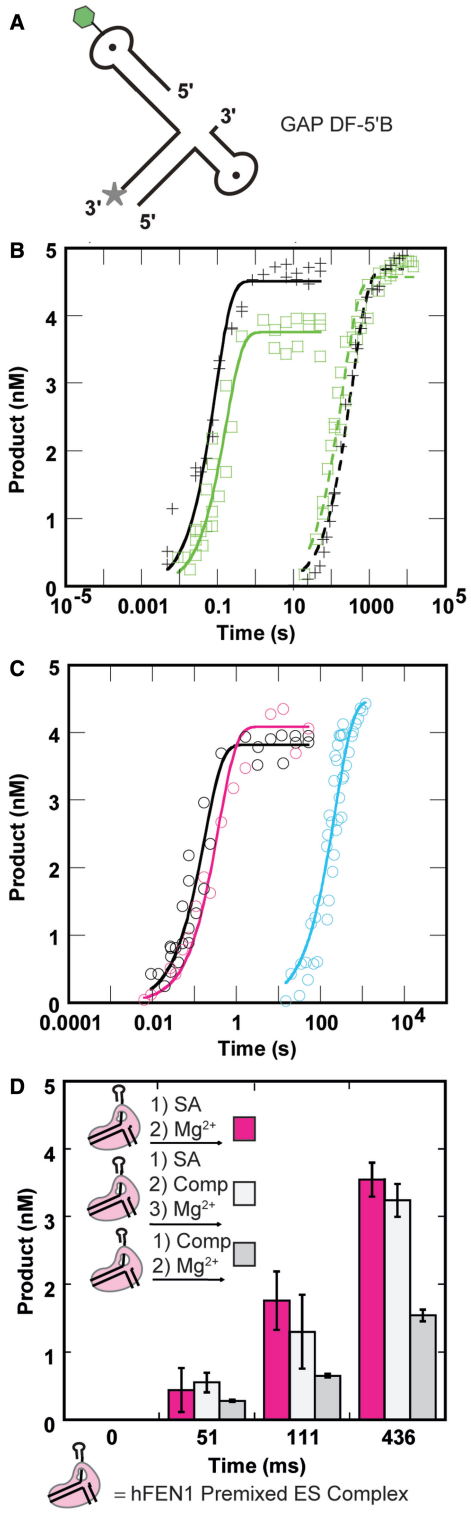


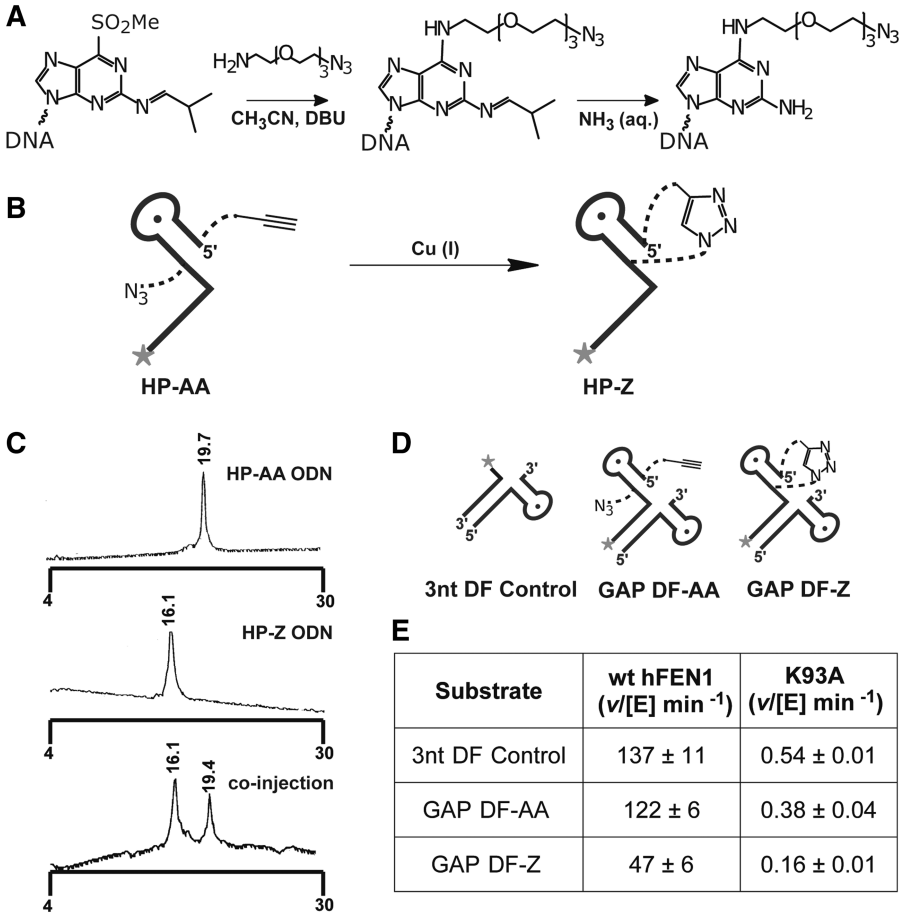
Figure 4. The effect of short stable hairpins in the 5′-flap on hFEN1 catalyzed reactions. (**A**) Schematic of the biotinylated gapped DF substrate (GAP DF-5′B) with biotin and a 3′-fluorescent label indicated by the green hexagon and gray star, respectively. (**B**) Plots of product versus time for wild-type (straight line) and K93A (dashed line) hFEN1 premixed GAP DF-5′B (green) and 21 nt DF-5′B (black) reactions showing the effects of the presence of dsDNA in the 5′-flap and K93A mutation on the rates of reaction. Note: the *X*-axis (time) is in log format. Curve fits using equation 1 yield first order rate constants 683 ± 56/min (WT hFEN1: 21 nt DF-5′B) 357 ± 34/min, (WT hFEN1:GAP DF-5′B) 0.3 ± 0.02/min (K93A hFEN1: 21 nt DF-5′B), 0.17 ± 0.01/min (K93A hFEN1: GAP DF-5′B). (**C**) Product versus time plots of hFEN1 reactions containing GAP DF-5′B illustrating the effects of order of the SA addition on the rates of reaction. Note: the *X*-axis (time) is in log format. Premixed (black), blocked (cyan) and trapped (pink) were assembled, and then the reaction was initiated as in Figure 3B. All data points are the result of three independent experiments (error bars omitted for clarity). Rate constants are summarized in Supplementary Table S3B. (**D**) The concentrations of product formed at various time intervals after initiation of reactions of trapped and premixed gapped flap complexes that were challenged with excess competitor as in Figure 3C, before addition of magnesium buffer to initiate reaction. Concentrations of product formed from trapped complexes with and without the addition of competitor are shown in white and pink, respectively. Concentration of product formed from premixed complexes challenged with competitor is shown in gray. All data points are the result of three independent experiments with standard errors shown. See also Supplementary Figure S3.

backbone of the flap to approach the backbone of the substrate duplex (Figure 6A, route 2). Departure of substrate past the other side of the arch involves inserting the ssDNA between α5 and α2 (Figure 6A, route 3) (3). Route 3 is not feasible in hFEN1 or hEXO1 since this path is blocked by interactions between the α5–α6 linker and α2, but has recently been suggested to be the way XPG would accommodate NER bubbles (3). Similar models can also be created with bacteriophage FENs, where arch architecture differs slightly (Supplementary Figure S7). Notably, all clamping models only predict close proximity of the 5′-nuclease and the first 2–4 nt joining the scissile phosphate diester at the base of 5′-flaps.

DISCUSSION

The way FENs accommodate 5′-flaps and whether this in turn confers substrate selectivity have been longstanding questions. This puzzle is further complicated by the 5′-nuclease superfamily, where structurally related proteins with conserved active sites catalyze the same reaction, but preferentially act on different nucleic acid structures (1). Furthermore, despite long-held notions that FEN specificity is mediated by initial recognition of a free 5′-ss flap, analyses show that FENs initially bind the two-way junctions of substrates largely by complementary strand interactions (2,3,11–13,16,17,34). Recent structural analyses imply that double nucleotide unpairing of the reacting duplex is required to form a catalytically competent complex with contact between active site ions and the scissile bond (1–3,30,31). Thus, determining how FENs accommodate the 5′-portion of substrate is essential to understand how the reaction competent complex is formed. Furthermore, this information together with known differences in specific regions of various FEN superfamily members may suggest how this mechanism is adapted by other 5′-nucleases to substrates that lack free 5′-termini.

Our experiments, tested with two FEN family members, were designed to differentiate between threading and clamping models, and functionally determine how FENs accommodate 5′-flaps. When the 53 kDa streptavidin tetramer is bound to the 5′-flap before the addition of enzyme, reactions catalyzed by T5FEN and hFEN1 are



E

Substrate	wt hFEN1 (v/[E] min^{-1})	K93A (v/[E] min^{-1})
3nt DF Control	137 ± 11	0.54 ± 0.01
GAP DF-AA	122 ± 6	0.38 ± 0.04
GAP DF-Z	47 ± 6	0.16 ± 0.01

Figure 5. Processing of gapped flaps by hFEN1 does not require resolution of secondary structure. (**A**) Schematic illustration of the procedure used to introduce an azide functionality after ODN synthesis but before removal from CPG. Cleavage from the support and deprotection produced a hairpin containing oligodeoxyribonucleotide (ODN) with a 5′-alkyne, internal azide and 3′-FAM (HP-AA). (**B**) 'Click' reaction using HP-AA ODN produced as in Figure 5A to yield the triazole cross-linked hairpin containing ODN HP-Z. (**C**) dHPLC traces of the reaction shown in Figure 5B of HP-AA ODN (top) to produce HP-Z ODN (middle). A co-injection of HP-AA and HP-Z ODNs is shown (bottom). (**D**) A 3 nt DF substrate (3 nt DF Control) used for comparison with gapped DF substrates made with HP-AA and HP-Z. (**E**) Normalized initial rates of reaction of WT hFEN1 and K93A hFEN1 with substrates as in Figure 5D. Reactions contained 50 nM substrate and either 5pM (WT) or 2.5 nM (K93A) in 50 mM HEPES pH 7.5, 100 mM KCl, 8 mM $MgCl_2$, 1 mM DTT and 0.1 mg/ml BSA.

inhibited consistent with streptavidin blocking the flap from threading though the arch (Figure 6A, route 1). Inhibition did not occur when streptavidin was added to the preformed enzyme–substrate complex, showing that streptavidin itself is not inhibitory and consistent with the hypothesis that the flap has already passed through the archway. Moreover, inhibition did not occur when streptavidin was added to the 3′-terminus of substrate before interaction with T5FEN, again showing that streptavidin itself is not inhibitory, but that its positioning on substrate is critical to the inhibition. The lack of displacement of labeled substrates by competitor from complexes where T5FEN or hFEN1 is added before 5′-streptavidin, contrasts with the ready competition of non-streptavidin conjugated complexes. This is consistent with a model where the 5′-flap is threaded through the arch and is trapped in this state by streptavidin.

In contrast, it is difficult to reconcile these results with clamping models where the substrate passes on either side of the arch (Figure 6A, routes 2 and 3). 5′-Streptavidin should not interfere with reactions proceeding by a clamping mechanism; indeed measured dissociation constants for hFEN1 5′-biotinylated substrates with long flaps ± streptavidin without cofactor ions are identical (16). The length of the 5′-flap in our substrates, particularly those with 21 nt flaps or gapped flaps, is much greater than the 2-4 nt that would be needed to pass the edge of the arch in either direction. An unlikely scenario where 5′-streptavidin produced non-exchangeable clamped (trapped) species would require an interaction between the tetramer and both T5FEN and hFEN1, regardless of the 5′-flap length. The possibility of a fortuitous interaction with streptavidin is ruled out by the lack of inhibition observed with the 3′-modified substrate where the length of the biotin linker could still allow interaction with the arch region of FENs. Together, our results strongly argue against a clamped structure and favor the threaded structure as the catalytically proficient form of T5FEN and hFEN1 that reacts on a biologically relevant timescale.

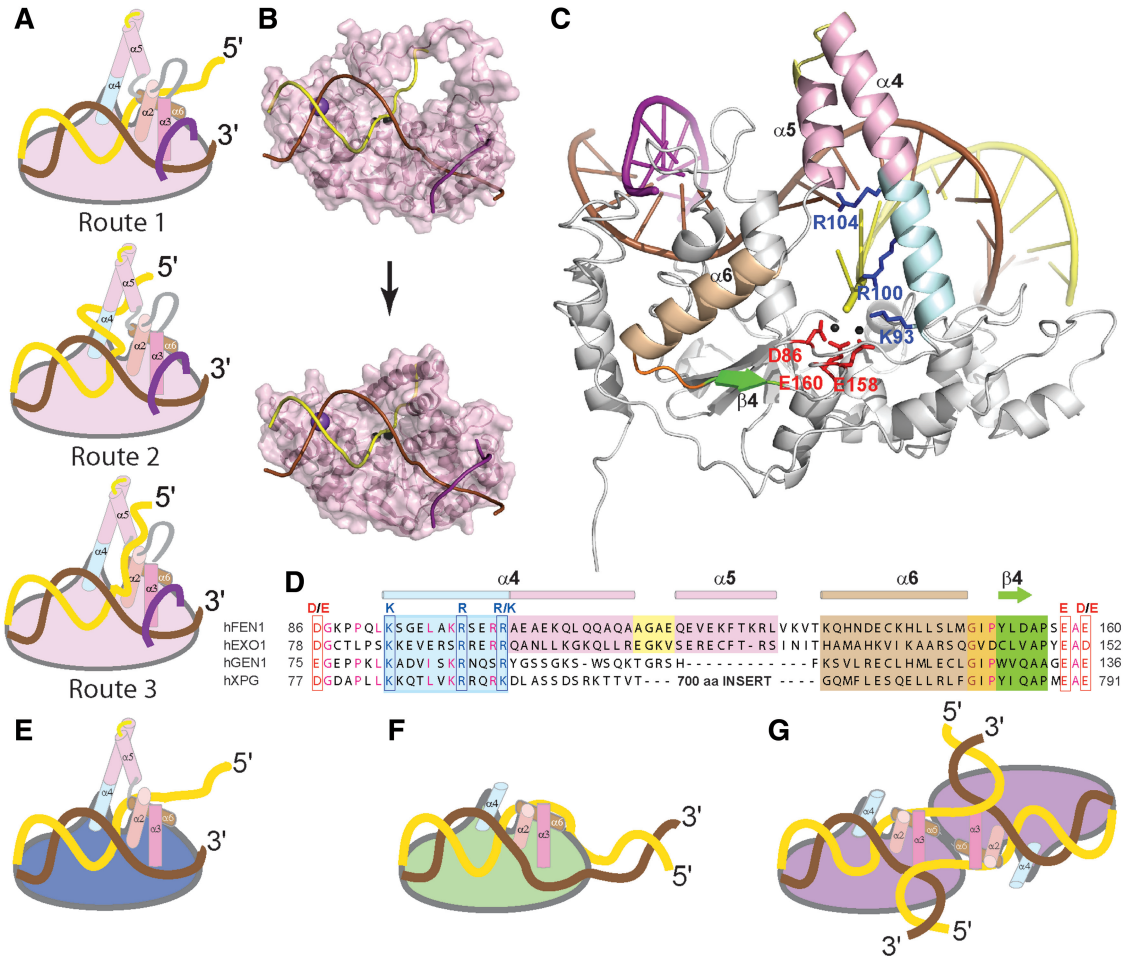
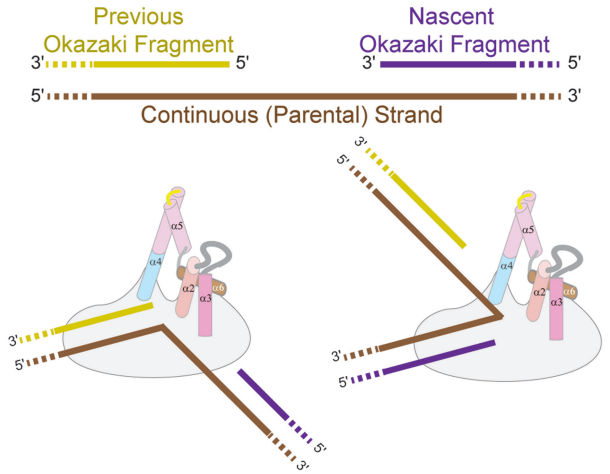


Figure 6. Disorder-thread-order model for hFEN1 and a unified model for the accommodation of substrates by the 5′-nuclease superfamily. (**A**) Schematic models for departure of the 5′-portion of hFEN1 substrates from the active site involving threading (route 1) or clamping (routes 2 and 3). Selected helical elements that define routes 1-3 are represented as cylinders. (**B**) Proposed disorder-thread-order model for coupling substrate selectivity to catalysis through the helical arch shown for hFEN1. The lower model is based on the structure of the enzyme-product (3Q8K). Bound DF substrate is modeled by extension of product DNA as in Figure 6A route 1. The upper panel is generated from the DNA-free hFEN1 (1UL1, Y chain) with missing regions modeled in with CHARMM and overlaying the substrate DNA from superimposition with the lower panel pdb model. (**C**) hFEN1-product structure (3Q8K) with secondary structure elements between D86 and D160 including the helical arch (amino acid residues 93–130), highlighted. The superfamily conserved lower part of α4, top of α4 and α5 (helical cap), α6, α6–β4 loop and β4 are colored pale cyan, pink, brown, orange and green, respectively. (**D**) Sequence alignment of hFEN1, hEXO1, hXPG and hGEN1 between D86 and D160 colored as in (C) showing conservation of catalytic residues and lack of conservation of the helical cap in hGEN1 and hXPG. (**E-G**) Schematic representations of the proposed unified substrate binding model for the 5′-nuclease superfamily where the 5′-portion of substrates departs from the active site following the same path. (E) EXO1 nuclease domain with pY DNA, (F) XPG nuclease domain with bubble and (G) GEN1 nuclease domain with a four-way junction. In (E) the junction portion of the substrate is positioned as in hEXO1 structures with the added 5′-flap passing through the arch. In (F) and (G), the FEN1/EXO1 helical cap has been removed allowing substrates without free 5′-ends to be accommodated. Reacting duplexes of substrates are positioned as in hFEN1 and hEXO1 structures and the path of 5′-portion of substrate past α4 (blue) is inferred from our data. The positioning of other substrate features is shown for illustrative purposes only.

Despite the dimensions of structured helical arches that will not permit ds nucleic acid to pass through, our data show that substrates with a short gapped 5′-duplex behave similarly to 5′-ss flaps and must also be passed through the arch. Furthermore, resolution of secondary structure is not required to achieve this. This observation rules out a mechanism where FENs thread DNA through a structured archway and implies a model where the 5′-portion of substrate is initially passed through the unstructured or partially structured arch seen in many FEN structures (Figure 6B) (8,9). Our comparisons of FEN crystal structures, which underscore the flexibility of the arch region even in the crystal environment, are consistent with the opening of the arch rather than the melting of the gapped DNA. Once 5′-flap or 5′-gapped duplex DNA is through the disordered loop, adopting the helical structure correctly positions the catalytic lysine and arginine residues for interaction with the unpaired scissile phosphate diester at the active site metals (Figures 1F and 6B and C), thereby linking reaction site selection to catalysis. Our disorder-thread-order model explains data that appeared at odds with the threading of substrates through a closed aperture, such as the ability of FENs to process 2′-branched flaps or those with moderately



Model	Could cleavage occur?	
Threading	Yes	No
Clamping	Yes	Yes

Figure 7. Models of FEN1 and EXO1 substrate interactions and their implications for genome integrity. The gapped junction between Okazaki fragments during lagging strand replication, or an equivalent structure produced as a response to damage, could be bound by FEN1 and EXO1 in two different ways. Both threading and clamping models could lead to reaction at the free 5′-termini after double nucleotide unpairing, with the 5′-nucleotide contained within or clamped by the arch (left). FENs readily process gapped flaps that contain short regions of duplex by threading them through the disordered arch. However, when the duplex is long or lacks 5′-termini, it could not pass through the arch and so reaction on the continuous strand is selected against, thus protecting genome integrity (right). In contrast if the 5′-portion of the substrate were clamped, the reaction of the continuous strand could occur.

bulky modifications such as (*cis*)-diamine dichloroplatinum (CPPD) adducts and thymine dimers (14,15). Threading though a disordered arch also eliminates concerns with an earlier bind and then thread hypothesis that proposed passing 5′-flaps through a structured archway (18). The act of binding the complementary DNA junction orients the 5′-portion of the substrate into the disordered/partially ordered arch region of the protein, and consequently there is no requirement for an energy source to push or pull a 5′-flap through a small aperture.

Considering the diverse range of substrates that the 5′-nuclease superfamily can process, devising a mechanism that is consistent with all family members and that can readily explain differences in overall specificity is a formidable challenge. Nevertheless, the junction binding abilities of 5′-nucleases moderated by hydrophobic wedges revealed by recent structures of hFEN1 and hEXO1 are consistent with adaptation to recognition of the bubble and four-way junction substrates of XPG and GEN1 (2,3). Similarly, as the site of reaction in all family members is predominantly 1 nt into a duplex, the double nucleotide unpairing mechanism implied by hFEN1 and hEXO1 structures seems likely to apply to all family members (1–3,30,31). As the hEXO1 helical arch so closely resembles that of hFEN1, it seems probable that this too uses a disorder-thread-order mechanism (Figure 6E). However, a considerable problem arises regarding how the 5′-portion of GEN1 and XPG substrates are accommodated. The data presented here do show that catalysis of reactions of blocked substrates can occur, albeit on a non-biologically significant time scale. Analogously, a chimera of the XPG junction binding domain with a FEN helical arch can cleave bubbles, albeit very slowly (2). Thus, if the features that encode specificity for free 5′-termini are removed, processing of bubbles and Holliday junctions could take place. Superfamily sequence comparisons of the arch subdomain and juxtaposed regions reveal that whilst all members conserve the base of α4 that contains conserved basic residues that interact with the scissile phosphate, the latter part of this helix and α5, called the 'helical cap', is not conserved by XPG and GEN1 (Figure 6C and D). If in XPG and GEN1 this region of the protein adopts a structure that does not close an aperture around substrates but instead creates a groove or cleft to accommodate a single non-base paired nucleotide, then the departure of all superfamily substrates from the active site can take a similar route past the catalytically relevant α4 residues (Figure 6A, route 1 and Figure 6E–G).

The disorder-thread-order mechanism revealed by the results presented here has implications for the roles of FEN1 and EXO1 during replication and repair and their ability to maintain genome fidelity. Gapped flaps containing short regions of duplex do not drastically alter FEN activity. When such gapped flaps form (e.g. from repeating sequences), they will not severely inhibit FEN1 activity protecting against genome expansions unless flap duplexes are bound stably by proteins (35). However, when duplexes in the 5′-portion of substrates are very long or lack a 5′-terminus, they will inhibit the activity of 5′-nucleases that thread substrates underneath a helical cap. This applies to gapped DNAs occurring between Okazaki fragments during lagging strand replication or as a response to damage. For these gapped DNAs, potentially two junction binding modes of FEN1 and EXO1 could result in reaction at a free 5′-terminus or of the continuous (in replication, parental) strand (Figure 7). The disorder-thread-order model predicts FEN1 or EXO1 catalyzed reaction of the continuous strand is strongly selected against because a long duplex could not pass through the aperture. Thus even when gaps are too small for ss binding proteins to associate with the continuous strand it remains protected from EXO1 and FEN1 action supporting genome fidelity. In contrast, a clamping mechanism would not protect the genome in this way.

The disorder-thread-order mechanism also has implications for targeting FENs for therapeutic intervention and for possible mechanisms for control of activity in other superfamily members including the more distantly related 5′-exoribonucleases. The importance of FEN activity in all organisms (4,33,36,37), differences between higher and lower evolutionary FENs and other superfamily members (1–3,6,9,17,19,30), and the high levels of FEN1 expression in cancer cells (38), all hold promise

for possible antibacterial, antiviral or anticancer therapies. Thus, the search for FEN-specific inhibitors has been initiated (39). A requirement for disorder-order transition provides new inhibition strategies directed at preventing structural changes to be explored. Extending the possibility of disorder-order transitions of the base of α4 to other family members, regardless of whether they thread substrates under the helical cap or place them in a groove, suggests the possibility for control of activity of larger family members (EXO1, XPG and GEN1) through intra- and/or inter-molecular protein–protein interactions (3). The 5′-nucleases of RNA degradation Xrn1 and Xrn2 (Rat1 in yeast) conserve the FEN active site including equivalent positioning of lysine and arginine residues on an arch-like domain, seen as partially disordered in current structures (40–42). However, unlike FENs, these enzymes produce only mononucleotide products not flaps from 5′-phosphorylated but not m^7GpppN-capped RNAs. It is interesting to note that exit from the back of the 'arch' is blocked in XRNs, and extending the threading model, this could explain both their specificity for 5′-phosphorylated species and the size of their products. Multiple rounds of disorder-thread-order transitions may be required for processive hydrolysis by these enzymes.

In contrast to other members of the superfamily, FENs are not activated by other proteins (3,43). Additionally, hFEN1 interaction partners have only modest effects on activity and do not downregulate it (44). Yet, *in vitro* hFEN1 is an extraordinarily efficient enzyme on DF substrates with second-order rate constants that approach those for the diffusional encounter of biomolecules ($10^7 - 10^9$/M/s) (13). Post-translational modifications may alter activity at certain cell cycle phases (45), but during replication FEN discrimination is paramount. FENs must only process substrates with free 5′-termini and not endanger the genome by cutting other junctions that are formed at replication forks. In particular, FENs must not destroy template DNA between Okazaki fragments. A single nucleotide 3′-flap binding site in hFEN1 affords some preference for the products of displacement synthesis (2,13,46). In concert, passing the 5′-portion of substrates through a disordered aperture, which then orders to position catalytic residues, couples substrate selection to catalysis and provides exquisite specificity for substrates that possess free non-protein bound 5′-termini.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–7, Supplementary Tables 1–3.

ACKNOWLEDGEMENTS

We thank Elaine Frary, Robert Hanson and Simon J. Thorpe for expert technical assistance.

FUNDING

Biotechnology and Biological Sciences Research Council (grant number BBF0147321 to J.A.G.); Marie Curie International Incoming Fellowship within the 7th European Community Framework Programme (project number PIIF-GA-2009-254386 to L.D.F.); National Cancer Institute (grant numbers RO1 CA081967, P01 CA092584 to J.A.T.). Funding for open access charge: BBSRC.

Conflict of interest statement. None declared.

REFERENCES

1. Tomlinson,C.G., Atack,J.M., Chapados,B., Tainer,J.A. and Grasby,J.A. (2010) Substrate recognition and catalysis by flap endonucleases and related enzymes. *Biochem. Soc. Trans.*, **38**, 433–437.
2. Tsutakawa,S.E., Classen,S., Chapados,B.R., Arvai,A., Finger,L.D., Guenther,G., Tomlinson,C.G., Thompson,P., Sarker,A.H., Shen,B. *et al.* (2011) Human flap endonuclease structures, DNA double-base flipping and a unified understanding of the FEN1 superfamily. *Cell*, **145**, 198–211.
3. Orans,J., McSweeney,E.A., Iyer,R.R., Hast,M.A., Hellinga,H.W., Modrich,P. and Beese,L.S. (2011) Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. *Cell*, **145**, 212–223.
4. Larsen,E., Gran,C., Saether,B.E., Seeberg,E. and Klungland,A. (2003) Proliferation failure and gamma radiation sensitivity of Fen1 null mutant mice at the blastocyst stage. *Mol. Cell. Biol.*, **23**, 5346–5353.
5. Lyamichev,V., Brow,M.A. and Dahlberg,J.E. (1993) Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerases. *Science*, **260**, 778–783.
6. Ceska,T.A., Sayers,J.R., Stier,G. and Suck,D. (1996) A helical arch allowing single-stranded DNA to thread through T5 5′-exonuclease. *Nature*, **382**, 90–93.
7. Chapados,B.R., Hosfield,D.J., Han,S., Qiu,J., Yelent,B., Shen,B. and Tainer,J.A. (2004) Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell*, **116**, 39–50.
8. Sakurai,S., Kitano,K., Yamaguchi,H., Hamada,K., Okada,K., Fukuda,K., Uchida,M., Ohtsuka,E., Morioka,H. and Hakoshima,T. (2005) Structural basis for recruitment of human flap endonuclease 1 to PCNA. *EMBO J.*, **24**, 683–693.
9. Devos,J.M., Tomanicek,S.J., Jones,C.E., Nossal,N.G. and Mueser,T.J. (2007) Crystal structure of bacteriophage T4 5′ nuclease in complex with a branched DNA reveals how FEN-1 family nucleases bind their substrates. *J. Biol. Chem.*, **282**, 31713–31724.
10. Mueser,T.C., Nossal,N.G. and Hyde,C.C. (1996) Structure of bacteriophage T4 RNase H, a 5′ to 3′ RNA-DNA and DNA-DNA exonuclease with sequence similarity to the Rad2 family of eukaryotic proteins. *Cell*, **85**, 1101–1112.
11. 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. Biol. Chem.*, **270**, 30377–30383.
12. Wu,X., Li,J., Li,X., Hsieh,C.L., Burgers,P.M. and Lieber,M.R. (1996) Processing of branched DNA intermediates by a complex of human FEN-1 and PCNA. *Nucleic Acids Res.*, **24**, 2036–2043.
13. Finger,L.D., Blanchard,M.S., Theimer,C.A., Sengerová,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. Biol. Chem.*, **284**, 22184–22194.
14. 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. *Biochemistry*, **38**, 13347–13354.
15. Barnes,C.J., Wahl,A.F., Shen,B.H., 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.

16. Gloor,J.W., Balakrishnan,L. and Bambara,R.A. (2011) Flap Endonuclease 1 Mechanism Analysis Indicates Flap Base Binding Prior to Threading. *J. Biol. Chem.*, **285**, 34922–34931.
17. Williams,R., Sengerová,B., Osborne,S., Syson,K., Ault,S., Kilgour,A., Chapados,B.R., Tainer,J.A., Sayers,J.R. and Grasby,J.A. (2007) Comparison of the catalytic parameters and reaction specificities of a phage and an archaeal flap endonuclease. *J. Mol. Biol.*, **371**, 34–48.
18. 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 substrate DNA. *J. Biol. Chem.*, **276**, 30167–30177.
19. Lieber,M.R. (1997) The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *Bioessays*, **19**, 233–240.
20. Lee,B.I. and Wilson,D.M. (1999) The RAD2 domain of human exonuclease 1 exhibits 5′ to 3′ exonuclease and flap structure-specific endonuclease activities. *J. Biol. Chem.*, **274**, 37763–37769.
21. Hohl,M., Thorel,F., Clarkson,S.G. and Scharer,O.D. (2003) Structural determinants for substrate binding and catalysis by the structure-specific endonuclease XPG. *J. Biol. Chem.*, **278**, 19500–19508.
22. Ip,S.C., Rass,U., Blanco,M.G., Flynn,H.R., Skehel,J.M. and West,S.C. (2008) Identification of Holliday junction resolvases from humans and yeast. *Nature*, **456**, 357–361.
23. Williams,R. and Kunkel,T. (2011) FEN nucleases: bind, bend, fray, cut. *Cell*, **145**, 171–172.
24. 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.
25. 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.
26. Patel,D., Tock,M.R., Frary,E., Feng,M., Pickering,T.J., Grasby,J.A. and Sayers,J.R. (2002) A conserved tyrosine residue aids ternary complex formation, but not catalysis, in phage T5 flap endonuclease. *J. Mol. Biol.*, **320**, 1025–1035.
27. Shibata,T., Glynn,N., McMurry,T.B., McElhinney,R.S., Margison,G.P. and Williams,D.M. (2006) Novel synthesis of O6-alkylguanine containing oligodeoxyribonucleotides as substrates for the human DNA repair protein, O6-methylguanine DNA methyltransferase (MGMT). *Nucleic Acids Res.*, **34**, 1884–1891.
28. Kocalka,P., El-Sagheer,A. and Brown,T. (2008) Rapid and efficient DNA strand cross-linking by click chemistry. *Chembiochem*, **9**, 1280–1285.
29. Chivers,C., Crozat,E., Chu,C., Moy,V., Sherratt,D. and Howarth,M. (2010) A streptavidin variant with slower biotin dissociation and increased mechanostability. *Nature Methods*, **7**, 391–393.
30. Tomlinson,C., Syson,K., Sengerová,B., Atack,J., Sayers,J., Swanson,L., Tainer,J., Williams,N. and Grasby,J. (2011) Neutralizing mutations of carboxylates that bind metal 2 in T5 flap endonuclease result in an enzyme that still requires two metal ions. *J. Biol. Chem.*, **286**, 30878–30887.
31. 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. Biol. Chem.*, **283**, 28741–28746.
32. Sengerová,B., Tomlinson,C., Atack,J.M., Williams,R., Sayers,J.R., Williams,N.H. and Grasby,J.A. (2010) Brønsted analysis and rate-limiting steps for the T5 flap endonuclease catalyzed hydrolysis of exonucleolytic substrates. *Biochemistry*, **49**, 8085–8093.
33. Storici,F., Henneke,G., Ferrari,E., Gordenin,D.A., Hübscher,U. and Resnick,M.A. (2002) The flexible loop of human FEN1 endonuclease is required for flap cleavage during DNA replication and repair. *EMBO J.*, **21**, 5930–5942.
34. Hitomi,K., Iwai,S. and Tainer,J.A. (2007) The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal, and repair. *DNA Repair*, **6**, 410–428.
35. McMurray,C. (2008) Hijacking of the mismatch repair system to cause CAG expansion and cell death in neurodegenerative disease. *DNA Repair*, **7**, 1121–1134.
36. Reagan,M.S., Pittenger,C., Siede,W. and Friedberg,E.C. (1995) Characterization of a mutant strain of Saccharomyces cerevisiae with a deletion of the RAD27 gene, a structural homolog of the RAD2 nucleotide excision repair gene. *J. Bacteriol.*, **177**, 364–371.
37. Matsuzaki,Y., Adachi,N. and Koyama,H. (2002) Vertebrate cells lacking FEN-1 endonuclease are viable but hypersensitive to methylating agents and H2O2. *Nucleic Acids Res.*, **30**, 3273–3277.
38. Finger,L.D. and Shen,B. (2010) FEN1 (flap structure-specific endonuclease 1). *Atlas Genet Cytogenet Oncol Haematol.*, http://AtlasGeneticsOncology.org/Genes/FEN1ID40543ch11q12.html (2 February 2012, date last accessed).
39. Tumey,L.N., Bom,D., Huck,B., Gleason,E., Wang,J., Silver,D., Brunden,K., Boozer,S., Rundlett,S., Sherf,B. *et al.* (2005) The identification and optimization of a N-hydroxy urea series of flap endonuclease 1 inhibitors. *Bioorg. Med. Chem. Lett.*, **15**, 277–281.
40. Chang,J.H., Xiang,S., Xiang,K., Manley,J.L. and Tong,L.A. (2011) Structural and biochemical studies of the 5′→3′ exoribonuclease Xrn1. *Nat. Struct. Mol. Biol.*, **18**, 270–276.
41. Jinek,M., Coyle,S.M. and Doudna,J.A. (2011) Coupled 5′ Nucleotide recognition and processivity in Xrn1-Mediated mRNA decay. *Mol. Cell*, **41**, 600–608.
42. Xiang,S., Cooper-Morgan,A., Jiao,X., Kiledjian,M., Manley,J.L. and Tong,L. (2009) Structure and function of the 5′→3′ exoribonuclease Rat1 and its activating partner Rai1. *Nature*, **458**, 784–788.
43. Staresincic,L., Fagbemi,A.F., Enzlin,J.H., Gourdin,A.M., Wijgers,N., Dunand-Sauthier,I., Giglia-Mari,G., Clarkson,S.G., Vermeulen,W. and Schärer,O.D. (2009) Coordination of dual incision and repair synthesis in human nucleotide excision repair. *EMBO J.*, **28**, 1111–1120.
44. Zheng,L., Jia,J., Finger,L.D., Guo,Z., Zer,C. and Shen,B.H. (2011) Functional regulation of FEN1 nuclease and its link to cancer. *Nucleic Acids Res.*, **39**, 781–794.
45. Guo,Z., Zheng,L., Xu,H., Dai,H., Zhou,M., Pascua,M.R., Chen,Q.M. and Shen,B.H. (2010) Methylation of FEN1 suppresses nearby phosphorylation and facilitates PCNA binding. *Nat. Chem. Biol.*, **6**, 766–773.
46. Friedrich-Heineken,E. and Hubscher,U. (2004) The Fen1 extrahelical 3′-flap pocket is conserved from archaea to human and regulates DNA substrate specificity. *Nucleic Acids Res.*, **32**, 2520–2528.